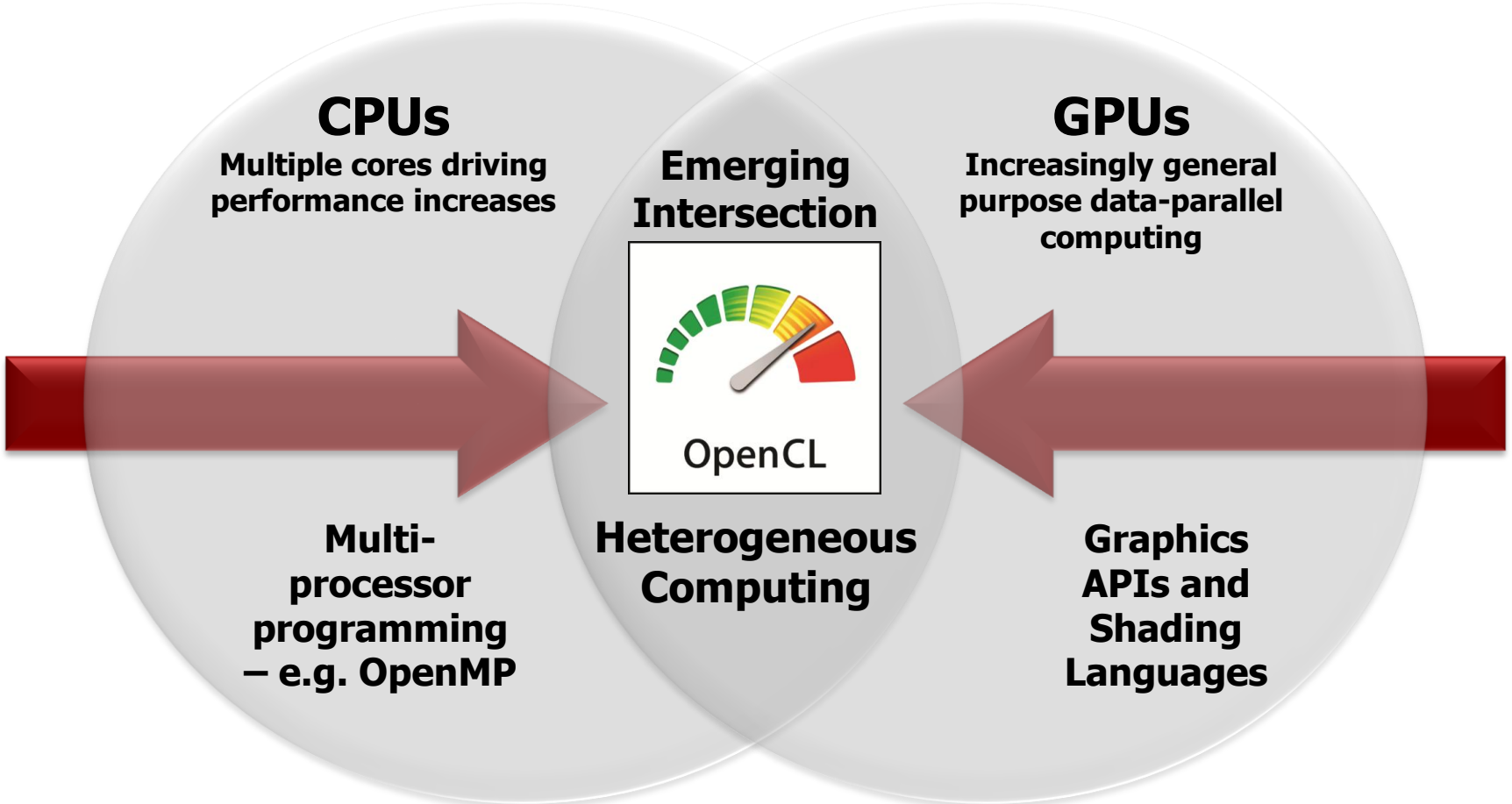
OpenCL WebGL and WebCL

Neil Trevett
Vice President Mobile Content, NVIDIA
President, The Khronos Group

Processor Parallelism



OpenCL is a programming framework for heterogeneous compute resources

The BIG Idea behind OpenCL

- **OpenCL execution model …**
  - Define N-dimensional computation domain
  - Execute a kernel at each point in computation domain
- **C Derivative to write kernels – based on ISO C99**
  - APIs to discover devices in a system and distribute work to them
- **Targeting many types of device**
  - GPUs, CPUs, DSPs, embedded systems, mobile phones.. Even FPGAs

Traditional loops

```
void
trad_mul(int n,
         const float *a,
         const float *b,
         float *c)
{
  int i;
  for (i=0; i<n; i++)
    c[i] = a[i] * b[i];       }
```

