



OpenCL Working Group Update

Neil Trevett

Khronos President

OpenCL Working Group Chair

NVIDIA VP Developer Ecosystems

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



Khronos Connects Software to Silicon



KHRONOS
GROUP

Open, royalty-free interoperability standards to harness the power of GPU, XR and multiprocessor hardware

3D graphics, augmented and virtual reality, parallel programming, inferencing and vision acceleration

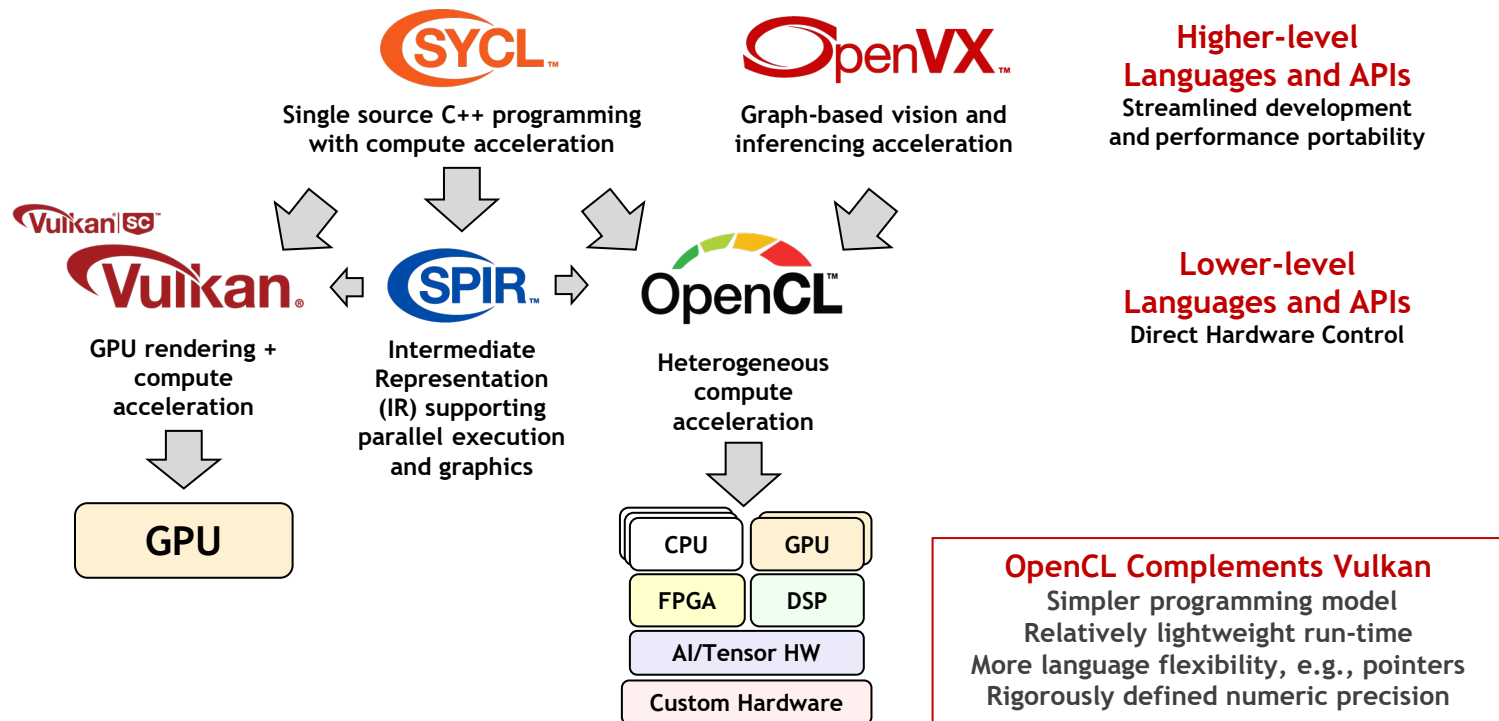
Non-profit, member-driven standards organization, open to any company

