AMD

AMD HETEROGENEOUS COMPUTING

MAY 29, 2014

BRUNO STEFANIZZI
DMITRY KOZLOV

MARKET TRENDS SHAPING AMD TECHNOLOGY INVESTMENTS

AMD

Consumer tech adoption is expanding, diversifying and evolving

New ecosystems and architectures emerging from the datacenter to embedded

Customers seek new approaches to innovation to differentiate and win

AMD SEES CLEAR DEMAND FOR COMPUTE INNOVATION…

GROWTH IN COMPUTE

- x86 and ARM® will be the dominant architectures moving forward
- Perpetual demand for more compute to propel new capabilities and experiences

GRAPHICS & ACCELERATION

- Graphics, visualization and video content on the rise in devices and the cloud

MERGING OF ECOSYSTEMS

- Data and application interoperability across all devices and ecosystems

HETEROGENEOUS COMPUTING PLATFORMS ARE BECOMING PERVASIVE

- EVERY Tablet and Smartphone shipping in 2012 and 2013 is based on a heterogeneous processor
- Of the Top 500 supercomputers in 2012, 52 were based on heterogeneous systems
- Both the major next generation consoles from SONY and Microsoft are shipping with heterogeneous processors
- AMD starting shipping embedded heterogeneous processors in 2011 and has announced that it will ship mainstream heterogeneous servers in 2014

Sources: See Reference slide

AMBIDEXTROUS COMPUTING

THE AMD VISION

Give Customers **New Paths to Innovation**

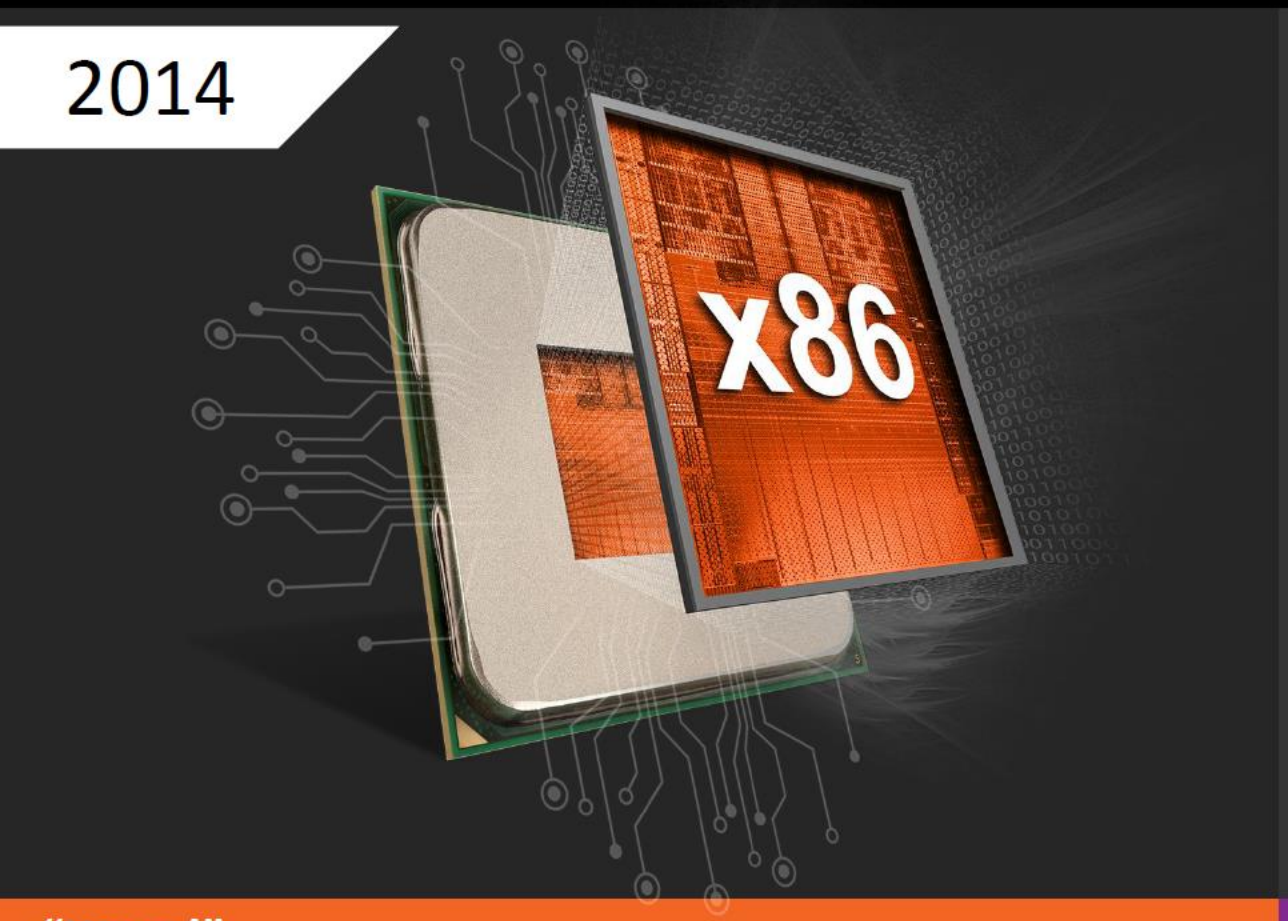
Harness the Best of the x86 and ARM Global Ecosystems

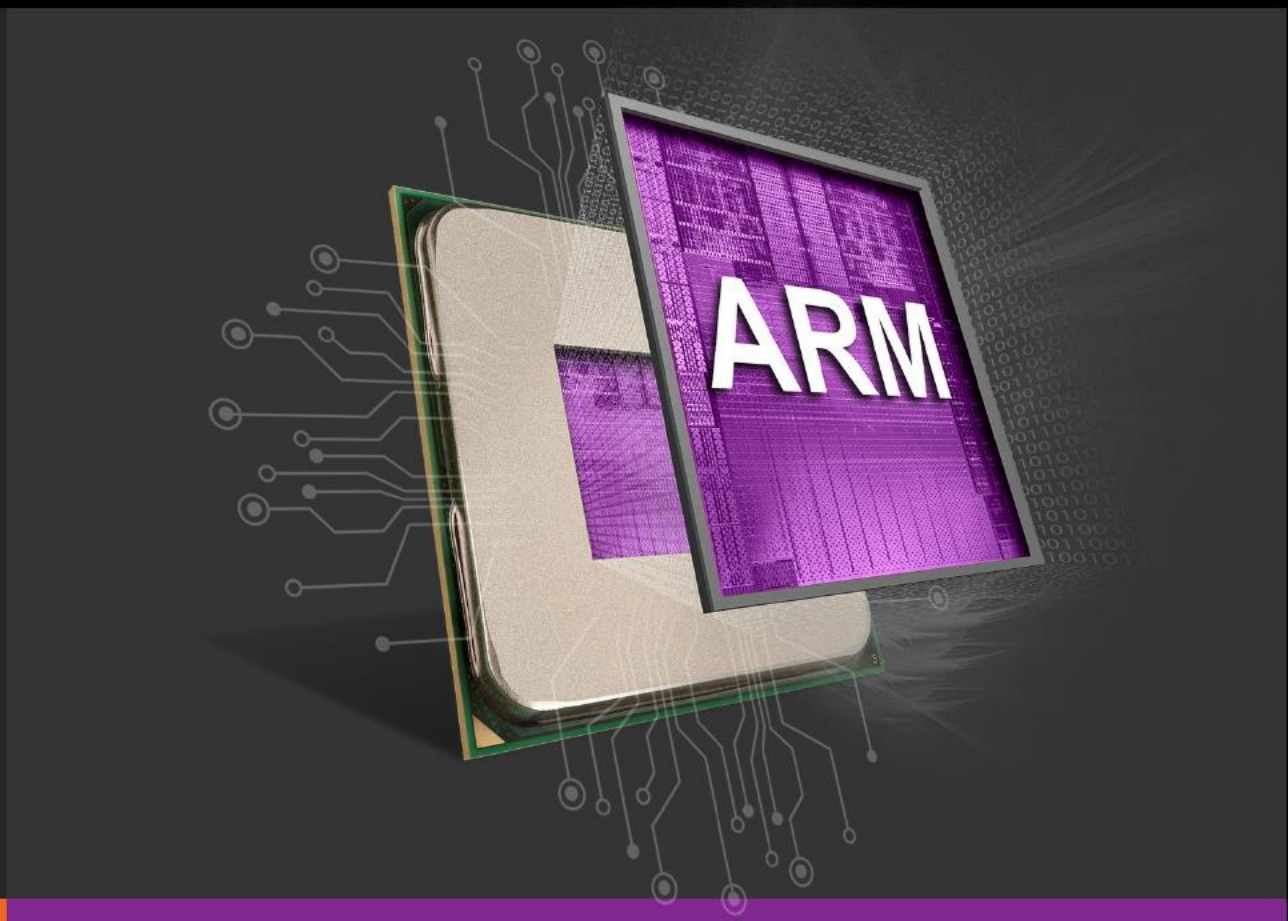
Offer **Compute Leadership** in Both x86 and 64-bit ARM

