# Table of Contents

1. Introduction ........................................................................................................... 3  
   1.1. Document Conventions .................................................................................. 3  
2. Fundamentals ......................................................................................................... 5  
   2.1. Host and Device Environment ...................................................................... 5  
   2.2. Execution Model .............................................................................................. 5  
   2.3. Object Model .................................................................................................. 7  
   2.4. Application Binary Interface ......................................................................... 10  
   2.5. Command Syntax and Duration .................................................................... 11  
   2.6. Threading Behavior ......................................................................................... 13  
   2.7. Errors ............................................................................................................. 19  
   2.8. Numeric Representation and Computation ..................................................... 27  
   2.9. Fixed-Point Data Conversions ....................................................................... 29  
   2.10. API Version Numbers and Semantics .......................................................... 30  
   2.11. Common Object Types .................................................................................. 31  
3. Initialization ........................................................................................................... 33  
   3.1. Command Function Pointers .......................................................................... 33  
   3.2. Instances ......................................................................................................... 35  
4. Devices and Queues .............................................................................................. 40  
   4.1. Physical Devices .............................................................................................. 40  
   4.2. Devices ............................................................................................................ 46  
   4.3. Queues ............................................................................................................. 52  
5. Command Buffers ................................................................................................ 57  
   5.1. Command Buffer Lifecycle ........................................................................... 57  
   5.2. Command Pools ............................................................................................... 59  
   5.3. Command Buffer Allocation and Management ............................................. 63  
   5.4. Command Buffer Recording ........................................................................... 67  
   5.5. Command Buffer Submission ......................................................................... 73  
   5.6. Queue Forward Progress ............................................................................... 78  
   5.7. Secondary Command Buffer Execution ......................................................... 79  
6. Synchronization and Cache Control .................................................................... 82  
   6.1. Execution and Memory Dependencies .......................................................... 82  
   6.2. Implicit Synchronization Guarantees ............................................................. 95  
   6.3. Fences ............................................................................................................. 96  
   6.4. Semaphores .................................................................................................... 103  
   6.5. Events ............................................................................................................ 107  
   6.6. Pipeline Barriers ............................................................................................ 118  
   6.7. Memory Barriers ............................................................................................ 124  
   6.8. Wait Idle Operations ..................................................................................... 131
Copyright 2014-2018 The Khronos Group Inc.

This Specification is protected by copyright laws and contains material proprietary to Khronos. Except as described by these terms, it or any components may not be reproduced, republished, distributed, transmitted, displayed, broadcast or otherwise exploited in any manner without the express prior written permission of Khronos. Khronos grants a conditional copyright license to use and reproduce the unmodified Specification for any purpose, without fee or royalty, EXCEPT no licenses to any patent, trademark or other intellectual property rights are granted under these terms.

Khronos makes no, and expressly disclaims any, representations or warranties, express or implied, regarding this Specification, including, without limitation: merchantability, fitness for a particular purpose, non-infringement of any intellectual property, correctness, accuracy, completeness, timeliness, and reliability. Under no circumstances will Khronos, or any of its Promoters, Contributors or Members, or their respective partners, officers, directors, employees, agents or representatives be liable for any damages, whether direct, indirect, special or consequential damages for lost revenues, lost profits, or otherwise, arising from or in connection with these materials.

This Specification has been created under the Khronos Intellectual Property Rights Policy, which is Attachment A of the Khronos Group Membership Agreement available at www.khronos.org/files/member_agreement.pdf, and which defines the terms 'Scope', 'Compliant Portion', and 'Necessary Patent Claims'. Parties desiring to implement the Specification and make use of Khronos trademarks in relation to that implementation, and receive reciprocal patent license protection under the Khronos Intellectual Property Rights Policy must become Adopters and confirm the implementation as conformant under the process defined by Khronos for this Specification; see https://www.khronos.org/adopters.

This Specification contains substantially unmodified functionality from, and is a successor to, Khronos specifications including OpenGL, OpenGL ES and OpenCL.

Some parts of this Specification are purely informative and so are EXCLUDED from the Scope of this Specification. The Document Conventions section of the Introduction defines how these parts of the Specification are identified.

Where this Specification uses technical terminology, defined in the [glossary] or otherwise, that refer to enabling technologies that are not expressly set forth in this Specification, those enabling technologies are EXCLUDED from the Scope of this Specification. For clarity, enabling technologies not disclosed with particularity in this Specification (e.g. semiconductor manufacturing technology, hardware architecture, processor architecture or microarchitecture, memory architecture, compiler technology, object oriented technology, basic operating system technology, compression technology, algorithms, and so on) are NOT to be considered expressly set forth; only those application program interfaces and data structures disclosed with particularity are included in the Scope of this Specification.

For purposes of the Khronos Intellectual Property Rights Policy as it relates to the definition of Necessary Patent Claims, all recommended or optional features, behaviors and functionality set forth in this Specification, if implemented, are considered to be included as Compliant Portions.
Where this Specification includes normative references to external documents, only the specifically identified sections of those external documents are INCLUDED in the Scope of this Specification. If not created by Khronos, those external documents may contain contributions from non-members of Khronos not covered by the Khronos Intellectual Property Rights Policy.

Vulkan is a registered trademark, and Khronos is a trademark of The Khronos Group Inc. ASTC is a trademark of ARM Holdings PLC; OpenCL is a trademark of Apple Inc.; and OpenGL is a registered trademark of Silicon Graphics International, all used under license by Khronos. All other product names, trademarks, and/or company names are used solely for identification and belong to their respective owners.
Chapter 1. Introduction

This document, referred to as the “Vulkan Specification” or just the “Specification” hereafter, describes the Vulkan Application Programming Interface (API). Vulkan is a C99 API designed for explicit control of low-level graphics and compute functionality.

The canonical version of the Specification is available in the official Vulkan Registry (http://www.khronos.org/registry/vulkan/). The source files used to generate the Vulkan specification are stored in the Vulkan Documentation Repository (https://github.com/KhronosGroup/Vulkan-Docs). The source repository additionally has a public issue tracker and allows the submission of pull requests that improve the specification.

1.1. Document Conventions

The Vulkan specification is intended for use by both implementors of the API and application developers seeking to make use of the API, forming a contract between these parties. Specification text may address either party; typically the intended audience can be inferred from context, though some sections are defined to address only one of these parties. (For example, Valid Usage sections only address application developers). Any requirements, prohibitions, recommendations or options defined by normative terminology are imposed only on the audience of that text.

Structure and enumerated types defined in extensions that were promoted to core in Vulkan 1.1 are now defined in terms of the equivalent Vulkan 1.1 interfaces. This affects the Vulkan Specification, the Vulkan header files, and the corresponding XML Registry.

1.1.1. Normative Terminology

Within this specification, the key words must, required, should, recommended, may, and optional are to be interpreted as described in RFC 2119 - Key words for use in RFCs to Indicate Requirement Levels (http://www.ietf.org/rfc/rfc2119.txt). These key words are highlighted in the specification for clarity. In text addressing application developers, their use expresses requirements that apply to application behavior. In text addressing implementors, their use expresses requirements that apply to implementations.

In text addressing application developers, the additional key words can and cannot are to be interpreted as describing the capabilities of an application, as follows:

**can**

This word means that the application is able to perform the action described.

**cannot**

This word means that the API and/or the execution environment provide no mechanism through which the application can express or accomplish the action described.

These key words are never used in text addressing implementors.
Note

There is an important distinction between cannot and must not, as used in this Specification. Cannot means something the application literally is unable to express or accomplish through the API, while must not means something that the application is capable of expressing through the API, but that the consequences of doing so are undefined and potentially unrecoverable for the implementation.

Unless otherwise noted in the section heading, all sections and appendices in this document are normative.

1.1.2. Technical Terminology

The Vulkan Specification makes use of common engineering and graphics terms such as Pipeline, Shader, and Host to identify and describe Vulkan API constructs and their attributes, states, and behaviors. The Glossary defines the basic meanings of these terms in the context of the Specification. The Specification text provides fuller definitions of the terms and may elaborate, extend, or clarify the Glossary definitions. When a term defined in the Glossary is used in normative language within the Specification, the definitions within the Specification govern and supersede any meanings the terms may have in other technical contexts (i.e. outside the Specification).

1.1.3. Normative References

References to external documents are considered normative references if the Specification uses any of the normative terms defined in Normative Terminology to refer to them or their requirements, either as a whole or in part.

The following documents are referenced by normative sections of the specification:


Chapter 2. Fundamentals

This chapter introduces fundamental concepts including the Vulkan architecture and execution model, API syntax, queues, pipeline configurations, numeric representation, state and state queries, and the different types of objects and shaders. It provides a framework for interpreting more specific descriptions of commands and behavior in the remainder of the Specification.

2.1. Host and Device Environment

The Vulkan Specification assumes and requires: the following properties of the host environment with respect to Vulkan implementations:

- The host **must** have runtime support for 8, 16, 32 and 64-bit signed and unsigned two's complement integers, all addressable at the granularity of their size in bytes.
- The host **must** have runtime support for 32- and 64-bit floating-point types satisfying the range and precision constraints in the Floating Point Computation section.
- The representation and endianness of these types on the host **must** match the representation and endianness of the same types on every physical device supported.

**Note**

Since a variety of data types and structures in Vulkan **may** be accessible by both host and physical device operations, the implementation **should** be able to access such data efficiently in both paths in order to facilitate writing portable and performant applications.

2.2. Execution Model

This section outlines the execution model of a Vulkan system.

Vulkan exposes one or more **devices**, each of which exposes one or more **queues** which **may** process work asynchronously to one another. The set of queues supported by a device is partitioned into **families**. Each family supports one or more types of functionality and **may** contain multiple queues with similar characteristics. Queues within a single family are considered **compatible** with one another, and work produced for a family of queues **can** be executed on any queue within that family. This Specification defines four types of functionality that queues **may** support: graphics, compute, transfer, and sparse memory management.

**Note**

A single device **may** report multiple similar queue families rather than, or as well as, reporting multiple members of one or more of those families. This indicates that while members of those families have similar capabilities, they are **not** directly compatible with one another.

Device memory is explicitly managed by the application. Each device **may** advertise one or more heaps, representing different areas of memory. Memory heaps are either device local or host local,
but are always visible to the device. Further detail about memory heaps is exposed via memory types available on that heap. Examples of memory areas that may be available on an implementation include:

- *device local* is memory that is physically connected to the device.
- *device local, host visible* is device local memory that is visible to the host.
- *host local, host visible* is memory that is local to the host and visible to the device and host.

On other architectures, there may only be a single heap that can be used for any purpose.

A Vulkan application controls a set of devices through the submission of command buffers which have recorded device commands issued via Vulkan library calls. The content of command buffers is specific to the underlying implementation and is opaque to the application. Once constructed, a command buffer can be submitted once or many times to a queue for execution. Multiple command buffers can be built in parallel by employing multiple threads within the application.

Command buffers submitted to different queues may execute in parallel or even out of order with respect to one another. Command buffers submitted to a single queue respect submission order, as described further in synchronization chapter. Command buffer execution by the device is also asynchronous to host execution. Once a command buffer is submitted to a queue, control may return to the application immediately. Synchronization between the device and host, and between different queues is the responsibility of the application.

### 2.2.1. Queue Operation

Vulkan queues provide an interface to the execution engines of a device. Commands for these execution engines are recorded into command buffers ahead of execution time. These command buffers are then submitted to queues with a queue submission command for execution in a number of batches. Once submitted to a queue, these commands will begin and complete execution without further application intervention, though the order of this execution is dependent on a number of implicit and explicit ordering constraints.

Work is submitted to queues using queue submission commands that typically take the form `vkQueue*` (e.g. `vkQueueSubmit`, `vkQueueBindSparse`), and optionally take a list of semaphores upon which to wait before work begins and a list of semaphores to signal once work has completed. The work itself, as well as signaling and waiting on the semaphores are all queue operations.

Queue operations on different queues have no implicit ordering constraints, and may execute in any order. Explicit ordering constraints between queues can be expressed with semaphores and fences.

Command buffer submissions to a single queue respect submission order and other implicit ordering guarantees, but otherwise may overlap or execute out of order. Other types of batches and queue submissions against a single queue (e.g. sparse memory binding) have no implicit ordering constraints with any other queue submission or batch. Additional explicit ordering constraints between queue submissions and individual batches can be expressed with semaphores and fences.

Before a fence or semaphore is signaled, it is guaranteed that any previously submitted queue operations have completed execution, and that memory writes from those queue operations are
available to future queue operations. Waiting on a signaled semaphore or fence guarantees that previous writes that are available are also visible to subsequent commands.

Command buffer boundaries, both between primary command buffers of the same or different batches or submissions as well as between primary and secondary command buffers, do not introduce any additional ordering constraints. In other words, submitting the set of command buffers (which can include executing secondary command buffers) between any semaphore or fence operations execute the recorded commands as if they had all been recorded into a single primary command buffer, except that the current state is reset on each boundary. Explicit ordering constraints can be expressed with explicit synchronization primitives.

There are a few implicit ordering guarantees between commands within a command buffer, but only covering a subset of execution. Additional explicit ordering constraints can be expressed with the various explicit synchronization primitives.

Note
Implementations have significant freedom to overlap execution of work submitted to a queue, and this is common due to deep pipelining and parallelism in Vulkan devices.

Commands recorded in command buffers either perform actions (draw, dispatch, clear, copy, query/timestamp operations, begin/end subpass operations), set state (bind pipelines, descriptor sets, and buffers, set dynamic state, push constants, set render pass/subpass state), or perform synchronization (set/wait events, pipeline barrier, render pass/subpass dependencies). Some commands perform more than one of these tasks. State setting commands update the current state of the command buffer. Some commands that perform actions (e.g. draw/dispatch) do so based on the current state set cumulatively since the start of the command buffer. The work involved in performing action commands is often allowed to overlap or to be reordered, but doing so must not alter the state to be used by each action command. In general, action commands are those commands that alter framebuffer attachments, read/write buffer or image memory, or write to query pools.

Synchronization commands introduce explicit execution and memory dependencies between two sets of action commands, where the second set of commands depends on the first set of commands. These dependencies enforce that both the execution of certain pipeline stages in the later set occur after the execution of certain stages in the source set, and that the effects of memory accesses performed by certain pipeline stages occur in order and are visible to each other. When not enforced by an explicit dependency or implicit ordering guarantees, action commands may overlap execution or execute out of order, and may not see the side effects of each other’s memory accesses.

The device executes queue operations asynchronously with respect to the host. Control is returned to an application immediately following command buffer submission to a queue. The application must synchronize work between the host and device as needed.

2.3. Object Model

The devices, queues, and other entities in Vulkan are represented by Vulkan objects. At the API
level, all objects are referred to by handles. There are two classes of handles, dispatchable and non-
dispatchable. Dispatchable handle types are a pointer to an opaque type. This pointer may be used
by layers as part of intercepting API commands, and thus each API command takes a dispatchable
type as its first parameter. Each object of a dispatchable type must have a unique handle value
during its lifetime.

Non-dispatchable handle types are a 64-bit integer type whose meaning is implementation-
dependent, and may encode object information directly in the handle rather than acting as a
reference to an underlying object. Objects of a non-dispatchable type may not have unique handle
values within a type or across types. If handle values are not unique, then destroying one such
handle must not cause identical handles of other types to become invalid, and must not cause
identical handles of the same type to become invalid if that handle value has been created more
times than it has been destroyed.

All objects created or allocated from a VkDevice (i.e. with a VkDevice as the first parameter) are
private to that device, and must not be used on other devices.

2.3.1. Object Lifetime

Objects are created or allocated by vkCreate* and vkAllocate* commands, respectively. Once an
object is created or allocated, its “structure” is considered to be immutable, though the contents of
certain object types is still free to change. Objects are destroyed or freed by vkDestroy* and vkFree*
commands, respectively.

Objects that are allocated (rather than created) take resources from an existing pool object or
memory heap, and when freed return resources to that pool or heap. While object creation and
destruction are generally expected to be low-frequency occurrences during runtime, allocating and
freeing objects can occur at high frequency. Pool objects help accommodate improved performance
of the allocations and frees.

It is an application’s responsibility to track the lifetime of Vulkan objects, and not to destroy them
while they are still in use.

The ownership of application-owned memory is immediately acquired by any Vulkan command it
is passed into. Ownership of such memory must be released back to the application at the end of
the duration of the command, so that the application can alter or free this memory as soon as all
the commands that acquired it have returned.

The following object types are consumed when they are passed into a Vulkan command and not
further accessed by the objects they are used to create. They must not be destroyed in the duration
of any API command they are passed into:

- VkShaderModule
- VkPipelineCache

A VkRenderPass object passed as a parameter to create another object is not further accessed by that
object after the duration of the command it is passed into. A VkRenderPass used in a command
buffer follows the rules described below.

A VkPipelineLayout object must not be destroyed while any command buffer that uses it is in the
-recording state.

`VkDescriptorSetLayout` objects may be accessed by commands that operate on descriptor sets allocated using that layout, and those descriptor sets must not be updated with `vkUpdateDescriptorSets` after the descriptor set layout has been destroyed. Otherwise, a `VkDescriptorSetLayout` object passed as a parameter to create another object is not further accessed by that object after the duration of the command it is passed into.

The application must not destroy any other type of Vulkan object until all uses of that object by the device (such as via command buffer execution) have completed.

The following Vulkan objects must not be destroyed while any command buffers using the object are in the pending state:

- `VkEvent`
- `VkQueryPool`
- `VkBuffer`
- `VkBufferView`
- `VkImage`
- `VkImageView`
- `VkPipeline`
- `VkSampler`
- `VkDescriptorPool`
- `VkFramebuffer`
- `VkRenderPass`
- `VkCommandBuffer`
- `VkCommandPool`
- `VkDeviceMemory`
- `VkDescriptorSet`

Destroying these objects will move any command buffers that are in the recording or executable state, and are using those objects, to the invalid state.

The following Vulkan objects must not be destroyed while any queue is executing commands that use the object:

- `VkFence`
- `VkSemaphore`
- `VkCommandBuffer`
- `VkCommandPool`

In general, objects can be destroyed or freed in any order, even if the object being freed is involved in the use of another object (e.g. use of a resource in a view, use of a view in a descriptor set, use of an object in a command buffer, binding of a memory allocation to a resource), as long as any object that uses the freed object is not further used in any way except to be destroyed or to be reset in such a way that it no longer uses the other object (such as resetting a command buffer). If the object has been reset, then it can be used as if it never used the freed object. An exception to this is when there is a parent/child relationship between objects. In this case, the application must not destroy a
parent object before its children, except when the parent is explicitly defined to free its children when it is destroyed (e.g. for pool objects, as defined below).

`VkCommandPool` objects are parents of `VkCommandBuffer` objects. `VkDescriptorPool` objects are parents of `VkDescriptorSet` objects. `VkDevice` objects are parents of many object types (all that take a `VkDevice` as a parameter to their creation).

The following Vulkan objects have specific restrictions for when they can be destroyed:

- **`VkQueue` objects cannot** be explicitly destroyed. Instead, they are implicitly destroyed when the `VkDevice` object they are retrieved from is destroyed.

- Destroying a pool object implicitly frees all objects allocated from that pool. Specifically, destroying `VkCommandPool` frees all `VkCommandBuffer` objects that were allocated from it, and destroying `VkDescriptorPool` frees all `VkDescriptorSet` objects that were allocated from it.

- `VkDevice` objects can be destroyed when all `VkQueue` objects retrieved from them are idle, and all objects created from them have been destroyed. This includes the following objects:
  - `VkFence`
  - `VkSemaphore`
  - `VkEvent`
  - `VkQueryPool`
  - `VkBuffer`
  - `VkBufferView`
  - `VkImage`
  - `VkImageView`
  - `VkShaderModule`
  - `VkPipelineCache`
  - `VkPipeline`
  - `VkPipelineLayout`
  - `VkSampler`
  - `VkDescriptorSetLayout`
  - `VkDescriptorPool`
  - `VkFramebuffer`
  - `VkRenderPass`
  - `VkCommandPool`
  - `VkCommandBuffer`
  - `VkDeviceMemory`  

- **`VkPhysicalDevice` objects cannot** be explicitly destroyed. Instead, they are implicitly destroyed when the `VkInstance` object they are retrieved from is destroyed.

- `VkInstance` objects can be destroyed once all `VkDevice` objects created from any of its `VkPhysicalDevice` objects have been destroyed.

### 2.4. Application Binary Interface

The mechanism by which Vulkan is made available to applications is platform- or implementation-
defined. On many platforms the C interface described in this Specification is provided by a shared library. Since shared libraries can be changed independently of the applications that use them, they present particular compatibility challenges, and this Specification places some requirements on them.

Shared library implementations **must** use the default Application Binary Interface (ABI) of the standard C compiler for the platform, or provide customized API headers that cause application code to use the implementation’s non-default ABI. An ABI in this context means the size, alignment, and layout of C data types; the procedure calling convention; and the naming convention for shared library symbols corresponding to C functions. Customizing the calling convention for a platform is usually accomplished by defining calling convention macros appropriately in `vk_platform.h`.

On platforms where Vulkan is provided as a shared library, library symbols beginning with “vk” and followed by a digit or uppercase letter are reserved for use by the implementation. Applications which use Vulkan **must** not provide definitions of these symbols. This allows the Vulkan shared library to be updated with additional symbols for new API versions or extensions without causing symbol conflicts with existing applications.

Shared library implementations **should** provide library symbols for commands in the highest version of this Specification they support, and for Window System Integration extensions relevant to the platform. They **may** also provide library symbols for commands defined by additional extensions.

**Note**
These requirements and recommendations are intended to allow implementors to take advantage of platform-specific conventions for SDKs, ABIs, library versioning mechanisms, etc. while still minimizing the code changes necessary to port applications or libraries between platforms. Platform vendors, or providers of the *de facto* standard Vulkan shared library for a platform, are encouraged to document what symbols the shared library provides and how it will be versioned when new symbols are added.

Applications **should** only rely on shared library symbols for commands in the minimum core version required by the application. `vkGetInstanceProcAddr` and `vkGetDeviceProcAddr` **should** be used to obtain function pointers for commands in core versions beyond the application’s minimum required version.

## 2.5. Command Syntax and Duration

The Specification describes Vulkan commands as functions or procedures using C99 syntax. Language bindings for other languages such as C++ and JavaScript **may** allow for stricter parameter passing, or object-oriented interfaces.

Vulkan uses the standard C types for the base type of scalar parameters (e.g. types from `<stdint.h>`), with exceptions described below, or elsewhere in the text when appropriate:

`VkBool32` represents boolean `True` and `False` values, since C does not have a sufficiently portable built-in boolean type:
**typedef** uint32_t VkBool32;

`VK_TRUE` represents a boolean `True` (integer 1) value, and `VK_FALSE` a boolean `False` (integer 0) value. All values returned from a Vulkan implementation in a `VkBool32` will be either `VK_TRUE` or `VK_FALSE`.

Applications **must** not pass any other values than `VK_TRUE` or `VK_FALSE` into a Vulkan implementation where a `VkBool32` is expected.

`VkDeviceSize` represents device memory size and offset values:

**typedef** uint64_t VkDeviceSize;

Commands that create Vulkan objects are of the form `vkCreate*` and take `Vk*CreateInfo` structures with the parameters needed to create the object. These Vulkan objects are destroyed with commands of the form `vkDestroy*`.

The last in-parameter to each command that creates or destroys a Vulkan object is `pAllocator`. The `pAllocator` parameter can be set to a non-NULL value such that allocations for the given object are delegated to an application provided callback; refer to the Memory Allocation chapter for further details.

Commands that allocate Vulkan objects owned by pool objects are of the form `vkAllocate*`, and take `Vk*AllocateInfo` structures. These Vulkan objects are freed with commands of the form `vkFree*`. These objects do not take allocators; if host memory is needed, they will use the allocator that was specified when their parent pool was created.

Commands are recorded into a command buffer by calling API commands of the form `vkCmd*`. Each such command may have different restrictions on where it can be used: in a primary and/or secondary command buffer, inside and/or outside a render pass, and in one or more of the supported queue types. These restrictions are documented together with the definition of each such command.

The **duration** of a Vulkan command refers to the interval between calling the command and its return to the caller.

**2.5.1. Lifetime of Retrieved Results**

Information is retrieved from the implementation with commands of the form `vkGet*` and `vkEnumerate*`.

Unless otherwise specified for an individual command, the results are *invariant*; that is, they will remain unchanged when retrieved again by calling the same command with the same parameters, so long as those parameters themselves all remain valid.
2.6. Threading Behavior

Vulkan is intended to provide scalable performance when used on multiple host threads. All commands support being called concurrently from multiple threads, but certain parameters, or components of parameters are defined to be **externally synchronized**. This means that the caller **must** guarantee that no more than one thread is using such a parameter at a given time.

More precisely, Vulkan commands use simple stores to update the state of Vulkan objects. A parameter declared as externally synchronized **may** have its contents updated at any time during the host execution of the command. If two commands operate on the same object and at least one of the commands declares the object to be externally synchronized, then the caller **must** guarantee not only that the commands do not execute simultaneously, but also that the two commands are separated by an appropriate memory barrier (if needed).

---

**Note**

Memory barriers are particularly relevant for hosts based on the ARM CPU architecture, which is more weakly ordered than many developers are accustomed to from x86/x64 programming. Fortunately, most higher-level synchronization primitives (like the pthread library) perform memory barriers as a part of mutual exclusion, so muting Vk Vulkan objects via these primitives will have the desired effect.

Similarly the application **must** avoid any potential data hazard of application-owned memory that has its ownership **temporarily acquired** by a Vulkan command. While the ownership of application-owned memory remains acquired by a command the implementation **may** read the memory at any point, and it **may** write non-**const** qualified memory at any point. Parameters referring to non-**const** qualified application-owned memory are not marked explicitly as **externally synchronized** in the Specification.

Many object types are **immutable**, meaning the objects **cannot** change once they have been created. These types of objects never need external synchronization, except that they **must** not be destroyed while they are in use on another thread. In certain special cases mutable object parameters are internally synchronized, making external synchronization unnecessary. One example of this is the use of a `VkPipelineCache` in `vkCreateGraphicsPipelines` and `vkCreateComputePipelines`, where external synchronization around such a heavyweight command would be impractical. The implementation **must** internally synchronize the cache in this example, and **may** be able to do so in the form of a much finer-grained mutex around the command. Any command parameters that are not labeled as externally synchronized are either not mutated by the command or are internally synchronized. Additionally, certain objects related to a command’s parameters (e.g. command pools and descriptor pools) **may** be affected by a command, and **must** also be externally synchronized. These implicit parameters are documented as described below.

Parameters of commands that are externally synchronized are listed below.
Externally Synchronized Parameters

- The `instance` parameter in `vkDestroyInstance`
- The `device` parameter in `vkDestroyDevice`
- The `queue` parameter in `vkQueueSubmit`
- The `fence` parameter in `vkQueueSubmit`
- The `memory` parameter in `vkFreeMemory`
- The `memory` parameter in `vkMapMemory`
- The `memory` parameter in `vkUnmapMemory`
- The `buffer` parameter in `vkBindBufferMemory`
- The `image` parameter in `vkBindImageMemory`
- The `queue` parameter in `vkQueueBindSparse`
- The `fence` parameter in `vkQueueBindSparse`
- The `fence` parameter in `vkDestroyFence`
- The `Semaphore` parameter in `vkDestroySemaphore`
- The `event` parameter in `vkDestroyEvent`
- The `event` parameter in `vkSetEvent`
- The `event` parameter in `vkResetEvent`
- The `queryPool` parameter in `vkDestroyQueryPool`
- The `buffer` parameter in `vkDestroyBuffer`
- The `bufferView` parameter in `vkDestroyBufferView`
- The `image` parameter in `vkDestroyImage`
- The `imageView` parameter in `vkDestroyImageView`
- The `descriptorSetLayout` parameter in `vkDestroyDescriptorSetLayout`
- The `descriptorPool` parameter in `vkDestroyShaderModule`
- The `pipelineCache` parameter in `vkDestroyShaderModule`
- The `dstCache` parameter in `vkMergePipelineCaches`
- The `pipeline` parameter in `vkDestroyPipeline`
- The `pipelineLayout` parameter in `vkDestroyPipelineLayout`
- The `sampler` parameter in `vkDestroySampler`
- The `descriptorSetLayout` parameter in `vkDestroyDescriptorSetLayout`
- The `descriptorPool` parameter in `vkDestroyDescriptorPool`
- The `descriptorPool` parameter in `vkResetDescriptorPool`
- The `descriptorPool` the `pAllocateInfo` parameter in `vkAllocateDescriptorSets`
- The `descriptorPool` parameter in `vkFreeDescriptorSets`
- The `framebuffer` parameter in `vkDestroyFramebuffer`
- The `renderPass` parameter in `vkDestroyRenderPass`
- The `commandPool` parameter in `vkDestroyCommandPool`
- The `commandPool` parameter in `vkResetCommandPool`
- The `commandPool` parameter in `vkAllocateCommandBuffers`
- The `commandPool` parameter in `vkFreeCommandBuffers`
- The `commandBuffer` parameter in `vkBeginCommandBuffer`
- The `commandBuffer` parameter in `vkEndCommandBuffer`
- The `commandBuffer` parameter in `vkResetCommandBuffer`
- The `commandBuffer` parameter in `vkCmdBindPipeline`
- The `commandBuffer` parameter in `vkCmdSetViewport`
- The `commandBuffer` parameter in `vkCmdSetScissor`
- The `commandBuffer` parameter in `vkCmdSetLineWidth`
- The `commandBuffer` parameter in `vkCmdSetDepthBias`
- The `commandBuffer` parameter in `vkCmdSetBlendConstants`
- The `commandBuffer` parameter in `vkCmdSetDepthBounds`
- The `commandBuffer` parameter in `vkCmdSetStencilCompareMask`
- The `commandBuffer` parameter in `vkCmdSetStencilWriteMask`
- The `commandBuffer` parameter in `vkCmdSetStencilReference`
- The `commandBuffer` parameter in `vkCmdBindDescriptorSets`
- The `commandBuffer` parameter in `vkCmdBindIndexBuffer`
- The `commandBuffer` parameter in `vkCmdBindVertexBuffers`
- The `commandBuffer` parameter in `vkCmdDraw`
- The `commandBuffer` parameter in `vkCmdDrawIndexed`
- The `commandBuffer` parameter in `vkCmdDrawIndirect`
- The `commandBuffer` parameter in `vkCmdDispatch`
- The `commandBuffer` parameter in `vkCmdDispatchIndirect`
- The `commandBuffer` parameter in `vkCmdCopyBuffer`
- The `commandBuffer` parameter in `vkCmdCopyImage`
- The `commandBuffer` parameter in `vkCmdUpdateBuffer`
- The `commandBuffer` parameter in `vkCmdFillBuffer`
- The `commandBuffer` parameter in `vkCmdClearColorImage`
There are also a few instances where a command **can** take in a user allocated list whose contents are externally synchronized parameters. In these cases, the caller **must** guarantee that at most one thread is using a given element within the list at a given time. These parameters are listed below.
Externally Synchronized Parameter Lists

- Each element of the `pWaitSemaphores` member of each element of the `pSubmits` parameter in `vkQueueSubmit`
- Each element of the `pSignalSemaphores` member of each element of the `pSubmits` parameter in `vkQueueSubmit`
- Each element of the `pWaitSemaphores` member of each element of the `pBindInfo` parameter in `vkQueueBindSparse`
- Each element of the `pSignalSemaphores` member of each element of the `pBindInfo` parameter in `vkQueueBindSparse`
- The `buffer` member of each element of the ` pBufferBinds` member of each element of the ` pBindInfo` parameter in `vkQueueBindSparse`
- The `image` member of each element of the ` pImageOpaqueBinds` member of each element of the ` pBindInfo` parameter in `vkQueueBindSparse`
- The `image` member of each element of the ` pImageBinds` member of each element of the ` pBindInfo` parameter in `vkQueueBindSparse`
- Each element of the `pFences` parameter in `vkResetFences`
- Each element of the ` pDescriptorSets` parameter in `vkFreeDescriptorSets`
- The `dstSet` member of each element of the ` pDescriptorWrites` parameter in `vkUpdateDescriptorSets`
- The `dstSet` member of each element of the ` pDescriptorCopies` parameter in `vkUpdateDescriptorSets`
- Each element of the ` pCommandBuffers` parameter in `vkFreeCommandBuffers`

In addition, there are some implicit parameters that need to be externally synchronized. For example, all `commandBuffer` parameters that need to be externally synchronized imply that the `commandPool` that was passed in when creating that command buffer also needs to be externally synchronized. The implicit parameters and their associated object are listed below.
Implicit Externally Synchronized Parameters

- All VkQueue objects created from device in vkDeviceWaitIdle
- Any VkDescriptorSet objects allocated from descriptorPool in vkResetDescriptorPool
- The VkCommandPool that commandBuffer was allocated from in vkBeginCommandBuffer
- The VkCommandPool that commandBuffer was allocated from in vkEndCommandBuffer
- The VkCommandPool that commandBuffer was allocated from, in vkCmdBindPipeline
- The VkCommandPool that commandBuffer was allocated from, in vkCmdSetViewport
- The VkCommandPool that commandBuffer was allocated from, in vkCmdSetScissor
- The VkCommandPool that commandBuffer was allocated from, in vkCmdSetLineWidth
- The VkCommandPool that commandBuffer was allocated from, in vkCmdSetDepthBias
- The VkCommandPool that commandBuffer was allocated from, in vkCmdSetBlendConstants
- The VkCommandPool that commandBuffer was allocated from, in vkCmdSetDepthBounds
- The VkCommandPool that commandBuffer was allocated from, in vkCmdSetStencilCompareMask
- The VkCommandPool that commandBuffer was allocated from, in vkCmdSetStencilWriteMask
- The VkCommandPool that commandBuffer was allocated from, in vkCmdBindDescriptorSets
- The VkCommandPool that commandBuffer was allocated from, in vkCmdBindVertexBuffer
- The VkCommandPool that commandBuffer was allocated from, in vkCmdDraw
- The VkCommandPool that commandBuffer was allocated from, in vkCmdDrawIndexed
- The VkCommandPool that commandBuffer was allocated from, in vkCmdDrawIndirect
- The VkCommandPool that commandBuffer was allocated from, in vkCmdDrawIndexedIndirect
- The VkCommandPool that commandBuffer was allocated from, in vkCmdDispatch
- The VkCommandPool that commandBuffer was allocated from, in vkCmdDispatchIndirect
- The VkCommandPool that commandBuffer was allocated from, in vkCmdCopyBuffer
- The VkCommandPool that commandBuffer was allocated from, in vkCmdCopyImage
- The VkCommandPool that commandBuffer was allocated from, in vkCmdBlitImage
- The VkCommandPool that commandBuffer was allocated from, in vkCmdCopyBufferToImage
- The VkCommandPool that commandBuffer was allocated from, in vkCmdCopyImageToBuffer
- The VkCommandPool that commandBuffer was allocated from, in vkCmdUpdateBuffer
- The VkCommandPool that commandBuffer was allocated from, in vkCmdClearColorImage

- The VkCommandPool that commandBuffer was allocated from, in...
vkCmdClearDepthStencilImage

- The VkCommandPool that commandBuffer was allocated from, in vkCmdClearAttachments
- The VkCommandPool that commandBuffer was allocated from, in vkCmdResolveImage
- The VkCommandPool that commandBuffer was allocated from, in vkCmdSetEvent
- The VkCommandPool that commandBuffer was allocated from, in vkCmdResetEvent
- The VkCommandPool that commandBuffer was allocated from, in vkCmdWaitEvents
- The VkCommandPool that commandBuffer was allocated from, in vkCmdPipelineBarrier
- The VkCommandPool that commandBuffer was allocated from, in vkCmdBeginQuery
- The VkCommandPool that commandBuffer was allocated from, in vkCmdEndQuery
- The VkCommandPool that commandBuffer was allocated from, in vkCmdResetQueryPool
- The VkCommandPool that commandBuffer was allocated from, in vkCmdWriteTimestamp
- The VkCommandPool that commandBuffer was allocated from, in vkCmdCopyQueryPoolResults
- The VkCommandPool that commandBuffer was allocated from, in vkCmdPushConstants
- The VkCommandPool that commandBuffer was allocated from, in vkCmdBeginRenderPass
- The VkCommandPool that commandBuffer was allocated from, in vkCmdNextSubpass
- The VkCommandPool that commandBuffer was allocated from, in vkCmdEndRenderPass
- The VkCommandPool that commandBuffer was allocated from, in vkCmdExecuteCommands

2.7. Errors

Vulkan is a layered API. The lowest layer is the core Vulkan layer, as defined by this Specification. The application can use additional layers above the core for debugging, validation, and other purposes.

One of the core principles of Vulkan is that building and submitting command buffers should be highly efficient. Thus error checking and validation of state in the core layer is minimal, although more rigorous validation can be enabled through the use of layers.

The core layer assumes applications are using the API correctly. Except as documented elsewhere in the Specification, the behavior of the core layer to an application using the API incorrectly is undefined, and may include program termination. However, implementations must ensure that incorrect usage by an application does not affect the integrity of the operating system, the Vulkan implementation, or other Vulkan client applications in the system. In particular, any guarantees made by an operating system about whether memory from one process can be visible to another process or not must not be violated by a Vulkan implementation for any memory allocation. Vulkan implementations are not required to make additional security or integrity guarantees beyond those provided by the OS unless explicitly directed by the application’s use of a particular feature or extension (e.g. via robust buffer access).
Note

For instance, if an operating system guarantees that data in all its memory allocations are set to zero when newly allocated, the Vulkan implementation **must** make the same guarantees for any allocations it controls (e.g. `VkDeviceMemory`).

Applications **can** request stronger robustness guarantees by enabling the `robustBufferAccess` feature as described in [features].

Validation of correct API usage is left to validation layers. Applications **should** be developed with validation layers enabled, to help catch and eliminate errors. Once validated, released applications **should** not enable validation layers by default.

### 2.7.1. Valid Usage

Valid usage defines a set of conditions which **must** be met in order to achieve well-defined run-time behavior in an application. These conditions depend only on Vulkan state, and the parameters or objects whose usage is constrained by the condition.

Some valid usage conditions have dependencies on run-time limits or feature availability. It is possible to validate these conditions against Vulkan’s minimum supported values for these limits and features, or some subset of other known values.

Valid usage conditions do not cover conditions where well-defined behavior (including returning an error code) exists.

Valid usage conditions **should** apply to the command or structure where complete information about the condition would be known during execution of an application. This is such that a validation layer or linter **can** be written directly against these statements at the point they are specified.

**Note**

This does lead to some non-obvious places for valid usage statements. For instance, the valid values for a structure might depend on a separate value in the calling command. In this case, the structure itself will not reference this valid usage as it is impossible to determine validity from the structure that it is invalid - instead this valid usage would be attached to the calling command.

Another example is draw state - the state setters are independent, and can cause a legitimately invalid state configuration between draw calls; so the valid usage statements are attached to the place where all state needs to be valid - at the draw command.

Valid usage conditions are described in a block labelled “Valid Usage” following each command or structure they apply to.

### 2.7.2. Implicit Valid Usage

Some valid usage conditions apply to all commands and structures in the API, unless explicitly
denoted otherwise for a specific command or structure. These conditions are considered *implicit*, and are described in a block labelled “Valid Usage (Implicit)” following each command or structure they apply to. Implicit valid usage conditions are described in detail below.

### Valid Usage for Object Handles

Any input parameter to a command that is an object handle **must** be a valid object handle, unless otherwise specified. An object handle is valid if:

- It has been created or allocated by a previous, successful call to the API. Such calls are noted in the Specification.
- It has not been deleted or freed by a previous call to the API. Such calls are noted in the Specification.
- Any objects used by that object, either as part of creation or execution, **must** also be valid.

The reserved values **VK_NULL_HANDLE** and **NULL** **can** be used in place of valid non-dispatchable handles and dispatchable handles, respectively, when *explicitly called out in the Specification*. Any command that creates an object successfully **must** not return these values. It is valid to pass these values to `vkDestroy*` or `vkFree*` commands, which will silently ignore these values.

### Valid Usage for Pointers

Any parameter that is a pointer **must** be a valid pointer only if it is explicitly called out by a Valid Usage statement.

A pointer is “valid” if it points at memory containing values of the number and type(s) expected by the command, and all fundamental types accessed through the pointer (e.g. as elements of an array or as members of a structure) satisfy the alignment requirements of the host processor.

### Valid Usage for Strings

Any parameter that is a pointer to `char` **must** be a finite sequence of values terminated by a null character, or if *explicitly called out in the Specification*, **can** be **NULL**.

### Valid Usage for Enumerated Types

Any parameter of an enumerated type **must** be a valid enumerant for that type. A enumerant is valid if:

- The enumerant is defined as part of the enumerated type.
- The enumerant is not one of the special values defined for the enumerated type, which are suffixed with `_BEGIN_RANGE`, `_END_RANGE`, `_RANGE_SIZE` or `_MAX_ENUM`.

The meaning of these special tokens is not exposed in the Vulkan Specification. They are not part of the API, and they **should** not be used by applications. Their original intended use was for internal consumption by Vulkan implementations. Even that use will no longer be supported in the future, but they will be retained for backwards compatibility reasons.
Any enumerated type returned from a query command or otherwise output from Vulkan to the application must not have a reserved value. Reserved values are values not defined by any extension for that enumerated type.

Note
This language is intended to accomodate cases such as “hidden” extensions known only to driver internals, or layers enabling extensions without knowledge of the application, without allowing return of values not defined by any extension.

Valid Usage for Flags

A collection of flags is represented by a bitmask using the type VkFlags:

```c
typedef uint32_t VkFlags;
```

Bitmasks are passed to many commands and structures to compactly represent options, but VkFlags is not used directly in the API. Instead, a Vk*Flags type which is an alias of VkFlags, and whose name matches the corresponding Vk*FlagBits that are valid for that type, is used.

Any Vk*Flags member or parameter used in the API as an input must be a valid combination of bit flags. A valid combination is either zero or the bitwise OR of valid bit flags. A bit flag is valid if:

- The bit flag is defined as part of the Vk*FlagBits type, where the bits type is obtained by taking the flag type and replacing the trailing Flags with FlagBits. For example, a flag value of type VkColorComponentFlags must contain only bit flags defined by VkColorComponentFlagBits.
- The flag is allowed in the context in which it is being used. For example, in some cases, certain bit flags or combinations of bit flags are mutually exclusive.

Any Vk*Flags member or parameter returned from a query command or otherwise output from Vulkan to the application may contain bit flags undefined in its corresponding Vk*FlagBits type. An application cannot rely on the state of these unspecified bits.

Valid Usage for Structure Types

Any parameter that is a structure containing a sType member must have a value of sType which is a valid VkStructureType value matching the type of the structure.

Structure types supported by the Vulkan API include:

```c
typedef enum VkStructureType {
    VK_STRUCTURE_TYPE_APPLICATION_INFO = 0,
    VK_STRUCTURE_TYPE_INSTANCE_CREATE_INFO = 1,
    VK_STRUCTURE_TYPE_DEVICE_QUEUE_CREATE_INFO = 2,
    VK_STRUCTURE_TYPE_DEVICE_CREATE_INFO = 3,
    VK_STRUCTURE_TYPE_SUBMIT_INFO = 4,
    VK_STRUCTURE_TYPE_MEMORY_ALLOCATE_INFO = 5,
    VK_STRUCTURE_TYPE_MAPPED_MEMORY_RANGE = 6,
    VK_STRUCTURE_TYPE_BIND_SPARSE_INFO = 7,
}
```
VK_STRUCTURE_TYPE_FENCE_CREATE_INFO = 8,
VK_STRUCTURE_TYPE_SEMAPHORE_CREATE_INFO = 9,
VK_STRUCTURE_TYPE_EVENT_CREATE_INFO = 10,
VK_STRUCTURE_TYPE_QUERY_POOL_CREATE_INFO = 11,
VK_STRUCTURE_TYPE_BUFFER_CREATE_INFO = 12,
VK_STRUCTURE_TYPE_BUFFER_VIEW_CREATE_INFO = 13,
VK_STRUCTURE_TYPE_IMAGE_CREATE_INFO = 14,
VK_STRUCTURE_TYPE_IMAGE_VIEW_CREATE_INFO = 15,
VK_STRUCTURE_TYPE_SHADER_MODULE_CREATE_INFO = 16,
VK_STRUCTURE_TYPE_PIPELINE_CACHE_CREATE_INFO = 17,
VK_STRUCTURE_TYPE_PIPELINE_SHADER_STAGE_CREATE_INFO = 18,
VK_STRUCTURE_TYPE_PIPELINE_VERTEX_INPUT_STATE_CREATE_INFO = 19,
VK_STRUCTURE_TYPE_PIPELINE_INPUT_ASSEMBLY_STATE_CREATE_INFO = 20,
VK_STRUCTURE_TYPE_PIPELINE_TESSELLATION_STATE_CREATE_INFO = 21,
VK_STRUCTURE_TYPE_PIPELINE_VIEWPORT_STATE_CREATE_INFO = 22,
VK_STRUCTURE_TYPE_PIPELINE_RASTERIZATION_STATE_CREATE_INFO = 23,
VK_STRUCTURE_TYPE_PIPELINE_MULTISAMPLE_STATE_CREATE_INFO = 24,
VK_STRUCTURE_TYPE_PIPELINE_DEPTH_STENCIL_STATE_CREATE_INFO = 25,
VK_STRUCTURE_TYPE_PIPELINE_COLOR_BLEND_STATE_CREATE_INFO = 26,
VK_STRUCTURE_TYPE_PIPELINE_DYNAMIC_STATE_CREATE_INFO = 27,
VK_STRUCTURE_TYPE_GRAPHICS_PIPELINE_CREATE_INFO = 28,
VK_STRUCTURE_TYPE_COMPUTE_PIPELINE_CREATE_INFO = 29,
VK_STRUCTURE_TYPE_PIPELINE_LAYOUT_CREATE_INFO = 30,
VK_STRUCTURE_TYPE_SAMPLER_CREATE_INFO = 31,
VK_STRUCTURE_TYPE_DESCRIPTOR_SET_LAYOUT_CREATE_INFO = 32,
VK_STRUCTURE_TYPE_DESCRIPTOR_POOL_CREATE_INFO = 33,
VK_STRUCTURE_TYPE_DESCRIPTOR_SET_ALLOCATE_INFO = 34,
VK_STRUCTURE_TYPE_WRITE_DESCRIPTOR_SET = 35,
VK_STRUCTURE_TYPE_COPY_DESCRIPTOR_SET = 36,
VK_STRUCTURE_TYPE_FRAMEBUFFER_CREATE_INFO = 37,
VK_STRUCTURE_TYPE_RENDER_PASS_CREATE_INFO = 38,
VK_STRUCTURE_TYPE_COMMAND_POOL_CREATE_INFO = 39,
VK_STRUCTURE_TYPE_COMMAND_BUFFER_ALLOCATE_INFO = 40,
VK_STRUCTURE_TYPE_COMMAND_BUFFER_INHERITANCE_INFO = 41,
VK_STRUCTURE_TYPE_COMMAND_BUFFER_BEGIN_INFO = 42,
VK_STRUCTURE_TYPE_RENDER_PASS_BEGIN_INFO = 43,
VK_STRUCTURE_TYPE_BUFFER_MEMORY_BARRIER = 44,
VK_STRUCTURE_TYPE_IMAGE_MEMORY_BARRIER = 45,
VK_STRUCTURE_TYPE_MEMORY_BARRIER = 46,
VK_STRUCTURE_TYPE_LOADER_INSTANCE_CREATE_INFO = 47,
VK_STRUCTURE_TYPE_LOADER_DEVICE_CREATE_INFO = 48,
} VkStructureType;

Each value corresponds to a particular structure with a sType member with a matching name. As a
general rule, the name of each VkStructureType value is obtained by taking the name of the
structure, stripping the leading Vk, prefixing each capital letter with _, converting the entire
resulting string to upper case, and prefixing it with VK_STRUCTURE_TYPE_. For example, structures of
type VkImageCreateInfo correspond to a VkStructureType of VK_STRUCTURE_TYPE_IMAGE_CREATE_INFO,
and thus its sType member must equal that when it is passed to the API.
The values `VK_STRUCTURE_TYPE_LOADER_INSTANCE_CREATE_INFO` and `VK_STRUCTURE_TYPE_LOADER_DEVICE_CREATE_INFO` are reserved for internal use by the loader, and do not have corresponding Vulkan structures in this Specification.

**Valid Usage for Structure Pointer Chains**

Any parameter that is a structure containing a `void* pNext` member must have a value of `pNext` that is either `NULL`, or points to a valid structure defined by an extension, containing `sType` and `pNext` members as described in the Vulkan Documentation and Extensions document in the section “Extension Interactions”. The set of structures connected by `pNext` pointers is referred to as a `pNext chain`. If that extension is supported by the implementation, then it must be enabled.

Each type of valid structure must not appear more than once in a `pNext` chain.

Any component of the implementation (the loader, any enabled layers, and drivers) must skip over, without processing (other than reading the `sType` and `pNext` members) any structures in the chain with `sType` values not defined by extensions supported by that component.

Extension structures are not described in the base Vulkan Specification, but either in layered Specifications incorporating those extensions, or in separate vendor-provided documents.

As a convenience to implementations and layers needing to iterate through a structure pointer chain, the Vulkan API provides the following base structures:

```c
typedef struct VkBaseInStructure {
    VkStructureType                    sType;
    const struct VkBaseInStructure*    pNext;
} VkBaseInStructure;
```

```c
typedef struct VkBaseOutStructure {
    VkStructureType               sType;
    struct VkBaseOutStructure*    pNext;
} VkBaseOutStructure;
```

`VkBaseInStructure` can be used to facilitate iterating through a read-only structure pointer chain. `VkBaseOutStructure` can be used to facilitate iterating through a structure pointer chain that returns data back to the application. These structures allow for some type safety and can be used by Vulkan API functions that operate on generic inputs and outputs.

**Valid Usage for Nested Structures**

The above conditions also apply recursively to members of structures provided as input to a command, either as a direct argument to the command, or themselves a member of another structure.

Specifics on valid usage of each command are covered in their individual sections.
Valid Usage for Extensions

Instance-level functionality or behavior added by an instance extension to the API must not be used unless that extension is supported by the instance as determined by `vkEnumerateInstanceExtensionProperties`, and that extension is enabled in `VkInstanceCreateInfo`.

Physical-device-level functionality or behavior added by an instance extension to the API must not be used unless that extension is supported by the instance as determined by `vkEnumerateInstanceExtensionProperties`, and that extension is enabled in `VkInstanceCreateInfo`.

Physical-device-level functionality or behavior added by a device extension to the API must not be used unless the conditions described in Extending Physical Device Core Functionality are met.

Device functionality or behavior added by a device extension to the API must not be used unless that extension is supported by the device as determined by `vkEnumerateDeviceExtensionProperties`, and that extension is enabled in `VkDeviceCreateInfo`.

Valid Usage for Newer Core Versions

Physical-device-level functionality or behavior added by a new core version of the API must not be used unless it is supported by the physical device as determined by `vkGetPhysicalDeviceProperties`.

Device-level functionality or behavior added by a new core version of the API must not be used unless it is supported by the device as determined by `vkGetPhysicalDeviceProperties`.

2.7.3. Return Codes

While the core Vulkan API is not designed to capture incorrect usage, some circumstances still require return codes. Commands in Vulkan return their status via return codes that are in one of two categories:

- Successful completion codes are returned when a command needs to communicate success or status information. All successful completion codes are non-negative values.
- Run time error codes are returned when a command needs to communicate a failure that could only be detected at run time. All run time error codes are negative values.

All return codes in Vulkan are reported via `VkResult` return values. The possible codes are:
typedef enum VkResult {
    VK_SUCCESS = 0,
    VK_NOT_READY = 1,
    VK_TIMEOUT = 2,
    VK_EVENT_SET = 3,
    VK_EVENT_RESET = 4,
    VK_INCOMPLETE = 5,
    VK_ERROR_OUT_OF_HOST_MEMORY = -1,
    VK_ERROR_OUT_OF_DEVICE_MEMORY = -2,
    VK_ERROR_INITIALIZATION_FAILED = -3,
    VK_ERROR_DEVICE_LOST = -4,
    VK_ERROR_MEMORY_MAP_FAILED = -5,
    VK_ERROR_LAYER_NOT_PRESENT = -6,
    VK_ERROR_EXTENSION_NOT_PRESENT = -7,
    VK_ERROR_FEATURE_NOT_PRESENT = -8,
    VK_ERROR_INCOMPATIBLE_DRIVER = -9,
    VK_ERROR_TOO_MANY_OBJECTS = -10,
    VK_ERROR_FORMAT_NOT_SUPPORTED = -11,
    VK_ERROR_FRAGMENTED_POOL = -12,
} VkResult;

**Success Codes**
- **VK_SUCCESS** Command successfully completed
- **VK_NOT_READY** A fence or query has not yet completed
- **VK_TIMEOUT** A wait operation has not completed in the specified time
- **VK_EVENT_SET** An event is signaled
- **VK_EVENT_RESET** An event is unsignaled
- **VK_INCOMPLETE** A return array was too small for the result

**Error codes**
- **VK_ERROR_OUT_OF_HOST_MEMORY** A host memory allocation has failed.
- **VK_ERROR_OUT_OF_DEVICE_MEMORY** A device memory allocation has failed.
- **VK_ERROR_INITIALIZATION_FAILED** Initialization of an object could not be completed for implementation-specific reasons.
- **VK_ERROR_DEVICE_LOST** The logical or physical device has been lost. See Lost Device
- **VK_ERROR_MEMORY_MAP_FAILED** Mapping of a memory object has failed.
- **VK_ERROR_LAYER_NOT_PRESENT** A requested layer is not present or could not be loaded.
- **VK_ERROR_EXTENSION_NOT_PRESENT** A requested extension is not supported.
- **VK_ERROR_FEATURE_NOT_PRESENT** A requested feature is not supported.
- **VK_ERROR_INCOMPATIBLE_DRIVER** The requested version of Vulkan is not supported by the driver or is otherwise incompatible for implementation-specific reasons.
- **VK_ERROR_TOO_MANY_OBJECTS** Too many objects of the type have already been created.
• **VK_ERROR_FORMAT_NOT_SUPPORTED** A requested format is not supported on this device.

• **VK_ERROR_FRAGMENTED_POOL** A pool allocation has failed due to fragmentation of the pool’s memory. This **must** only be returned if no attempt to allocate host or device memory was made to accommodate the new allocation.

If a command returns a run time error, unless otherwise specified any output parameters will have undefined contents, except that if the output parameter is a structure with **sType** and **pNext** fields, those fields will be unmodified. Any structures chained from **pNext** will also have undefined contents, except that **sType** and **pNext** will be unmodified.

Out of memory errors do not damage any currently existing Vulkan objects. Objects that have already been successfully created **can** still be used by the application.

Performance-critical commands generally do not have return codes. If a run time error occurs in such commands, the implementation will defer reporting the error until a specified point. For commands that record into command buffers (**vkCmd***s*) run time errors are reported by **vkEndCommandBuffer**.

### 2.8. Numeric Representation and Computation

Implementations normally perform computations in floating-point, and **must** meet the range and precision requirements defined under “Floating-Point Computation” below.

These requirements only apply to computations performed in Vulkan operations outside of shader execution, such as texture image specification and sampling, and per-fragment operations. Range and precision requirements during shader execution differ and are specified by the **Precision and Operation of SPIR-V Instructions** section.

In some cases, the representation and/or precision of operations is implicitly limited by the specified format of vertex or texel data consumed by Vulkan. Specific floating-point formats are described later in this section.

#### 2.8.1. Floating-Point Computation

Most floating-point computation is performed in SPIR-V shader modules. The properties of computation within shaders are constrained as defined by the **Precision and Operation of SPIR-V Instructions** section.

Some floating-point computation is performed outside of shaders, such as viewport and depth range calculations. For these computations, we do not specify how floating-point numbers are to be represented, or the details of how operations on them are performed, but only place minimal requirements on representation and precision as described in the remainder of this section.

We require simply that numbers’ floating-point parts contain enough bits and that their exponent fields are large enough so that individual results of floating-point operations are accurate to about 1 part in $10^5$. The maximum representable magnitude for all floating-point values **must** be at least $2^{32}$.

\[ x \times 0 = 0 \times x = 0 \] for any non-infinite and non-NaN \( x \).
\[ 1 \times x = x \times 1 = x. \]
\[ x + 0 = 0 + x = x. \]
\[ 0^0 = 1. \]

Occasionally, further requirements will be specified. Most single-precision floating-point formats meet these requirements.

The special values Inf and -Inf encode values with magnitudes too large to be represented; the special value NaN encodes “Not A Number” values resulting from undefined arithmetic operations such as 0 / 0. Implementations may support Inf and NaN in their floating-point computations.