Proven multi-company governance and Intellectual Property Framework

Founded in 2000

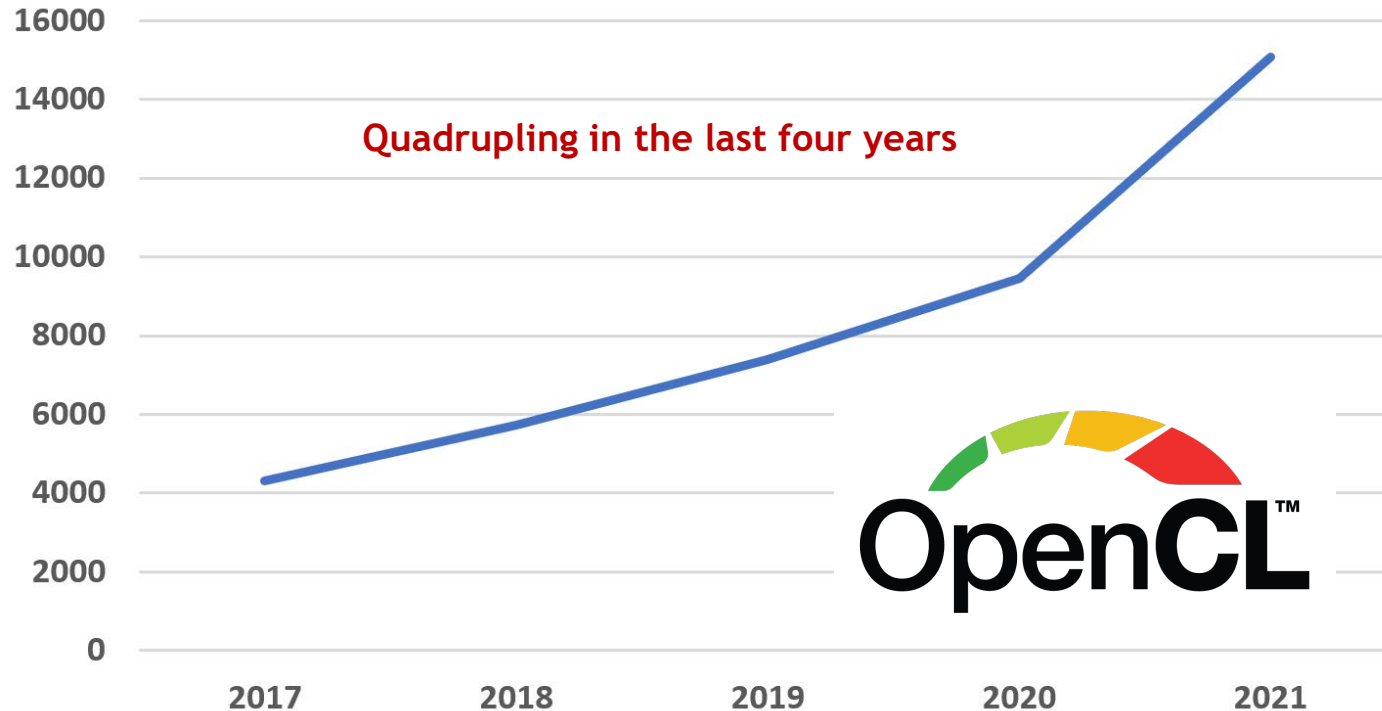
~ 200 Members | ~ 40% US, 30% Europe, 30% Asia

Khronos Compute Acceleration Standards



OpenCL Open-Source Project Momentum

OpenCL-based GitHib Repos



OpenCL 3.0 Released September 2020

Increased Ecosystem Flexibility

All functionality beyond OpenCL 1.2 is queryable
Macros for optional OpenCL C language features
Widely adopted new extensions are integrated into new core specifications

OpenCL C++ for OpenCL

Open-source C++ for OpenCL front end compiler
Combines OpenCL C and C++17
Replaces OpenCL C++ language specification

Unified Specification

All versions of OpenCL in one specification
Easier maintenance, evolution and accessibility
Specification Source on Khronos GitHub for requests, feedback and bugs

Moving Applications to OpenCL 3.0

OpenCL 1.2 applications - no change
OpenCL 2.X applications - no code changes if all used functionality is present
Queries recommended for future portability