AMBIDEXTROUS COMPUTING ROADMAP

AMD

2014





“Kaveri”

- AMD Introduces World’s first APU with HSA Features

“Beema” and “Mullins"

- Mainstream and Low-Power APUs deliver performance leadership

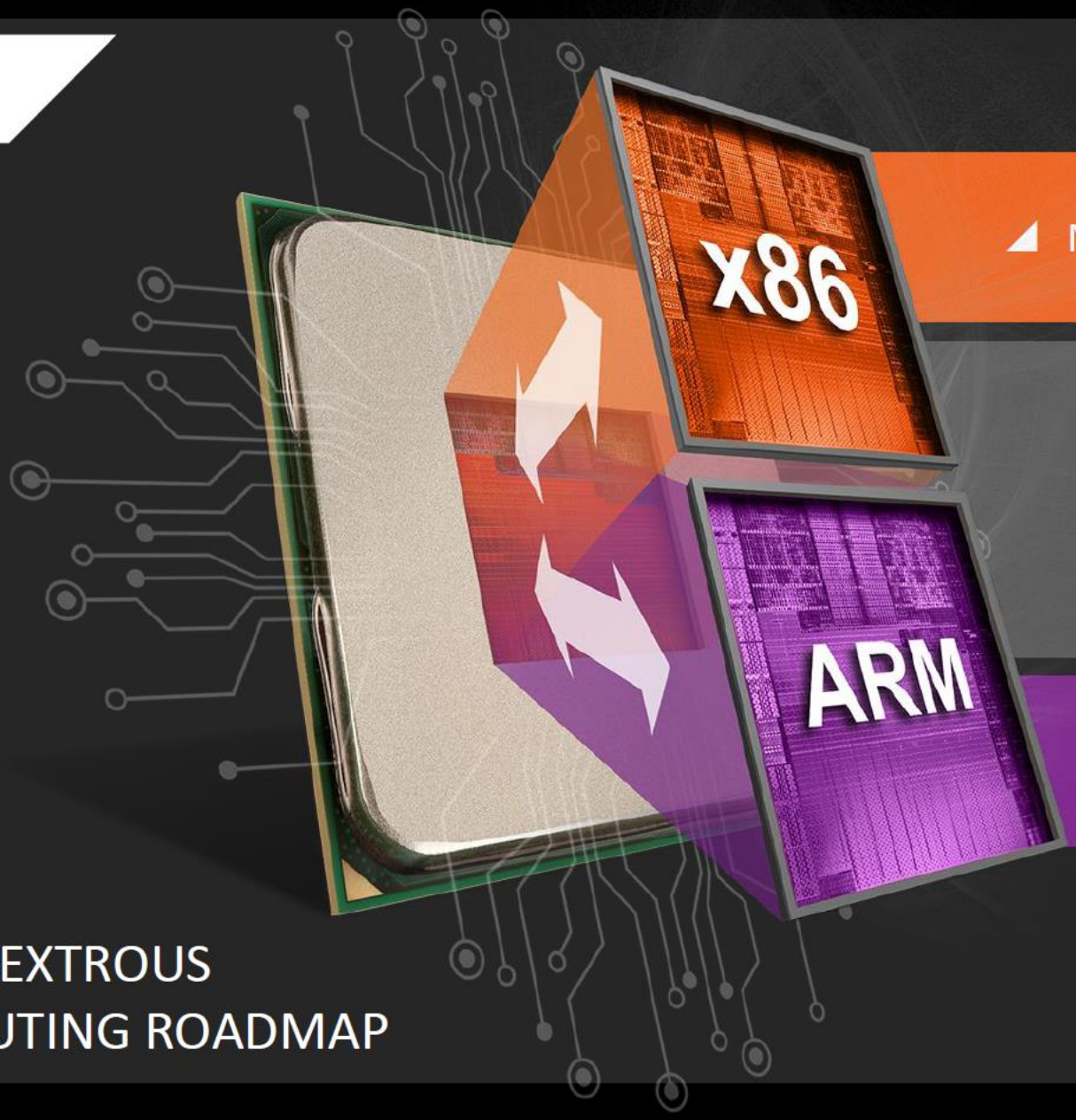
“Seattle”

- AMD samples world’s first 28nm 64-bit ARM server processor

AMD “PROJECT SKYBRIDGE” AMBIDEXTROUS DESIGN FRAMEWORK

AMD

2015



- Next-Gen “Puma+” x86 Cores

- Starting in 2015, families of 20nm APUs and SoCs with pin-compatible x86/ARM compute
- Designed for Full HSA support
- AMD Graphics Core Next (GCN)

- Low-power A57 64-bit ARM Cores
- AMD‘s first HSA Android platform

AMBIDEXTROUS
COMPUTING ROADMAP

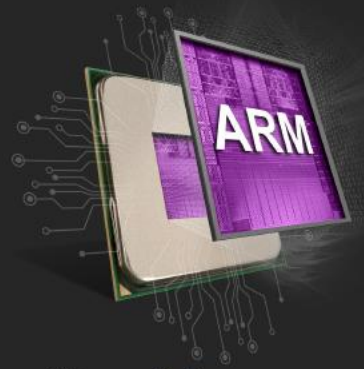
AMBIDEXTROUS COMPUTING ROADMAP

AMD

2014

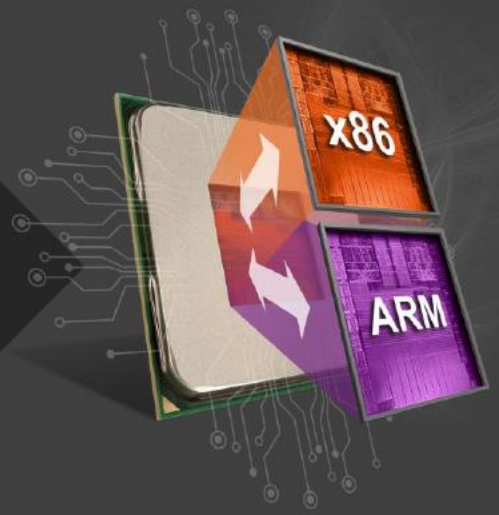
"Kaveri"

"Beema" and "Mullins"



