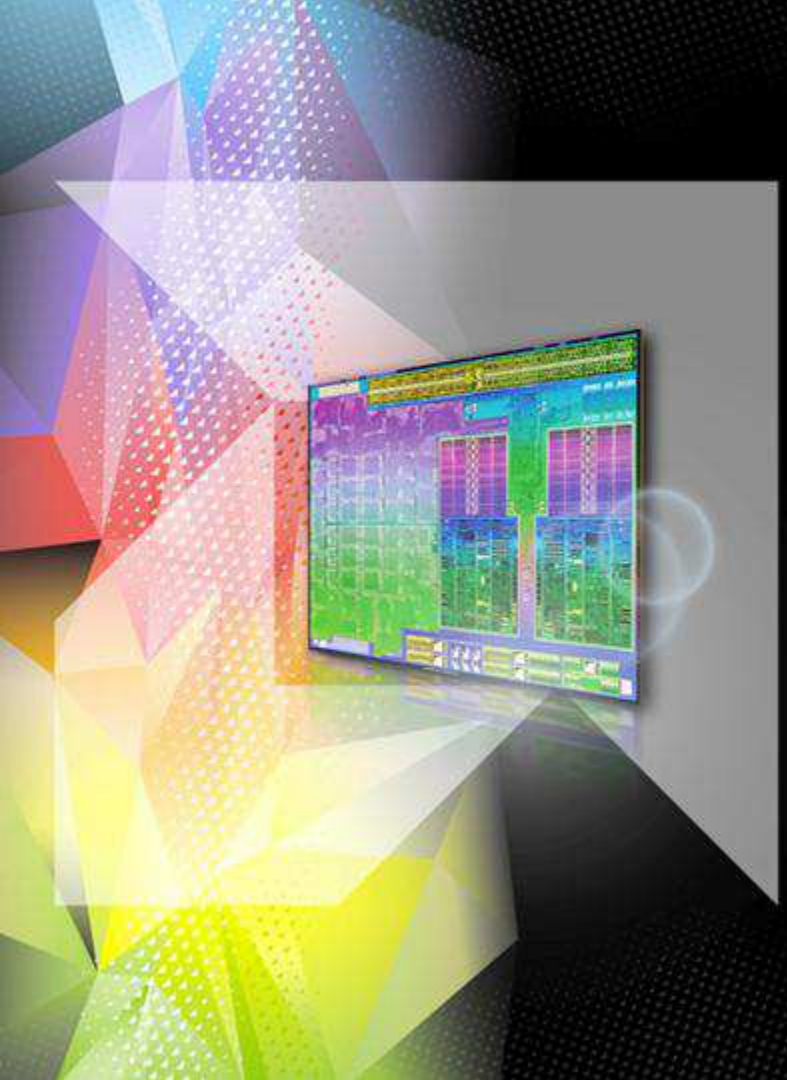


AMD

HOT CHIPS 2012

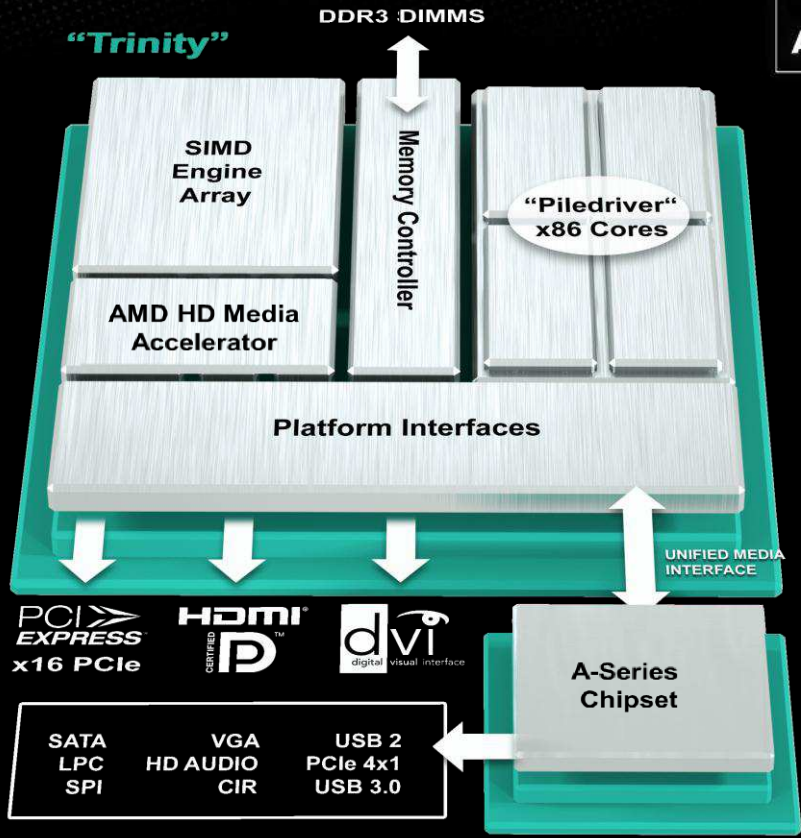
AMD “TRINITY” APU

Sebastien Nussbaum
AMD Fellow
Trinity SOC Architect