Data Parallel OpenCL

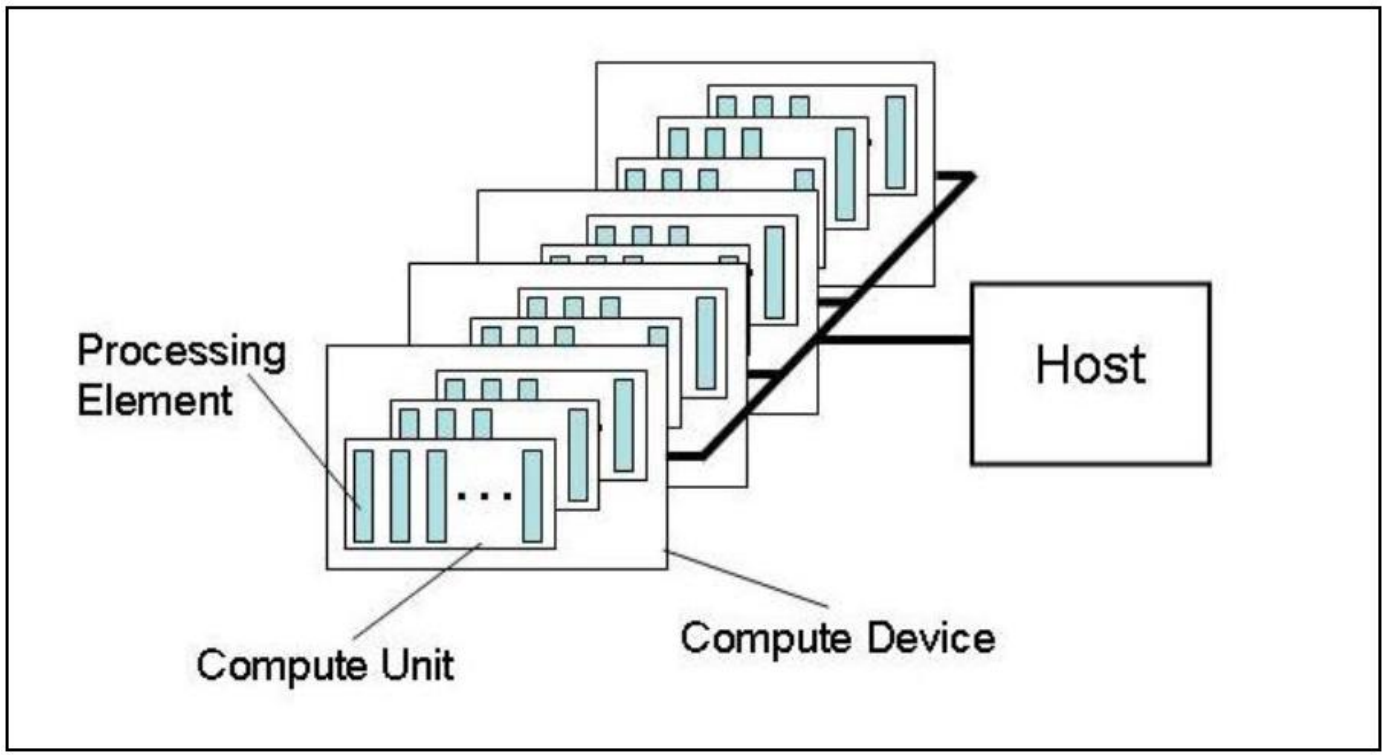
```
kernel void
dp_mul(global const float *a,
       global const float *b,
       global float *c)
{
  int id = get_global_id(0);

  c[id] = a[id] * b[id];

} // execute over "n" work-items
```

OpenCL Platform Model

- **One Host + one or more Compute Devices**
  - Each Compute Device is composed of one or more Compute Units
    - Each Compute Unit is further divided into one or more Processing Elements



OpenCL – Heterogeneous Computing

- **Framework for programming diverse parallel computing resources in a system**

- **Platform Layer API**
  - Query, select and initialize compute devices
- **Kernel Language Specification**
  - Subset of ISO C99 with language extensions
- **Runtime API**
  - Execute compute kernels – gather results
- **OpenCL has Embedded profile**
  - No need for a separate "ES" spec

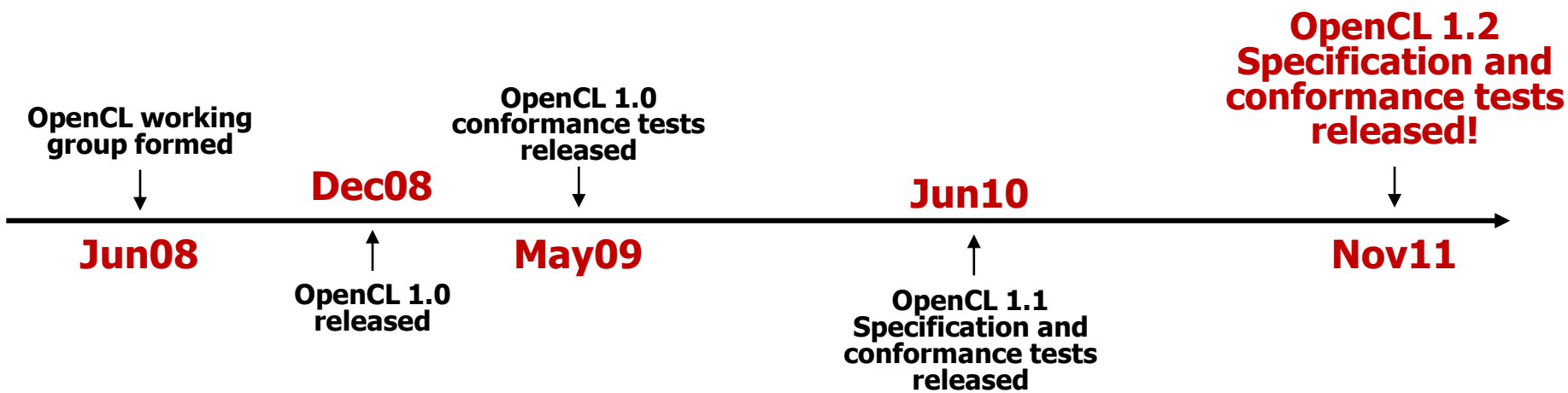
KHRONOS GROUP

OpenCL Working Group Members

- **Diverse industry participation – many industry experts**
  - Processor vendors, system OEMs, middleware vendors, application developers
  - Academia and research labs, FPGA vendors
- **NVIDIA is chair, Apple is specification editor**

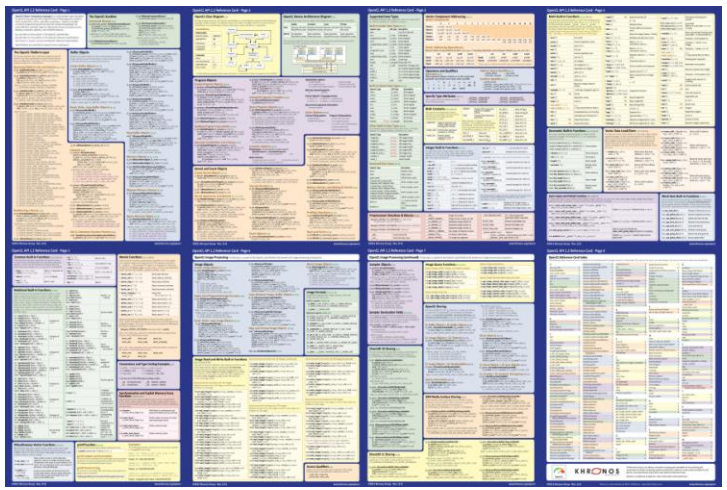
OpenCL Milestones

- **Six months from proposal to released OpenCL 1.0 specification**
  - Due to a strong initial proposal and a shared commercial incentive
- **Multiple conformant implementations shipping**
  - For CPUs and GPUs on multiple OS
- **18 month cadence between OpenCL 1.0, OpenCL 1.1 and now OpenCL 1.2**
  - Backwards compatibility protect software investment