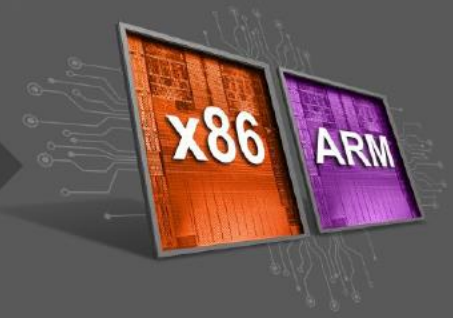
"Seattle"

2015



"Project SkyBridge" family of 20nm APUs and SoCs

2016+



Developing 64-bit ARM cores alongside new 64-bit x86 cores

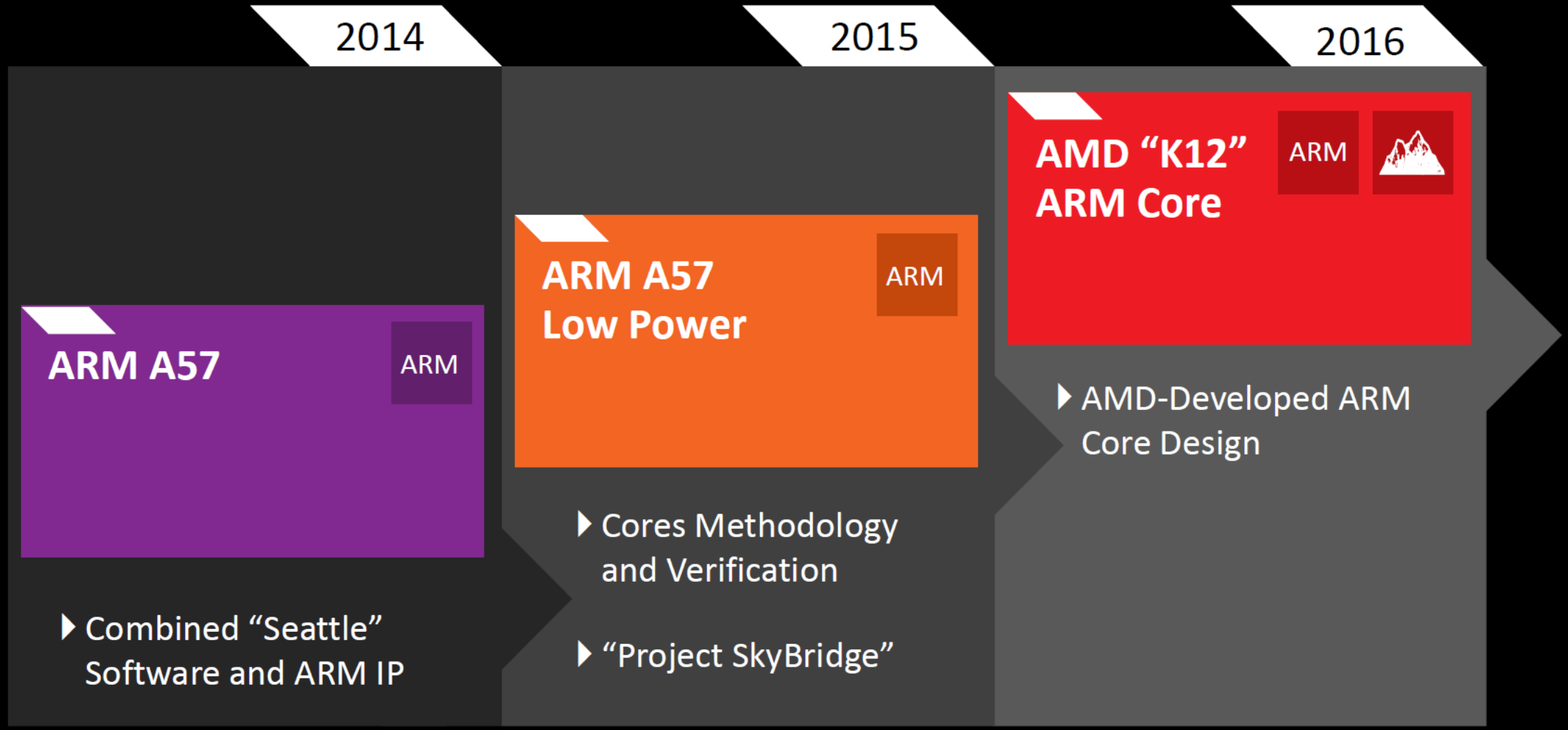
Ambidextrous Computing Leadership

x86 and 64-bit ARM for:

- ☑ **Dense Server**
- ☑ **Embedded**
- ☑ **Semi-Custom**
- ☑ **Ultra-Low Power Client**

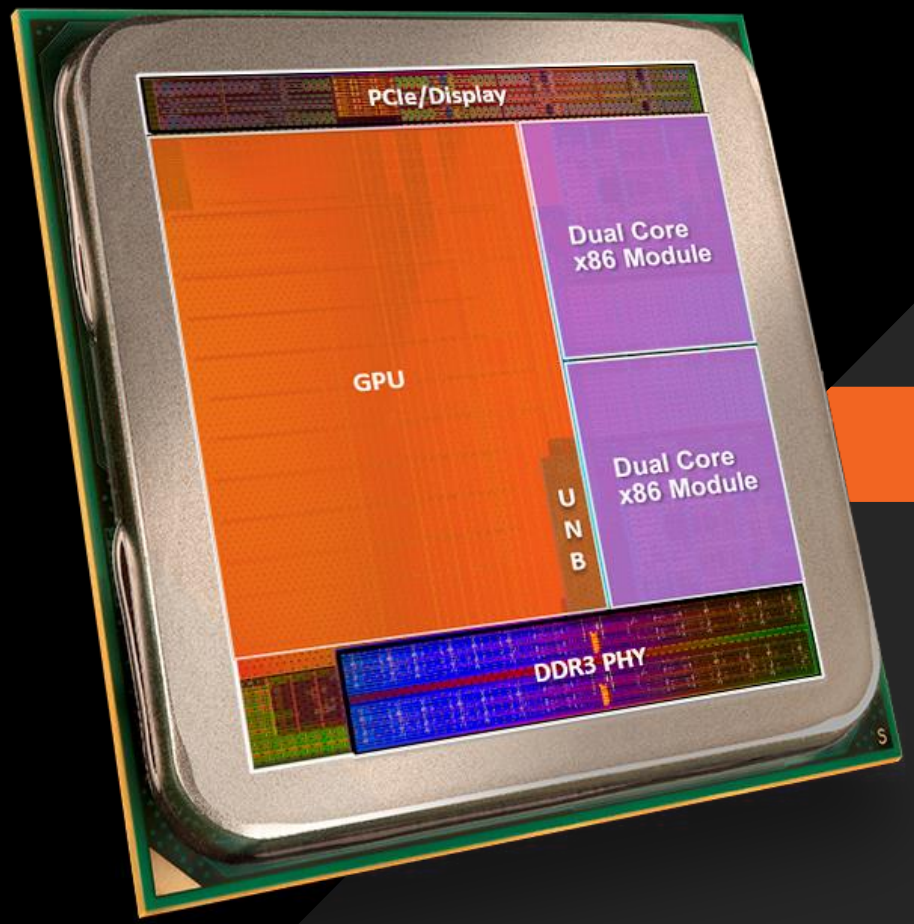
AMD'S ARM CORE ROADMAP

AMD



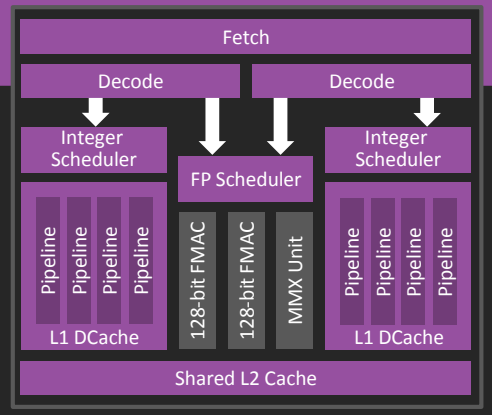
AMD roadmaps are subject to change without notice or obligations to notify of changes.
Placement of boxes intended to represent first year of production shipments.