A Common Baseline for Ecosystem Evolution



OpenCL 3.0 Adoption



Product Conformance Status

<https://www.khronos.org/conformance/adopters/conformant-products/opencl>

OpenCL 3.0 Adopters



OpenCL 3.0 Adopters
Shipping Conformant Implementations

Apps, Libraries and Engines using OpenCL

The industry's most pervasive, cross-vendor, open standard for
low-level heterogeneous parallel programming

https://en.wikipedia.org/wiki/List_of_OpenCL_applications

Desktop Creative Apps



Parallel Languages



Machine Learning Libraries and Frameworks



Molecular Modelling Libraries



Machine Learning Compilers



Vision, Imaging and Video Libraries



Math and Physics Libraries



Linear Algebra Libraries



OpenCL and Machine Learning

Machine Learning
Compilers



Import Formats

Caffe, Keras,
MXNet, ONNX

TensorFlow Graph,
MXNet, PaddlePaddle,
Keras, ONNX

PyTorch, ONNX

TensorFlow Graph,
PyTorch, ONNX

Front-end / IR

NNVM / Relay IR

nGraph / Stripe IR

Glow Core / Glow IR

XLA HLO

Output

OpenCL, LLVM,
CUDA, Metal

OpenCL,
LLVM, CUDA

OpenCL
LLVM

LLVM, TPU IR, XLA IR
TensorFlow Lite / NNAPI
(inc. HW accel)



Common Steps

1.Import Trained
Network Description

2. Graph-level optimizations
e.g., node fusion, node
lowering and memory tiling

3. Decompose to primitive
instructions and emit programs
for accelerated run-times

Machine Learning Compilers and Frameworks using OpenCL Acceleration

Inferencing Libraries and Frameworks

Alibaba MNN
Arm Compute Library
Baidu PaddlePaddle/Paddle-Lite
Caffe
Intel cIDNN and OpenVINO

Google TensorFlow and NNAPI

SYCL-DNN
Synopsis MetaWare EV
Texas Instruments DL Library (TIDL)
VeriSilicon Acuity
Xiaomi Mace

Embedded NN Compilers

CEVA Deep Neural Network (CDNN)
Cadence Xtensa
Neural Network Compiler (XNNC)



C++ for OpenCL

Open-Source Compiler Front-end

Replaces the OpenCL C++ kernel language spec

[Official release](#) published in OpenCL-Docs repo

Enables full OpenCL C and most C++17 capabilities

OpenCL C code is valid and fully compatible

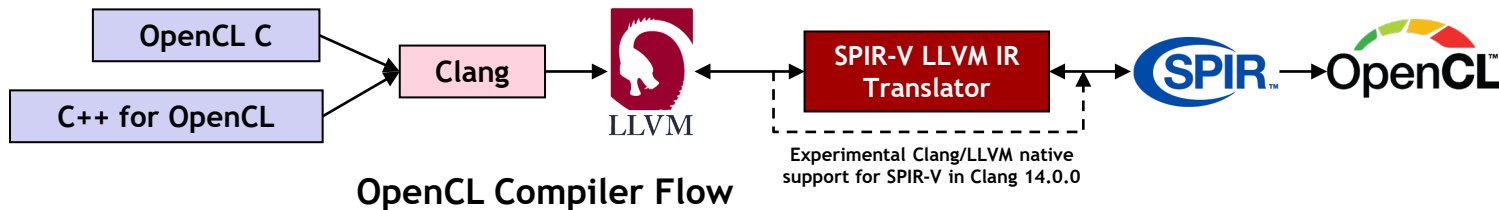
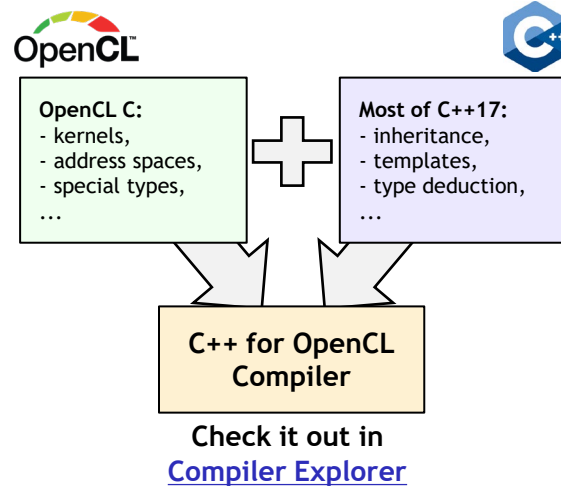
Enables gradual transition to C++ for existing apps

