



Fast Forward

August 2018

Neil Trevett | Khronos President
NVIDIA | VP Developer Ecosystem

ntrevett@nvidia.com | [@neilt3d](https://twitter.com/neilt3d)
www.khronos.org



GENERATIONS / VANCOUVER
12-16 AUGUST
SIGGRAPH2018

Khronos Mission

PROMOTER MEMBERS

KHRONOS™ GROUP

Over 140 members worldwide
Any company is welcome to join



3D Incorporated Autodesk AMOTIVE Almalence Amazon.com antiLatency Au-Zone AXELL CORPORATION AXIS COMMUNICATIONS BASE MARK

BINOMIAL BLIZZARD ENTERTAINMENT BROADCOM BRENTWILL cadence CAICT codeplay codeweavers CTO continental

Coordinate COREAVI DASSAULT SYSTEMES DELL CATAPULT Digital DisplayLink DMP Esperanto Technologies ETRI Fraunhofer Futuremark

DAIICHI HITACHI HTC igalia IKEA Communications AB IKEA Imperial College London 財團法人資訊工業策進會 ITRI KALRAY AKDAB

KNU KYUNGPOOK NATIONAL UNIVERSITY LG Linaro logitech LUNAR magic leap matrox MAXON MEDIATEK mercury systems Mentor Graphics

Microsoft MIT Lincoln Laboratory MITSUBISHI ELECTRIC mobica moz://a NEC NIHON UNIVERSITY Nintendo NOKIA NXP

OSU oculus otoy owlii PEAKHILLS GROUP Pico PIXAR pluto Qt RAZER

RENESAS Rockwell Collins 서울 대 학 교 兆芯 SIRU INNOVATIONS SOCIONEXT SPREADTRUM STAR VR STREAM High Performance Computing

Synopsys TAKUMI TAMPERE UNIVERSITY OF TECHNOLOGY TEXAS INSTRUMENTS thinci Think Silicon tobii umbra unity UCL

University of BRISTOL UNIVERSITY OF TORONTO University of Windsor UX3D VIZIO Visteon VMware vrg. XILINX ZERO LATENCY zSpace

Khronos is an International Industry Consortium creating royalty-free, **open standards** to enable software to access hardware acceleration for **3D Graphics, Virtual and Augmented Reality, Parallel Computing, Neural Networks and Vision Processing**

Software



Silicon

Khronos *Fast Forward*

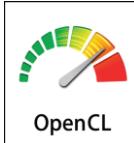
3D Graphics

Run-time APIs



Heterogenous Compute
(Parallel Processing)

SYCL™



Vulkan.

VR and AR
Vision and Inferencing

OpenVX™

OpenXR™

File Formats



Vulkan and New Generation GPU APIs

Non-proprietary, royalty-free open standard ‘By the industry for the industry’
Portable across multiple platforms - desktop and mobile
Modern architecture | Low overhead | Multi-thread friendly
EXPLICIT GPU access for EFFICIENT, LOW-LATENCY,
PREDICTABLE performance



Vulkan is available on Android 7.0+

Pervasive Vulkan

Major GPU Companies supporting Vulkan for Desktop and Mobile Platforms



<http://vulkan.gpuinfo.org/>

Platforms



Desktop



Mobile
(Android 7.0+)



Media Players



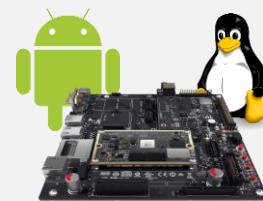
Consoles



Virtual Reality



Cloud Services



Embedded



CRYENGINE®



Vulkan 1.1 Launch and Ongoing Momentum

Strengthening the Ecosystem

Improved developer tools (SDK, validation/debug layers)

More rigorous conformance testing

Shader toolchain improvements (size, speed, robustness)

Shading language flexibility - HLSL and OpenCL C support

Vulkan Public Ecosystem Forum



February 2016
Vulkan 1.0

Explicit Access to
GPU Acceleration

Vulkan 1.0 Extensions

Maintenance updates plus additional functionality

Explicit Building Blocks for VR:

e.g. multiview

Explicit Building Blocks for
Homogeneous Multi-GPU

Enhanced Windows System Integration

Increased Shader Flexibility:

16 bit storage, Variable Pointers

Enhanced Cross-Process and
Cross-API Sharing



March 2018
Vulkan 1.1

Integration of 1.0 Extensions.
New Technology into Core e.g.
Subgroup Operations

Widening Platform Support

Pervasive GPU vendor native driver availability

Open source drivers - ANV (Intel), AMDVLK/RADV (AMD)

Port Vulkan apps to macOS/iOS and DX12

Building Vulkan's Future

Deliver complete ecosystem - not just specs

Listen and prioritize developer needs

Drive GPU technology

Released Vulkan 1.1 Extensions

KHR_draw_indirect_count

Source draw count parameter from a buffer in GPU-writable
memory for greater flexibility for GPU-generated work

KHR_8bit_storage

8-bit types in uniform and storage buffers for improved compute
support in apps such as inferencing and vision

EXT_descriptor_indexing

Dynamically non-uniform (aka bindless) resource access
Required by some modern game engine architectures

Discussions in Flight

Reduced precision arithmetic types in shaders

Detailed driver property queries

Variable-resolution rendering

Cross-vendor performance counter queries

Memory residency management

Depth/stencil resolve

Ray tracing

Video

New synchronization primitives

Content is shipping on desktop...



Vulkan-only AAA
Titles on PC



Dota 2 on PC
and macOS

AAA titles on Linux



...and Mobile



Plus....
Lineage 2 Revolution
Heroes of Incredible Tales
Dream League Soccer...



CODEMASTERS®



Fortnite on Android!



FORTNITE

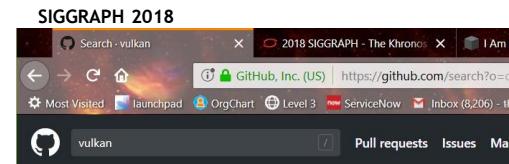
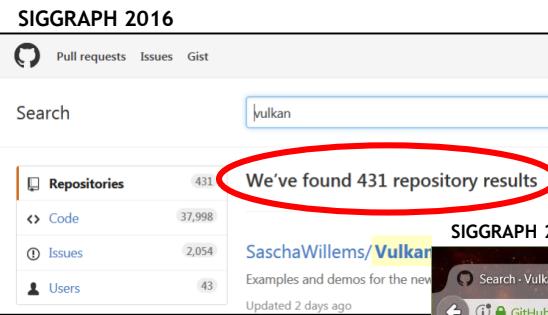
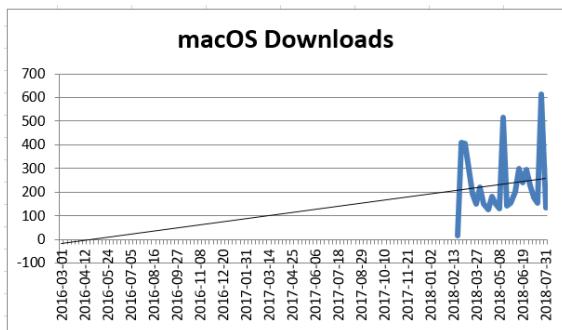
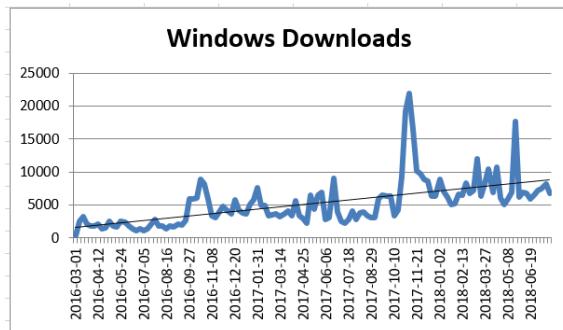
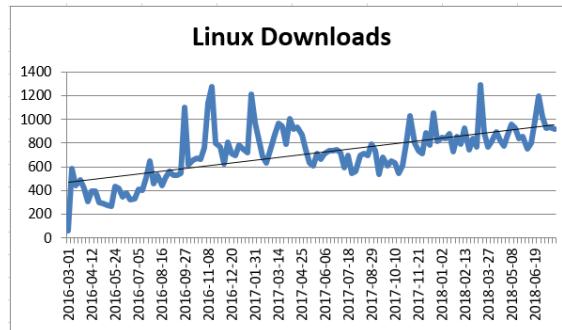
Vulkan.

Fortnite on Android uses Vulkan on select phones for optimal performance, including the best-performing Samsung - the Galaxy Note9

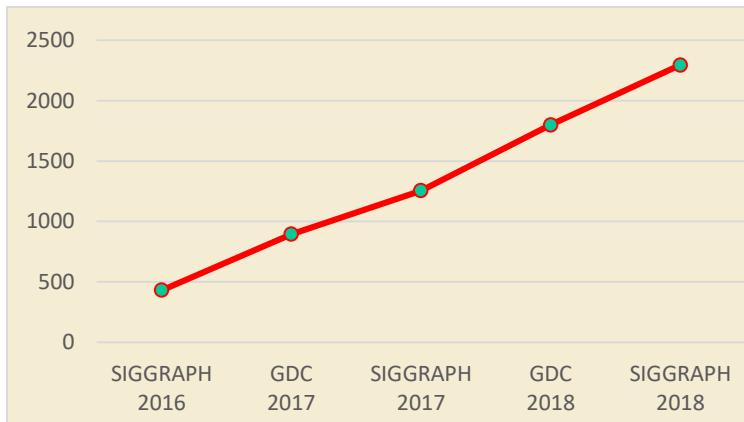
Vulkan Developer Activity - SDK and GitHub



LunarG Vulkan SDK
Download rate increases every year since launch
<http://vulkan.lunarg.com>

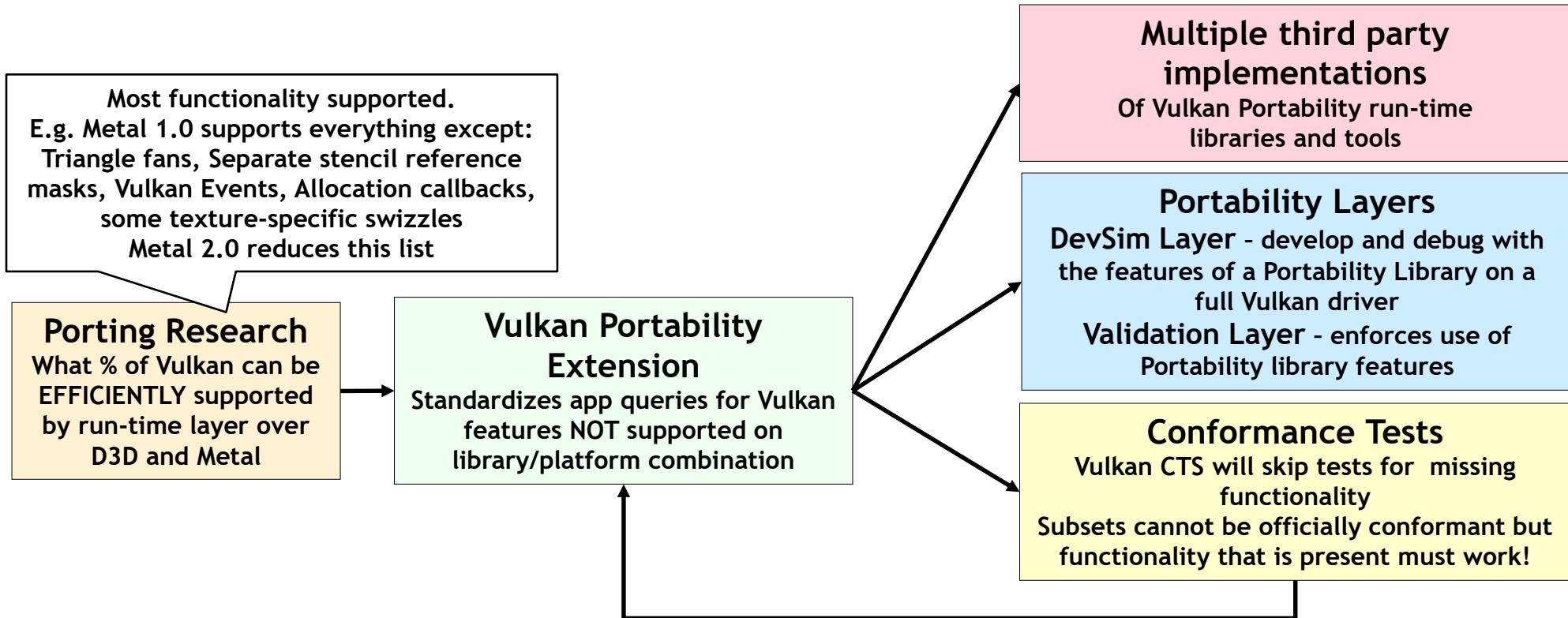


Vulkan Related GitHub Repos



Vulkan Portability Initiative

Enabling and accelerating the creation of tools and run-time libraries for Vulkan applications to run on platforms supporting only Metal or Direct3D