“KAVERI” FEATURING UP TO 12 COMPUTE CORES (4 CPU+ 8 GPU)



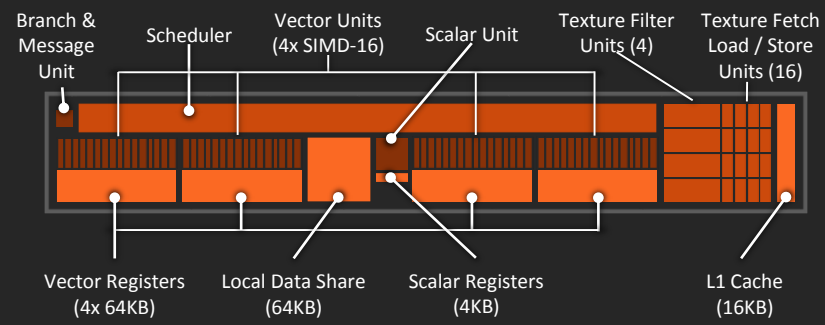
CPU COMPUTE CORES

Up to four new multi-threaded AMD “Steamroller” CPU CORES



GPU COMPUTE CORES

Up to eight GCN GPU CORES powering parallel compute and next-gen gaming



PROFESSIONAL USERS DEMAND

A NEW CLASS OF GPU FOR A NEW ERA

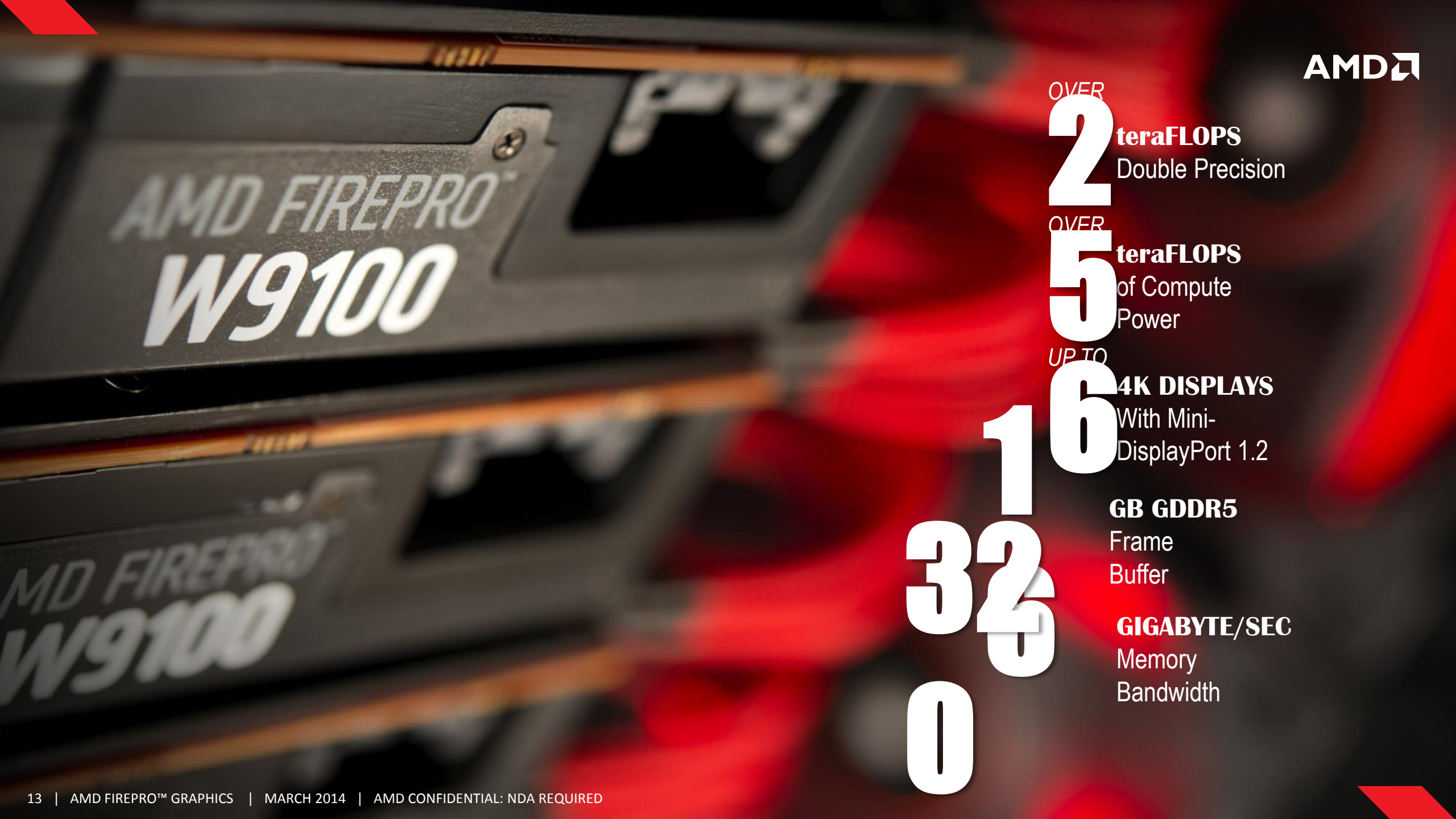
…OF CONTENT CREATION & PRODUCTIVITY

INTRODUCING

AMD FIREPRO™ W9100 GRAPHICS

DESIGNED FOR 4K

POWERED BY OPENCL™



AMD

OVER
2 **teraFLOPS**
Double Precision

OVER
5 **teraFLOPS**
of Compute
Power

UP TO
16 **4K DISPLAYS**
With Mini-
DisplayPort 1.2

32 **GB GDDR5**
Frame
Buffer

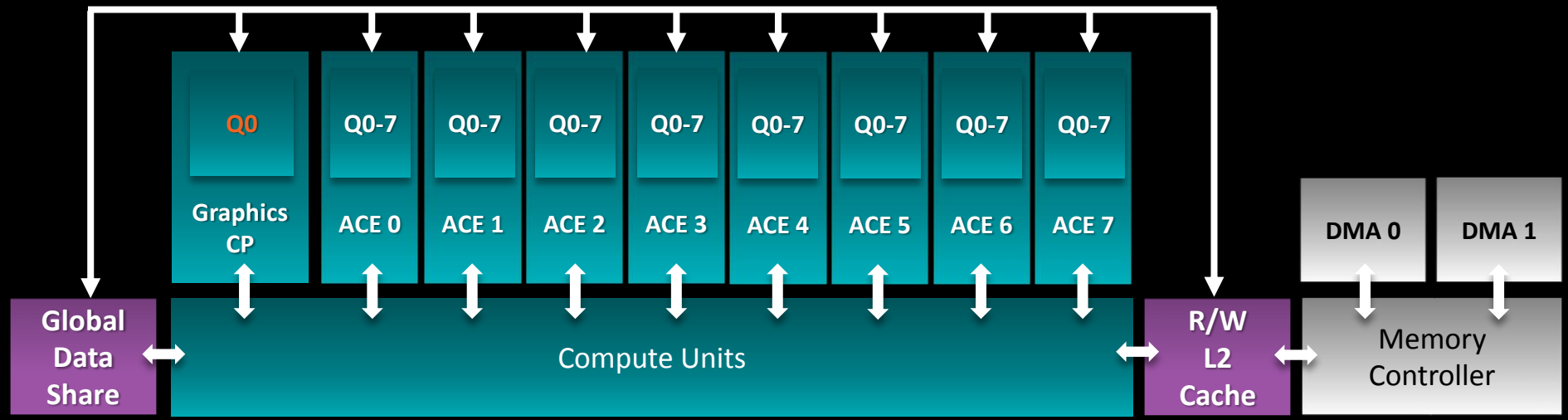
320 **GIGABYTE/SEC**
Memory
Bandwidth

AMD FirePro™ Next Gen "Hawaii" Series

- **44 Compute Units**
  - **½ Rate DPSP**
- **176 Texture Units**
- **704 32b Load/Store Units**
- **1MB R/W L2 Cache**
- **512-bit GDDR5 memory interface**
- **Up to 16 GB video memory**
- **6.2 billion transistors**
  - **438 mm² on 28nm process node**