Supported in Clang since release 9.0

Generates SPIR-V 1.0 plus SPIR-V 1.2 where necessary

Online compilation via [cl_ext_cxx_for_opengl](#) extension

New! C++ for OpenCL 2021 provisional documentation and experimental native support in [Clang 14.0.0](#)

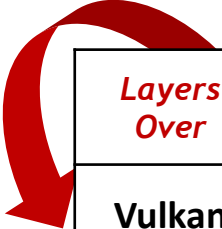


KHRONOS[®] GROUP



API Layering

Enabled by growing robustness of open-source compiler ecosystem using SPIR-V



<i>Layers Over</i>	Vulkan	OpenGL	OpenCL	OpenGL ES	DX12	DX9-11
Vulkan		Zink	clspv clvk	GLOVE Angle	vkd3d-Proton vkd3d	DXVK WineD3D
OpenGL	gfx-rs Ashes			Angle		WineD3D
DX12	Dozen gfx-rs	Microsoft 'GLOn12'	Microsoft 'CLOn12'			Microsoft D3D11On12
DX9-11	gfx-rs Ashes			Angle		
Metal	MoltenVK gfx-rs			MoltenGL Angle		

ROWS Benefit Platforms by adding APIs

COLUMNS Benefit ISVs by making an API available everywhere

Layered OpenCL Implementations

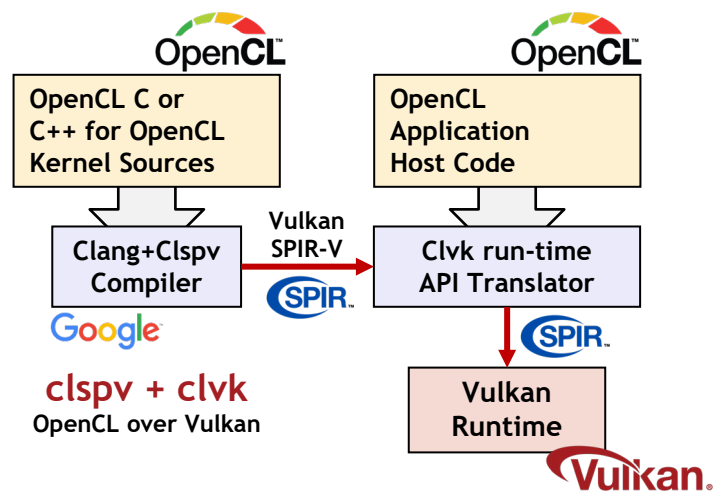
clspv + clvk

clspv - Google's open-source OpenCL kernel to Vulkan SPIR-V compiler

Tracks top-of-tree LLVM and Clang - not a fork

Clvk - prototype open-source OpenCL to Vulkan run-time API translator

Used by shipping apps and engines on Android
e.g., Adobe Premiere Rush video editor - 200K lines of OpenCL C kernel code