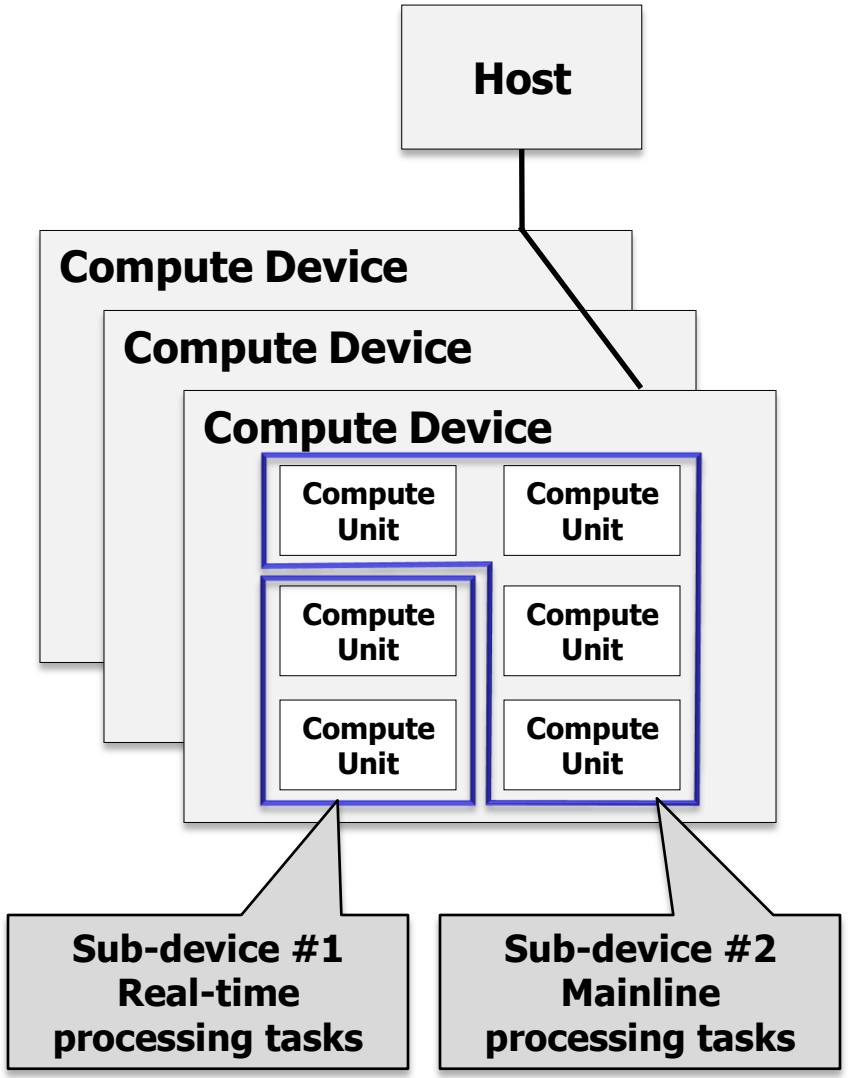
OpenCL 1.2 Announced in December

- **Significant updates - Khronos being responsive to developer requests**
  - Updated OpenCL 1.2 conformance tests available
  - Multiple implementations underway
- **Backward compatible upgrade to OpenCL 1.1**
  - OpenCL 1.2 will run any OpenCL 1.0 and OpenCL 1.1 programs
  - OpenCL 1.2 platform can contain 1.0, 1.1 and 1.2 devices
  - Maintains embedded profile for mobile and embedded devices



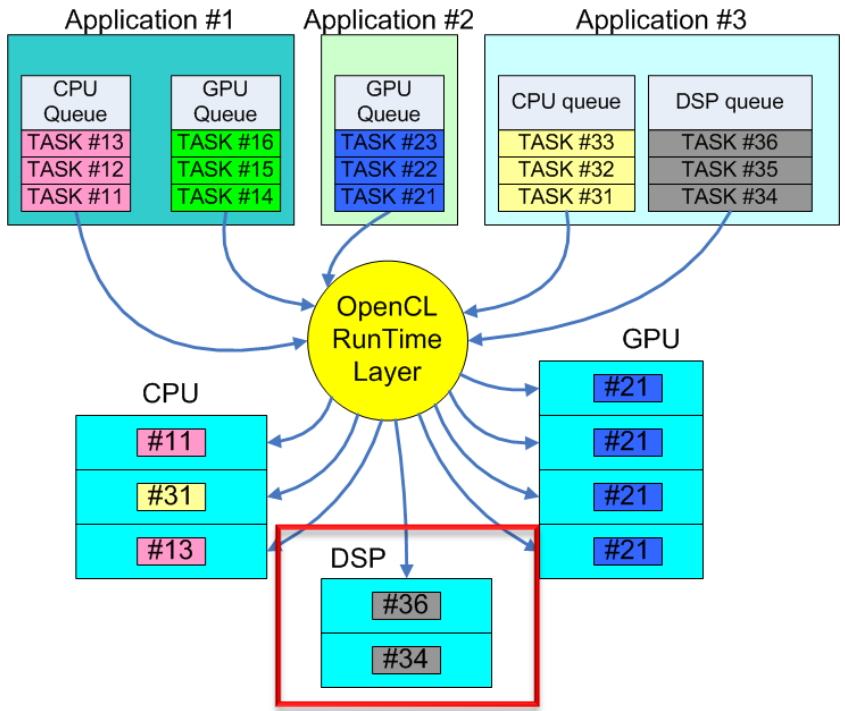
Partitioning Devices

- **Devices can be partitioned into sub-devices**
  - More control over how computation is assigned to compute units
- **Sub-devices may be used just like a normal device**
  - Create contexts, building programs, further partitioning and creating command-queues
- **Three ways to partition a device**
  - Split into equal-size groups
  - Provide list of group sizes
  - Group devices sharing a part of a cache hierarchy