Asynchronous Compute

- Up to 8 Asynchronous Compute Engines (ACE)
  - Independent scheduling and work item dispatch for efficient multi-tasking
  - Operate in parallel with graphics command processor
  - Each can manage up to 8 queues
  - L2 cache and GDS access
  - Fast context switching
- Dual DMA engines
  - Can saturate PCIe 3.0 x16 bus bandwidth (16 GB/sec bidirectional)



EXECUTING CLEAR VISION
FOR AMBIDEXTROUS COMPUTING

Ambidextrous:
Comprehensive **ARM a**nd **x86** Roadmap Across Fastest Growing Markets

Leveraging AMD Leadership In **Graphics** and **Heterogeneous Compute**

Paving New Paths Of **Innovation** Unique to AMD Customers and Partners

AMBIDEXTROUS COMPUTING

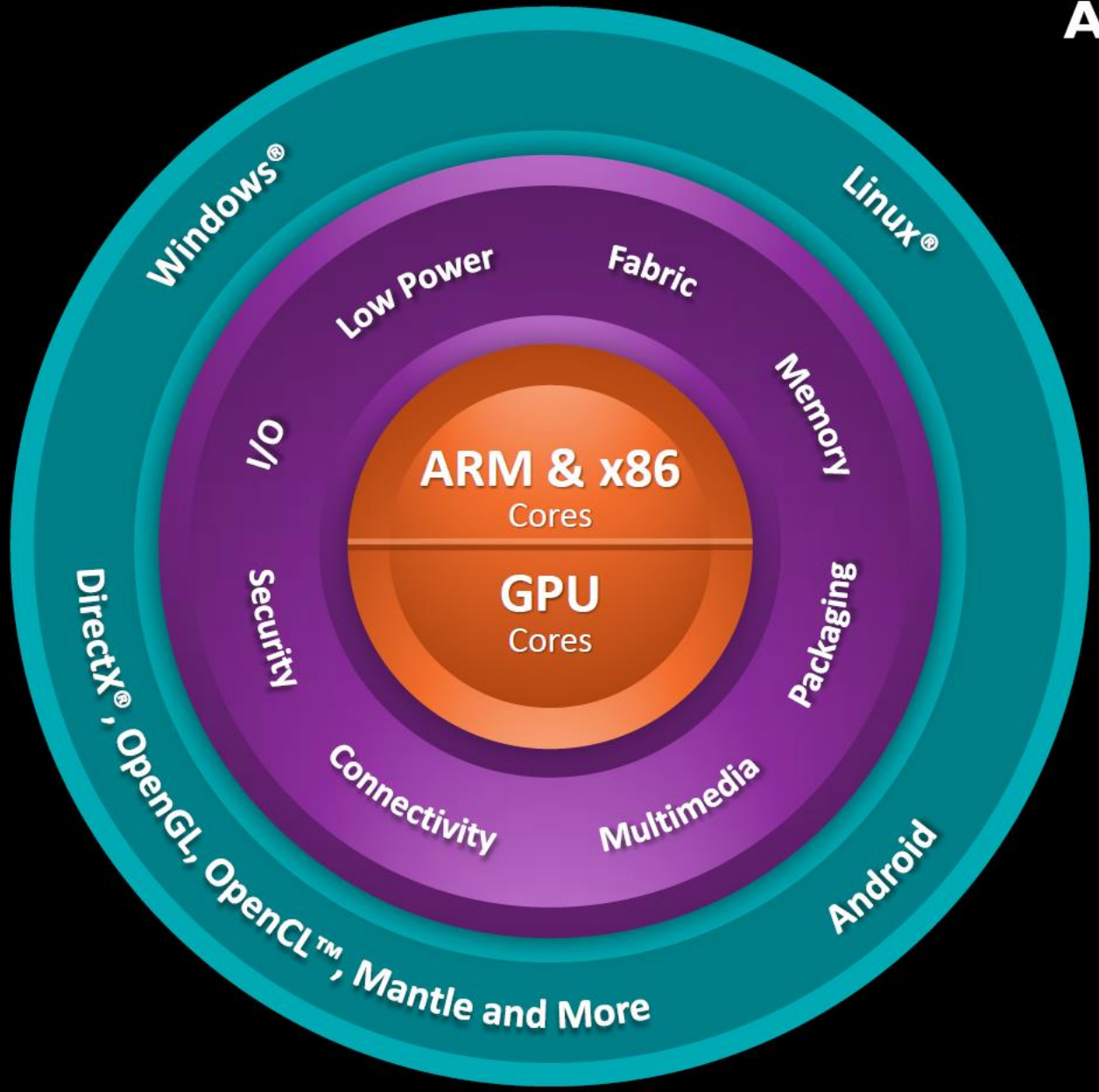
TURNING VISION INTO REALITY

Reengineered Design Process for ARM and x86

Shared, Modular IP and Components

AMD IP
BUILDING BLOCKS

- CROWN JEWEL CORES
- COMPLEMENTARY SoC IP
- OPTIMIZED SOFTWARE



AMD

PROBLEM

PROGRAMMING HETEROGENEOUS SOCs/APUs

IS STILL HARD

From dGPU to SoC

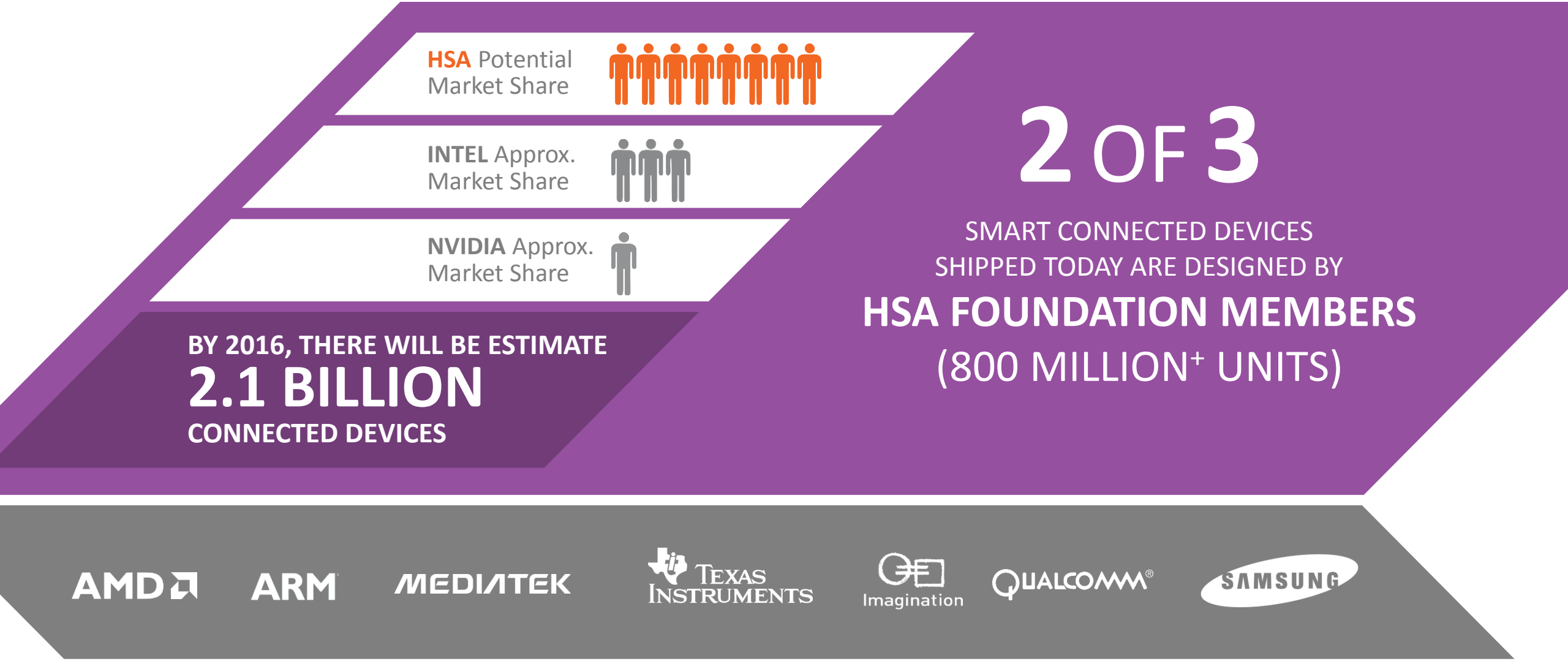
Need for open programming framework for legacy and future-proof codes