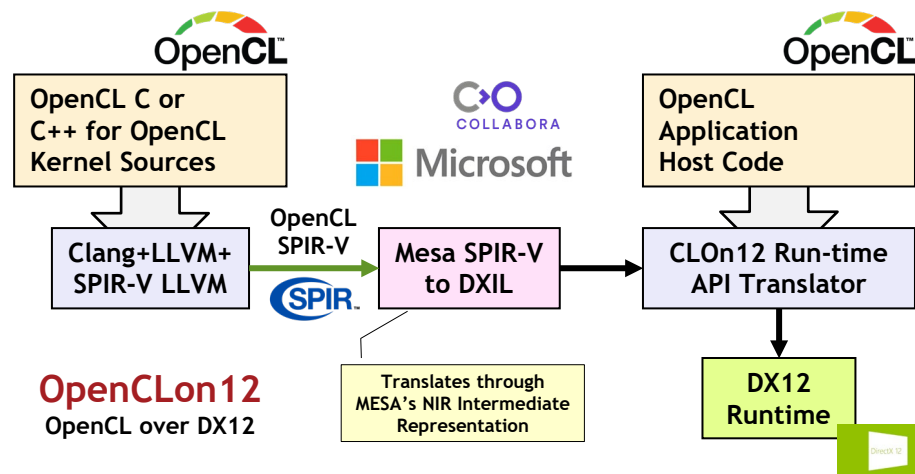
OpenCLon12

Microsoft and COLLABORA

GPU-accelerated OpenCL on any DX12 PC and Cloud instance (x86 or Arm)

Leverages Clang/LLVM AND MESA

OpenCLon12 - OpenGL 3.3 over DX12 is already conformant



Regular OpenCL Specification Releases

OpenCL 3.0.11 shipped on May 6th, 2022

Continues the regular release cadence for new functionality and bug fixes

New OpenCL extensions shipped since IWOCL 2021

Subgroup rotate extension for efficient data exchange among work-items

Workgroup Uniform Arithmetic for new work-group scan and reduction operators

Command Buffers Record and Replay (provisional)

Asynchronous DMA

Expect Assume Hints

Enhanced subgroup functionality

Extended bit-level operations

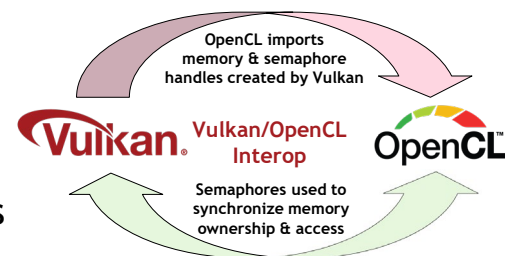
Universally unique device identifier query

Enhanced queries for platform and device versions

SPIR-V support for C++ linkage types

Integer Dot Product for Faster Neural Network Inferencing

External memory objects and semaphores for external sharing and Interop (provisional)



OpenCL Added to GPUInfo.org



Name	Coverage
CL_DEVICE_MAX_SAMPLERS	100%
CL_DEVICE_MAX_WORK_GROUP_SIZE	100%
CL_DEVICE_MAX_WORK_ITEM_DIMENSIONS	100%
CL_DEVICE_MAX_WORK_ITEM_SIZES	100%
CL_DEVICE_MAX_WRITE_IMAGE_ARGS	100%
CL_DEVICE_MEM_BASE_ADDR_ALIGN	100%
CL_DEVICE_MIN_DATA_TYPE_ALIGN_SIZE	100%
CL_DEVICE_NAME	100%
CL_DEVICE_PREFERRED_VECTOR_WIDTH_CHAR	100%
CL_DEVICE_PREFERRED_VECTOR_WIDTH_DOUBLE	100%
CL_DEVICE_PREFERRED_VECTOR_WIDTH_FLOAT	100%
CL_DEVICE_PREFERRED_VECTOR_WIDTH_INT	100%
CL_DEVICE_PREFERRED_VECTOR_WIDTH_LONG	100%

Device	Max. API version	Latest Driver version	Last submission	Count
NVIDIA GeForce RTX 3080 Laptop GPU	3.0	515.66	2022-04-03 18:13:50	15
NVIDIA TITAN V	3.0	515.66	2022-05-01 22:32:25	11
Intel(R) HD Graphics 530	3.0	31.0.101.1956	2022-05-02 02:58:31	9
AMD Radeon Pro WX 5100 Graphics (POLA...	1.1	21.3.7	2022-03-14 17:20:42	9
Intel(R) UHD Graphics 620 [0x3ea0]	3.0	22.17.23034	2022-05-02 19:39:18	8
NVIDIA GeForce RTX 2060	3.0	511.79	2022-03-31 21:53:58	8
Intel(R) UHD Graphics 630	3.0	30.0.101.1191	2022-01-19 17:37:11	7
NVIDIA GeForce MX110	3.0	512.59	2022-05-02 19:39:08	7
NVIDIA GeForce RTX 3090	3.0	512.59	2022-05-01 01:31:41	6
AMD Radeon Pro WX 5100 Graphics (pola...	1.1	22.0.1	2022-04-14 10:01:58	6
Intel(R) HD Graphics 520	3.0	30.0.101.1660	2022-04-03 05:40:55	6
Intel(R) Iris(R) Xe Graphics	3.0	30.0.101.1934	2022-04-27 03:22:55	5