### 2.8.2. Floating-Point Format Conversions

When a value is converted to a defined floating-point representation, finite values falling between two representable finite values are rounded to one or the other. The rounding mode is not defined. Finite values whose magnitude is larger than that of any representable finite value may be rounded either to the closest representable finite value or to the appropriately signed infinity. For unsigned destination formats any negative values are converted to zero. Positive infinity is converted to positive infinity; negative infinity is converted to negative infinity in signed formats and to zero in unsigned formats; and any NaN is converted to a NaN.

### 2.8.3. 16-Bit Floating-Point Numbers

16-bit floating point numbers are defined in the “16-bit floating point numbers” section of the Khronos Data Format Specification.

### 2.8.4. Unsigned 11-Bit Floating-Point Numbers

Unsigned 11-bit floating point numbers are defined in the “Unsigned 11-bit floating point numbers” section of the Khronos Data Format Specification.

### 2.8.5. Unsigned 10-Bit Floating-Point Numbers

Unsigned 10-bit floating point numbers are defined in the “Unsigned 10-bit floating point numbers” section of the Khronos Data Format Specification.

### 2.8.6. General Requirements

Any representable floating-point value in the appropriate format is legal as input to a Vulkan command that requires floating-point data. The result of providing a value that is not a floating-point number to such a command is unspecified, but must not lead to Vulkan interruption or termination. For example, providing a negative zero (where applicable) or a denormalized number to an Vulkan command must yield deterministic results, while providing a NaN or Inf yields unspecified results.

Some calculations require division. In such cases (including implied divisions performed by vector normalization), division by zero produces an unspecified result but must not lead to Vulkan

---

28 | Chapter 2. Fundamentals
2.9. Fixed-Point Data Conversions

When generic vertex attributes and pixel color or depth components are represented as integers, they are often (but not always) considered to be normalized. Normalized integer values are treated specially when being converted to and from floating-point values, and are usually referred to as normalized fixed-point.

In the remainder of this section, b denotes the bit width of the fixed-point integer representation. When the integer is one of the types defined by the API, b is the bit width of that type. When the integer comes from an image containing color or depth component texels, b is the number of bits allocated to that component in its specified image format.

The signed and unsigned fixed-point representations are assumed to be b-bit binary two’s-complement integers and binary unsigned integers, respectively.

2.9.1. Conversion from Normalized Fixed-Point to Floating-Point

Unsigned normalized fixed-point integers represent numbers in the range [0,1]. The conversion from an unsigned normalized fixed-point value c to the corresponding floating-point value f is defined as

\[ f = \frac{c}{2^b - 1} \]

Signed normalized fixed-point integers represent numbers in the range [-1,1]. The conversion from a signed normalized fixed-point value c to the corresponding floating-point value f is performed using

\[ f = \max\left(\frac{c}{2^b - 1}, -1.0\right) \]

Only the range \([-2^{b-1} + 1, 2^{b-1} - 1]\) is used to represent signed fixed-point values in the range [-1,1]. For example, if \(b = 8\), then the integer value -127 corresponds to -1.0 and the value 127 corresponds to 1.0. Note that while zero is exactly expressible in this representation, one value (-128 in the example) is outside the representable range, and **must** be clamped before use. This equation is used everywhere that signed normalized fixed-point values are converted to floating-point.

2.9.2. Conversion from Floating-Point to Normalized Fixed-Point

The conversion from a floating-point value f to the corresponding unsigned normalized fixed-point value c is defined by first clamping f to the range [0,1], then computing

\[ c = \text{convertFloatToUint}(f \times (2^b - 1), b) \]

where convertFloatToUint(r,b) returns one of the two unsigned binary integer values with exactly b bits which are closest to the floating-point value r. Implementations **should** round to nearest. If r is equal to an integer, then that integer value **must** be returned. In particular, if f is equal to 0.0 or 1.0, then c **must** be assigned 0 or \(2^b - 1\), respectively.
The conversion from a floating-point value \( f \) to the corresponding signed normalized fixed-point value \( c \) is performed by clamping \( f \) to the range \([-1,1]\), then computing

\[
c = \text{convertFloatToInt}(f \times (2^b - 1), b)
\]

where \( \text{convertFloatToInt}(r,b) \) returns one of the two signed two's-complement binary integer values with exactly \( b \) bits which are closest to the floating-point value \( r \). Implementations should round to nearest. If \( r \) is equal to an integer, then that integer value must be returned. In particular, if \( f \) is equal to -1.0, 0.0, or 1.0, then \( c \) must be assigned \(-(2^b - 1)\), 0, or \( 2^b - 1 \), respectively.

This equation is used everywhere that floating-point values are converted to signed normalized fixed-point.

### 2.10. API Version Numbers and Semantics

The Vulkan version number is used in several places in the API. In each such use, the API major version number, minor version number, and patch version number are packed into a 32-bit integer as follows:

- The major version number is a 10-bit integer packed into bits 31-22.
- The minor version number is a 10-bit integer packed into bits 21-12.
- The patch version number is a 12-bit integer packed into bits 11-0.

Differences in any of the Vulkan version numbers indicates a change to the API in some way, with each part of the version number indicating a different scope of changes.

A difference in patch version numbers indicates that some usually small part of the Specification or header has been modified, typically to fix a bug, and may have an impact on the behavior of existing functionality. Differences in this version number should not affect either full compatibility or backwards compatibility between two versions, or add additional interfaces to the API.

A difference in minor version numbers indicates that some amount of new functionality has been added. This will usually include new interfaces in the header, and may also include behavior changes and bug fixes. Functionality may be deprecated in a minor revision, but will not be removed. The patch version will continue to increment through minor version number changes since all minor versions are generated from the same source files, and changes to the source files may affect all minor versions within a major version. Differences in the patch version should not affect backwards compatibility, but will affect full compatibility. The patch version of the Specification is taken from `VK_HEADER_VERSION`.

A difference in major version numbers indicates a large set of changes to the API, potentially including new functionality and header interfaces, behavioral changes, removal of deprecated features, modification or outright replacement of any feature, and is thus very likely to break any and all compatibility. Differences in this version will typically require significant modification to an application in order for it to function.

C language macros for manipulating version numbers are defined in the `Version Number Macros` appendix.
2.11. Common Object Types

Some types of Vulkan objects are used in many different structures and command parameters, and are described here. These types include offsets, extents, and rectangles.

2.11.1. Offsets

Offsets are used to describe a pixel location within an image or framebuffer, as an (x,y) location for two-dimensional images, or an (x,y,z) location for three-dimensional images.

A two-dimensional offsets is defined by the structure:

```c
typedef struct VkOffset2D {
    int32_t    x;
    int32_t    y;
} VkOffset2D;
```

- x is the x offset.
- y is the y offset.

A three-dimensional offset is defined by the structure:

```c
typedef struct VkOffset3D {
    int32_t    x;
    int32_t    y;
    int32_t    z;
} VkOffset3D;
```

- x is the x offset.
- y is the y offset.
- z is the z offset.

2.11.2. Extents

Extents are used to describe the size of a rectangular region of pixels within an image or framebuffer, as (width,height) for two-dimensional images, or as (width,height,depth) for three-dimensional images.

A two-dimensional extent is defined by the structure:

```c
typedef struct VkExtent2D {
    uint32_t    width;
    uint32_t    height;
} VkExtent2D;
```
• \textit{width} is the width of the extent.
• \textit{height} is the height of the extent.

A three-dimensional extent is defined by the structure:

\begin{verbatim}
typedef struct VkExtent3D {
    uint32_t    width;
    uint32_t    height;
    uint32_t    depth;
} VkExtent3D;
\end{verbatim}

• \textit{width} is the width of the extent.
• \textit{height} is the height of the extent.
• \textit{depth} is the depth of the extent.

2.11.3. Rectangles

Rectangles are used to describe a specified rectangular region of pixels within an image or framebuffer. Rectangles include both an offset and an extent of the same dimensionality, as described above. Two-dimensional rectangles are defined by the structure:

\begin{verbatim}
typedef struct VkRect2D {
    VkOffset2D    offset;
    VkExtent2D    extent;
} VkRect2D;
\end{verbatim}

• \textit{offset} is a \texttt{VkOffset2D} specifying the rectangle offset.
• \textit{extent} is a \texttt{VkExtent2D} specifying the rectangle extent.
Chapter 3. Initialization

Before using Vulkan, an application must initialize it by loading the Vulkan commands, and creating a VkInstance object.

3.1. Command Function Pointers

Vulkan commands are not necessarily exposed statically on a platform. Function pointers for all Vulkan commands can be obtained with the command:

```c
PFN_vkVoidFunction vkGetInstanceProcAddr(
    VkInstance instance,
    const char* pName);
```

- `instance` is the instance that the function pointer will be compatible with, or NULL for commands not dependent on any instance.
- `pName` is the name of the command to obtain.

`vkGetInstanceProcAddr` itself is obtained in a platform- and loader- specific manner. Typically, the loader library will export this command as a function symbol, so applications can link against the loader library, or load it dynamically and look up the symbol using platform-specific APIs.

The table below defines the various use cases for `vkGetInstanceProcAddr` and expected return value ("fp" is “function pointer”) for each case.

The returned function pointer is of type `PFN_vkVoidFunction`, and must be cast to the type of the command being queried.

<table>
<thead>
<tr>
<th><code>instance</code></th>
<th><code>pName</code></th>
<th><code>return value</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>NULL</td>
<td>undefined</td>
</tr>
<tr>
<td>invalid instance</td>
<td>*</td>
<td>undefined</td>
</tr>
<tr>
<td>NULL</td>
<td><code>vkEnumerateInstanceExtensionProperties</code></td>
<td>fp</td>
</tr>
<tr>
<td>NULL</td>
<td><code>vkEnumerateInstanceLayerProperties</code></td>
<td>fp</td>
</tr>
<tr>
<td>NULL</td>
<td><code>vkCreateInstance</code></td>
<td>fp</td>
</tr>
<tr>
<td>NULL</td>
<td>* (any <code>pName</code> not covered above)</td>
<td>NULL</td>
</tr>
<tr>
<td>instance</td>
<td>core Vulkan command</td>
<td>fp</td>
</tr>
<tr>
<td>instance</td>
<td>enabled instance extension commands for <code>instance</code></td>
<td>fp</td>
</tr>
</tbody>
</table>

Table 1. `vkGetInstanceProcAddr` behavior
The returned function pointer **must** only be called with a dispatchable object (the first parameter) that is `instance` or a child of `instance`, e.g. `VkInstance`, `VkPhysicalDevice`, `VkDevice`, `VkQueue`, or `VkCommandBuffer`.

An “available device extension” is a device extension supported by any physical device enumerated by `instance`.

### Valid Usage (Implicit)

- If `instance` is not `NULL`, `instance` **must** be a valid `VkInstance` handle
- `pName` **must** be a null-terminated UTF-8 string

In order to support systems with multiple Vulkan implementations, the function pointers returned by `vkGetInstanceProcAddr` **may** point to dispatch code that calls a different real implementation for different `VkDevice` objects or their child objects. The overhead of the internal dispatch for `VkDevice` objects can be avoided by obtaining device-specific function pointers for any commands that use a device or device-child object as their dispatchable object. Such function pointers **can** be obtained with the command:

```c
PFN_vkVoidFunction vkGetDeviceProcAddr(VkDevice device, const char* pName);
```

The table below defines the various use cases for `vkGetDeviceProcAddr` and expected return value for each case.

The returned function pointer is of type `PFN_vkVoidFunction`, and must be cast to the type of the command being queried. The function pointer **must** only be called with a dispatchable object (the first parameter) that is `device` or a child of `device`.

**Table 2. vkGetDeviceProcAddr behavior**

<table>
<thead>
<tr>
<th>device</th>
<th>pName</th>
<th>return value</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>NULL</code></td>
<td>*</td>
<td>undefined</td>
</tr>
<tr>
<td>invalid device</td>
<td>*</td>
<td>undefined</td>
</tr>
<tr>
<td>device</td>
<td><code>NULL</code></td>
<td>undefined</td>
</tr>
</tbody>
</table>
### Valid Usage (Implicit)

- **device** must be a valid `VkDevice` handle
- **pName** must be a null-terminated UTF-8 string

The definition of `PFN_vkVoidFunction` is:

```c
typedef void (VKAPI_PTR *PFN_vkVoidFunction)(void);
```

## 3.2. Instances

There is no global state in Vulkan and all per-application state is stored in a `VkInstance` object. Creating a `VkInstance` object initializes the Vulkan library and allows the application to pass information about itself to the implementation.

Instances are represented by `VkInstance` handles:

```c
VK_DEFINE_HANDLE(VkInstance)
```

To create an instance object, call:

```c
VkResult vkCreateInstance(
    const VkInstanceCreateInfo* pCreateInfo,
    const VkAllocationCallbacks* pAllocator,
    VkInstance* pInstance);
```

- **pCreateInfo** points to an instance of `VkInstanceCreateInfo` controlling creation of the instance.
- **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.
- **pInstance** points a `VkInstance` handle in which the resulting instance is returned.

`vkCreateInstance` verifies that the requested layers exist. If not, `vkCreateInstance` will return `VK_ERROR_LAYER_NOT_PRESENT`. Next `vkCreateInstance` verifies that the requested extensions are supported (e.g. in the implementation or in any enabled instance layer) and if any requested extension is not supported, `vkCreateInstance` must return `VK_ERROR_EXTENSION_NOT_PRESENT`. After
verifying and enabling the instance layers and extensions the VkInstance object is created and returned to the application. If a requested extension is only supported by a layer, both the layer and the extension need to be specified at vkCreateInstance time for the creation to succeed.

**Valid Usage**

- All required extensions for each extension in the VkInstanceCreateInfo::ppEnabledExtensionNames list must also be present in that list.

**Valid Usage (Implicit)**

- pCreateInfo must be a valid pointer to a valid VkInstanceCreateInfo structure
- If pAllocator is not NULL, pAllocator must be a valid pointer to a valid VkAllocationCallbacks structure
- pInstance must be a valid pointer to a VkInstance handle

**Return Codes**

**Success**
- VK_SUCCESS

**Failure**
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY
- VK_ERROR_INITIALIZATION_FAILED
- VK_ERROR_LAYER_NOT_PRESENT
- VK_ERROR_EXTENSION_NOT_PRESENT
- VK_ERROR_INCOMPATIBLE_DRIVER

The VkInstanceCreateInfo structure is defined as:

```c
typedef struct VkInstanceCreateInfo {
    VkStructureType             sType;
    const void*                 pNext;
    VkInstanceCreateFlags       flags;
    const VkApplicationInfo*    pApplicationInfo;
    uint32_t                    enabledLayerCount;
    const char* const*          ppEnabledLayerNames;
    uint32_t                    enabledExtensionCount;
    const char* const*          ppEnabledExtensionNames;
} VkInstanceCreateInfo;
```

- sType is the type of this structure.
• `pNext` is `NULL` or a pointer to an extension-specific structure.

• `flags` is reserved for future use.

• `pApplicationInfo` is `NULL` or a pointer to an instance of `VkApplicationInfo`. If not `NULL`, this information helps implementations recognize behavior inherent to classes of applications. `VkApplicationInfo` is defined in detail below.

• `enabledLayerCount` is the number of global layers to enable.

• `ppEnabledLayerNames` is a pointer to an array of `enabledLayerCount` null-terminated UTF-8 strings containing the names of layers to enable for the created instance. See the Layers section for further details.

• `enabledExtensionCount` is the number of global extensions to enable.

• `ppEnabledExtensionNames` is a pointer to an array of `enabledExtensionCount` null-terminated UTF-8 strings containing the names of extensions to enable.

### Valid Usage (Implicit)

- `sType` must be `VK_STRUCTURE_TYPE_INSTANCE_CREATE_INFO`
- `pNext` must be `NULL`
- `flags` must be `0`
  - If `pApplicationInfo` is not `NULL`, `pApplicationInfo` must be a valid pointer to a valid `VkApplicationInfo` structure
  - If `enabledLayerCount` is not `0`, `ppEnabledLayerNames` must be a valid pointer to an array of `enabledLayerCount` null-terminated UTF-8 strings
  - If `enabledExtensionCount` is not `0`, `ppEnabledExtensionNames` must be a valid pointer to an array of `enabledExtensionCount` null-terminated UTF-8 strings

```c
typedef VkFlags VkInstanceCreateFlags;
```

`VkInstanceCreateFlags` is a bitmask type for setting a mask, but is currently reserved for future use.

The `VkApplicationInfo` structure is defined as:

```c
typedef struct VkApplicationInfo {
    VkStructureType    sType;
    const void*        pNext;
    const char*        pApplicationName;
    uint32_t           applicationVersion;
    const char*        pEngineName;
    uint32_t           engineVersion;
    uint32_t           apiVersion;
} VkApplicationInfo;
```
• **sType** is the type of this structure.

• **pNext** is **NULL** or a pointer to an extension-specific structure.

• **pApplicationName** is **NULL** or is a pointer to a null-terminated UTF-8 string containing the name of the application.

• **applicationVersion** is an unsigned integer variable containing the developer-supplied version number of the application.

• **pEngineName** is **NULL** or is a pointer to a null-terminated UTF-8 string containing the name of the engine (if any) used to create the application.

• **engineVersion** is an unsigned integer variable containing the developer-supplied version number of the engine used to create the application.

• **apiVersion** is the version of the Vulkan API against which the application expects to run, encoded as described in the API Version Numbers and Semantics section. If **apiVersion** is **0** the implementation **must** ignore it, otherwise if the implementation does not support the requested **apiVersion**, or an effective substitute for **apiVersion**, it **must** return **VK_ERROR_INCOMPATIBLE_DRIVER**. The patch version number specified in **apiVersion** is ignored when creating an instance object. Only the major and minor versions of the instance **must** match those requested in **apiVersion**.

---

### Valid Usage (Implicit)

- **sType** **must** be **VK_STRUCTURE_TYPE_APPLICATION_INFO**

- **pNext** **must** be **NULL**

- If **pApplicationName** is not **NULL**, **pApplicationName** **must** be a null-terminated UTF-8 string

- If **pEngineName** is not **NULL**, **pEngineName** **must** be a null-terminated UTF-8 string

To destroy an instance, call:

```c
void vkDestroyInstance(
    VkInstance                 instance,
    const VkAllocationCallbacks* pAllocator);
```

- **instance** is the handle of the instance to destroy.

- **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.
Valid Usage

- All child objects created using `instance` must have been destroyed prior to destroying `instance`.
- If `VkAllocationCallbacks` were provided when `instance` was created, a compatible set of callbacks must be provided here.
- If no `VkAllocationCallbacks` were provided when `instance` was created, `pAllocator` must be `NULL`.

Valid Usage (Implicit)

- If `instance` is not `NULL`, `instance` must be a valid `VkInstance` handle.
- If `pAllocator` is not `NULL`, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure.

Host Synchronization

- Host access to `instance` must be externally synchronized.
Chapter 4. Devices and Queues

Once Vulkan is initialized, devices and queues are the primary objects used to interact with a Vulkan implementation.

Vulkan separates the concept of physical and logical devices. A physical device usually represents a single complete implementation of Vulkan (excluding instance-level functionality) available to the host, of which there are a finite number. A logical device represents an instance of that implementation with its own state and resources independent of other logical devices.

Physical devices are represented by VkPhysicalDevice handles:

```
VK_DEFINE_HANDLE(VkPhysicalDevice)
```

4.1. Physical Devices

To retrieve a list of physical device objects representing the physical devices installed in the system, call:

```
VkResult vkEnumeratePhysicalDevices(
    VkInstance instance,
    uint32_t* pPhysicalDeviceCount,
    VkPhysicalDevice* pPhysicalDevices);
```

- `instance` is a handle to a Vulkan instance previously created with `vkCreateInstance`.
- `pPhysicalDeviceCount` is a pointer to an integer related to the number of physical devices available or queried, as described below.
- `pPhysicalDevices` is either `NULL` or a pointer to an array of `VkPhysicalDevice` handles.

If `pPhysicalDevices` is `NULL`, then the number of physical devices available is returned in `pPhysicalDeviceCount`. Otherwise, `pPhysicalDeviceCount` must point to a variable set by the user to the number of elements in the `pPhysicalDevices` array, and on return the variable is overwritten with the number of handles actually written to `pPhysicalDevices`. If `pPhysicalDeviceCount` is less than the number of physical devices available, `VK_INCOMPLETE` will be returned instead of `VK_SUCCESS`, to indicate that not all the available physical devices were returned.
Valid Usage (Implicit)

- **instance** must be a valid `VkInstance` handle
- **pPhysicalDeviceCount** must be a valid pointer to a `uint32_t` value
- If the value referenced by `pPhysicalDeviceCount` is not 0, and `pPhysicalDevices` is not NULL, `pPhysicalDevices` must be a valid pointer to an array of `pPhysicalDeviceCount` `VkPhysicalDevice` handles

Return Codes

**Success**

- VK_SUCCESS
- VK_INCOMPLETE

**Failure**

- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY
- VK_ERROR_INITIALIZATION_FAILED

To query general properties of physical devices once enumerated, call:

```c
void vkGetPhysicalDeviceProperties(
    VkPhysicalDevice                            physicalDevice,
    VkPhysicalDeviceProperties*                 pProperties);
```

- **physicalDevice** is the handle to the physical device whose properties will be queried.
- **pProperties** points to an instance of the `VkPhysicalDeviceProperties` structure, that will be filled with returned information.

Valid Usage (Implicit)

- **physicalDevice** must be a valid `VkPhysicalDevice` handle
- **pProperties** must be a valid pointer to a `VkPhysicalDeviceProperties` structure

The `VkPhysicalDeviceProperties` structure is defined as:
typedef struct VkPhysicalDeviceProperties {
    uint32_t                            apiVersion;
    uint32_t                            driverVersion;
    uint32_t                            vendorID;
    uint32_t                            deviceID;
    VkPhysicalDeviceType                deviceType;
    char                                 deviceName[VK_MAX_PHYSICAL_DEVICE_NAME_SIZE];
    uint8_t                             pipelineCacheUUID[VK_UUID_SIZE];
    VkPhysicalDeviceLimits              limits;
    VkPhysicalDeviceSparseProperties    sparseProperties;
} VkPhysicalDeviceProperties;

- `apiVersion` is the version of Vulkan supported by the device, encoded as described in the API Version Numbers and Semantics section.
- `driverVersion` is the vendor-specified version of the driver.
- `vendorID` is a unique identifier for the vendor (see below) of the physical device.
- `deviceID` is a unique identifier for the physical device among devices available from the vendor.
- `deviceType` is a `VkPhysicalDeviceType` specifying the type of device.
- `deviceName` is a null-terminated UTF-8 string containing the name of the device.
- `pipelineCacheUUID` is an array of size `VK_UUID_SIZE`, containing 8-bit values that represent a universally unique identifier for the device.
- `limits` is the `VkPhysicalDeviceLimits` structure which specifies device-specific limits of the physical device. See Limits for details.
- `sparseProperties` is the `VkPhysicalDeviceSparseProperties` structure which specifies various sparse related properties of the physical device. See Sparse Properties for details.

The `vendorID` and `deviceID` fields are provided to allow applications to adapt to device characteristics that are not adequately exposed by other Vulkan queries.

**Note**
These may include performance profiles, hardware errata, or other characteristics.

The `vendor` identified by `vendorID` is the entity responsible for the most salient characteristics of the underlying implementation of the `VkPhysicalDevice` being queried.

**Note**
For example, in the case of a discrete GPU implementation, this should be the GPU chipset vendor. In the case of a hardware accelerator integrated into a system-on-chip (SoC), this should be the supplier of the silicon IP used to create the accelerator.

If the vendor has a PCI vendor ID, the low 16 bits of `vendorID` must contain that PCI vendor ID, and the remaining bits must be set to zero. Otherwise, the value returned must be a valid Khronos...
vendor ID, obtained as described in the Vulkan Documentation and Extensions: Procedures and Conventions document in the section “Registering a Vendor ID with Khronos”. Khronos vendor IDs are allocated starting at 0x10000, to distinguish them from the PCI vendor ID namespace. Khronos vendor IDs are symbolically defined in the VkVendorId type.

The vendor is also responsible for the value returned in deviceID. If the implementation is driven primarily by a PCI device with a PCI device ID, the low 16 bits of deviceID must contain that PCI device ID, and the remaining bits must be set to zero. Otherwise, the choice of what values to return may be dictated by operating system or platform policies - but should uniquely identify both the device version and any major configuration options (for example, core count in the case of multicore devices).

Note
The same device ID should be used for all physical implementations of that device version and configuration. For example, all uses of a specific silicon IP GPU version and configuration should use the same device ID, even if those uses occur in different SoCs.

Khronos vendor IDs which may be returned in VkPhysicalDeviceProperties::vendorID are:

```c
typedef enum VkVendorId {
    VK_VENDOR_ID_VIV = 0x10001,
    VK_VENDOR_ID_VSI = 0x10002,
    VK_VENDOR_ID_KAZAN = 0x10003,
} VkVendorId;
```

Note
Khronos vendor IDs may be allocated by vendors at any time. Only the latest canonical versions of this Specification, of the corresponding vk.xml API Registry, and of the corresponding vulkan_core.h header file must contain all reserved Khronos vendor IDs.

Only Khronos vendor IDs are given symbolic names at present. PCI vendor IDs returned by the implementation can be looked up in the PCI-SIG database.

The physical device types which may be returned in VkPhysicalDeviceProperties::deviceType are:

```c
typedef enum VkPhysicalDeviceType {
    VK_PHYSICAL_DEVICE_TYPE_OTHER = 0,
    VK_PHYSICAL_DEVICE_TYPE_INTEGRATED_GPU = 1,
    VK_PHYSICAL_DEVICE_TYPE_DISCRETE_GPU = 2,
    VK_PHYSICAL_DEVICE_TYPE_VIRTUAL_GPU = 3,
    VK_PHYSICAL_DEVICE_TYPE_CPU = 4,
} VkPhysicalDeviceType;
```

- **VK_PHYSICAL_DEVICE_TYPE_OTHER** - the device does not match any other available types.
• **VK_PHYSICAL_DEVICE_TYPE_INTEGRATED_GPU** - the device is typically one embedded in or tightly coupled with the host.

• **VK_PHYSICAL_DEVICE_TYPE_DISCRETE_GPU** - the device is typically a separate processor connected to the host via an interlink.

• **VK_PHYSICAL_DEVICE_TYPE_VIRTUAL_GPU** - the device is typically a virtual node in a virtualization environment.

• **VK_PHYSICAL_DEVICE_TYPE_CPU** - the device is typically running on the same processors as the host.

The physical device type is advertised for informational purposes only, and does not directly affect the operation of the system. However, the device type may correlate with other advertised properties or capabilities of the system, such as how many memory heaps there are.

To query properties of queues available on a physical device, call:

```c
void vkGetPhysicalDeviceQueueFamilyProperties(
    VkPhysicalDevice physicalDevice,
    uint32_t* pQueueFamilyPropertyCount,
    VkQueueFamilyProperties* pQueueFamilyProperties);
```

- **physicalDevice** is the handle to the physical device whose properties will be queried.
- **pQueueFamilyPropertyCount** is a pointer to an integer related to the number of queue families available or queried, as described below.
- **pQueueFamilyProperties** is either NULL or a pointer to an array of VkQueueFamilyProperties structures.

If **pQueueFamilyProperties** is NULL, then the number of queue families available is returned in **pQueueFamilyPropertyCount**. Otherwise, **pQueueFamilyPropertyCount** must point to a variable set by the user to the number of elements in the **pQueueFamilyProperties** array, and on return the variable is overwritten with the number of structures actually written to **pQueueFamilyProperties**. If **pQueueFamilyPropertyCount** is less than the number of queue families available, at most **pQueueFamilyPropertyCount** structures will be written.

---

**Valid Usage (Implicit)**

- **physicalDevice** must be a valid VkPhysicalDevice handle
- **pQueueFamilyPropertyCount** must be a valid pointer to a uint32_t value
- If the value referenced by **pQueueFamilyPropertyCount** is not 0, and **pQueueFamilyProperties** is not NULL, **pQueueFamilyProperties** must be a valid pointer to an array of **pQueueFamilyPropertyCount** VkQueueFamilyProperties structures

The **VkQueueFamilyProperties** structure is defined as:
typedef struct VkQueueFamilyProperties {
    VkQueueFlags    queueFlags;
    uint32_t        queueCount;
    uint32_t        timestampValidBits;
    VkExtent3D      minImageTransferGranularity;
} VkQueueFamilyProperties;

- **queueFlags** is a bitmask of VkQueueFlagBits indicating capabilities of the queues in this queue family.
- **queueCount** is the unsigned integer count of queues in this queue family.
- **timestampValidBits** is the unsigned integer count of meaningful bits in the timestamps written via vkCmdWriteTimestamp. The valid range for the count is 36..64 bits, or a value of 0, indicating no support for timestamps. Bits outside the valid range are guaranteed to be zeros.
- **minImageTransferGranularity** is the minimum granularity supported for image transfer operations on the queues in this queue family.

The value returned in **minImageTransferGranularity** has a unit of compressed texel blocks for images having a block-compressed format, and a unit of texels otherwise.

Possible values of **minImageTransferGranularity** are:

- (0,0,0) which indicates that only whole mip levels **must** be transferred using the image transfer operations on the corresponding queues. In this case, the following restrictions apply to all offset and extent parameters of image transfer operations:
  - The x, y, and z members of a VkOffset3D parameter **must** always be zero.
  - The width, height, and depth members of a VkExtent3D parameter **must** always match the width, height, and depth of the image subresource corresponding to the parameter, respectively.

- (A_x, A_y, A_z) where A_x, A_y, and A_z are all integer powers of two. In this case the following restrictions apply to all image transfer operations:
  - x, y, and z of a VkOffset3D parameter **must** be integer multiples of A_x, A_y, and A_z, respectively.
  - width of a VkExtent3D parameter **must** be an integer multiple of A_x, or else x + width **must** equal the width of the image subresource corresponding to the parameter.
  - height of a VkExtent3D parameter **must** be an integer multiple of A_y, or else y + height **must** equal the height of the image subresource corresponding to the parameter.
  - depth of a VkExtent3D parameter **must** be an integer multiple of A_z, or else z + depth **must** equal the depth of the image subresource corresponding to the parameter.
  - If the format of the image corresponding to the parameters is one of the block-compressed formats then for the purposes of the above calculations the granularity **must** be scaled up by the compressed texel block dimensions.

Queues supporting graphics and/or compute operations **must** report (1,1,1) in **minImageTransferGranularity**, meaning that there are no additional restrictions on the granularity of
image transfer operations for these queues. Other queues supporting image transfer operations are only **required** to support whole mip level transfers, thus \texttt{minImageTransferGranularity} for queues belonging to such queue families **may** be \((0,0,0)\).

The **Device Memory** section describes memory properties queried from the physical device.

For physical device feature queries see the **Features** chapter.

Bits which **may** be set in \texttt{VkQueueFamilyProperties::queueFlags} indicating capabilities of queues in a queue family are:

```c
typedef enum VkQueueFlagBits {
    VK_QUEUE_GRAPHICS_BIT = 0x00000001,
    VK_QUEUE_COMPUTE_BIT = 0x00000002,
    VK_QUEUE_TRANSFER_BIT = 0x00000004,
    VK_QUEUE_SPARSE_BINDING_BIT = 0x00000008,
} VkQueueFlagBits;
```

- **VK_QUEUE_GRAPHICS_BIT** specifies that queues in this queue family support graphics operations.
- **VK_QUEUE_COMPUTE_BIT** specifies that queues in this queue family support compute operations.
- **VK_QUEUE_TRANSFER_BIT** specifies that queues in this queue family support transfer operations.
- **VK_QUEUE_SPARSE_BINDING_BIT** specifies that queues in this queue family support sparse memory management operations (see **Sparse Resources**). If any of the sparse resource features are enabled, then at least one queue family **must** support this bit.

If an implementation exposes any queue family that supports graphics operations, at least one queue family of at least one physical device exposed by the implementation **must** support both graphics and compute operations.

```
Note
All commands that are allowed on a queue that supports transfer operations are also allowed on a queue that supports either graphics or compute operations. Thus, if the capabilities of a queue family include **VK_QUEUE_GRAPHICS_BIT** or **VK_QUEUE_COMPUTE_BIT**, then reporting the **VK_QUEUE_TRANSFER_BIT** capability separately for that queue family is **optional**.
```

For further details see **Queues**.

```c
ttypedef VkFlags VkQueueFlags;
```

\texttt{VkQueueFlags} is a bitmask type for setting a mask of zero or more \texttt{VkQueueFlagBits}.

## 4.2. Devices

Device objects represent logical connections to physical devices. Each device exposes a number of **queue families** each having one or more **queues**. All queues in a queue family support the same
operations.

As described in Physical Devices, a Vulkan application will first query for all physical devices in a system. Each physical device can then be queried for its capabilities, including its queue and queue family properties. Once an acceptable physical device is identified, an application must create a corresponding logical device. An application must create a separate logical device for each physical device it will use. The created logical device is then the primary interface to the physical device.

How to enumerate the physical devices in a system and query those physical devices for their queue family properties is described in the Physical Device Enumeration section above.

### 4.2.1. Device Creation

Logical devices are represented by VkDevice handles:

```cpp
VK_DEFINE_HANDLE(VkDevice)
```

A logical device is created as a connection to a physical device. To create a logical device, call:

```cpp
VkResult vkCreateDevice(
    VkPhysicalDevice                            physicalDevice,
    const VkDeviceCreateInfo*                   pCreateInfo,
    const VkAllocationCallbacks*                pAllocator,
    VkDevice*                                   pDevice);
```

- physicalDevice must be one of the device handles returned from a call to `vkEnumeratePhysicalDevices` (see Physical Device Enumeration).
- pCreateInfo is a pointer to a VkDeviceCreateInfo structure containing information about how to create the device.
- pAllocator controls host memory allocation as described in the Memory Allocation chapter.
- pDevice points to a handle in which the created VkDevice is returned.

`vkCreateDevice` verifies that extensions and features requested in the `ppEnabledExtensionNames` and `pEnabledFeatures` members of pCreateInfo, respectively, are supported by the implementation. If any requested extension is not supported, `vkCreateDevice` must return VK_ERROR_EXTENSION_NOT_PRESENT. If any requested feature is not supported, `vkCreateDevice` must return VK_ERROR_FEATURE_NOT_PRESENT. Support for extensions can be checked before creating a device by querying `vkEnumerateDeviceExtensionProperties`. Support for features can similarly be checked by querying `vkGetPhysicalDeviceFeatures`.

After verifying and enabling the extensions the VkDevice object is created and returned to the application. If a requested extension is only supported by a layer, both the layer and the extension need to be specified at `vkCreateInstance` time for the creation to succeed.

Multiple logical devices can be created from the same physical device. Logical device creation may fail due to lack of device-specific resources (in addition to the other errors). If that occurs,
vkCreateDevice will return VK_ERROR_TOO_MANY_OBJECTS.

Valid Usage

- All required extensions for each extension in the VkDeviceCreateInfo::ppEnabledExtensionNames list must also be present in that list.

Valid Usage (Implicit)

- physicalDevice must be a valid VkPhysicalDevice handle
- pCreateInfo must be a valid pointer to a valid VkDeviceCreateInfo structure
- If pAllocator is not NULL, pAllocator must be a valid pointer to a valid VkAllocationCallbacks structure
- pDevice must be a valid pointer to a VkDevice handle

Return Codes

Success
- VK_SUCCESS

Failure
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY
- VK_ERROR_INITIALIZATION_FAILED
- VK_ERROR_EXTENSION_NOT_PRESENT
- VK_ERROR_FEATURE_NOT_PRESENT
- VK_ERROR_TOO_MANY_OBJECTS
- VK_ERROR_DEVICE_LOST

The VkDeviceCreateInfo structure is defined as:

```c
typedef struct VkDeviceCreateInfo {
    VkStructureType                   sType;
    const void*                       pNext;
    VkDeviceCreateFlags               flags;
    uint32_t                           queueCreateInfoCount;
    const VkDeviceQueueCreateInfo*    pQueueCreateInfos;
    uint32_t                           enabledLayerCount;
    const char* const*                 ppEnabledLayerNames;
    uint32_t                           enabledExtensionCount;
    const char* const*                 ppEnabledExtensionNames;
    const VkPhysicalDeviceFeatures*    pEnabledFeatures;
} VkDeviceCreateInfo;
```
• **sType** is the type of this structure.

• **pNext** is NULL or a pointer to an extension-specific structure.

• **flags** is reserved for future use.

• **queueCreateInfoCount** is the unsigned integer size of the **pQueueCreateInfos** array. Refer to the Queue Creation section below for further details.

• **pQueueCreateInfos** is a pointer to an array of **VkDeviceQueueCreateInfo** structures describing the queues that are requested to be created along with the logical device. Refer to the Queue Creation section below for further details.

• **enabledLayerCount** is deprecated and ignored.

• **ppEnabledLayerNames** is deprecated and ignored. See Device Layer Deprecation.

• **enabledExtensionCount** is the number of device extensions to enable.

• **ppEnabledExtensionNames** is a pointer to an array of **enabledExtensionCount** null-terminated UTF-8 strings containing the names of extensions to enable for the created device. See the Extensions section for further details.

• **pEnabledFeatures** is NULL or a pointer to a **VkPhysicalDeviceFeatures** structure that contains boolean indicators of all the features to be enabled. Refer to the Features section for further details.

### Valid Usage

- The **queueFamilyIndex** member of each element of **pQueueCreateInfos** must be unique within **pQueueCreateInfos**

### Valid Usage (Implicit)

- **sType** must be **VK_STRUCTURE_TYPE_DEVICE_CREATE_INFO**
- **pNext** must be NULL
- **flags** must be 0
- **pQueueCreateInfos** must be a valid pointer to an array of **queueCreateInfoCount** valid **VkDeviceQueueCreateInfo** structures
- If **enabledLayerCount** is not 0, **ppEnabledLayerNames** must be a valid pointer to an array of **enabledLayerCount** null-terminated UTF-8 strings
- If **enabledExtensionCount** is not 0, **ppEnabledExtensionNames** must be a valid pointer to an array of **enabledExtensionCount** null-terminated UTF-8 strings
- If **pEnabledFeatures** is not NULL, **pEnabledFeatures** must be a valid pointer to a valid **VkPhysicalDeviceFeatures** structure
- **queueCreateInfoCount** must be greater than 0
**typedef** VkFlags VkDeviceCreateFlags;

\texttt{VkDeviceCreateFlags} is a bitmask type for setting a mask, but is currently reserved for future use.

### 4.2.2. Device Use

The following is a high-level list of \texttt{VkDevice} uses along with references on where to find more information:

- Creation of queues. See the \texttt{Queues} section below for further details.
- Creation and tracking of various synchronization constructs. See \texttt{Synchronization and Cache Control} for further details.
- Allocating, freeing, and managing memory. See \texttt{Memory Allocation} and \texttt{Resource Creation} for further details.
- Creation and destruction of command buffers and command buffer pools. See \texttt{Command Buffers} for further details.
- Creation, destruction, and management of graphics state. See \texttt{Pipelines} and \texttt{Resource Descriptors}, among others, for further details.

### 4.2.3. Lost Device

A logical device \textbf{may} become \textit{lost} for a number of implementation-specific reasons, indicating that pending and future command execution \textbf{may} fail and cause resources and backing memory to become undefined.

\begin{quote}
\textbf{Note}

Typical reasons for device loss will include things like execution timing out (to prevent denial of service), power management events, platform resource management, or implementation errors.
\end{quote}

When this happens, certain commands will return \texttt{VK_ERROR_DEVICE_LOST} (see \texttt{Error Codes} for a list of such commands). After any such event, the logical device is considered \textit{lost}. It is not possible to reset the logical device to a non-lost state, however the lost state is specific to a logical device (\texttt{VkDevice}), and the corresponding physical device (\texttt{VkPhysicalDevice}) \textbf{may} be otherwise unaffected.

In some cases, the physical device \textbf{may} also be lost, and attempting to create a new logical device will fail, returning \texttt{VK_ERROR_DEVICE_LOST}. This is usually indicative of a problem with the underlying implementation, or its connection to the host. If the physical device has not been lost, and a new logical device is successfully created from that physical device, it \textbf{must} be in the non-lost state.
Whilst logical device loss may be recoverable, in the case of physical device loss, it is unlikely that an application will be able to recover unless additional, unaffected physical devices exist on the system. The error is largely informational and intended only to inform the user that a platform issue has occurred, and should be investigated further. For example, underlying hardware may have developed a fault or become physically disconnected from the rest of the system. In many cases, physical device loss may cause other more serious issues such as the operating system crashing; in which case it may not be reported via the Vulkan API.

Undefined behavior caused by an application error may cause a device to become lost. However, such undefined behavior may also cause unrecoverable damage to the process, and it is then not guaranteed that the API objects, including the VkPhysicalDevice or the VkInstance are still valid or that the error is recoverable.

When a device is lost, its child objects are not implicitly destroyed and their handles are still valid. Those objects must still be destroyed before their parents or the device can be destroyed (see the Object Lifetime section). The host address space corresponding to device memory mapped using vkMapMemory is still valid, and host memory accesses to these mapped regions are still valid, but the contents are undefined. It is still legal to call any API command on the device and child objects.

Once a device is lost, command execution may fail, and commands that return a VkResult may return VK_ERROR_DEVICE_LOST. Commands that do not allow run-time errors must still operate correctly for valid usage and, if applicable, return valid data.

Commands that wait indefinitely for device execution (namely vkDeviceWaitIdle, vkQueueWaitIdle, vkWaitForFences with a maximum timeout, and vkGetQueryPoolResults with the VK_QUERY_RESULT_WAIT_BIT bit set in flags) must return in finite time even in the case of a lost device, and return either VK_SUCCESS or VK_ERROR_DEVICE_LOST. For any command that may return VK_ERROR_DEVICE_LOST, for the purpose of determining whether a command buffer is in the pending state, or whether resources are considered in-use by the device, a return value of VK_ERROR_DEVICE_LOST is equivalent to VK_SUCCESS.

### 4.2.4. Device Destruction

To destroy a device, call:

```c
void vkDestroyDevice(
    VkDevice device, 
    const VkAllocationCallbacks* pAllocator); 
```

- `device` is the logical device to destroy.
- `pAllocator` controls host memory allocation as described in the Memory Allocation chapter.
To ensure that no work is active on the device, `vkDeviceWaitIdle` can be used to gate the destruction of the device. Prior to destroying a device, an application is responsible for destroying/freeing any Vulkan objects that were created using that device as the first parameter of the corresponding `vkCreate*` or `vkAllocate*` command.

**Note**
The lifetime of each of these objects is bound by the lifetime of the `VkDevice` object. Therefore, to avoid resource leaks, it is critical that an application explicitly free all of these resources prior to calling `vkDestroyDevice`.

### Valid Usage
- All child objects created on `device` must have been destroyed prior to destroying `device`
- If `VkAllocationCallbacks` were provided when `device` was created, a compatible set of callbacks must be provided here
- If no `VkAllocationCallbacks` were provided when `device` was created, `pAllocator` must be `NULL`

### Valid Usage (Implicit)
- If `device` is not `NULL`, `device` must be a valid `VkDevice` handle
- If `pAllocator` is not `NULL`, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure

### Host Synchronization
- Host access to `device` must be externally synchronized

## 4.3. Queues

### 4.3.1. Queue Family Properties
As discussed in the Physical Device Enumeration section above, the `vkGetPhysicalDeviceQueueFamilyProperties` command is used to retrieve details about the queue families and queues supported by a device.

Each index in the `pQueueFamilyProperties` array returned by `vkGetPhysicalDeviceQueueFamilyProperties` describes a unique queue family on that physical device. These indices are used when creating queues, and they correspond directly with the `queueFamilyIndex` that is passed to the `vkCreateDevice` command via the `VkDeviceQueueCreateInfo` structure as described in the Queue Creation section below.

Grouping of queue families within a physical device is implementation-dependent.
The general expectation is that a physical device groups all queues of matching capabilities into a single family. However, while implementations should do this, it is possible that a physical device may return two separate queue families with the same capabilities.

Once an application has identified a physical device with the queue(s) that it desires to use, it will create those queues in conjunction with a logical device. This is described in the following section.

### 4.3.2. Queue Creation

Creating a logical device also creates the queues associated with that device. The queues to create are described by a set of `VkDeviceQueueCreateInfo` structures that are passed to `vkCreateDevice` in `pQueueCreateInfos`.

Queues are represented by `VkQueue` handles:

```
VK_DEFINE_HANDLE(VkQueue)
```

The `VkDeviceQueueCreateInfo` structure is defined as:

```c
typedef struct VkDeviceQueueCreateInfo {
    VkStructureType             sType;
    const void*                 pNext;
    VkDeviceQueueCreateFlags    flags;
    uint32_t                    queueFamilyIndex;
    uint32_t                    queueCount;
    const float*                pQueuePriorities;
} VkDeviceQueueCreateInfo;
```

- **sType** is the type of this structure.
- **pNext** is `NULL` or a pointer to an extension-specific structure.
- **flags** is reserved for future use.
- **queueFamilyIndex** is an unsigned integer indicating the index of the queue family to create on this device. This index corresponds to the index of an element of the `pQueueFamilyProperties` array that was returned by `vkGetPhysicalDeviceQueueFamilyProperties`.
- **queueCount** is an unsigned integer specifying the number of queues to create in the queue family indicated by `queueFamilyIndex`.
- **pQueuePriorities** is an array of `queueCount` normalized floating point values, specifying priorities of work that will be submitted to each created queue. See Queue Priority for more information.
**Valid Usage**

- `queueFamilyIndex` must be less than `pQueueFamilyPropertyCount` returned by `vkGetPhysicalDeviceQueueFamilyProperties`
- `queueCount` must be less than or equal to the `queueCount` member of the `VkQueueFamilyProperties` structure, as returned by `vkGetPhysicalDeviceQueueFamilyProperties` in the `pQueueFamilyProperties[queueFamilyIndex]`
- Each element of `pQueuePriorities` must be between 0.0 and 1.0 inclusive

**Valid Usage (Implicit)**

- `sType` must be `VK_STRUCTURE_TYPE_DEVICE_QUEUE_CREATE_INFO`
- `pNext` must be `NULL`
- `flags` must be a valid combination of `VkDeviceQueueCreateFlagBits` values
- `pQueuePriorities` must be a valid pointer to an array of `queueCount` `float` values
- `queueCount` must be greater than 0

```c
typedef VkFlags VkDeviceQueueCreateFlags;
```

`VkDeviceQueueCreateFlags` is a bitmask type for setting a mask of zero or more `VkDeviceQueueCreateFlagBits`.

To retrieve a handle to a `VkQueue` object, call:

```c
void vkGetDeviceQueue(
    VkDevice device,
    uint32_t queueFamilyIndex,
    uint32_t queueIndex,
    VkQueue* pQueue);
```

- `device` is the logical device that owns the queue.
- `queueFamilyIndex` is the index of the queue family to which the queue belongs.
- `queueIndex` is the index within this queue family of the queue to retrieve.
- `pQueue` is a pointer to a `VkQueue` object that will be filled with the handle for the requested queue.
Valid Usage

- `queueFamilyIndex must` be one of the queue family indices specified when `device` was created, via the `VkDeviceQueueCreateInfo` structure
- `queueIndex must` be less than the number of queues created for the specified queue family index when `device` was created, via the `queueCount` member of the `VkDeviceQueueCreateInfo` structure
- `VkDeviceQueueCreateInfo::flags must` have been set to zero when `device` was created

Valid Usage (Implicit)

- `device must` be a valid `VkDevice` handle
- `pQueue must` be a valid pointer to a `VkQueue` handle

4.3.3. Queue Family Index

The queue family index is used in multiple places in Vulkan in order to tie operations to a specific family of queues.

When retrieving a handle to the queue via `vkGetDeviceQueue`, the queue family index is used to select which queue family to retrieve the `VkQueue` handle from as described in the previous section.

When creating a `VkCommandPool` object (see Command Pools), a queue family index is specified in the `VkCommandPoolCreateInfo` structure. Command buffers from this pool can only be submitted on queues corresponding to this queue family.

When creating `VkImage` (see Images) and `VkBuffer` (see Buffers) resources, a set of queue families is included in the `VkImageCreateInfo` and `VkBufferCreateInfo` structures to specify the queue families that can access the resource.

When inserting a `VkBufferMemoryBarrier` or `VkImageMemoryBarrier` (see Events) a source and destination queue family index is specified to allow the ownership of a buffer or image to be transferred from one queue family to another. See the Resource Sharing section for details.

4.3.4. Queue Priority

Each queue is assigned a priority, as set in the `VkDeviceQueueCreateInfo` structures when creating the device. The priority of each queue is a normalized floating point value between 0.0 and 1.0, which is then translated to a discrete priority level by the implementation. Higher values indicate a higher priority, with 0.0 being the lowest priority and 1.0 being the highest.

Within the same device, queues with higher priority may be allotted more processing time than queues with lower priority. The implementation makes no guarantees with regards to ordering or scheduling among queues with the same priority, other than the constraints defined by any explicit synchronization primitives. The implementation make no guarantees with regards to queues across
An implementation may allow a higher-priority queue to starve a lower-priority queue on the same VkDevice until the higher-priority queue has no further commands to execute. The relationship of queue priorities must not cause queues on one VkDevice to starve queues on another VkDevice.

No specific guarantees are made about higher priority queues receiving more processing time or better quality of service than lower priority queues.

### 4.3.5. Queue Submission

Work is submitted to a queue via *queue submission* commands such as `vkQueueSubmit`. Queue submission commands define a set of *queue operations* to be executed by the underlying physical device, including synchronization with semaphores and fences.

Submission commands take as parameters a target queue, zero or more *batches* of work, and an *optional* fence to signal upon completion. Each batch consists of three distinct parts:

1. Zero or more semaphores to wait on before execution of the rest of the batch.
   - If present, these describe a *semaphore wait operation*.
2. Zero or more work items to execute.
   - If present, these describe a *queue operation* matching the work described.
3. Zero or more semaphores to signal upon completion of the work items.
   - If present, these describe a *semaphore signal operation*.

If a fence is present in a queue submission, it describes a *fence signal operation*.

All work described by a queue submission command must be submitted to the queue before the command returns.

#### Sparse Memory Binding

In Vulkan it is possible to sparsely bind memory to buffers and images as described in the *Sparse Resource* chapter. Sparse memory binding is a queue operation. A queue whose flags include the `VK_QUEUE_SPARSE_BINDING_BIT` must be able to support the mapping of a virtual address to a physical address on the device. This causes an update to the page table mappings on the device. This update must be synchronized on a queue to avoid corrupting page table mappings during execution of graphics commands. By binding the sparse memory resources on queues, all commands that are dependent on the updated bindings are synchronized to only execute after the binding is updated. See the *Synchronization and Cache Control* chapter for how this synchronization is accomplished.

### 4.3.6. Queue Destruction

Queues are created along with a logical device during `vkCreateDevice`. All queues associated with a logical device are destroyed when `vkDestroyDevice` is called on that device.
Chapter 5. Command Buffers

Command buffers are objects used to record commands which can be subsequently submitted to a device queue for execution. There are two levels of command buffers - primary command buffers, which can execute secondary command buffers, and which are submitted to queues, and secondary command buffers, which can be executed by primary command buffers, and which are not directly submitted to queues.

Command buffers are represented by VkCommandBuffer handles:

```c
VK_DEFINE_HANDLE(VkCommandBuffer)
```

Recorded commands include commands to bind pipelines and descriptor sets to the command buffer, commands to modify dynamic state, commands to draw (for graphics rendering), commands to dispatch (for compute), commands to execute secondary command buffers (for primary command buffers only), commands to copy buffers and images, and other commands.

Each command buffer manages state independently of other command buffers. There is no inheritance of state across primary and secondary command buffers, or between secondary command buffers. When a command buffer begins recording, all state in that command buffer is undefined. When secondary command buffer(s) are recorded to execute on a primary command buffer, the secondary command buffer inherits no state from the primary command buffer, and all state of the primary command buffer is undefined after an execute secondary command buffer command is recorded. There is one exception to this rule - if the primary command buffer is inside a render pass instance, then the render pass and subpass state is not disturbed by executing secondary command buffers. Whenever the state of a command buffer is undefined, the application must set all relevant state on the command buffer before any state dependent commands such as draws and dispatches are recorded, otherwise the behavior of executing that command buffer is undefined.

Unless otherwise specified, and without explicit synchronization, the various commands submitted to a queue via command buffers may execute in arbitrary order relative to each other, and/or concurrently. Also, the memory side-effects of those commands may not be directly visible to other commands without explicit memory dependencies. This is true within a command buffer, and across command buffers submitted to a given queue. See the synchronization chapter for information on implicit and explicit synchronization between commands.

5.1. Command Buffer Lifecycle

Each command buffer is always in one of the following states:

**Initial**

When a command buffer is allocated, it is in the initial state. Some commands are able to reset a command buffer, or a set of command buffers, back to this state from any of the executable, recording or invalid state. Command buffers in the initial state can only be moved to the recording state, or freed.
**Recording**

`vkBeginCommandBuffer` changes the state of a command buffer from the initial state to the *recording state*. Once a command buffer is in the recording state, `vkCmd*` commands *can* be used to record to the command buffer.

**Executable**

`vkEndCommandBuffer` ends the recording of a command buffer, and moves it from the recording state to the *executable state*. Executable command buffers *can* be submitted, reset, or recorded to another command buffer.

**Pending**

*Queue submission* of a command buffer changes the state of a command buffer from the executable state to the *pending state*. Whilst in the pending state, applications *must* not attempt to modify the command buffer in any way - as the device *may* be processing the commands recorded to it. Once execution of a command buffer completes, the command buffer reverts back to either the *executable state*, or the *invalid state* if it was recorded with `VK_COMMAND_BUFFER_USAGE_ONE_TIME_SUBMIT_BIT`. A synchronization command *should* be used to detect when this occurs.

**Invalid**

Some operations, such as modifying or deleting a resource that was used in a command recorded to a command buffer, will transition the state of that command buffer into the *invalid state*. Command buffers in the invalid state *can* only be reset or freed.

![Figure 1. Lifecycle of a command buffer](image-url)

*Any given command that operates on a command buffer has its own requirements on what state a command buffer *must* be in, which are detailed in the valid usage constraints for that command.*

Resetting a command buffer is an operation that discards any previously recorded commands and puts a command buffer in the initial state. Resetting occurs as a result of `vkResetCommandBuffer` or `vkResetCommandPool`, or as part of `vkBeginCommandBuffer` (which additionally puts the command buffer in the recording state).

*Secondary command buffers *can* be recorded to a primary command buffer via `vkCmdExecuteCommands`. This partially ties the lifecycle of the two command buffers together - if*
the primary is submitted to a queue, both the primary and any secondaries recorded to it move to
the pending state. Once execution of the primary completes, so does any secondary recorded within
it, and once all executions of each command buffer complete, they move to the executable state. If a
secondary moves to any other state whilst it is recorded to another command buffer, the primary
moves to the invalid state. A primary moving to any other state does not affect the state of the
secondary. Resetting or freeing a primary command buffer removes the linkage to any secondary
command buffers that were recorded to it.

5.2. Command Pools

Command pools are opaque objects that command buffer memory is allocated from, and which
allow the implementation to amortize the cost of resource creation across multiple command
buffers. Command pools are externally synchronized, meaning that a command pool must not be
used concurrently in multiple threads. That includes use via recording commands on any
command buffers allocated from the pool, as well as operations that allocate, free, and reset
command buffers or the pool itself.

Command pools are represented by VkCommandPool handles:

```
VK_DEFINE_NON_DISPATCHABLE_HANDLE(VkCommandPool)
```

To create a command pool, call:

```
VkResult vkCreateCommandPool(
    VkDevice                                    device,
    const VkCommandPoolCreateInfo*              pCreateInfo,
    const VkAllocationCallbacks*                pAllocator,
    VkCommandPool*                              pCommandPool);
```

- `device` is the logical device that creates the command pool.
- `pCreateInfo` is a pointer to an instance of the VkCommandPoolCreateInfo structure specifying
  the state of the command pool object.
- `pAllocator` controls host memory allocation as described in the Memory Allocation chapter.
- `pCommandPool` points to a VkCommandPool handle in which the created pool is returned.