HSA Foundation

Write-once-run-everywhere for heterogeneous systems

- Industry leaders setting an industry standard
- Designed for developers, by developers
- One architecture, differentiated 'IP' vendors
  - Multiple hardware solutions to be exposed to software through a common standard low-level interface layer
- Support high-level parallel programming languages and models, including C++, C#, FORTRAN, **OpenCL**, **OpenMP**, Java and Python

AMD JOINED BY FELLOW INDUSTRY LEADERS TO ADVANCE HETEROGENEOUS DESIGN



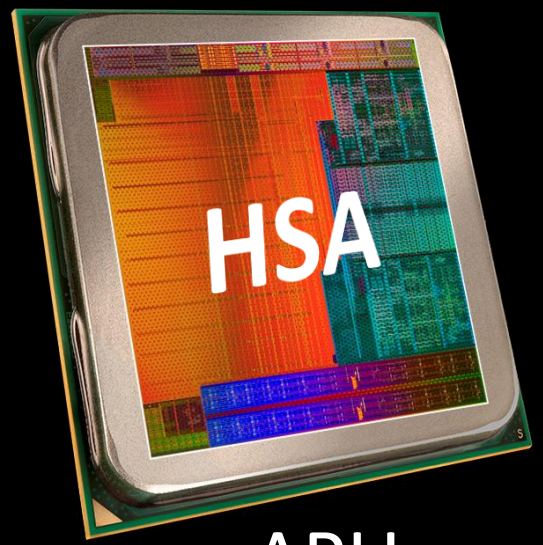
For more information go to: http://hsafoundation.Com/ Source http://pinterest.Com/pin/193021534001931884

TERMINOLOGY

- **APU = CPU + GPU**
- **Heterogeneous Computing**
  - Heterogeneous computing refers to systems that use more than one kind of processor.
- **COMPUTE = HARDWARE ACCELERATION = GPU ACCELERATION = GPGPU**
  - GPU-acceleration is the term used when the task that a portion of the task that was traditionally processed on CPU is now processed on a GPU i.e. GPU ‘offloads’ the CPU.
- **Heterogeneous System Architecture = HSA**
  - **APU Processor design** that makes it easy to harness the **entire computing power of an APU** for faster and more power-efficient devices, including personal computers, tablets, smartphones and cloud servers
- **Programming Languages (OpenCL, Java, OpenMP, Python)**
  - OpenCL™ (**Open C**omputing **L**anguage) is the first truly open and royalty-free programming standard forgeneral-purpose computations on heterogeneous systems. OpenCL™ allows programmers to preserve their expensive source code investment and easily target multi-core CPUs, GPUs, and the new APUs.
  - Java , a general-purpose computer programming language, allows you to play online games, chat with people around the world, calculate your mortgage interest, and view images in 3D, just to name a few. It's also integral to the intranet applications and other e-business solutions that are the foundation of corporate computing.

TERMINOLOGY, VISUALIZED

DEVELOPERS



APU

APPLICATIONS

Java

OpenCL

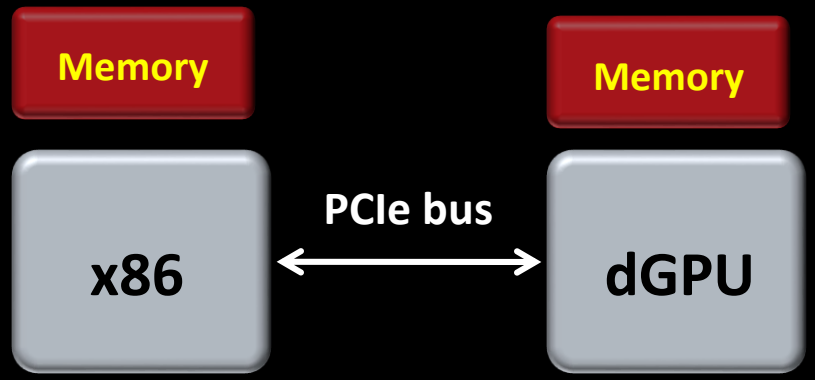
C++

PROGRAMMING LANGUAGES

AMD

SOLUTION

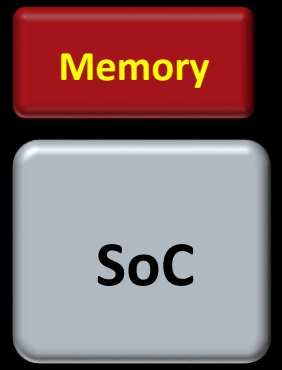
GPU Computing Today and Tomorrow



DATA TO WORK

Since 2007

Not efficient!



WORK TO DATA

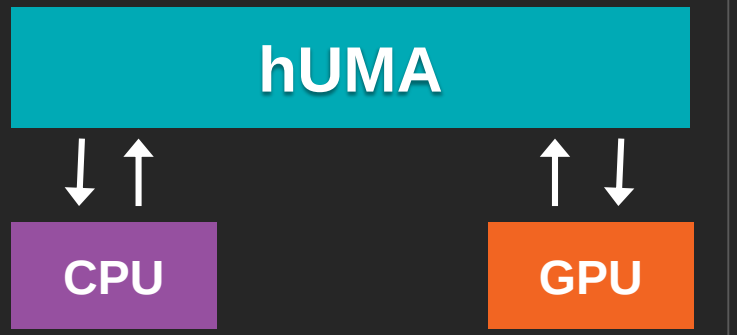
Beyond 2015

Run indifferently serial and parallel codes

HSA FEATURES FEED AND EQUALIZE "KAVERI 'COMPUTE UNITS'"