Custom Devices and Built-in Kernels

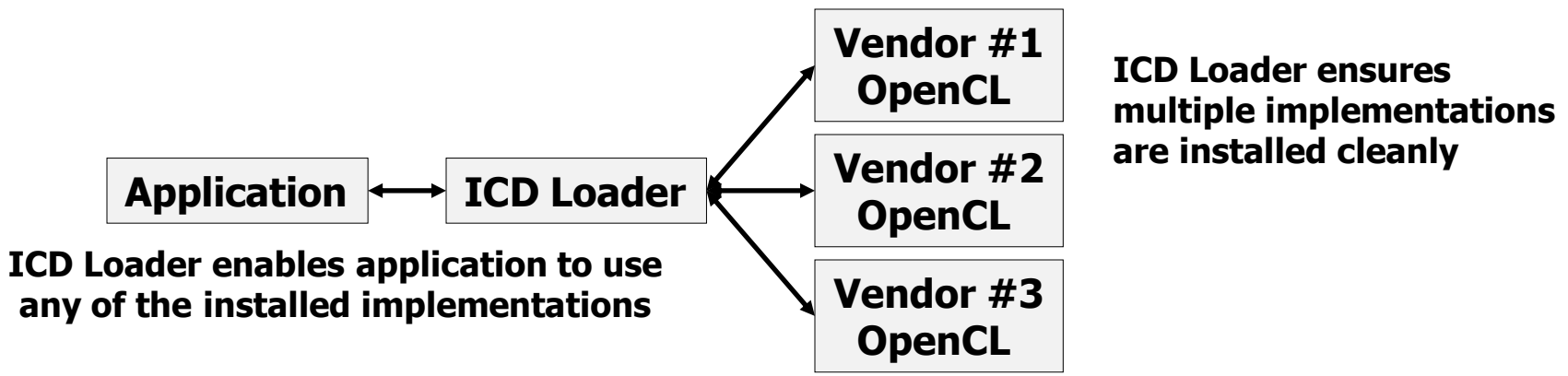
- **Embedded platforms often contain specialized hardware and firmware**
  - That cannot support OpenCL C
- **Built-in kernels can represent these hardware and firmware capabilities**
  - Such as video encode/decode
- **Hardware can be integrated and controlled from the OpenCL framework**
  - Can enqueue built-in kernels to custom devices alongside OpenCL kernels
- **FPGAs are one example of device that can expose built-in kernels**
  - Latest FPGAs can support full OpenCL C as well
- **OpenCL becomes a powerful coordinating framework for diverse resources**
  - Programmable and non-programmable devices controlled by one run-time



Built-in kernels enable control of specialized processors and hardware from OpenCL run-time

Installable Client Driver

- **Analogous to OpenGL ICDs in use for many years**
  - Used to handle multiple OpenGL implementations installed on a system
- **Optional extension**
  - Platform vendor will choose whether to use ICD mechanisms
- **Khronos OpenCL installable client driver loader**
  - Exposes multiple separate vendor installable client drivers (Vendor ICDs)
- **Application can access all vendor implementations**
  - The ICD Loader acts as a de-multiplexor



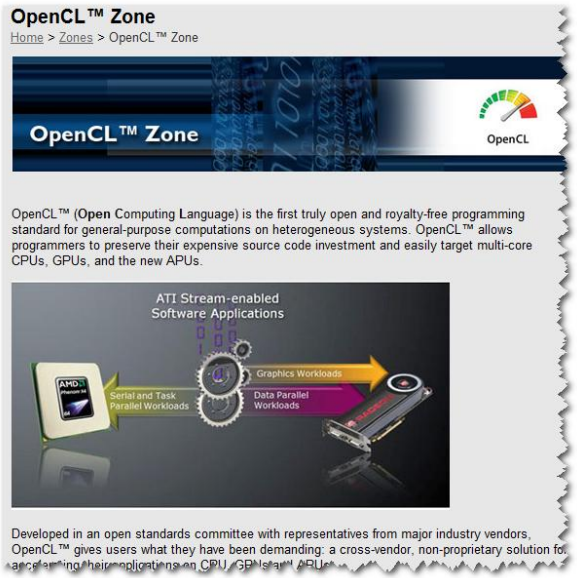
Other Major New Features in OpenCL 1.2

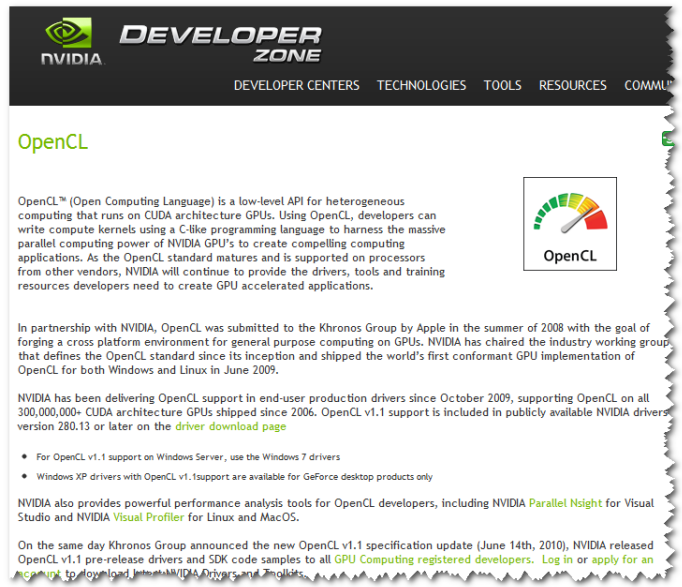
- **Separate compilation and linking of objects**
  - Provides the capabilities and flexibility of traditional compilers
  - Create a library of OpenCL programs that other programs can link to
- **Enhanced Image Support**
  - Added support for 1D images, 1D & 2D image arrays
  - OpenGL sharing extension now enables an OpenCL image to be created from an OpenGL 1D texture, 1D and 2D texture arrays
- **DX9 Media Surface Sharing**
  - Efficient sharing between OpenCL and DirectX 9 or DXVA media surfaces
- **DX11 surface sharing**
  - Efficient sharing between OpenCL and DirectX 11 surfaces
- **And many other updates and additions..**

KHRONOS GROUP

OpenCL Desktop Implementations

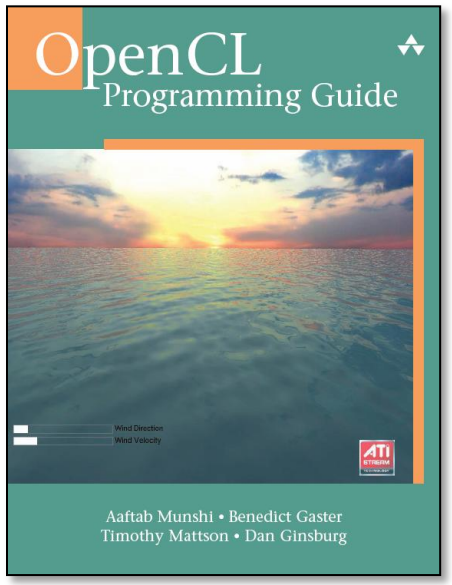
- http://developer.amd.com/zones/OpenCLZone/
- http://software.intel.com/en-us/articles/opencl-sdk/
- http://developer.nvidia.com/opencl