Valid Usage

- `pCreateInfo::queueFamilyIndex` must be the index of a queue family available in the logical
device `device`.
Valid Usage (Implicit)

- **device** must be a valid `VkDevice` handle
- **pCreateInfo** must be a valid pointer to a valid `VkCommandPoolCreateInfo` structure
- If **pAllocator** is not `NULL`, **pAllocator** must be a valid pointer to a valid `VkAllocationCallbacks` structure
- **pCommandPool** must be a valid pointer to a `VkCommandPool` handle

Return Codes

**Success**
- `VK_SUCCESS`

**Failure**
- `VK_ERROR_OUT_OF_HOST_MEMORY`
- `VK_ERROR_OUT_OF_DEVICE_MEMORY`

The `VkCommandPoolCreateInfo` structure is defined as:

```c
typedef struct VkCommandPoolCreateInfo {
    VkStructureType          sType;
    const void*              pNext;
    VkCommandPoolCreateFlags flags;
    uint32_t                  queueFamilyIndex;
} VkCommandPoolCreateInfo;
```

- **sType** is the type of this structure.
- **pNext** is `NULL` or a pointer to an extension-specific structure.
- **flags** is a bitmask of `VkCommandPoolCreateFlagBits` indicating usage behavior for the pool and command buffers allocated from it.
- **queueFamilyIndex** designates a queue family as described in section Queue Family Properties. All command buffers allocated from this command pool must be submitted on queues from the same queue family.

Valid Usage (Implicit)

- **sType** must be `VK_STRUCTURE_TYPE_COMMAND_POOL_CREATE_INFO`
- **pNext** must be `NULL`
- **flags** must be a valid combination of `VkCommandPoolCreateFlagBits` values

Bits which can be set in `VkCommandPoolCreateInfo::flags` to specify usage behavior for a
command pool are:

typedef enum VkCommandPoolCreateFlagBits {
    VK_COMMAND_POOL_CREATE_TRANSIENT_BIT = 0x00000001,
    VK_COMMAND_POOL_CREATE_RESET_COMMAND_BUFFER_BIT = 0x00000002,
} VkCommandPoolCreateFlagBits;

- **VK_COMMAND_POOL_CREATE_TRANSIENT_BIT** specifies that command buffers allocated from the pool will be short-lived, meaning that they will be reset or freed in a relatively short timeframe. This flag may be used by the implementation to control memory allocation behavior within the pool.

- **VK_COMMAND_POOL_CREATE_RESET_COMMAND_BUFFER_BIT** allows any command buffer allocated from a pool to be individually reset to the initial state; either by calling `vkResetCommandBuffer`, or via the implicit reset when calling `vkBeginCommandBuffer`. If this flag is not set on a pool, then `vkResetCommandBuffer` must not be called for any command buffer allocated from that pool.

typedef VkFlags VkCommandPoolCreateFlags;

**VkCommandPoolCreateFlags** is a bitmask type for setting a mask of zero or more `VkCommandPoolCreateFlagBits`.

To reset a command pool, call:

```c
VkResult vkResetCommandPool(
    VkDevice device,
    VkCommandPool commandPool,
    VkCommandPoolResetFlags flags);
```

- **device** is the logical device that owns the command pool.
- **commandPool** is the command pool to reset.
- **flags** is a bitmask of `VkCommandPoolResetFlagBits` controlling the reset operation.

Resetting a command pool recycles all of the resources from all of the command buffers allocated from the command pool back to the command pool. All command buffers that have been allocated from the command pool are put in the initial state.

Any primary command buffer allocated from another `VkCommandPool` that is in the recording or executable state and has a secondary command buffer allocated from `commandPool` recorded into it, becomes invalid.

**Valid Usage**

- All `VkCommandBuffer` objects allocated from `commandPool` must not be in the pending state
Valid Usage (Implicit)

- `device` **must** be a valid `VkDevice` handle
- `commandPool` **must** be a valid `VkCommandPool` handle
- `flags` **must** be a valid combination of `VkCommandPoolResetFlagBits` values
- `commandPool` **must** have been created, allocated, or retrieved from `device`

Host Synchronization

- Host access to `commandPool` **must** be externally synchronized

Return Codes

**Success**
- `VK_SUCCESS`

**Failure**
- `VK_ERROR_OUT_OF_HOST_MEMORY`
- `VK_ERROR_OUT_OF_DEVICE_MEMORY`

Bits which **can** be set in `vkResetCommandPool::flags` to control the reset operation are:

```c
typedef enum VkCommandPoolResetFlagBits {
    VK_COMMAND_POOL_RESET_RELEASE_RESOURCES_BIT = 0x00000001,
} VkCommandPoolResetFlagBits;
```

- `VK_COMMAND_POOL_RESET_RELEASE_RESOURCES_BIT` specifies that resetting a command pool recycles all of the resources from the command pool back to the system.

```c
typedef VkFlags VkCommandPoolResetFlags;
```

`VkCommandPoolResetFlags` is a bitmask type for setting a mask of zero or more `VkCommandPoolResetFlagBits`.

To destroy a command pool, call:

```c
void vkDestroyCommandPool(
    VkDevice                                    device, 
    VkCommandPool                               commandPool, 
    const VkAllocationCallbacks*                pAllocator);
```
device is the logical device that destroys the command pool.

commandPool is the handle of the command pool to destroy.

pAllocator controls host memory allocation as described in the Memory Allocation chapter.

When a pool is destroyed, all command buffers allocated from the pool are freed.

Any primary command buffer allocated from another VkCommandPool that is in the recording or executable state and has a secondary command buffer allocated from commandPool recorded into it, becomes invalid.

Valid Usage

- All VkCommandBuffer objects allocated from commandPool must not be in the pending state.
- If VkAllocationCallbacks were provided when commandPool was created, a compatible set of callbacks must be provided here.
- If no VkAllocationCallbacks were provided when commandPool was created, pAllocator must be NULL.

Valid Usage (Implicit)

- device must be a valid VkDevice handle.
- If commandPool is not VK_NULL_HANDLE, commandPool must be a valid VkCommandPool handle.
- If pAllocator is not NULL, pAllocator must be a valid pointer to a valid VkAllocationCallbacks structure.
- If commandPool is a valid handle, it must have been created, allocated, or retrieved from device.

Host Synchronization

- Host access to commandPool must be externally synchronized.

5.3. Command Buffer Allocation and Management

To allocate command buffers, call:

```c
VkResult vkAllocateCommandBuffers(
    VkDevice
    const VkCommandBufferAllocateInfo*          pAllocateInfo,
    VkCommandBuffer*                            pCommandBuffers);
```

- device is the logical device that owns the command pool.
• *pAllocateInfo* is a pointer to an instance of the *VkCommandBufferAllocateInfo* structure describing parameters of the allocation.

• *pCommandBuffers* is a pointer to an array of *VkCommandBuffer* handles in which the resulting command buffer objects are returned. The array **must** be at least the length specified by the *commandBufferCount* member of *pAllocateInfo*. Each allocated command buffer begins in the initial state.

When command buffers are first allocated, they are in the initial state.

### Valid Usage (Implicit)

- **device** **must** be a valid *VkDevice* handle
- **pAllocateInfo** **must** be a valid pointer to a valid *VkCommandBufferAllocateInfo* structure
- **pCommandBuffers** **must** be a valid pointer to an array of *pAllocateInfo::*commandBufferCount* *VkCommandBuffer* handles

### Host Synchronization

- Host access to *pAllocateInfo::*commandPool* **must** be externally synchronized

### Return Codes

**Success**

- **VK_SUCCESS**

**Failure**

- **VK_ERROR_OUT_OF_HOST_MEMORY**
- **VK_ERROR_OUT_OF_DEVICE_MEMORY**

The *VkCommandBufferAllocateInfo* structure is defined as:

```c
typedef struct VkCommandBufferAllocateInfo {
    VkStructureType         sType;
    const void*             pNext;
    VkCommandPool           commandPool;
    VkCommandBufferLevel    level;
    uint32_t                commandBufferCount;
} VkCommandBufferAllocateInfo;
```

• **sType** is the type of this structure.

• **pNext** is **NULL** or a pointer to an extension-specific structure.

• **commandPool** is the command pool from which the command buffers are allocated.
• **level** is a `VkCommandBufferLevel` value specifying the command buffer level.

• **commandBufferCount** is the number of command buffers to allocate from the pool.

### Valid Usage

- **commandBufferCount** must be greater than 0

### Valid Usage (Implicit)

- **sType** must be `VK_STRUCTURE_TYPE_COMMAND_BUFFER_ALLOCATE_INFO`
- **pNext** must be NULL
- **commandPool** must be a valid `VkCommandPool` handle
- **level** must be a valid `VkCommandBufferLevel` value

Possible values of `VkCommandBufferAllocateInfo::level`, specifying the command buffer level, are:

```c
typedef enum VkCommandBufferLevel {
    VK_COMMAND_BUFFER_LEVEL_PRIMARY = 0,
    VK_COMMAND_BUFFER_LEVEL_SECONDARY = 1,
} VkCommandBufferLevel;
```

- **VK_COMMAND_BUFFER_LEVEL_PRIMARY** specifies a primary command buffer.
- **VK_COMMAND_BUFFER_LEVEL_SECONDARY** specifies a secondary command buffer.

To reset command buffers, call:

```c
VkResult vkResetCommandBuffer(
    VkCommandBuffer commandBuffer,
    VkCommandBufferResetFlags flags);
```

- **commandBuffer** is the command buffer to reset. The command buffer can be in any state other than pending, and is moved into the initial state.

- **flags** is a bitmask of `VkCommandBufferResetFlagBits` controlling the reset operation.

Any primary command buffer that is in the recording or executable state and has `commandBuffer` recorded into it, becomes invalid.
Valid Usage

- `commandBuffer` **must** not be in the pending state
- `commandBuffer` **must** have been allocated from a pool that was created with the `VK_COMMAND_POOL_CREATE_RESET_COMMAND_BUFFER_BIT`

Valid Usage (Implicit)

- `commandBuffer` **must** be a valid `VkCommandBuffer` handle
- flags **must** be a valid combination of `VkCommandBufferResetFlagBits` values

Host Synchronization

- Host access to `commandBuffer` **must** be externally synchronized

Return Codes

**Success**
- `VK_SUCCESS`

**Failure**
- `VK_ERROR_OUT_OF_HOST_MEMORY`
- `VK_ERROR_OUT_OF_DEVICE_MEMORY`

Bits which **can** be set in `vkResetCommandBuffer::flags` to control the reset operation are:

```c
typedef enum VkCommandBufferResetFlagBits {
    VK_COMMAND_BUFFER_RESET_RELEASE_RESOURCES_BIT = 0x00000001,
} VkCommandBufferResetFlagBits;
```

- `VK_COMMAND_BUFFER_RESET_RELEASE_RESOURCES_BIT` specifies that most or all memory resources currently owned by the command buffer **should** be returned to the parent command pool. If this flag is not set, then the command buffer **may** hold onto memory resources and reuse them when recording commands. `commandBuffer` is moved to the initial state.

```c
typedef VkFlags VkCommandBufferResetFlags;
```

`VkCommandBufferResetFlags` is a bitmask type for setting a mask of zero or more `VkCommandBufferResetFlagBits`.

To free command buffers, call:
void vkFreeCommandBuffers(
    VkDevice                                    device,
    VkCommandPool                               commandPool,
    uint32_t                                    commandBufferCount,
    const VkCommandBuffer*                      pCommandBuffers);

- device is the logical device that owns the command pool.
- commandPool is the command pool from which the command buffers were allocated.
- commandBufferCount is the length of the pCommandBuffers array.
- pCommandBuffers is an array of handles of command buffers to free.

Any primary command buffer that is in the recording or executable state and has any element of pCommandBuffers recorded into it, becomes invalid.

Valid Usage

- All elements of pCommandBuffers must not be in the pending state
- pCommandBuffers must be a valid pointer to an array of commandBufferCount VkCommandBuffer handles, each element of which must either be a valid handle or NULL

Valid Usage (Implicit)

- device must be a valid VkDevice handle
- commandPool must be a valid VkCommandPool handle
- commandBufferCount must be greater than 0
- commandPool must have been created, allocated, or retrieved from device
- Each element of pCommandBuffers that is a valid handle must have been created, allocated, or retrieved from commandPool

Host Synchronization

- Host access to commandPool must be externally synchronized
- Host access to each member of pCommandBuffers must be externally synchronized

5.4. Command Buffer Recording

To begin recording a command buffer, call:
VkResult vkBeginCommandBuffer(
    VkCommandBuffer                             commandBuffer,  // commandBuffer is the handle of the command buffer which is to be put in the recording state.
    const VkCommandBufferBeginInfo*             pBeginInfo)        // pBeginInfo is an instance of the VkCommandBufferBeginInfo structure, which defines additional information about how the command buffer begins recording.

**Valid Usage**

- **commandBuffer** must not be in the recording or pending state.
- If **commandBuffer** was allocated from a VkCommandPool which did not have the VK_COMMAND_POOL_CREATE_RESET_COMMAND_BUFFER_BIT flag set, **commandBuffer** must be in the initial state.
- If **commandBuffer** is a secondary command buffer, the pInheritanceInfo member of **pBeginInfo** must be a valid VkCommandBufferInheritanceInfo structure.
- If **commandBuffer** is a secondary command buffer and either the occlusionQueryEnable member of the pInheritanceInfo member of **pBeginInfo** is VK_FALSE, or the precise occlusion queries feature is not enabled, the queryFlags member of the pInheritanceInfo member **pBeginInfo** must not contain VK_QUERY_CONTROL_PRECISE_BIT

**Valid Usage (Implicit)**

- **commandBuffer** must be a valid VkCommandBuffer handle
- **pBeginInfo** must be a valid pointer to a valid VkCommandBufferBeginInfo structure

**Host Synchronization**

- Host access to **commandBuffer** must be externally synchronized
- Host access to the VkCommandPool that **commandBuffer** was allocated from must be externally synchronized

**Return Codes**

**Success**
- VK_SUCCESS

**Failure**
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY
The \texttt{VkCommandBufferBeginInfo} structure is defined as:

```c
typedef struct VkCommandBufferBeginInfo {
    VkStructureType                          sType;
    const void*                              pNext;
    VkCommandBufferUsageFlags                flags;
    const VkCommandBufferInheritanceInfo*    pInheritanceInfo;
} VkCommandBufferBeginInfo;
```

- \texttt{sType} is the type of this structure.
- \texttt{pNext} is \texttt{NULL} or a pointer to an extension-specific structure.
- \texttt{flags} is a bitmask of \texttt{VkCommandBufferUsageFlagBits} specifying usage behavior for the command buffer.
- \texttt{pInheritanceInfo} is a pointer to a \texttt{VkCommandBufferInheritanceInfo} structure, which is used if \texttt{commandBuffer} is a secondary command buffer. If this is a primary command buffer, then this value is ignored.

### Valid Usage

- If \texttt{flags} contains \texttt{VK_COMMAND_BUFFER_USAGE_RENDER_PASS_CONTINUE_BIT}, the \texttt{renderPass} member of \texttt{pInheritanceInfo} must be a valid \texttt{VkRenderPass}.
- If \texttt{flags} contains \texttt{VK_COMMAND_BUFFER_USAGE_RENDER_PASS_CONTINUE_BIT}, the \texttt{subpass} member of \texttt{pInheritanceInfo} must be a valid subpass index within the \texttt{renderPass} member of \texttt{pInheritanceInfo}.
- If \texttt{flags} contains \texttt{VK_COMMAND_BUFFER_USAGE_RENDER_PASS_CONTINUE_BIT}, the \texttt{framebuffer} member of \texttt{pInheritanceInfo} must be either \texttt{VK_NULL_HANDLE}, or a valid \texttt{VkFramebuffer} that is compatible with the \texttt{renderPass} member of \texttt{pInheritanceInfo}.

### Valid Usage (Implicit)

- \texttt{sType} must be \texttt{VK_STRUCTURE_TYPE_COMMAND_BUFFER_BEGIN_INFO}.
- \texttt{pNext} must be \texttt{NULL}.
- \texttt{flags} must be a valid combination of \texttt{VkCommandBufferUsageFlagBits} values.

Bits which \textit{can} be set in \texttt{VkCommandBufferBeginInfo::flags} to specify usage behavior for a command buffer are:

```c
typedef enum VkCommandBufferUsageFlagBits {
    VK_COMMAND_BUFFER_USAGE_ONE_TIME_SUBMIT_BIT = 0x00000001,
    VK_COMMAND_BUFFER_USAGE_RENDER_PASS_CONTINUE_BIT = 0x00000002,
    VK_COMMAND_BUFFER_USAGE_SIMULTANEOUS_USE_BIT = 0x00000004,
} VkCommandBufferUsageFlagBits;
```
• **VK_COMMAND_BUFFER_USAGE_ONE_TIME_SUBMIT_BIT** specifies that each recording of the command buffer will only be submitted once, and the command buffer will be reset and recorded again between each submission.

• **VK_COMMAND_BUFFER_USAGE_RENDER_PASS_CONTINUE_BIT** specifies that a secondary command buffer is considered to be entirely inside a render pass. If this is a primary command buffer, then this bit is ignored.

• **VK_COMMAND_BUFFER_USAGE_SIMULTANEOUS_USE_BIT** specifies that a command buffer can be resubmitted to a queue while it is in the *pending state*, and recorded into multiple primary command buffers.

```cpp
typedef VkFlags VkCommandBufferUsageFlags;
```

*VkCommandBufferUsageFlags* is a bitmask type for setting a mask of zero or more *VkCommandBufferUsageFlagBits*.

If the command buffer is a secondary command buffer, then the *VkCommandBufferInheritanceInfo* structure defines any state that will be inherited from the primary command buffer:

```cpp
typedef struct VkCommandBufferInheritanceInfo {
    VkStructureType                  sType;
    const void*                      pNext;
    VkRenderPass                     renderPass;
    uint32_t                         subpass;
    VkFramebuffer                    framebuffer;
    VkBool32                         occlusionQueryEnable;
    VkQueryControlFlags              queryFlags;
    VkQueryPipelineStatisticFlags    pipelineStatistics;
} VkCommandBufferInheritanceInfo;
```

- **sType** is the type of this structure.
- **pNext** is *NULL* or a pointer to an extension-specific structure.
- **renderPass** is a *VkRenderPass* object defining which render passes the *VkCommandBuffer* will be compatible with and can be executed within. If the *VkCommandBuffer* will not be executed within a render pass instance, **renderPass** is ignored.
- **subpass** is the index of the subpass within the render pass instance that the *VkCommandBuffer* will be executed within. If the *VkCommandBuffer* will not be executed within a render pass instance, **subpass** is ignored.
- **framebuffer** optionally refers to the *VkFramebuffer* object that the *VkCommandBuffer* will be rendering to if it is executed within a render pass instance. It can be **VK_NULL_HANDLE** if the framebuffer is not known, or if the *VkCommandBuffer* will not be executed within a render pass instance.
Note
Specifying the exact framebuffer that the secondary command buffer will be executed with may result in better performance at command buffer execution time.

• **occlusionQueryEnable** specifies whether the command buffer can be executed while an occlusion query is active in the primary command buffer. If this is **VK_TRUE**, then this command buffer can be executed whether the primary command buffer has an occlusion query active or not. If this is **VK_FALSE**, then the primary command buffer must not have an occlusion query active.

• **queryFlags** specifies the query flags that can be used by an active occlusion query in the primary command buffer when this secondary command buffer is executed. If this value includes the **VK_QUERY_CONTROL_PRECISE_BIT** bit, then the active query can return boolean results or actual sample counts. If this bit is not set, then the active query must not use the **VK_QUERY_CONTROL_PRECISE_BIT** bit.

• **pipelineStatistics** is a bitmask of **VkQueryPipelineStatisticFlagBits** specifying the set of pipeline statistics that can be counted by an active query in the primary command buffer when this secondary command buffer is executed. If this value includes a given bit, then this command buffer can be executed whether the primary command buffer has a pipeline statistics query active that includes this bit or not. If this value excludes a given bit, then the active pipeline statistics query must not be from a query pool that counts that statistic.

---

**Valid Usage**

- If the **inherited queries** feature is not enabled, **occlusionQueryEnable** must be **VK_FALSE**
- If the **inherited queries** feature is enabled, **queryFlags** must be a valid combination of **VkQueryControlFlagBits** values
- If the **pipeline statistics queries** feature is not enabled, **pipelineStatistics** must be **0**

**Valid Usage (Implicit)**

- **sType** must be **VK_STRUCTURE_TYPE_COMMAND BUFFER_INHERITANCE_INFO**
- **pNext** must be **NULL**
- Both of **framebuffer**, and **renderPass** that are valid handles must have been created, allocated, or retrieved from the same **VkDevice**

If **VK_COMMAND BUFFER USAGE_SIMULTANEOUS_USE_BIT** was not set when creating a command buffer, that command buffer must not be submitted to a queue whilst it is already in the pending state. If **VK_COMMAND BUFFER USAGE_SIMULTANEOUS_USE_BIT** is not set on a secondary command buffer, that command buffer must not be used more than once in a given primary command buffer.
Note

On some implementations, not using the `VK_COMMAND_BUFFER_USAGE_SIMULTANEOUS_USE_BIT` bit enables command buffers to be patched in-place if needed, rather than creating a copy of the command buffer.

If a command buffer is in the invalid, or executable state, and the command buffer was allocated from a command pool with the `VK_COMMAND_POOL_CREATE_RESET_COMMAND_BUFFER_BIT` flag set, then `vkBeginCommandBuffer` implicitly resets the command buffer, behaving as if `vkResetCommandBuffer` had been called with `VK_COMMAND_BUFFER_RESET_RELEASE_RESOURCES_BIT` not set. After the implicit reset, `commandBuffer` is moved to the recording state.

Once recording starts, an application records a sequence of commands (vkCmd*) to set state in the command buffer, draw, dispatch, and other commands.

To complete recording of a command buffer, call:

```c
VkResult vkEndCommandBuffer(
    VkCommandBuffer commandBuffer);
```

- `commandBuffer` is the command buffer to complete recording.

If there was an error during recording, the application will be notified by an unsuccessful return code returned by `vkEndCommandBuffer`. If the application wishes to further use the command buffer, the command buffer must be reset. The command buffer must have been in the recording state, and is moved to the executable state.

Valid Usage

- `commandBuffer` must be in the recording state.
- If `commandBuffer` is a primary command buffer, there must not be an active render pass instance
- All queries made active during the recording of `commandBuffer` must have been made inactive

Valid Usage (Implicit)

- `commandBuffer` must be a valid VkCommandBuffer handle

Host Synchronization

- Host access to `commandBuffer` must be externally synchronized
- Host access to the `VkCommandPool` that `commandBuffer` was allocated from must be externally synchronized
Return Codes

Success
- VK_SUCCESS

Failure
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY

When a command buffer is in the executable state, it can be submitted to a queue for execution.

5.5. Command Buffer Submission

To submit command buffers to a queue, call:

```c
VkResult vkQueueSubmit(
    VkQueue                                     queue,
    uint32_t                                    submitCount,
    const VkSubmitInfo*                         pSubmits,
    VkFence                                     fence);
```

- `queue` is the queue that the command buffers will be submitted to.
- `submitCount` is the number of elements in the `pSubmits` array.
- `pSubmits` is a pointer to an array of `VkSubmitInfo` structures, each specifying a command buffer submission batch.
- `fence` is an optional handle to a fence to be signaled once all submitted command buffers have completed execution. If `fence` is not `VK_NULL_HANDLE`, it defines a fence signal operation.

**Note**
Submission can be a high overhead operation, and applications should attempt to batch work together into as few calls to `vkQueueSubmit` as possible.

`vkQueueSubmit` is a queue submission command, with each batch defined by an element of `pSubmits` as an instance of the `VkSubmitInfo` structure. Batches begin execution in the order they appear in `pSubmits`, but may complete out of order.

Fence and semaphore operations submitted with `vkQueueSubmit` have additional ordering constraints compared to other submission commands, with dependencies involving previous and subsequent queue operations. Information about these additional constraints can be found in the semaphore and fence sections of the synchronization chapter.

Details on the interaction of `pWaitDstStageMask` with synchronization are described in the semaphore wait operation section of the synchronization chapter.

The order that batches appear in `pSubmits` is used to determine submission order, and thus all the
implicit ordering guarantees that respect it. Other than these implicit ordering guarantees and any explicit synchronization primitives, these batches may overlap or otherwise execute out of order.

If any command buffer submitted to this queue is in the executable state, it is moved to the pending state. Once execution of all submissions of a command buffer complete, it moves from the pending state, back to the executable state. If a command buffer was recorded with the VK_COMMAND_BUFFER_USAGE_ONE_TIME_SUBMIT_BIT flag, it instead moves back to the invalid state.

If vkQueueSubmit fails, it may return VK_ERROR_OUT_OF_HOST_MEMORY or VK_ERROR_OUT_OF_DEVICE_MEMORY. If it does, the implementation must ensure that the state and contents of any resources or synchronization primitives referenced by the submitted command buffers and any semaphores referenced by pSubmits is unaffected by the call or its failure. If vkQueueSubmit fails in such a way that the implementation is unable to make that guarantee, the implementation must return VK_ERROR_DEVICE_LOST. See Lost Device.
Valid Usage

• If fence is not VK_NULL_HANDLE, fence must be unsignaled.

• If fence is not VK_NULL_HANDLE, fence must not be associated with any other queue command that has not yet completed execution on that queue.

• Any calls to vkCmdSetEvent, vkCmdResetEvent or vkCmdWaitEvents that have been recorded into any of the command buffer elements of the pCommandBuffers member of any element of pSubmits, must not reference any VkEvent that is referenced by any of those commands in a command buffer that has been submitted to another queue and is still in the pending state.

• Any stage flag included in any element of the pWaitDstStageMask member of any element of pSubmits must be a pipeline stage supported by one of the capabilities of queue, as specified in the table of supported pipeline stages.

• Each element of the pSignalSemaphores member of any element of pSubmits must be unsignaled when the semaphore signal operation it defines is executed on the device.

• When a semaphore unsignal operation defined by any element of the pWaitSemaphores member of any element of pSubmits executes on queue, no other queue must be waiting on the same semaphore.

• All elements of the pWaitSemaphores member of all elements of pSubmits must be semaphores that are signaled, or have semaphore signal operations previously submitted for execution.

• Each element of the pCommandBuffers member of each element of pSubmits must be in the pending or executable state.

• If any element of the pCommandBuffers member of any element of pSubmits was not recorded with the VK_COMMAND_BUFFER_USAGE_SIMULTANEOUS_USE_BIT, it must not be in the pending state.

• Any secondary command buffers recorded into any element of the pCommandBuffers member of any element of pSubmits must be in the pending or executable state.

• If any secondary command buffers recorded into any element of the pCommandBuffers member of any element of pSubmits was not recorded with the VK_COMMAND_BUFFER_USAGE_SIMULTANEOUS_USE_BIT, it must not be in the pending state.

• Each element of the pCommandBuffers member of each element of pSubmits must have been allocated from a VkCommandPool that was created for the same queue family queue belongs to.

• If any element of pSubmits->pCommandBuffers includes a Queue Family Transfer Acquire Operation, there must exist a previously submitted Queue Family Transfer Release Operation on a queue in the queue family identified by the acquire operation, with parameters matching the acquire operation as defined in the definition of such acquire operations, and which happens before the acquire operation.
Valid Usage (Implicit)

- **queue** must be a valid VkQueue handle
- If `submitCount` is not 0, `pSubmits` must be a valid pointer to an array of `submitCount` valid VkSubmitInfo structures
- If `fence` is not VK_NULL_HANDLE, `fence` must be a valid VkFence handle
- Both of `fence`, and `queue` that are valid handles must have been created, allocated, or retrieved from the same VkDevice

Host Synchronization

- Host access to `queue` must be externally synchronized
- Host access to `pSubmits[].pWaitSemaphores[]` must be externally synchronized
- Host access to `pSubmits[].pSignalSemaphores[]` must be externally synchronized
- Host access to `fence` must be externally synchronized

Command Properties

<table>
<thead>
<tr>
<th>Command Buffer Levels</th>
<th>Render Pass Scope</th>
<th>Supported Queue Types</th>
<th>Pipeline Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>Any</td>
<td>-</td>
</tr>
</tbody>
</table>

Return Codes

Success
- VK_SUCCESS

Failure
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY
- VK_ERROR_DEVICE_LOST

The VkSubmitInfo structure is defined as:
typedef struct VkSubmitInfo {
  VkStructureType sType;
  const void* pNext;
  uint32_t waitSemaphoreCount;
  const VkSemaphore* pWaitSemaphores;
  const VkPipelineStageFlags* pWaitDstStageMask;
  uint32_t commandBufferCount;
  const VkCommandBuffer* pCommandBuffers;
  uint32_t signalSemaphoreCount;
  const VkSemaphore* pSignalSemaphores;
} VkSubmitInfo;

• **sType** is the type of this structure.
• **pNext** is NULL or a pointer to an extension-specific structure.
• **waitSemaphoreCount** is the number of semaphores upon which to wait before executing the command buffers for the batch.
• **pWaitSemaphores** is a pointer to an array of semaphores upon which to wait before the command buffers for this batch begin execution. If semaphores to wait on are provided, they define a semaphore wait operation.
• **pWaitDstStageMask** is a pointer to an array of pipeline stages at which each corresponding semaphore wait will occur.
• **commandBufferCount** is the number of command buffers to execute in the batch.
• **pCommandBuffers** is a pointer to an array of command buffers to execute in the batch.
• **signalSemaphoreCount** is the number of semaphores to be signaled once the commands specified in **pCommandBuffers** have completed execution.
• **pSignalSemaphores** is a pointer to an array of semaphores which will be signaled when the command buffers for this batch have completed execution. If semaphores to be signaled are provided, they define a semaphore signal operation.

The order that command buffers appear in **pCommandBuffers** is used to determine submission order, and thus all the implicit ordering guarantees that respect it. Other than these implicit ordering guarantees and any explicit synchronization primitives, these command buffers may overlap or otherwise execute out of order.
Valid Usage

- Each element of `pCommandBuffers` must not have been allocated with `VK_COMMAND_BUFFER_LEVEL_SECONDARY`.
- If the geometry shaders feature is not enabled, each element of `pWaitDstStageMask` must not contain `VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT`.
- If the tessellation shaders feature is not enabled, each element of `pWaitDstStageMask` must not contain `VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT` or `VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT`.
- Each element of `pWaitDstStageMask` must not include `VK_PIPELINE_STAGE_HOST_BIT`.

Valid Usage (Implicit)

- `sType` must be `VK_STRUCTURE_TYPE_SUBMIT_INFO`.
- `pNext` must be `NULL`.
- If `waitSemaphoreCount` is not 0, `pWaitSemaphores` must be a valid pointer to an array of `waitSemaphoreCount` valid `VkSemaphore` handles.
- If `waitSemaphoreCount` is not 0, `pWaitDstStageMask` must be a valid pointer to an array of `waitSemaphoreCount` valid combinations of `VkPipelineStageFlagBits` values.
- Each element of `pWaitDstStageMask` must not be 0.
- If `commandBufferCount` is not 0, `pCommandBuffers` must be a valid pointer to an array of `commandBufferCount` valid `VkCommandBuffer` handles.
- If `signalSemaphoreCount` is not 0, `pSignalSemaphores` must be a valid pointer to an array of `signalSemaphoreCount` valid `VkSemaphore` handles.
- Each of the elements of `pCommandBuffers`, the elements of `pSignalSemaphores`, and the elements of `pWaitSemaphores` that are valid handles must have been created, allocated, or retrieved from the same `VkDevice`.

5.6. Queue Forward Progress

The application must ensure that command buffer submissions will be able to complete without any subsequent operations by the application on any queue. After any call to `vkQueueSubmit`, for every queued wait on a semaphore there must be a prior signal of that semaphore that will not be consumed by a different wait on the semaphore.

Command buffers in the submission can include `vkCmdWaitEvents` commands that wait on events that will not be signaled by earlier commands in the queue. Such events must be signaled by the application using `vkSetEvent`, and the `vkCmdWaitEvents` commands that wait upon them must not be inside a render pass instance. Implementations may have limits on how long the command buffer will wait, in order to avoid interfering with progress of other clients of the device. If the event is not signaled within these limits, results are undefined and may include device loss.
5.7. Secondary Command Buffer Execution

A secondary command buffer must not be directly submitted to a queue. Instead, secondary command buffers are recorded to execute as part of a primary command buffer with the command:

```c
void vkCmdExecuteCommands(
    VkCommandBuffer commandBuffer,  // A handle to a primary command buffer.
    uint32_t commandBufferCount,    // Length of the pCommandBuffers array.
    const VkCommandBuffer* pCommandBuffers);  // Array of secondary command buffer handles.
```

- `commandBuffer` is a handle to a primary command buffer that the secondary command buffers are executed in.
- `commandBufferCount` is the length of the `pCommandBuffers` array.
- `pCommandBuffers` is an array of secondary command buffer handles, which are recorded to execute in the primary command buffer in the order they are listed in the array.

If any element of `pCommandBuffers` was not recorded with the `VK_COMMAND_BUFFER_USAGE_SIMULTANEOUS_USE_BIT` flag, and it was recorded into any other primary command buffer which is currently in the executable or recording state, that primary command buffer becomes invalid.
Valid Usage

• commandBuffer must have been allocated with a level of VK_COMMAND_BUFFER_LEVEL_PRIMARY

• Each element of pCommandBuffers must have been allocated with a level of VK_COMMAND_BUFFER_LEVEL_SECONDARY

• Each element of pCommandBuffers must be in the pending or executable state.

• If any element of pCommandBuffers was not recorded with the VK_COMMAND_BUFFER_USAGE_SIMULTANEOUS_USE_BIT flag, and it was recorded into any other primary command buffer, that primary command buffer must not be in the pending state.

• If any element of pCommandBuffers was not recorded with the VK_COMMAND_BUFFER_USAGE_SIMULTANEOUS_USE_BIT flag, it must not be in the pending state.

• If any element of pCommandBuffers was not recorded with the VK_COMMAND_BUFFER_USAGE_SIMULTANEOUS_USE_BIT flag, it must not have already been recorded to commandBuffer.

• If any element of pCommandBuffers was not recorded with the VK_COMMAND_BUFFER_USAGE_SIMULTANEOUS_USE_BIT flag, it must not appear more than once in pCommandBuffers.

• Each element of pCommandBuffers must have been allocated from a VkCommandPool that was created for the same queue family as the VkCommandPool from which commandBuffer was allocated.

• If vkCmdExecuteCommands is being called within a render pass instance, that render pass instance must have been begun with the contents parameter of vkCmdBeginRenderPass set to VK_SUBPASS_CONTENTS_SECONDARY_COMMAND_BUFFERS.

• If vkCmdExecuteCommands is being called within a render pass instance, each element of pCommandBuffers must have been recorded with the VK_COMMAND_BUFFER_USAGE_RENDER_PASS_CONTINUE_BIT.

• If vkCmdExecuteCommands is being called within a render pass instance, each element of pCommandBuffers must have been recorded with VkCommandBufferInheritanceInfo::subpass set to the index of the subpass which the given command buffer will be executed in.

• If vkCmdExecuteCommands is being called within a render pass instance, the render passes specified in the pBeginInfo::pInheritanceInfo::renderPass members of the vkBeginCommandBuffer commands used to begin recording each element of pCommandBuffers must be compatible with the current render pass.

• If vkCmdExecuteCommands is being called within a render pass instance, and any element of pCommandBuffers was recorded with VkCommandBufferInheritanceInfo::Framebuffer not equal to VK_NULL_HANDLE, that VkFramebuffer must match the VkFramebuffer used in the current render pass instance.

• If vkCmdExecuteCommands is not being called within a render pass instance, each element of pCommandBuffers must not have been recorded with the VK_COMMAND_BUFFER_USAGE_RENDER_PASS_CONTINUE_BIT.

• If the inherited queries feature is not enabled, commandBuffer must not have any queries active.
• If `commandBuffer` has a `VK_QUERY_TYPE_OCCLUSION` query active, then each element of `pCommandBuffers` must have been recorded with `VkCommandBufferInheritanceInfo::occlusionQueryEnable` set to `VK_TRUE`

• If `commandBuffer` has a `VK_QUERY_TYPE_OCCLUSION` query active, then each element of `pCommandBuffers` must have been recorded with `VkCommandBufferInheritanceInfo::queryFlags` having all bits set that are set for the query

• If `commandBuffer` has a `VK_QUERY_TYPE_PIPELINE_STATISTICS` query active, then each element of `pCommandBuffers` must have been recorded with `VkCommandBufferInheritanceInfo::pipelineStatistics` having all bits set that are set in the `VkQueryPool` the query uses

• Each element of `pCommandBuffers` must not begin any query types that are active in `commandBuffer`

---

### Valid Usage (Implicit)

• `commandBuffer` must be a valid `VkCommandBuffer` handle

• `pCommandBuffers` must be a valid pointer to an array of `commandBufferCount` valid `VkCommandBuffer` handles

• `commandBuffer` must be in the recording state

• The `VkCommandPool` that `commandBuffer` was allocated from must support transfer, graphics, or compute operations

• `commandBuffer` must be a primary `VkCommandBuffer`

• `commandBufferCount` must be greater than 0

• Both of `commandBuffer`, and the elements of `pCommandBuffers` must have been created, allocated, or retrieved from the same `VkDevice`

---

### Host Synchronization

• Host access to `commandBuffer` must be externally synchronized

• Host access to the `VkCommandPool` that `commandBuffer` was allocated from must be externally synchronized

---

### Command Properties

<table>
<thead>
<tr>
<th>Command Buffer Levels</th>
<th>Render Pass Scope</th>
<th>Supported Queue Types</th>
<th>Pipeline Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>Both</td>
<td>Transfer, Graphics, Compute</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 6. Synchronization and Cache Control

Synchronization of access to resources is primarily the responsibility of the application in Vulkan. The order of execution of commands with respect to the host and other commands on the device has few implicit guarantees, and needs to be explicitly specified. Memory caches and other optimizations are also explicitly managed, requiring that the flow of data through the system is largely under application control.

Whilst some implicit guarantees exist between commands, five explicit synchronization mechanisms are exposed by Vulkan:

Fences
Fences can be used to communicate to the host that execution of some task on the device has completed.

Semaphores
Semaphores can be used to control resource access across multiple queues.

Events
Events provide a fine-grained synchronization primitive which can be signaled either within a command buffer or by the host, and can be waited upon within a command buffer or queried on the host.

Pipeline Barriers
Pipeline barriers also provide synchronization control within a command buffer, but at a single point, rather than with separate signal and wait operations.

Render Passes
Render passes provide a useful synchronization framework for most rendering tasks, built upon the concepts in this chapter. Many cases that would otherwise need an application to use other synchronization primitives can be expressed more efficiently as part of a render pass.

6.1. Execution and Memory Dependencies

An operation is an arbitrary amount of work to be executed on the host, a device, or an external entity such as a presentation engine. Synchronization commands introduce explicit execution dependencies, and memory dependencies between two sets of operations defined by the command’s two synchronization scopes.

The synchronization scopes define which other operations a synchronization command is able to create execution dependencies with. Any type of operation that is not in a synchronization command’s synchronization scopes will not be included in the resulting dependency. For example, for many synchronization commands, the synchronization scopes can be limited to just operations executing in specific pipeline stages, which allows other pipeline stages to be excluded from a dependency. Other scoping options are possible, depending on the particular command.
An execution dependency is a guarantee that for two sets of operations, the first set must happen-before the second set. If an operation happens-before another operation, then the first operation must complete before the second operation is initiated. More precisely:

- Let \( A \) and \( B \) be separate sets of operations.
- Let \( S \) be a synchronization command.
- Let \( A_s \) and \( B_s \) be the synchronization scopes of \( S \).
- Let \( A' \) be the intersection of sets \( A \) and \( A_s \).
- Let \( B' \) be the intersection of sets \( B \) and \( B_s \).
- Submitting \( A, S \) and \( B \) for execution, in that order, will result in execution dependency \( E \) between \( A' \) and \( B' \).
- Execution dependency \( E \) guarantees that \( A' \) happens-before \( B' \).

An execution dependency chain is a sequence of execution dependencies that form a happens-before relation between the first dependency's \( A' \) and the final dependency's \( B' \). For each consecutive pair of execution dependencies, a chain exists if the intersection of \( B_s \) in the first dependency and \( A_s \) in the second dependency is not an empty set. The formation of a single execution dependency from an execution dependency chain can be described by substituting the following in the description of execution dependencies:

- Let \( S \) be a set of synchronization commands that generate an execution dependency chain.
- Let \( A_s \) be the first synchronization scope of the first command in \( S \).
- Let \( B_s \) be the second synchronization scope of the last command in \( S \).

Note

An execution dependency is inherently also multiple execution dependencies - a dependency exists between each subset of \( A' \) and each subset of \( B' \), and the same is true for execution dependency chains. For example, a synchronization command with multiple pipeline stages in its stage masks effectively generates one dependency between each source stage and each destination stage. This can be useful to think about when considering how execution chains are formed if they do not involve all parts of a synchronization command’s dependency. Similarly, any set of adjacent dependencies in an execution dependency chain can be considered an execution dependency chain in its own right.

Execution dependencies alone are not sufficient to guarantee that values resulting from writes in one set of operations can be read from another set of operations.

Three additional types of operation are used to control memory access. Availability operations cause the values generated by specified memory write accesses to become available to a memory domain for future access. Any available value remains available until a subsequent write to the same memory location occurs (whether it is made available or not) or the memory is freed. Memory domain operations cause writes that are available to a source memory domain to become available to a destination memory domain (an example of this is making writes available to the host domain available to the device domain). Visibility operations cause values available to a memory domain to
become visible to specified memory accesses.

A memory dependency is an execution dependency which includes availability and visibility operations such that:

- The first set of operations happens-before the availability operation.
- The availability operation happens-before the visibility operation.
- The visibility operation happens-before the second set of operations.

Once written values are made visible to a particular type of memory access, they can be read or written by that type of memory access. Most synchronization commands in Vulkan define a memory dependency.

The specific memory accesses that are made available and visible are defined by the access scopes of a memory dependency. Any type of access that is in a memory dependency’s first access scope and occurs in A’ is made available. Any type of access that is in a memory dependency’s second access scope and occurs in B’ has any available writes made visible to it. Any type of operation that is not in a synchronization command’s access scopes will not be included in the resulting dependency.

A memory dependency enforces availability and visibility of memory accesses and execution order between two sets of operations. Adding to the description of execution dependency chains:

- Let \( a \) be the set of memory accesses performed by A’.
- Let \( b \) be the set of memory accesses performed by B’.
- Let \( a_s \) be the first access scope of the first command in S.
- Let \( b_s \) be the second access scope of the last command in S.
- Let \( a' \) be the intersection of sets \( a \) and \( a_s \).
- Let \( b' \) be the intersection of sets \( b \) and \( b_s \).
- Submitting A, S and B for execution, in that order, will result in a memory dependency \( m \) between A’ and B’.
- Memory dependency \( m \) guarantees that:
  - Memory writes in \( a' \) are made available.
  - Available memory writes, including those from \( a' \), are made visible to \( b' \).

**Note**

Execution and memory dependencies are used to solve data hazards, i.e. to ensure that read and write operations occur in a well-defined order. Write-after-read hazards can be solved with just an execution dependency, but read-after-write and write-after-write hazards need appropriate memory dependencies to be included between them. If an application does not include dependencies to solve these hazards, the results and execution orders of memory accesses are undefined.
6.1.1. Image Layout Transitions

Image subresources can be transitioned from one layout to another as part of a memory dependency (e.g. by using an image memory barrier). When a layout transition is specified in a memory dependency, it happens-after the availability operations in the memory dependency, and happens-before the visibility operations. Image layout transitions may perform read and write accesses on all memory bound to the image subresource range, so applications must ensure that all memory writes have been made available before a layout transition is executed. Available memory is automatically made visible to a layout transition, and writes performed by a layout transition are automatically made available.

Layout transitions always apply to a particular image subresource range, and specify both an old layout and new layout. If the old layout does not match the new layout, a transition occurs. The old layout must match the current layout of the image subresource range, with one exception. The old layout can always be specified as VK_IMAGE_LAYOUT_UNDEFINED, though doing so invalidates the contents of the image subresource range.

**Note**

Setting the old layout to VK_IMAGE_LAYOUT_UNDEFINED implies that the contents of the image subresource need not be preserved. Implementations may use this information to avoid performing expensive data transition operations.

**Note**

Applications must ensure that layout transitions happen-after all operations accessing the image with the old layout, and happen-before any operations that will access the image with the new layout. Layout transitions are potentially read/write operations, so not defining appropriate memory dependencies to guarantee this will result in a data race.

Image layout transitions interact with memory aliasing.

6.1.2. Pipeline Stages

The work performed by an action or synchronization command consists of multiple operations, which are performed as a sequence of logically independent steps known as pipeline stages. The exact pipeline stages executed depend on the particular command that is used, and current command buffer state when the command was recorded. Drawing commands, dispatching commands, copy commands, clear commands, and synchronization commands all execute in different sets of pipeline stages. Synchronization commands do not execute in a defined pipeline, but do execute VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT and VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT.

**Note**

Operations performed by synchronization commands (e.g. availability and visibility operations) are not executed by a defined pipeline stage. However other commands can still synchronize with them via the VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT and VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT pipeline stages.
Execution of operations across pipeline stages must adhere to implicit ordering guarantees, particularly including pipeline stage order. Otherwise, execution across pipeline stages may overlap or execute out of order with regards to other stages, unless otherwise enforced by an execution dependency.

Several of the synchronization commands include pipeline stage parameters, restricting the synchronization scopes for that command to just those stages. This allows fine grained control over the exact execution dependencies and accesses performed by action commands. Implementations should use these pipeline stages to avoid unnecessary stalls or cache flushing.

Bits which can be set, specifying pipeline stages, are:

```c
typedef enum VkPipelineStageFlagBits {
    VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT = 0x00000001,
    VK_PIPELINE_STAGE_DRAW_INDIRECT_BIT = 0x00000002,
    VK_PIPELINE_STAGE_VERTEX_INPUT_BIT = 0x00000004,
    VK_PIPELINE_STAGE_VERTEX_SHADER_BIT = 0x00000008,
    VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT = 0x00000010,
    VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT = 0x00000020,
    VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT = 0x00000040,
    VK_PIPELINE_STAGE_FRAGMENT_SHADER_BIT = 0x00000080,
    VK_PIPELINE_STAGE_EARLY_FRAGMENT_TESTS_BIT = 0x00000100,
    VK_PIPELINE_STAGE_LATE_FRAGMENT_TESTS_BIT = 0x00000200,
    VK_PIPELINE_STAGE_COLOR_ATTACHMENT_OUTPUT_BIT = 0x00000400,
    VK_PIPELINE_STAGE_COMPUTE_SHADER_BIT = 0x00000800,
    VK_PIPELINE_STAGE_TRANSFER_BIT = 0x00001000,
    VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT = 0x00002000,
    VK_PIPELINE_STAGE_HOST_BIT = 0x00004000,
    VK_PIPELINE_STAGE_ALL_GRAPHICS_BIT = 0x00008000,
    VK_PIPELINE_STAGE_ALL_COMMANDS_BIT = 0x00010000,
} VkPipelineStageFlagBits;
```

- **VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT** specifies the stage of the pipeline where any commands are initially received by the queue.
- **VK_PIPELINE_STAGE_DRAW_INDIRECT_BIT** specifies the stage of the pipeline where Draw/DispatchIndirect data structures are consumed.
- **VK_PIPELINE_STAGE_VERTEX_INPUT_BIT** specifies the stage of the pipeline where vertex and index buffers are consumed.
- **VK_PIPELINE_STAGE_VERTEX_SHADER_BIT** specifies the vertex shader stage.
- **VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT** specifies the tessellation control shader stage.
- **VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT** specifies the tessellation evaluation shader stage.
- **VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT** specifies the geometry shader stage.
- **VK_PIPELINE_STAGE_FRAGMENT_SHADER_BIT** specifies the fragment shader stage.
• **VK_PIPELINE_STAGE_EARLY_FRAGMENT_TESTS_BIT** specifies the stage of the pipeline where early fragment tests (depth and stencil tests before fragment shading) are performed. This stage also includes subpass load operations for framebuffer attachments with a depth/stencil format.

• **VK_PIPELINE_STAGE_LATE_FRAGMENT_TESTS_BIT** specifies the stage of the pipeline where late fragment tests (depth and stencil tests after fragment shading) are performed. This stage also includes subpass store operations for framebuffer attachments with a depth/stencil format.

• **VK_PIPELINE_STAGE_COLOR_ATTACHMENT_OUTPUT_BIT** specifies the stage of the pipeline after blending where the final color values are output from the pipeline. This stage also includes subpass load and store operations and multisample resolve operations for framebuffer attachments with a color format.

• **VK_PIPELINE_STAGE_COMPUTE_SHADER_BIT** specifies the execution of a compute shader.

• **VK_PIPELINE_STAGE_TRANSFER_BIT** specifies the execution of copy commands. This includes the operations resulting from all copy commands, clear commands (with the exception of vkCmdClearAttachments), and vkCmdCopyQueryPoolResults.

• **VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT** specifies the final stage in the pipeline where operations generated by all commands complete execution.

• **VK_PIPELINE_STAGE_HOST_BIT** specifies a pseudo-stage indicating execution on the host of reads/writes of device memory. This stage is not invoked by any commands recorded in a command buffer.

• **VK_PIPELINE_STAGE_ALL_GRAPHICS_BIT** specifies the execution of all graphics pipeline stages, and is equivalent to the logical OR of:
  - **VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT**
  - **VK_PIPELINE_STAGE_DRAW_INDIRECT_BIT**
  - **VK_PIPELINE_STAGE_VERTEX_INPUT_BIT**
  - **VK_PIPELINE_STAGE_VERTEX_SHADER_BIT**
  - **VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT**
  - **VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT**
  - **VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT**
  - **VK_PIPELINE_STAGE_FRAGMENT_SHADER_BIT**
  - **VK_PIPELINE_STAGE_EARLY_FRAGMENT_TESTS_BIT**
  - **VK_PIPELINE_STAGE_LATE_FRAGMENT_TESTS_BIT**
  - **VK_PIPELINE_STAGE_COLOR_ATTACHMENT_OUTPUT_BIT**
  - **VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT**

• **VK_PIPELINE_STAGE_ALL_COMMANDS_BIT** is equivalent to the logical OR of every other pipeline stage flag that is supported on the queue it is used with.
**Note**

An execution dependency with only `VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT` in the destination stage mask will only prevent that stage from executing in subsequently submitted commands. As this stage does not perform any actual execution, this is not observable - in effect, it does not delay processing of subsequent commands. Similarly an execution dependency with only `VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT` in the source stage mask will effectively not wait for any prior commands to complete.

When defining a memory dependency, using only `VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT` or `VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT` would never make any accesses available and/or visible because these stages do not access memory.

`VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT` and `VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT` are useful for accomplishing layout transitions and queue ownership operations when the required execution dependency is satisfied by other means - for example, semaphore operations between queues.

```c
typedef VkFlags VkPipelineStageFlags;
```

`VkPipelineStageFlags` is a bitmask type for setting a mask of zero or more `VkPipelineStageFlagBits`.

If a synchronization command includes a source stage mask, its first *synchronization scope* only includes execution of the pipeline stages specified in that mask, and its first *access scope* only includes memory access performed by pipeline stages specified in that mask. If a synchronization command includes a destination stage mask, its second *synchronization scope* only includes execution of the pipeline stages specified in that mask, and its second *access scope* only includes memory access performed by pipeline stages specified in that mask.

**Note**

Including a particular pipeline stage in the first *synchronization scope* of a command implicitly includes logically earlier pipeline stages in the synchronization scope. Similarly, the second *synchronization scope* includes logically later pipeline stages.

However, note that *access scopes* are not affected in this way - only the precise stages specified are considered part of each access scope.

Certain pipeline stages are only available on queues that support a particular set of operations. The following table lists, for each pipeline stage flag, which queue capability flag must be supported by the queue. When multiple flags are enumerated in the second column of the table, it means that the pipeline stage is supported on the queue if it supports any of the listed capability flags. For further details on queue capabilities see Physical Device Enumeration and Queues.

<table>
<thead>
<tr>
<th>Table 3. Supported pipeline stage flags</th>
</tr>
</thead>
</table>

88 | Chapter 6. Synchronization and Cache Control
<table>
<thead>
<tr>
<th>Pipeline stage flag</th>
<th>Required queue capability flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT</td>
<td>None required</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_DRAW_INDIRECT_BIT</td>
<td>VK_QUEUE_GRAPHICS_BIT or VK_QUEUE_COMPUTE_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_VERTEX_INPUT_BIT</td>
<td>VK_QUEUE_GRAPHICS_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_VERTEX_SHADER_BIT</td>
<td>VK_QUEUE_GRAPHICS_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT</td>
<td>VK_QUEUE_GRAPHICS_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT</td>
<td>VK_QUEUE_GRAPHICS_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT</td>
<td>VK_QUEUE_GRAPHICS_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_FRAGMENT_SHADER_BIT</td>
<td>VK_QUEUE_GRAPHICS_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_EARLY_FRAGMENT_TESTS_BIT</td>
<td>VK_QUEUE_GRAPHICS_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_LATE_FRAGMENT_TESTS_BIT</td>
<td>VK_QUEUE_GRAPHICS_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_COLOR_ATTACHMENT_OUTPUT_BIT</td>
<td>VK_QUEUE_GRAPHICS_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_COMPUTE_SHADER_BIT</td>
<td>VK_QUEUE_COMPUTE_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_TRANSFER_BIT</td>
<td>VK_QUEUE_GRAPHICS_BIT, VK_QUEUE_COMPUTE_BIT or VK_QUEUE_TRANSFER_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT</td>
<td>None required</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_HOST_BIT</td>
<td>None required</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_ALL_GRAPHICS_BIT</td>
<td>VK_QUEUE_GRAPHICS_BIT</td>
</tr>
<tr>
<td>VK_PIPELINE_STAGE_ALL_COMMANDS_BIT</td>
<td>None required</td>
</tr>
</tbody>
</table>

Pipeline stages that execute as a result of a command logically complete execution in a specific order, such that completion of a logically later pipeline stage **must** not happen-before completion of a logically earlier stage. This means that including any stage in the source stage mask for a particular synchronization command also implies that any logically earlier stages are included in A, for that command.

Similarly, initiation of a logically earlier pipeline stage **must** not happen-after initiation of a logically later pipeline stage. Including any given stage in the destination stage mask for a particular synchronization command also implies that any logically later stages are included in B, for that command.
Implementations **may** not support synchronization at every pipeline stage for every synchronization operation. If a pipeline stage that an implementation does not support synchronization for appears in a source stage mask, it **may** substitute any logically later stage in its place for the first synchronization scope. If a pipeline stage that an implementation does not support synchronization for appears in a destination stage mask, it **may** substitute any logically earlier stage in its place for the second synchronization scope.

For example, if an implementation is unable to signal an event immediately after vertex shader execution is complete, it **may** instead signal the event after color attachment output has completed.

If an implementation makes such a substitution, it **must** not affect the semantics of execution or memory dependencies or image and buffer memory barriers.

The order and set of pipeline stages executed by a given command is determined by the command’s pipeline type, as described below:

For the graphics pipeline, the following stages occur in this order:

- `VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT`
- `VK_PIPELINE_STAGE_DRAW_INDIRECT_BIT`
- `VK_PIPELINE_STAGE_VERTEX_INPUT_BIT`
- `VK_PIPELINE_STAGE_VERTEX_SHADER_BIT`
- `VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT`
- `VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT`
- `VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT`
- `VK_PIPELINE_STAGE_EARLY_FRAGMENT_TESTS_BIT`
- `VK_PIPELINE_STAGE_FRAGMENT_SHADER_BIT`
- `VK_PIPELINE_STAGE_LATE_FRAGMENT_TESTS_BIT`
- `VK_PIPELINE_STAGE_COLOR_ATTACHMENT_OUTPUT_BIT`
- `VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT`

For the compute pipeline, the following stages occur in this order:

- `VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT`
- `VK_PIPELINE_STAGE_DRAW_INDIRECT_BIT`
- `VK_PIPELINE_STAGE_COMPUTE_SHADER_BIT`
- `VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT`

For the transfer pipeline, the following stages occur in this order:

- `VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT`
- `VK_PIPELINE_STAGE_TRANSFER_BIT`
- `VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT`

For host operations, only one pipeline stage occurs, so no order is guaranteed:
6.1.3. Access Types

Memory in Vulkan can be accessed from within shader invocations and via some fixed-function stages of the pipeline. The access type is a function of the descriptor type used, or how a fixed-function stage accesses memory. Each access type corresponds to a bit flag in VkAccessFlagBits.

Some synchronization commands take sets of access types as parameters to define the access scopes of a memory dependency. If a synchronization command includes a source access mask, its first access scope only includes accesses via the access types specified in that mask. Similarly, if a synchronization command includes a destination access mask, its second access scope only includes accesses via the access types specified in that mask.