AMD APU "TRINITY" WITH AMD DISCRETE CLASS GRAPHICS

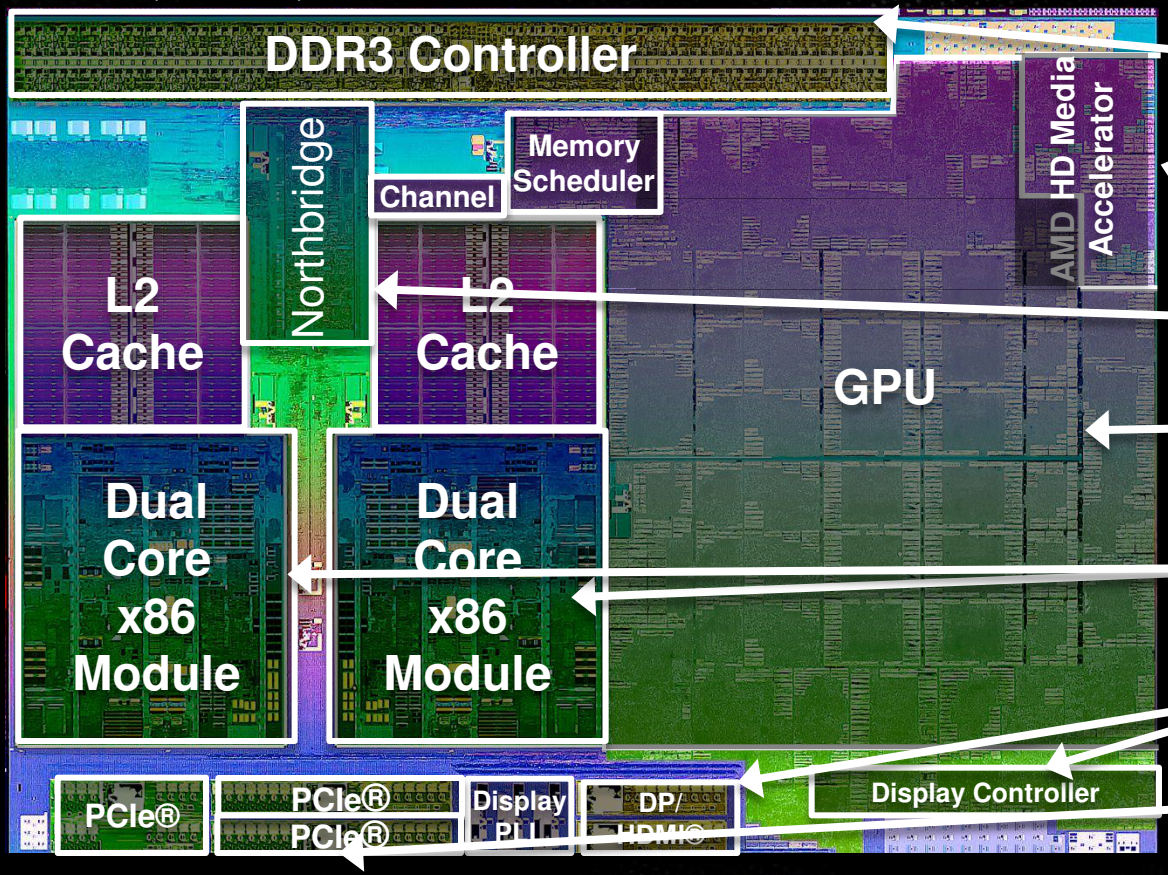
ALL NEW ARCHITECTURE FOR UP TO 50% GPU[1] AND UP TO 25% BETTER X86 PERFORMANCE[2]