Bringing Vulkan Apps to Apple Platforms Today



Dota 2 running on Mac up to 50% faster than native OpenGL



Applications

First production apps using MoltenVK already shipping on macOS and iOS

Vulkan
macOS SDK

Open source SDK to build, run, and debug applications on macOS including validation layer support



SPIRV-Cross

Convert SPIR-V shaders to platform source formats

KHRONOS™
GROUP

macOS / iOS
Run-time

Maps Vulkan to Metal

MoltenVK for macOS and iOS
For macOS 10.11, iOS 9.0 and up



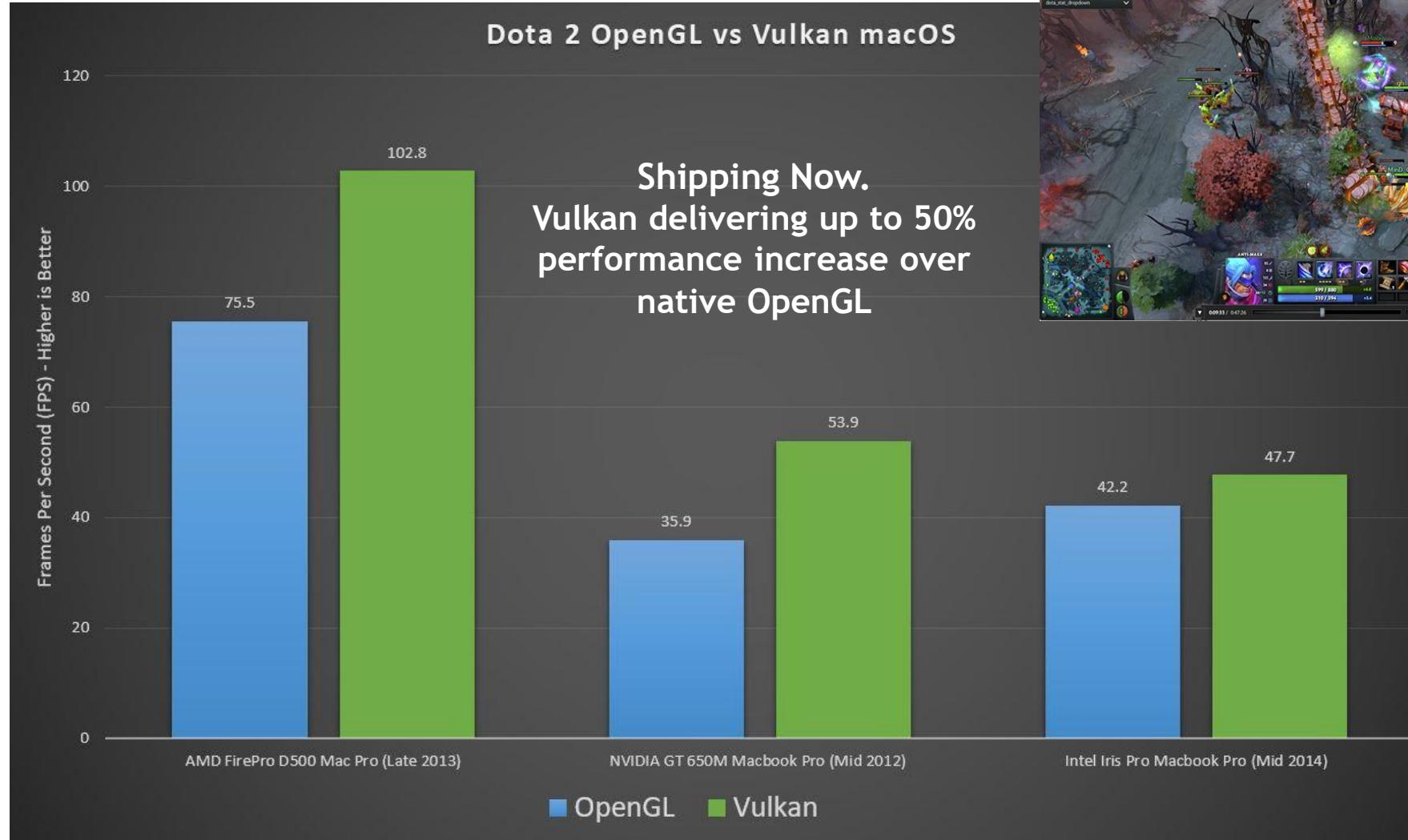
Beta release - working to pass all applicable conformance tests

Previously a paid product

Now released into OPEN SOURCE

Completely free to use - no fees or royalties - including commercial applications

Valve - Vulkan Dota 2 on macOS



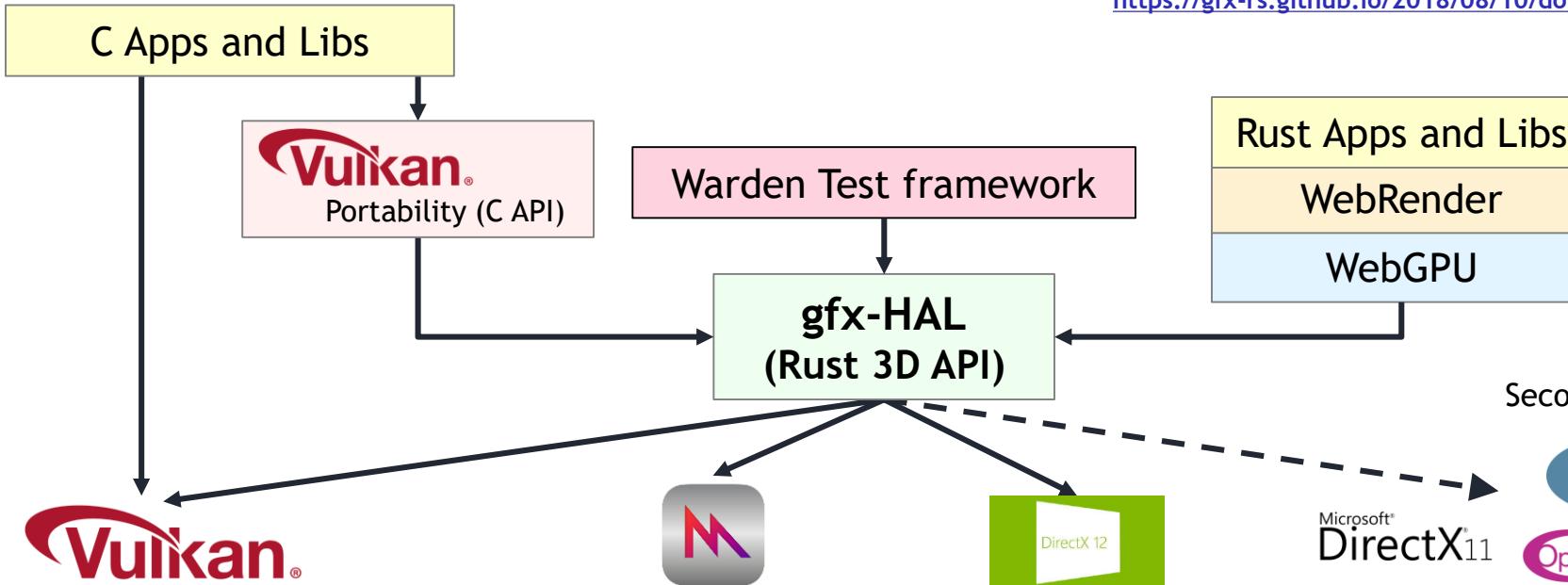
gfx-rs

- Vulkan Portability over D3D, Metal, and OpenGL
 - D3D layer useful for Vulkan on UWP platforms such as Windows 10 S, Polaris, Xbox One
 - <https://github.com/gfx-rs/gfx>
 - <https://github.com/gfx-rs/portability>



Efficiently running Dota 2 on Mac - working to increase conformance coverage

<https://gfx-rs.github.io/2018/08/10/dota2-macos-performance.html>



Secondary Backends

OpenGL.
Microsoft
DirectX 11
OpenGL|ES.

OpenGL and OpenGL ES



January 2018

OpenGL 4.6 conformance test suite
released in open source

Intel and NVIDIA released conformant
OpenGL 4.6 drivers

April 2018

OpenGL 4.6.0.1 CTS bugfix update
released in April

June 2018

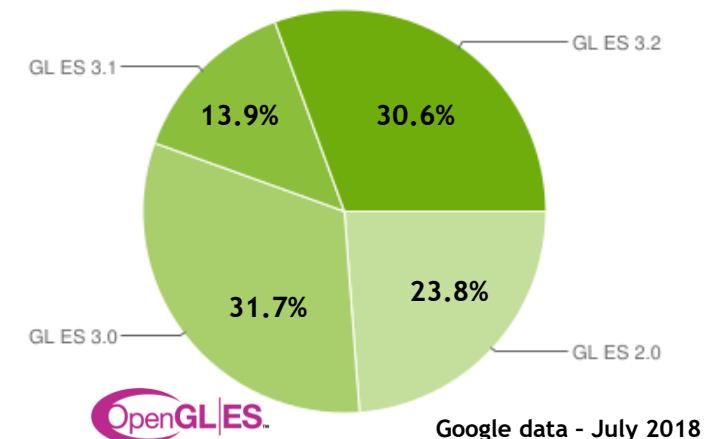
OpenGL ES CTS 3.2.5.0
released in open source

Raises the quality bar for OpenGL
3.2 implementations

OpenGL ES still the most prevalent
3D API (billions of units!)
More conformant products added
OpenGL ES 3.2 adoption increasing

**Working Group Meetings Merged under one Chairperson
for Improved Efficiency**

GLSL and ESSL specs merged and migrated from LibreOffice to
AsciiDoctor to improve maintainability and reduce divergence
OpenGL 4.6, OpenGL ES 3.2, GLSL 4.60 and ESSL 3.20 specs June 2018
Lots of bug fixes - many leveraged from open GitHub projects



OpenGL ES and WebGL Evolution

Pervasive OpenGL ES 2.0
OpenGL and OpenGL ES ships on every desktop and mobile OS 3D on the Web is enabled!