Access types that can be set in an access mask include:

```c
typedef enum VkAccessFlagBits {
    VK_ACCESS_INDIRECT_COMMAND_READ_BIT = 0x00000001,
    VK_ACCESS_INDEX_READ_BIT = 0x00000002,
    VK_ACCESS_VERTEX_ATTRIBUTE_READ_BIT = 0x00000004,
    VK_ACCESS_UNIFORM_READ_BIT = 0x00000008,
    VK_ACCESS_INPUT_ATTACHMENT_READ_BIT = 0x00000010,
    VK_ACCESS_SHADER_READ_BIT = 0x00000020,
    VK_ACCESS_SHADER_WRITE_BIT = 0x00000040,
    VK_ACCESS_COLOR_ATTACHMENT_READ_BIT = 0x00000080,
    VK_ACCESS_COLOR_ATTACHMENT_WRITE_BIT = 0x00000100,
    VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_READ_BIT = 0x00000200,
    VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_WRITE_BIT = 0x00000400,
    VK_ACCESS_TRANSFER_READ_BIT = 0x00000800,
    VK_ACCESS_TRANSFER_WRITE_BIT = 0x00001000,
    VK_ACCESS_HOST_READ_BIT = 0x00002000,
    VK_ACCESS_HOST_WRITE_BIT = 0x00004000,
    VK_ACCESS_MEMORY_READ_BIT = 0x00008000,
    VK_ACCESS_MEMORY_WRITE_BIT = 0x00010000,
} VkAccessFlagBits;
```

- **VK_ACCESS_INDIRECT_COMMAND_READ_BIT** specifies read access to an indirect command structure read as part of an indirect drawing or dispatch command.
- **VK_ACCESS_INDEX_READ_BIT** specifies read access to an index buffer as part of an indexed drawing command, bound by `vkCmdBindIndexBuffer`.
- **VK_ACCESS_VERTEX_ATTRIBUTE_READ_BIT** specifies read access to a vertex buffer as part of a drawing command, bound by `vkCmdBindVertexBuffers`.
- **VK_ACCESS_UNIFORM_READ_BIT** specifies read access to a uniform buffer.
- **VK_ACCESS_INPUT_ATTACHMENT_READ_BIT** specifies read access to an input attachment within a render pass during fragment shading.
- **VK_ACCESS_SHADER_READ_BIT** specifies read access to a storage buffer, uniform texel buffer, storage texel buffer, sampled image, or storage image.
VK_ACCESS_SHADER_WRITE_BIT specifies write access to a storage buffer, storage texel buffer, or storage image.

VK_ACCESS_COLOR_ATTACHMENT_READ_BIT specifies read access to a color attachment, such as via blending, logic operations, or via certain subpass load operations.

VK_ACCESS_COLOR_ATTACHMENT_WRITE_BIT specifies write access to a color or resolve attachment during a render pass or via certain subpass load and store operations.

VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_READ_BIT specifies read access to a depth/stencil attachment, via depth or stencil operations or via certain subpass load operations.

VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_WRITE_BIT specifies write access to a depth/stencil attachment, via depth or stencil operations or via certain subpass load and store operations.

VK_ACCESS_TRANSFER_READ_BIT specifies read access to an image or buffer in a copy operation.

VK_ACCESS_TRANSFER_WRITE_BIT specifies write access to an image or buffer in a clear or copy operation.

VK_ACCESS_HOST_READ_BIT specifies read access by a host operation. Accesses of this type are not performed through a resource, but directly on memory.

VK_ACCESS_HOST_WRITE_BIT specifies write access by a host operation. Accesses of this type are not performed through a resource, but directly on memory.

VK_ACCESS_MEMORY_READ_BIT specifies read access via non-specific entities. These entities include the Vulkan device and host, but may also include entities external to the Vulkan device or otherwise not part of the core Vulkan pipeline. When included in a destination access mask, makes all available writes visible to all future read accesses on entities known to the Vulkan device.

VK_ACCESS_MEMORY_WRITE_BIT specifies write access via non-specific entities. These entities include the Vulkan device and host, but may also include entities external to the Vulkan device or otherwise not part of the core Vulkan pipeline. When included in a source access mask, all writes that are performed by entities known to the Vulkan device are made available. When included in a destination access mask, makes all available writes visible to all future write accesses on entities known to the Vulkan device.

Certain access types are only performed by a subset of pipeline stages. Any synchronization command that takes both stage masks and access masks uses both to define the access scopes - only the specified access types performed by the specified stages are included in the access scope. An application must not specify an access flag in a synchronization command if it does not include a pipeline stage in the corresponding stage mask that is able to perform accesses of that type. The following table lists, for each access flag, which pipeline stages can perform that type of access.

Table 4. Supported access types

<table>
<thead>
<tr>
<th>Access flag</th>
<th>Supported pipeline stages</th>
</tr>
</thead>
<tbody>
<tr>
<td>VK_ACCESS_INDIRECT_COMMAND_READ_BIT</td>
<td>VK_PIPELINE_STAGE_DRAW_INDIRECT_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_INDEX_READ_BIT</td>
<td>VK_PIPELINE_STAGE_VERTEX_INPUT_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_VERTEX_ATTRIBUTE_READ_BIT</td>
<td>VK_PIPELINE_STAGE_VERTEX_INPUT_BIT</td>
</tr>
</tbody>
</table>

<p>| VK_ACCESS_INDIRECT_COMMAND_READ_BIT           | VK_PIPELINE_STAGE_DRAW_INDIRECT_BIT       |
| VK_ACCESS_INDEX_READ_BIT                      | VK_PIPELINE_STAGE_VERTEX_INPUT_BIT        |
| VK_ACCESS_VERTEX_ATTRIBUTE_READ_BIT           | VK_PIPELINE_STAGE_VERTEX_INPUT_BIT        |</p>
<table>
<thead>
<tr>
<th>Access flag</th>
<th>Supported pipeline stages</th>
</tr>
</thead>
<tbody>
<tr>
<td>VK_ACCESS_UNIFORM_READ_BIT</td>
<td>VK_PIPELINE_STAGE_VERTEX_SHADER_BIT, VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT, VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT, VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT, VK_PIPELINE_STAGE_FRAGMENT_SHADER_BIT, or VK_PIPELINE_STAGE_COMPUTE_SHADER_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_SHADER_READ_BIT</td>
<td>VK_PIPELINE_STAGE_VERTEX_SHADER_BIT, VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT, VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT, VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT, VK_PIPELINE_STAGE_FRAGMENT_SHADER_BIT, or VK_PIPELINE_STAGE_COMPUTE_SHADER_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_SHADER_WRITE_BIT</td>
<td>VK_PIPELINE_STAGE_VERTEX_SHADER_BIT, VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT, VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT, VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT, VK_PIPELINE_STAGE_FRAGMENT_SHADER_BIT, or VK_PIPELINE_STAGE_COMPUTE_SHADER_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_INPUT_ATTACHMENT_READ_BIT</td>
<td>VK_PIPELINE_STAGE_FRAGMENT_SHADER_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_COLOR_ATTACHMENT_READ_BIT</td>
<td>VK_PIPELINE_STAGE_COLOR_ATTACHMENT_OUTPUT_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_COLOR_ATTACHMENT_WRITE_BIT</td>
<td>VK_PIPELINE_STAGE_COLOR_ATTACHMENT_OUTPUT_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_READ_BIT</td>
<td>VK_PIPELINE_STAGE_EARLY_FRAGMENT_TESTS_BIT, or VK_PIPELINE_STAGE_LATE_FRAGMENT_TESTS_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_WRITE_BIT</td>
<td>VK_PIPELINE_STAGE_EARLY_FRAGMENT_TESTS_BIT, or VK_PIPELINE_STAGE_LATE_FRAGMENT_TESTS_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_TRANSFER_READ_BIT</td>
<td>VK_PIPELINE_STAGE_TRANSFER_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_TRANSFER_WRITE_BIT</td>
<td>VK_PIPELINE_STAGE_TRANSFER_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_HOST_READ_BIT</td>
<td>VK_PIPELINE_STAGE_HOST_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_HOST_WRITE_BIT</td>
<td>VK_PIPELINE_STAGE_HOST_BIT</td>
</tr>
<tr>
<td>VK_ACCESS_MEMORY_READ_BIT</td>
<td>N/A</td>
</tr>
<tr>
<td>VK_ACCESS_MEMORY_WRITE_BIT</td>
<td>N/A</td>
</tr>
</tbody>
</table>

If a memory object does not have the `VK_MEMORY_PROPERTY_HOST_COHERENT_BIT` property, then `vkFlushMappedMemoryRanges` **must** be called in order to guarantee that writes to the memory object from the host are made available to the host domain, where they **can** be further made available to the device domain via a domain operation. Similarly, `vkInvalidateMappedMemoryRanges` **must** be called to guarantee that writes which are available to the host domain are made visible to host operations.

If the memory object does have the `VK_MEMORY_PROPERTY_HOST_COHERENT_BIT` property flag, writes to the memory object from the host are automatically made available to the host domain. Similarly, writes made available to the host domain are automatically made visible to the host.
Note
The `vkQueueSubmit` command automatically performs a domain operation from **host to device** for all writes performed before the command executes, so in most cases an explicit memory barrier is not needed for this case. In the few circumstances where a submit does not occur between the host write and the device read access, writes can be made available by using an explicit memory barrier.

```c
typedef VkFlags VkAccessFlags;
```

`VkAccessFlags` is a bitmask type for setting a mask of zero or more `VkAccessFlagBits`.

### 6.1.4. Framebuffer Region Dependencies

**Pipeline stages** that operate on, or with respect to, the framebuffer are collectively the **framebuffer-space** pipeline stages. These stages are:

- `VK_PIPELINE_STAGE_FRAGMENT_SHADER_BIT`
- `VK_PIPELINE_STAGE_EARLY_FRAGMENT_TESTS_BIT`
- `VK_PIPELINE_STAGE_LATE_FRAGMENT_TESTS_BIT`
- `VK_PIPELINE_STAGE_COLOR_ATTACHMENT_OUTPUT_BIT`

For these pipeline stages, an execution or memory dependency from the first set of operations to the second set can either be a single **framebuffer-global** dependency, or split into multiple **framebuffer-local** dependencies. A dependency with non-framebuffer-space pipeline stages is neither framebuffer-global nor framebuffer-local.

A **framebuffer region** is a set of sample (x, y, layer, sample) coordinates that is a subset of the entire framebuffer.

Both **synchronization scopes** of a framebuffer-local dependency include only the operations performed within corresponding framebuffer regions (as defined below). No ordering guarantees are made between different framebuffer regions for a framebuffer-local dependency.

Both **synchronization scopes** of a framebuffer-global dependency include operations on all framebuffer-regions.

If the first synchronization scope includes operations on pixels/fragments with N samples and the second synchronization scope includes operations on pixels/fragments with M samples, where N does not equal M, then a framebuffer region containing all samples at a given (x, y, layer) coordinate in the first synchronization scope corresponds to a region containing all samples at the same coordinate in the second synchronization scope. In other words, it is a pixel granularity dependency. If N equals M, then a framebuffer region containing a single (x, y, layer, sample) coordinate in the first synchronization scope corresponds to a region containing the same sample at the same coordinate in the second synchronization scope. In other words, it is a sample granularity dependency.
Since fragment invocations are not specified to run in any particular groupings, the size of a framebuffer region is implementation-dependent, not known to the application, and must be assumed to be no larger than specified above.

Practically, the pixel vs sample granularity dependency means that if an input attachment has a different number of samples than the pipeline's rasterizationSamples, then a fragment can access any sample in the input attachment's pixel even if it only uses framebuffer-local dependencies. If the input attachment has the same number of samples, then the fragment can only access the covered samples in its input SampleMask (i.e. the fragment operations happen-after a framebuffer-local dependency for each sample the fragment covers). To access samples that are not covered, a framebuffer-global dependency is required.

If a synchronization command includes a dependencyFlags parameter, and specifies the VK_DEPENDENCY_BY_REGION_BIT flag, then it defines framebuffer-local dependencies for the framebuffer-space pipeline stages in that synchronization command, for all framebuffer regions. If no dependencyFlags parameter is included, or the VK_DEPENDENCY_BY_REGION_BIT flag is not specified, then a framebuffer-global dependency is specified for those stages. The VK_DEPENDENCY_BY_REGION_BIT flag does not affect the dependencies between non-framebuffer-space pipeline stages, nor does it affect the dependencies between framebuffer-space and non-framebuffer-space pipeline stages.

Framebuffer-local dependencies are more optimal for most architectures; particularly tile-based architectures - which can keep framebuffer-regions entirely in on-chip registers and thus avoid external bandwidth across such a dependency. Including a framebuffer-global dependency in your rendering will usually force all implementations to flush data to memory, or to a higher level cache, breaking any potential locality optimizations.

6.2. Implicit Synchronization Guarantees

A small number of implicit ordering guarantees are provided by Vulkan, ensuring that the order in which commands are submitted is meaningful, and avoiding unnecessary complexity in common operations.

Submission order is a fundamental ordering in Vulkan, giving meaning to the order in which action and synchronization commands are recorded and submitted to a single queue. Explicit and implicit ordering guarantees between commands in Vulkan all work on the premise that this ordering is meaningful. This order does not itself define any execution or memory dependencies; synchronization commands and other orderings within the API use this ordering to define their scopes.

Submission order for any given set of commands is based on the order in which they were
recorded to command buffers and then submitted. This order is determined as follows:

1. The initial order is determined by the order in which \texttt{vkQueueSubmit} commands are executed on the host, for a single queue, from first to last.

2. The order in which \texttt{VkSubmitInfo} structures are specified in the \texttt{pSubmits} parameter of \texttt{vkQueueSubmit}, from lowest index to highest.

3. The order in which command buffers are specified in the \texttt{pCommandBuffers} member of \texttt{VkSubmitInfo}, from lowest index to highest.

4. The order in which commands were recorded to a command buffer on the host, from first to last:
   - For commands recorded outside a render pass, this includes all other commands recorded outside a render pass, including \texttt{vkCmdBeginRenderPass} and \texttt{vkCmdEndRenderPass} commands; it does not directly include commands inside a render pass.
   - For commands recorded inside a render pass, this includes all other commands recorded inside the same subpass, including the \texttt{vkCmdBeginRenderPass} and \texttt{vkCmdEndRenderPass} commands that delimit the same render pass instance; it does not include commands recorded to other subpasses.

   \textbf{Action and synchronization commands} recorded to a command buffer execute the \texttt{VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT} pipeline stage in submission order - forming an implicit execution dependency between this stage in each command.

   \textbf{State commands} do not execute any operations on the device, instead they set the state of the command buffer when they execute on the host, in the order that they are recorded. \textbf{Action commands} consume the current state of the command buffer when they are recorded, and will execute state changes on the device as required to match the recorded state.

   \textbf{Query commands}, the order of primitives passing through the graphics pipeline and image layout transitions as part of an image memory barrier provide additional guarantees based on submission order.

   Execution of \textbf{pipeline stages} within a given command also has a loose ordering, dependent only on a single command.

\section*{6.3. Fences}

Fences are a synchronization primitive that can be used to insert a dependency from a queue to the host. Fences have two states - signaled and unsignaled. A fence can be signaled as part of the execution of a queue submission command. Fences can be unsignaled on the host with \texttt{vkResetFences}. Fences can be waited on by the host with the \texttt{vkWaitForFences} command, and the current state can be queried with \texttt{vkGetFenceStatus}.

Fences are represented by \texttt{VkFence} handles:

\begin{verbatim}
VK_DEFINE_NON_DISPATCHABLE_HANDLE(VkFence)
\end{verbatim}
To create a fence, call:

```c
VkResult vkCreateFence(
    VkDevice                                    device,
    const VkFenceCreateInfo*                    pCreateInfo,
    const VkAllocationCallbacks*                pAllocator,
    VkFence*                                    pFence);
```

- `device` is the logical device that creates the fence.
- `pCreateInfo` is a pointer to an instance of the `VkFenceCreateInfo` structure which contains information about how the fence is to be created.
- `pAllocator` controls host memory allocation as described in the `Memory Allocation` chapter.
- `pFence` points to a handle in which the resulting fence object is returned.

### Valid Usage (Implicit)

- `device` **must** be a valid `VkDevice` handle
- `pCreateInfo` **must** be a valid pointer to a valid `VkFenceCreateInfo` structure
- If `pAllocator` is not `NULL`, `pAllocator` **must** be a valid pointer to a valid `VkAllocationCallbacks` structure
- `pFence` **must** be a valid pointer to a `VkFence` handle

### Return Codes

**Success**
- `VK_SUCCESS`

**Failure**
- `VK_ERROR_OUT_OF_HOST_MEMORY`
- `VK_ERROR_OUT_OF_DEVICE_MEMORY`

The `VkFenceCreateInfo` structure is defined as:

```c
typedef struct VkFenceCreateInfo {
    VkStructureType       sType;
    const void*           pNext;
    VkFenceCreateFlags    flags;
} VkFenceCreateInfo;
```

- `sType` is the type of this structure.
- `pNext` is `NULL` or a pointer to an extension-specific structure.
- `flags` is a bitmask of `VkFenceCreateFlagBits` specifying the initial state and behavior of the
fence.

**Valid Usage (Implicit)**

- `sType` **must** be `VK_STRUCTURE_TYPE_FENCE_CREATE_INFO`
- `pNext` **must** be `NULL`
- `flags` **must** be a valid combination of `VkFenceCreateFlagBits` values

```c
typedef enum VkFenceCreateFlagBits {
    VK_FENCE_CREATE_SIGNALED_BIT = 0x00000001,
} VkFenceCreateFlagBits;
```

- `VK_FENCE_CREATE_SIGNALED_BIT` specifies that the fence object is created in the signaled state. Otherwise, it is created in the unsignaled state.

```c
typedef VkFlags VkFenceCreateFlags;
```

`VkFenceCreateFlags` is a bitmask type for setting a mask of zero or more `VkFenceCreateFlagBits`.

To destroy a fence, call:

```c
void vkDestroyFence(
    VkDevice device,
    VkFence fence,
    const VkAllocationCallbacks* pAllocator);
```

- `device` is the logical device that destroys the fence.
- `fence` is the handle of the fence to destroy.
- `pAllocator` controls host memory allocation as described in the Memory Allocation chapter.

**Valid Usage**

- All queue submission commands that refer to `fence` **must** have completed execution
- If `VkAllocationCallbacks` were provided when `fence` was created, a compatible set of callbacks **must** be provided here
- If no `VkAllocationCallbacks` were provided when `fence` was created, `pAllocator` **must** be `NULL`
Valid Usage (Implicit)

- **device** must be a valid `VkDevice` handle.
- If `fence` is not `VK_NULL_HANDLE`, `fence` must be a valid `VkFence` handle.
- If `pAllocator` is not `NULL`, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure.
- If `fence` is a valid handle, it must have been created, allocated, or retrieved from `device`.

Host Synchronization

- Host access to `fence` must be externally synchronized.

To query the status of a fence from the host, call:

```c
VkResult vkGetFenceStatus(
    VkDevice device,            // device is the logical device that owns the fence.
    VkFence fence);            // fence is the handle of the fence to query.
```

Upon success, `vkGetFenceStatus` returns the status of the fence object, with the following return codes:

<table>
<thead>
<tr>
<th>Status</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>VK_SUCCESS</td>
<td>The fence specified by <code>fence</code> is signaled.</td>
</tr>
<tr>
<td>VK_NOT_READY</td>
<td>The fence specified by <code>fence</code> is unsignaled.</td>
</tr>
<tr>
<td>VK_ERROR_DEVICE_LOST</td>
<td>The device has been lost. See Lost Device.</td>
</tr>
</tbody>
</table>

If a queue submission command is pending execution, then the value returned by this command may immediately be out of date.

If the device has been lost (see Lost Device), `vkGetFenceStatus` may return any of the above status codes. If the device has been lost and `vkGetFenceStatus` is called repeatedly, it will eventually return either `VK_SUCCESS` or `VK_ERROR_DEVICE_LOST`.
Valid Usage (Implicit)

- **device must** be a valid VkDevice handle
- **fence must** be a valid VkFence handle
- **fence must** have been created, allocated, or retrieved from device

Return Codes

**Success**
- VK_SUCCESS
- VK_NOT_READY

**Failure**
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY
- VK_ERROR_DEVICE_LOST

To set the state of fences to unsignaled from the host, call:

```c
VkResult vkResetFences(
    VkDevice device,
    uint32_t fenceCount,
    const VkFence* pFences);
```

- **device** is the logical device that owns the fences.
- **fenceCount** is the number of fences to reset.
- **pFences** is a pointer to an array of fence handles to reset.

When `vkResetFences` is executed on the host, it defines a *fence unsignal operation* for each fence, which resets the fence to the unsignaled state.

If any member of **pFences** is already in the unsignaled state when `vkResetFences` is executed, then `vkResetFences` has no effect on that fence.

Valid Usage

- Each element of **pFences must** not be currently associated with any queue command that has not yet completed execution on that queue.
Valid Usage (Implicit)

- **device** must be a valid VkDevice handle
- **pFences** must be a valid pointer to an array of **fenceCount** valid VkFence handles
- **fenceCount** must be greater than 0
- Each element of **pFences** must have been created, allocated, or retrieved from **device**

Host Synchronization

- Host access to each member of **pFences** must be externally synchronized

Return Codes

**Success**
- VK_SUCCESS

**Failure**
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY

When a fence is submitted to a queue as part of a queue submission command, it defines a memory dependency on the batches that were submitted as part of that command, and defines a fence signal operation which sets the fence to the signaled state.

The first synchronization scope includes every batch submitted in the same queue submission command. Fence signal operations that are defined by vkQueueSubmit additionally include in the first synchronization scope all commands that occur earlier in submission order.

The second synchronization scope only includes the fence signal operation.

The first access scope includes all memory access performed by the device.

The second access scope is empty.

To wait for one or more fences to enter the signaled state on the host, call:

```c
VkResult vkWaitForFences(
    VkDevice device,
    uint32_t fenceCount,
    const VkFence* pFences,
    VkBool32 waitAll,
    uint64_t timeout);
```

- **device** is the logical device that owns the fences.
• fenceCount is the number of fences to wait on.
• pFences is a pointer to an array of fenceCount fence handles.
• waitAll is the condition that must be satisfied to successfully unblock the wait. If waitAll is VK_TRUE, then the condition is that all fences in pFences are signaled. Otherwise, the condition is that at least one fence in pFences is signaled.
• timeout is the timeout period in units of nanoseconds. timeout is adjusted to the closest value allowed by the implementation-dependent timeout accuracy, which may be substantially longer than one nanosecond, and may be longer than the requested period.

If the condition is satisfied when vkWaitForFences is called, then vkWaitForFences returns immediately. If the condition is not satisfied at the time vkWaitForFences is called, then vkWaitForFences will block and wait up to timeout nanoseconds for the condition to become satisfied.

If timeout is zero, then vkWaitForFences does not wait, but simply returns the current state of the fences. VK_TIMEOUT will be returned in this case if the condition is not satisfied, even though no actual wait was performed.

If the specified timeout period expires before the condition is satisfied, vkWaitForFences returns VK_TIMEOUT. If the condition is satisfied before timeout nanoseconds has expired, vkWaitForFences returns VK_SUCCESS.

If device loss occurs (see Lost Device) before the timeout has expired, vkWaitForFences must return in finite time with either VK_SUCCESS or VK_ERROR_DEVICE_LOST.

<table>
<thead>
<tr>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>While we guarantee that vkWaitForFences must return in finite time, no guarantees are made that it returns immediately upon device loss. However, the client can reasonably expect that the delay will be on the order of seconds and that calling vkWaitForFences will not result in a permanently (or seemingly permanently) dead process.</td>
</tr>
</tbody>
</table>

Valid Usage (Implicit)

• device must be a valid VkDevice handle
• pFences must be a valid pointer to an array of fenceCount valid VkFence handles
• fenceCount must be greater than 0
• Each element of pFences must have been created, allocated, or retrieved from device
An execution dependency is defined by waiting for a fence to become signaled, either via `vkWaitForFences` or by polling on `vkGetFenceStatus`.

The first synchronization scope includes only the fence signal operation.

The second synchronization scope includes the host operations of `vkWaitForFences` or `vkGetFenceStatus` indicating that the fence has become signaled.

**Note**

Signaling a fence and waiting on the host does not guarantee that the results of memory accesses will be visible to the host, as the access scope of a memory dependency defined by a fence only includes device access. A memory barrier or other memory dependency must be used to guarantee this. See the description of host access types for more information.

### 6.4. Semaphores

Semaphores are a synchronization primitive that can be used to insert a dependency between batches submitted to queues. Semaphores have two states - signaled and unsignaled. The state of a semaphore can be signaled after execution of a batch of commands is completed. A batch can wait for a semaphore to become signaled before it begins execution, and the semaphore is also unsignaled before the batch begins execution.

Semaphores are represented by `VkSemaphore` handles:

```c
VK_DEFINE_NON_DISPATCHABLE_HANDLE(VkSemaphore)
```

To create a semaphore, call:

```c
VkResult vkCreateSemaphore(
    VkDevice                                    device,
    const VkSemaphoreCreateInfo*                pCreateInfo,
    const VkAllocationCallbacks*                pAllocator,
    VkSemaphore*                                pSemaphore);
```
• **device** is the logical device that creates the semaphore.

• **pCreateInfo** is a pointer to an instance of the `VkSemaphoreCreateInfo` structure which contains information about how the semaphore is to be created.

• **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.

• **pSemaphore** points to a handle in which the resulting semaphore object is returned.

When created, the semaphore is in the unsignaled state.

---

### Valid Usage (Implicit)

- **device** must be a valid `VkDevice` handle
- **pCreateInfo** must be a valid pointer to a valid `VkSemaphoreCreateInfo` structure
- If **pAllocator** is not NULL, **pAllocator** must be a valid pointer to a valid `VkAllocationCallbacks` structure
- **pSemaphore** must be a valid pointer to a `VkSemaphore` handle

---

### Return Codes

**Success**

- **VK_SUCCESS**

**Failure**

- **VK_ERROR_OUT_OF_HOST_MEMORY**
- **VK_ERROR_OUT_OF_DEVICE_MEMORY**

The `VkSemaphoreCreateInfo` structure is defined as:

```c
typedef struct VkSemaphoreCreateInfo {
    VkStructureType           sType;
    const void*               pNext;
    VkSemaphoreCreateFlags    flags;
} VkSemaphoreCreateInfo;
```

- **sType** is the type of this structure.
- **pNext** is **NULL** or a pointer to an extension-specific structure.
- **flags** is reserved for future use.
Valid Usage (Implicit)

- sType must be VK_STRUCTURE_TYPE_SEMAPHORE_CREATE_INFO
- pNext must be NULL
- flags must be 0

```c
typedef VkFlags VkSemaphoreCreateFlags;
```

VkSemaphoreCreateFlags is a bitmask type for setting a mask, but is currently reserved for future use.

To destroy a semaphore, call:

```c
void vkDestroySemaphore(
    VkDevice                                    device,
    VkSemaphore                                 semaphore,
    const VkAllocationCallbacks*                pAllocator);
```

- device is the logical device that destroys the semaphore.
- semaphore is the handle of the semaphore to destroy.
- pAllocator controls host memory allocation as described in the Memory Allocation chapter.

Valid Usage

- All submitted batches that refer to semaphore must have completed execution
- If VkAllocationCallbacks were provided when semaphore was created, a compatible set of callbacks must be provided here
- If no VkAllocationCallbacks were provided when semaphore was created, pAllocator must be NULL

Valid Usage (Implicit)

- device must be a valid VkDevice handle
- If semaphore is not VK_NULL_HANDLE, semaphore must be a valid VkSemaphore handle
- If pAllocator is not NULL, pAllocator must be a valid pointer to a valid VkAllocationCallbacks structure
- If semaphore is a valid handle, it must have been created, allocated, or retrieved from device
Host Synchronization

• Host access to semaphore must be externally synchronized

6.4.1. Semaphore Signaling

When a batch is submitted to a queue via a queue submission, and it includes semaphores to be signaled, it defines a memory dependency on the batch, and defines semaphore signal operations which set the semaphores to the signaled state.

The first synchronization scope includes every command submitted in the same batch. Semaphore signal operations that are defined by vkQueueSubmit additionally include all commands that occur earlier in submission order.

The second synchronization scope includes only the semaphore signal operation.

The first access scope includes all memory access performed by the device.

The second access scope is empty.

6.4.2. Semaphore Waiting & Unsignaling

When a batch is submitted to a queue via a queue submission, and it includes semaphores to be waited on, it defines a memory dependency between prior semaphore signal operations and the batch, and defines semaphore unsignal operations which set the semaphores to the unsignaled state.

The first synchronization scope includes all semaphore signal operations that operate on semaphores waited on in the same batch, and that happen-before the wait completes.

The second synchronization scope includes every command submitted in the same batch. In the case of vkQueueSubmit, the second synchronization scope is limited to operations on the pipeline stages determined by the destination stage mask specified by the corresponding element of pWaitDstStageMask. Also, in the case of vkQueueSubmit, the second synchronization scope additionally includes all commands that occur later in submission order.

The first access scope is empty.

The second access scope includes all memory access performed by the device.

The semaphore unsignal operation happens-after the first set of operations in the execution dependency, and happens-before the second set of operations in the execution dependency.
Note
Unlike fences or events, the act of waiting for a semaphore also unsignals that semaphore. If two operations are separately specified to wait for the same semaphore, and there are no other execution dependencies between those operations, behaviour is undefined. An execution dependency must be present that guarantees that the semaphore unsignal operation for the first of those waits, happens-before the semaphore is signalled again, and before the second unsignal operation. Semaphore waits and signals should thus occur in discrete 1:1 pairs.

6.4.3. Semaphore State Requirements For Wait Operations
Before waiting on a semaphore, the application must ensure the semaphore is in a valid state for a wait operation. Specifically, when a semaphore wait and unsignal operation is submitted to a queue:

- The semaphore must be signaled, or have an associated semaphore signal operation that is pending execution.
- There must be no other queue waiting on the same semaphore when the operation executes.

6.5. Events
Events are a synchronization primitive that can be used to insert a fine-grained dependency between commands submitted to the same queue, or between the host and a queue. Events must not be used to insert a dependency between commands submitted to different queues. Events have two states - signaled and unsignaled. An application can signal an event, or unsignal it, on either the host or the device. A device can wait for an event to become signaled before executing further operations. No command exists to wait for an event to become signaled on the host, but the current state of an event can be queried.

Events are represented by VkEvent handles:

```c
VK_DEFINE_NON_DISPATCHABLE_HANDLE(VkEvent)
```

To create an event, call:

```c
VkResult vkCreateEvent(
    VkDevice device,
    const VkEventCreateInfo* pCreateInfo,
    const VkAllocationCallbacks* pAllocator,
    VkEvent* pEvent);
```

- `device` is the logical device that creates the event.
- `pCreateInfo` is a pointer to an instance of the `VkEventCreateInfo` structure which contains information about how the event is to be created.
- `pAllocator` controls host memory allocation as described in the Memory Allocation chapter.
• `pEvent` points to a handle in which the resulting event object is returned.

When created, the event object is in the unsignaled state.

### Valid Usage (Implicit)

- `device` must be a valid `VkDevice` handle
- `pCreateInfo` must be a valid pointer to a valid `VkEventCreateInfo` structure
- If `pAllocator` is not `NULL`, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure
- `pEvent` must be a valid pointer to a `VkEvent` handle

### Return Codes

**Success**
- `VK_SUCCESS`

**Failure**
- `VK_ERROR_OUT_OF_HOST_MEMORY`
- `VK_ERROR_OUT_OF_DEVICE_MEMORY`

The `VkEventCreateInfo` structure is defined as:

```c
typedef struct VkEventCreateInfo {
    VkStructureType       sType;
    const void*           pNext;
    VkEventCreateFlags    flags;
} VkEventCreateInfo;
```

- `sType` is the type of this structure.
- `pNext` is `NULL` or a pointer to an extension-specific structure.
- `flags` is reserved for future use.

### Valid Usage (Implicit)

- `sType` must be `VK_STRUCTURE_TYPE_EVENT_CREATE_INFO`
- `pNext` must be `NULL`
- `flags` must be `0`

```c
typedef VkFlags VkEventCreateFlags;
```
VkEventCreateFlags is a bitmask type for setting a mask, but is currently reserved for future use.

To destroy an event, call:

```c
void vkDestroyEvent(
    VkDevice device,
    VkEvent event,
    const VkAllocationCallbacks* pAllocator);
```

- **device** is the logical device that destroys the event.
- **event** is the handle of the event to destroy.
- **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.

### Valid Usage

- All submitted commands that refer to **event** must have completed execution.
- If **VkAllocationCallbacks** were provided when **event** was created, a compatible set of callbacks must be provided here.
- If no **VkAllocationCallbacks** were provided when **event** was created, **pAllocator** must be NULL.

### Valid Usage (Implicit)

- **device** must be a valid VkDevice handle.
- If **event** is not **VK_NULL_HANDLE**, **event** must be a valid VkEvent handle.
- If **pAllocator** is not NULL, **pAllocator** must be a valid pointer to a valid VkAllocationCallbacks structure.
- If **event** is a valid handle, it must have been created, allocated, or retrieved from **device**.

### Host Synchronization

- Host access to **event** must be externally synchronized.

To query the state of an event from the host, call:

```c
VkResult vkGetEventStatus(
    VkDevice device,
    VkEvent event);
```

- **device** is the logical device that owns the event.
- **event** is the handle of the event to query.
Upon success, `vkGetEventStatus` returns the state of the event object with the following return codes:

<table>
<thead>
<tr>
<th>Status</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>VK_EVENT_SET</td>
<td>The event specified by <code>event</code> is signaled.</td>
</tr>
<tr>
<td>VK_EVENT_RESET</td>
<td>The event specified by <code>event</code> is unsignaled.</td>
</tr>
</tbody>
</table>

If a `vkCmdSetEvent` or `vkCmdResetEvent` command is in a command buffer that is in the pending state, then the value returned by this command may immediately be out of date.

The state of an event can be updated by the host. The state of the event is immediately changed, and subsequent calls to `vkGetEventStatus` will return the new state. If an event is already in the requested state, then updating it to the same state has no effect.

---

**Valid Usage (Implicit)**

- `device` must be a valid `VkDevice` handle
- `event` must be a valid `VkEvent` handle
- `event` must have been created, allocated, or retrieved from `device`

---

**Return Codes**

**Success**
- VK_EVENT_SET
- VK_EVENT_RESET

**Failure**
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY
- VK_ERROR_DEVICE_LOST

To set the state of an event to signaled from the host, call:

```c
VkResult vkSetEvent(
    VkDevice device,
    VkEvent event);
```

- `device` is the logical device that owns the event.
- `event` is the event to set.

When `vkSetEvent` is executed on the host, it defines an event signal operation which sets the event to the signaled state.
If `event` is already in the signaled state when `vkSetEvent` is executed, then `vkSetEvent` has no effect, and no event signal operation occurs.

**Valid Usage (Implicit)**

- `device` must be a valid `VkDevice` handle
- `event` must be a valid `VkEvent` handle
- `event` must have been created, allocated, or retrieved from `device`

**Host Synchronization**

- Host access to `event` must be externally synchronized

**Return Codes**

**Success**
- `VK_SUCCESS`

**Failure**
- `VK_ERROR_OUT_OF_HOST_MEMORY`
- `VK_ERROR_OUT_OF_DEVICE_MEMORY`

To set the state of an event to unsignaled from the host, call:

```c
VkResult vkResetEvent(
    VkDevice   device,   
    VkEvent    event);   
```

- `device` is the logical device that owns the event.
- `event` is the event to reset.

When `vkResetEvent` is executed on the host, it defines an event unsignal operation which resets the event to the unsignaled state.

If `event` is already in the unsignaled state when `vkResetEvent` is executed, then `vkResetEvent` has no effect, and no event unsignal operation occurs.

**Valid Usage**

- `event` must not be waited on by a `vkCmdWaitEvents` command that is currently executing
Valid Usage (Implicit)

- **device** must be a valid `VkDevice` handle
- **event** must be a valid `VkEvent` handle
- **event** must have been created, allocated, or retrieved from `device`

Host Synchronization

- Host access to **event** must be externally synchronized

Return Codes

**Success**
- `VK_SUCCESS`

**Failure**
- `VK_ERROR_OUT_OF_HOST_MEMORY`
- `VK_ERROR_OUT_OF_DEVICE_MEMORY`

The state of an event can also be updated on the device by commands inserted in command buffers.

To set the state of an event to signaled from a device, call:

```c
void vkCmdSetEvent(
    VkCommandBuffer commandBuffer,
    VkEvent event,
    VkPipelineStageFlags stageMask);
```

- **commandBuffer** is the command buffer into which the command is recorded.
- **event** is the event that will be signaled.
- **stageMask** specifies the source stage mask used to determine when the **event** is signaled.

When `vkCmdSetEvent` is submitted to a queue, it defines an execution dependency on commands that were submitted before it, and defines an event signal operation which sets the event to the signaled state.

The first synchronization scope includes all commands that occur earlier in submission order. The synchronization scope is limited to operations on the pipeline stages determined by the source stage mask specified by `stageMask`.

The second synchronization scope includes only the event signal operation.

If **event** is already in the signaled state when `vkCmdSetEvent` is executed on the device, then
vkCmdSetEvent has no effect, no event signal operation occurs, and no execution dependency is generated.

### Valid Usage

- **stageMask** must not include `VK_PIPELINE_STAGE_HOST_BIT`
- If the **geometry shaders** feature is not enabled, **stageMask** must not contain `VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT`
- If the **tessellation shaders** feature is not enabled, **stageMask** must not contain `VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT` or `VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT`

### Valid Usage (Implicit)

- **commandBuffer** must be a valid `VkCommandBuffer` handle
- **event** must be a valid `VkEvent` handle
- **stageMask** must be a valid combination of `VkPipelineStageFlagBits` values
- **stageMask** must not be `0`
- **commandBuffer** must be in the **recording state**
- The **VkCommandPool** that **commandBuffer** was allocated from must support graphics, or compute operations
- This command must only be called outside of a render pass instance
- Both of **commandBuffer**, and **event** must have been created, allocated, or retrieved from the same **VkDevice**

### Host Synchronization

- Host access to **commandBuffer** must be externally synchronized
- Host access to the **VkCommandPool** that **commandBuffer** was allocated from must be externally synchronized

### Command Properties

<table>
<thead>
<tr>
<th>Command Buffer Levels</th>
<th>Render Pass Scope</th>
<th>Supported Queue Types</th>
<th>Pipeline Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>Outside</td>
<td>Graphics</td>
<td>Compute</td>
</tr>
<tr>
<td>Secondary</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To set the state of an event to unsignaled from a device, call:
void vkCmdResetEvent(
    VkCommandBuffer commandBuffer,
    VkEvent event,
    VkPipelineStageFlags stageMask);

- **commandBuffer** is the command buffer into which the command is recorded.
- **event** is the event that will be unsignaled.
- **stageMask** is a bitmask of `VkPipelineStageFlagBits` specifying the source stage mask used to determine when the event is unsignaled.

When `vkCmdResetEvent` is submitted to a queue, it defines an execution dependency on commands that were submitted before it, and defines an event unsignal operation which resets the event to the unsignaled state.

The first synchronization scope includes all commands that occur earlier in submission order. The synchronization scope is limited to operations on the pipeline stages determined by the source stage mask specified by `stageMask`.

The second synchronization scope includes only the event unsignal operation.

If **event** is already in the unsignaled state when `vkCmdResetEvent` is executed on the device, then `vkCmdResetEvent` has no effect, no event unsignal operation occurs, and no execution dependency is generated.

### Valid Usage

- **stageMask** must not include `VK_PIPELINE_STAGE_HOST_BIT`
- If the geometry shaders feature is not enabled, **stageMask** must not contain `VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT`
- If the tessellation shaders feature is not enabled, **stageMask** must not contain `VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT` or `VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT`
- When this command executes, **event** must not be waited on by a `vkCmdWaitEvents` command that is currently executing
Valid Usage (Implicit)

• `commandBuffer` must be a valid `VkCommandBuffer` handle
• `event` must be a valid `VkEvent` handle
• `stageMask` must be a valid combination of `VkPipelineStageFlagBits` values
• `stageMask` must not be 0
• `commandBuffer` must be in the recording state
• The `VkCommandPool` that `commandBuffer` was allocated from must support graphics, or compute operations
• This command must only be called outside of a render pass instance
• Both of `commandBuffer`, and `event` must have been created, allocated, or retrieved from the same `VkDevice`

Host Synchronization

• Host access to `commandBuffer` must be externally synchronized
• Host access to the `VkCommandPool` that `commandBuffer` was allocated from must be externally synchronized

Command Properties

<table>
<thead>
<tr>
<th>Command Buffer Levels</th>
<th>Render Pass Scope</th>
<th>Supported Queue Types</th>
<th>Pipeline Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>Outside</td>
<td>Graphics</td>
<td>Graphics</td>
</tr>
<tr>
<td>Secondary</td>
<td></td>
<td>Compute</td>
<td>Compute</td>
</tr>
</tbody>
</table>

To wait for one or more events to enter the signaled state on a device, call:

```c
void vkCmdWaitEvents(
    VkCommandBuffer commandBuffer,
    uint32_t eventCount,
    const VkEvent* pEvents,
    VkPipelineStageFlags srcStageMask,
    VkPipelineStageFlags dstStageMask,
    uint32_t memoryBarrierCount,
    const VkMemoryBarrier* pMemoryBarriers,
    uint32_t bufferMemoryBarrierCount,
    const VkBufferMemoryBarrier* pBufferMemoryBarriers,
    uint32_t imageMemoryBarrierCount,
    const VkImageMemoryBarrier* pImageMemoryBarriers);
```
- `commandBuffer` is the command buffer into which the command is recorded.
- `eventCount` is the length of the `pEvents` array.
- `pEvents` is an array of event object handles to wait on.
- `srcStageMask` is a bitmask of `VkPipelineStageFlagBits` specifying the source stage mask.
- `dstStageMask` is a bitmask of `VkPipelineStageFlagBits` specifying the destination stage mask.
- `memoryBarrierCount` is the length of the `pMemoryBarriers` array.
- `pMemoryBarriers` is a pointer to an array of `VkMemoryBarrier` structures.
- `bufferMemoryBarrierCount` is the length of the `pBufferMemoryBarriers` array.
- `pBufferMemoryBarriers` is a pointer to an array of `VkBufferMemoryBarrier` structures.
- `imageMemoryBarrierCount` is the length of the `pImageMemoryBarriers` array.
- `pImageMemoryBarriers` is a pointer to an array of `VkImageMemoryBarrier` structures.

When `vkCmdWaitEvents` is submitted to a queue, it defines a memory dependency between prior event signal operations on the same queue or the host, and subsequent commands. `vkCmdWaitEvents` must not be used to wait on event signal operations occurring on other queues.

The first synchronization scope only includes event signal operations that operate on members of `pEvents`, and the operations that happened-before the event signal operations. Event signal operations performed by `vkCmdSetEvent` that occur earlier in submission order are included in the first synchronization scope, if the logically latest pipeline stage in their `stageMask` parameter is logically earlier than or equal to the logically latest pipeline stage in `srcStageMask`. Event signal operations performed by `vkSetEvent` are only included in the first synchronization scope if `VK_PIPELINE_STAGE_HOST_BIT` is included in `srcStageMask`.

The second synchronization scope includes all commands that occur later in submission order. The second synchronization scope is limited to operations on the pipeline stages determined by the destination stage mask specified by `dstStageMask`.

The first access scope is limited to access in the pipeline stages determined by the source stage mask specified by `srcStageMask`. Within that, the first access scope only includes the first access scopes defined by elements of the `pMemoryBarriers`, `pBufferMemoryBarriers` and `pImageMemoryBarriers` arrays, which each define a set of memory barriers. If no memory barriers are specified, then the first access scope includes no accesses.

The second access scope is limited to access in the pipeline stages determined by the destination stage mask specified by `dstStageMask`. Within that, the second access scope only includes the second access scopes defined by elements of the `pMemoryBarriers`, `pBufferMemoryBarriers` and `pImageMemoryBarriers` arrays, which each define a set of memory barriers. If no memory barriers are specified, then the second access scope includes no accesses.
vkCmdWaitEvents is used with vkCmdSetEvent to define a memory dependency between two sets of action commands, roughly in the same way as pipeline barriers, but split into two commands such that work between the two may execute unhindered.

Applications should be careful to avoid race conditions when using events. There is no direct ordering guarantee between a vkCmdResetEvent command and a vkCmdWaitEvents command submitted after it, so some other execution dependency must be included between these commands (e.g. a semaphore).

**Valid Usage**

- **srcStageMask must** be the bitwise OR of the stageMask parameter used in previous calls to vkCmdSetEvent with any of the members of pEvents and VK_PIPELINE_STAGE_HOST_BIT if any of the members of pEvents was set using vkSetEvent
- If the geometry shaders feature is not enabled, **srcStageMask must** not contain VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT
- If the geometry shaders feature is not enabled, **dstStageMask must** not contain VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT
- If the tessellation shaders feature is not enabled, **srcStageMask must** not contain VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT or VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT
- If the tessellation shaders feature is not enabled, **dstStageMask must** not contain VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT or VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT
- If pEvents includes one or more events that will be signaled by vkSetEvent after commandBuffer has been submitted to a queue, then vkCmdWaitEvents must not be called inside a render pass instance
- Any pipeline stage included in srcStageMask or dstStageMask must be supported by the capabilities of the queue family specified by the queueFamilyIndex member of the VkCommandPoolCreateInfo structure that was used to create the VkCommandPool that commandBuffer was allocated from, as specified in the table of supported pipeline stages.
- Each element of pMemoryBarriers, pBufferMemoryBarriers or pImageMemoryBarriers must not have any access flag included in its srcAccessMask member if that bit is not supported by any of the pipeline stages in srcStageMask, as specified in the table of supported access types.
- Each element of pMemoryBarriers, pBufferMemoryBarriers or pImageMemoryBarriers must not have any access flag included in its dstAccessMask member if that bit is not supported by any of the pipeline stages in dstStageMask, as specified in the table of supported access types.
Valid Usage (Implicit)

- `commandBuffer` must be a valid `VkCommandBuffer` handle
- `pEvents` must be a valid pointer to an array of `eventCount` valid `VkEvent` handles
- `srcStageMask` must be a valid combination of `VkPipelineStageFlagBits` values
- `srcStageMask` must not be 0
- `dstStageMask` must be a valid combination of `VkPipelineStageFlagBits` values
- `dstStageMask` must not be 0
- If `memoryBarrierCount` is not 0, `pMemoryBarriers` must be a valid pointer to an array of `memoryBarrierCount` valid `VkMemoryBarrier` structures
- If `bufferMemoryBarrierCount` is not 0, ` pBufferMemoryBarriers` must be a valid pointer to an array of `bufferMemoryBarrierCount` valid `VkBufferMemoryBarrier` structures
- If `imageMemoryBarrierCount` is not 0, `pImageMemoryBarriers` must be a valid pointer to an array of `imageMemoryBarrierCount` valid `VkImageMemoryBarrier` structures
- `commandBuffer` must be in the recording state
- The `VkCommandPool` that `commandBuffer` was allocated from must support graphics, or compute operations
- `eventCount` must be greater than 0
- Both of `commandBuffer`, and the elements of `pEvents` must have been created, allocated, or retrieved from the same `VkDevice`

Host Synchronization

- Host access to `commandBuffer` must be externally synchronized
- Host access to the `VkCommandPool` that `commandBuffer` was allocated from must be externally synchronized

Command Properties

<table>
<thead>
<tr>
<th>Command Buffer Levels</th>
<th>Render Pass Scope</th>
<th>Supported Queue Types</th>
<th>Pipeline Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Secondary</td>
<td>Both</td>
<td>Graphics</td>
<td>Compute</td>
</tr>
</tbody>
</table>

6.6. Pipeline Barriers

`vkCmdPipelineBarrier` is a synchronization command that inserts a dependency between commands submitted to the same queue, or between commands in the same subpass.
To record a pipeline barrier, call:

```c
void vkCmdPipelineBarrier(
    VkCommandBuffer commandBuffer,
    VkPipelineStageFlags srcStageMask,
    VkPipelineStageFlags dstStageMask,
    VkDependencyFlags dependencyFlags,
    uint32_t memoryBarrierCount,
    const VkMemoryBarrier* pMemoryBarriers,
    uint32_t bufferMemoryBarrierCount,
    const VkBufferMemoryBarrier* pBufferMemoryBarriers,
    uint32_t imageMemoryBarrierCount,
    const VkImageMemoryBarrier* pImageMemoryBarriers);
```

- `commandBuffer` is the command buffer into which the command is recorded.
- `srcStageMask` is a bitmask of `VkPipelineStageFlagBits` specifying the source stage mask.
- `dstStageMask` is a bitmask of `VkPipelineStageFlagBits` specifying the destination stage mask.
- `dependencyFlags` is a bitmask of `VkDependencyFlagBits` specifying how execution and memory dependencies are formed.
- `memoryBarrierCount` is the length of the `pMemoryBarriers` array.
- `pMemoryBarriers` is a pointer to an array of `VkMemoryBarrier` structures.
- `bufferMemoryBarrierCount` is the length of the ` pBufferMemoryBarriers` array.
- ` pBufferMemoryBarriers` is a pointer to an array of `VkBufferMemoryBarrier` structures.
- `imageMemoryBarrierCount` is the length of the `pImageMemoryBarriers` array.
- `pImageMemoryBarriers` is a pointer to an array of `VkImageMemoryBarrier` structures.

When `vkCmdPipelineBarrier` is submitted to a queue, it defines a memory dependency between commands that were submitted before it, and those submitted after it.

If `vkCmdPipelineBarrier` was recorded outside a render pass instance, the first synchronization scope includes all commands that occur earlier in submission order. If `vkCmdPipelineBarrier` was recorded inside a render pass instance, the first synchronization scope includes only commands that occur earlier in submission order within the same subpass. In either case, the first synchronization scope is limited to operations on the pipeline stages determined by the source stage mask specified by `srcStageMask`.

If `vkCmdPipelineBarrier` was recorded outside a render pass instance, the second synchronization scope includes all commands that occur later in submission order. If `vkCmdPipelineBarrier` was recorded inside a render pass instance, the second synchronization scope includes only commands that occur later in submission order within the same subpass. In either case, the second synchronization scope is limited to operations on the pipeline stages determined by the destination stage mask specified by `dstStageMask`.

The first access scope is limited to access in the pipeline stages determined by the source stage mask specified by `srcStageMask`. Within that, the first access scope only includes the first access
scopes defined by elements of the `pMemoryBarriers`, `pBufferMemoryBarriers` and `pImageMemoryBarriers` arrays, which each define a set of memory barriers. If no memory barriers are specified, then the first access scope includes no accesses.

The second access scope is limited to access in the pipeline stages determined by the destination stage mask specified by `dstStageMask`. Within that, the second access scope only includes the second access scopes defined by elements of the `pMemoryBarriers`, `pBufferMemoryBarriers` and `pImageMemoryBarriers` arrays, which each define a set of memory barriers. If no memory barriers are specified, then the second access scope includes no accesses.

If `dependencyFlags` includes `VK_DEPENDENCY_BY_REGION_BIT`, then any dependency between framebuffer-space pipeline stages is framebuffer-local - otherwise it is framebuffer-global.
Valid Usage

- If the geometry shaders feature is not enabled, `srcStageMask` must not contain `VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT`.
- If the geometry shaders feature is not enabled, `dstStageMask` must not contain `VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT`.
- If the tessellation shaders feature is not enabled, `srcStageMask` must not contain `VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT` or `VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT`.
- If the tessellation shaders feature is not enabled, `dstStageMask` must not contain `VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT` or `VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT`.
- If `vkCmdPipelineBarrier` is called within a render pass instance, the render pass must have been created with at least one `VkSubpassDependency` instance in `VkRenderPassCreateInfo::pDependencies` that expresses a dependency from the current subpass to itself, and for which `srcStageMask` contains a subset of the bit values in `VkSubpassDependency::srcStageMask`, `dstStageMask` contains a subset of the bit values in `VkSubpassDependency::dstStageMask`, and `dependencyFlags` is equal to `VkSubpassDependency::dependencyFlags`.
- If `vkCmdPipelineBarrier` is called within a render pass instance, for each element of `pMemoryBarriers` and `pImageMemoryBarriers`, the render pass must have been defined with a `VkSubpassDependency` self-dependency for the current subpass with valid `srcStageMask`, `dstStageMask`, and `dependencyFlags` values such that `Vk*Barrier::srcAccessMask` contains a subset of the bit values in `VkSubpassDependency::srcAccessMask` and `Vk*Barrier::dstAccessMask` contains a subset of the bit values in `VkSubpassDependency::dstAccessMask`.
- If `vkCmdPipelineBarrier` is called within a render pass instance, `bufferMemoryBarrierCount` must be 0.
- If `vkCmdPipelineBarrier` is called within a render pass instance, the `image` member of any element of `pImageMemoryBarriers` must be equal to one of the elements of `pAttachments` that the current framebuffer was created with, that is also referred to by one of the elements of the `pColorAttachments`, `pResolveAttachments` or `pDepthStencilAttachment` members of the `VkSubpassDescription` instance that the current subpass was created with.
- If `vkCmdPipelineBarrier` is called within a render pass instance, the `oldLayout` and `newLayout` members of any element of `pImageMemoryBarriers` must be equal to the `layout` member of an element of the `pColorAttachments`, `pResolveAttachments` or `pDepthStencilAttachment` members of the `VkSubpassDescription` instance that the current subpass was created with, that refers to the same `image`.
- If `vkCmdPipelineBarrier` is called within a render pass instance, the `oldLayout` and `newLayout` members of an element of `pImageMemoryBarriers` must be equal.
- If `vkCmdPipelineBarrier` is called within a render pass instance, the `srcQueueFamilyIndex` and `dstQueueFamilyIndex` members of any element of `pImageMemoryBarriers` must be `VK_QUEUE_FAMILY_IGNORED`.
- Any pipeline stage included in `srcStageMask` or `dstStageMask` must be supported by the capabilities of the queue family specified by the `queueFamilyIndex` member of the `VkCommandPoolCreateInfo` structure that was used to create the `VkCommandPool` that
commandBuffer was allocated from, as specified in the table of supported pipeline stages.

- Each element of pMemoryBarriers, pBufferMemoryBarriers and pImageMemoryBarriers must not have any access flag included in its srcAccessMask member if that bit is not supported by any of the pipeline stages in srcStageMask, as specified in the table of supported access types.

- Each element of pMemoryBarriers, pBufferMemoryBarriers and pImageMemoryBarriers must not have any access flag included in its dstAccessMask member if that bit is not supported by any of the pipeline stages in dstStageMask, as specified in the table of supported access types.

Valid Usage (Implicit)

- commandBuffer must be a valid VkCommandBuffer handle
- srcStageMask must be a valid combination of VkPipelineStageFlagBits values
- srcStageMask must not be 0
- dstStageMask must be a valid combination of VkPipelineStageFlagBits values
- dstStageMask must not be 0
- dependencyFlags must be a valid combination of VkDependencyFlagBits values
- If memoryBarrierCount is not 0, pMemoryBarriers must be a valid pointer to an array of memoryBarrierCount valid VkMemoryBarrier structures
- If bufferMemoryBarrierCount is not 0, pBufferMemoryBarriers must be a valid pointer to an array of bufferMemoryBarrierCount valid VkBufferMemoryBarrier structures
- If imageMemoryBarrierCount is not 0, pImageMemoryBarriers must be a valid pointer to an array of imageMemoryBarrierCount valid VkImageMemoryBarrier structures
- commandBuffer must be in the recording state
- The VkCommandPool that commandBuffer was allocated from must support transfer, graphics, or compute operations

Host Synchronization

- Host access to commandBuffer must be externally synchronized
- Host access to the VkCommandPool that commandBuffer was allocated from must be externally synchronized
Bits which can be set in `vkCmdPipelineBarrier::dependencyFlags`, specifying how execution and memory dependencies are formed, are:

```c
typedef enum VkDependencyFlagBits {
    VK_DEPENDENCY_BY_REGION_BIT = 0x00000001,
} VkDependencyFlagBits;
```

- `VK_DEPENDENCY_BY_REGION_BIT` specifies that dependencies will be framebuffer-local.

```c
typedef VkFlags VkDependencyFlags;
```

`VkDependencyFlags` is a bitmask type for setting a mask of zero or more `VkDependencyFlagBits`.

### 6.6.1. Subpass Self-dependency

If `vkCmdPipelineBarrier` is called inside a render pass instance, the following restrictions apply. For a given subpass to allow a pipeline barrier, the render pass must declare a self-dependency from that subpass to itself. That is, there must exist a `VkSubpassDependency` in the subpass dependency list for the render pass with `srcSubpass` and `dstSubpass` equal to that subpass index. More than one self-dependency can be declared for each subpass. Self-dependencies must only include pipeline stage bits that are graphics stages. Self-dependencies must not have any earlier pipeline stages depend on any later pipeline stages (according to the order of graphics pipeline stages), unless all of the stages are framebuffer-space stages. If the source and destination stage masks both include framebuffer-space stages, then `dependencyFlags` must include `VK_DEPENDENCY_BY_REGION_BIT`.