- **"Piledriver" Cores**
  - Improved performance and power efficiency
  - 3rd-Gen Turbo Core technology
  - Quad CPU Core with total of 4MB L2
- **2nd-Gen AMD Radeon™ with DirectX® 11 support**
  - 384 Radeon™ Cores 2.0
- **HD Media Accelerator**
  - Accelerates and improves HD playback
  - Accelerates media conversion
  - Improves streaming media
  - Allows for smooth wireless video
- **Enhanced Display Support**
  - AMD Eyefinity Technology[3]
  - 3 Simultaneous DisplayPort 1.2 or HDMI/DVI links
  - Up to 4 display heads with display multi-streaming



"TRINITY" FLOORPLAN

32nm SOI, 246mm2, 1.303BN TRANSISTORS



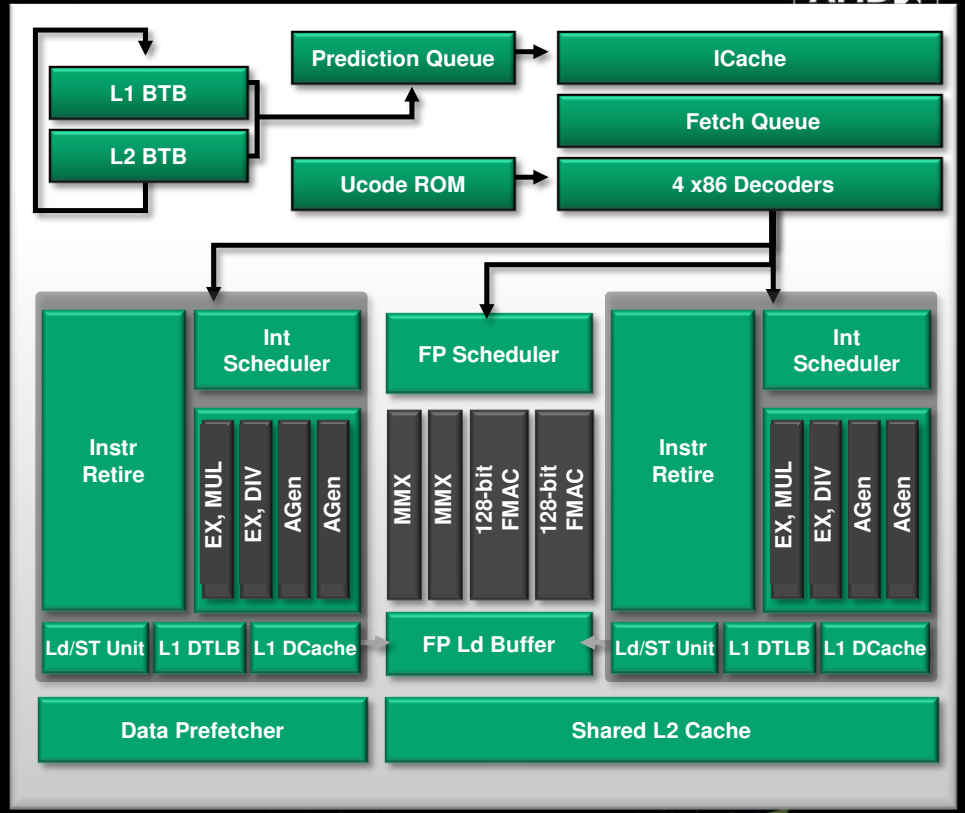
- Dual Channel DDR3 Memory Controller
- AMD HD Media Accelerator (UVD, AMD Accelerated Video Converter)
- Unified Northbridge
- AMD Radeon™ GPU
- Up to 4 "Piledriver" Cores with total 4MB L2
- HDMI, DisplayPort 1.2, DVI controllers
- PCI Express® I/O — 24 lanes, optional digital display interfaces