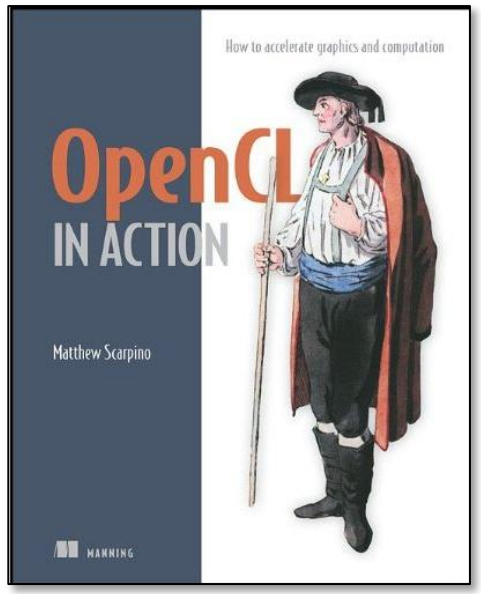


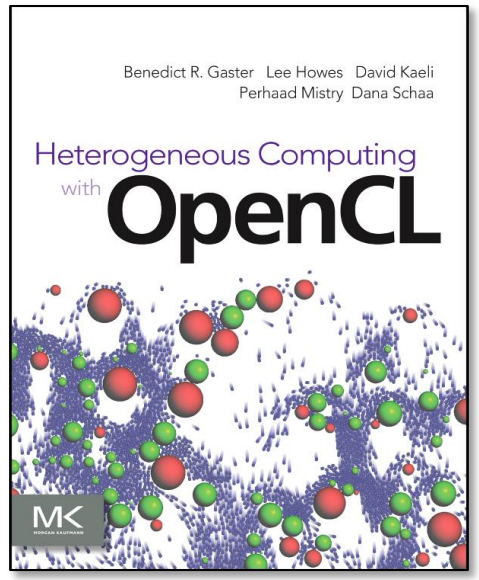


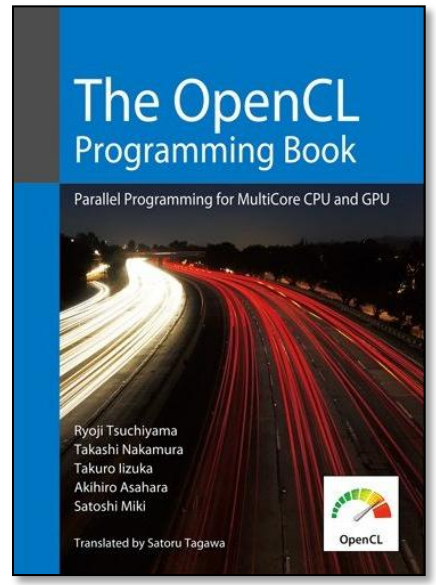
OpenCL Books – Available Now!

- **OpenCL Programming Guide - The "Red Book" of OpenCL**
  - http://www.amazon.com/OpenCL-Programming-Guide-Aaftab-Munshi/dp/0321749642
- **OpenCL in Action**
  - http://www.amazon.com/OpenCL-Action-Accelerate-Graphics-Computations/dp/1617290173/
- **Heterogeneous Computing with OpenCL**
  - http://www.amazon.com/Heterogeneous-Computing-with-OpenCL-ebook/dp/B005JRHYUS
- **The OpenCL Programming Book**
  - http://www.fixstars.com/en/opencl/book/









Spec Translations

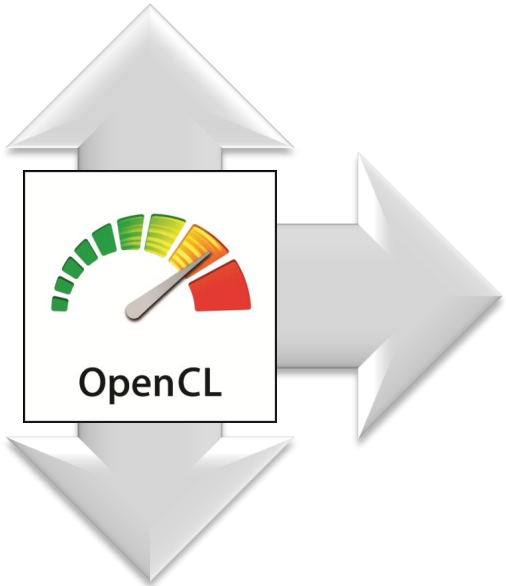
- **Japanese OpenCL 1.1 spec translation available today**
  - http://www.cutt.co.jp/book/978-4-87783-256-8.html
  - Valued partnership between Khronos and CUTT in Japan
- **Working on OpenCL 1.2 specification translations**
  - Japanese, Korean and Chinese

Khronos OpenCL Resources

- **OpenCL is 100% free for developers**
  - Download drivers from your silicon vendor
- **OpenCL Registry**
  - www.khronos.org/registry/cl/
- **OpenCL 1.2 Reference Card**
  - PDF version
  - http://www.khronos.org/files/opencl-1-2-quick-reference-card.pdf
- **Online Man pages**
  - http://www.khronos.org/registry/cl/sdk/1.2/docs/man/xhtml/
- **OpenCL Developer Forums**
  - Give us your feedback!
  - www.khronos.org/message_boards/

Looking Forward

OpenCL-HLM (High Level Model)
Exploring high-level programming model, unifying host and device execution environments through language syntax for increased usability and broader optimization opportunities



Long-term Core Roadmap
Exploring enhanced memory and execution model flexibility to catalyze and expose emerging hardware capabilities

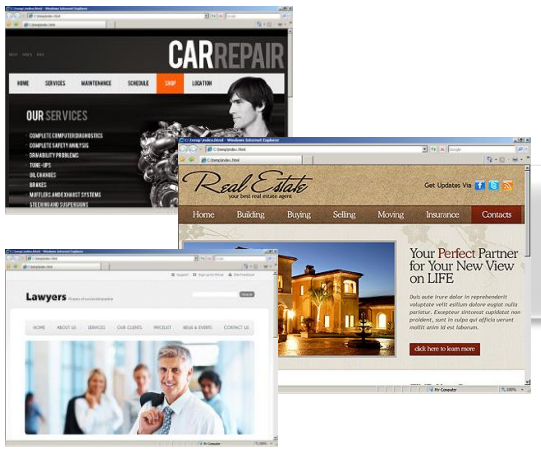
WebCL
Bring parallel computation to the Web through a JavaScript binding to OpenCL