Mobile Graphics
Programmable Vertex and Fragment shaders



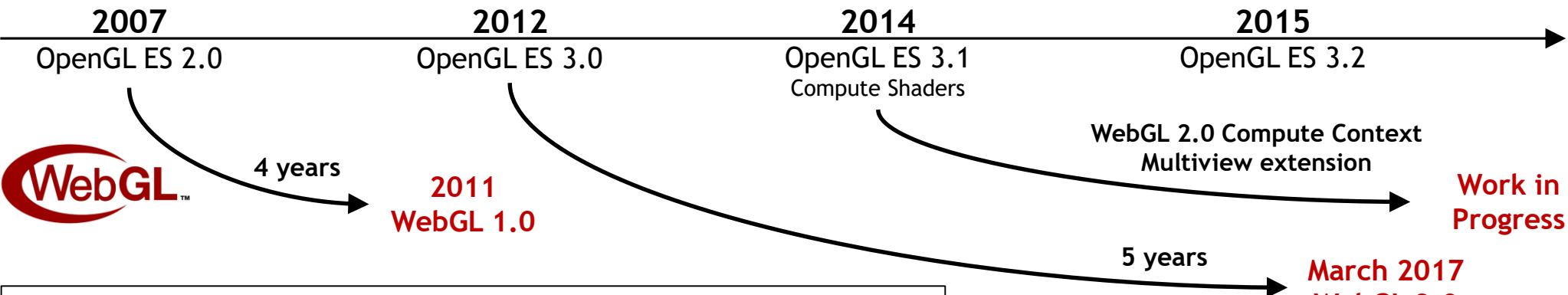
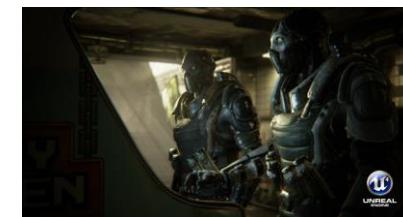
Desktop Graphics
Textures: NPOT, 3D, Depth, Arrays, Int/float
Objects: Query, Sync, Samplers
Seamless Cubemaps, Integer vertex attributes
Multiple Render Targets, Instanced rendering
Transform feedback, Uniform blocks
Vertex array objects, GLSL ES 3.0 shaders



Compute Shaders



Advanced Graphics
Tessellation and geometry shaders
ASTC Texture Compression
Floating point render targets
Debug and robustness for security



Conformance Testing is vital for Cross-Platform Reliability
WebGL 2.0 conformance tests are very thorough 10x more tests than WebGL 1.0 tests

WebGL Momentum - WebGL 2.0 is Here!



FLASH & THE FUTURE OF INTERACTIVE CONTENT

POSTED BY ADOBE CORPORATE COMMUNICATIONS ON JULY 25, 2017

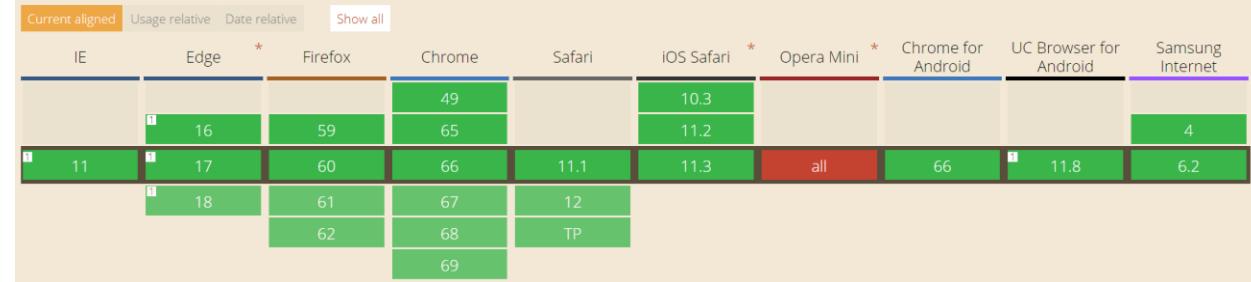
Subscribe

Adobe has long played a leadership role in advancing interactivity and creative content – from video, to games and more – on the web. Where we've seen a need to push content and interactivity forward, we've innovated to meet those needs. Where a format didn't exist, we invented one – such as with Flash and Shockwave. And over time, as the web evolved, these new formats were adopted by the community. In some cases formed the basis for open standards, and became an essential part of the web.

But as open standards like [HTML5](#), [WebGL](#) and [WebAssembly](#) have matured over the past several years, most now provide many of the capabilities and functionalities that plugins pioneered and have become a viable alternative for content on the web. Over time, we've seen helper apps evolve to become plugins, and more recently, have seen many of these plugin capabilities get incorporated into open web standards. Today, most browser vendors are integrating capabilities once provided by plugins directly into browsers and deprecating plugins.

WebGL - 3D Canvas graphics

Method of generating dynamic 3D graphics using JavaScript, accelerated through hardware



<http://caniuse.com/#feat=webgl>

62.85% Globally

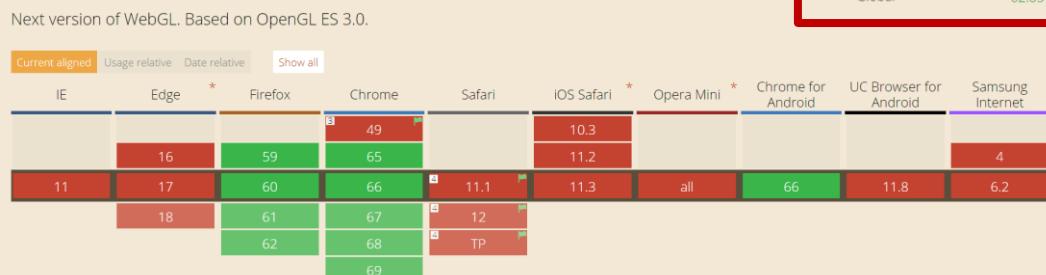
WebGL 2.0

Next version of WebGL. Based on OpenGL ES 3.0.

Usage: % of all users

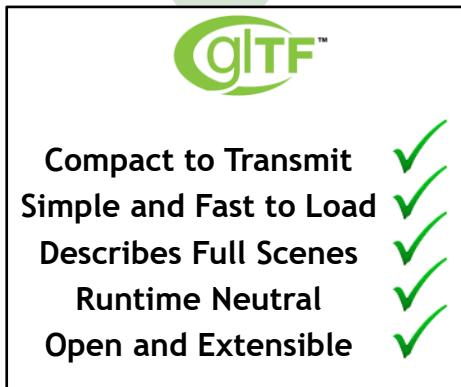
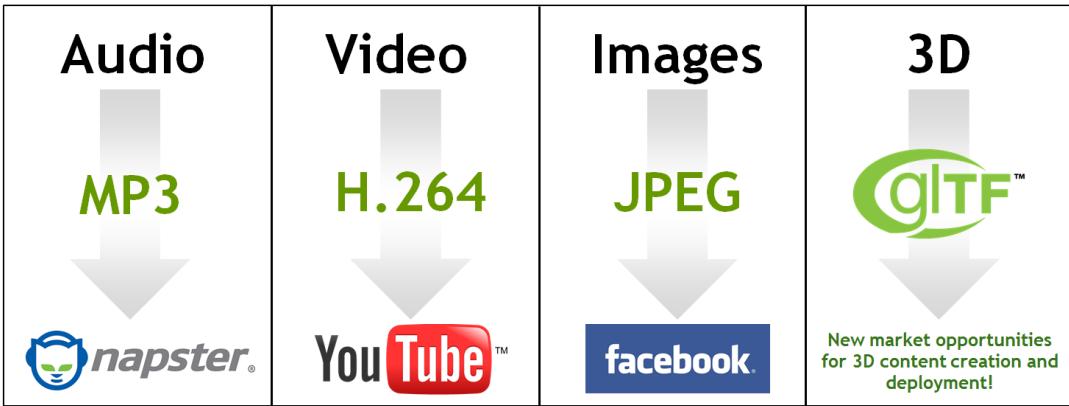
Global

62.85%



WebGL 2.0 brings Desktop-class
graphics to the Web
The time to create a new class of
Web-based 3D Apps is now!

glTF - Cross-Platform 3D Asset Transmission



Efficient, reliable transmission
Bring 3D assets into 1000s of apps
and engines - NOT an
Authoring Interchange Format



glTF 1.0 - December 2015
Primarily for WebGL
Uses GLSL for materials

glTF spec development
on open GitHub - get involved!
<https://github.com/KhronosGroup/glTF>



glTF 2.0 - June 2017
Native AND Web APIs
Physically Based Rendering
Metallic-Roughness and Specular-Glossiness

glTF Ecosystem

Creation Tools



modo



Assimp

Open Asset Import Library



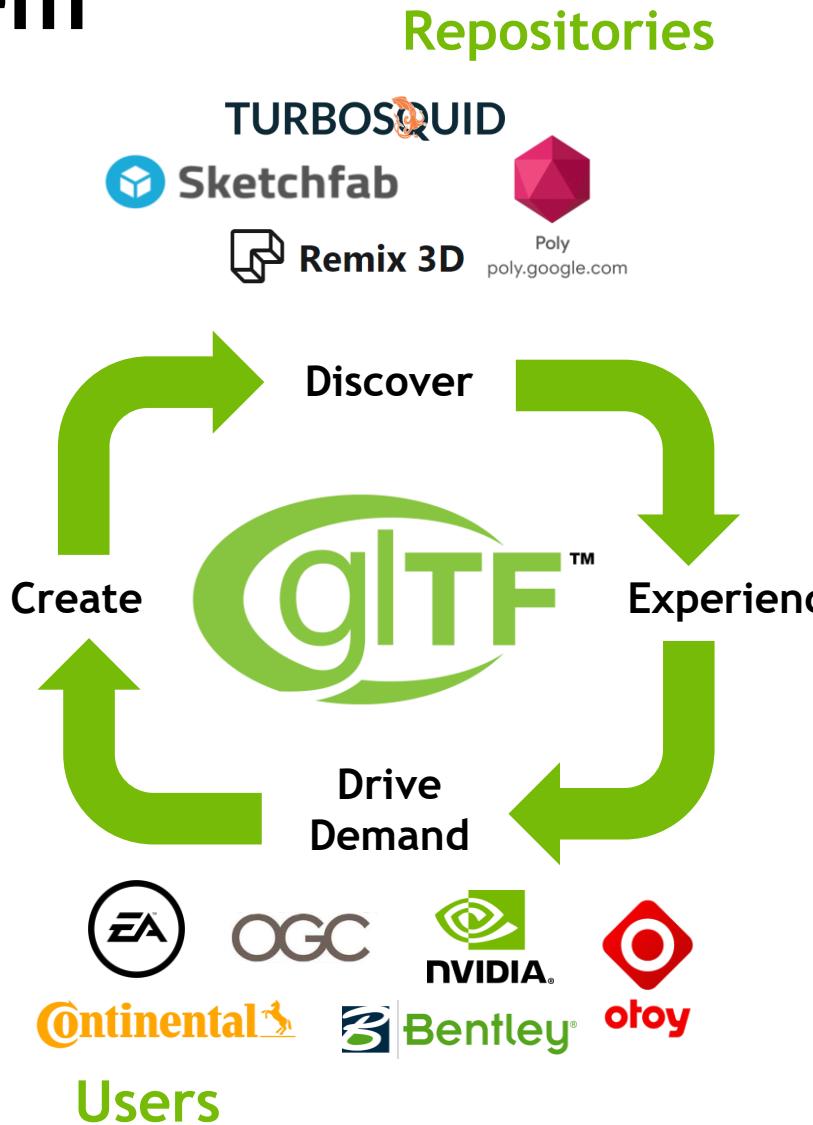
Collada2gltf



FBX2glTF

glTF-validator

glTF-asset-generator



glTF Recent Highlights



Microsoft makes glTF files as pervasively usable as JPGs in Windows 10
Open sources entire glTF SDK! <https://github.com/Microsoft/glTF-SDK>