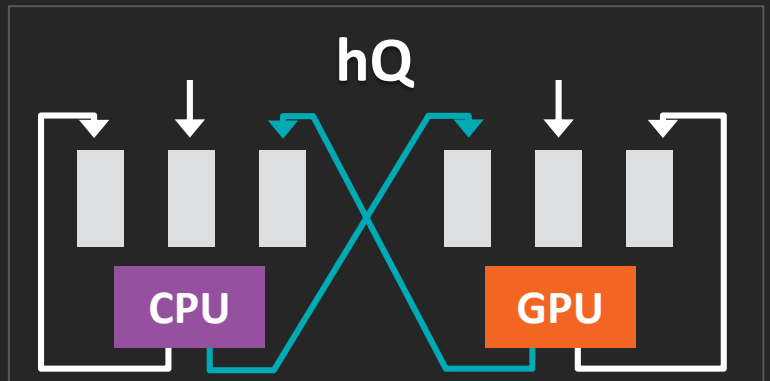
AMD

EQUAL ACCESS TO ENTIRE MEMORY



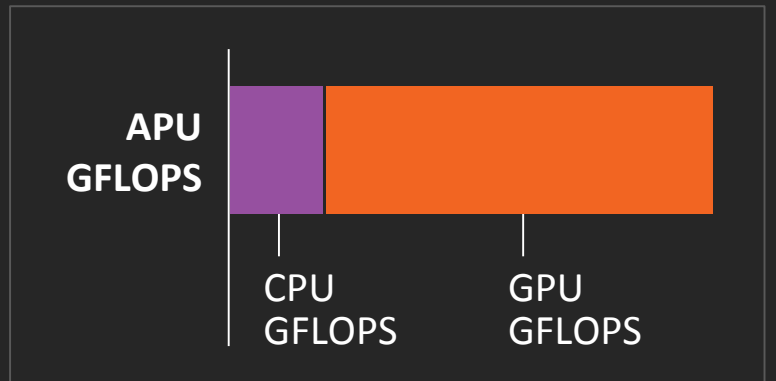
- First time ever: GPU and CPU have uniform visibility into entire memory space (up to 32 GB)

EQUAL FLEXIBILITY TO DISPATCH



- Heterogeneous queuing (hQ) defines how processors interact equally
- GPU and CPU have equal flexibility to create/dispatch work

UNLOCKING All APU GFLOPS



- Access to full potential of APU compute power

UNLOCKS THE COMPUTE POTENTIAL AND EFFICIENCY OF APUs

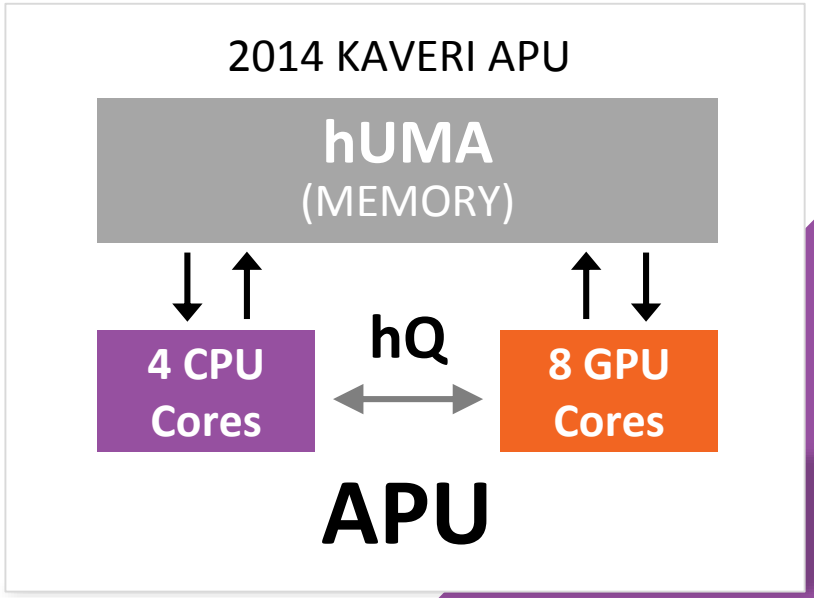
UNLEASHING DEVELOPER INNOVATION WITH HSA FEATURES

AMD

Up to

12

COMPUTE CORES



*12 Compute Cores (4 CPU + 8 GPU)

- hUMA advances how processors inside an APU **access memory**
- hQ revolutionizes how processors inside an APU **interact** with each other to handle computational tasks

WITH HSA, BOTH THE CPU AND GPU CORES ARE **COMPUTE CORES**

BENEFITS OF HETEROGENEOUS SYSTEM ARCHITECTURE

UNLEASHING THE GPU

AMD

PERFORMANCE

- Reduced latency when performing work on the GPU
- Opens up many more opportunities to accelerate application performance

POWER

- Easier utilization of the GPU leads to power reduction on parallel workloads
- Reducing software layers saves power both from instruction execution and memory copies

PORTABILITY

- Because HSA is a broadly endorsed industry standard. benefits will be available across a broad range of platforms including handheld devices, PCs and servers

PROGRAMMABILITY

- Programming models for the GPU can become equivalent to familiar models available for the CPU

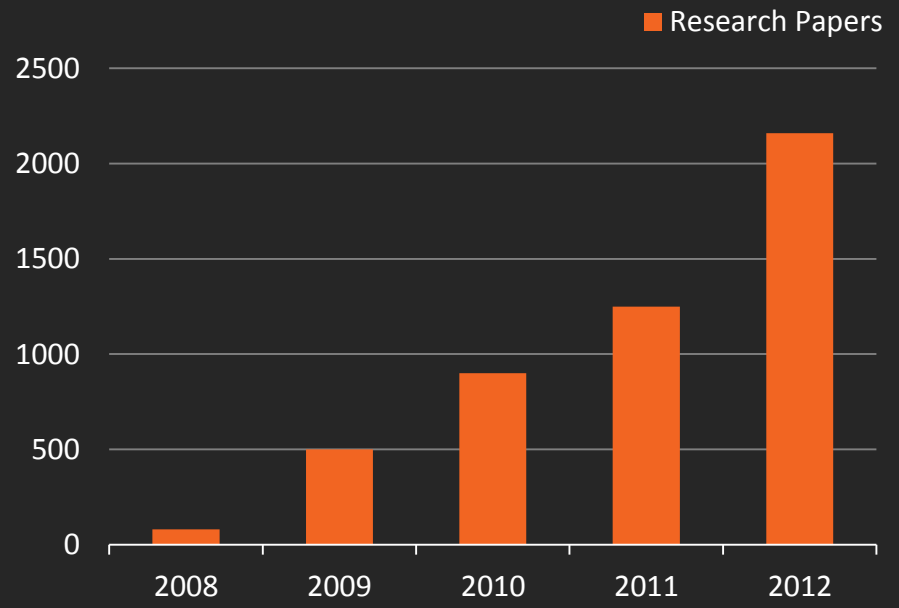

AMD
DEVELOPER
ECOSYSTEM

OpenCL, THE CURRENT STANDARD FOR HC, IS GETTING ACADEMIC TRACTION

COMPLETE UNIVERSITY KIT AVAILABLE INCLUDING

- OpenCL textbooks
- OpenCL presentation w/instructor & speaker notes, example code, & sample applications

OPENCL RESEARCH PAPERS ON GOOGLE SCHOLAR



http://developer.amd.com/Resources/library/Pages/default.aspx for of select recent OpenCL™ papers