OpenCL-SPIR (Standard Parallel Intermediate Representation)
Exploring low-level Intermediate Representation for code obfuscation/security and to provide target back-end for alternative high-level languages

HTML5 – Cross OS App Platform

- **Increasing diversity of devices creates a demand for a true cross OS programming platform**
- **BUT need more than "more HTML"**

Traditional Web-content



HTML

Rich Experiential Processing

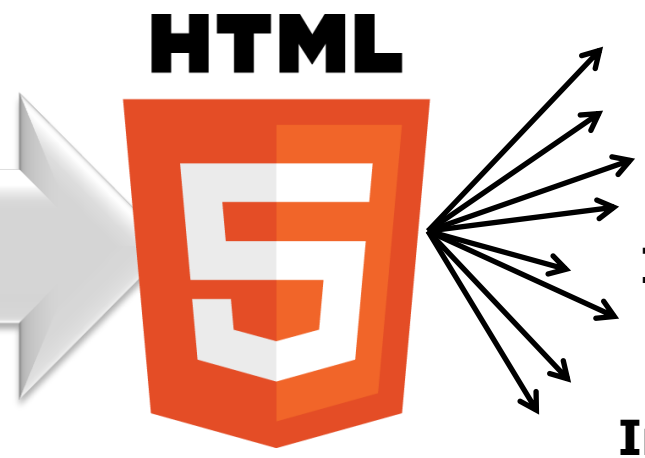
Multi-core CPUs
Rich 2D and 3D GPU
GPU Computing
Multiple HD cameras
Image and vision processing
Video encode/decode
Audio encode/decode
Inertial and positional sensors

How can the Browser rapidly assimilate such diverse functionality?

Web Apps versus Native Apps

- **Mobile Apps have functional and aesthetic appeal**
  - Beautiful, responsive, focused
- **HTML5 with accelerated APIs can provide the same level of "App Appeal"**
  - Highly interactive, rich visual design
- **Using HTML5 to create 'Web Apps' has many advantages**
  - Portable to any browser enabled system
  - Same code can run as app or as web page
  - Web app is searchable and discoverable through the web
  - Not a closed app store – no app store 'tax'

A Lot More than Just "More HTML"

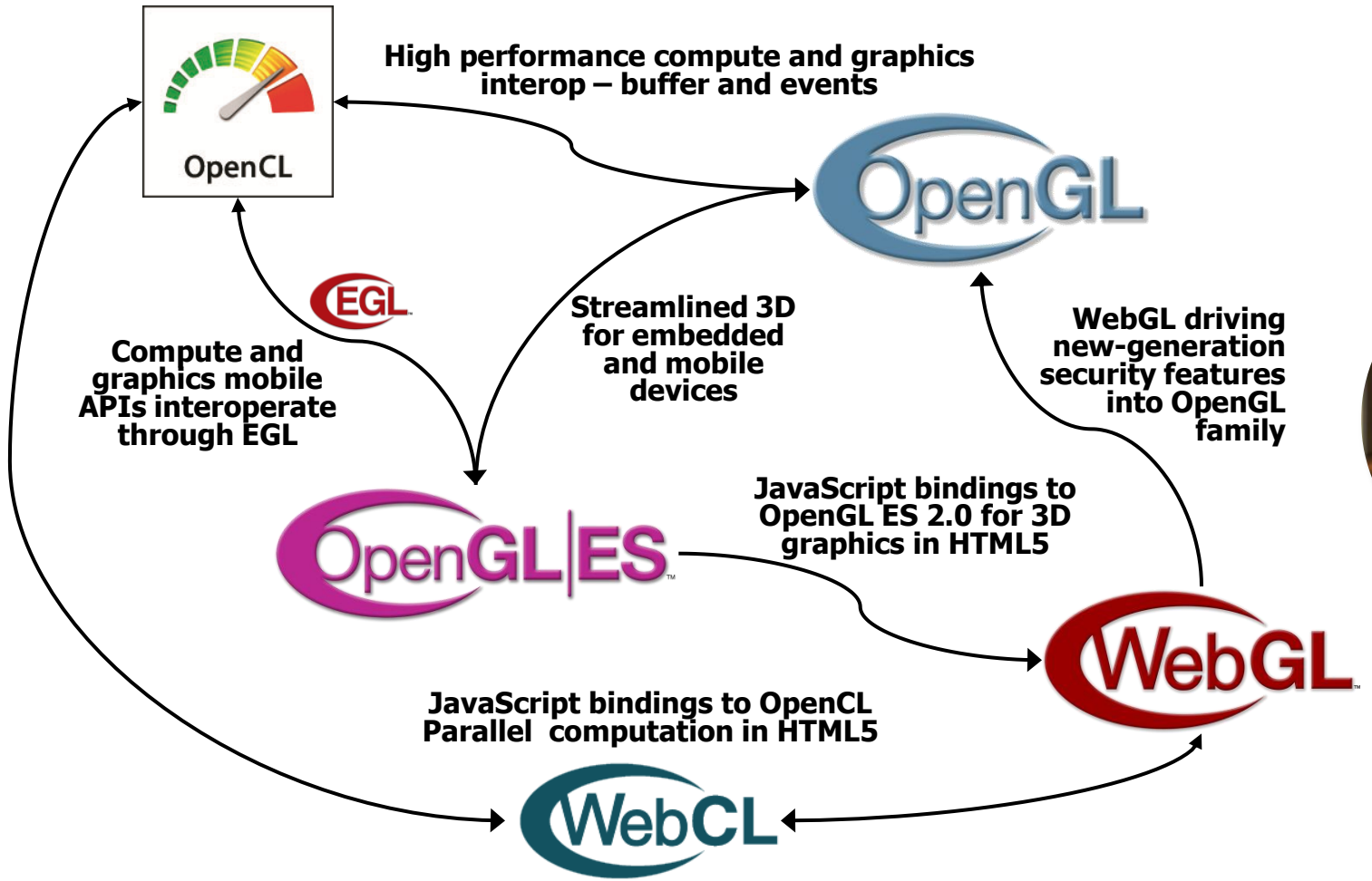


Rich Experiential Processing

Multi-core CPUs
Rich 2D and 3D GPU
GPU Computing
Multiple HD cameras
Image and vision processing
Video encode/decode
Audio encode/decode
Inertial and positional sensors

How can the Browser rapidly assimilate such diverse functionality?