The online GPUInfo.org database is populated using the [OpenCL Hardware Capability Viewer](#) application
Available for Windows, Linux and Android
Reads and displays OpenCL information and uploads to the database
Please download and run to help populate the database!

OpenCL SDK Upgrades

Open-source OpenCL SDK includes all components to develop OpenCL applications

OpenCL Headers (include/api)
OpenCL C++ bindings (include/cpp)
OpenCL Utility Libraries (include/utls)
Build system and CI

Documentation and Sample Code

OpenCL Guide
Code samples (samples/)
Documentation (docs/)

Loader and Layers
SDK and Layers Tutorial

Khronos funds SDK upgrades
Community contributions also welcome!



Spring 2022 SDK Updates

More details in the [SDK Blog](#)

Enhanced Cmake-based build system
Subprojects and components

Binary releases
Tagged SDK versions

Enhanced SDK documentation
In OpenCL Guide

OpenCL 3.0 Samples
C, C++, Python and Ruby

Utility Libraries
For loading kernel source and binary files

What's Coming!

Upstream to Kitware's FindOpenCL.cmake
Enhances OpenCL:: namespace

Packaging and Distribution Support
Build packages from the SDK
Package newer versions of OpenCL
Ease cross-platform installation, including PPAs

Enhanced SDK Validation Layers
Object lifetime, Input parameters, SPIR-V

OpenCL Roadmap



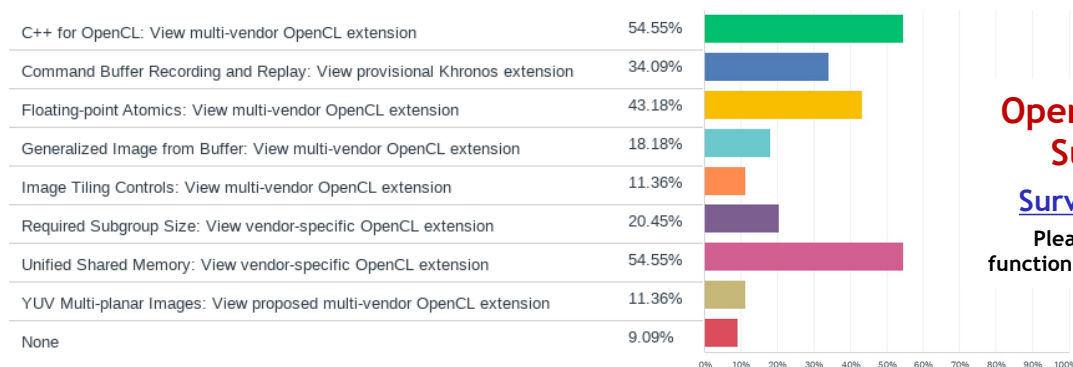
OpenCL Extension Pipeline

Provisional and Vendor Extensions - Candidates for Ratification

We are listening to your input!