Adds glTF to StemCell - 60K+ 3D artists and 700K 3D models

<https://www.khronos.org/blog/turbosquid-adds-gltf-to-supported-formats-for-its-stemcell-initiative>



Supports drag and drop of glTF models to your feed
<https://developers.facebook.com/blog/post/2018/02/20/3d-posts-facebook/>

Over 150K royalty-free glTF models

<https://sketchfab.com/features/gltf>



Dimension generates glTF for delivery of 3D marketing assets
https://www.khronos.org/assets/uploads/developers/library/2018-gdc-webgl-and-gltf/glTF-Adobe-Dimension-GDC_Mar18.pdf

Integrating glTF into HUBS Web VR Meeting Space

<https://www.roadtovr.com/mozillas-hubs-one-click-vr-meeting-space-ive-waiting/>



3D Tiles community standard proposal references glTF for streaming 3D geospatial datasets
<http://www.opengeospatial.org/pressroom/pressreleases/2829>

Import of glTF into AR Core apps via the Google Sceneform Tools plugin

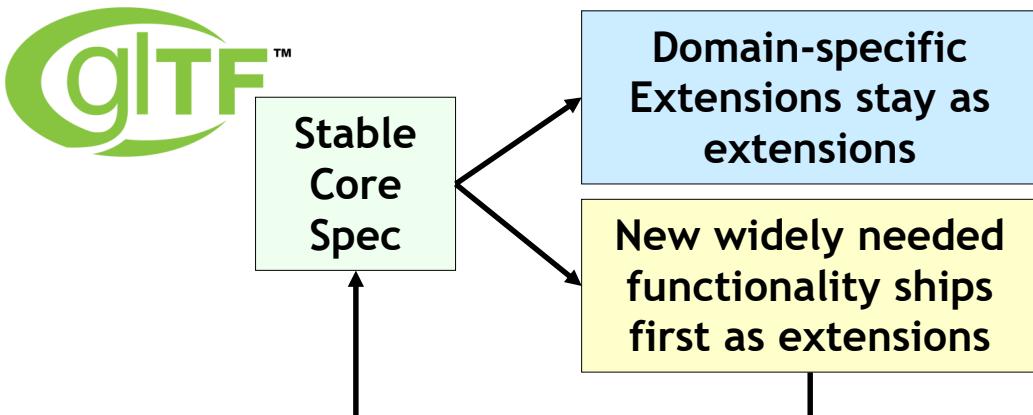
Draco Mesh Compression Technology Contributed as an extension

<https://www.khronos.org/news/press/khronos-announces-gltf-geometry-compression-extension-google-draco>



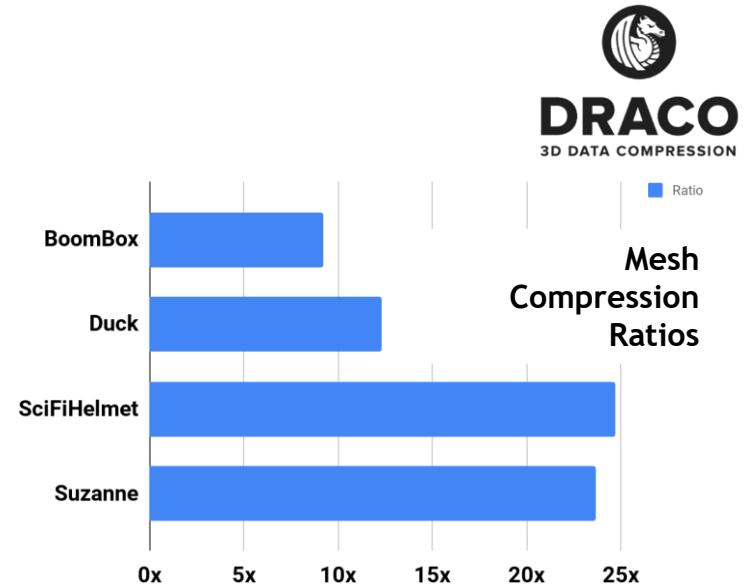
glTF Roadmap

- glTF manages its roadmap very carefully - complexity is the enemy
 - Mission #1: ensure widespread, consistent, reliable usage
- Rollout new functionality first as extensions: e.g. Draco Mesh Compression
 - Open source encoders and decoders are available
 - Extension is now shipping in more and more tools and engines
- New extensions...
 - Texture transform (for texture atlases) has shipped, Unlit Materials is close..

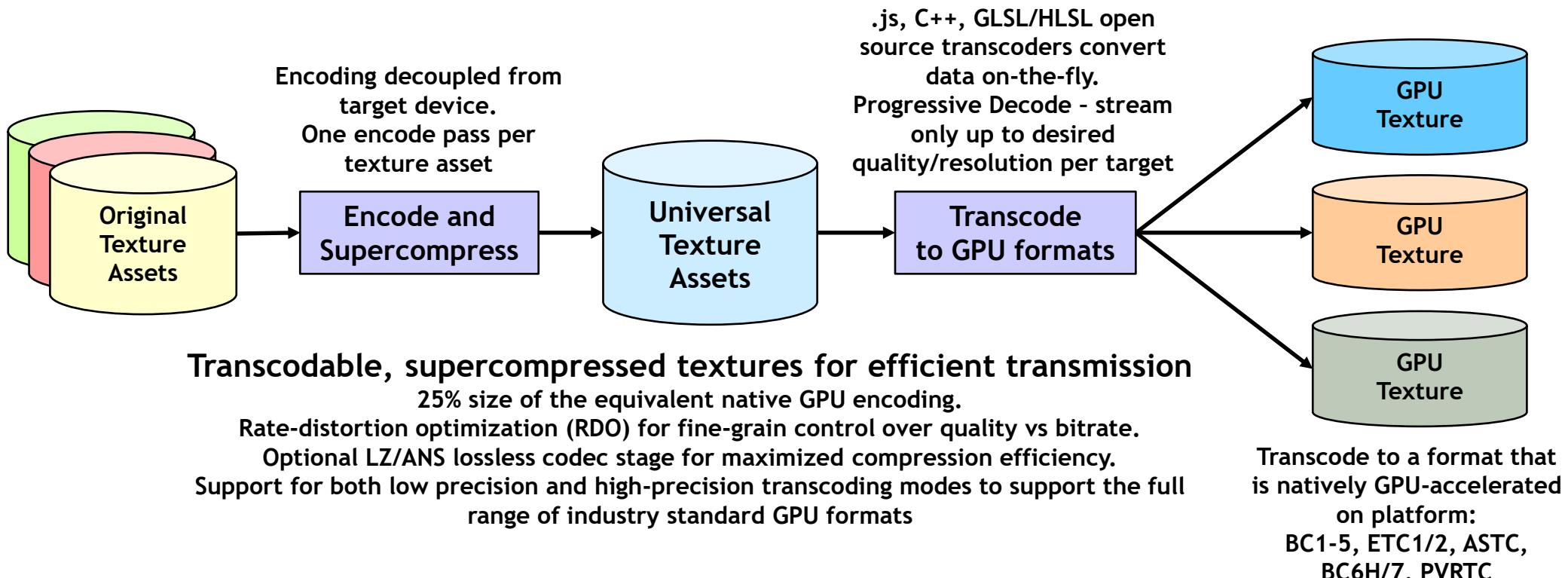


Integrate extensions into new core spec only when:

- 1) Widespread need is confirmed by the industry
- 2) Widespread reliable implementation is enabled (e.g. open source)



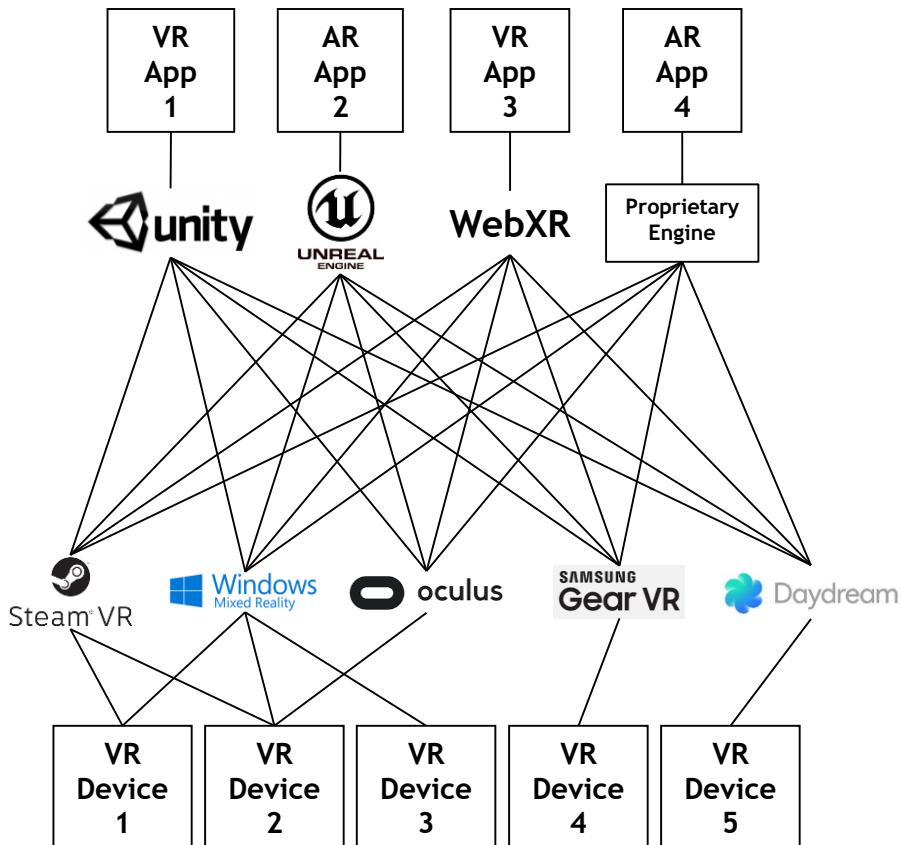
Texture Transmission Extension in Progress



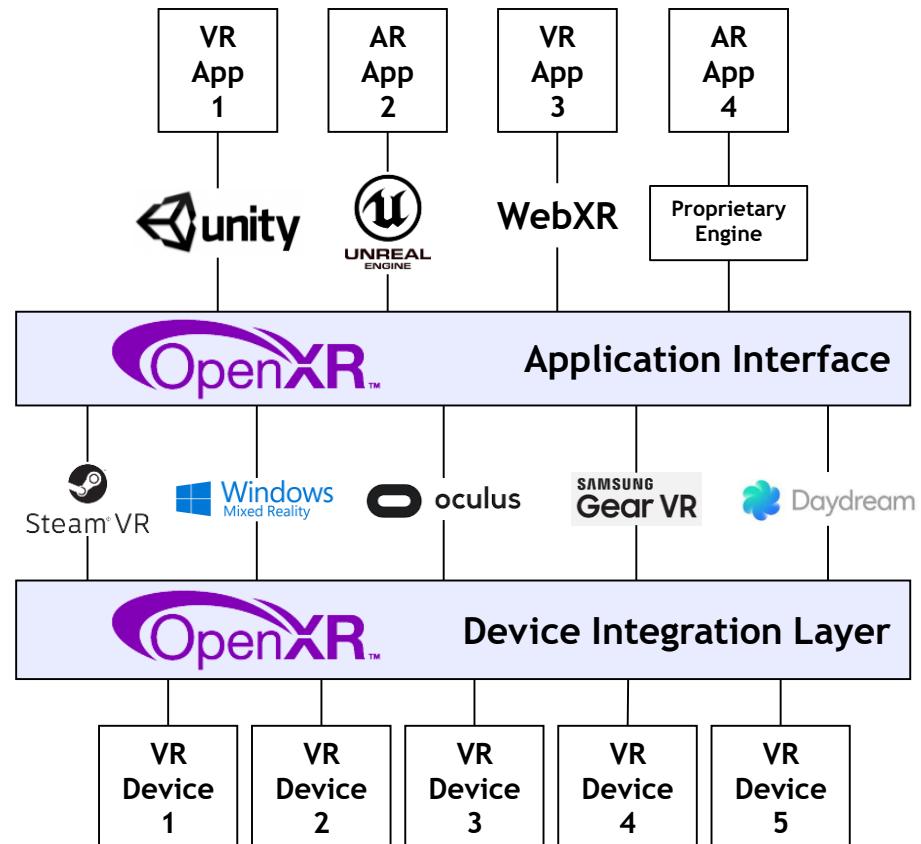
Extension in design - welcome industry feedback