Visual Computing Ecosystem



WebGL – 3D on the Web – No Plug-in!

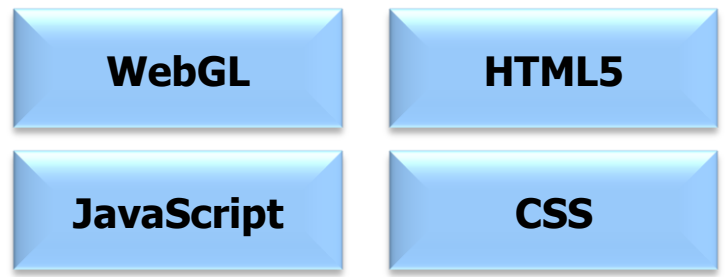
- **Historic opportunity to bring accelerated 3D graphics to the Web**
  - WebGL defines JavaScript binding to OpenGL ES 2.0
- **Leveraging HTML5 and uses <canvas> element**
  - Enables a 3D context for the canvas
- **JavaScript is easily fast enough now for visual computing**
  - Plus OpenGL ES 2.0 enables local geometry caching and GPGPU computation

Being defined by major browsers and GPU vendors working together

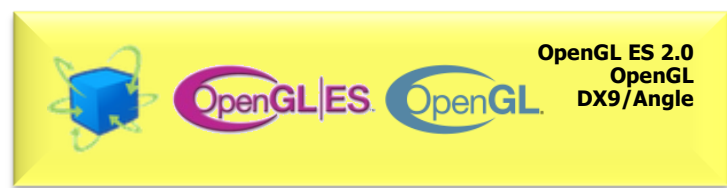
WebGL Implementation Anatomy

Content downloaded from the Web. Middleware can make WebGL accessible to non-expert 3D programmers

Browser provides WebGL functionality alongside other HTML5 specs - no plug-in required

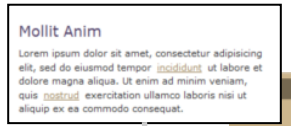


OS Provided Drivers. WebGL on Windows can use Google Angle to create conformant OpenGL ES 2.0 over DX9



HTML5 Content Architecture

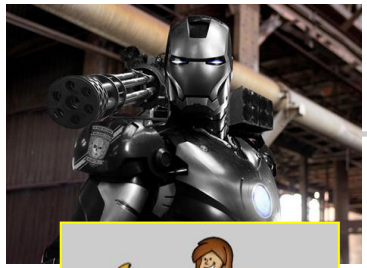
HTML content generated by layout engine 'on page'



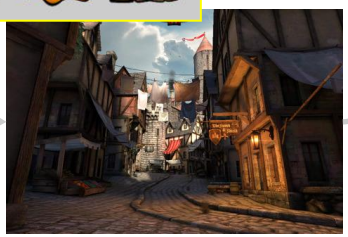
All content passes through CSS layout

CSS Layout and Transforms

<video> tag



<canvas> tag



JavaScript drives interactivity for 2D and 3D graphics

Composition of off-screen buffers

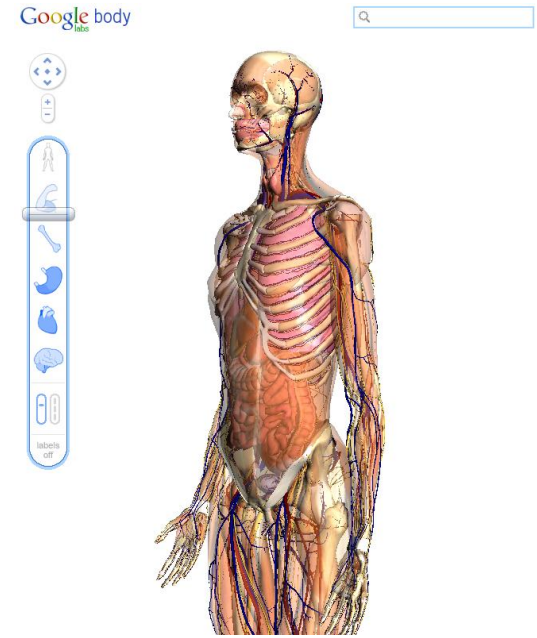
Composition needs to be GPU accelerated

Video, Vector Graphics and 3D created off-screen buffers

WebGL

WebGL and HTML Interaction

- **3D is not trapped in a rectangular window**
  - 3D can overlay and underlay HTML content
  - Easy to make 2D HTML HUDs or 3D user interfaces
- **Strong ties with other advanced HTML5**
  - WebGL can use HTML5 <video>
    or canvas as a texture
- **Can use 3D for core Web UI – as well as content**
  - Advanced transforms and special effects
- **Render HTML DOM sub-tree as texture**
  - Mozilla and Google prototyping as extension
  - Support user interaction when in 3D



WebGL Deployment

- **WebGL 1.0 Released at GDC March 2011**
  - Mozilla, Apple, Google and Opera working closely with GPU vendors
- **Typed array 1.0 spec ratified by Khronos in May**
  - Supporting bulk data transfer between threads (workers)
  - Many use cases - background mesh loading, generation, deformation, physics ...
- **1.0.1 release of WebGL spec and conformance suite imminent**
  - 100% robust stance on security
  - Fixing bugs in 1.0.0 conformance suite
  - Implementations will report getContext("webgl") (not experimental)

Show all versions	IE	Firefox	Chrome	Safari	Opera	iOS Safari	Opera Mini	Opera Mobile	Android Browser	
3 versions back	6.0	5.0	12.0	3.2	11.0	3.2		10.0		
2 versions back	7.0	6.0	13.0	4.0	11.1	4.0-4.1		11.0	2.1	
Previous version	8.0	7.0	14.0	5.0	11.5	4.2-4.3		11.1	2.2	
Current	9.0	8.0	15.0	5.1	11.6	5.0	5.0-6.0	11.5	2.3	3.0
Near future		9.0	16.0		12.0				4.0	
Farther future	10.0	10.0	17.0	6.0	12.1					

WebGL is not enabled by default in Safari

http://caniuse.com/#search=webgl

Frameworks and Tools

- **WebGL is deliberately low level to enable the full power and flexibility of OpenGL ES 2.0**
- **If you are not an expert 3D programmer – don't panic!**
- **WebGL is perfect foundational layer for JavaScript middleware frameworks**
- **Lots of utilities and tools already appearing**

http://www.khronos.org/webgl/wiki/User_Contributions

WebGL Security

- **Any new functionality in the browser increases exposure to attack**
  - True since the beginning of the web – the new functionality becomes hardened
- **ANY graphics in the browser need the GPU drivers to be hardened**
  - HTML, Canvas, WebGL, Adobe Molehill, Silverlight 5 …
- **WebGL is designed with security as the highest priority**
  - Hardening is being strongly promoted and enabled
- **Short term – browser vendors will maintain white and black lists**
  - Compromised system can have WebGL disabled until mitigation developed
- **Longer term – GPUs provide increasingly robust security and multi-tasking**
  - GPU becoming a first-class computing platform alongside CPU

KHRONOS GROUP

WebGL Security in the Press!