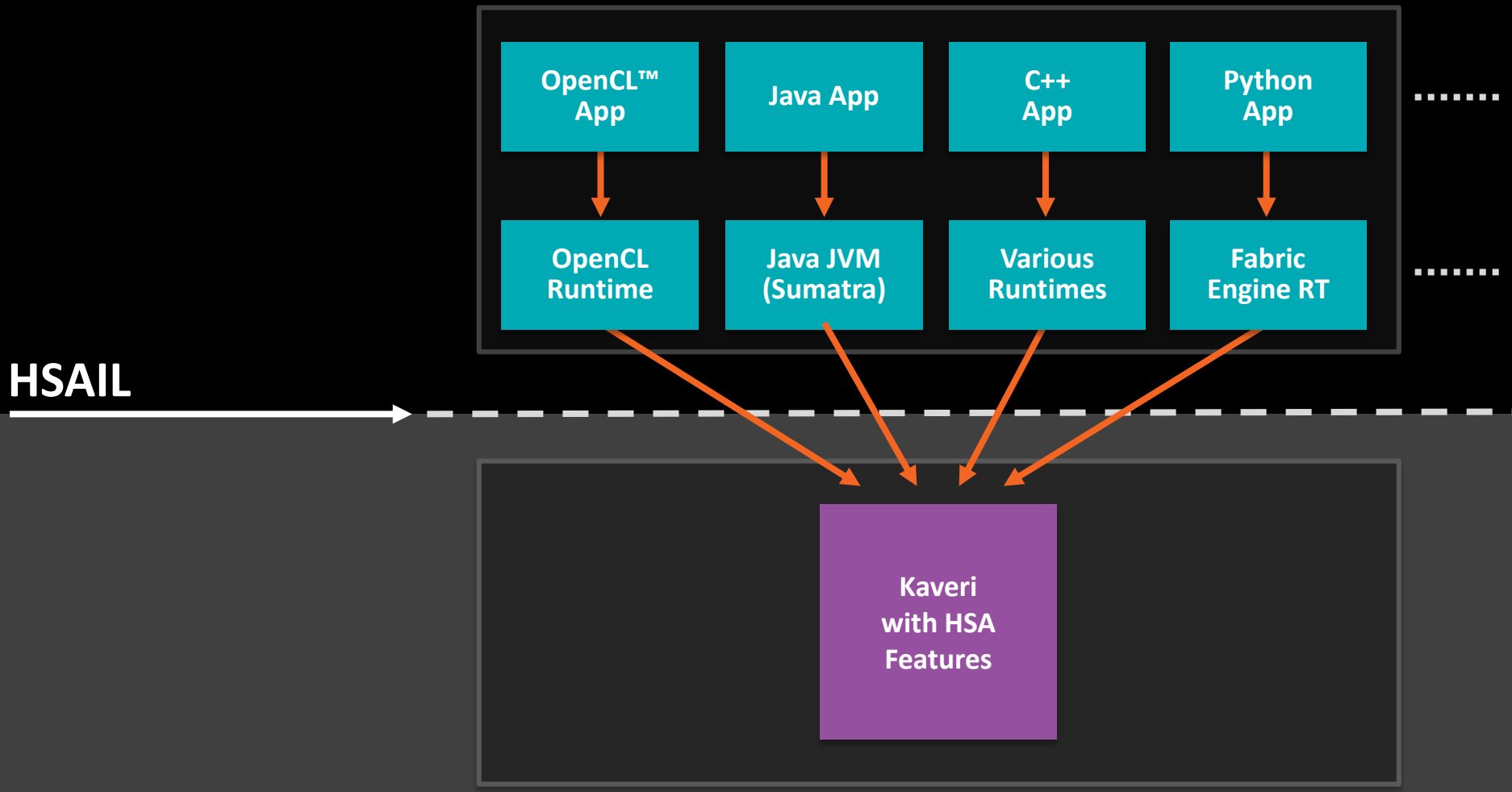
AMD FIREPRO™, OPENCL™ & APPLE MAC PRO

"This is a machine unlike anything we've ever made both inside and out."

Phil Schiller,
Senior Vice President, Apple

http://www.abs-cbnnews.com/business/06/11/13/look-radically-re-imagined-mac-pro

OTHER PROGRAMMING LANGUAGES WILL WORK WELL ON KAVERI



AMD FIREPRO™ FOR GPU COMPUTE – SW ECOSYSTEM

AMD

Developer Tools

- CodeXL
- APP SDK

Programming Languages

- OpenCL 2.0
- OpenMP 4.0
- C++
- Python
- Java

Optimized Libraries

- BLAS, FFT, RNG
- Solvers
- SpMV
- Language Specific
- Domain Specific

Comprehensive SDKs

Web Resources and Developer Forums


GAMING EFFECTS


NEW USER EXPERIENCES


SMALL FORM FACTOR


ULTRA HD MEDIA

THANK YOU

ENDNOTES

Testing by AMD Performance Labs using AMD systems consisting of either an

- AMD A10-7850K, A10-7700K A8-7600, A10-6800K, A8-6500, A8-6500T

with 2x8GB DDR3-2133 memory, 512 GB SSD, Windows 8.1, Catalyst 13.30 Beta 11

Intel systems using

- Intel Core i5 4670K, i5-4440S, i3-4330

with HD4000 graphics, 2x8GB DDR3-1600 memory, 512 GB SSD, Windows 8.1, Driver 3345 (some tests with NV GT630 with GK208 and Driver 331.93)

FOOTNOTES

AMD

1. *A GPU Core is a GCN-based hardware block containing a dedicated scheduler that feeds four 16-wide SIMD vector processors, a scalar processor, local data registers and data share memory, a branch & message processor, 16 texture fetch or load/store units, four texture filter units, and a texture cache. GPU Core can independently execute work-groups consisting of 64 work items in parallel.*

CAUTIONARY STATEMENT

This presentation contains forward-looking statements concerning Advanced Micro Devices, Inc. (“AMD” or the “Company”) including, among other things, the timing, availability, features and functionality of AMD’s A-Series APUs, which are made pursuant to the safe harbor provisions of the Private Securities Litigation Reform Act. Forward-looking statements are commonly identified by words such as "would," "may," "expects," "believes," "plans," "intends," "projects," and other terms with similar meaning. Investors are cautioned that the forward-looking statements in this presentation are based on current beliefs, assumptions and expectations, speak only as of the date of this presentation and involve risks and uncertainties that could cause actual results to differ materially from current expectations. Risks include the possibility that that Intel Corporation’s pricing, marketing and rebating programs, product bundling, standard setting, new product introductions or other activities may negatively impact the Company’s plans; that the Company will require additional funding and may be unable to raise sufficient capital on favorable terms, or at all; that customers stop buying the Company’s products or materially reduce their operations or demand for its products; that the Company may be unable to develop, launch and ramp new products and technologies in the volumes that are required by the market at mature yields on a timely basis; that the company’s third-party foundry suppliers will be unable to transition the Company’s products to advanced manufacturing process technologies in a timely and effective way or to manufacture the Company’s products on a timely basis in sufficient quantities and using competitive process technologies; that the Company will be unable to obtain sufficient manufacturing capacity or components to meet demand for its products or will not fully utilize the Company’s projected manufacturing capacity needs at GLOBALFOUNDRIES Inc. (GF) microprocessor manufacturing facilities; that the Company’s requirements for wafers will be less than the fixed number of wafers that the Company agreed to purchase from GF or GF encounters problems that significantly reduce the number of functional die the Company receives from each wafer; that the Company is unable to successfully implement its long-term business strategy; that the Company inaccurately estimates the quantity or type of products that its customers will want in the future or will ultimately end up purchasing, resulting in excess or obsolete inventory; that the Company is unable to manage the risks related to the use of its third-party distributors and add-in-board (AIB) partners or offer the appropriate incentives to focus them on the sale of the Company’s products; that the Company may be unable to maintain the level of investment in research and development that is required to remain competitive; that there may be unexpected variations in market growth and demand for the Company’s products and technologies in light of the product mix that it may have available at any particular time; that global business and economic conditions, including consumer PC market conditions, will not improve or will worsen; and the effect of political or economic instability, domestically or internationally, on our sales or supply chain. Investors are urged to review in detail the risks and uncertainties in the Company's Securities and Exchange Commission filings, including but not limited to the Quarterly Report on Form 10-Q for the quarter ended Sept. 28, 2013.