<https://github.com/KhronosGroup/gltf-Texture-Transmission-Tools>

OpenXR - Solving VR/AR Fragmentation



Before OpenXR
XR Market
Fragmentation



After OpenXR
Wide interoperability of
XR apps and devices

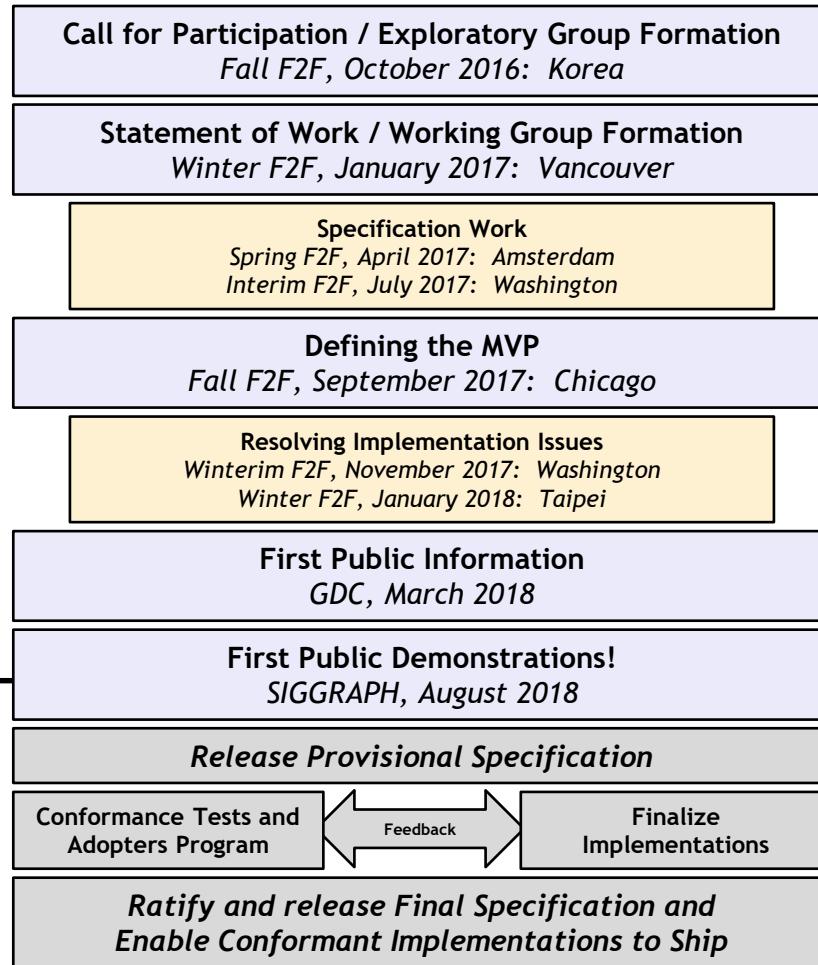
Companies Publicly Supporting OpenXR



OpenXR is a collaborative design

- 1) For cross-platform XR portability - VR in V1.0, then add AR
- 2) Integrating many lessons from proprietary 'first-generation' API designs

OpenXR Development Process



Present Day
Coming Soon



Much more detailed specification
overview and GDC session videos:
<https://www.khronos.org/developers/library/2018-gdc>

Multiple implementations Underway!
Final specification will incorporate
implementation feedback

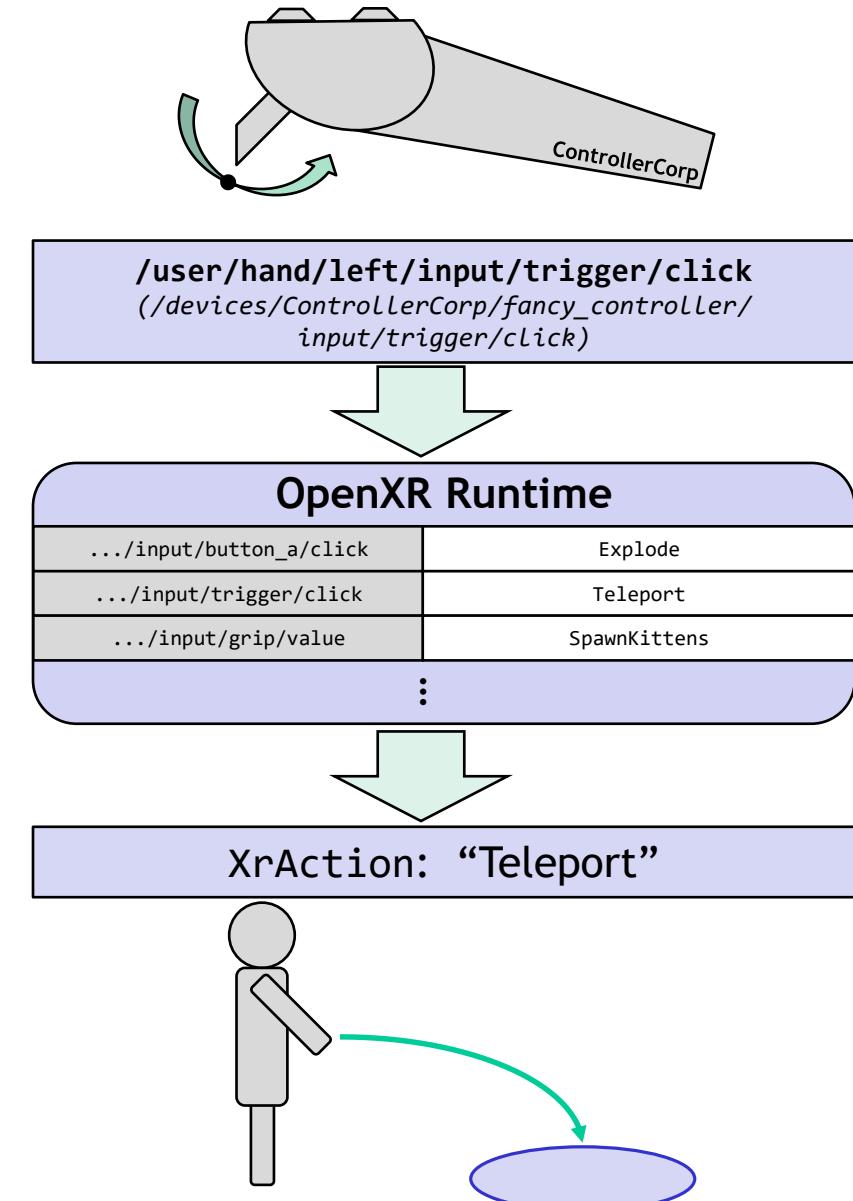
Epic ‘Showdown’ VR Demo at SIGGRAPH



Demo runs portably across StarVR and Microsoft Windows Mixed Reality headsets through the OpenXR APIs via an Unreal Engine 4 plugin

Input and Haptics

- **Input uses abstracted Input Actions**
 - E.g. “Move,” “Jump,” “Teleport”
- **Many advantages**
 - Existing content can easily use new devices
 - Mix-and-match multiple input sources to create a unified UI
 - Easy optional feature support (e.g. eye and body tracking)
 - Future-proofing for innovation in input devices and form factors



OpenXR Viewport Configurations

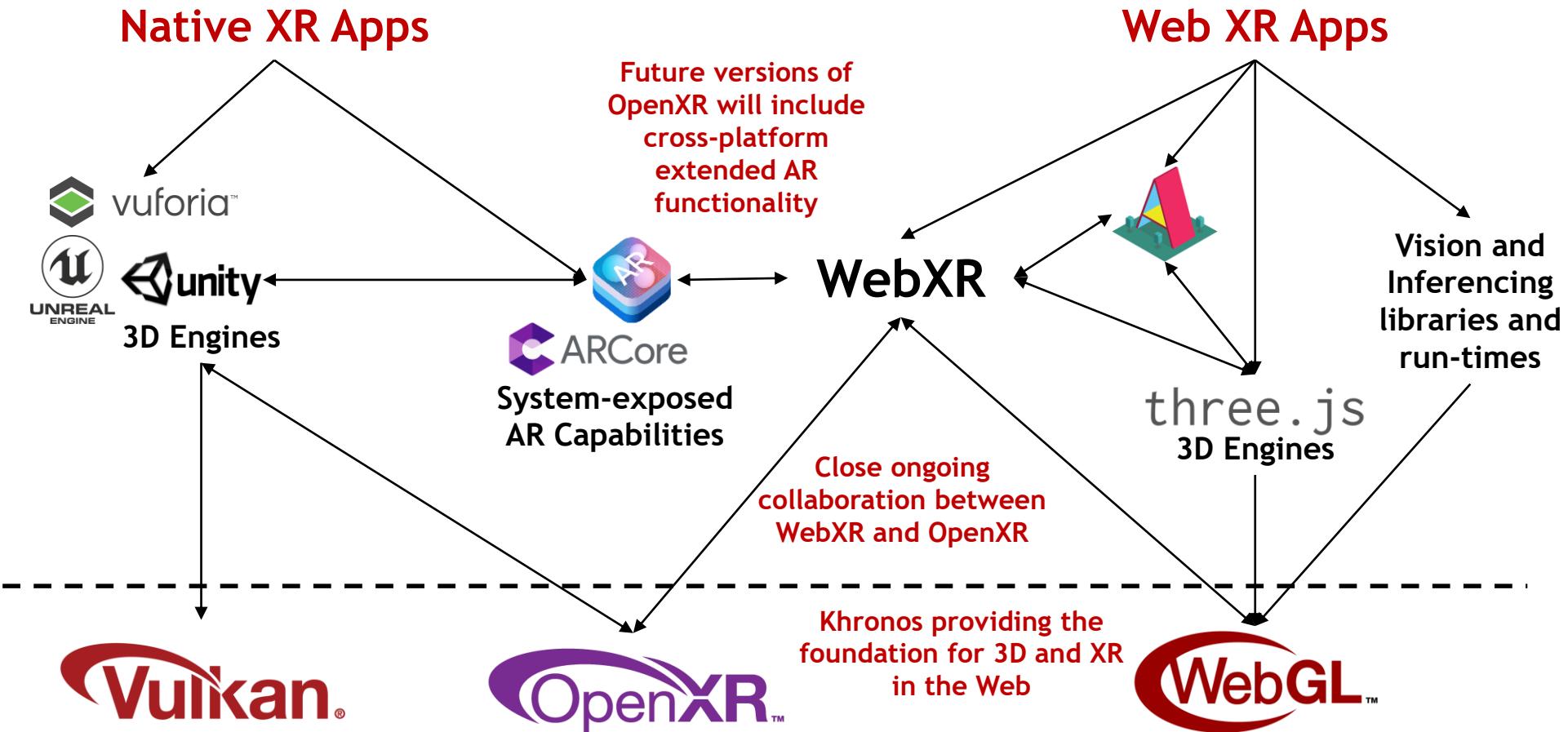
- Applications can:
 - Query for runtime supported Viewport Configurations
 - Applications can then set the Viewport Configurations that they plan to use
 - Select and change their active configuration over the lifetime of the session