- **Confusion in the industry as we start this hardening process**
  - Shader programs *cannot* access general system resources or perform out of range memory access!
- **Issues in the Press**
  - Cross domain image access – timed loop attack – SOLVED!
    - WebGL *and HTML* spec updates - mandating CORS for video, images and audio
    - Servers have to grant cross-domain access to media resources
  - DOS Attacks and general hardening
    - ARB_robustness extensions that provide additional protection being mandated
    - New robustness spec limits the side-effects of a GPU reset after a DOS attack
    - ANGLE shader validator improved; more improvements coming

WebCL – Parallel Computing for the Web

- **Khronos launching new WebCL initiative**
  - First announced in March 2011
  - API definition already underway
- **JavaScript binding to OpenCL**
  - Security is top priority
- **Many use cases**
  - Physics engines to complement WebGL
  - Image and video editing in browser
- **Stay close to the OpenCL standard**
  - Maximum flexibility
  - Foundation for higher-level middleware

KHRONOS GROUP

WebCL Open Process and Resources

- **Khronos open process to engage Web community**
  - Public specification drafts, mailing lists, forums
  - http://www.khronos.org/webcl/
  - webcl_public@khronos.org
- **Khronos welcomes new members to define and drive WebCL**
  - info@khronos.org
- **Nokia open sourced prototype for Firefox in May 2011 (LGPL)**
  - http://webcl.nokiaresearch.com
- **Samsung open sourced prototype for WebKit in July 2011 (BSD)**
  - http://code.google.com/p/webcl/

- **Deformation Demo:**
- Calculates and renders transparent and reflective deformed spheres on top of photo background
- Performance comparison
  - JS: ~1 FPS
  - WebCL: 87-116 FPS
- http://www.youtube.com/user/SamsungSISA#p/a/u/1/9Ttux1A-Nuc

Web Apps - Wider Ecosystem

- **OS capability access before in HTML5**
  - Execution with no browser UI
  - Packaging standalone apps

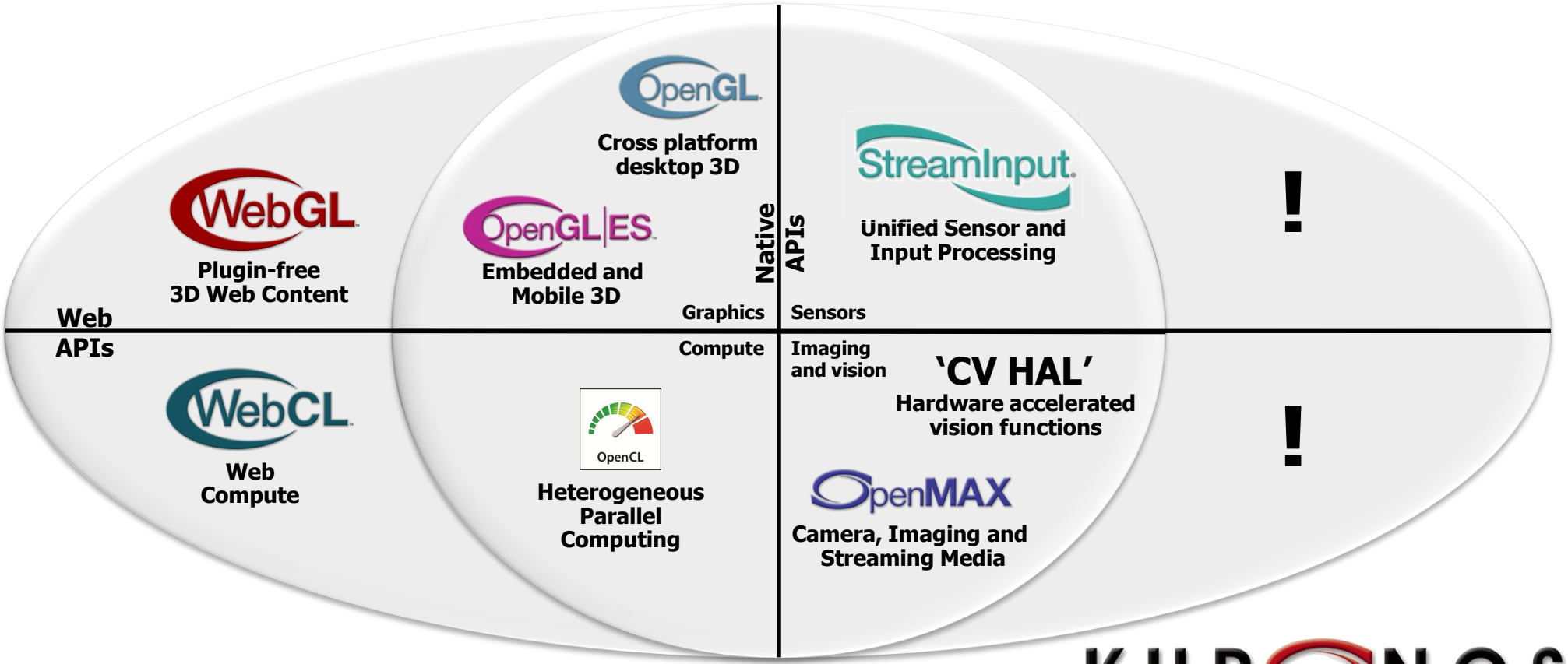
- **OS Independent App stores**
  - Discovery and payment

- **Language and JavaScript Tools**
  - Native code compilation to JavaScript
  - JavaScript libraries

- **Authoring Tools**
  - Bringing Flash-grade authoring to HTML5

Industry HTML5 Ecosystem

Khronos Ecosystem of Standards



KHRONOS GROUP