Support C++ for OpenCL
External Memory Export (in provisional release)
Command Buffer Record/Replay (in provisional release)
Unified Shared Memory
Floating Point Atomics
Required Subgroup Size
Generalized Image from buffer

Image Tiling Controls
YUV Multi-planar Images
Cross-workgroup Barriers
Cooperative Matrices
Timeline Semaphores
32 and 64-length vectors
Indirect Dispatch

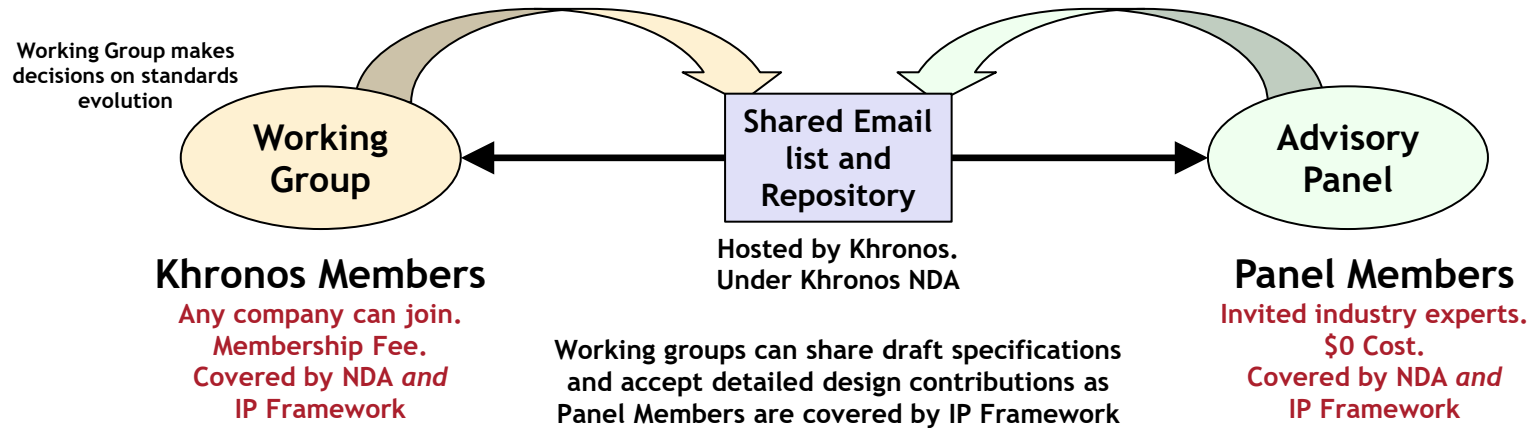


**OpenCL Developer
Survey 2022**

Survey is still open!

Please let us know what
functionality is important to you!

OpenCL Advisory Panel



Chaired by Máté Ferenc Nagy-Egri at StreamHPC

Regular meetings to give feedback on roadmap and draft specifications

Please reach out to opencl-chair@lists.khronos.org if you wish to apply

Developers - Please Give Us Feedback!

- Give us your feedback on the OpenCL spec GitHub
 - What could be added to the OpenCL ecosystem to make you more productive?
 - What API and Language features do you most need?
 - <https://github.com/KhronosGroup/OpenCL-Docs>
- Please download and run the GPUinfo OpenCL Hardware Capability Viewer
 - <https://opencl.gpuinfo.org/download.php>
- Consider applying to join the OpenCL Advisory Panel!
 - Email opencl-chair@lists.khronos.org
- Take the 2022 OpenCL Developer Survey!
 - <https://www.surveymonkey.com/r/J2BCQJN>



OpenCL Resources

- OpenCL Home Page
 - <https://www.khronos.org/opencvl/>
- OpenCL Registry for OpenCL core and extension specifications
 - <https://www.khronos.org/registry/OpenCL/>
- C++ for OpenCL Documentation
 - https://github.com/KhronosGroup/Khronosdotorg/blob/master/api/opencvl/assets/CXX_for_OpenCL.pdf
- OpenCL SDK
 - <https://github.com/KhronosGroup/OpenCL-SDK>
- OpenCL Guide
 - <https://github.com/KhronosGroup/OpenCL-Guide>
- OpenCL Specification Source
 - <https://github.com/KhronosGroup/OpenCL-Docs>
- OpenCL Conformant Products
 - <https://www.khronos.org/conformance/adopters/conformant-products/opencvl>
- GPUinfo.org Hardware Database
 - <https://www.gpuinfo.org/>
- Layered OpenCL implementations - clspv/clvk and OpenCLon12
 - <https://github.com/google/clspv>
 - <https://github.com/kpet/clvk>
 - <https://github.com/microsoft/OpenCLon12>