A `vkCmdPipelineBarrier` command inside a render pass instance must be a subset of one of the self-dependencies of the subpass it is used in, meaning that the stage masks and access masks must each include only a subset of the bits of the corresponding mask in that self-dependency. If the self-dependency has `VK_DEPENDENCY_BY_REGION_BIT` set, then so must the pipeline barrier. Pipeline barriers within a render pass instance can only be types `VkMemoryBarrier` or `VkImageMemoryBarrier`. If a `VkImageMemoryBarrier` is used, the image and image subresource range specified in the barrier must be a subset of one of the image views used by the framebuffer in the current subpass. Additionally, `oldLayout` must be equal to `newLayout`, and both the `srcQueueFamilyIndex` and `dstQueueFamilyIndex` must be `VK_QUEUE_FAMILY_IGNORED`.

---

Chapter 6. Synchronization and Cache Control | 123
6.7. Memory Barriers

Memory barriers are used to explicitly control access to buffer and image subresource ranges. Memory barriers are used to transfer ownership between queue families, change image layouts, and define availability and visibility operations. They explicitly define the access types and buffer and image subresource ranges that are included in the access scopes of a memory dependency that is created by a synchronization command that includes them.

6.7.1. Global Memory Barriers

Global memory barriers apply to memory accesses involving all memory objects that exist at the time of its execution.

The VkMemoryBarrier structure is defined as:

```c
typedef struct VkMemoryBarrier {
    VkStructureType    sType;
    const void*        pNext;
    VkAccessFlags      srcAccessMask;
    VkAccessFlags      dstAccessMask;
} VkMemoryBarrier;
```

- **sType** is the type of this structure.
- **pNext** is NULL or a pointer to an extension-specific structure.
- **srcAccessMask** is a bitmask of VkAccessFlagBits specifying a source access mask.
- **dstAccessMask** is a bitmask of VkAccessFlagBits specifying a destination access mask.

The first access scope is limited to access types in the source access mask specified by srcAccessMask.

The second access scope is limited to access types in the destination access mask specified by dstAccessMask.

**Valid Usage (Implicit)**

- **sType** must be VK_STRUCTURE_TYPE_MEMORY_BARRIER
- **pNext** must be NULL
- **srcAccessMask** must be a valid combination of VkAccessFlagBits values
- **dstAccessMask** must be a valid combination of VkAccessFlagBits values

6.7.2. Buffer Memory Barriers

Buffer memory barriers only apply to memory accesses involving a specific buffer range. That is, a memory dependency formed from a buffer memory barrier is scoped to access via the specified buffer range. Buffer memory barriers can also be used to define a queue family ownership transfer...
for the specified buffer range.

The `VkBufferMemoryBarrier` structure is defined as:

```plaintext
typedef struct VkBufferMemoryBarrier {
    VkStructureType    sType;
    const void*        pNext;
    VkAccessFlags      srcAccessMask;
    VkAccessFlags      dstAccessMask;
    uint32_t           srcQueueFamilyIndex;
    uint32_t           dstQueueFamilyIndex;
    VkBuffer           buffer;
    VkDeviceSize       offset;
    VkDeviceSize       size;
} VkBufferMemoryBarrier;
```

- **sType** is the type of this structure.
- **pNext** is **NULL** or a pointer to an extension-specific structure.
- **srcAccessMask** is a bitmask of `VkAccessFlagBits` specifying a source access mask.
- **dstAccessMask** is a bitmask of `VkAccessFlagBits` specifying a destination access mask.
- **srcQueueFamilyIndex** is the source queue family for a queue family ownership transfer.
- **dstQueueFamilyIndex** is the destination queue family for a queue family ownership transfer.
- **buffer** is a handle to the buffer whose backing memory is affected by the barrier.
- **offset** is an offset in bytes into the backing memory for `buffer`; this is relative to the base offset as bound to the buffer (see `vkBindBufferMemory`).
- **size** is a size in bytes of the affected area of backing memory for `buffer`, or `VK_WHOLE_SIZE` to use the range from `offset` to the end of the buffer.

The first access scope is limited to access to memory through the specified buffer range, via access types in the source access mask specified by `srcAccessMask`. If `srcAccessMask` includes `VK_ACCESS_HOST_WRITE_BIT`, memory writes performed by that access type are also made visible, as that access type is not performed through a resource.

The second access scope is limited to access to memory through the specified buffer range, via access types in the destination access mask specified by `dstAccessMask`. If `dstAccessMask` includes `VK_ACCESS_HOST_WRITE_BIT` or `VK_ACCESS_HOST_READ_BIT`, available memory writes are also made visible to accesses of those types, as those access types are not performed through a resource.

If `srcQueueFamilyIndex` is not equal to `dstQueueFamilyIndex`, and `srcQueueFamilyIndex` is equal to the current queue family, then the memory barrier defines a queue family release operation for the specified buffer range, and the second access scope includes no access, as if `dstAccessMask` was `0`.

If `dstQueueFamilyIndex` is not equal to `srcQueueFamilyIndex`, and `dstQueueFamilyIndex` is equal to the current queue family, then the memory barrier defines a queue family acquire operation for the specified buffer range, and the first access scope includes no access, as if `srcAccessMask` was `0`. 
Valid Usage

- **offset must** be less than the size of buffer
- If `size` is not equal to `VK_WHOLE_SIZE`, `size must` be greater than 0
- If `size` is not equal to `VK_WHOLE_SIZE`, `size must` be less than or equal to than the size of `buffer minus offset`

- If `buffer` was created with a sharing mode of `VK_SHARING_MODE_CONCURRENT`, `srcQueueFamilyIndex` and `dstQueueFamilyIndex` must both be `VK_QUEUE_FAMILY_IGNORED`
- If `buffer` was created with a sharing mode of `VK_SHARING_MODE_EXCLUSIVE`, `srcQueueFamilyIndex` and `dstQueueFamilyIndex` must either both be `VK_QUEUE_FAMILY_IGNORED`, or both be a valid queue family (see Queue Family Properties)

- If `buffer` was created with a sharing mode of `VK_SHARING_MODE_EXCLUSIVE`, and `srcQueueFamilyIndex` and `dstQueueFamilyIndex` are not `VK_QUEUE_FAMILY_IGNORED`, at least one of them must be the same as the family of the queue that will execute this barrier
- If `buffer` is non-sparse then it must be bound completely and contiguously to a single `VkDeviceMemory` object

Valid Usage (Implicit)

- **sType** must be `VK_STRUCTURE_TYPE_BUFFER_MEMORY_BARRIER`
- **pNext** must be `NULL`
- **srcAccessMask** must be a valid combination of `VkAccessFlagBits` values
- **dstAccessMask** must be a valid combination of `VkAccessFlagBits` values
- **buffer** must be a valid `VkBuffer` handle

6.7.3. Image Memory Barriers

Image memory barriers only apply to memory accesses involving a specific image subresource range. That is, a memory dependency formed from an image memory barrier is scoped to access via the specified image subresource range. Image memory barriers can also be used to define image layout transitions or a queue family ownership transfer for the specified image subresource range.

The `VkImageMemoryBarrier` structure is defined as:
typedef struct VkImageMemoryBarrier {
    VkStructureType sType;
    const void* pNext;
    VkAccessFlags srcAccessMask;
    VkAccessFlags dstAccessMask;
    VkImageLayout oldLayout;
    VkImageLayout newLayout;
    uint32_t srcQueueFamilyIndex;
    uint32_t dstQueueFamilyIndex;
    VkImage image;
    VkImageSubresourceRange subresourceRange;
} VkImageMemoryBarrier;

- **sType** is the type of this structure.
- **pNext** is NULL or a pointer to an extension-specific structure.
- **srcAccessMask** is a bitmask of VkAccessFlagBits specifying a source access mask.
- **dstAccessMask** is a bitmask of VkAccessFlagBits specifying a destination access mask.
- **oldLayout** is the old layout in an image layout transition.
- **newLayout** is the new layout in an image layout transition.
- **srcQueueFamilyIndex** is the source queue family for a queue family ownership transfer.
- **dstQueueFamilyIndex** is the destination queue family for a queue family ownership transfer.
- **image** is a handle to the image affected by this barrier.
- **subresourceRange** describes the image subresource range within image that is affected by this barrier.

The first access scope is limited to access to memory through the specified image subresource range, via access types in the source access mask specified by srcAccessMask. If srcAccessMask includes VK_ACCESS_HOST_WRITE_BIT, memory writes performed by that access type are also made visible, as that access type is not performed through a resource.

The second access scope is limited to access to memory through the specified image subresource range, via access types in the destination access mask specified by dstAccessMask. If dstAccessMask includes VK_ACCESS_HOST_WRITE_BIT or VK_ACCESS_HOST_READ_BIT, available memory writes are also made visible to accesses of those types, as those access types are not performed through a resource.

If srcQueueFamilyIndex is not equal to dstQueueFamilyIndex, and srcQueueFamilyIndex is equal to the current queue family, then the memory barrier defines a queue family release operation for the specified image subresource range, and the second access scope includes no access, as if dstAccessMask was 0.

If dstQueueFamilyIndex is not equal to srcQueueFamilyIndex, and dstQueueFamilyIndex is equal to the current queue family, then the memory barrier defines a queue family acquire operation for the specified image subresource range, and the first access scope includes no access, as if srcAccessMask was 0.
If `oldLayout` is not equal to `newLayout`, then the memory barrier defines an image layout transition for the specified image subresource range.

Layout transitions that are performed via image memory barriers execute in their entirety in submission order, relative to other image layout transitions submitted to the same queue, including those performed by render passes. In effect there is an implicit execution dependency from each such layout transition to all layout transitions previously submitted to the same queue.
Valid Usage

- `oldLayout` must be `VK_IMAGE_LAYOUT_UNDEFINED` or the current layout of the image subresources affected by the barrier.

- `newLayout` must not be `VK_IMAGE_LAYOUT_UNDEFINED` or `VK_IMAGE_LAYOUT_PREINITIALIZED`.

- If `image` was created with a sharing mode of `VK_SHARING_MODE_CONCURRENT`, `srcQueueFamilyIndex` and `dstQueueFamilyIndex` must both be `VK_QUEUE_FAMILY_IGNORED`.

- If `image` was created with a sharing mode of `VK_SHARING_MODE_EXCLUSIVE`, `srcQueueFamilyIndex` and `dstQueueFamilyIndex` must either both be `VK_QUEUE_FAMILY_IGNORED`, or both be a valid queue family (see Queue Family Properties).

- If `image` was created with a sharing mode of `VK_SHARING_MODE_EXCLUSIVE`, and `srcQueueFamilyIndex` and `dstQueueFamilyIndex` are not `VK_QUEUE_FAMILY_IGNORED`, at least one of them must be the same as the family of the queue that will execute this barrier.

- `subresourceRange.baseMipLevel` must be less than the `mipLevels` specified in `VkImageCreateInfo` when `image` was created.

- If `subresourceRange.levelCount` is not `VK_REMAINING_MIP_LEVELS`, `subresourceRange.baseMipLevel + subresourceRange.levelCount` must be less than or equal to the `mipLevels` specified in `VkImageCreateInfo` when `image` was created.

- `subresourceRange.baseArrayLayer` must be less than the `arrayLayers` specified in `VkImageCreateInfo` when `image` was created.

- If `subresourceRange.layerCount` is not `VK_REMAINING_ARRAY_LAYERS`, `subresourceRange.baseArrayLayer + subresourceRange.layerCount` must be less than or equal to the `arrayLayers` specified in `VkImageCreateInfo` when `image` was created.

- If `image` has a depth/stencil format with both depth and stencil components, then the `aspectMask` member of `subresourceRange` must include both `VK_IMAGE_ASPECT_DEPTH_BIT` and `VK_IMAGE_ASPECT_STENCIL_BIT`.

- If either `oldLayout` or `newLayout` is `VK_IMAGE_LAYOUT_COLOR_ATTACHMENT_OPTIMAL` then `image` must have been created with `VK_IMAGE_USAGE_COLOR_ATTACHMENT_BIT` set.

- If either `oldLayout` or `newLayout` is `VK_IMAGE_LAYOUTDEPTH_STENCIL_ATTACHMENT_OPTIMAL` then `image` must have been created with `VK_IMAGE_USAGE_DEPTH_STENCIL_ATTACHMENT_BIT` set.

- If either `oldLayout` or `newLayout` is `VK_IMAGE_LAYOUTDEPTH_STENCIL_READ_ONLY_OPTIMAL` then `image` must have been created with `VK_IMAGE_USAGE_DEPTH_STENCIL_ATTACHMENT_BIT` set.

- If either `oldLayout` or `newLayout` is `VK_IMAGE_LAYOUT_SHADER_READ_ONLY_OPTIMAL` then `image` must have been created with `VK_IMAGE_USAGE_SAMPLED_BIT` or `VK_IMAGE_USAGE_INPUT_ATTACHMENT_BIT` set.

- If either `oldLayout` or `newLayout` is `VK_IMAGE_LAYOUT_TRANSFER_SRC_OPTIMAL` then `image` must have been created with `VK_IMAGE_USAGE_TRANSFER_SRC_BIT` set.

- If either `oldLayout` or `newLayout` is `VK_IMAGE_LAYOUT_TRANSFER_DST_OPTIMAL` then `image` must have been created with `VK_IMAGE_USAGE_TRANSFER_DST_BIT` set.

- If `image` is non-sparse then it must be bound completely and contiguously to a single `VkDeviceMemory` object.
Valid Usage (Implicit)

- **sType** must be `VK_STRUCTURE_TYPE_IMAGE_MEMORY_BARRIER`.
- **pNext** must be `NULL`.
- **srcAccessMask** must be a valid combination of `VkAccessFlagBits` values.
- **dstAccessMask** must be a valid combination of `VkAccessFlagBits` values.
- **oldLayout** must be a valid `VkImageLayout` value.
- **newLayout** must be a valid `VkImageLayout` value.
- **image** must be a valid `VkImage` handle.
- **subresourceRange** must be a valid `VkImageSubresourceRange` structure.

### 6.7.4. Queue Family Ownership Transfer

Resources created with a `VkSharingMode` of `VK_SHARING_MODE_EXCLUSIVE` must have their ownership explicitly transferred from one queue family to another in order to access their content in a well-defined manner on a queue in a different queue family. If memory dependencies are correctly expressed between uses of such a resource between two queues in different families, but no ownership transfer is defined, the contents of that resource are undefined for any read accesses performed by the second queue family.

*Note*

If an application does not need the contents of a resource to remain valid when transferring from one queue family to another, then the ownership transfer should be skipped.

A queue family ownership transfer consists of two distinct parts:

1. Release exclusive ownership from the source queue family
2. Acquire exclusive ownership for the destination queue family

An application must ensure that these operations occur in the correct order by defining an execution dependency between them, e.g. using a semaphore.

A release operation is used to release exclusive ownership of a range of a buffer or image subresource range. A release operation is defined by executing a buffer memory barrier (for a buffer range) or an image memory barrier (for an image subresource range), on a queue from the source queue family. The `srcQueueFamilyIndex` parameter of the barrier must be set to the source queue family index, and the `dstQueueFamilyIndex` parameter to the destination queue family index. `dstStageMask` is ignored for such a barrier, such that no visibility operation is executed - the value of this mask does not affect the validity of the barrier. The release operation happens-after the availability operation.

An acquire operation is used to acquire exclusive ownership of a range of a buffer or image subresource range. An acquire operation is defined by executing a buffer memory barrier (for a
buffer range) or an image memory barrier (for an image subresource range), on a queue from the destination queue family. The buffer range or image subresource range specified in an acquire operation must match exactly that of a previous release operation. The srcQueueFamilyIndex parameter of the barrier must be set to the source queue family index, and the dstQueueFamilyIndex parameter to the destination queue family index. srcStageMask is ignored for such a barrier, such that no availability operation is executed - the value of this mask does not affect the validity of the barrier. The acquire operation happens-before the visibility operation.

**Note**
Whilst it is not invalid to provide destination or source access masks for memory barriers used for release or acquire operations, respectively, they have no practical effect. Access after a release operation has undefined results, and so visibility for those accesses has no practical effect. Similarly, write access before an acquire operation will produce undefined results for future access, so availability of those writes has no practical use. In an earlier version of the specification, these were required to match on both sides - but this was subsequently relaxed. These masks should be set to 0.

If the transfer is via an image memory barrier, and an image layout transition is desired, then the values of oldLayout and newLayout in the release memory barrier must be equal to values of oldLayout and newLayout in the acquire memory barrier. Although the image layout transition is submitted twice, it will only be executed once. A layout transition specified in this way happens-after the release operation and happens-before the acquire operation.

If the values of srcQueueFamilyIndex and dstQueueFamilyIndex are equal, no ownership transfer is performed, and the barrier operates as if they were both set to VK_QUEUE_FAMILY_IGNORED.

Queue family ownership transfers may perform read and write accesses on all memory bound to the image subresource or buffer range, so applications must ensure that all memory writes have been made available before a queue family ownership transfer is executed. Available memory is automatically made visible to queue family release and acquire operations, and writes performed by those operations are automatically made available.

Once a queue family has acquired ownership of a buffer range or image subresource range of an VK_SHARING_MODE_EXCLUSIVE resource, its contents are undefined to other queue families unless ownership is transferred. The contents of any portion of another resource which aliases memory that is bound to the transferred buffer or image subresource range are undefined after a release or acquire operation.

### 6.8. Wait Idle Operations

To wait on the host for the completion of outstanding queue operations for a given queue, call:

```
VkResult vkQueueWaitIdle(
    VkQueue queue);
```

- queue is the queue on which to wait.
vkQueueWaitIdle is equivalent to submitting a fence to a queue and waiting with an infinite timeout for that fence to signal.

### Valid Usage (Implicit)

- queue must be a valid VkQueue handle

### Command Properties

<table>
<thead>
<tr>
<th>Command Buffer Levels</th>
<th>Render Pass Scope</th>
<th>Supported Queue Types</th>
<th>Pipeline Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>Any</td>
<td>-</td>
</tr>
</tbody>
</table>

### Return Codes

**Success**
- VK_SUCCESS

**Failure**
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY
- VK_ERROR_DEVICE_LOST

To wait on the host for the completion of outstanding queue operations for all queues on a given logical device, call:

```c
VkResult vkDeviceWaitIdle(
    VkDevice device);
```

- device is the logical device to idle.

vkDeviceWaitIdle is equivalent to calling vkQueueWaitIdle for all queues owned by device.

### Valid Usage (Implicit)

- device must be a valid VkDevice handle

### Host Synchronization

- Host access to all VkQueue objects created from device must be externally synchronized
Return Codes

**Success**
- VK_SUCCESS

**Failure**
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY
- VK_ERROR_DEVICE_LOST

6.9. Host Write Ordering Guarantees

When batches of command buffers are submitted to a queue via `vkQueueSubmit`, it defines a memory dependency with prior host operations, and execution of command buffers submitted to the queue.

The first **synchronization scope** is defined by the host execution model, but includes execution of `vkQueueSubmit` on the host and anything that happened-before it.

The second **synchronization scope** includes all commands submitted in the same **queue submission**, and all commands that occur later in **submission order**.

The first **access scope** includes all host writes to mappable device memory that are available to the host memory domain.

The second **access scope** includes all memory access performed by the device.
Chapter 7. Render Pass

A render pass represents a collection of attachments, subpasses, and dependencies between the subpasses, and describes how the attachments are used over the course of the subpasses. The use of a render pass in a command buffer is a render pass instance.

Render passes are represented by VkRenderPass handles:

```cpp
VK_DEFINE_NON_DISPATCHABLE_HANDLE(VkRenderPass)
```

An attachment description describes the properties of an attachment including its format, sample count, and how its contents are treated at the beginning and end of each render pass instance.

A subpass represents a phase of rendering that reads and writes a subset of the attachments in a render pass. Rendering commands are recorded into a particular subpass of a render pass instance.

A subpass description describes the subset of attachments that is involved in the execution of a subpass. Each subpass can read from some attachments as input attachments, write to some as color attachments or depth/stencil attachments, and perform multisample resolve operations to resolve attachments. A subpass description can also include a set of preserve attachments, which are attachments that are not read or written by the subpass but whose contents must be preserved throughout the subpass.

A subpass uses an attachment if the attachment is a color, depth/stencil, resolve, or input attachment for that subpass (as determined by the pColorAttachments, pDepthStencilAttachment, pResolveAttachments, and pInputAttachments members of VkSubpassDescription, respectively). A subpass does not use an attachment if that attachment is preserved by the subpass. The first use of an attachment is in the lowest numbered subpass that uses that attachment. Similarly, the last use of an attachment is in the highest numbered subpass that uses that attachment.

The subpasses in a render pass all render to the same dimensions, and fragments for pixel (x,y,layer) in one subpass can only read attachment contents written by previous subpasses at that same (x,y,layer) location.

**Note**

By describing a complete set of subpasses in advance, render passes provide the implementation an opportunity to optimize the storage and transfer of attachment data between subpasses.

In practice, this means that subpasses with a simple framebuffer-space dependency may be merged into a single tiled rendering pass, keeping the attachment data on-chip for the duration of a render pass instance. However, it is also quite common for a render pass to only contain a single subpass.

Subpass dependencies describe execution and memory dependencies between subpasses.

A subpass dependency chain is a sequence of subpass dependencies in a render pass, where the source subpass of each subpass dependency (after the first) equals the destination subpass of the
previous dependency.

Execution of subpasses may overlap or execute out of order with regards to other subpasses, unless otherwise enforced by an execution dependency. Each subpass only respects submission order for commands recorded in the same subpass, and the `vkCmdBeginRenderPass` and `vkCmdEndRenderPass` commands that delimit the render pass - commands within other subpasses are not included. This affects most other implicit ordering guarantees.

A render pass describes the structure of subpasses and attachments independent of any specific image views for the attachments. The specific image views that will be used for the attachments, and their dimensions, are specified in `VkFramebuffer` objects. Framebuffers are created with respect to a specific render pass that the framebuffer is compatible with (see Render Pass Compatibility). Collectively, a render pass and a framebuffer define the complete render target state for one or more subpasses as well as the algorithmic dependencies between the subpasses.

The various pipeline stages of the drawing commands for a given subpass may execute concurrently and/or out of order, both within and across drawing commands, whilst still respecting pipeline order. However for a given (x,y,layer,sample) sample location, certain per-sample operations are performed in rasterization order.

### 7.1. Render Pass Creation

To create a render pass, call:

```c
VkResult vkCreateRenderPass(  
    VkDevice device,            
    const VkRenderPassCreateInfo* pCreateInfo,  
    const VkAllocationCallbacks* pAllocator,  
    VkRenderPass* pRenderPass);  
```

- `device` is the logical device that creates the render pass.
- `pCreateInfo` is a pointer to an instance of the `VkRenderPassCreateInfo` structure that describes the parameters of the render pass.
- `pAllocator` controls host memory allocation as described in the Memory Allocation chapter.
- `pRenderPass` points to a `VkRenderPass` handle in which the resulting render pass object is returned.

**Valid Usage (Implicit)**

- `device` must be a valid `VkDevice` handle
- `pCreateInfo` must be a valid pointer to a valid `VkRenderPassCreateInfo` structure
- If `pAllocator` is not NULL, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure
- `pRenderPass` must be a valid pointer to a `VkRenderPass` handle
Return Codes

Success
- VK_SUCCESS

Failure
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY

The **VkRenderPassCreateInfo** structure is defined as:

```c
typedef struct VkRenderPassCreateInfo {
    VkStructureType                     sType;
    const void*                         pNext;
    VkRenderPassCreateFlags             flags;
    uint32_t                            attachmentCount;
    const VkAttachmentDescription*      pAttachments;
    uint32_t                            subpassCount;
    const VkSubpassDescription*         pSubpasses;
    uint32_t                            dependencyCount;
    const VkSubpassDependency*          pDependencies;
} VkRenderPassCreateInfo;
```

- **sType** is the type of this structure.
- **pNext** is NULL or a pointer to an extension-specific structure.
- **flags** is reserved for future use.
- **attachmentCount** is the number of attachments used by this render pass, or zero indicating no attachments. Attachments are referred to by zero-based indices in the range [0, attachmentCount).
- **pAttachments** points to an array of attachmentCount number of **VkAttachmentDescription** structures describing properties of the attachments, or NULL if attachmentCount is zero.
- **subpassCount** is the number of subpasses to create for this render pass. Subpasses are referred to by zero-based indices in the range [0, subpassCount). A render pass **must** have at least one subpass.
- **pSubpasses** points to an array of subpassCount number of **VkSubpassDescription** structures describing properties of the subpasses.
- **dependencyCount** is the number of dependencies between pairs of subpasses, or zero indicating no dependencies.
- **pDependencies** points to an array of dependencyCount number of **VkSubpassDependency** structures describing dependencies between pairs of subpasses, or NULL if dependencyCount is zero.
Valid Usage

- If any two subpasses operate on attachments with overlapping ranges of the same VkDeviceMemory object, and at least one subpass writes to that area of VkDeviceMemory, a subpass dependency must be included (either directly or via some intermediate subpasses) between them.

- If the attachment member of any element of pInputAttachments, pColorAttachments, pResolveAttachments or pDepthStencilAttachment, or the attachment indexed by any element of pPreserveAttachments in any element of pSubpasses is bound to a range of a VkDeviceMemory object that overlaps with any other attachment in any subpass (including the same subpass), the VkAttachmentDescription structures describing them must include VK_ATTACHMENT_DESCRIPTION_MAY_ALIAS_BIT in flags.

- If the attachment member of any element of pInputAttachments, pColorAttachments, pResolveAttachments or pDepthStencilAttachment, or any element of pPreserveAttachments in any element of pSubpasses is not VK_ATTACHMENT_UNUSED, it must be less than attachmentCount.

- The value of each element of the pPreserveAttachments member in each element of pSubpasses must not be VK_ATTACHMENT_UNUSED.

- For any member of pAttachments with a loadOp equal to VK_ATTACHMENT_LOAD_OP_CLEAR, the first use of that attachment must not specify a layout equal to VK_IMAGE_LAYOUT_SHADER_READ_ONLY_OPTIMAL or VK_IMAGE_LAYOUT_DEPTH_STENCIL_READ_ONLY_OPTIMAL.

- For any element of pDependencies, if the srcSubpass is not VK_SUBPASS_EXTERNAL, all stage flags included in the srcStageMask member of that dependency must be a pipeline stage supported by the pipeline identified by the pipelineBindPoint member of the source subpass.

- For any element of pDependencies, if the dstSubpass is not VK_SUBPASS_EXTERNAL, all stage flags included in the dstStageMask member of that dependency must be a pipeline stage supported by the pipeline identified by the pipelineBindPoint member of the source subpass.
Valid Usage (Implicit)

- **sType** must be `VK_STRUCTURE_TYPE_RENDER_PASS_CREATE_INFO`
- **pNext** must be `NULL`
- **flags** must be a valid combination of `VkRenderPassCreateFlagBits` values
- If `attachmentCount` is not 0, **pAttachments** must be a valid pointer to an array of `attachmentCount` valid `VkAttachmentDescription` structures
- **pSubpasses** must be a valid pointer to an array of `subpassCount` valid `VkSubpassDescription` structures
- If `dependencyCount` is not 0, **pDependencies** must be a valid pointer to an array of `dependencyCount` valid `VkSubpassDependency` structures
- **subpassCount** must be greater than 0

```c
typedef VkFlags VkRenderPassCreateFlags;

VkRenderPassCreateFlags` is a bitmask type for setting a mask, but is currently reserved for future use.

The `VkAttachmentDescription` structure is defined as:

```c
typedef struct VkAttachmentDescription {
    VkAttachmentDescriptionFlags flags;
    VkFormat format;
    VkSampleCountFlagBits samples;
    VkAttachmentLoadOp loadOp;
    VkAttachmentStoreOp storeOp;
    VkAttachmentLoadOp stencilLoadOp;
    VkAttachmentStoreOp stencilStoreOp;
    VkImageLayout initialLayout;
    VkImageLayout finalLayout;
} VkAttachmentDescription;
```

- **flags** is a bitmask of `VkAttachmentDescriptionFlagBits` specifying additional properties of the attachment.
- **format** is a `VkFormat` value specifying the format of the image view that will be used for the attachment.
- **samples** is the number of samples of the image as defined in `VkSampleCountFlagBits`.
- **loadOp** is a `VkAttachmentLoadOp` value specifying how the contents of color and depth components of the attachment are treated at the beginning of the subpass where it is first used.
- **storeOp** is a `VkAttachmentStoreOp` value specifying how the contents of color and depth components of the attachment are treated at the end of the subpass where it is last used.
• **stencilLoadOp** is a `VkAttachmentLoadOp` value specifying how the contents of stencil components of the attachment are treated at the beginning of the subpass where it is first used.

• **stencilStoreOp** is a `VkAttachmentStoreOp` value specifying how the contents of stencil components of the attachment are treated at the end of the last subpass where it is used.

• **initialLayout** is the layout the attachment image subresource will be in when a render pass instance begins.

• **finalLayout** is the layout the attachment image subresource will be transitioned to when a render pass instance ends. During a render pass instance, an attachment can use a different layout in each subpass, if desired.

If the attachment uses a color format, then **loadOp** and **storeOp** are used, and **stencilLoadOp** and **stencilStoreOp** are ignored. If the format has depth and/or stencil components, **loadOp** and **storeOp** apply only to the depth data, while **stencilLoadOp** and **stencilStoreOp** define how the stencil data is handled. **loadOp** and **stencilLoadOp** define the **load operations** that execute as part of the first subpass that uses the attachment. **storeOp** and **stencilStoreOp** define the **store operations** that execute as part of the last subpass that uses the attachment.

The load operation for each sample in an attachment happens-before any recorded command which accesses the sample in the first subpass where the attachment is used. Load operations for attachments with a depth/stencil format execute in the `VK_PIPELINE_STAGE_EARLY_FRAGMENT_TESTS_BIT` pipeline stage. Load operations for attachments with a color format execute in the `VK_PIPELINE_STAGE_COLOR_ATTACHMENT_OUTPUT_BIT` pipeline stage.

The store operation for each sample in an attachment happens-after any recorded command which accesses the sample in the last subpass where the attachment is used. Store operations for attachments with a depth/stencil format execute in the `VK_PIPELINE_STAGE_LATE_FRAGMENT_TESTS_BIT` pipeline stage. Store operations for attachments with a color format execute in the `VK_PIPELINE_STAGE_COLOR_ATTACHMENT_OUTPUT_BIT` pipeline stage.

If an attachment is not used by any subpass, then **loadOp**, **storeOp**, **stencilStoreOp**, and **stencilLoadOp** are ignored, and the attachment's memory contents will not be modified by execution of a render pass instance.

During a render pass instance, input/color attachments with color formats that have a component size of 8, 16, or 32 bits **must** be represented in the attachment's format throughout the instance. Attachments with other floating- or fixed-point color formats, or with depth components **may** be represented in a format with a precision higher than the attachment format, but **must** be represented with the same range. When such a component is loaded via the **loadOp**, it will be converted into an implementation-dependent format used by the render pass. Such components **must** be converted from the render pass format, to the format of the attachment, before they are resolved or stored at the end of a render pass instance via **storeOp**. Conversions occur as described in Numeric Representation and Computation and Fixed-Point Data Conversions.

If **flags** includes `VK_ATTACHMENT_DESCRIPTION_MAY_ALIAS_BIT`, then the attachment is treated as if it shares physical memory with another attachment in the same render pass. This information limits the ability of the implementation to reorder certain operations (like layout transitions and the **loadOp**) such that it is not improperly reordered against other uses of the same physical memory via a different attachment. This is described in more detail below.
Valid Usage

- `finalLayout` must not be `VK_IMAGE_LAYOUT_UNDEFINED` or `VK_IMAGE_LAYOUT_PREINITIALIZED`

Valid Usage (Implicit)

- `flags` must be a valid combination of `VkAttachmentDescriptionFlagBits` values
- `format` must be a valid `VkFormat` value
- `samples` must be a valid `VkSampleCountFlagBits` value
- `loadOp` must be a valid `VkAttachmentLoadOp` value
- `storeOp` must be a valid `VkAttachmentStoreOp` value
- `stencilLoadOp` must be a valid `VkAttachmentLoadOp` value
- `stencilStoreOp` must be a valid `VkAttachmentStoreOp` value
- `initialLayout` must be a valid `VkImageLayout` value
- `finalLayout` must be a valid `VkImageLayout` value

Bits which can be set in `VkAttachmentDescription::flags` describing additional properties of the attachment are:

```c
typedef enum VkAttachmentDescriptionFlagBits {
    VK_ATTACHMENT_DESCRIPTION_MAY_ALIAS_BIT = 0x00000001,
} VkAttachmentDescriptionFlagBits;
```

- `VK_ATTACHMENT_DESCRIPTION_MAY_ALIAS_BIT` specifies that the attachment aliases the same device memory as other attachments.

```c
typedef VkFlags VkAttachmentDescriptionFlags;
```

`VkAttachmentDescriptionFlags` is a bitmask type for setting a mask of zero or more `VkAttachmentDescriptionFlagBits`.

Possible values of `VkAttachmentDescription::loadOp` and `stencilLoadOp`, specifying how the contents of the attachment are treated, are:

```c
typedef enum VkAttachmentLoadOp {
    VK_ATTACHMENT_LOAD_OP_LOAD = 0,
    VK_ATTACHMENT_LOAD_OP_CLEAR = 1,
    VK_ATTACHMENT_LOAD_OP_DONT_CARE = 2,
} VkAttachmentLoadOp;
```
• **VK_ATTACHMENT_LOAD_OP_LOAD** specifies that the previous contents of the image within the render area will be preserved. For attachments with a depth/stencil format, this uses the access type **VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_READ_BIT**. For attachments with a color format, this uses the access type **VK_ACCESS_COLOR_ATTACHMENT_READ_BIT**.

• **VK_ATTACHMENT_LOAD_OP_CLEAR** specifies that the contents within the render area will be cleared to a uniform value, which is specified when a render pass instance is begun. For attachments with a depth/stencil format, this uses the access type **VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_WRITE_BIT**. For attachments with a color format, this uses the access type **VK_ACCESS_COLOR_ATTACHMENT_WRITE_BIT**.

• **VK_ATTACHMENT_LOAD_OP_DONT_CARE** specifies that the previous contents within the area need not be preserved; the contents of the attachment will be undefined inside the render area. For attachments with a depth/stencil format, this uses the access type **VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_WRITE_BIT**. For attachments with a color format, this uses the access type **VK_ACCESS_COLOR_ATTACHMENT_WRITE_BIT**.

Possible values of **VkAttachmentDescription::storeOp** and **stencilStoreOp**, specifying how the contents of the attachment are treated, are:

```c
typedef enum VkAttachmentStoreOp {
    VK_ATTACHMENT_STORE_OP_STORE = 0,
    VK_ATTACHMENT_STORE_OP_DONT_CARE = 1,
} VkAttachmentStoreOp;
```

• **VK_ATTACHMENT_STORE_OP_STORE** specifies the contents generated during the render pass and within the render area are written to memory. For attachments with a depth/stencil format, this uses the access type **VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_WRITE_BIT**. For attachments with a color format, this uses the access type **VK_ACCESS_COLOR_ATTACHMENT_WRITE_BIT**.

• **VK_ATTACHMENT_STORE_OP_DONT_CARE** specifies the contents within the render area are not needed after rendering, and may be discarded; the contents of the attachment will be undefined inside the render area. For attachments with a depth/stencil format, this uses the access type **VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_WRITE_BIT**. For attachments with a color format, this uses the access type **VK_ACCESS_COLOR_ATTACHMENT_WRITE_BIT**.

If a render pass uses multiple attachments that alias the same device memory, those attachments must each include the **VK_ATTACHMENT_DESCRIPTION_MAY_ALIAS_BIT** bit in their attachment description flags. Attachments aliasing the same memory occurs in multiple ways:

• Multiple attachments being assigned the same image view as part of framebuffer creation.

• Attachments using distinct image views that correspond to the same image subresource of an image.

• Attachments using views of distinct image subresources which are bound to overlapping memory ranges.
Note
Render passes **must** include subpass dependencies (either directly or via a subpass dependency chain) between any two subpasses that operate on the same attachment or aliasing attachments and those subpass dependencies **must** include execution and memory dependencies separating uses of the aliases, if at least one of those subpasses writes to one of the aliases. These dependencies **must** not include the `VK_DEPENDENCY_BY_REGION_BIT` if the aliases are views of distinct image subresources which overlap in memory.

Multiple attachments that alias the same memory **must** not be used in a single subpass. A given attachment index **must** not be used multiple times in a single subpass, with one exception: two subpass attachments **can** use the same attachment index if at least one use is as an input attachment and neither use is as a resolve or preserve attachment. In other words, the same view **can** be used simultaneously as an input and color or depth/stencil attachment, but **must** not be used as multiple color or depth/stencil attachments nor as resolve or preserve attachments. The precise set of valid scenarios is described in more detail below.

If a set of attachments alias each other, then all except the first to be used in the render pass **must** use an **initialLayout** of `VK_IMAGE_LAYOUT_UNDEFINED`, since the earlier uses of the other aliases make their contents undefined. Once an alias has been used and a different alias has been used after it, the first alias **must** not be used in any later subpasses. However, an application **can** assign the same image view to multiple aliasing attachment indices, which allows that image view to be used multiple times even if other aliases are used in between.

Note
Once an attachment needs the `VK_ATTACHMENT_DESCRIPTION_MAY_ALIAS_BIT` bit, there **should** be no additional cost of introducing additional aliases, and using these additional aliases **may** allow more efficient clearing of the attachments on multiple uses via `VK_ATTACHMENT_LOAD_OP_CLEAR`.

The `VkSubpassDescription` structure is defined as:

```c
typedef struct VkSubpassDescription {
    VkSubpassDescriptionFlags flags;
    VkPipelineBindPoint pipelineBindPoint;
    uint32_t inputAttachmentCount;
    const VkAttachmentReference* pInputAttachments;
    uint32_t colorAttachmentCount;
    const VkAttachmentReference* pColorAttachments;
    const VkAttachmentReference* pResolveAttachments;
    const VkAttachmentReference* pDepthStencilAttachment;
    uint32_t preserveAttachmentCount;
    const uint32_t* pPreserveAttachments;
    const uint32_t* pPreserveAttachments;
} VkSubpassDescription;
```

- **flags** is a bitmask of `VkSubpassDescriptionFlagBits` specifying usage of the subpass.
- **pipelineBindPoint** is a `VkPipelineBindPoint` value specifying whether this is a compute or
graphics subpass. Currently, only graphics subpasses are supported.

- **inputAttachmentCount** is the number of input attachments.

- **pInputAttachments** is an array of `VkAttachmentReference` structures (defined below) that lists which of the render pass's attachments can be read in the fragment shader stage during the subpass, and what layout each attachment will be in during the subpass. Each element of the array corresponds to an input attachment unit number in the shader, i.e. if the shader declares an input variable `layout(input_attachment_index=X, set=Y, binding=Z)` then it uses the attachment provided in `pInputAttachments[X]`. Input attachments must also be bound to the pipeline with a descriptor set, with the input attachment descriptor written in the location (set=Y, binding=Z). Fragment shaders can use subpass input variables to access the contents of an input attachment at the fragment's (x, y, layer) framebuffer coordinates.

- **colorAttachmentCount** is the number of color attachments.

- **pColorAttachments** is an array of `colorAttachmentCount` `VkAttachmentReference` structures that lists which of the render pass's attachments will be used as color attachments in the subpass, and what layout each attachment will be in during the subpass. Each element of the array corresponds to a fragment shader output location, i.e. if the shader declared an output variable `layout(location=X)` then it uses the attachment provided in `pColorAttachments[X]`.

- **pResolveAttachments** is NULL or an array of `colorAttachmentCount` `VkAttachmentReference` structures that lists which of the render pass's attachments are resolved to at the end of the subpass, and what layout each attachment will be in during the multisample resolve operation. If `pResolveAttachments` is not NULL, each of its elements corresponds to a color attachment (the element in `pColorAttachments` at the same index), and a multisample resolve operation is defined for each attachment. At the end of each subpass, multisample resolve operations read the subpass's color attachments, and resolve the samples for each pixel to the same pixel location in the corresponding resolve attachments, unless the resolve attachment index is `VK_ATTACHMENT_UNUSED`. If the first use of an attachment in a render pass is as a resolve attachment, then the `loadOp` is effectively ignored as the resolve is guaranteed to overwrite all pixels in the render area.

- **pDepthStencilAttachment** is a pointer to a `VkAttachmentReference` specifying which attachment will be used for depth/stencil data and the layout it will be in during the subpass. Setting the attachment index to `VK_ATTACHMENT_UNUSED` or leaving this pointer as NULL indicates that no depth/stencil attachment will be used in the subpass.

- **preserveAttachmentCount** is the number of preserved attachments.

- **pPreserveAttachments** is an array of `preserveAttachmentCount` render pass attachment indices describing the attachments that are not used by a subpass, but whose contents must be preserved throughout the subpass.

The contents of an attachment within the render area become undefined at the start of a subpass S if all of the following conditions are true:

- The attachment is used as a color, depth/stencil, or resolve attachment in any subpass in the render pass.

- There is a subpass S₁ that uses or preserves the attachment, and a subpass dependency from S₁ to S.
• The attachment is not used or preserved in subpass $S$.

Once the contents of an attachment become undefined in subpass $S$, they remain undefined for subpasses in subpass dependency chains starting with subpass $S$ until they are written again. However, they remain valid for subpasses in other subpass dependency chains starting with subpass $S_i$ if those subpasses use or preserve the attachment.

**Valid Usage**

- **pipelineBindPoint** must be `VK_PIPELINE_BIND_POINT_GRAPHICS`
- **colorAttachmentCount** must be less than or equal to `VkPhysicalDeviceLimits::maxColorAttachments`
- If the first use of an attachment in this render pass is as an input attachment, and the attachment is not also used as a color or depth/stencil attachment in the same subpass, then **loadOp** must not be `VK_ATTACHMENT_LOAD_OP_CLEAR`
- If `pResolveAttachments` is not NULL, for each resolve attachment that does not have the value `VK_ATTACHMENT_UNUSED`, the corresponding color attachment must not have the value `VK_ATTACHMENT_UNUSED`
- If `pResolveAttachments` is not NULL, the sample count of each element of `pColorAttachments` must be anything other than `VK_SAMPLE_COUNT_1_BIT`
- Each element of `pResolveAttachments` must have a sample count of `VK_SAMPLE_COUNT_1_BIT`
- Each element of `pResolveAttachments` must have the same `VkFormat` as its corresponding color attachment
- All attachments in `pColorAttachments` that are not `VK_ATTACHMENT_UNUSED` must have the same sample count
- If `pDepthStencilAttachment` is not `VK_ATTACHMENT_UNUSED` and any attachments in `pColorAttachments` are not `VK_ATTACHMENT_UNUSED`, they must have the same sample count
- If any input attachments are `VK_ATTACHMENT_UNUSED`, then any pipelines bound during the subpass must not access those input attachments from the fragment shader
- The attachment member of each element of `pPreserveAttachments` must not be `VK_ATTACHMENT_UNUSED`
- Each element of `pPreserveAttachments` must not also be an element of any other member of the subpass description
- If any attachment is used as both an input attachment and a color or depth/stencil attachment, then each use must use the same layout
### Valid Usage (Implicit)

- **flags** **must** be a valid combination of `VkSubpassDescriptionFlagBits` values.
- **pipelineBindPoint** **must** be a valid `VkPipelineBindPoint` value.
- If `inputAttachmentCount` is not 0, `pInputAttachments` **must** be a valid pointer to an array of `inputAttachmentCount` valid `VkAttachmentReference` structures.
- If `colorAttachmentCount` is not 0, `pColorAttachments` **must** be a valid pointer to an array of `colorAttachmentCount` valid `VkAttachmentReference` structures.
- If `colorAttachmentCount` is not 0, and `pResolveAttachments` is not NULL, `pResolveAttachments` **must** be a valid pointer to an array of `colorAttachmentCount` valid `VkAttachmentReference` structures.
- If `pDepthStencilAttachment` is not NULL, `pDepthStencilAttachment` **must** be a valid pointer to a valid `VkAttachmentReference` structure.
- If `preserveAttachmentCount` is not 0, `pPreserveAttachments` **must** be a valid pointer to an array of `preserveAttachmentCount` uint32_t values.

Bits which **can** be set in `VkSubpassDescription::flags`, specifying usage of the subpass, are:

```cpp
typedef enum VkSubpassDescriptionFlagBits {
} VkSubpassDescriptionFlagBits;
```

**Note**
All bits for this type are defined by extensions, and none of those extensions are enabled in this build of the specification.

```cpp
typedef VkFlags VkSubpassDescriptionFlags;
```

`VkSubpassDescriptionFlags` is a bitmask type for setting a mask of zero or more `VkSubpassDescriptionFlagBits`.

The `VkAttachmentReference` structure is defined as:

```cpp
typedef struct VkAttachmentReference {
    uint32_t attachment;
    VkImageLayout layout;
} VkAttachmentReference;
```

- **attachment** is the index of the attachment of the render pass, and corresponds to the index of the corresponding element in the `pAttachments` array of the `VkRenderPassCreateInfo` structure. If any color or depth/stencil attachments are `VK_ATTACHMENT_UNUSED`, then no writes occur for those attachments.
• `layout` is a `VkImageLayout` value specifying the layout the attachment uses during the subpass.

---

**Valid Usage**

• `layout` must not be `VK_IMAGE_LAYOUT_UNDEFINED` or `VK_IMAGE_LAYOUT_PREINITIALIZED`

---

**Valid Usage (Implicit)**

• `layout` must be a valid `VkImageLayout` value

---

The `VkSubpassDependency` structure is defined as:

```
typedef struct VkSubpassDependency {
    uint32_t                srcSubpass;
    uint32_t                dstSubpass;
    VkPipelineStageFlags    srcStageMask;
    VkPipelineStageFlags    dstStageMask;
    VkAccessFlags           srcAccessMask;
    VkAccessFlags           dstAccessMask;
    VkDependencyFlags       dependencyFlags;
} VkSubpassDependency;
```

• `srcSubpass` is the subpass index of the first subpass in the dependency, or `VK_SUBPASS_EXTERNAL`.

• `dstSubpass` is the subpass index of the second subpass in the dependency, or `VK_SUBPASS_EXTERNAL`.

• `srcStageMask` is a bitmask of `VkPipelineStageFlagBits` specifying the source stage mask.

• `dstStageMask` is a bitmask of `VkPipelineStageFlagBits` specifying the destination stage mask.

• `srcAccessMask` is a bitmask of `VkAccessFlagBits` specifying a source access mask.

• `dstAccessMask` is a bitmask of `VkAccessFlagBits` specifying a destination access mask.

• `dependencyFlags` is a bitmask of `VkDependencyFlagBits`.

If `srcSubpass` is equal to `dstSubpass` then the `VkSubpassDependency` describes a subpass self-dependency, and only constrains the pipeline barriers allowed within a subpass instance. Otherwise, when a render pass instance which includes a subpass dependency is submitted to a queue, it defines a memory dependency between the subpasses identified by `srcSubpass` and `dstSubpass`.

If `srcSubpass` is equal to `VK_SUBPASS_EXTERNAL`, the first synchronization scope includes commands that occur earlier in submission order than the `vkCmdBeginRenderPass` used to begin the render pass instance. Otherwise, the first set of commands includes all commands submitted as part of the subpass instance identified by `srcSubpass` and any load, store or multisample resolve operations on attachments used in `srcSubpass`. In either case, the first synchronization scope is limited to operations on the pipeline stages determined by the source stage mask specified by `srcStageMask`.  

---

146 | Chapter 7. Render Pass
If `dstSubpass` is equal to `VK_SUBPASS_EXTERNAL`, the second synchronization scope includes commands that occur later in submission order than the `vkCmdEndRenderPass` used to end the render pass instance. Otherwise, the second set of commands includes all commands submitted as part of the subpass instance identified by `dstSubpass` and any load, store or multisample resolve operations on attachments used in `dstSubpass`. In either case, the second synchronization scope is limited to operations on the pipeline stages determined by the destination stage mask specified by `dstStageMask`.

The first access scope is limited to access in the pipeline stages determined by the source stage mask specified by `srcStageMask`. It is also limited to access types in the source access mask specified by `srcAccessMask`.

The second access scope is limited to access in the pipeline stages determined by the destination stage mask specified by `dstStageMask`. It is also limited to access types in the destination access mask specified by `dstAccessMask`.

The availability and visibility operations defined by a subpass dependency affect the execution of image layout transitions within the render pass.

**Note**

For non-attachment resources, the memory dependency expressed by subpass dependency is nearly identical to that of a `VkMemoryBarrier` (with matching `srcAccessMask/dstAccessMask` parameters) submitted as a part of a `vkCmdPipelineBarrier` (with matching `srcStageMask/dstStageMask` parameters). The only difference being that its scopes are limited to the identified subpasses rather than potentially affecting everything before and after.

For attachments however, subpass dependencies work more like a `VkImageMemoryBarrier` defined similarly to the `VkMemoryBarrier` above, the queue family indices set to `VK_QUEUE_FAMILY_IGNORED`, and layouts as follows:

- The equivalent to `oldLayout` is the attachment's layout according to the subpass description for `srcSubpass`.
- The equivalent to `newLayout` is the attachment's layout according to the subpass description for `dstSubpass`.
Valid Usage

• If `srcSubpass` is not `VK_SUBPASS_EXTERNAL`, `srcStageMask` must not include `VK_PIPELINE_STAGE_HOST_BIT`.

• If `dstSubpass` is not `VK_SUBPASS_EXTERNAL`, `dstStageMask` must not include `VK_PIPELINE_STAGE_HOST_BIT`.

• If the `geometry shaders` feature is not enabled, `srcStageMask` must not contain `VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT`.

• If the `geometry shaders` feature is not enabled, `dstStageMask` must not contain `VK_PIPELINE_STAGE_GEOMETRY_SHADER_BIT`.

• If the `tessellation shaders` feature is not enabled, `srcStageMask` must not contain `VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT` or `VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT`.

• If the `tessellation shaders` feature is not enabled, `dstStageMask` must not contain `VK_PIPELINE_STAGE_TESSELLATION_CONTROL_SHADER_BIT` or `VK_PIPELINE_STAGE_TESSELLATION_EVALUATION_SHADER_BIT`.

• `srcSubpass` must be less than or equal to `dstSubpass`, unless one of them is `VK_SUBPASS_EXTERNAL`, to avoid cyclic dependencies and ensure a valid execution order.

• `srcSubpass` and `dstSubpass` must not both be equal to `VK_SUBPASS_EXTERNAL`.

• If `srcSubpass` is equal to `dstSubpass`, `srcStageMask` and `dstStageMask` must not set any bits that are neither `VK_PIPELINE_STAGE_ALL_GRAPHICS_BIT`, nor one of the graphics pipeline stages.

• If `srcSubpass` is equal to `dstSubpass` and not all of the stages in `srcStageMask` and `dstStageMask` are framebuffer-space stages, the logically latest pipeline stage in `srcStageMask` must be logically earlier than or equal to the logically earliest pipeline stage in `dstStageMask`.

• Any access flag included in `srcAccessMask` must be supported by one of the pipeline stages in `srcStageMask`, as specified in the table of supported access types.

• Any access flag included in `dstAccessMask` must be supported by one of the pipeline stages in `dstStageMask`, as specified in the table of supported access types.

• If `srcSubpass` equals `dstSubpass`, and `srcStageMask` and `dstStageMask` both include a framebuffer-space stage, then `dependencyFlags` must include `VK_DEPENDENCY_BY_REGION_BIT`.
Valid Usage (Implicit)

- `srcStageMask` must be a valid combination of `VkPipelineStageFlagBits` values
- `srcStageMask` must not be 0
- `dstStageMask` must be a valid combination of `VkPipelineStageFlagBits` values
- `dstStageMask` must not be 0
- `srcAccessMask` must be a valid combination of `VkAccessFlagBits` values
- `dstAccessMask` must be a valid combination of `VkAccessFlagBits` values
- `dependencyFlags` must be a valid combination of `VkDependencyFlagBits` values

If there is no subpass dependency from `VK_SUBPASS_EXTERNAL` to the first subpass that uses an attachment, then an implicit subpass dependency exists from `VK_SUBPASS_EXTERNAL` to the first subpass it is used in. The subpass dependency operates as if defined with the following parameters:

```c
VkSubpassDependency implicitDependency = {
   .srcSubpass = VK_SUBPASS_EXTERNAL;
   .dstSubpass = firstSubpass; // First subpass attachment is used in
   .srcStageMask = VK_PIPELINE_STAGE_TOP_OF_PIPE_BIT;
   .dstStageMask = VK_PIPELINE_STAGE_ALL_COMMANDS_BIT;
   .srcAccessMask = 0;
   .dstAccessMask = VK_ACCESS_INPUT_ATTACHMENT_READ_BIT |
                    VK_ACCESS_COLOR_ATTACHMENT_READ_BIT |
                    VK_ACCESS_COLOR_ATTACHMENT_WRITE_BIT |
                    VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_READ_BIT |
                    VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_WRITE_BIT;
   .dependencyFlags = 0;
};
```

Similarly, if there is no subpass dependency from the last subpass that uses an attachment to `VK_SUBPASS_EXTERNAL`, then an implicit subpass dependency exists from the last subpass it is used in to `VK_SUBPASS_EXTERNAL`. The subpass dependency operates as if defined with the following parameters:
As subpasses may overlap or execute out of order with regards to other subpasses unless a subpass dependency chain describes otherwise, the layout transitions required between subpasses cannot be known to an application. Instead, an application provides the layout that each attachment must be in at the start and end of a render pass, and the layout it must be in during each subpass it is used in. The implementation then must execute layout transitions between subpasses in order to guarantee that the images are in the layouts required by each subpass, and in the final layout at the end of the render pass.

Automatic layout transitions apply to the entire image subresource attached to the framebuffer.

Automatic layout transitions away from the layout used in a subpass happen-after the availability operations for all dependencies with that subpass as the srcSubpass.

Automatic layout transitions into the layout used in a subpass happen-before the visibility operations for all dependencies with that subpass as the dstSubpass.

Automatic layout transitions away from initialLayout happens-after the availability operations for all dependencies with a srcSubpass equal to VK_SUBPASS_EXTERNAL, where dstSubpass uses the attachment that will be transitioned. For attachments created with VK_ATTACHMENT_DESCRIPTION_MAY_ALIAS_BIT, automatic layout transitions away from initialLayout happen-after the availability operations for all dependencies with a srcSubpass equal to VK_SUBPASS_EXTERNAL, where dstSubpass uses any aliased attachment.

Automatic layout transitions into finalLayout happens-before the visibility operations for all dependencies with a dstSubpass equal to VK_SUBPASS_EXTERNAL, where srcSubpass uses the attachment that will be transitioned. For attachments created with VK_ATTACHMENT_DESCRIPTION_MAY_ALIAS_BIT, automatic layout transitions into finalLayout happen-before the visibility operations for all dependencies with a dstSubpass equal to VK_SUBPASS_EXTERNAL, where srcSubpass uses any aliased attachment.

If two subpasses use the same attachment in different layouts, and both layouts are read-only, no subpass dependency needs to be specified between those subpasses. If an implementation treats those layouts separately, it must insert an implicit subpass dependency between those subpasses to separate the uses in each layout. The subpass dependency operates as if defined with the following parameters:

```cpp
VkSubpassDependency implicitDependency = {
    .srcSubpass = lastSubpass; // Last subpass attachment is used in
    .dstSubpass = VK_SUBPASS_EXTERNAL;
    .srcStageMask = VK_PIPELINE_STAGE_ALL_COMMANDS_BIT;
    .dstStageMask = VK_PIPELINE_STAGE_BOTTOM_OF_PIPE_BIT;
    .srcAccessMask = VK_ACCESS_INPUT_ATTACHMENT_READ_BIT |
                     VK_ACCESS_COLOR_ATTACHMENT_READ_BIT |
                     VK_ACCESS_COLOR_ATTACHMENT_WRITE_BIT |
                     VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_READ_BIT |
                     VK_ACCESS_DEPTH_STENCIL_ATTACHMENT_WRITE_BIT;
    .dstAccessMask = 0;
    .dependencyFlags = 0;
};
```
If a subpass uses the same attachment as both an input attachment and either a color attachment or a depth/stencil attachment, writes via the color or depth/stencil attachment are not automatically made visible to reads via the input attachment, causing a feedback loop, except in any of the following conditions:

- If the color components or depth/stencil components read by the input attachment are mutually exclusive with the components written by the color or depth/stencil attachments, then there is no feedback loop. This requires the graphics pipelines used by the subpass to disable writes to color components that are read as inputs via the `colorWriteMask`, and to disable writes to depth/stencil components that are read as inputs via `depthWriteEnable` or `stencilTestEnable`.

- If the attachment is used as an input attachment and depth/stencil attachment only, and the depth/stencil attachment is not written to.

- If a memory dependency is inserted between when the attachment is written and when it is subsequently read by later fragments. Pipeline barriers expressing a subpass self-dependency are the only way to achieve this, and one must be inserted every time a fragment will read values at a particular sample (x, y, layer, sample) coordinate, if those values have been written since the most recent pipeline barrier; or the since start of the subpass if there have been no pipeline barriers since the start of the subpass.

An attachment used as both an input attachment and a color attachment must be in the `VK_IMAGE_LAYOUT_GENERAL` layout. An attachment used as an input attachment and depth/stencil attachment must be in the `VK_IMAGE_LAYOUT_DEPTH_STENCIL_READ_ONLY_OPTIMAL`, or `VK_IMAGE_LAYOUT_GENERAL` layout. An attachment must not be used as both a depth/stencil attachment and a color attachment.

To destroy a render pass, call:
void vkDestroyRenderPass(
    VkDevice        device,
    VkRenderPass    renderPass,
    const VkAllocationCallbacks* pAllocator);

- **device** is the logical device that destroys the render pass.
- **renderPass** is the handle of the render pass to destroy.
- **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.

### Valid Usage

- All submitted commands that refer to **renderPass** must have completed execution.
- If **VkAllocationCallbacks** were provided when **renderPass** was created, a compatible set of callbacks must be provided here.
- If no **VkAllocationCallbacks** were provided when **renderPass** was created, **pAllocator** must be **NULL**.

### Valid Usage (Implicit)

- **device** must be a valid **VkDevice** handle.
- If **renderPass** is not **VK_NULL_HANDLE**, **renderPass** must be a valid **VkRenderPass** handle.
- If **pAllocator** is not **NULL**, **pAllocator** must be a valid pointer to a valid **VkAllocationCallbacks** structure.
- If **renderPass** is a valid handle, it must have been created, allocated, or retrieved from **device**.

### Host Synchronization

- Host access to **renderPass** must be externally synchronized.

### 7.2. Render Pass Compatibility

Framebuffers and graphics pipelines are created based on a specific render pass object. They must only be used with that render pass object, or one compatible with it.

Two attachment references are compatible if they have matching format and sample count, or are both **VK_ATTACHMENT_UNUSED** or the pointer that would contain the reference is **NULL**.

Two arrays of attachment references are compatible if all corresponding pairs of attachments are compatible. If the arrays are of different lengths, attachment references not present in the smaller array are treated as **VK_ATTACHMENT_UNUSED**.
Two render passes are compatible if their corresponding color, input, resolve, and depth/stencil attachment references are compatible and if they are otherwise identical except for:

- Initial and final image layout in attachment descriptions
- Load and store operations in attachment descriptions
- Image layout in attachment references

A framebuffer is compatible with a render pass if it was created using the same render pass or a compatible render pass.

### 7.3. Framebuffers

Render passes operate in conjunction with *framebuffers*. Framebuffers represent a collection of specific memory attachments that a render pass instance uses.

Framebuffers are represented by `VkFramebuffer` handles:

```cpp
VK_DEFINE_NON_DISPATCHABLE_HANDLE(VkFramebuffer)
```

To create a framebuffer, call:

```cpp
VkResult vkCreateFramebuffer(
    VkDevice                                    device,
    const VkFramebufferCreateInfo*              pCreateInfo,
    const VkAllocationCallbacks*                pAllocator,
    VkFramebuffer*                              pFramebuffer);
```

- `device` is the logical device that creates the framebuffer.
- `pCreateInfo` points to a `VkFramebufferCreateInfo` structure which describes additional information about framebuffer creation.
- `pAllocator` controls host memory allocation as described in the *Memory Allocation* chapter.
- `pFramebuffer` points to a `VkFramebuffer` handle in which the resulting framebuffer object is returned.

### Valid Usage (Implicit)

- `device` must be a valid `VkDevice` handle
- `pCreateInfo` must be a valid pointer to a valid `VkFramebufferCreateInfo` structure
- If `pAllocator` is not NULL, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure
- `pFramebuffer` must be a valid pointer to a `VkFramebuffer` handle
Return Codes

Success
- VK_SUCCESS

Failure
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY

The `VkFramebufferCreateInfo` structure is defined as:

```c
typedef struct VkFramebufferCreateInfo {
    VkStructureType             sType;
    const void*                 pNext;
    VkFramebufferCreateFlags    flags;
    VkRenderPass                renderPass;
    uint32_t                    attachmentCount;
    const VkImageView*          pAttachments;
    uint32_t                    width;
    uint32_t                    height;
    uint32_t                    layers;
} VkFramebufferCreateInfo;
```

- `sType` is the type of this structure.
- `pNext` is NULL or a pointer to an extension-specific structure.
- `flags` is reserved for future use.
- `renderPass` is a render pass that defines what render passes the framebuffer will be compatible with. See Render Pass Compatibility for details.
- `attachmentCount` is the number of attachments.
- `pAttachments` is an array of `VkImageView` handles, each of which will be used as the corresponding attachment in a render pass instance.
- `width`, `height` and `layers` define the dimensions of the framebuffer.

Applications must ensure that all accesses to memory that backs image subresources used as attachments in a given renderpass instance either happen-before the load operations for those attachments, or happen-after the store operations for those attachments.

Note

This restriction means that the render pass has full knowledge of all uses of all of the attachments, so that the implementation is able to make correct decisions about when and how to perform layout transitions, when to overlap execution of subpasses, etc.

It is legal for a subpass to use no color or depth/stencil attachments, and rather use shader side
effects such as image stores and atomics to produce an output. In this case, the subpass continues to use the width, height, and layers of the framebuffer to define the dimensions of the rendering area, and the rasterizationSamples from each pipeline's VkPipelineMultisampleStateCreateInfo to define the number of samples used in rasterization; however, if VkPhysicalDeviceFeatures::variableMultisampleRate is VK_FALSE, then all pipelines to be bound with a given zero-attachment subpass must have the same value for VkPipelineMultisampleStateCreateInfo::rasterizationSamples.

### Valid Usage

- attachmentCount must be equal to the attachment count specified in renderPass

- Each element of pAttachments that is used as a color attachment or resolve attachment by renderPass must have been created with a usage value including VK_IMAGE_USAGE_COLOR_ATTACHMENT_BIT

- Each element of pAttachments that is used as a depth/stencil attachment by renderPass must have been created with a usage value including VK_IMAGE_USAGE_DEPTH_STENCIL_ATTACHMENT_BIT

- Each element of pAttachments that is used as an input attachment by renderPass must have been created with a usage value including VK_IMAGE_USAGE_INPUT_ATTACHMENT_BIT

- Each element of pAttachments must have been created with a VkFormat value that matches the VkFormat specified by the corresponding VkAttachmentDescription in renderPass

- Each element of pAttachments must have been created with a samples value that matches the samples value specified by the corresponding VkAttachmentDescription in renderPass

- Each element of pAttachments must have dimensions at least as large as the corresponding framebuffer dimension

- Each element of pAttachments must only specify a single mip level

- Each element of pAttachments must have been created with the identity swizzle

- width must be greater than 0.

- width must be less than or equal to VkPhysicalDeviceLimits::maxFramebufferWidth

- height must be greater than 0.

- height must be less than or equal to VkPhysicalDeviceLimits::maxFramebufferHeight

- layers must be greater than 0.

- layers must be less than or equal to VkPhysicalDeviceLimits::maxFramebufferLayers
Valid Usage (Implicit)

- **sType** must be `VK_STRUCTURE_TYPE_FRAMEBUFFER_CREATE_INFO`
- **pNext** must be NULL
- **flags** must be 0
- **renderPass** must be a valid `VkRenderPass` handle
- If `attachmentCount` is not 0, **pAttachments** must be a valid pointer to an array of `attachmentCount` valid `VkImageView` handles
- Both of **renderPass**, and the elements of **pAttachments** that are valid handles must have been created, allocated, or retrieved from the same `VkDevice`

```typedef VkFlags VkFramebufferCreateFlags;
```

`VkFramebufferCreateFlags` is a bitmask type for setting a mask, but is currently reserved for future use.

To destroy a framebuffer, call:

```void vkDestroyFramebuffer(
    VkDevice device, 
    VkFramebuffer framebuffer, 
    const VkAllocationCallbacks* pAllocator);
```

- **device** is the logical device that destroys the framebuffer.
- **framebuffer** is the handle of the framebuffer to destroy.
- **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.

Valid Usage

- All submitted commands that refer to **framebuffer** must have completed execution
- If `VkAllocationCallbacks` were provided when **framebuffer** was created, a compatible set of callbacks must be provided here
- If no `VkAllocationCallbacks` were provided when **framebuffer** was created, **pAllocator** must be NULL
Valid Usage (Implicit)

- device must be a valid VkDevice handle
- If framebuffer is not VK_NULL_HANDLE, framebuffer must be a valid VkFramebuffer handle
- If pAllocator is not NULL, pAllocator must be a valid pointer to a valid VkAllocationCallbacks structure
- If framebuffer is a valid handle, it must have been created, allocated, or retrieved from device

Host Synchronization

- Host access to framebuffer must be externally synchronized

7.4. Render Pass Commands

An application records the commands for a render pass instance one subpass at a time, by beginning a render pass instance, iterating over the subpasses to record commands for that subpass, and then ending the render pass instance.

To begin a render pass instance, call:

```c
void vkCmdBeginRenderPass(
    VkCommandBuffer commandBuffer,
    const VkRenderPassBeginInfo* pRenderPassBegin,
    VkSubpassContents contents);
```

- commandBuffer is the command buffer in which to record the command.
- pRenderPassBegin is a pointer to a VkRenderPassBeginInfo structure (defined below) which specifies the render pass to begin an instance of, and the framebuffer the instance uses.
- contents is a VkSubpassContents value specifying how the commands in the first subpass will be provided.

After beginning a render pass instance, the command buffer is ready to record the commands for the first subpass of that render pass.
Valid Usage

- If any of the initialLayout or finalLayout member of the VkAttachmentDescription structures or the layout member of the VkAttachmentReference structures specified when creating the render pass specified in the renderPass member of pRenderPassBegin is VK_IMAGE_LAYOUT_COLOR_ATTACHMENT_OPTIMAL then the corresponding attachment image subresource of the framebuffer specified in the framebuffer member of pRenderPassBegin must have been created with VK_IMAGE_USAGE_COLOR_ATTACHMENT_BIT set.

- If any of the initialLayout or finalLayout member of the VkAttachmentDescription structures or the layout member of the VkAttachmentReference structures specified when creating the render pass specified in the renderPass member of pRenderPassBegin is VK_IMAGE_LAYOUT_DEPTH_STENCIL_ATTACHMENT_OPTIMAL, or VK_IMAGE_LAYOUT_DEPTH_STENCIL_READ_ONLY_OPTIMAL then the corresponding attachment image subresource of the framebuffer specified in the framebuffer member of pRenderPassBegin must have been created with VK_IMAGE_USAGE_DEPTH_STENCIL_ATTACHMENT_BIT set.

- If any of the initialLayout or finalLayout member of the VkAttachmentDescription structures or the layout member of the VkAttachmentReference structures specified when creating the render pass specified in the renderPass member of pRenderPassBegin is VK_IMAGE_LAYOUT_SHADER_READ_ONLY_OPTIMAL then the corresponding attachment image subresource of the framebuffer specified in the framebuffer member of pRenderPassBegin must have been created with VK_IMAGE_USAGE_SAMPLED_BIT or VK_IMAGE_USAGE_INPUT_ATTACHMENT_BIT set.

- If any of the initialLayout members of the VkAttachmentDescription structures specified when creating the render pass specified in the renderPass member of pRenderPassBegin is not VK_IMAGE_LAYOUT_UNDEFINED, then each such initialLayout must be equal to the current layout of the corresponding attachment image subresource of the framebuffer specified in the framebuffer member of pRenderPassBegin.

- The srcStageMask and dstStageMask members of any element of the pDependencies member of VkRenderPassCreateInfo used to create renderPass must be supported by the capabilities of the queue family identified by the queueFamilyIndex member of the VkCommandPoolCreateInfo used to create the command pool which commandBuffer was allocated from.
Valid Usage (Implicit)

- `commandBuffer` must be a valid `VkCommandBuffer` handle
- `pRenderPassBegin` must be a valid pointer to a valid `VkRenderPassBeginInfo` structure
- `contents` must be a valid `VkSubpassContents` value
- `commandBuffer` must be in the recording state
- The `VkCommandPool` that `commandBuffer` was allocated from must support graphics operations
- This command must only be called outside of a render pass instance
- `commandBuffer` must be a primary `VkCommandBuffer`

Host Synchronization

- Host access to `commandBuffer` must be externally synchronized
- Host access to the `VkCommandPool` that `commandBuffer` was allocated from must be externally synchronized

Command Properties

<table>
<thead>
<tr>
<th>Command Buffer Levels</th>
<th>Render Pass Scope</th>
<th>Supported Queue Types</th>
<th>Pipeline Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>Outside</td>
<td>Graphics</td>
<td>Graphics</td>
</tr>
</tbody>
</table>

The `VkRenderPassBeginInfo` structure is defined as:

```c
typedef struct VkRenderPassBeginInfo {
    VkStructureType        sType;
    const void*            pNext;  
    VkRenderPass           renderPass;
    VkFramebuffer          framebuffer;
    VkRect2D               renderArea;
    uint32_t               clearValueCount;
    const VkClearColor*    pClearValues;
} VkRenderPassBeginInfo;
```

- `sType` is the type of this structure.
- `pNext` is `NULL` or a pointer to an extension-specific structure.
- `renderPass` is the render pass to begin an instance of.
- `framebuffer` is the framebuffer containing the attachments that are used with the render pass.
• **renderArea** is the render area that is affected by the render pass instance, and is described in more detail below.

• **clearValueCount** is the number of elements in `pClearValues`.

• **pClearValues** is an array of `VkClearValue` structures that contains clear values for each attachment, if the attachment uses a `loadOp` value of `VK_ATTACHMENT_LOAD_OP_CLEAR` or if the attachment has a depth/stencil format and uses a `stencilLoadOp` value of `VK_ATTACHMENT_LOAD_OP_CLEAR`. The array is indexed by attachment number. Only elements corresponding to cleared attachments are used. Other elements of `pClearValues` are ignored.

`renderArea` is the render area that is affected by the render pass instance. The effects of attachment load, store and multisample resolve operations are restricted to the pixels whose x and y coordinates fall within the render area on all attachments. The render area extends to all layers of `framebuffer`. The application **must** ensure (using scissor if necessary) that all rendering is contained within the render area, otherwise the pixels outside of the render area become undefined and shader side effects **may** occur for fragments outside the render area. The render area **must** be contained within the framebuffer dimensions.

**Note**

There **may** be a performance cost for using a render area smaller than the framebuffer, unless it matches the render area granularity for the render pass.

### Valid Usage

- **clearValueCount** **must** be greater than the largest attachment index in `renderPass` that specifies a `loadOp` (or `stencilLoadOp`, if the attachment has a depth/stencil format) of `VK_ATTACHMENT_LOAD_OP_CLEAR`

- If `clearValueCount` is not 0, `pClearValues` **must** be a valid pointer to an array of `clearValueCount` valid `VkClearValue` unions

- `renderPass` **must** be compatible with the `renderPass` member of the `VkFramebufferCreateInfo` structure specified when creating `framebuffer`.

### Valid Usage (Implicit)

- **sType** **must** be `VK_STRUCTURE_TYPE_RENDER_PASS_BEGIN_INFO`
- **pNext** **must** be `NULL`
- `renderPass` **must** be a valid `VkRenderPass` handle
- `framebuffer` **must** be a valid `VkFramebuffer` handle
- Both of `framebuffer`, and `renderPass` **must** have been created, allocated, or retrieved from the same `VkDevice`

Possible values of `vkCmdBeginRenderPass::contents`, specifying how the commands in the first subpass will be provided, are:
typedef enum VkSubpassContents {
    VK_SUBPASS_CONTENTS_INLINE = 0,
    VK_SUBPASS_CONTENTS_SECONDARY_COMMAND_BUFFERS = 1,
} VkSubpassContents;

- **VK_SUBPASS_CONTENTS_INLINE** specifies that the contents of the subpass will be recorded inline in the primary command buffer, and secondary command buffers must not be executed within the subpass.

- **VK_SUBPASS_CONTENTS_SECONDARY_COMMAND_BUFFERS** specifies that the contents are recorded in secondary command buffers that will be called from the primary command buffer, and `vkCmdExecuteCommands` is the only valid command on the command buffer until `vkCmdNextSubpass` or `vkCmdEndRenderPass`.

To query the render area granularity, call:

```c
void vkGetRenderAreaGranularity(
    VkDevice device,       
    VkRenderPass renderPass, 
    VkExtent2D* pGranularity);
```

- **device** is the logical device that owns the render pass.
- **renderPass** is a handle to a render pass.
- **pGranularity** points to a `VkExtent2D` structure in which the granularity is returned.

The conditions leading to an optimal `renderArea` are:

- The `offset.x` member in `renderArea` is a multiple of the `width` member of the returned `VkExtent2D` (the horizontal granularity).

- The `offset.y` member in `renderArea` is a multiple of the `height` of the returned `VkExtent2D` (the vertical granularity).

- Either the `offset.width` member in `renderArea` is a multiple of the horizontal granularity or `offset.x+offset.width` is equal to the `width` of the framebuffer in the `VkRenderPassBeginInfo`.

- Either the `offset.height` member in `renderArea` is a multiple of the vertical granularity or `offset.y+offset.height` is equal to the `height` of the framebuffer in the `VkRenderPassBeginInfo`.

Subpass dependencies are not affected by the render area, and apply to the entire image subresources attached to the framebuffer as specified in the description of automatic layout transitions. Similarly, pipeline barriers are valid even if their effect extends outside the render area.
Valid Usage (Implicit)

- **device** **must** be a valid `VkDevice` handle
- **renderPass** **must** be a valid `VkRenderPass` handle
- **pGranularity** **must** be a valid pointer to a `VkExtent2D` structure
- **renderPass** **must** have been created, allocated, or retrieved from `device`

To transition to the next subpass in the render pass instance after recording the commands for a subpass, call:

```c
void vkCmdNextSubpass(
    VkCommandBuffer commandBuffer,
    VkSubpassContents contents);
```

- **commandBuffer** is the command buffer in which to record the command.
- **contents** specifies how the commands in the next subpass will be provided, in the same fashion as the corresponding parameter of `vkCmdBeginRenderPass`.

The subpass index for a render pass begins at zero when `vkCmdBeginRenderPass` is recorded, and increments each time `vkCmdNextSubpass` is recorded.

Moving to the next subpass automatically performs any multisample resolve operations in the subpass being ended. End-of-subpass multisample resolves are treated as color attachment writes for the purposes of synchronization. That is, they are considered to execute in the `VK_PIPELINE_STAGE_COLOR_ATTACHMENT_OUTPUT_BIT` pipeline stage and their writes are synchronized with `VK_ACCESS_COLOR_ATTACHMENT_WRITE_BIT`. Synchronization between rendering within a subpass and any resolve operations at the end of the subpass occurs automatically, without need for explicit dependencies or pipeline barriers. However, if the resolve attachment is also used in a different subpass, an explicit dependency is needed.

After transitioning to the next subpass, the application **can** record the commands for that subpass.

Valid Usage

- The current subpass index **must** be less than the number of subpasses in the render pass minus one
Valid Usage (Implicit)

- `commandBuffer` must be a valid `VkCommandBuffer` handle
- `contents` must be a valid `VkSubpassContents` value
- `commandBuffer` must be in the recording state
- The `VkCommandPool` that `commandBuffer` was allocated from must support graphics operations
- This command must only be called inside of a render pass instance
- `commandBuffer` must be a primary `VkCommandBuffer`

Host Synchronization

- Host access to `commandBuffer` must be externally synchronized
- Host access to the `VkCommandPool` that `commandBuffer` was allocated from must be externally synchronized

Command Properties

<table>
<thead>
<tr>
<th>Command Buffer Levels</th>
<th>Render Pass Scope</th>
<th>Supported Queue Types</th>
<th>Pipeline Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>Inside</td>
<td>Graphics</td>
<td>Graphics</td>
</tr>
</tbody>
</table>

To record a command to end a render pass instance after recording the commands for the last subpass, call:

```c
void vkCmdEndRenderPass(
    VkCommandBuffer commandBuffer);
```

- `commandBuffer` is the command buffer in which to end the current render pass instance.

Ending a render pass instance performs any multisample resolve operations on the final subpass.

Valid Usage

- The current subpass index must be equal to the number of subpasses in the render pass minus one
Valid Usage (Implicit)

- `commandBuffer` must be a valid `VkCommandBuffer` handle
- `commandBuffer` must be in the recording state
- The `VkCommandPool` that `commandBuffer` was allocated from must support graphics operations
- This command must only be called inside of a render pass instance
- `commandBuffer` must be a primary `VkCommandBuffer`

Host Synchronization

- Host access to `commandBuffer` must be externally synchronized
- Host access to the `VkCommandPool` that `commandBuffer` was allocated from must be externally synchronized

Command Properties

<table>
<thead>
<tr>
<th>Command Buffer Levels</th>
<th>Render Pass Scope</th>
<th>Supported Queue Types</th>
<th>Pipeline Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>Inside</td>
<td>Graphics</td>
<td>Graphics</td>
</tr>
</tbody>
</table>
Chapter 8. Shaders

A shader specifies programmable operations that execute for each vertex, control point, tessellated vertex, primitive, fragment, or workgroup in the corresponding stage(s) of the graphics and compute pipelines.

Graphics pipelines include vertex shader execution as a result of primitive assembly, followed, if enabled, by tessellation control and evaluation shaders operating on patches, geometry shaders, if enabled, operating on primitives, and fragment shaders, if present, operating on fragments generated by Rasterization. In this specification, vertex, tessellation control, tessellation evaluation and geometry shaders are collectively referred to as vertex processing stages and occur in the logical pipeline before rasterization. The fragment shader occurs logically after rasterization.

Only the compute shader stage is included in a compute pipeline. Compute shaders operate on compute invocations in a workgroup.

Shaders can read from input variables, and read from and write to output variables. Input and output variables can be used to transfer data between shader stages, or to allow the shader to interact with values that exist in the execution environment. Similarly, the execution environment provides constants that describe capabilities.

Shader variables are associated with execution environment-provided inputs and outputs using built-in decorations in the shader. The available decorations for each stage are documented in the following subsections.

8.1. Shader Modules

Shader modules contain shader code and one or more entry points. Shaders are selected from a shader module by specifying an entry point as part of pipeline creation. The stages of a pipeline can use shaders that come from different modules. The shader code defining a shader module must be in the SPIR-V format, as described by the Vulkan Environment for SPIR-V appendix.

Shader modules are represented by VkShaderModule handles:

```
VK_DEFINE_NON_DISPATCHABLE_HANDLE(VkShaderModule)
```

To create a shader module, call:

```
VkResult vkCreateShaderModule(
    VkDevice device,
    const VkShaderModuleCreateInfo* pCreateInfo,
    const VkAllocationCallbacks* pAllocator,
    VkShaderModule* pShaderModule);
```

- device is the logical device that creates the shader module.
- pCreateInfo is a pointer to an instance of the VkShaderModuleCreateInfo structure.
• `pAllocator` controls host memory allocation as described in the Memory Allocation chapter.

• `pShaderModule` points to a `VkShaderModule` handle in which the resulting shader module object is returned.

Once a shader module has been created, any entry points it contains can be used in pipeline shader stages as described in Compute Pipelines and Graphics Pipelines.

**Valid Usage (Implicit)**

• `device` must be a valid `VkDevice` handle

• `pCreateInfo` must be a valid pointer to a valid `VkShaderModuleCreateInfo` structure

• If `pAllocator` is not `NULL`, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure

• `pShaderModule` must be a valid pointer to a `VkShaderModule` handle

**Return Codes**

**Success**

• `VK_SUCCESS`

**Failure**

• `VK_ERROR_OUT_OF_HOST_MEMORY`

• `VK_ERROR_OUT_OF_DEVICE_MEMORY`

The `VkShaderModuleCreateInfo` structure is defined as:

```c
typedef struct VkShaderModuleCreateInfo {
    VkStructureType              sType;
    const void*                  pNext;
    VkShaderModuleCreateFlags    flags;
    size_t                       codeSize;
    const uint32_t*              pCode;
} VkShaderModuleCreateInfo;
```

• `sType` is the type of this structure.

• `pNext` is `NULL` or a pointer to an extension-specific structure.

• `flags` is reserved for future use.

• `codeSize` is the size, in bytes, of the code pointed to by `pCode`.

• `pCode` points to code that is used to create the shader module. The type and format of the code is determined from the content of the memory addressed by `pCode`. 
Valid Usage

- codeSize must be greater than 0
- codeSize must be a multiple of 4
- pCode must point to valid SPIR-V code, formatted and packed as described by the Khronos SPIR-V Specification
- pCode must adhere to the validation rules described by the Validation Rules within a Module section of the SPIR-V Environment appendix
- pCode must declare the Shader capability for SPIR-V code
- pCode must not declare any capability that is not supported by the API, as described by the Capabilities section of the SPIR-V Environment appendix
- If pCode declares any of the capabilities listed as optional in the SPIR-V Environment appendix, the corresponding feature(s) must be enabled.

Valid Usage (Implicit)

- sType must be VK_STRUCTURE_TYPE_SHADER_MODULE_CREATE_INFO
- pNext must be NULL
- flags must be 0
- pCode must be a valid pointer to an array of \( \text{codeSize} \times 4 \) uint32_t values

```c
typedef VkFlags VkShaderModuleCreateFlags;
```

VkShaderModuleCreateFlags is a bitmask type for setting a mask, but is currently reserved for future use.

To destroy a shader module, call:

```c
void vkDestroyShaderModule(
    VkDevice device,  // device, the logical device that destroys the shader module.
    VkShaderModule shaderModule,  // shaderModule, the handle of the shader module to destroy.
    const VkAllocationCallbacks* pAllocator);  // pAllocator controls host memory allocation as described in the Memory Allocation chapter.
```

A shader module can be destroyed while pipelines created using its shaders are still in use.
Valid Usage

• If `VkAllocationCallbacks` were provided when `shaderModule` was created, a compatible set of callbacks **must** be provided here.

• If no `VkAllocationCallbacks` were provided when `shaderModule` was created, `pAllocator` **must** be `NULL`.

Valid Usage (Implicit)

• `device` **must** be a valid `VkDevice` handle.

• If `shaderModule` is not `VK_NULL_HANDLE`, `shaderModule` **must** be a valid `VkShaderModule` handle.

• If `pAllocator` is not `NULL`, `pAllocator` **must** be a valid pointer to a valid `VkAllocationCallbacks` structure.

• If `shaderModule` is a valid handle, it **must** have been created, allocated, or retrieved from `device`.

Host Synchronization

• Host access to `shaderModule` **must** be externally synchronized.

8.2. Shader Execution

At each stage of the pipeline, multiple invocations of a shader **may** execute simultaneously. Further, invocations of a single shader produced as the result of different commands **may** execute simultaneously. The relative execution order of invocations of the same shader type is undefined. Shader invocations **may** complete in a different order than that in which the primitives they originated from were drawn or dispatched by the application. However, fragment shader outputs are written to attachments in rasterization order.

The relative order of invocations of different shader types is largely undefined. However, when invoking a shader whose inputs are generated from a previous pipeline stage, the shader invocations from the previous stage are guaranteed to have executed far enough to generate input values for all required inputs.

8.3. Shader Memory Access Ordering

The order in which image or buffer memory is read or written by shaders is largely undefined. For some shader types (vertex, tessellation evaluation, and in some cases, fragment), even the number of shader invocations that **may** perform loads and stores is undefined.

In particular, the following rules apply:
• **Vertex** and **tessellation evaluation** shaders will be invoked at least once for each unique vertex, as defined in those sections.

• **Fragment** shaders will be invoked zero or more times, as defined in that section.

• The relative order of invocations of the same shader type are undefined. A store issued by a shader when working on primitive B might complete prior to a store for primitive A, even if primitive A is specified prior to primitive B. This applies even to fragment shaders; while fragment shader outputs are always written to the framebuffer in **rasterization order**, stores executed by fragment shader invocations are not.

• The relative order of invocations of different shader types is largely undefined.

  **Note**
  The above limitations on shader invocation order make some forms of synchronization between shader invocations within a single set of primitives unimplementable. For example, having one invocation poll memory written by another invocation assumes that the other invocation has been launched and will complete its writes in finite time.

Stores issued to different memory locations within a single shader invocation **may** not be visible to other invocations, or **may** not become visible in the order they were performed.

The **OpMemoryBarrier** instruction **can** be used to provide stronger ordering of reads and writes performed by a single invocation. **OpMemoryBarrier** guarantees that any memory transactions issued by the shader invocation prior to the instruction complete prior to the memory transactions issued after the instruction. Memory barriers are needed for algorithms that require multiple invocations to access the same memory and require the operations to be performed in a partially-defined relative order. For example, if one shader invocation does a series of writes, followed by an **OpMemoryBarrier** instruction, followed by another write, then the results of the series of writes before the barrier become visible to other shader invocations at a time earlier or equal to when the results of the final write become visible to those invocations. In practice it means that another invocation that sees the results of the final write would also see the previous writes. Without the memory barrier, the final write **may** be visible before the previous writes.

Writes that are the result of shader stores through a variable decorated with **Coherent** automatically have available writes to the same buffer, buffer view, or image view made visible to them, and are themselves automatically made available to access by the same buffer, buffer view, or image view. Reads that are the result of shader loads through a variable decorated with **Coherent** automatically have available writes to the same buffer, buffer view, or image view made visible to them. The order that coherent writes to different locations become available is undefined, unless enforced by a memory barrier instruction or other memory dependency.

  **Note**
  Explicit memory dependencies **must** still be used to guarantee availability and visibility for access via other buffers, buffer views, or image views.

The built-in atomic memory transaction instructions **can** be used to read and write a given memory address atomically. While built-in atomic functions issued by multiple shader invocations are
executed in undefined order relative to each other, these functions perform both a read and a write of a memory address and guarantee that no other memory transaction will write to the underlying memory between the read and write. Atomic operations ensure automatic availability and visibility for writes and reads in the same way as those to Coherent variables.

**Note**

Memory accesses performed on different resource descriptors with the same memory backing may not be well-defined even with the Coherent decoration or via atomics, due to things such as image layouts or ownership of the resource - as described in the Synchronization and Cache Control chapter.

**Note**

Atomics allow shaders to use shared global addresses for mutual exclusion or as counters, among other uses.

### 8.4. Shader Inputs and Outputs

Data is passed into and out of shaders using variables with input or output storage class, respectively. User-defined inputs and outputs are connected between stages by matching their Location decorations. Additionally, data can be provided by or communicated to special functions provided by the execution environment using BuiltIn decorations.

In many cases, the same BuiltIn decoration can be used in multiple shader stages with similar meaning. The specific behavior of variables decorated as BuiltIn is documented in the following sections.

### 8.5. Vertex Shaders

Each vertex shader invocation operates on one vertex and its associated vertex attribute data, and outputs one vertex and associated data. Graphics pipelines must include a vertex shader, and the vertex shader stage is always the first shader stage in the graphics pipeline.

#### 8.5.1. Vertex Shader Execution

A vertex shader must be executed at least once for each vertex specified by a draw command. During execution, the shader is presented with the index of the vertex and instance for which it has been invoked. Input variables declared in the vertex shader are filled by the implementation with the values of vertex attributes associated with the invocation being executed.

If the same vertex is specified multiple times in a draw command (e.g. by including the same index value multiple times in an index buffer) the implementation may reuse the results of vertex shading if it can statically determine that the vertex shader invocations will produce identical results.
Note
It is implementation-dependent when and if results of vertex shading are reused, and thus how many times the vertex shader will be executed. This is true also if the vertex shader contains stores or atomic operations (see `vertexPipelineStoresAndAtomics`).

8.6. Tessellation Control Shaders

The tessellation control shader is used to read an input patch provided by the application and to produce an output patch. Each tessellation control shader invocation operates on an input patch (after all control points in the patch are processed by a vertex shader) and its associated data, and outputs a single control point of the output patch and its associated data, and can also output additional per-patch data. The input patch is sized according to the `patchControlPoints` member of `VkPipelineTessellationStateCreateInfo`, as part of input assembly. The size of the output patch is controlled by the `OpExecutionMode OutputVertices` specified in the tessellation control or tessellation evaluation shaders, which must be specified in at least one of the shaders. The size of the input and output patches must each be greater than zero and less than or equal to `VkPhysicalDeviceLimits::maxTessellationPatchSize`.

8.6.1. Tessellation Control Shader Execution

A tessellation control shader is invoked at least once for each output vertex in a patch.

Inputs to the tessellation control shader are generated by the vertex shader. Each invocation of the tessellation control shader can read the attributes of any incoming vertices and their associated data. The invocations corresponding to a given patch execute logically in parallel, with undefined relative execution order. However, the `OpControlBarrier` instruction can be used to provide limited control of the execution order by synchronizing invocations within a patch, effectively dividing tessellation control shader execution into a set of phases. Tessellation control shaders will read undefined values if one invocation reads a per-vertex or per-patch attribute written by another invocation at any point during the same phase, or if two invocations attempt to write different values to the same per-patch output in a single phase.

8.7. Tessellation Evaluation Shaders

The Tessellation Evaluation Shader operates on an input patch of control points and their associated data, and a single input barycentric coordinate indicating the invocation’s relative position within the subdivided patch, and outputs a single vertex and its associated data.

8.7.1. Tessellation Evaluation Shader Execution

A tessellation evaluation shader is invoked at least once for each unique vertex generated by the tessellator.
8.8. Geometry Shaders

The geometry shader operates on a group of vertices and their associated data assembled from a single input primitive, and emits zero or more output primitives and the group of vertices and their associated data required for each output primitive.

8.8.1. Geometry Shader Execution

A geometry shader is invoked at least once for each primitive produced by the tessellation stages, or at least once for each primitive generated by primitive assembly when tessellation is not in use. A shader can request that the geometry shader runs multiple instances. A geometry shader is invoked at least once for each instance.

8.9. Fragment Shaders

Fragment shaders are invoked as the result of rasterization in a graphics pipeline. Each fragment shader invocation operates on a single fragment and its associated data. With few exceptions, fragment shaders do not have access to any data associated with other fragments and are considered to execute in isolation of fragment shader invocations associated with other fragments.

8.9.1. Fragment Shader Execution

For each fragment generated by rasterization, a fragment shader may be invoked. A fragment shader must not be invoked if the Early Per-Fragment Tests cause it to have no coverage.

Furthermore, if it is determined that a fragment generated as the result of rasterizing a first primitive will have its outputs entirely overwritten by a fragment generated as the result of rasterizing a second primitive in the same subpass, and the fragment shader used for the fragment has no other side effects, then the fragment shader may not be executed for the fragment from the first primitive.

Relative ordering of execution of different fragment shader invocations is not defined.

For each fragment generated by a primitive, the number of times the fragment shader is invoked is implementation-dependent, but must obey the following constraints:

- Each covered sample is included in a single fragment shader invocation.
- When sample shading is not enabled, there is at least one fragment shader invocation.
- When sample shading is enabled, the minimum number of fragment shader invocations is as defined in Sample Shading.

When there is more than one fragment shader invocation per fragment, the association of samples to invocations is implementation-dependent.

In addition to the conditions outlined above for the invocation of a fragment shader, a fragment shader invocation may be produced as a helper invocation. A helper invocation is a fragment shader invocation that is created solely for the purposes of evaluating derivatives for use in non-helper fragment shader invocations. Stores and atomics performed by helper invocations must not
have any effect on memory, and values returned by atomic instructions in helper invocations are undefined.

8.9.2. Early Fragment Tests

An explicit control is provided to allow fragment shaders to enable early fragment tests. If the fragment shader specifies the `EarlyFragmentTests OpExecutionMode`, the per-fragment tests described in Early Fragment Test Mode are performed prior to fragment shader execution. Otherwise, they are performed after fragment shader execution.

8.10. Compute Shaders

Compute shaders are invoked via `vkCmdDispatch` and `vkCmdDispatchIndirect` commands. In general, they have access to similar resources as shader stages executing as part of a graphics pipeline.

Compute workloads are formed from groups of work items called workgroups and processed by the compute shader in the current compute pipeline. A workgroup is a collection of shader invocations that execute the same shader, potentially in parallel. Compute shaders execute in global workgroups which are divided into a number of local workgroups with a size that can be set by assigning a value to the `LocalSize` execution mode or via an object decorated by the `WorkgroupSize` decoration. An invocation within a local workgroup can share data with other members of the local workgroup through shared variables and issue memory and control flow barriers to synchronize with other members of the local workgroup.

8.11. Interpolation Decorations

Interpolation decorations control the behavior of attribute interpolation in the fragment shader stage. Interpolation decorations can be applied to Input storage class variables in the fragment shader stage's interface, and control the interpolation behavior of those variables.

Inputs that could be interpolated can be decorated by at most one of the following decorations:

- **Flat**: no interpolation
- **NoPerspective**: linear interpolation (for lines and polygons)

Fragment input variables decorated with neither Flat nor NoPerspective use perspective-correct interpolation (for lines and polygons).

The presence of and type of interpolation is controlled by the above interpolation decorations as well as the auxiliary decorations Centroid and Sample.

A variable decorated with Flat will not be interpolated. Instead, it will have the same value for every fragment within a triangle. This value will come from a single provoking vertex. A variable decorated with Flat can also be decorated with Centroid or Sample, which will mean the same thing as decorating it only as Flat.

For fragment shader input variables decorated with neither Centroid nor Sample, the assigned
variable may be interpolated anywhere within the fragment and a single value may be assigned to each sample within the fragment.

If a fragment shader input is decorated with **Centroid**, a single value may be assigned to that variable for all samples in the fragment, but that value must be interpolated to a location that lies in both the fragment and in the primitive being rendered, including any of the fragment's samples covered by the primitive. Because the location at which the variable is interpolated may be different in neighboring fragments, and derivatives may be computed by computing differences between neighboring fragments, derivatives of centroid-sampled inputs may be less accurate than those for non-centroid interpolated variables.

If a fragment shader input is decorated with **Sample**, a separate value must be assigned to that variable for each covered sample in the fragment, and that value must be sampled at the location of the individual sample. When `rasterizationSamples` is `VK_SAMPLE_COUNT_1_BIT`, the fragment center must be used for **Centroid**, **Sample**, and undecorated attribute interpolation.

Fragment shader inputs that are signed or unsigned integers, integer vectors, or any double-precision floating-point type must be decorated with **Flat**.

### 8.12. Static Use

A SPIR-V module declares a global object in memory using the `OpVariable` instruction, which results in a pointer `x` to that object. A specific entry point in a SPIR-V module is said to *statically use* that object if that entry point's call tree contains a function that contains a memory instruction or image instruction with `x` as an *id* operand. See the “Memory Instructions” and “Image Instructions” subsections of section 3 “Binary Form” of the SPIR-V specification for the complete list of SPIR-V memory instructions.

Static use is not used to control the behavior of variables with **Input** and **Output** storage. The effects of those variables are applied based only on whether they are present in a shader entry point’s interface.

### 8.13. Invocation and Derivative Groups

An *invocation group* (see the subsection “Control Flow” of section 2 of the SPIR-V specification) for a compute shader is the set of invocations in a single local workgroup. For graphics shaders, an invocation group is an implementation-dependent subset of the set of shader invocations of a given shader stage which are produced by a single drawing command. For indirect drawing commands with `drawCount` greater than one, invocations from separate draws are in distinct invocation groups.

> **Note**
> Because the partitioning of invocations into invocation groups is implementation-dependent and not observable, applications generally need to assume the worst case of all invocations in a draw belonging to a single invocation group.

A *derivative group* (see the subsection “Control Flow” of section 2 of the SPIR-V 1.00 Revision 4 specification) for a fragment shader is the set of invocations generated by a single primitive (point, line, or triangle), including any helper invocations generated by that primitive. Derivatives are
undefined for a sampled image instruction if the instruction is in flow control that is not uniform across the derivative group.
Chapter 9. Pipelines

The following figure shows a block diagram of the Vulkan pipelines. Some Vulkan commands specify geometric objects to be drawn or computational work to be performed, while others specify state controlling how objects are handled by the various pipeline stages, or control data transfer between memory organized as images and buffers. Commands are effectively sent through a processing pipeline, either a graphics pipeline or a compute pipeline.

The first stage of the graphics pipeline (Input Assembler) assembles vertices to form geometric primitives such as points, lines, and triangles, based on a requested primitive topology. In the next stage (Vertex Shader) vertices can be transformed, computing positions and attributes for each vertex. If tessellation and/or geometry shaders are supported, they can then generate multiple primitives from a single input primitive, possibly changing the primitive topology or generating additional attribute data in the process.

The final resulting primitives are clipped to a clip volume in preparation for the next stage, Rasterization. The rasterizer produces a series of framebuffer addresses and values using a two-dimensional description of a point, line segment, or triangle. Each fragment so produced is fed to the next stage (Fragment Shader) that performs operations on individual fragments before they finally alter the framebuffer. These operations include conditional updates into the framebuffer based on incoming and previously stored depth values (to effect depth buffering), blending of incoming fragment colors with stored colors, as well as masking, stenciling, and other logical operations on fragment values.

Framebuffer operations read and write the color and depth/stencil attachments of the framebuffer for a given subpass of a render pass instance. The attachments can be used as input attachments in the fragment shader in a later subpass of the same render pass.

The compute pipeline is a separate pipeline from the graphics pipeline, which operates on one-, two-, or three-dimensional workgroups which can read from and write to buffer and image memory.

This ordering is meant only as a tool for describing Vulkan, not as a strict rule of how Vulkan is implemented, and we present it only as a means to organize the various operations of the pipelines. Actual ordering guarantees between pipeline stages are explained in detail in the synchronization chapter.
Each pipeline is controlled by a monolithic object created from a description of all of the shader stages and any relevant fixed-function stages. Linking the whole pipeline together allows the optimization of shaders based on their input/outputs and eliminates expensive draw time state validation.

A pipeline object is bound to the current state using `vkCmdBindPipeline`. Any pipeline object state that is specified as `dynamic` is not applied to the current state when the pipeline object is bound, but is instead set by dynamic state setting commands.

No state, including dynamic state, is inherited from one command buffer to another.

Compute and graphics pipelines are each represented by `VkPipeline` handles:

```c
VK_DEFINE_NON_DISPATCHABLE_HANDLE(VkPipeline)
```

### 9.1. Compute Pipelines

Compute pipelines consist of a single static compute shader stage and the pipeline layout.

The compute pipeline represents a compute shader and is created by calling `vkCreateComputePipelines` with `module` and `pName` selecting an entry point from a shader module, where that entry point defines a valid compute shader, in the `VkPipelineShaderStageCreateInfo` structure contained within the `VkComputePipelineCreateInfo` structure.

To create compute pipelines, call:
VkResult vkCreateComputePipelines(
    VkDevice                                    device,
    VkPipelineCache                             pipelineCache,
    uint32_t                                    createInfoCount,
    const VkComputePipelineCreateInfo*          pCreateInfos,
    const VkAllocationCallbacks*                pAllocator,
    VkPipeline*                                 pPipelines);

• **device** is the logical device that creates the compute pipelines.

• **pipelineCache** is either VK_NULL_HANDLE, indicating that pipeline caching is disabled; or the handle of a valid pipeline cache object, in which case use of that cache is enabled for the duration of the command.

• **createInfoCount** is the length of the pCreateInfos and pPipelines arrays.

• **pCreateInfos** is an array of VkComputePipelineCreateInfo structures.

• **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.

• **pPipelines** is a pointer to an array in which the resulting compute pipeline objects are returned.

### Valid Usage

- If the **flags** member of any element of pCreateInfos contains the VK_PIPELINE_CREATE_DERIVATIVE_BIT flag, and the basePipelineIndex member of that same element is not -1, basePipelineIndex **must** be less than the index into pCreateInfos that corresponds to that element.

- If the **flags** member of any element of pCreateInfos contains the VK_PIPELINE_CREATE_DERIVATIVE_BIT flag, the base pipeline **must** have been created with the VK_PIPELINE_CREATE_ALLOW_DERIVATIVES_BIT flag set.

### Valid Usage (Implicit)

- **device** **must** be a valid VkDevice handle

- If **pipelineCache** is not VK_NULL_HANDLE, **pipelineCache** **must** be a valid VkPipelineCache handle

- **pCreateInfos** **must** be a valid pointer to an array of createInfoCount valid VkComputePipelineCreateInfo structures

- If **pAllocator** is not NULL, **pAllocator** **must** be a valid pointer to a valid VkAllocationCallbacks structure

- **pPipelines** **must** be a valid pointer to an array of createInfoCount VkPipeline handles

- **createInfoCount** **must** be greater than 0

- If **pipelineCache** is a valid handle, it **must** have been created, allocated, or retrieved from **device**
Return Codes

Success
- VK_SUCCESS

Failure
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY

The `VkComputePipelineCreateInfo` structure is defined as:

```c
typedef struct VkComputePipelineCreateInfo {
    VkStructureType sType;
    const void* pNext;
    VkPipelineCreateFlags flags;
    VkPipelineShaderStageCreateInfo stage;
    VkPipelineLayout layout;
    VkPipeline basePipelineHandle;
    int32_t basePipelineIndex;
} VkComputePipelineCreateInfo;
```

- `sType` is the type of this structure.
- `pNext` is NULL or a pointer to an extension-specific structure.
- `flags` is a bitmask of `VkPipelineCreateFlagBits` specifying how the pipeline will be generated.
- `stage` is a `VkPipelineShaderStageCreateInfo` describing the compute shader.
- `layout` is the description of binding locations used by both the pipeline and descriptor sets used with the pipeline.
- `basePipelineHandle` is a pipeline to derive from
- `basePipelineIndex` is an index into the `pCreateInfos` parameter to use as a pipeline to derive from

The parameters `basePipelineHandle` and `basePipelineIndex` are described in more detail in *Pipeline Derivatives*.

`stage` points to a structure of type `VkPipelineShaderStageCreateInfo`. 
Valid Usage

- If `flags` contains the `VK_PIPELINE_CREATE_DERIVATIVE_BIT` flag, and `basePipelineIndex` is -1, `basePipelineHandle` must be a valid handle to a compute `VkPipeline`
- If `flags` contains the `VK_PIPELINE_CREATE_DERIVATIVE_BIT` flag, and `basePipelineHandle` is `VK_NULL_HANDLE`, `basePipelineIndex` must be a valid index into the calling command's `pCreateInfos` parameter
- If `flags` contains the `VK_PIPELINE_CREATE_DERIVATIVE_BIT` flag, and `basePipelineIndex` is not -1, `basePipelineHandle` must be `VK_NULL_HANDLE`
- If `flags` contains the `VK_PIPELINE_CREATE_DERIVATIVE_BIT` flag, and `basePipelineHandle` is not `VK_NULL_HANDLE`, `basePipelineIndex` must be -1
- The `stage` member of `stage` must be `VK_SHADER_STAGE_COMPUTE_BIT`
- The shader code for the entry point identified by `stage` and the rest of the state identified by this structure must adhere to the pipeline linking rules described in the Shader Interfaces chapter
- `layout` must be consistent with the layout of the compute shader specified in `stage`
- The number of resources in `layout` accessible to the compute shader stage must be less than or equal to `VkPhysicalDeviceLimits::maxPerStageResources`

Valid Usage (Implicit)

- `sType` must be `VK_STRUCTURE_TYPE_COMPUTE_PIPELINE_CREATE_INFO`
- `pNext` must be NULL
- `flags` must be a valid combination of `VkPipelineCreateFlagBits` values
- `stage` must be a valid `VkPipelineShaderStageCreateInfo` structure
- `layout` must be a valid `VkPipelineLayout` handle
- Both of `basePipelineHandle`, and `layout` that are valid handles must have been created, allocated, or retrieved from the same `VkDevice`

The `VkPipelineShaderStageCreateInfo` structure is defined as:

```c
typedef struct VkPipelineShaderStageCreateInfo {
    VkStructureType                     sType;
    const void*                         pNext;
    VkPipelineShaderStageCreateFlags    flags;
    VkShaderStageFlagBits               stage;
    VkShaderModule                      module;
    const char*                         pName;
    const VkSpecializationInfo*         pSpecializationInfo;
} VkPipelineShaderStageCreateInfo;
```
- **sType** is the type of this structure.
- **pNext** is **NULL** or a pointer to an extension-specific structure.
- **flags** is reserved for future use.
- **stage** is a `VkShaderStageFlagBits` value specifying a single pipeline stage.
- **module** is a `VkShaderModule` object that contains the shader for this stage.
- **pName** is a pointer to a null-terminated UTF-8 string specifying the entry point name of the shader for this stage.
- **pSpecializationInfo** is a pointer to `VkSpecializationInfo`, as described in Specialization Constants, and can be **NULL**.
Valid Usage

- If the geometry shaders feature is not enabled, `stage` must not be `VK_SHADER_STAGE_GEOMETRY_BIT`.
- If the tessellation shaders feature is not enabled, `stage` must not be `VK_SHADER_STAGE_TESSELLATION_CONTROL_BIT` or `VK_SHADER_STAGE_TESSELLATION_EVALUATION_BIT`.
- `stage` must not be `VK_SHADER_STAGE_ALL_GRAPHICS`, or `VK_SHADER_STAGE_ALL`.
- `pName` must be the name of an `OpEntryPoint` in `module` with an execution model that matches `stage`.
- If the identified entry point includes any variable in its interface that is declared with the `ClipDistance BuiltIn` decoration, that variable must not have an array size greater than `VkPhysicalDeviceLimits::maxClipDistances`.
- If the identified entry point includes any variable in its interface that is declared with the `CullDistance BuiltIn` decoration, that variable must not have an array size greater than `VkPhysicalDeviceLimits::maxCullDistances`.
- If the identified entry point includes any variables in its interface that are declared with the `ClipDistance` or `CullDistance BuiltIn` decoration, those variables must not have array sizes which sum to more than `VkPhysicalDeviceLimits::maxCombinedClipAndCullDistances`.
- If the identified entry point includes any variable in its interface that is declared with the `SampleMask BuiltIn` decoration, that variable must not have an array size greater than `VkPhysicalDeviceLimits::maxSampleMaskWords`.
- If `stage` is `VK_SHADER_STAGE_VERTEX_BIT`, the identified entry point must not include any input variable in its interface that is decorated with `CullDistance`.
- If `stage` is `VK_SHADER_STAGE_TESSELLATION_CONTROL_BIT` or `VK_SHADER_STAGE_TESSELLATION_EVALUATION_BIT`, and the identified entry point has an `OpExecutionMode` instruction that specifies a patch size with `OutputVertices`, the patch size must be greater than 0 and less than or equal to `VkPhysicalDeviceLimits::maxTessellationPatchSize`.
- If `stage` is `VK_SHADER_STAGE_GEOMETRY_BIT`, the identified entry point must have an `OpExecutionMode` instruction that specifies a maximum output vertex count that is greater than 0 and less than or equal to `VkPhysicalDeviceLimits::maxGeometryOutputVertices`.
- If `stage` is `VK_SHADER_STAGE_GEOMETRY_BIT`, the identified entry point must have an `OpExecutionMode` instruction that specifies an invocation count that is greater than 0 and less than or equal to `VkPhysicalDeviceLimits::maxGeometryShaderInvocations`.
- If `stage` is `VK_SHADER_STAGE_GEOMETRY_BIT`, and the identified entry point writes to `Layer` for any primitive, it must write the same value to `Layer` for all vertices of a given primitive.
- If `stage` is `VK_SHADER_STAGE_GEOMETRY_BIT`, and the identified entry point writes to `ViewportIndex` for any primitive, it must write the same value to `ViewportIndex` for all vertices of a given primitive.
- If `stage` is `VK_SHADER_STAGE_FRAGMENT_BIT`, the identified entry point must not include any output variables in its interface decorated with `CullDistance`. 
• If `stage` is `VK_SHADER_STAGE_FRAGMENT_BIT`, and the identified entry point writes to `FragDepth` in any execution path, it **must** write to `FragDepth` in all execution paths.

### Valid Usage (Implicit)

- `sType` **must** be `VK_STRUCTURE_TYPE_PIPELINE_SHADER_STAGE_CREATE_INFO`
- `pNext` **must** be `NULL`
- `flags` **must** be `0`
- `stage` **must** be a valid `VkShaderStageFlagBits` value
- `module` **must** be a valid `VkShaderModule` handle
- `pName` **must** be a null-terminated UTF-8 string

If `pSpecializationInfo` is not `NULL`, `pSpecializationInfo` **must** be a valid pointer to a valid `VkSpecializationInfo` structure

```c
typedef VkFlags VkPipelineShaderStageCreateFlags;
```

`VkPipelineShaderStageCreateFlags` is a bitmask type for setting a mask, but is currently reserved for future use.

Commands and structures which need to specify one or more shader stages do so using a bitmask whose bits correspond to stages. Bits which **can** be set to specify shader stages are:

```c
typedef enum VkShaderStageFlagBits {
    VK_SHADER_STAGE_VERTEX_BIT = 0x00000001,
    VK_SHADER_STAGE_TESSELLATION_CONTROL_BIT = 0x00000002,
    VK_SHADER_STAGE_TESSELLATION_EVALUATION_BIT = 0x00000004,
    VK_SHADER_STAGE_GEOMETRY_BIT = 0x00000008,
    VK_SHADER_STAGE_FRAGMENT_BIT = 0x00000010,
    VK_SHADER_STAGE_COMPUTE_BIT = 0x00000020,
    VK_SHADER_STAGE_ALL_GRAPHICS = 0x0000001F,
    VK_SHADER_STAGE_ALL = 0x7FFFFFFF,
} VkShaderStageFlagBits;
```

- `VK_SHADER_STAGE_VERTEX_BIT` specifies the vertex stage.
- `VK_SHADER_STAGE_TESSELLATION_CONTROL_BIT` specifies the tessellation control stage.
- `VK_SHADER_STAGE_TESSELLATION_EVALUATION_BIT` specifies the tessellation evaluation stage.
- `VK_SHADER_STAGE_GEOMETRY_BIT` specifies the geometry stage.
- `VK_SHADER_STAGE_FRAGMENT_BIT` specifies the fragment stage.
- `VK_SHADER_STAGE_COMPUTE_BIT` specifies the compute stage.
- `VK_SHADER_STAGE_ALL_GRAPHICS` is a combination of bits used as shorthand to specify all graphics.
stages defined above (excluding the compute stage).

- **VK_SHADER_STAGE_ALL** is a combination of bits used as shorthand to specify all shader stages supported by the device, including all additional stages which are introduced by extensions.

  \[\text{Note}\]
  
  **VK_SHADER_STAGE_ALL_GRAPHICS** only includes the original five graphics stages included in Vulkan 1.0, and not any stages added by extensions. Thus, it may not have the desired effect in all cases.

\[
\text{typedef} \ Vk\text{Flags} \ Vk\text{ShaderStageFlags};
\]

**VkShaderStageFlags** is a bitmask type for setting a mask of zero or more **VkShaderStageFlagBits**.

### 9.2. Graphics Pipelines

Graphics pipelines consist of multiple shader stages, multiple fixed-function pipeline stages, and a pipeline layout.

To create graphics pipelines, call:

\[
\text{VkResult} \ vk\text{CreateGraphicsPipelines}(\text{VkDevice} \ device, \text{VkPipelineCache} \ pipelineCache, \text{uint32}\_t \ createInfoCount, \text{const} \ Vk\text{GraphicsPipelineCreateInfo}\_* \ p\text{CreateInfos}, \text{const} \ Vk\text{AllocationCallbacks}\_* \ p\text{Allocator}, \text{VkPipeline}\_* \ p\text{Pipelines});
\]

- **device** is the logical device that creates the graphics pipelines.
- **pipelineCache** is either **VK_NULL_HANDLE**, indicating that pipeline caching is disabled; or the handle of a valid pipeline cache object, in which case use of that cache is enabled for the duration of the command.
- **createInfoCount** is the length of the **pCreateInfos** and **pPipelines** arrays.
- **pCreateInfos** is an array of **VkGraphicsPipelineCreateInfo** structures.
- **pAllocator** controls host memory allocation as described in the **Memory Allocation** chapter.
- **pPipelines** is a pointer to an array in which the resulting graphics pipeline objects are returned.

The **VkGraphicsPipelineCreateInfo** structure includes an array of shader create info structures containing all the desired active shader stages, as well as creation info to define all relevant fixed-function stages, and a pipeline layout.
**Valid Usage**

- If the `flags` member of any element of `pCreateInfos` contains the `VK_PIPELINE_CREATE_DERIVATIVE_BIT` flag, and the `basePipelineIndex` member of that same element is not -1, `basePipelineIndex` must be less than the index into `pCreateInfos` that corresponds to that element.

- If the `flags` member of any element of `pCreateInfos` contains the `VK_PIPELINE_CREATE_DERIVATIVE_BIT` flag, the base pipeline must have been created with the `VK_PIPELINE_CREATE_ALLOW_DERIVATIVES_BIT` flag set.

**Valid Usage (Implicit)**

- `device` must be a valid `VkDevice` handle.
- If `pipelineCache` is not `VK_NULL_HANDLE`, `pipelineCache` must be a valid `VkPipelineCache` handle.
- `pCreateInfos` must be a valid pointer to an array of `createInfoCount` valid `VkGraphicsPipelineCreateInfo` structures.
- If `pAllocator` is not NULL, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure.
- `pPipelines` must be a valid pointer to an array of `createInfoCount` `VkPipeline` handles.
- `createInfoCount` must be greater than 0.
- If `pipelineCache` is a valid handle, it must have been created, allocated, or retrieved from `device`.

**Return Codes**

**Success**

- `VK_SUCCESS`

**Failure**

- `VK_ERROR_OUT_OF_HOST_MEMORY`
- `VK_ERROR_OUT_OF_DEVICE_MEMORY`

The `VkGraphicsPipelineCreateInfo` structure is defined as:
typedef struct VkGraphicsPipelineCreateInfo {
    VkStructureType sType;
    const void* pNext;
    VkPipelineCreateFlags flags;
    uint32_t stageCount;
    const VkPipelineShaderStageCreateInfo* pStages;
    const VkPipelineVertexInputStateCreateInfo* pVertexInputState;
    const VkPipelineInputAssemblyStateCreateInfo* pInputAssemblyState;
    const VkPipelineTessellationStateCreateInfo* pTessellationState;
    const VkPipelineViewportStateCreateInfo* pViewportState;
    const VkPipelineRasterizationStateCreateInfo* pRasterizationState;
    const VkPipelineMultisampleStateCreateInfo* pMultisampleState;
    const VkPipelineDepthStencilStateCreateInfo* pDepthStencilState;
    const VkPipelineColorBlendStateCreateInfo* pColorBlendState;
    const VkPipelineDynamicStateCreateInfo* pDynamicState;
    VkPipelineLayout layout;
    VkRenderPass renderPass;
    uint32_t subpass;
    VkPipeline basePipelineHandle;
    int32_t basePipelineIndex;
} VkGraphicsPipelineCreateInfo;

- **sType** is the type of this structure.
- **pNext** is `NULL` or a pointer to an extension-specific structure.
- **flags** is a bitmask of `VkPipelineCreateFlagBits` specifying how the pipeline will be generated.
- **stageCount** is the number of entries in the **pStages** array.
- **pStages** is an array of size **stageCount** structures of type `VkPipelineShaderStageCreateInfo` describing the set of the shader stages to be included in the graphics pipeline.
- **pVertexInputState** is a pointer to an instance of the `VkPipelineVertexInputStateCreateInfo` structure.
- **pInputAssemblyState** is a pointer to an instance of the `VkPipelineInputAssemblyStateCreateInfo` structure which determines input assembly behavior, as described in Drawing Commands.
- **pTessellationState** is a pointer to an instance of the `VkPipelineTessellationStateCreateInfo` structure, and is ignored if the pipeline does not include a tessellation control shader stage and tessellation evaluation shader stage.
- **pViewportState** is a pointer to an instance of the `VkPipelineViewportStateCreateInfo` structure, and is ignored if the pipeline has rasterization disabled.
- **pRasterizationState** is a pointer to an instance of the `VkPipelineRasterizationStateCreateInfo` structure.
- **pMultisampleState** is a pointer to an instance of the `VkPipelineMultisampleStateCreateInfo`, and is ignored if the pipeline has rasterization disabled.
- **pDepthStencilState** is a pointer to an instance of the `VkPipelineDepthStencilStateCreateInfo` structure, and is ignored if the pipeline has rasterization disabled or if the subpass of the render pass the pipeline is created against does not use a depth/stencil attachment.
- **pColorBlendState** is a pointer to an instance of the `VkPipelineColorBlendStateCreateInfo` structure, and is ignored if the pipeline has rasterization disabled or if the subpass of the render pass the pipeline is created against does not use any color attachments.

- **pDynamicState** is a pointer to `VkPipelineDynamicStateCreateInfo` and is used to indicate which properties of the pipeline state object are dynamic and can be changed independently of the pipeline state. This can be `NULL`, which means no state in the pipeline is considered dynamic.

- **layout** is the description of binding locations used by both the pipeline and descriptor sets used with the pipeline.

- **renderPass** is a handle to a render pass object describing the environment in which the pipeline will be used; the pipeline must only be used with an instance of any render pass compatible with the one provided. See Render Pass Compatibility for more information.

- **subpass** is the index of the subpass in the render pass where this pipeline will be used.

- **basePipelineHandle** is a pipeline to derive from.

- **basePipelineIndex** is an index into the `pCreateInfos` parameter to use as a pipeline to derive from.

The parameters **basePipelineHandle** and **basePipelineIndex** are described in more detail in Pipeline Derivatives.

- **pStages** points to an array of `VkPipelineShaderStageCreateInfo` structures, which were previously described in Compute Pipelines.

- **pDynamicState** points to a structure of type `VkPipelineDynamicStateCreateInfo`.


Valid Usage

• If flags contains the VK_PIPELINE_CREATE_DERIVATIVE_BIT flag, and basePipelineIndex is -1, basePipelineHandle must be a valid handle to a graphics VkPipeline.

• If flags contains the VK_PIPELINE_CREATE_DERIVATIVE_BIT flag, and basePipelineHandle is VK_NULL_HANDLE, basePipelineIndex must be a valid index into the calling command’s pCreateInfos parameter.

• If flags contains the VK_PIPELINE_CREATE_DERIVATIVE_BIT flag, and basePipelineHandle is not VK_NULL_HANDLE, basePipelineIndex must be -1.

• The stage member of each element of pStages must be unique.

• The stage member of one element of pStages must be VK_SHADER_STAGE_VERTEX_BIT.

• The stage member of each element of pStages must not be VK_SHADER_STAGE_COMPUTE_BIT.

• If pStages includes a tessellation control shader stage, it must include a tessellation evaluation shader stage.

• If pStages includes a tessellation evaluation shader stage, it must include a tessellation control shader stage.

• If pStages includes a tessellation control shader stage and a tessellation evaluation shader stage, pTessellationState must be a valid pointer to a valid VkPipelineTessellationStateCreateInfo structure.

• If pStages includes tessellation shader stages, the shader code of at least one stage must contain an OpExecutionMode instruction that specifies the type of subdivision in the pipeline.

• If pStages includes tessellation shader stages, and the shader code of both stages contain an OpExecutionMode instruction that specifies the type of subdivision in the pipeline, they must both specify the same subdivision mode.

• If pStages includes tessellation shader stages, the shader code of at least one stage must contain an OpExecutionMode instruction that specifies the output patch size in the pipeline.

• If pStages includes tessellation shader stages, and the shader code of both contain an OpExecutionMode instruction that specifies the output patch size in the pipeline, they must both specify the same patch size.

• If pStages includes tessellation shader stages, the topology member of pInputAssembly must be VK_PRIMITIVE_TOPOLOGY_PATCH_LIST.

• If the topology member of pInputAssembly is VK_PRIMITIVE_TOPOLOGY_PATCH_LIST, pStages must include tessellation shader stages.

• If pStages includes a geometry shader stage, and does not include any tessellation shader stages, its shader code must contain an OpExecutionMode instruction that specifies an input primitive type that is compatible with the primitive topology specified in pInputAssembly.

• If pStages includes a geometry shader stage, and also includes tessellation shader stages, its shader code must contain an OpExecutionMode instruction that specifies an input...
primitive type that is **compatible** with the primitive topology that is output by the tessellation stages

- If `pStages` includes a fragment shader stage and a geometry shader stage, and the fragment shader code reads from an input variable that is decorated with `PrimitiveID`, then the geometry shader code must write to a matching output variable, decorated with `PrimitiveID`, in all execution paths.

- If `pStages` includes a fragment shader stage, its shader code must not read from any input attachment that is defined as `VK_ATTACHMENT_UNUSED` in `subpass`.

- The shader code for the entry points identified by `pStages`, and the rest of the state identified by this structure must adhere to the pipeline linking rules described in the **Shader Interfaces** chapter.

- If rasterization is not disabled and `subpass` uses a depth/stencil attachment in `renderPass` that has a layout of `VK_IMAGE_LAYOUT_DEPTH_STENCIL_READ_ONLY_OPTIMAL` in the `VkAttachmentReference` defined by `subpass`, the `depthWriteEnable` member of `pDepthStencilState` must be `VK_FALSE`.

- If rasterization is not disabled and `subpass` uses a depth/stencil attachment in `renderPass` that has a layout of `VK_IMAGE_LAYOUT_DEPTH_STENCIL_READ_ONLY_OPTIMAL` in the `VkAttachmentReference` defined by `subpass`, the `failOp`, `passOp` and `depthFailOp` members of each of the `front` and `back` members of `pDepthStencilState` must be `VK_STENCIL_OP_KEEP`.

- If rasterization is not disabled and the `subpass` uses color attachments, then for each color attachment in the `subpass` the `blendEnable` member of the corresponding element of the `pAttachment` member of `pColorBlendState` must be `VK_FALSE` if the attached image's format features does not contain the `VK_FORMAT_FEATURE_COLOR_ATTACHMENT_BLEND_BIT`.

- If rasterization is not disabled and the `subpass` uses color attachments, the `attachmentCount` member of `pColorBlendState` must be equal to the `colorAttachmentCount` used to create `subpass`.

- If no element of the `pDynamicStates` member of `pDynamicState` is `VK_DYNAMIC_STATE_VIEWPORT`, the `pViewports` member of `pViewportState` must be a valid pointer to an array of `pViewportState::viewportCount` valid `VkViewport` structures.

- If no element of the `pDynamicStates` member of `pDynamicState` is `VK_DYNAMIC_STATE_SCISSOR`, the `pScissors` member of `pViewportState` must be a valid pointer to an array of `pViewportState::scissorCount` `VkRect2D` structures.

- If the wide lines feature is not enabled, and no element of the `pDynamicStates` member of `pDynamicState` is `VK_DYNAMIC_STATE_LINE_WIDTH`, the `lineWidth` member of `pRasterizationState` must be `1.0`.

- If the `rasterizerDiscardEnable` member of `pRasterizationState` is `VK_FALSE`, `pViewportState` must be a valid pointer to a valid `VkPipelineViewportStateCreateInfo` structure.

- If the `rasterizerDiscardEnable` member of `pRasterizationState` is `VK_FALSE`, `pMultisampleState` must be a valid pointer to a valid `VkPipelineMultisampleStateCreateInfo` structure.

- If the `rasterizerDiscardEnable` member of `pRasterizationState` is `VK_FALSE`, and `subpass` uses a depth/stencil attachment, `pDepthStencilState` must be a valid pointer to a valid `VkPipelineDepthStencilStateCreateInfo` structure.
• If the `rasterizerDiscardEnable` member of `pRasterizationState` is `VK_FALSE`, and `subpass` uses color attachments, `pColorBlendState` must be a valid pointer to a valid `VkPipelineColorBlendStateCreateInfo` structure.

• If the depth bias clamping feature is not enabled, no element of the `pDynamicStates` member of `pDynamicState` is `VK_DYNAMIC_STATE_DEPTH_BIAS`, and the `depthBiasEnable` member of `pRasterizationState` is `VK_TRUE`, the `depthBiasClamp` member of `pRasterizationState` must be 0.0.

• If no element of the `pDynamicStates` member of `pDynamicState` is `VK_DYNAMIC_STATE_DEPTH_BOUNDS`, and the `depthBoundsTestEnable` member of `pDepthStencilState` is `VK_TRUE`, the `minDepthBounds` and `maxDepthBounds` members of `pDepthStencilState` must be between 0.0 and 1.0, inclusive.

• `layout` must be consistent with all shaders specified in `pStages`.

• If `subpass` uses color and/or depth/stencil attachments, then the `rasterizationSamples` member of `pMultisampleState` must be the same as the sample count for those subpass attachments.

• If `subpass` does not use any color and/or depth/stencil attachments, then the `rasterizationSamples` member of `pMultisampleState` must follow the rules for a zero-attachment subpass.

• `subpass` must be a valid subpass within `renderPass`.

• The number of resources in `layout` accessible to each shader stage that is used by the pipeline must be less than or equal to `VkPhysicalDeviceLimits::maxPerStageResources`.

• If `pStages` includes a vertex shader stage, `pVertexInputState` must be a valid pointer to a valid `VkPipelineVertexInputStateCreateInfo` structure.

• If `pStages` includes a vertex shader stage, `pInputAssemblyState` must be a valid pointer to a valid `VkPipelineInputAssemblyStateCreateInfo` structure.
Valid Usage (Implicit)

- **sType** must be `VK_STRUCTURE_TYPE_GRAPHICS_PIPELINE_CREATE_INFO`
- **pNext** must be `NULL`
- **flags** must be a valid combination of `VkPipelineCreateFlagBits` values
- **pStages** must be a valid pointer to an array of `stageCount` valid `VkPipelineShaderStageCreateInfo` structures
- **pRasterizationState** must be a valid pointer to a valid `VkPipelineRasterizationStateCreateInfo` structure
- If `pDynamicState` is not `NULL`, `pDynamicState` must be a valid pointer to a valid `VkPipelineDynamicStateCreateInfo` structure
- **layout** must be a valid `VkPipelineLayout` handle
- **renderPass** must be a valid `VkRenderPass` handle
- **stageCount** must be greater than 0
- Each of `basePipelineHandle`, `layout`, and `renderPass` that are valid handles must have been created, allocated, or retrieved from the same `VkDevice`

Possible values of the **flags** member of `VkGraphicsPipelineCreateInfo` and `VkComputePipelineCreateInfo`, specifying how a pipeline is created, are:

```c
typedef enum VkPipelineCreateFlagBits {
    VK_PIPELINE_CREATE_DISABLE_OPTIMIZATION_BIT = 0x00000001,
    VK_PIPELINE_CREATE_ALLOW_DERIVATIVES_BIT = 0x00000002,
    VK_PIPELINE_CREATE_DERIVATIVE_BIT = 0x00000004,
} VkPipelineCreateFlagBits;
```

- **VK_PIPELINE_CREATE_DISABLE_OPTIMIZATION_BIT** specifies that the created pipeline will not be optimized. Using this flag may reduce the time taken to create the pipeline.
- **VK_PIPELINE_CREATE_ALLOW_DERIVATIVES_BIT** specifies that the pipeline to be created is allowed to be the parent of a pipeline that will be created in a subsequent call to `vkCreateGraphicsPipelines` or `vkCreateComputePipelines`.
- **VK_PIPELINE_CREATE_DERIVATIVE_BIT** specifies that the pipeline to be created will be a child of a previously created parent pipeline.

It is valid to set both **VK_PIPELINE_CREATE_ALLOW_DERIVATIVES_BIT** and **VK_PIPELINE_CREATE_DERIVATIVE_BIT**. This allows a pipeline to be both a parent and possibly a child in a pipeline hierarchy. See **Pipeline Derivatives** for more information.

```c
typedef VkFlags VkPipelineCreateFlags;
```

`VkPipelineCreateFlags` is a bitmask type for setting a mask of zero or more
The `VkPipelineDynamicStateCreateInfo` structure is defined as:

```c
typedef struct VkPipelineDynamicStateCreateInfo {
    VkStructureType                      sType;  
    const void*                          pNext;  
    VkPipelineDynamicStateCreateFlags    flags;  
    uint32_t                             dynamicStateCount;  
    const VkDynamicState*                pDynamicStates;  
} VkPipelineDynamicStateCreateInfo;
```

- **sType** is the type of this structure.
- **pNext** is `NULL` or a pointer to an extension-specific structure.
- **flags** is reserved for future use.
- **dynamicStateCount** is the number of elements in the **pDynamicStates** array.
- **pDynamicStates** is an array of `VkDynamicState` values specifying which pieces of pipeline state will use the values from dynamic state commands rather than from pipeline state creation info.

**Valid Usage**

- Each element of **pDynamicStates** must be unique.

**Valid Usage (Implicit)**

- **sType** must be `VK_STRUCTURE_TYPE_PIPELINE_DYNAMIC_STATE_CREATE_INFO`
- **pNext** must be `NULL`
- **flags** must be `0`
- **pDynamicStates** must be a valid pointer to an array of `dynamicStateCount` valid `VkDynamicState` values
- **dynamicStateCount** must be greater than `0`

```c
typedef VkFlags VkPipelineDynamicStateCreateFlags;
```

`VkPipelineDynamicStateCreateFlags` is a bitmask type for setting a mask, but is currently reserved for future use.

The source of different pieces of dynamic state is specified by the `VkPipelineDynamicStateCreateInfo::pDynamicStates` property of the currently active pipeline, each of whose elements must be one of the values:
typedef enum VkDynamicState {
    VK_DYNAMIC_STATE_VIEWPORT = 0,
    VK_DYNAMIC_STATE_SCISSOR = 1,
    VK_DYNAMIC_STATE_LINE_WIDTH = 2,
    VK_DYNAMIC_STATE_DEPTH_BIAS = 3,
    VK_DYNAMIC_STATE_BLEND_CONSTANTS = 4,
    VK_DYNAMIC_STATE_DEPTH_BOUNDS = 5,
    VK_DYNAMIC_STATE_STENCIL_COMPARE_MASK = 6,
    VK_DYNAMIC_STATE_STENCIL_WRITE_MASK = 7,
    VK_DYNAMIC_STATE_STENCIL_REFERENCE = 8,
} VkDynamicState;

- **VK_DYNAMIC_STATE_VIEWPORT** specifies that the *pViewports* state in `VkPipelineViewportStateCreateInfo` will be ignored and **must** be set dynamically with `vkCmdSetViewport` before any draw commands. The number of viewports used by a pipeline is still specified by the `viewportCount` member of `VkPipelineViewportStateCreateInfo`.

- **VK_DYNAMIC_STATE_SCISSOR** specifies that the *pScissors* state in `VkPipelineViewportStateCreateInfo` will be ignored and **must** be set dynamically with `vkCmdSetScissor` before any draw commands. The number of scissor rectangles used by a pipeline is still specified by the `scissorCount` member of `VkPipelineViewportStateCreateInfo`.

- **VK_DYNAMIC_STATE_LINE_WIDTH** specifies that the *lineWidth* state in `VkPipelineRasterizationStateCreateInfo` will be ignored and **must** be set dynamically with `vkCmdSetLineWidth` before any draw commands that generate line primitives for the rasterizer.

- **VK_DYNAMIC_STATE_DEPTH_BIAS** specifies that the *depthBiasConstantFactor*, *depthBiasClamp* and *depthBiasSlopeFactor* states in `VkPipelineRasterizationStateCreateInfo` will be ignored and **must** be set dynamically with `vkCmdSetDepthBias` before any draws are performed with *depthBiasEnable* in `VkPipelineRasterizationStateCreateInfo` set to **VK_TRUE**.

- **VK_DYNAMIC_STATE_BLEND_CONSTANTS** specifies that the *blendConstants* state in `VkPipelineColorBlendStateCreateInfo` will be ignored and **must** be set dynamically with `vkCmdSetBlendConstants` before any draws are performed with a pipeline state with *blendEnable* in `VkPipelineColorBlendAttachmentState` member *blendEnable* set to **VK_TRUE** and any of the blend functions using a constant blend color.

- **VK_DYNAMIC_STATE_DEPTH_BOUNDS** specifies that the *minDepthBounds* and *maxDepthBounds* states of `VkPipelineDepthStencilStateCreateInfo` will be ignored and **must** be set dynamically with `vkCmdSetDepthBounds` before any draws are performed with a pipeline state with `VkPipelineDepthStencilStateCreateInfo` member *depthBoundsTestEnable* set to **VK_TRUE**.

- **VK_DYNAMIC_STATE_STENCIL_COMPARE_MASK** specifies that the *compareMask* state in `VkPipelineDepthStencilStateCreateInfo` for both *front* and *back* will be ignored and **must** be set dynamically with `vkCmdSetStencilCompareMask` before any draws are performed with a pipeline state with `VkPipelineDepthStencilStateCreateInfo` member *stencilTestEnable* set to **VK_TRUE**.

- **VK_DYNAMIC_STATE_STENCIL_WRITE_MASK** specifies that the *writeMask* state in `VkPipelineDepthStencilStateCreateInfo` for both *front* and *back* will be ignored and **must** be set dynamically with `vkCmdSetStencilWriteMask` before any draws are performed with a pipeline...
state with `VkPipelineDepthStencilStateCreateInfo` member `stencilTestEnable` set to `VK_TRUE`

- `VK_DYNAMIC_STATE_STENCIL_REFERENCE` specifies that the reference state in `VkPipelineDepthStencilStateCreateInfo` for both `front` and `back` will be ignored and must be set dynamically with `vkCmdSetStencilReference` before any draws are performed with a pipeline state with `VkPipelineDepthStencilStateCreateInfo` member `stencilTestEnable` set to `VK_TRUE`.

### 9.2.1. Valid Combinations of Stages for Graphics Pipelines

If tessellation shader stages are omitted, the tessellation shading and fixed-function stages of the pipeline are skipped.

If a geometry shader is omitted, the geometry shading stage is skipped.

If a fragment shader is omitted, the results of fragment processing are undefined. Specifically, any fragment color outputs are considered to have undefined values, and the fragment depth is considered to be unmodified. This can be useful for depth-only rendering.

Presence of a shader stage in a pipeline is indicated by including a valid `VkPipelineShaderStageCreateInfo` with `module` and `pName` selecting an entry point from a shader module, where that entry point is valid for the stage specified by `stage`.

Presence of some of the fixed-function stages in the pipeline is implicitly derived from enabled shaders and provided state. For example, the fixed-function tessellator is always present when the pipeline has valid Tessellation Control and Tessellation Evaluation shaders.

**For example:**

- Depth/stencil-only rendering in a subpass with no color attachments
  - Active Pipeline Shader Stages
    - Vertex Shader
  - Required: Fixed-Function Pipeline Stages
    - `VkPipelineVertexInputStateCreateInfo`
    - `VkPipelineInputAssemblyStateCreateInfo`
    - `VkPipelineViewportStateCreateInfo`
    - `VkPipelineRasterizationStateCreateInfo`
    - `VkPipelineMultisampleStateCreateInfo`
    - `VkPipelineDepthStencilStateCreateInfo`
- Color-only rendering in a subpass with no depth/stencil attachment
  - Active Pipeline Shader Stages
    - Vertex Shader
    - Fragment Shader
  - Required: Fixed-Function Pipeline Stages
    - `VkPipelineVertexInputStateCreateInfo`
• VkPipelineInputAssemblyStateCreateInfo
• VkPipelineViewportStateCreateInfo
• VkPipelineRasterizationStateCreateInfo
• VkPipelineMultisampleStateCreateInfo
• VkPipelineColorBlendStateCreateInfo

• Rendering pipeline with tessellation and geometry shaders
  ◦ Active Pipeline Shader Stages
    • Vertex Shader
    • Tessellation Control Shader
    • Tessellation Evaluation Shader
    • Geometry Shader
    • Fragment Shader
  ◦ Required: Fixed-Function Pipeline Stages
    • VkPipelineVertexInputStateCreateInfo
    • VkPipelineInputAssemblyStateCreateInfo
    • VkPipelineTessellationStateCreateInfo
    • VkPipelineViewportStateCreateInfo
    • VkPipelineRasterizationStateCreateInfo
    • VkPipelineMultisampleStateCreateInfo
    • VkPipelineColorBlendStateCreateInfo

9.3. Pipeline destruction

To destroy a graphics or compute pipeline, call:

```c
void vkDestroyPipeline(
    VkDevice device,                // device, pipeline, pAllocator);
    VkPipeline pipeline,
    const VkAllocationCallbacks* pAllocator);
```

• `device` is the logical device that destroys the pipeline.
• `pipeline` is the handle of the pipeline to destroy.
• `pAllocator` controls host memory allocation as described in the Memory Allocation chapter.
Valid Usage

- All submitted commands that refer to `pipeline` must have completed execution.
- If `VkAllocationCallbacks` were provided when `pipeline` was created, a compatible set of callbacks must be provided here.
- If no `VkAllocationCallbacks` were provided when `pipeline` was created, `pAllocator` must be NULL.

Valid Usage (Implicit)

- `device` must be a valid `VkDevice` handle.
- If `pipeline` is not `VK_NULL_HANDLE`, `pipeline` must be a valid `VkPipeline` handle.
- If `pAllocator` is not NULL, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure.
- If `pipeline` is a valid handle, it must have been created, allocated, or retrieved from `device`.

Host Synchronization

- Host access to `pipeline` must be externally synchronized.

9.4. Multiple Pipeline Creation

Multiple pipelines can be created simultaneously by passing an array of `VkGraphicsPipelineCreateInfo` or `VkComputePipelineCreateInfo` structures into the `vkCreateGraphicsPipelines` and `vkCreateComputePipelines` commands, respectively. Applications can group together similar pipelines to be created in a single call, and implementations are encouraged to look for reuse opportunities within a group-create.

When an application attempts to create many pipelines in a single command, it is possible that some subset may fail creation. In that case, the corresponding entries in the `pPipelines` output array will be filled with `VK_NULL_HANDLE` values. If any pipeline fails creation (for example, due to out of memory errors), the `vkCreate*Pipelines` commands will return an error code. The implementation will attempt to create all pipelines, and only return `VK_NULL_HANDLE` values for those that actually failed.

9.5. Pipeline Derivatives

A pipeline derivative is a child pipeline created from a parent pipeline, where the child and parent are expected to have much commonality. The goal of derivative pipelines is that they be cheaper to create using the parent as a starting point, and that it be more efficient (on either host or device) to switch/bind between children of the same parent.
A derivative pipeline is created by setting the `VK_PIPELINE_CREATE_DERIVATIVE_BIT` flag in the `VkPipelineCreateInfo` structure. If this is set, then exactly one of `basePipelineHandle` or `basePipelineIndex` members of the structure **must** have a valid handle/index, and specifies the parent pipeline. If `basePipelineHandle` is used, the parent pipeline **must** have already been created. If `basePipelineIndex` is used, then the parent is being created in the same command. `VK_NULL_HANDLE` acts as the invalid handle for `basePipelineHandle`, and `-1` is the invalid index for `basePipelineIndex`. If `basePipelineIndex` is used, the base pipeline **must** appear earlier in the array. The base pipeline **must** have been created with the `VK_PIPELINE_CREATE_ALLOW_DERIVATIVES_BIT` flag set.

### 9.6. Pipeline Cache

Pipeline cache objects allow the result of pipeline construction to be reused between pipelines and between runs of an application. Reuse between pipelines is achieved by passing the same pipeline cache object when creating multiple related pipelines. Reuse across runs of an application is achieved by retrieving pipeline cache contents in one run of an application, saving the contents, and using them to preinitialize a pipeline cache on a subsequent run. The contents of the pipeline cache objects are managed by the implementation. Applications **can** manage the host memory consumed by a pipeline cache object and control the amount of data retrieved from a pipeline cache object.

Pipeline cache objects are represented by `VkPipelineCache` handles:

```cpp
VK_DEFINE_NON_DISPATCHABLE_HANDLE(VkPipelineCache)
```

To create pipeline cache objects, call:

```cpp
VkResult vkCreatePipelineCache(
    VkDevice device, 
    const VkPipelineCacheCreateInfo* pCreateInfo, 
    const VkAllocationCallbacks* pAllocator, 
    VkPipelineCache* pPipelineCache);
```

- **device** is the logical device that creates the pipeline cache object.
- **pCreateInfo** is a pointer to a `VkPipelineCacheCreateInfo` structure that contains the initial parameters for the pipeline cache object.
- **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.
- **pPipelineCache** is a pointer to a `VkPipelineCache` handle in which the resulting pipeline cache object is returned.
Applications can track and manage the total host memory size of a pipeline cache object using the pAllocator. Applications can limit the amount of data retrieved from a pipeline cache object in vkGetPipelineCacheData. Implementations should not internally limit the total number of entries added to a pipeline cache object or the total host memory consumed.

Once created, a pipeline cache can be passed to the vkCreateGraphicsPipelines and vkCreateComputePipelines commands. If the pipeline cache passed into these commands is not VK_NULL_HANDLE, the implementation will query it for possible reuse opportunities and update it with new content. The use of the pipeline cache object in these commands is internally synchronized, and the same pipeline cache object can be used in multiple threads simultaneously.

Implementations should make every effort to limit any critical sections to the actual accesses to the cache, which is expected to be significantly shorter than the duration of the vkCreateGraphicsPipelines and vkCreateComputePipelines commands.

Valid Usage (Implicit)

- **device** must be a valid VkDevice handle
- **pCreateInfo** must be a valid pointer to a valid VkPipelineCacheCreateInfo structure
- If **pAllocator** is not NULL, **pAllocator** must be a valid pointer to a valid VkAllocationCallbacks structure
- **pPipelineCache** must be a valid pointer to a VkPipelineCache handle

Return Codes

**Success**
- VK_SUCCESS

**Failure**
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY

The VkPipelineCacheCreateInfo structure is defined as:
```c
typedef struct VkPipelineCacheCreateInfo {
    VkStructureType sType;
    const void* pNext;
    VkPipelineCacheCreateFlags flags;
    size_t initialDataSize;
    const void* pInitialData;
} VkPipelineCacheCreateInfo;
```

- `sType` is the type of this structure.
- `pNext` is `NULL` or a pointer to an extension-specific structure.
- `flags` is reserved for future use.
- `initialDataSize` is the number of bytes in `pInitialData`. If `initialDataSize` is zero, the pipeline cache will initially be empty.
- `pInitialData` is a pointer to previously retrieved pipeline cache data. If the pipeline cache data is incompatible (as defined below) with the device, the pipeline cache will be initially empty. If `initialDataSize` is zero, `pInitialData` is ignored.

### Valid Usage

- If `initialDataSize` is not 0, it must be equal to the size of `pInitialData`, as returned by `vkGetPipelineCacheData` when `pInitialData` was originally retrieved.
- If `initialDataSize` is not 0, `pInitialData` must have been retrieved from a previous call to `vkGetPipelineCacheData`.

### Valid Usage (Implicit)

- `sType` must be `VK_STRUCTURE_TYPE_PIPELINE_CACHE_CREATE_INFO`.
- `pNext` must be `NULL`.
- `flags` must be 0.
- If `initialDataSize` is not 0, `pInitialData` must be a valid pointer to an array of `initialDataSize` bytes.

```c
typedef VkFlags VkPipelineCacheCreateFlags;
```

`VkPipelineCacheCreateFlags` is a bitmask type for setting a mask, but is currently reserved for future use.

Pipeline cache objects can be merged using the command:
VkResult vkMergePipelineCaches(
    VkDevice device,
    VkPipelineCache dstCache,
    uint32_t srcCacheCount,
    const VkPipelineCache* pSrcCaches);

• **device** is the logical device that owns the pipeline cache objects.
• **dstCache** is the handle of the pipeline cache to merge results into.
• **srcCacheCount** is the length of the **pSrcCaches** array.
• **pSrcCaches** is an array of pipeline cache handles, which will be merged into **dstCache**. The previous contents of **dstCache** are included after the merge.

<table>
<thead>
<tr>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>The details of the merge operation are implementation dependent, but implementations <strong>should</strong> merge the contents of the specified pipelines and prune duplicate entries.</td>
</tr>
</tbody>
</table>

### Valid Usage

• **dstCache** must not appear in the list of source caches

### Valid Usage (Implicit)

• **device** must be a valid **VkDevice** handle
• **dstCache** must be a valid **VkPipelineCache** handle
• **pSrcCaches** must be a valid pointer to an array of **srcCacheCount** valid **VkPipelineCache** handles
• **srcCacheCount** must be greater than 0
• **dstCache** must have been created, allocated, or retrieved from **device**
• Each element of **pSrcCaches** must have been created, allocated, or retrieved from **device**

### Host Synchronization

• Host access to **dstCache** must be externally synchronized
Raw Text:

**Return Codes**

**Success**
- VK_SUCCESS

**Failure**
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY

Data can be retrieved from a pipeline cache object using the command:

```c
VkResult vkGetPipelineCacheData(
    VkDevice                                    device,
    VkPipelineCache                             pipelineCache,
    size_t*                                     pDataSize,
    void*                                       pData);
```

- `device` is the logical device that owns the pipeline cache.
- `pipelineCache` is the pipeline cache to retrieve data from.
- `pDataSize` is a pointer to a value related to the amount of data in the pipeline cache, as described below.
- `pData` is either `NULL` or a pointer to a buffer.

If `pData` is `NULL`, then the maximum size of the data that can be retrieved from the pipeline cache, in bytes, is returned in `pDataSize`. Otherwise, `pDataSize` must point to a variable set by the user to the size of the buffer, in bytes, pointed to by `pData`, and on return the variable is overwritten with the amount of data actually written to `pData`.

If `pDataSize` is less than the maximum size that can be retrieved by the pipeline cache, at most `pDataSize` bytes will be written to `pData`, and `vkGetPipelineCacheData` will return VK_INCOMPLETE. Any data written to `pData` is valid and can be provided as the `pInitialData` member of the `VkPipelineCacheCreateInfo` structure passed to `vkCreatePipelineCache`.

Two calls to `vkGetPipelineCacheData` with the same parameters must retrieve the same data unless a command that modifies the contents of the cache is called between them.

Applications can store the data retrieved from the pipeline cache, and use these data, possibly in a future run of the application, to populate new pipeline cache objects. The results of pipeline compiles, however, may depend on the vendor ID, device ID, driver version, and other details of the device. To enable applications to detect when previously retrieved data is incompatible with the device, the initial bytes written to `pData` must be a header consisting of the following members:

**Table 7. Layout for pipeline cache header version VK_PIPELINE_CACHE_HEADER_VERSION_ONE**
<table>
<thead>
<tr>
<th>Offset</th>
<th>Size</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4</td>
<td>length in bytes of the entire pipeline cache header written as a stream of bytes, with the least significant byte first</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>a <code>VkPipelineCacheHeaderVersion</code> value written as a stream of bytes, with the least significant byte first</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>a vendor ID equal to <code>VkPhysicalDeviceProperties::vendorID</code> written as a stream of bytes, with the least significant byte first</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>a device ID equal to <code>VkPhysicalDeviceProperties::deviceID</code> written as a stream of bytes, with the least significant byte first</td>
</tr>
<tr>
<td>16</td>
<td><code>VK_UUID_SIZE</code></td>
<td>a pipeline cache ID equal to <code>VkPhysicalDeviceProperties::pipelineCacheUUID</code></td>
</tr>
</tbody>
</table>

The first four bytes encode the length of the entire pipeline cache header, in bytes. This value includes all fields in the header including the pipeline cache version field and the size of the length field.

The next four bytes encode the pipeline cache version, as described for `VkPipelineCacheHeaderVersion`. A consumer of the pipeline cache should use the cache version to interpret the remainder of the cache header.

If `pDataSize` is less than what is necessary to store this header, nothing will be written to `pData` and zero will be written to `pDataSize`.

**Valid Usage (Implicit)**

- `device` must be a valid `VkDevice` handle
- `pipelineCache` must be a valid `VkPipelineCache` handle
- `pDataSize` must be a valid pointer to a `size_t` value
- If the value referenced by `pDataSize` is not 0, and `pData` is not NULL, `pData` must be a valid pointer to an array of `pDataSize` bytes
- `pipelineCache` must have been created, allocated, or retrieved from `device`

**Return Codes**

**Success**
- `VK_SUCCESS`
- `VK_INCOMPLETE`

**Failure**
- `VK_ERROR_OUT_OF_HOST_MEMORY`
- `VK_ERROR_OUT_OF_DEVICE_MEMORY`
Possible values of the second group of four bytes in the header returned by `vkGetPipelineCacheData`, encoding the pipeline cache version, are:

```c
typedef enum VkPipelineCacheHeaderVersion {
    VK_PIPELINE_CACHE_HEADER_VERSION_ONE = 1,
} VkPipelineCacheHeaderVersion;
```

- `VK_PIPELINE_CACHE_HEADER_VERSION_ONE` specifies version one of the pipeline cache.

To destroy a pipeline cache, call:

```c
void vkDestroyPipelineCache(
    VkDevice device, 
    VkPipelineCache pipelineCache, 
    const VkAllocationCallbacks* pAllocator);
```

- `device` is the logical device that destroys the pipeline cache object.
- `pipelineCache` is the handle of the pipeline cache to destroy.
- `pAllocator` controls host memory allocation as described in the Memory Allocation chapter.

**Valid Usage**

- If `VkAllocationCallbacks` were provided when `pipelineCache` was created, a compatible set of callbacks **must** be provided here
- If no `VkAllocationCallbacks` were provided when `pipelineCache` was created, `pAllocator` **must** be `NULL`

**Valid Usage (Implicit)**

- `device` **must** be a valid `VkDevice` handle
- If `pipelineCache` is not `VK_NULL_HANDLE`, `pipelineCache` **must** be a valid `VkPipelineCache` handle
- If `pAllocator` is not `NULL`, `pAllocator` **must** be a valid pointer to a valid `VkAllocationCallbacks` structure
- If `pipelineCache` is a valid handle, it **must** have been created, allocated, or retrieved from `device`

**Host Synchronization**

- Host access to `pipelineCache` **must** be externally synchronized
9.7. Specialization Constants

Specialization constants are a mechanism whereby constants in a SPIR-V module can have their constant value specified at the time the VkPipeline is created. This allows a SPIR-V module to have constants that can be modified while executing an application that uses the Vulkan API.

Note
Specialization constants are useful to allow a compute shader to have its local workgroup size changed at runtime by the user, for example.

Each instance of the VkPipelineShaderStageCreateInfo structure contains a parameter pSpecializationInfo, which can be NULL to indicate no specialization constants, or point to a VkSpecializationInfo structure.

The VkSpecializationInfo structure is defined as:

```c
typedef struct VkSpecializationInfo {
    uint32_t mapEntryCount;
    const VkSpecializationMapEntry* pMapEntries;
    size_t dataSize;
    const void* pData;
} VkSpecializationInfo;
```

- **mapEntryCount** is the number of entries in the pMapEntries array.
- **pMapEntries** is a pointer to an array of VkSpecializationMapEntry which maps constant IDs to offsets in pData.
- **dataSize** is the byte size of the pData buffer.
- **pData** contains the actual constant values to specialize with.

pMapEntries points to a structure of type VkSpecializationMapEntry.

**Valid Usage**

- The offset member of each element of pMapEntries must be less than dataSize.
- The size member of each element of pMapEntries must be less than or equal to dataSize minus offset.

**Valid Usage (Implicit)**

- If mapEntryCount is not 0, pMapEntries must be a valid pointer to an array of mapEntryCount valid VkSpecializationMapEntry structures.
- If dataSize is not 0, pData must be a valid pointer to an array of dataSize bytes.
The `VkSpecializationMapEntry` structure is defined as:

```c
typedef struct VkSpecializationMapEntry {
    uint32_t    constantID;
    uint32_t    offset;
    size_t      size;
} VkSpecializationMapEntry;
```

- **constantID** is the ID of the specialization constant in SPIR-V.
- **offset** is the byte offset of the specialization constant value within the supplied data buffer.
- **size** is the byte size of the specialization constant value within the supplied data buffer.

If a `constantID` value is not a specialization constant ID used in the shader, that map entry does not affect the behavior of the pipeline.

### Valid Usage

- For a `constantID` specialization constant declared in a shader, `size` **must** match the byte size of the `constantID`. If the specialization constant is of type `boolean`, `size` **must** be the byte size of `VkBool32`.

In human readable SPIR-V:

```
OpDecorate %x SpecId 13 ; decorate .x component of WorkgroupSize with ID 13
OpDecorate %y SpecId 42 ; decorate .y component of WorkgroupSize with ID 42
OpDecorate %z SpecId 3  ; decorate .z component of WorkgroupSize with ID 3
OpDecorate %wgsize BuiltIn WorkgroupSize ; decorate WorkgroupSize onto constant
  %i32 = OpTypeInt 32 0 ; declare an unsigned 32-bit type
  %uvec3 = OpTypeVector %i32 3 ; declare a 3 element vector type of unsigned 32-bit
  %x = OpSpecConstant %i32 1 ; declare the .x component of WorkgroupSize
  %y = OpSpecConstant %i32 1 ; declare the .y component of WorkgroupSize
  %z = OpSpecConstant %i32 1 ; declare the .z component of WorkgroupSize
  %wgsize = OpSpecConstantComposite %uvec3 %x %y %z ; declare WorkgroupSize
```

From the above we have three specialization constants, one for each of the x, y & z elements of the `WorkgroupSize` vector.

Now to specialize the above via the specialization constants mechanism:
const VkSpecializationMapEntry entries[] =
{
    {
        13, // constantID
        0 * sizeof(uint32_t), // offset
        sizeof(uint32_t) // size
    },
    {
        42, // constantID
        1 * sizeof(uint32_t), // offset
        sizeof(uint32_t) // size
    },
    {
        3, // constantID
        2 * sizeof(uint32_t), // offset
        sizeof(uint32_t) // size
    }
};

const uint32_t data[] = { 16, 8, 4 }; // our workgroup size is 16x8x4

const VkSpecializationInfo info =
{
    3, // mapEntryCount
    entries, // pMapEntries
    3 * sizeof(uint32_t), // dataSize
    data, // pData
};

Then when calling `vkCreateComputePipelines`, and passing the `VkSpecializationInfo` we defined as the `pSpecializationInfo` parameter of `VkPipelineShaderStageCreateInfo`, we will create a compute pipeline with the runtime specified local workgroup size.

Another example would be that an application has a SPIR-V module that has some platform-dependent constants they wish to use.

In human readable SPIR-V:

```
OpDecorate %1 SpecId 0  ; decorate our signed 32-bit integer constant
OpDecorate %2 SpecId 12 ; decorate our 32-bit floating-point constant
%i32 = OpTypeInt 32 1  ; declare a signed 32-bit type
%float = OpTypeFloat 32 ; declare a 32-bit floating-point type
%1 = OpSpecConstant %i32 -1 ; some signed 32-bit integer constant
%2 = OpSpecConstant %float 0.5 ; some 32-bit floating-point constant
```

From the above we have two specialization constants, one is a signed 32-bit integer and the second is a 32-bit floating-point.
Now to specialize the above via the specialization constants mechanism:

```c
struct SpecializationData {
    int32_t data0;
    float data1;
};

const VkSpecializationMapEntry entries[] = {
    {0,             // constantID
     offsetof(SpecializationData, data0), // offset
     sizeof(SpecializationData::data0)   // size
    },
    {12,            // constantID
     offsetof(SpecializationData, data1), // offset
     sizeof(SpecializationData::data1)   // size
    }
};

SpecializationData data;
data.data0 = -42;     // set the data for the 32-bit integer
data.data1 = 42.0f;  // set the data for the 32-bit floating-point

const VkSpecializationInfo info = {
    2,             // mapEntryCount
    entries,       // pMapEntries
    sizeof(data),  // dataSize
    &data,         // pData
};
```

It is legal for a SPIR-V module with specializations to be compiled into a pipeline where no specialization info was provided. SPIR-V specialization constants contain default values such that if a specialization is not provided, the default value will be used. In the examples above, it would be valid for an application to only specialize some of the specialization constants within the SPIR-V module, and let the other constants use their default values encoded within the OpSpecConstant declarations.

### 9.8. Pipeline Binding

Once a pipeline has been created, it can be bound to the command buffer using the command:
```c
void vkCmdBindPipeline(
    VkCommandBuffer                             commandBuffer,
    VkPipelineBindPoint                         pipelineBindPoint,
    VkPipeline                                  pipeline);
```

- `commandBuffer` is the command buffer that the pipeline will be bound to.
- `pipelineBindPoint` is a `VkPipelineBindPoint` value specifying whether to bind to the compute or graphics bind point. Binding one does not disturb the other.
- `pipeline` is the pipeline to be bound.

Once bound, a pipeline binding affects subsequent graphics or compute commands in the command buffer until a different pipeline is bound to the bind point. The pipeline bound to `VK_PIPELINE_BIND_POINT_COMPUTE` controls the behavior of `vkCmdDispatch` and `vkCmdDispatchIndirect`. The pipeline bound to `VK_PIPELINE_BIND_POINT_GRAPHICS` controls the behavior of all drawing commands. No other commands are affected by the pipeline state.

### Valid Usage

- If `pipelineBindPoint` is `VK_PIPELINE_BIND_POINT_COMPUTE`, the `VkCommandPool` that `commandBuffer` was allocated from must support compute operations
- If `pipelineBindPoint` is `VK_PIPELINE_BIND_POINT_GRAPHICS`, the `VkCommandPool` that `commandBuffer` was allocated from must support graphics operations
- If `pipelineBindPoint` is `VK_PIPELINE_BIND_POINT_COMPUTE`, `pipeline` must be a compute pipeline
- If `pipelineBindPoint` is `VK_PIPELINE_BIND_POINT_GRAPHICS`, `pipeline` must be a graphics pipeline
- If the variable multisample rate feature is not supported, `pipeline` is a graphics pipeline, the current subpass has no attachments, and this is not the first call to this function with a graphics pipeline after transitioning to the current subpass, then the sample count specified by this pipeline must match that set in the previous pipeline

### Valid Usage (Implicit)

- `commandBuffer` must be a valid `VkCommandBuffer` handle
- `pipelineBindPoint` must be a valid `VkPipelineBindPoint` value
- `pipeline` must be a valid `VkPipeline` handle
- `commandBuffer` must be in the recording state
- The `VkCommandPool` that `commandBuffer` was allocated from must support graphics, or compute operations
- Both of `commandBuffer`, and `pipeline` must have been created, allocated, or retrieved from the same `VkDevice`
Host Synchronization

- Host access to commandBuffer must be externally synchronized.
- Host access to the VkCommandPool that commandBuffer was allocated from must be externally synchronized.

Command Properties

<table>
<thead>
<tr>
<th>Command Buffer Levels</th>
<th>Render Pass Scope</th>
<th>Supported Queue Types</th>
<th>Pipeline Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>Both</td>
<td>Graphics</td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td></td>
<td>Compute</td>
<td></td>
</tr>
</tbody>
</table>

Possible values of `vkCmdBindPipeline::pipelineBindPoint`, specifying the bind point of a pipeline object, are:

```c
typedef enum VkPipelineBindPoint {
    VK_PIPELINE_BIND_POINT_GRAPHICS = 0,
    VK_PIPELINE_BIND_POINT_COMPUTE = 1,
} VkPipelineBindPoint;
```

- `VK_PIPELINE_BIND_POINT_COMPUTE` specifies binding as a compute pipeline.
- `VK_PIPELINE_BIND_POINT_GRAPHICS` specifies binding as a graphics pipeline.

9.9. Dynamic State

When a pipeline object is bound, any pipeline object state that is not specified as dynamic is applied to the command buffer state. Pipeline object state that is specified as dynamic is not applied to the command buffer state at this time. Instead, dynamic state can be modified at any time and persists for the lifetime of the command buffer, or until modified by another dynamic state setting command or another pipeline bind.

When a pipeline object is bound, the following applies to each state parameter:

- If the state is not specified as dynamic in the new pipeline object, then that command buffer state is overwritten by the state in the new pipeline object.
- If the state is specified as dynamic in both the new and the previous pipeline object, then that command buffer state is not disturbed.
- If the state is specified as dynamic in the new pipeline object but is not specified as dynamic in the previous pipeline object, then that command buffer state becomes undefined. If the state is an array, then the entire array becomes undefined.
- If the state is an array specified as dynamic in both the new and the previous pipeline object, and the array size is not the same in both pipeline objects, then that command buffer state...
becomes undefined.

Dynamic state setting commands **must** not be issued for state that is not specified as dynamic in the bound pipeline object.

Dynamic state that does not affect the result of operations **can** be left undefined.

**Note**

For example, if blending is disabled by the pipeline object state then the dynamic color blend constants do not need to be specified in the command buffer, even if this state is specified as dynamic in the pipeline object.
Chapter 10. Memory Allocation

Vulkan memory is broken up into two categories, *host memory* and *device memory*.

10.1. Host Memory

Host memory is memory needed by the Vulkan implementation for non-device-visible storage.

*Note*

This memory *may* be used to store the implementation's representation and state of Vulkan objects.

Vulkan provides applications the opportunity to perform host memory allocations on behalf of the Vulkan implementation. If this feature is not used, the implementation will perform its own memory allocations. Since most memory allocations are off the critical path, this is not meant as a performance feature. Rather, this *can* be useful for certain embedded systems, for debugging purposes (e.g. putting a guard page after all host allocations), or for memory allocation logging.

Allocators are provided by the application as a pointer to a `VkAllocationCallbacks` structure:

```c
typedef struct VkAllocationCallbacks {
    void* pUserData;
    PFN_vkAllocationFunction pfnAllocation;
    PFN_vkReallocationFunction pfnReallocation;
    PFN_vkFreeFunction pfnFree;
    PFN_vkInternalAllocationNotification pfnInternalAllocation;
    PFN_vkInternalFreeNotification pfnInternalFree;
} VkAllocationCallbacks;
```

- *pUserData* is a value to be interpreted by the implementation of the callbacks. When any of the callbacks in `VkAllocationCallbacks` are called, the Vulkan implementation will pass this value as the first parameter to the callback. This value *can* vary each time an allocator is passed into a command, even when the same object takes an allocator in multiple commands.

- *pfnAllocation* is a pointer to an application-defined memory allocation function of type `PFN_vkAllocationFunction`.

- *pfnReallocation* is a pointer to an application-defined memory reallocation function of type `PFN_vkReallocationFunction`.

- *pfnFree* is a pointer to an application-defined memory free function of type `PFN_vkFreeFunction`.

- *pfnInternalAllocation* is a pointer to an application-defined function that is called by the implementation when the implementation makes internal allocations, and it is of type `PFN_vkInternalAllocationNotification`.

- *pfnInternalFree* is a pointer to an application-defined function that is called by the implementation when the implementation frees internal allocations, and it is of type `PFN_vkInternalFreeNotification`.
PFN_vkInternalFreeNotification.

Valid Usage

- **pfnAllocation** must be a valid pointer to a valid user-defined `PFN_vkAllocationFunction`
- **pfnReallocation** must be a valid pointer to a valid user-defined `PFN_vkReallocationFunction`
- **pfnFree** must be a valid pointer to a valid user-defined `PFN_vkFreeFunction`
- If either of **pfnInternalAllocation** or **pfnInternalFree** is not **NULL**, both must be valid callbacks

The type of **pfnAllocation** is:

```
typedef void* (VKAPI_PTR *PFN_vkAllocationFunction)(
    void* pUserData,
    size_t size,
    size_t alignment,
    VkSystemAllocationScope allocationScope);
```

- **pUserData** is the value specified for `VkAllocationCallbacks::pUserData` in the allocator specified by the application.
- **size** is the size in bytes of the requested allocation.
- **alignment** is the requested alignment of the allocation in bytes and must be a power of two.
- **allocationScope** is a `VkSystemAllocationScope` value specifying the allocation scope of the lifetime of the allocation, as described here.

If **pfnAllocation** is unable to allocate the requested memory, it must return **NULL**. If the allocation was successful, it must return a valid pointer to memory allocation containing at least **size** bytes, and with the pointer value being a multiple of **alignment**.

**Note**
Correct Vulkan operation cannot be assumed if the application does not follow these rules.

For example, **pfnAllocation** (or **pfnReallocation**) could cause termination of running Vulkan instance(s) on a failed allocation for debugging purposes, either directly or indirectly. In these circumstances, it cannot be assumed that any part of any affected `VkInstance` objects are going to operate correctly (even `vkDestroyInstance`), and the application must ensure it cleans up properly via other means (e.g. process termination).

If **pfnAllocation** returns **NULL**, and if the implementation is unable to continue correct processing of the current command without the requested allocation, it must treat this as a run-time error, and generate `VK_ERROR_OUT_OF_HOST_MEMORY` at the appropriate time for the command in which the
condition was detected, as described in Return Codes.

If the implementation is able to continue correct processing of the current command without the requested allocation, then it may do so, and must not generate VK_ERROR_OUT_OF_HOST_MEMORY as a result of this failed allocation.

The type of pfnReallocation is:

```
typedef void* (VKAPI_PTR *PFN_vkReallocationFunction)(
    void* pUserData,
    void* pOriginal,
    size_t size,
    size_t alignment,
    VkSystemAllocationScope allocationScope);
```

- **pUserData** is the value specified for VkAllocationCallbacks::pUserData in the allocator specified by the application.
- **pOriginal** must be either NULL or a pointer previously returned by pfnReallocation or pfnAllocation of the same allocator.
- **size** is the size in bytes of the requested allocation.
- **alignment** is the requested alignment of the allocation in bytes and must be a power of two.
- **allocationScope** is a VkSystemAllocationScope value specifying the allocation scope of the lifetime of the allocation, as described here.

pfnReallocation must return an allocation with enough space for size bytes, and the contents of the original allocation from bytes zero to min(original size, new size) - 1 must be preserved in the returned allocation. If size is larger than the old size, the contents of the additional space are undefined. If satisfying these requirements involves creating a new allocation, then the old allocation should be freed.

If pOriginal is NULL, then pfnReallocation must behave equivalently to a call to PFN_vkAllocationFunction with the same parameter values (without pOriginal).

If size is zero, then pfnReallocation must behave equivalently to a call to PFN_vkFreeFunction with the same pUserData parameter value, and pMemory equal to pOriginal.

If pOriginal is non-NULL, the implementation must ensure that alignment is equal to the alignment used to originally allocate pOriginal.

If this function fails and pOriginal is non-NULL the application must not free the old allocation.

pfnReallocation must follow the same rules for return values as PFN_vkAllocationFunction.

The type of pfnFree is:
typedef void (VKAPI_PTR *PFN_vkFreeFunction)(
    void* pUserData,
    void* pMemory);

• **pUserData** is the value specified for **VkAllocationCallbacks::pUserData** in the allocator specified by the application.

• **pMemory** is the allocation to be freed.

**pMemory may be NULL**, which the callback **must** handle safely. If **pMemory** is non-**NULL**, it **must** be a pointer previously allocated by **PFN_allocation** or **PFN_reallocation**. The application **should** free this memory.

The type of **PFN_internalAllocation** is:

typedef void (VKAPI_PTR *PFN_vkInternalAllocationNotification)(
    void* pUserData,
    size_t size,
    VkInternalAllocationType allocationType,
    VkSystemAllocationScope allocationScope);

• **pUserData** is the value specified for **VkAllocationCallbacks::pUserData** in the allocator specified by the application.

• **size** is the requested size of an allocation.

• **allocationType** is a **VkInternalAllocationType** value specifying the requested type of an allocation.

• **allocationScope** is a **VkSystemAllocationScope** value specifying the allocation scope of the lifetime of the allocation, as described here.

This is a purely informational callback.

The type of **PFN_internalFree** is:

typedef void (VKAPI_PTR *PFN_vkInternalFreeNotification)(
    void* pUserData,
    size_t size,
    VkInternalAllocationType allocationType,
    VkSystemAllocationScope allocationScope);

• **pUserData** is the value specified for **VkAllocationCallbacks::pUserData** in the allocator specified by the application.

• **size** is the requested size of an allocation.

• **allocationType** is a **VkInternalAllocationType** value specifying the requested type of an allocation.

• **allocationScope** is a **VkSystemAllocationScope** value specifying the allocation scope of the
lifetime of the allocation, as described here.

Each allocation has an allocation scope which defines its lifetime and which object it is associated with. Possible values passed to the allocationScope parameter of the callback functions specified by VkAllocationCallbacks, indicating the allocation scope, are:

```c
typedef enum VkSystemAllocationScope {
    VK_SYSTEM_ALLOCATION_SCOPE_COMMAND = 0,
    VK_SYSTEM_ALLOCATION_SCOPE_OBJECT = 1,
    VK_SYSTEM_ALLOCATION_SCOPE_CACHE = 2,
    VK_SYSTEM_ALLOCATION_SCOPE_DEVICE = 3,
    VK_SYSTEM_ALLOCATION_SCOPE_INSTANCE = 4,
} VkSystemAllocationScope;
```

- **VK_SYSTEM_ALLOCATION_SCOPE_COMMAND** specifies that the allocation is scoped to the duration of the Vulkan command.
- **VK_SYSTEM_ALLOCATION_SCOPE_OBJECT** specifies that the allocation is scoped to the lifetime of the Vulkan object that is being created or used.
- **VK_SYSTEM_ALLOCATION_SCOPE_CACHE** specifies that the allocation is scoped to the lifetime of a VkPipelineCache object.
- **VK_SYSTEM_ALLOCATION_SCOPE_DEVICE** specifies that the allocation is scoped to the lifetime of the Vulkan device.
- **VK_SYSTEM_ALLOCATION_SCOPE_INSTANCE** specifies that the allocation is scoped to the lifetime of the Vulkan instance.

Most Vulkan commands operate on a single object, or there is a sole object that is being created or manipulated. When an allocation uses an allocation scope of **VK_SYSTEM_ALLOCATION_SCOPE_OBJECT** or **VK_SYSTEM_ALLOCATION_SCOPE_CACHE**, the allocation is scoped to the object being created or manipulated.

When an implementation requires host memory, it will make callbacks to the application using the most specific allocator and allocation scope available:

- If an allocation is scoped to the duration of a command, the allocator will use the **VK_SYSTEM_ALLOCATION_SCOPE_COMMAND** allocation scope. The most specific allocator available is used: if the object being created or manipulated has an allocator, that object’s allocator will be used, else if the parent VkDevice has an allocator it will be used, else if the parent VkInstance has an allocator it will be used. Else,

- If an allocation is associated with an object of type VkPipelineCache, the allocator will use the **VK_SYSTEM_ALLOCATION_SCOPE_CACHE** allocation scope. The most specific allocator available is used (cache, else device, else instance). Else,

- If an allocation is scoped to the lifetime of an object, that object is being created or manipulated by the command, and that object’s type is not VkDevice or VkInstance, the allocator will use an allocation scope of **VK_SYSTEM_ALLOCATION_SCOPE_OBJECT**. The most specific allocator available is used (object, else device, else instance). Else,
• If an allocation is scoped to the lifetime of a device, the allocator will use an allocation scope of `VK_SYSTEM_ALLOCATION_SCOPE_DEVICE`. The most specific allocator available is used (device, else instance). Else,

• If the allocation is scoped to the lifetime of an instance and the instance has an allocator, its allocator will be used with an allocation scope of `VK_SYSTEM_ALLOCATION_SCOPE_INSTANCE`.

• Otherwise an implementation will allocate memory through an alternative mechanism that is unspecified.

Objects that are allocated from pools do not specify their own allocator. When an implementation requires host memory for such an object, that memory is sourced from the object’s parent pool’s allocator.

The application is not expected to handle allocating memory that is intended for execution by the host due to the complexities of differing security implementations across multiple platforms. The implementation will allocate such memory internally and invoke an application provided informational callback when these *internal allocations* are allocated and freed. Upon allocation of executable memory, `PFN_INTERNAL_ALLOCATION` will be called. Upon freeing executable memory, `PFN_INTERNAL_FREE` will be called. An implementation will only call an informational callback for executable memory allocations and frees.

The `allocationType` parameter to the `PFN_INTERNAL_ALLOCATION` and `PFN_INTERNAL_FREE` functions may be one of the following values:

```c
typedef enum VkInternalAllocationType {
    VK_INTERNAL_ALLOCATION_TYPE_EXECUTABLE = 0,
} VkInternalAllocationType;
```

• `VK_INTERNAL_ALLOCATION_TYPE_EXECUTABLE` specifies that the allocation is intended for execution by the host.

An implementation must only make calls into an application-provided allocator during the execution of an API command. An implementation must only make calls into an application-provided allocator from the same thread that called the provoking API command. The implementation should not synchronize calls to any of the callbacks. If synchronization is needed, the callbacks must provide it themselves. The informational callbacks are subject to the same restrictions as the allocation callbacks.

If an implementation intends to make calls through a `VkAllocationCallbacks` structure between the time a `vkCreate*` command returns and the time a corresponding `vkDestroy*` command begins, that implementation must save a copy of the allocator before the `vkCreate*` command returns. The callback functions and any data structures they rely upon must remain valid for the lifetime of the object they are associated with.

If an allocator is provided to a `vkCreate*` command, a compatible allocator must be provided to the corresponding `vkDestroy*` command. Two `VkAllocationCallbacks` structures are compatible if memory allocated with `PFN_ALLOCATION` or `PFN_REALLOCATION` in each can be freed with `PFN_REALLOCATION` or `PFN_FREE` in the other. An allocator must not be provided to a `vkDestroy*` command if an allocator was not provided to the corresponding `vkCreate*` command.
If a non-NULL allocator is used, the `PFNAllocation`, `PFNReallocation` and `PFNFree` members **must** be non-NULL and point to valid implementations of the callbacks. An application **can** choose to not provide informational callbacks by setting both `PFNInternalAllocation` and `PFNInternalFree` to `NULL`. `PFNInternalAllocation` and `PFNInternalFree` **must** either both be `NULL` or both be non-NULL.

If `PFNAllocation` or `PFNReallocation` fail, the implementation **may** fail object creation and/or generate an `VK_ERROR_OUT_OF_HOST_MEMORY` error, as appropriate.

Allocation callbacks **must** not call any Vulkan commands.

The following sets of rules define when an implementation is permitted to call the allocator callbacks.

`PFNAllocation` or `PFNReallocation` **may** be called in the following situations:

- Allocations scoped to a `VkDevice` or `VkInstance` **may** be allocated from any API command.
- Allocations scoped to a command **may** be allocated from any API command.
- Allocations scoped to a `VkPipelineCache` **may** only be allocated from:
  - `vkCreatePipelineCache`
  - `vkMergePipelineCaches` for `dstCache`
  - `vkCreateGraphicsPipelines` for `pipelineCache`
  - `vkCreateComputePipelines` for `pipelineCache`
- Allocations scoped to a `VkDescriptorPool` **may** only be allocated from:
  - any command that takes the pool as a direct argument
  - `vkAllocateDescriptorSets` for the `descriptorPool` member of its `pAllocateInfo` parameter
  - `vkCreateDescriptorPool`
- Allocations scoped to a `VkCommandPool` **may** only be allocated from:
  - any command that takes the pool as a direct argument
  - `vkCreateCommandPool`
  - `vkAllocateCommandBuffers` for the `commandPool` member of its `pAllocateInfo` parameter
  - any `vkCmd*` command whose `commandBuffer` was allocated from that `VkCommandPool`
- Allocations scoped to any other object **may** only be allocated in that object's `vkCreate*` command.

`PFNFree` **may** be called in the following situations:

- Allocations scoped to a `VkDevice` or `VkInstance` **may** be freed from any API command.
- Allocations scoped to a command **must** be freed by any API command which allocates such memory.
- Allocations scoped to a `VkPipelineCache` **may** be freed from `vkDestroyPipelineCache`.
- Allocations scoped to a `VkDescriptorPool` **may** be freed from
  - any command that takes the pool as a direct argument
• Allocations scoped to a VkCommandPool may be freed from:
  ◦ any command that takes the pool as a direct argument
  ◦ vkResetCommandBuffer whose commandBuffer was allocated from that VkCommandPool
• Allocations scoped to any other object may be freed in that object’s vkDestroy* command.
• Any command that allocates host memory may also free host memory of the same scope.

10.2. Device Memory

Device memory is memory that is visible to the device — for example the contents of the image or buffer objects, which can be natively used by the device.

Memory properties of a physical device describe the memory heaps and memory types available.

To query memory properties, call:

```c
void vkGetPhysicalDeviceMemoryProperties(
    VkPhysicalDevice                            physicalDevice,
    VkPhysicalDeviceMemoryProperties*           pMemoryProperties);
```

• physicalDevice is the handle to the device to query.
• pMemoryProperties points to an instance of VkPhysicalDeviceMemoryProperties structure in which the properties are returned.

**Valid Usage (Implicit)**

• physicalDevice must be a valid VkPhysicalDevice handle
• pMemoryProperties must be a valid pointer to a VkPhysicalDeviceMemoryProperties structure

The VkPhysicalDeviceMemoryProperties structure is defined as:

```c
typedef struct VkPhysicalDeviceMemoryProperties {
    uint32_t        memoryTypeCount;
    VkMemoryType    memoryTypes[VK_MAX_MEMORY_TYPES];
    uint32_t        memoryHeapCount;
    VkMemoryHeap    memoryHeaps[VK_MAX_MEMORY_HEAPs];
} VkPhysicalDeviceMemoryProperties;
```

• memoryTypeCount is the number of valid elements in the memoryTypes array.
• memoryTypes is an array of VkMemoryType structures describing the memory types that can be used to access memory allocated from the heaps specified by memoryHeaps.
• memoryHeapCount is the number of valid elements in the memoryHeaps array.
• memoryHeaps is an array of VkMemoryHeap structures describing the memory heaps from which
memory can be allocated.

The `VkPhysicalDeviceMemoryProperties` structure describes a number of *memory heaps* as well as a number of *memory types* that can be used to access memory allocated in those heaps. Each heap describes a memory resource of a particular size, and each memory type describes a set of memory properties (e.g. host cached vs uncached) that can be used with a given memory heap. Allocations using a particular memory type will consume resources from the heap indicated by that memory type's heap index. More than one memory type may share each heap, and the heaps and memory types provide a mechanism to advertise an accurate size of the physical memory resources while allowing the memory to be used with a variety of different properties.

The number of memory heaps is given by `memoryHeapCount` and is less than or equal to `VK_MAX_MEMORY_HEAPS`. Each heap is described by an element of the `memoryHeaps` array as a `VkMemoryHeap` structure. The number of memory types available across all memory heaps is given by `memoryTypeCount` and is less than or equal to `VK_MAX_MEMORY_TYPES`. Each memory type is described by an element of the `memoryTypes` array as a `VkMemoryType` structure.

At least one heap must include `VK_MEMORY_HEAP_DEVICE_LOCAL_BIT` in `VkMemoryHeap::flags`. If there are multiple heaps that all have similar performance characteristics, they may all include `VK_MEMORY_HEAP_DEVICE_LOCAL_BIT`. In a unified memory architecture (UMA) system there is often only a single memory heap which is considered to be equally “local” to the host and to the device, and such an implementation must advertise the heap as device-local.

Each memory type returned by `vkGetPhysicalDeviceMemoryProperties` must have its `propertyFlags` set to one of the following values:

- 0
- `VK_MEMORY_PROPERTY_HOST_VISIBLE_BIT` | `VK_MEMORY_PROPERTY_HOST_COHERENT_BIT`
- `VK_MEMORY_PROPERTY_HOST_VISIBLE_BIT` | `VK_MEMORY_PROPERTY_HOST_CACHED_BIT`
- `VK_MEMORY_PROPERTY_HOST_VISIBLE_BIT` | `VK_MEMORY_PROPERTY_HOST_CACHED_BIT` | `VK_MEMORY_PROPERTY_HOST_COHERENT_BIT`
- `VK_MEMORY_PROPERTY_DEVICE_LOCAL_BIT` | `VK_MEMORY_PROPERTY_HOST_VISIBLE_BIT` | `VK_MEMORY_PROPERTY_HOST_COHERENT_BIT`
- `VK_MEMORY_PROPERTY_DEVICE_LOCAL_BIT` | `VK_MEMORY_PROPERTY_HOST_VISIBLE_BIT` | `VK_MEMORY_PROPERTY_HOST_CACHED_BIT`
- `VK_MEMORY_PROPERTY_DEVICE_LOCAL_BIT` | `VK_MEMORY_PROPERTY_HOST_VISIBLE_BIT` | `VK_MEMORY_PROPERTY_HOST_CACHED_BIT` | `VK_MEMORY_PROPERTY_HOST_COHERENT_BIT`
- `VK_MEMORY_PROPERTY_DEVICE_LOCAL_BIT`
**VK_MEMORY_PROPERTY_LAZILY_ALLOCATED_BIT**

There must be at least one memory type with both the `VK_MEMORY_PROPERTY_HOST_VISIBLE_BIT` and `VK_MEMORY_PROPERTY_HOST_COHERENT_BIT` bits set in its `propertyFlags`. There must be at least one memory type with the `VK_MEMORY_PROPERTY_DEVICE_LOCAL_BIT` bit set in its `propertyFlags`.

For each pair of elements X and Y returned in `memoryTypes`, X must be placed at a lower index position than Y if:

- either the set of bit flags returned in the `propertyFlags` member of X is a strict subset of the set of bit flags returned in the `propertyFlags` member of Y.
- or the `propertyFlags` members of X and Y are equal, and X belongs to a memory heap with greater performance (as determined in an implementation-specific manner).

**Note**

There is no ordering requirement between X and Y elements for the case their `propertyFlags` members are not in a subset relation. That potentially allows more than one possible way to order the same set of memory types. Notice that the list of all allowed memory property flag combinations is written in a valid order. But if instead `VK_MEMORYPROPERTY_DEVICE_LOCAL_BIT` was before `VK_MEMORYPROPERTY_HOST_VISIBLE_BIT | VK_MEMORYPROPERTY_HOST_COHERENT_BIT`, the list would still be in a valid order.

This ordering requirement enables applications to use a simple search loop to select the desired memory type along the lines of:
```
int32_t findProperties(const VkPhysicalDeviceMemoryProperties* pMemoryProperties,
    uint32_t memoryTypeBitsRequirement,
    VkMemoryPropertyFlags requiredProperties) {
    const uint32_t memoryCount = pMemoryProperties->memoryTypeCount;
    for (uint32_t memoryIndex = 0; memoryIndex < memoryCount; ++memoryIndex) {
        const uint32_t memoryTypeBits = (1 << memoryIndex);
        const bool isRequiredMemoryType = memoryTypeBitsRequirement & memoryTypeBits;
        const VkMemoryPropertyFlags properties =
            pMemoryProperties->memoryTypes[memoryIndex].propertyFlags;
        const bool hasRequiredProperties =
            (properties & requiredProperties) == requiredProperties;
        if (isRequiredMemoryType && hasRequiredProperties) return static_cast<int32_t>(memoryIndex);
    }
    // failed to find memory type
    return -1;
}
```

The `VkMemoryHeap` structure is defined as:
```
typedef struct VkMemoryHeap {
    VkDeviceSize size;
    VkMemoryHeapFlags flags;
} VkMemoryHeap;
```

- `size` is the total memory size in bytes in the heap.
• **flags** is a bitmask of `VkMemoryHeapFlagBits` specifying attribute flags for the heap.

Bits which **may** be set in `VkMemoryHeap::flags`, indicating attribute flags for the heap, are:

```cpp
typedef enum VkMemoryHeapFlagBits {
    VK_MEMORY_HEAP_DEVICE_LOCAL_BIT = 0x00000001,
} VkMemoryHeapFlagBits;
```

• **VK_MEMORY_HEAP_DEVICE_LOCAL_BIT** specifies that the heap corresponds to device local memory. Device local memory **may** have different performance characteristics than host local memory, and **may** support different memory property flags.

```cpp
typedef VkFlags VkMemoryHeapFlags;
```

`VkMemoryHeapFlags` is a bitmask type for setting a mask of zero or more `VkMemoryHeapFlagBits`.

The `VkMemoryType` structure is defined as:

```cpp
typedef struct VkMemoryType {
    VkMemoryPropertyFlags propertyFlags;
    uint32_t heapIndex;
} VkMemoryType;
```

• **heapIndex** describes which memory heap this memory type corresponds to, and **must** be less than `memoryHeapCount` from the `VkPhysicalDeviceMemoryProperties` structure.

• **propertyFlags** is a bitmask of `VkMemoryPropertyFlagBits` of properties for this memory type.

Bits which **may** be set in `VkMemoryType::propertyFlags`, indicating properties of a memory heap, are:

```cpp
typedef enum VkMemoryPropertyFlagBits {
    VK_MEMORY_PROPERTY_DEVICE_LOCAL_BIT = 0x00000001,
    VK_MEMORY_PROPERTY_HOST_VISIBLE_BIT = 0x00000002,
    VK_MEMORY_PROPERTY_HOST_COHERENT_BIT = 0x00000004,
    VK_MEMORY_PROPERTY_HOST_CACHED_BIT = 0x00000008,
    VK_MEMORY_PROPERTY_LAZILY_ALLOCATED_BIT = 0x00000010,
} VkMemoryPropertyFlagBits;
```

• **VK_MEMORY_PROPERTY_DEVICE_LOCAL_BIT** bit specifies that memory allocated with this type is the most efficient for device access. This property will be set if and only if the memory type belongs to a heap with the `VK_MEMORY_HEAP_DEVICE_LOCAL_BIT` set.

• **VK_MEMORY_PROPERTY_HOST_VISIBLE_BIT** bit specifies that memory allocated with this type **can** be mapped for host access using `vkMapMemory`.

• **VK_MEMORY_PROPERTY_HOST_COHERENT_BIT** bit specifies that the host cache management commands
vkFlushMappedMemoryRanges and vkInvalidateMappedMemoryRanges are not needed to flush host writes to the device or make device writes visible to the host, respectively.

- **VK_MEMORY_PROPERTY_HOST_CACHED_BIT** bit specifies that memory allocated with this type is cached on the host. Host memory accesses to uncached memory are slower than to cached memory, however uncached memory is always host coherent.

- **VK_MEMORY_PROPERTY_LAZILY_ALLOCATED_BIT** bit specifies that the memory type only allows device access to the memory. Memory types must not have both **VK_MEMORY_PROPERTY_LAZILY_ALLOCATED_BIT** and **VK_MEMORY_PROPERTY_HOST_VISIBLE_BIT** set. Additionally, the object’s backing memory may be provided by the implementation lazily as specified in *Lazily Allocated Memory*.

```plaintext
typedef VkFlags VkMemoryPropertyFlags;
```

**VkMemoryPropertyFlags** is a bitmask type for setting a mask of zero or more **VkMemoryPropertyFlagBits**.

A Vulkan device operates on data in device memory via memory objects that are represented in the API by a **VkDeviceMemory** handle:

```c
VK_DEFINE_NON_DISPATCHABLE_HANDLE(VkDeviceMemory)
```

To allocate memory objects, call:

```c
VkDevice vkAllocateMemory(
    VkDevice device,
    const VkMemoryAllocateInfo* pAllocateInfo,
    const VkAllocationCallbacks* pAllocator,
    VkDeviceMemory* pMemory);
```

- **device** is the logical device that owns the memory.

- **pAllocateInfo** is a pointer to an instance of the **VkMemoryAllocateInfo** structure describing parameters of the allocation. A successful returned allocation must use the requested parameters — no substitution is permitted by the implementation.

- **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.

- **pMemory** is a pointer to a **VkDeviceMemory** handle in which information about the allocated memory is returned.

Allocations returned by **vkAllocateMemory** are guaranteed to meet any alignment requirement of the implementation. For example, if an implementation requires 128 byte alignment for images and 64 byte alignment for buffers, the device memory returned through this mechanism would be 128-byte aligned. This ensures that applications can correctly suballocate objects of different types (with potentially different alignment requirements) in the same memory object.

When memory is allocated, its contents are undefined.
The maximum number of valid memory allocations that can exist simultaneously within a VkDevice may be restricted by implementation- or platform-dependent limits. If a call to vkAllocateMemory would cause the total number of allocations to exceed these limits, such a call will fail and must return VK_ERROR_TOO_MANY_OBJECTS. The maxMemoryAllocationCount feature describes the number of allocations that can exist simultaneously before encountering these internal limits.

Some platforms may have a limit on the maximum size of a single allocation. For example, certain systems may fail to create allocations with a size greater than or equal to 4GB. Such a limit is implementation-dependent, and if such a failure occurs then the error VK_ERROR_OUT_OF_DEVICE_MEMORY must be returned.

Valid Usage

- pAllocateInfo->allocationSize must be less than or equal to VkPhysicalDeviceMemoryProperties::memoryHeaps[pAllocateInfo->memoryTypeIndex].size as returned by vkGetPhysicalDeviceMemoryProperties for the VkPhysicalDevice that device was created from.
- pAllocateInfo->memoryTypeIndex must be less than VkPhysicalDeviceMemoryProperties::memoryTypeCount as returned by vkGetPhysicalDeviceMemoryProperties for the VkPhysicalDevice that device was created from.

Valid Usage (Implicit)

- device must be a valid VkDevice handle
- pAllocateInfo must be a valid pointer to a valid VkMemoryAllocateInfo structure
- If pAllocator is not NULL, pAllocator must be a valid pointer to a valid VkAllocationCallbacks structure
- pMemory must be a valid pointer to a VkDeviceMemory handle

Return Codes

Success

- VK_SUCCESS

Failure

- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY
- VK_ERROR_TOO_MANY_OBJECTS

The VkMemoryAllocateInfo structure is defined as:
typedef struct VkMemoryAllocateInfo {
    VkStructureType    sType;
    const void*        pNext;
    VkDeviceSize       allocationSize;
    uint32_t           memoryTypeIndex;
} VkMemoryAllocateInfo;

- **sType** is the type of this structure.
- **pNext** is NULL or a pointer to an extension-specific structure.
- **allocationSize** is the size of the allocation in bytes
- **memoryTypeIndex** is an index identifying a memory type from the `memoryTypes` array of the `VkPhysicalDeviceMemoryProperties` structure

### Valid Usage

- **allocationSize** **must** be greater than 0

### Valid Usage (Implicit)

- **sType** **must** be `VK_STRUCTURE_TYPE_MEMORY_ALLOCATE_INFO`
- **pNext** **must** be NULL

To free a memory object, call:

```c
void vkFreeMemory(
    VkDevice                                    device,
    VkDeviceMemory                              memory,
    const VkAllocationCallbacks*                pAllocator);
```

- **device** is the logical device that owns the memory.
- **memory** is the `VkDeviceMemory` object to be freed.
- **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.

Before freeing a memory object, an application **must** ensure the memory object is no longer in use by the device—for example by command buffers in the pending state. The memory **can** remain bound to images or buffers at the time the memory object is freed, but any further use of them (on host or device) for anything other than destroying those objects will result in undefined behavior. If there are still any bound images or buffers, the memory **may** not be immediately released by the implementation, but **must** be released by the time all bound images and buffers have been destroyed. Once memory is released, it is returned to the heap from which it was allocated.

How memory objects are bound to Images and Buffers is described in detail in the Resource
If a memory object is mapped at the time it is freed, it is implicitly unmapped.

Note
As described below, host writes are not implicitly flushed when the memory object is unmapped, but the implementation must guarantee that writes that have not been flushed do not affect any other memory.

Valid Usage
• All submitted commands that refer to memory (via images or buffers) must have completed execution

Valid Usage (Implicit)
• device must be a valid VkDevice handle
• If memory is not VK_NULL_HANDLE, memory must be a valid VkDeviceMemory handle
• If pAllocator is not NULL, pAllocator must be a valid pointer to a valid VkAllocationCallbacks structure
• If memory is a valid handle, it must have been created, allocated, or retrieved from device

Host Synchronization
• Host access to memory must be externally synchronized

10.2.1. Host Access to Device Memory Objects
Memory objects created with vkAllocateMemory are not directly host accessible.

Memory objects created with the memory property VK_MEMORY_PROPERTY_HOST_VISIBLE_BIT are considered mappable. Memory objects must be mappable in order to be successfully mapped on the host.

To retrieve a host virtual address pointer to a region of a mappable memory object, call:

VkResult vkMapMemory(
    VkDevice device,
    VkDeviceMemory memory,
    VkDeviceSize offset,
    VkDeviceSize size,
    VkMemoryMapFlags flags,
    void** ppData);
• **device** is the logical device that owns the memory.
• **memory** is the VkDeviceMemory object to be mapped.
• **offset** is a zero-based byte offset from the beginning of the memory object.
• **size** is the size of the memory range to map, or VK_WHOLE_SIZE to map from offset to the end of the allocation.
• **flags** is reserved for future use.
• **pData** points to a pointer in which is returned a host-accessible pointer to the beginning of the mapped range. This pointer minus offset must be aligned to at least VkPhysicalDeviceLimits::minMemoryMapAlignment.

It is an application error to call *vkMapMemory* on a memory object that is already mapped.

---

**Note**

*vkMapMemory* will fail if the implementation is unable to allocate an appropriately sized contiguous virtual address range, e.g. due to virtual address space fragmentation or platform limits. In such cases, *vkMapMemory* must return VK_ERROR_MEMORY_MAP_FAILED. The application can improve the likelihood of success by reducing the size of the mapped range and/or removing unneeded mappings using *VkUnmapMemory*.

*vkMapMemory* does not check whether the device memory is currently in use before returning the host-accessible pointer. The application must guarantee that any previously submitted command that writes to this range has completed before the host reads from or writes to that range, and that any previously submitted command that reads from that range has completed before the host writes to that region (see here for details on fulfilling such a guarantee). If the device memory was allocated without the VK_MEMORY_PROPERTY_HOST_COHERENT_BIT set, these guarantees must be made for an extended range: the application must round down the start of the range to the nearest multiple of VkPhysicalDeviceLimits::nonCoherentAtomSize, and round the end of the range up to the nearest multiple of VkPhysicalDeviceLimits::nonCoherentAtomSize.

While a range of device memory is mapped for host access, the application is responsible for synchronizing both device and host access to that memory range.

---

**Note**

It is important for the application developer to become meticulously familiar with all of the mechanisms described in the chapter on Synchronization and Cache Control as they are crucial to maintaining memory access ordering.
Valid Usage

- memory must not be currently mapped
- offset must be less than the size of memory
- If size is not equal to VK_WHOLE_SIZE, size must be greater than 0
- If size is not equal to VK_WHOLE_SIZE, size must be less than or equal to the size of the memory minus offset
- memory must have been created with a memory type that reports VK_MEMORYPROPERTY_HOST_VISIBLE_BIT

Valid Usage (Implicit)

- device must be a valid VkDevice handle
- memory must be a valid VkDeviceMemory handle
- flags must be 0
- ppData must be a valid pointer to a pointer value
- memory must have been created, allocated, or retrieved from device

Host Synchronization

- Host access to memory must be externally synchronized

Return Codes

Success

- VK_SUCCESS

Failure

- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY
- VK_ERROR_MEMORY_MAP_FAILED

typedef VkFlags VkMemoryMapFlags;

VkMemoryMapFlags is a bitmask type for setting a mask, but is currently reserved for future use.

Two commands are provided to enable applications to work with non-coherent memory allocations: vkFlushMappedMemoryRanges and vkInvalidateMappedMemoryRanges.
Note
If the memory object was created with the `VK_MEMORY_PROPERTY_HOST_COHERENT_BIT` set, `vkFlushMappedMemoryRanges` and `vkInvalidateMappedMemoryRanges` are unnecessary and may have a performance cost. However, availability and visibility operations still need to be managed on the device. See the description of host access types for more information.

To flush ranges of non-coherent memory from the host caches, call:

```c
VkResult vkFlushMappedMemoryRanges(
    VkDevice                                    device,
    uint32_t                                    memoryRangeCount,
    const VkMappedMemoryRange*                  pMemoryRanges);
```

- **device** is the logical device that owns the memory ranges.
- **memoryRangeCount** is the length of the `pMemoryRanges` array.
- **pMemoryRanges** is a pointer to an array of `VkMappedMemoryRange` structures describing the memory ranges to flush.

`vkFlushMappedMemoryRanges` guarantees that host writes to the memory ranges described by `pMemoryRanges` are made available to the host memory domain, such that they can be made available to the device memory domain via memory domain operations using the `VK_ACCESS_HOST_WRITE_BIT` access type.

Within each range described by `pMemoryRanges`, each set of `nonCoherentAtomSize` bytes in that range is flushed if any byte in that set has been written by the host since it was first mapped, or the last time it was flushed. If `pMemoryRanges` includes sets of `nonCoherentAtomSize` bytes where no bytes have been written by the host, those bytes must not be flushed.

Unmapping non-coherent memory does not implicitly flush the mapped memory, and host writes that have not been flushed may not ever be visible to the device. However, implementations must ensure that writes that have not been flushed do not become visible to any other memory.

Note
The above guarantee avoids a potential memory corruption in scenarios where host writes to a mapped memory object have not been flushed before the memory is unmapped (or freed), and the virtual address range is subsequently reused for a different mapping (or memory allocation).
Valid Usage (Implicit)

- device must be a valid VkDevice handle
- pMemoryRanges must be a valid pointer to an array of memoryRangeCount valid VkMappedMemoryRange structures
- memoryRangeCount must be greater than 0

Return Codes

Success
- VK_SUCCESS

Failure
- VK_ERROR_OUT_OF_HOST_MEMORY
- VK_ERROR_OUT_OF_DEVICE_MEMORY

To invalidate ranges of non-coherent memory from the host caches, call:

```c
VkResult vkInvalidateMappedMemoryRanges(
  VkDevice                                    device,
  uint32_t                                    memoryRangeCount,
  const VkMappedMemoryRange*                  pMemoryRanges);
```

- device is the logical device that owns the memory ranges.
- memoryRangeCount is the length of the pMemoryRanges array.
- pMemoryRanges is a pointer to an array of VkMappedMemoryRange structures describing the memory ranges to invalidate.

vkInvalidateMappedMemoryRanges guarantees that device writes to the memory ranges described by pMemoryRanges, which have been made available to the host memory domain using the VK_ACCESS_HOST_WRITE_BIT and VK_ACCESS_HOST_READ_BIT access types, are made visible to the host. If a range of non-coherent memory is written by the host and then invalidated without first being flushed, its contents are undefined.

Within each range described by pMemoryRanges, each set of nonCoherentAtomSize bytes in that range is invalidated if any byte in that set has been written by the device since it was first mapped, or the last time it was invalidated.

**Note**

Mapping non-coherent memory does not implicitly invalidate the mapped memory, and device writes that have not been invalidated must be made visible before the host reads or overwrites them.
Valid Usage (Implicit)

- `device` must be a valid `VkDevice` handle
- `pMemoryRanges` must be a valid pointer to an array of `memoryRangeCount` valid `VkMappedMemoryRange` structures
- `memoryRangeCount` must be greater than 0

Return Codes

Success
- `VK_SUCCESS`

Failure
- `VK_ERROR_OUT_OF_HOST_MEMORY`
- `VK_ERROR_OUT_OF_DEVICE_MEMORY`

The `VkMappedMemoryRange` structure is defined as:

```c
typedef struct VkMappedMemoryRange {
    VkStructureType     sType;
    const void*          pNext;
    VkDeviceMemory   memory;
    VkDeviceSize       offset;
    VkDeviceSize       size;
} VkMappedMemoryRange;
```

- `sType` is the type of this structure.
- `pNext` is `NULL` or a pointer to an extension-specific structure.
- `memory` is the memory object to which this range belongs.
- `offset` is the zero-based byte offset from the beginning of the memory object.
- `size` is either the size of range, or `VK_WHOLE_SIZE` to affect the range from `offset` to the end of the current mapping of the allocation.
Valid Usage

- `memory` must be currently mapped
- If `size` is not equal to `VK_WHOLE_SIZE`, `offset` and `size` must specify a range contained within the currently mapped range of `memory`
- If `size` is equal to `VK_WHOLE_SIZE`, `offset` must be within the currently mapped range of `memory`
- If `size` is equal to `VK_WHOLE_SIZE`, the end of the current mapping of `memory` must be a multiple of `VkPhysicalDeviceLimits::nonCoherentAtomSize` bytes from the beginning of the memory object.
- `offset` must be a multiple of `VkPhysicalDeviceLimits::nonCoherentAtomSize`
- If `size` is not equal to `VK_WHOLE_SIZE`, `size` must either be a multiple of `VkPhysicalDeviceLimits::nonCoherentAtomSize`, or `offset` plus `size` must equal the size of `memory`.

Valid Usage (Implicit)

- `sType` must be `VK_STRUCTURE_TYPE_MAPPED_MEMORY_RANGE`
- `pNext` must be `NULL`
- `memory` must be a valid `VkDeviceMemory` handle

To unmap a memory object once host access to it is no longer needed by the application, call:

```c
void vkUnmapMemory(
    VkDevice device,  // device is the logical device that owns the memory.
    VkDeviceMemory memory);  // memory is the memory object to be unmapped.
```

Valid Usage

- `memory` must be currently mapped

Valid Usage (Implicit)

- `device` must be a valid `VkDevice` handle
- `memory` must be a valid `VkDeviceMemory` handle
- `memory` must have been created, allocated, or retrieved from `device`
10.2.2. Lazily Allocated Memory

If the memory object is allocated from a heap with the `VK_MEMORY_PROPERTY_LAZILY_ALLOCATED_BIT` bit set, that object's backing memory may be provided by the implementation lazily. The actual committed size of the memory may initially be as small as zero (or as large as the requested size), and monotonically increases as additional memory is needed.

A memory type with this flag set is only allowed to be bound to a `VkImage` whose usage flags include `VK_IMAGE_USAGE_TRANSIENT_ATTACHMENT_BIT`.

Note

Using lazily allocated memory objects for framebuffer attachments that are not needed once a render pass instance has completed may allow some implementations to never allocate memory for such attachments.

To determine the amount of lazily-allocated memory that is currently committed for a memory object, call:

```c
void vkGetDeviceMemoryCommitment(
    VkDevice device,                      // device is the logical device that owns the memory.
    VkDeviceMemory memory,               // memory is the memory object being queried.
    VkDeviceSize* pCommittedMemoryInBytes); // pCommittedMemoryInBytes is a pointer to a VkDeviceSize value in which the number of bytes currently committed is returned, on success.
```

The implementation may update the commitment at any time, and the value returned by this query may be out of date.

The implementation guarantees to allocate any committed memory from the `heapIndex` indicated by the memory type that the memory object was created with.

Valid Usage

- memory must have been created with a memory type that reports `VK_MEMORY_PROPERTY_LAZILY_ALLOCATED_BIT`
Valid Usage (Implicit)

- **device** must be a valid VkDevice handle
- **memory** must be a valid VkDeviceMemory handle
- **pCommittedMemoryInBytes** must be a valid pointer to a VkDeviceSize value
- **memory** must have been created, allocated, or retrieved from **device**
Chapter 11. Resource Creation

Vulkan supports two primary resource types: *buffers* and *images*. Resources are views of memory with associated formatting and dimensionality. Buffers are essentially unformatted arrays of bytes whereas images contain format information, can be multidimensional and may have associated metadata.

### 11.1. Buffers

Buffers represent linear arrays of data which are used for various purposes by binding them to a graphics or compute pipeline via descriptor sets or via certain commands, or by directly specifying them as parameters to certain commands.

Buffers are represented by *VkBuffer* handles:

```cpp
to create buffers, call:

```

```cpp
VkResult vkCreateBuffer(S
  VkDevice device,               pCreateInfo,  pCreateInfo,
  const VkBufferCreateInfo* pCreateInfo,
  const VkAllocationCallbacks* pAllocator,
  VkBuffer* pBuffer);
```

- *device* is the logical device that creates the buffer object.
- *pCreateInfo* is a pointer to an instance of the *VkBufferCreateInfo* structure containing parameters affecting creation of the buffer.
- *pAllocator* controls host memory allocation as described in the Memory Allocation chapter.
- *pBuffer* points to a *VkBuffer* handle in which the resulting buffer object is returned.

#### Valid Usage

- If the *flags* member of *pCreateInfo* includes `VK_BUFFER_CREATE_SPARSE_BINDING_BIT`, creating this *VkBuffer* must not cause the total required sparse memory for all currently valid sparse resources on the device to exceed *VkPhysicalDeviceLimits::sparseAddressSpaceSize*
Valid Usage (Implicit)

- **device** must be a valid `VkDevice` handle
- **pCreateInfo** must be a valid pointer to a valid `VkBufferCreateInfo` structure
- If **pAllocator** is not `NULL`, **pAllocator** must be a valid pointer to a valid `VkAllocationCallbacks` structure
- **pBuffer** must be a valid pointer to a `VkBuffer` handle

Return Codes

**Success**
- **VK_SUCCESS**

**Failure**
- **VK_ERROR_OUT_OF_HOST_MEMORY**
- **VK_ERROR_OUT_OF_DEVICE_MEMORY**

The `VkBufferCreateInfo` structure is defined as:

```c
typedef struct VkBufferCreateInfo {
    VkStructureType sType;
    const void* pNext;
    VkBufferCreateFlags flags;
    VkDeviceSize size;
    VkBufferUsageFlags usage;
    VkSharingMode sharingMode;
    uint32_t queueFamilyIndexCount;
    const uint32_t* pQueueFamilyIndices;
} VkBufferCreateInfo;
```

- **sType** is the type of this structure.
- **pNext** is `NULL` or a pointer to an extension-specific structure.
- **flags** is a bitmask of `VkBufferCreateFlagBits` specifying additional parameters of the buffer.
- **size** is the size in bytes of the buffer to be created.
- **usage** is a bitmask of `VkBufferUsageFlagBits` specifying allowed usages of the buffer.
- **sharingMode** is a `VkSharingMode` value specifying the sharing mode of the buffer when it will be accessed by multiple queue families.
- **queueFamilyIndexCount** is the number of entries in the `pQueueFamilyIndices` array.
- **pQueueFamilyIndices** is a list of queue families that will access this buffer (ignored if `sharingMode` is not `VK_SHARING_MODE_CONCURRENT`).
Valid Usage

- **size** must be greater than 0

- If `sharingMode` is `VK_SHARING_MODE_CONCURRENT`, `pQueueFamilyIndices` must be a valid pointer to an array of `queueFamilyIndexCount uint32_t` values

- If `sharingMode` is `VK_SHARING_MODE_CONCURRENT`, `queueFamilyIndexCount` must be greater than 1

- If `sharingMode` is `VK_SHARING_MODE_CONCURRENT`, each element of `pQueueFamilyIndices` must be unique and must be less than `pQueueFamilyPropertyCount` returned by `vkGetPhysicalDeviceQueueFamilyProperties` for the `physicalDevice` that was used to create device

- If the sparse bindings feature is not enabled, `flags` must not contain `VK_BUFFER_CREATE_SPARSE_BINDING_BIT`

- If the sparse buffer residency feature is not enabled, `flags` must not contain `VK_BUFFER_CREATE_SPARSE_RESIDENCY_BIT`

- If the sparse aliased residency feature is not enabled, `flags` must not contain `VK_BUFFER_CREATE_SPARSE_ALIASED_BIT`

- If `flags` contains `VK_BUFFER_CREATE_SPARSE_RESIDENCY_BIT` or `VK_BUFFER_CREATE_SPARSE_ALIASED_BIT`, it must also contain `VK_BUFFER_CREATE_SPARSE_BINDING_BIT`

Valid Usage (Implicit)

- **sType** must be `VK_STRUCTURE_TYPE_BUFFER_CREATE_INFO`

- **pNext** must be `NULL`

- **flags** must be a valid combination of `VkBufferCreateFlagBits` values

- **usage** must be a valid combination of `VkBufferUsageFlagBits` values

- **usage** must not be 0

- **sharingMode** must be a valid `VkSharingMode` value

Bits which can be set in `VkBufferCreateInfo::usage`, specifying usage behavior of a buffer, are:
typedef enum VkBufferUsageFlagBits {
    VK_BUFFER_USAGE_TRANSFER_SRC_BIT = 0x00000001,
    VK_BUFFER_USAGE_TRANSFER_DST_BIT = 0x00000002,
    VK_BUFFER_USAGE_UNIFORM_TEXEL_BUFFER_BIT = 0x00000004,
    VK_BUFFER_USAGE_STORAGE_TEXEL_BUFFER_BIT = 0x00000008,
    VK_BUFFER_USAGE_UNIFORM_BUFFER_BIT = 0x00000010,
    VK_BUFFER_USAGE_STORAGE_BUFFER_BIT = 0x00000020,
    VK_BUFFER_USAGE_INDEX_BUFFER_BIT = 0x00000040,
    VK_BUFFER_USAGE_VERTEX_BUFFER_BIT = 0x00000080,
    VK_BUFFER_USAGE_INDIRECT_BUFFER_BIT = 0x00000100,
} VkBufferUsageFlagBits;

• VK_BUFFER_USAGE_TRANSFER_SRC_BIT specifies that the buffer can be used as the source of a transfer command (see the definition of VK_PIPELINE_STAGE_TRANSFER_BIT).
• VK_BUFFER_USAGE_TRANSFER_DST_BIT specifies that the buffer can be used as the destination of a transfer command.
• VK_BUFFER_USAGE_UNIFORM_TEXEL_BUFFER_BIT specifies that the buffer can be used to create a VkBufferView suitable for occupying a VkDescriptorSet slot of type VK_DESCRIPTOR_TYPE_UNIFORM_TEXEL_BUFFER.
• VK_BUFFER_USAGE_STORAGE_TEXEL_BUFFER_BIT specifies that the buffer can be used to create a VkBufferView suitable for occupying a VkDescriptorSet slot of type VK_DESCRIPTOR_TYPE_STORAGE_TEXEL_BUFFER.
• VK_BUFFER_USAGE_UNIFORM_BUFFER_BIT specifies that the buffer can be used in a VkDescriptorBufferInfo suitable for occupying a VkDescriptorSet slot either of type VK_DESCRIPTOR_TYPE_UNIFORM_BUFFER or VK_DESCRIPTOR_TYPE_UNIFORM_BUFFER_DYNAMIC.
• VK_BUFFER_USAGE_STORAGE_BUFFER_BIT specifies that the buffer can be used in a VkDescriptorBufferInfo suitable for occupying a VkDescriptorSet slot either of type VK_DESCRIPTOR_TYPE_STORAGE_BUFFER or VK_DESCRIPTOR_TYPE_STORAGE_BUFFER_DYNAMIC.
• VK_BUFFER_USAGE_INDEX_BUFFER_BIT specifies that the buffer is suitable for passing as the buffer parameter to vkCmdBindIndexBuffer.
• VK_BUFFER_USAGE_VERTEX_BUFFER_BIT specifies that the buffer is suitable for passing as an element of the pBuffers array to vkCmdBindVertexBuffers.
• VK_BUFFER_USAGE_INDIRECT_BUFFER_BIT specifies that the buffer is suitable for passing as the buffer parameter to vkCmdDrawIndirect, vkCmdDrawIndexedIndirect, or vkCmdDispatchIndirect.

typedef VkFlags VkBufferUsageFlags;

VkBufferUsageFlags is a bitmask type for setting a mask of zero or more VkBufferUsageFlagBits.

Bits which can be set in VkBufferCreateInfo::flags, specifying additional parameters of a buffer, are:
typedef enum VkBufferCreateFlagBits {
    VK_BUFFER_CREATE_SPARSE_BINDING_BIT = 0x00000001,
    VK_BUFFER_CREATE_SPARSE_RESIDENCY_BIT = 0x00000002,
    VK_BUFFER_CREATE_SPARSE_ALIASED_BIT = 0x00000004,
} VkBufferCreateFlagBits;

• **VK_BUFFER_CREATE_SPARSE_BINDING_BIT** specifies that the buffer will be backed using sparse memory binding.

• **VK_BUFFER_CREATE_SPARSE_RESIDENCY_BIT** specifies that the buffer can be partially backed using sparse memory binding. Buffers created with this flag must also be created with the **VK_BUFFER_CREATE_SPARSE_BINDING_BIT** flag.

• **VK_BUFFER_CREATE_SPARSE_ALIASED_BIT** specifies that the buffer will be backed using sparse memory binding with memory ranges that might also simultaneously be backing another buffer (or another portion of the same buffer). Buffers created with this flag must also be created with the **VK_BUFFER_CREATE_SPARSE_BINDING_BIT** flag.

See [Sparse Resource Features](#) and [Physical Device Features](#) for details of the sparse memory features supported on a device.

typedef VkFlags VkBufferCreateFlags;

**VkBufferCreateFlags** is a bitmask type for setting a mask of zero or more **VkBufferCreateFlagBits**.

To destroy a buffer, call:

```c
void vkDestroyBuffer(
    VkDevice device,           
    VkBuffer buffer,           
    const VkAllocationCallbacks* pAllocator);
```

• **device** is the logical device that destroys the buffer.

• **buffer** is the buffer to destroy.

• **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.

**Valid Usage**

• All submitted commands that refer to **buffer**, either directly or via a **VkBufferView**, must have completed execution.

• If **VkAllocationCallbacks** were provided when **buffer** was created, a compatible set of callbacks must be provided here.

• If no **VkAllocationCallbacks** were provided when **buffer** was created, **pAllocator** must be **NULL**.
Valid Usage (Implicit)

- **device** must be a valid `VkDevice` handle
- If `buffer` is not `VK_NULL_HANDLE`, `buffer` must be a valid `VkBuffer` handle
- If `pAllocator` is not NULL, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure
- If `buffer` is a valid handle, it must have been created, allocated, or retrieved from `device`

Host Synchronization

- Host access to `buffer` must be externally synchronized

### 11.2. Buffer Views

A **buffer view** represents a contiguous range of a buffer and a specific format to be used to interpret the data. Buffer views are used to enable shaders to access buffer contents interpreted as formatted data. In order to create a valid buffer view, the buffer must have been created with at least one of the following usage flags:

- `VK_BUFFER_USAGE_UNIFORM_TEXEL_BUFFER_BIT`
- `VK_BUFFER_USAGE_STORAGE_TEXEL_BUFFER_BIT`

Buffer views are represented by `VkBufferView` handles:

```c
VK_DEFINE_NON_DISPATCHABLE_HANDLE(VkBufferView)
```

To create a buffer view, call:

```c
VkResult vkCreateBufferView(
    VkDevice device,  // device
    const VkBufferViewCreateInfo* pCreateInfo,  // pCreateInfo
    const VkAllocationCallbacks* pAllocator,  // pAllocator
    VkBufferView* pView);  // pView
```

- **device** is the logical device that creates the buffer view.
- **pCreateInfo** is a pointer to an instance of the `VkBufferViewCreateInfo` structure containing parameters to be used to create the buffer.
- **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.
- **pView** points to a `VkBufferView` handle in which the resulting buffer view object is returned.
Valid Usage (Implicit)

- **device** must be a valid `VkDevice` handle
- **pCreateInfo** must be a valid pointer to a valid `VkBufferViewCreateInfo` structure
- If `pAllocator` is not `NULL`, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure
- **pView** must be a valid pointer to a `VkBufferView` handle

Return Codes

**Success**
- `VK_SUCCESS`

**Failure**
- `VK_ERROR_OUT_OF_HOST_MEMORY`
- `VK_ERROR_OUT_OF_DEVICE_MEMORY`

The `VkBufferViewCreateInfo` structure is defined as:

```c
typedef struct VkBufferViewCreateInfo {
    VkStructureType sType;
    const void* pNext;
    VkBufferViewCreateFlags flags;
    VkBuffer buffer;
    VkFormat format;
    VkDeviceSize offset;
    VkDeviceSize range;
} VkBufferViewCreateInfo;
```

- **sType** is the type of this structure.
- **pNext** is `NULL` or a pointer to an extension-specific structure.
- **flags** is reserved for future use.
- **buffer** is a `VkBuffer` on which the view will be created.
- **format** is a `VkFormat` describing the format of the data elements in the buffer.
- **offset** is an offset in bytes from the base address of the buffer. Accesses to the buffer view from shaders use addressing that is relative to this starting offset.
- **range** is a size in bytes of the buffer view. If **range** is equal to `VK_WHOLE_SIZE`, the range from **offset** to the end of the buffer is used. If `VK_WHOLE_SIZE` is used and the remaining size of the buffer is not a multiple of the element size of **format**, then the nearest smaller multiple is used.
Valid Usage

• **offset** must be less than the size of **buffer**

• **offset** must be a multiple of **VkPhysicalDeviceLimits::minTexelBufferOffsetAlignment**

• If **range** is not equal to **VK_WHOLE_SIZE**, **range** must be greater than 0

• If **range** is not equal to **VK_WHOLE_SIZE**, **range** must be a multiple of the element size of **format**

• If **range** is not equal to **VK_WHOLE_SIZE**, **range** divided by the element size of **format** must be less than or equal to **VkPhysicalDeviceLimits::maxTexelBufferElements**

• If **range** is not equal to **VK_WHOLE_SIZE**, the sum of **offset** and **range** must be less than or equal to the size of **buffer**

• **buffer** must have been created with a **usage** value containing at least one of **VK_BUFFER_USAGE_UNIFORM_TEXEL_BUFFER_BIT** or **VK_BUFFER_USAGE_STORAGE_TEXEL_BUFFER_BIT**

• If **buffer** was created with **usage** containing **VK_BUFFER_USAGE_UNIFORM_TEXEL_BUFFER_BIT**, **format** must be supported for uniform texel buffers, as specified by the **VK_FORMAT_FEATURE_UNIFORM_TEXEL_BUFFER_BIT** flag in **VkFormatProperties::bufferFeatures** returned by **vkGetPhysicalDeviceFormatProperties**

• If **buffer** was created with **usage** containing **VK_BUFFER_USAGE_STORAGE_TEXEL_BUFFER_BIT**, **format** must be supported for storage texel buffers, as specified by the **VK_FORMAT_FEATURE_STORAGE_TEXEL_BUFFER_BIT** flag in **VkFormatProperties::bufferFeatures** returned by **vkGetPhysicalDeviceFormatProperties**

• If **buffer** is non-sparse then it must be bound completely and contiguously to a single **VkDeviceMemory** object

Valid Usage (Implicit)

• **sType** must be **VK_STRUCTURE_TYPE_BUFFER_VIEW_CREATE_INFO**

• **pNext** must be **NULL**

• **flags** must be 0

• **buffer** must be a valid **VkBuffer** handle

• **format** must be a valid **VkFormat** value

```
typedef VkFlags VkBufferViewCreateFlags;
```

**VkBufferViewCreateFlags** is a bitmask type for setting a mask, but is currently reserved for future use.

To destroy a buffer view, call:
void vkDestroyBufferView(
    VkDevice device, 
    VkBufferView bufferView, 
    const VkAllocationCallbacks* pAllocator);

- **device** is the logical device that destroys the buffer view.
- **bufferView** is the buffer view to destroy.
- **pAllocator** controls host memory allocation as described in the Memory Allocation chapter.

### Valid Usage

- All submitted commands that refer to `bufferView` must have completed execution.
- If `VkAllocationCallbacks` were provided when `bufferView` was created, a compatible set of callbacks must be provided here.
- If no `VkAllocationCallbacks` were provided when `bufferView` was created, `pAllocator` must be `NULL`.

### Valid Usage (Implicit)

- `device` must be a valid `VkDevice` handle.
- If `bufferView` is not `VK_NULL_HANDLE`, `bufferView` must be a valid `VkBufferView` handle.
- If `pAllocator` is not `NULL`, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure.
- If `bufferView` is a valid handle, it must have been created, allocated, or retrieved from `device`.

### Host Synchronization

- Host access to `bufferView` must be externally synchronized.

## 11.3. Images

Images represent multidimensional - up to 3 - arrays of data which can be used for various purposes (e.g. attachments, textures), by binding them to a graphics or compute pipeline via descriptor sets, or by directly specifying them as parameters to certain commands.

Images are represented by `VkImage` handles:

```
VK_DEFINE_NON_DISPATCHABLE_HANDLE(VkImage)
```
To create images, call:

```c
VkResult vkCreateImage(
    VkDevice                                    device,
    const VkImageCreateInfo*                    pCreateInfo,
    const VkAllocationCallbacks*                pAllocator,
    VkImage*                                    pImage);
```

- `device` is the logical device that creates the image.
- `pCreateInfo` is a pointer to an instance of the `VkImageCreateInfo` structure containing parameters to be used to create the image.
- `pAllocator` controls host memory allocation as described in the Memory Allocation chapter.
- `pImage` points to a `VkImage` handle in which the resulting image object is returned.

### Valid Usage

- If the `flags` member of `pCreateInfo` includes `VK_IMAGE_CREATE_SPARSE_BINDING_BIT`, creating this `VkImage` must not cause the total required sparse memory for all currently valid sparse resources on the device to exceed `VkPhysicalDeviceLimits::sparseAddressSpaceSize`.

### Valid Usage (Implicit)

- `device` must be a valid `VkDevice` handle
- `pCreateInfo` must be a valid pointer to a valid `VkImageCreateInfo` structure
- If `pAllocator` is not NULL, `pAllocator` must be a valid pointer to a valid `VkAllocationCallbacks` structure
- `pImage` must be a valid pointer to a `VkImage` handle

### Return Codes

**Success**
- `VK_SUCCESS`

**Failure**
- `VK_ERROR_OUT_OF_HOST_MEMORY`
- `VK_ERROR_OUT_OF_DEVICE_MEMORY`

The `VkImageCreateInfo` structure is defined as:
```c
typedef struct VkImageCreateInfo {
    VkStructureType          sType;
    const void*              pNext;
    VkImageCreateFlags       flags;
    VkImageType              imageType;
    VkFormat                 format;
    VkExtent3D               extent;
    uint32_t                 mipLevels;
    uint32_t                 arrayLayers;
    VkSampleCountFlagBits    samples;
    VkImageTiling            tiling;
    VkImageUsageFlags        usage;
    VkSharingMode            sharingMode;
    uint32_t                 queueFamilyIndexCount;
    const uint32_t*          pQueueFamilyIndices;
    VkImageLayout            initialLayout;
} VkImageCreateInfo;
```

- **sType** is the type of this structure.
- **pNext** is **NULL** or a pointer to an extension-specific structure.
- **flags** is a bitmask of `VkImageCreateFlagBits` describing additional parameters of the image.
- **imageType** is a `VkImageType` value specifying the basic dimensionality of the image. Layers in array textures do not count as a dimension for the purposes of the image type.
- **format** is a `VkFormat` describing the format and type of the data elements that will be contained in the image.
- **extent** is a `VkExtent3D` describing the number of data elements in each dimension of the base level.
- **mipLevels** describes the number of levels of detail available for minified sampling of the image.
- **arrayLayers** is the number of layers in the image.
- **samples** is the number of sub-data element samples in the image as defined in `VkSampleCountFlagBits`. See Multisampling.
- **tiling** is a `VkImageTiling` value specifying the tiling arrangement of the data elements in memory.
- **usage** is a bitmask of `VkImageUsageFlagBits` describing the intended usage of the image.
- **sharingMode** is a `VkSharingMode` value specifying the sharing mode of the image when it will be accessed by multiple queue families.
- **queueFamilyIndexCount** is the number of entries in the `pQueueFamilyIndices` array.
- **pQueueFamilyIndices** is a list of queue families that will access this image (ignored if `sharingMode` is not `VK_SHARING_MODE_CONCURRENT`).
- **initialLayout** is a `VkImageLayout` value specifying the initial `VkImageLayout` of all image subresources of the image. See Image Layouts.
Images created with \texttt{tiling} equal to \texttt{VK\_IMAGE\_TILING\_LINEAR} have further restrictions on their limits and capabilities compared to images created with \texttt{tiling} equal to \texttt{VK\_IMAGE\_TILING\_OPTIMAL}. Creation of images with tiling \texttt{VK\_IMAGE\_TILING\_LINEAR} \textbf{may} not be supported unless other parameters meet all of the constraints:

- \texttt{imageType} is \texttt{VK\_IMAGE\_TYPE\_2D}
- \texttt{format} is not a depth/stencil format
- \texttt{mipLevels} is 1
- \texttt{arrayLayers} is 1
- \texttt{samples} is \texttt{VK\_SAMPLE\_COUNT\_1\_BIT}
- \texttt{usage} only includes \texttt{VK\_IMAGE\_USAGE\_TRANSFER\_SRC\_BIT} and/or \texttt{VK\_IMAGE\_USAGE\_TRANSFER\_DST\_BIT}

Implementations \textbf{may} support additional limits and capabilities beyond those listed above.

To determine the set of valid \texttt{usage} bits for a given format, call \texttt{vkGetPhysicalDeviceFormatProperties}.

If the size of the resultant image would exceed \texttt{maxResourceSize}, then \texttt{vkCreateImage} \textbf{must} fail and return \texttt{VK\_ERROR\_OUT\_OF\_DEVICE\_MEMORY}. This failure \textbf{may} occur even when all image creation parameters satisfy their valid usage requirements.
Valid values for some image creation parameters are limited by a numerical upper bound or by inclusion in a bitset. For example, `VkImageCreateInfo::arrayLayers` is limited by `imageCreateMaxArrayLayers`, defined below; and `VkImageCreateInfo::samples` is limited by `imageCreateSampleCounts`, also defined below.

Several limiting values are defined below, as well as assisting values from which the limiting values are derived. The limiting values are referenced by the relevant valid usage statements of `VkImageCreateInfo`.

- Let `VkBool32 imageCreateMaybeLinear` indicate if the resultant image may be linear. (The definition below is trivial because certain extensions are disabled in this build of the specification).
  - If `tiling` is `VK_IMAGE_TILING_LINEAR`, then `imageCreateMaybeLinear` is `true`.
  - If `tiling` is `VK_IMAGE_TILING_OPTIMAL`, then `imageCreateMaybeLinear` is `false`.

- Let `VkFormatFeatureFlags imageCreateFormatFeatures` be the set of format features available during image creation.
  - If `tiling` is `VK_IMAGE_TILING_LINEAR`, then `imageCreateFormatFeatures` is the value of `VkImageFormatProperties::linearTilingFeatures` found by calling `vkGetPhysicalDeviceFormatProperties` with parameter `format` equal to `VkImageCreateInfo::format`.
  - If `tiling` is `VK_IMAGE_TILING_OPTIMAL`, then `imageCreateFormatFeatures` is value of `VkImageFormatProperties::optimalTilingFeatures` found by calling `vkGetPhysicalDeviceFormatProperties` with parameter `format` equal to `VkImageCreateInfo::format`. 