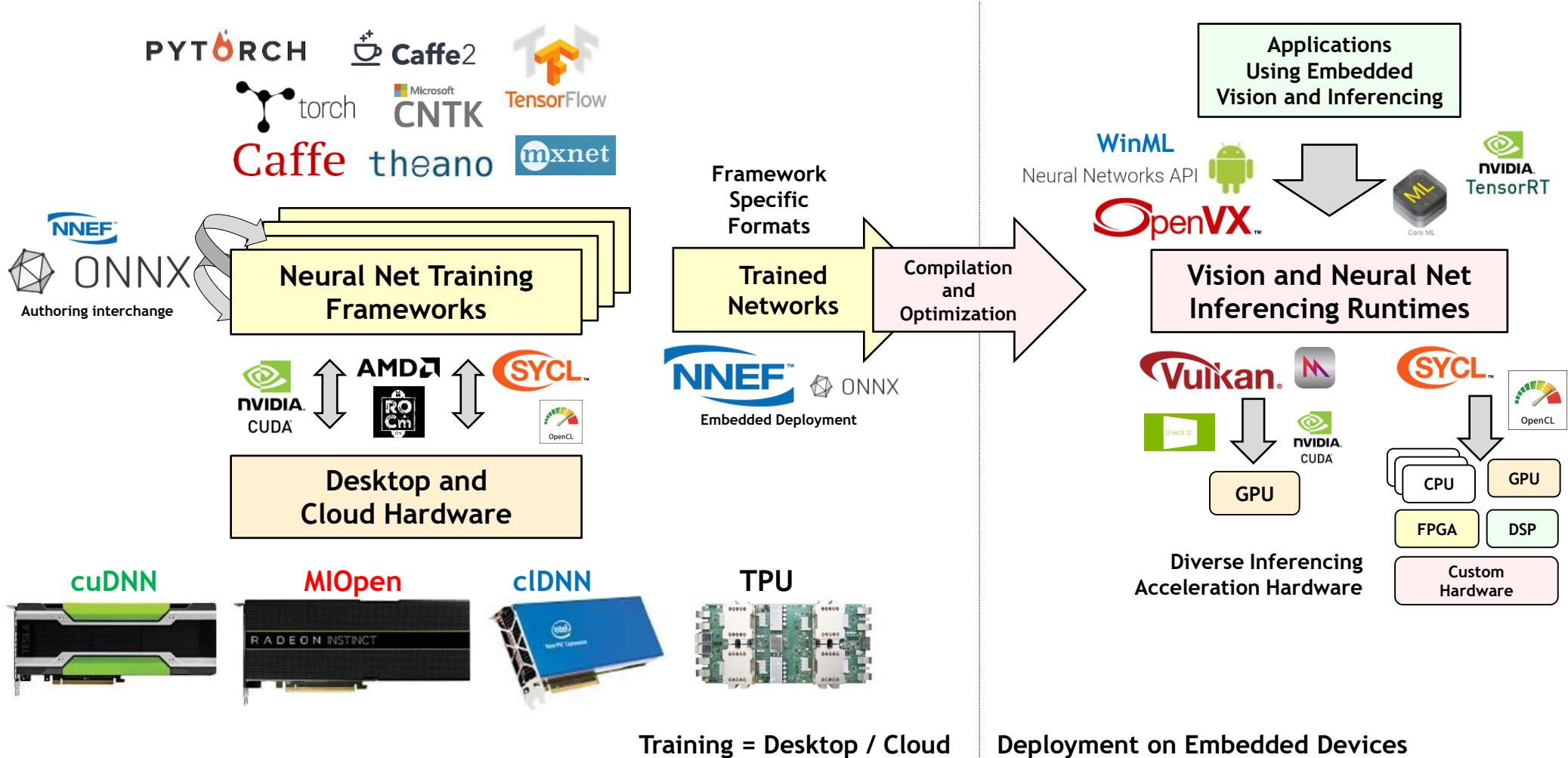
Photo Credit: Dave Pape

One Viewport	Two Viewports (one per eye)	Twelve Viewports (six per eye)
/viewport_configuration/ar_mono/magic_window	/viewport_configuration/vr/hmd	/viewport_configuration/vr_cube/cave_vr

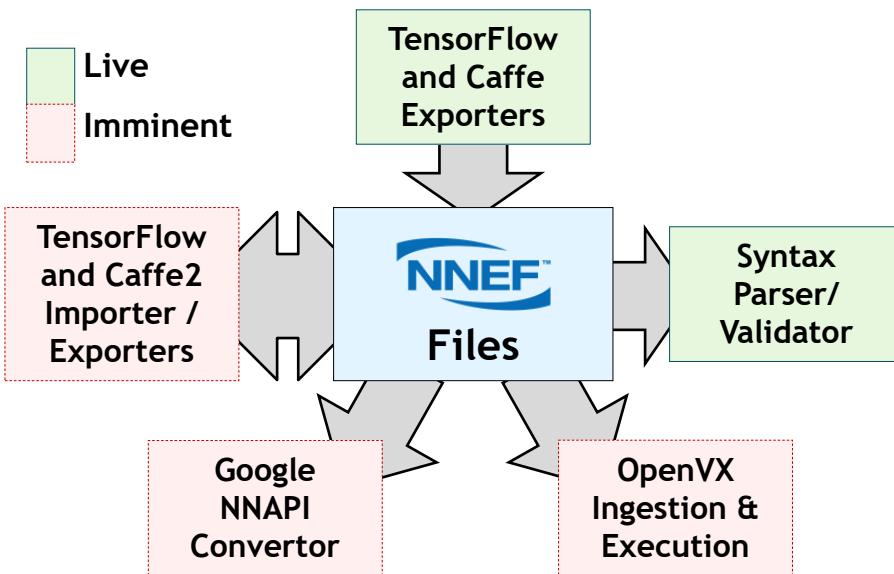
Layered XR Ecosystems



Neural Network Workflow



NNEF Ecosystem



NNEF open source projects hosted on
Khronos NNEF GitHub repository
Apache 2.0 license
<https://github.com/KhronosGroup/NNEF-Tools>

NNEF = Neural Network Exchange Format

NNEF V1.0 released at SIGGRAPH!!
After positive industry feedback on Provisional specification
released in December 2017

Comparing Neural Network Exchange Industry Initiatives



Defined Specification

Stability for hardware deployment

Multi-company Governance

Flexible Precision

Flat and Compound Ops



Open Source Project

Software stack flexibility

Initiated by Facebook

32-bit Floating Point only

Flat Ops Only

OpenVX

Wide range of vision hardware architectures
OpenVX provides a high-level Graph-based abstraction

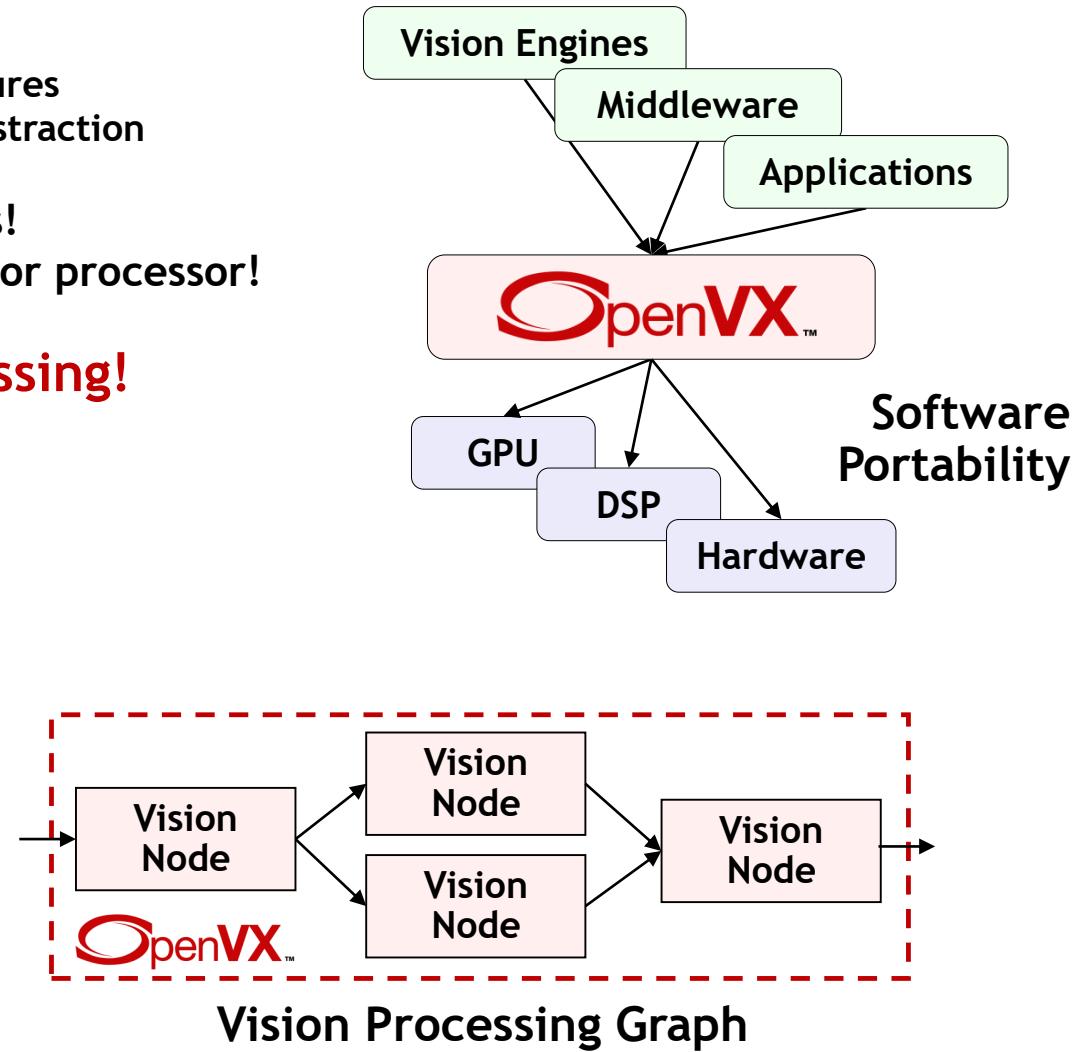
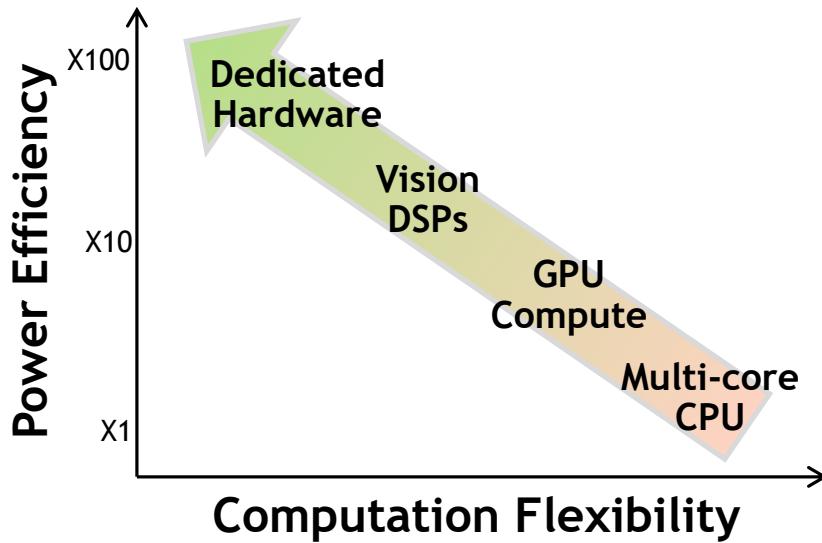
->

Enables Graph-level optimizations!

Can be implemented on almost any hardware or processor!

->

Portable, Efficient Vision Processing!



OpenVX Evolution

Conformant Implementations



SOCIONEXT



OpenVX 1.0

Spec released October 2014

Conformant Implementations

cadence



SYNOPSYS

TEXAS INSTRUMENTS

New Functionality

Expanded Nodes Functionality
Enhanced Graph Framework

AMD OpenVX Tools

- Open source, highly optimized for x86 CPU and OpenCL for GPU
 - “Graph Optimizer”
- Scripting for rapid prototyping, without re-compiling, at production performance levels

<http://gpuopen.com/compute-product/amd-openvx/>

OpenVX 1.1

Spec released May 2016



New Functionality

Conditional node execution
Feature detection
Classification operators
Expanded imaging operations

Extensions

Neural Network Acceleration
Graph Save and Restore
16-bit image operation

Safety Critical

OpenVX 1.1 SC for
safety-certifiable systems

OpenVX

Roadmap under Development

NNEF Import

Programmable user
kernels with
accelerator offload

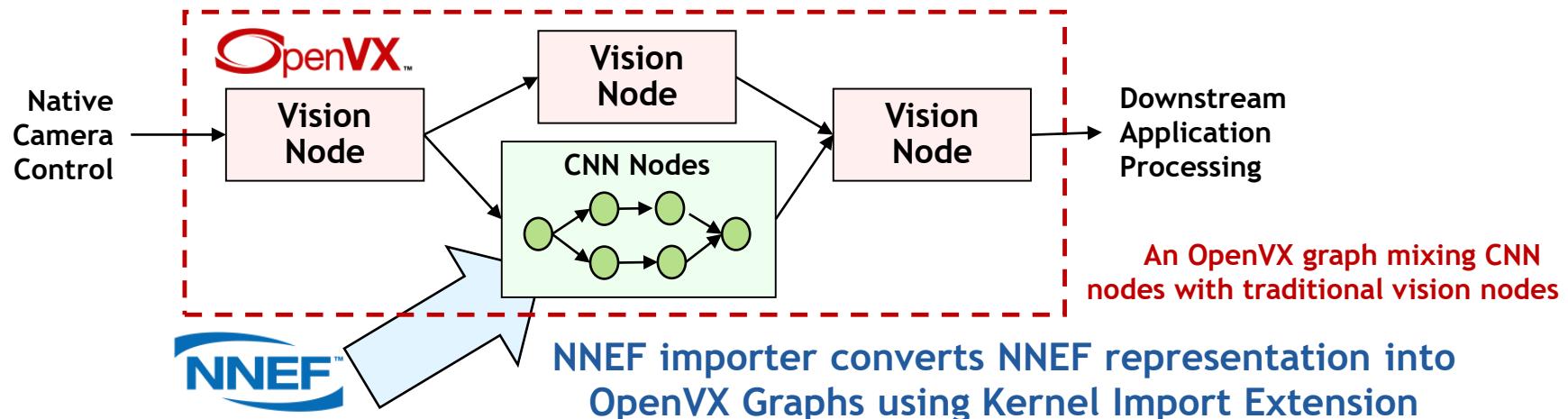
OpenVX 1.2

Spec released May 2017

Adopters Program November 2017

OpenVX NNEF Import

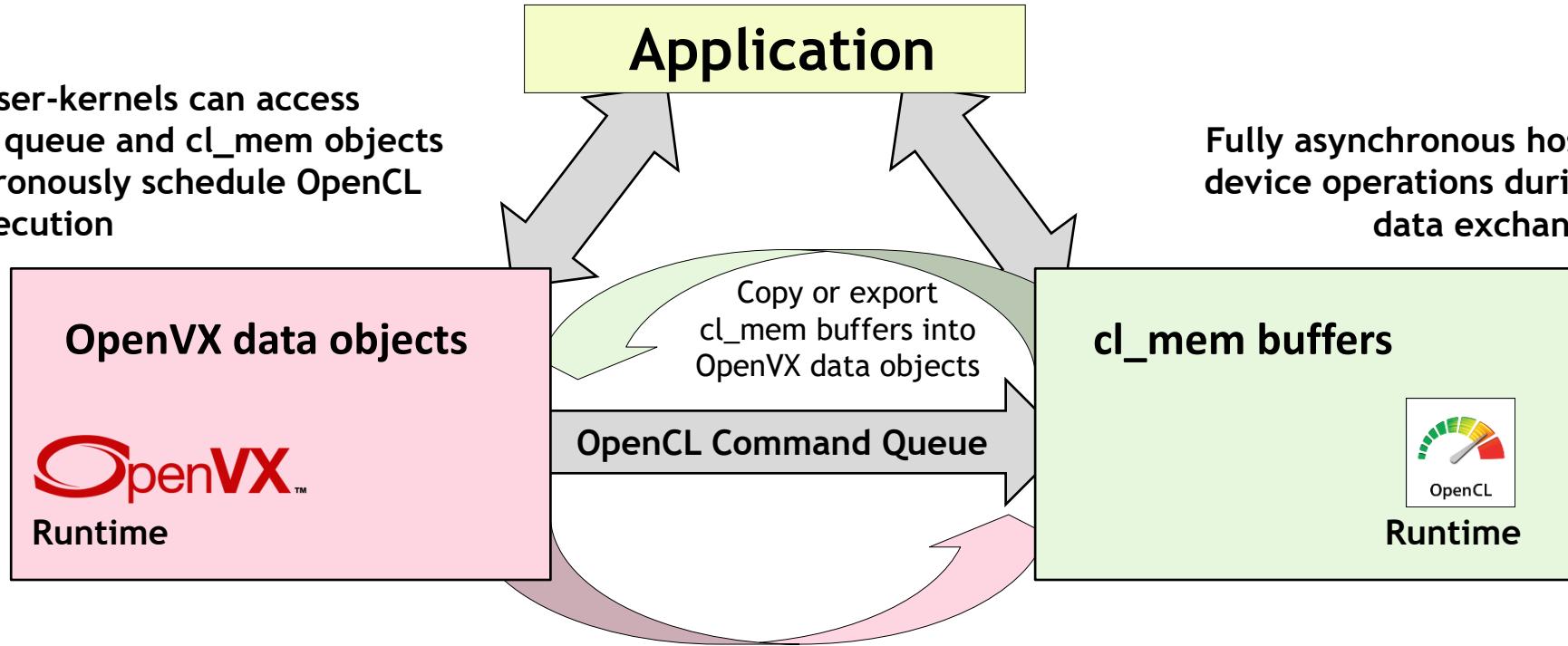
- Convolution Neural Network topologies can be represented as OpenVX graphs
 - Can also combine traditional vision and neural network operations
- OpenVX Neural Network Extension
 - Defines OpenVX nodes to represent many common NN layer types
 - Layer types include convolution, activation, pooling, fully-connected, soft-max
 - Defines multi-dimensional tensors objects to connect layers
- Kernel Import Extension
 - Enables loading of external program representations into OpenVX graphs



OpenVX / OpenCL Interop Extension

OpenVX user-kernels can access command queue and `cl_mem` objects to asynchronously schedule OpenCL kernel execution

Fully asynchronous host-device operations during data exchange

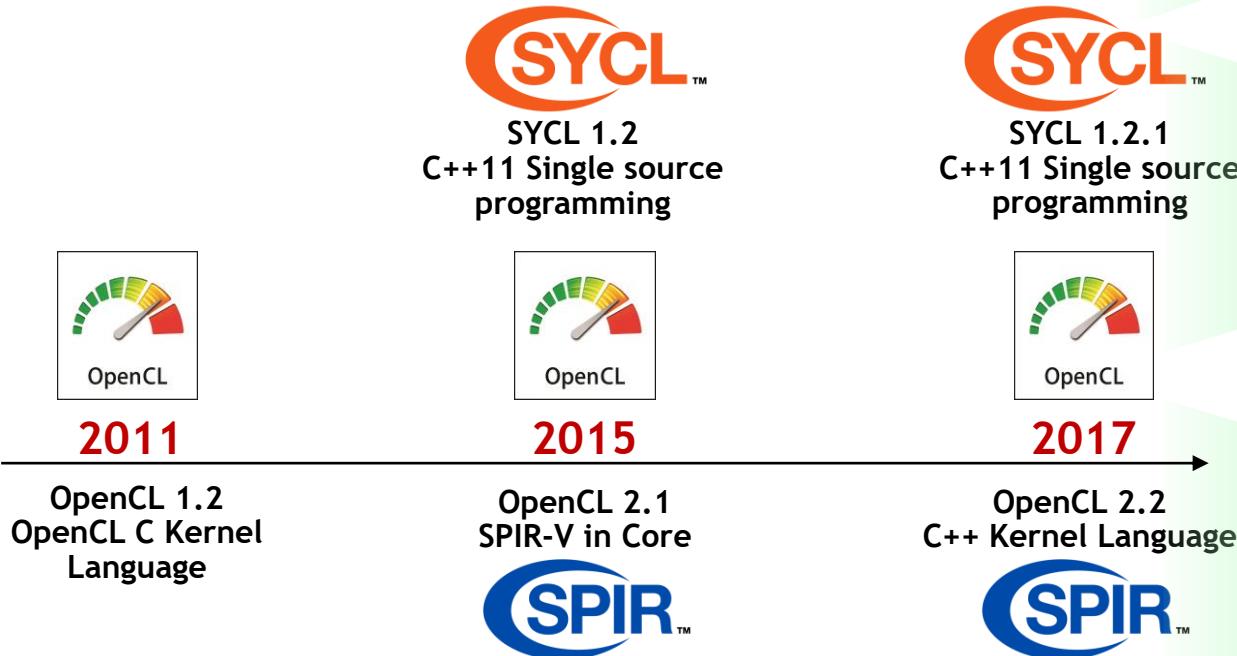


Enables custom OpenCL acceleration to be used within OpenVX User Kernels

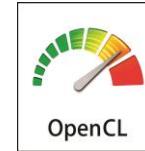
OpenCL Ecosystem Roadmap



OpenCL has an active three track roadmap



Bringing Heterogeneous compute to standard ISO C++
Khronos hosting C++17 Parallel STL
C++20 Parallel STL with Ranges Proposal



Processor Deployment Flexibility

Parallel computation across diverse processor architectures

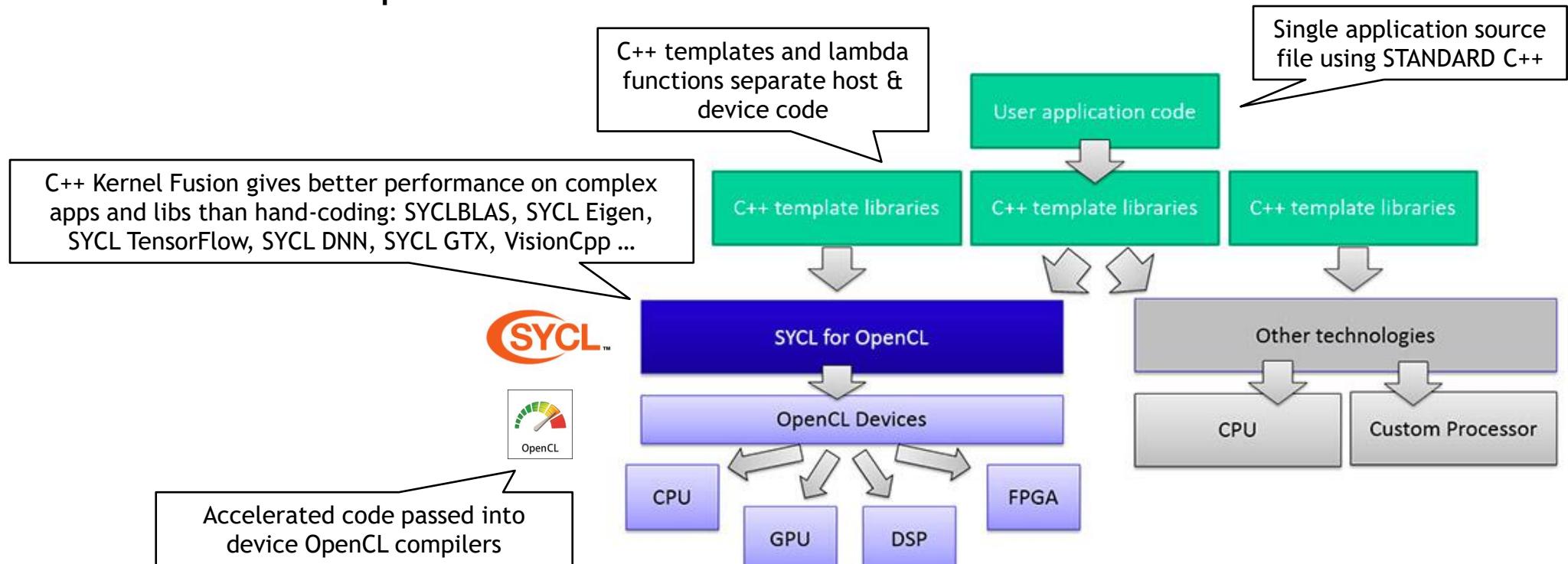


Kernel Deployment Flexibility

Execute OpenCL C kernels on Vulkan GPU runtimes

SYCL Single Source C++ Parallel Programming

- SYCL 1.2.1 Adopters Program released in July 2018 with open source conformance tests
 - <https://www.khronos.org/news/press/khronos-releases-conformance-test-suite-for-sycl-1.2.1>
- Multiple Implementations shipping: triSYCL, ComputeCpp, ComputeCpp SDK
 - <http://sycl.tech>
- SYCL 2.2 Provisional Specification Released

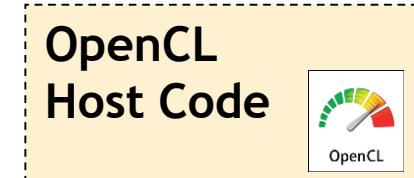
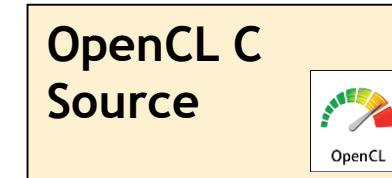


Bringing OpenCL Compute to Vulkan



- Experimental Clspv Compiler from Google, Adobe and Codeplay
 - Compiles OpenCL C to Vulkan's SPIR-V execution environment
Successfully tested on over 200K lines of Adobe OpenCL C production code
 - Open source - tracks top-of-tree LLVM and clang, not a fork

Layering and cross-compilation is making Vulkan a Universally Portable API for Graphics AND Compute



Prototype open source project

<https://github.com/google/clspv>

Clspv Compiler



Run-time API Translator



Possible future project - if interest

Increasing deployment options for OpenCL kernel developers
e.g. Vulkan is a supported API on Android

Refining clspv with Diverse Workloads

Public test kernel repository so
developers can upload kernels to be used
in long-term perf/regression testing
<https://github.com/KhronosGroup/OpenCL-Sample-Kernels>

Interesting domains to explore include
existing OpenCL compute libraries -
including vision and inferencing

Vulkan is already expanding its compute
model e.g. 16-bit storage, Variable
Pointers, Subgroups. Compact memory
types and operations coming

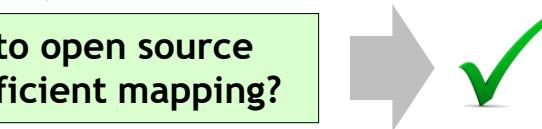
Compile diverse OpenCL C
kernel workloads



Efficient mapping to
Vulkan SPIR-V?



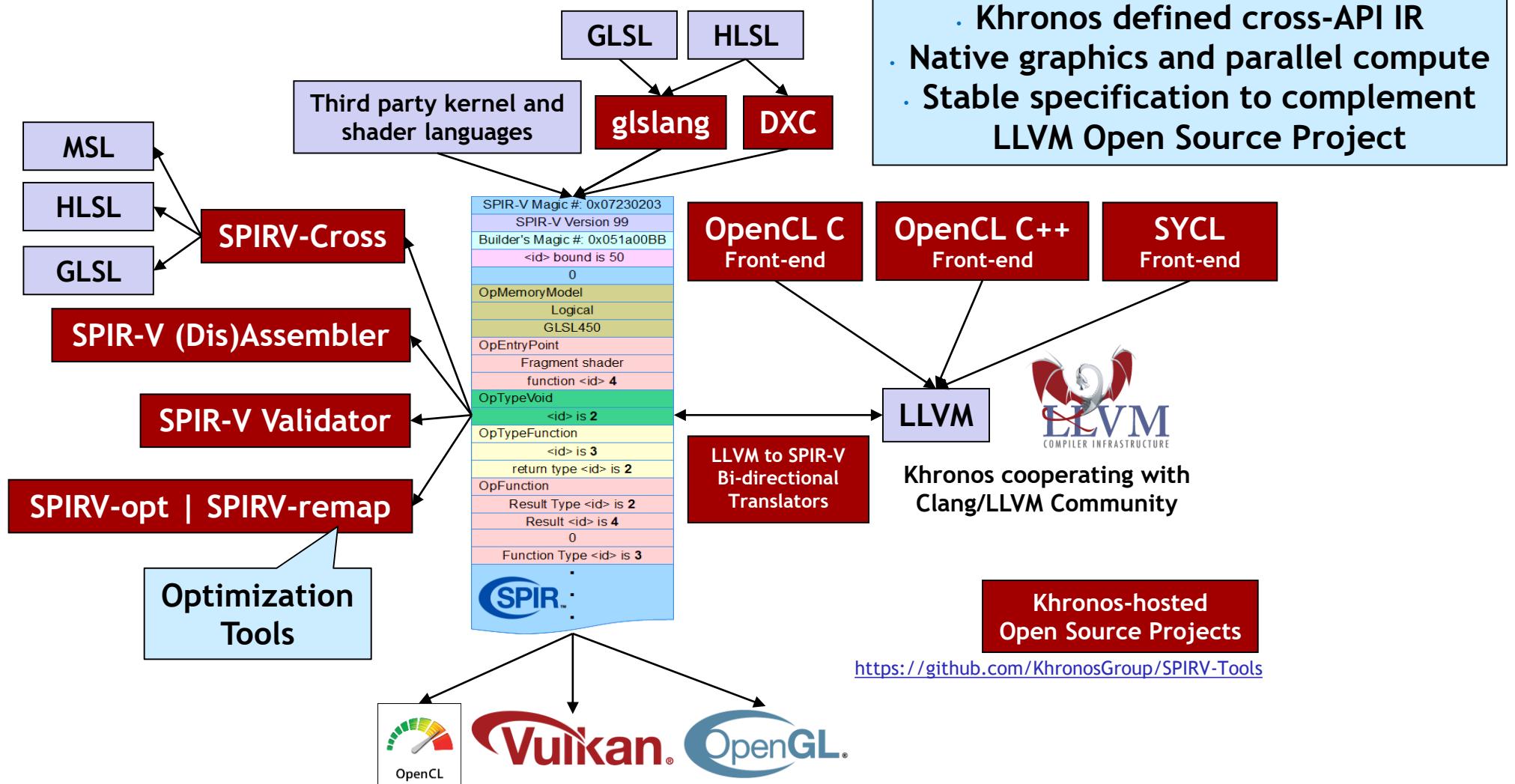
Updates to open source
achieve efficient mapping?



Propose updates to Vulkan
programming model

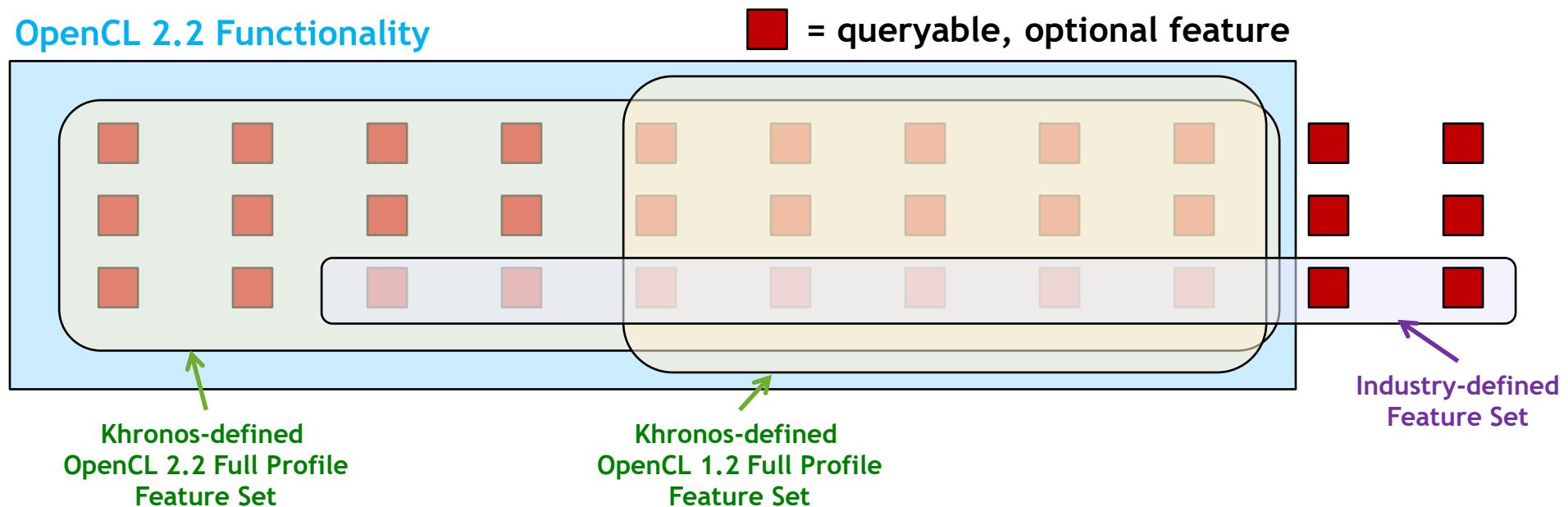


SPIR-V Ecosystem



OpenCL Next - Feature Set Flexibility

- Defining OpenCL features that become optional for enhanced deployment flexibility
 - API and language features e.g. floating point precisions
- Feature Sets avoid fragmentation
 - Defined to suit specific markets - e.g. desktop, embedded vision and inferencing
- Implementations are conformant if fully support feature set functionality



Khronos Education Forum

- Supporting educators to teach Khronos technologies
 - A common platform on which to openly share and discuss course materials
 - Khronos will manage an Open Educational Resource hub
 - <https://www.khronos.org/education>
- Feedback on materials from the related Khronos Working Groups
 - Direct contact with the designers of the standards
- Any educator is welcome to get involved!
 - Just send an email to education@khronos.org
 - Name, institution, and the Khronos standards you want to teach





Come Back at **5:30** for the
NETWORKING RECEPTION

Sponsored by:



NVIDIA®



[WWW.KHRONOS.ORG](http://www.khronos.org)