AMD 2ND GENERATION "BULLDOZER" CORE: "PILEDRIVER"

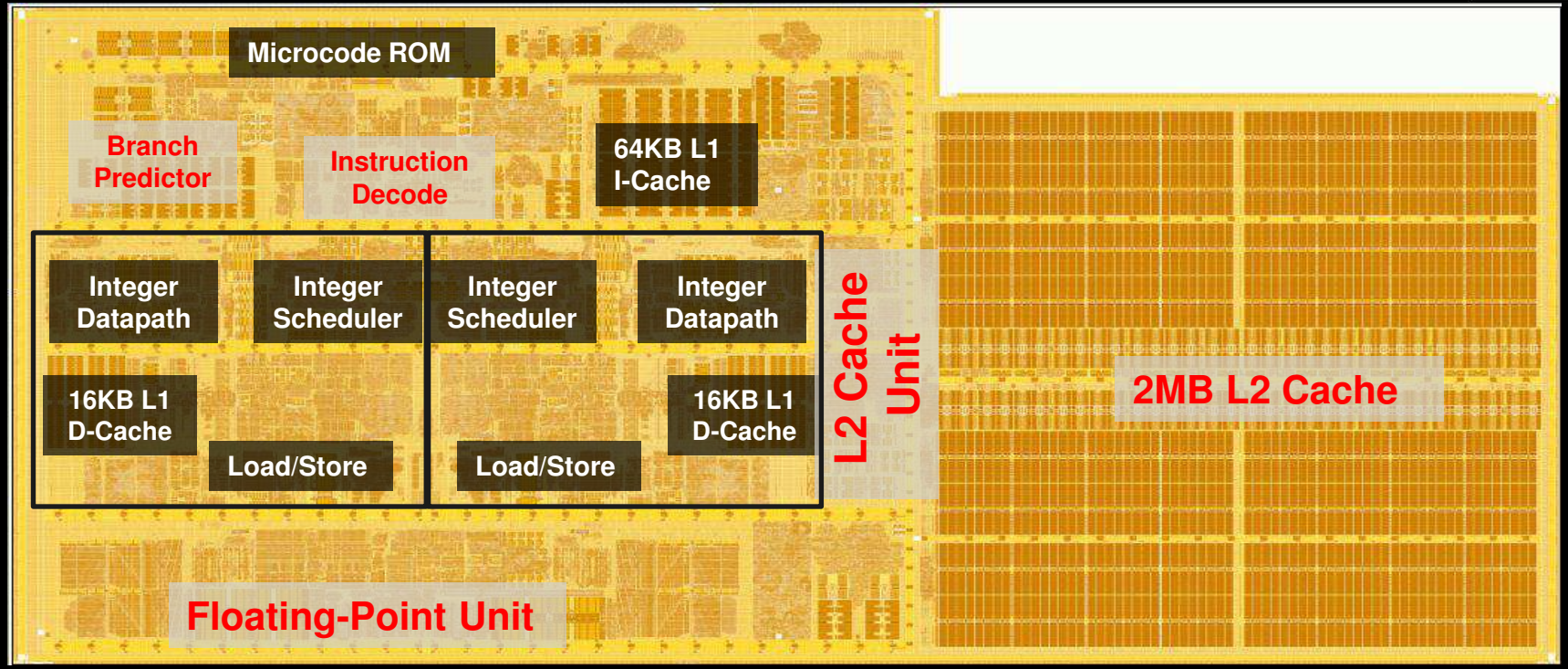
32nm "PILEDRIVER" COMPUTE MODULE

x86 CORE REDESIGN

- **Shared Fetcher / prediction pipeline - 64KB I-Cache**
- **Shared 4-way x86 decoder**
- **Shared Floating Point Unit - dual 128-bit FMA pipes**
- **Shared 16-way 2MB L2;**
- **Dedicated integer cores**
  - Register renaming based on physical register file
  - Unified scheduler per core
  - Way-predicted 16KB L1 D-cache
  - Out-of-order Load-Store Unit
- **ISA additions: FMA3, F16C**
- **Lightweight profiling support in HW**
- **"Piledriver" performance increase over "Stars"**
  - 14% improvement for desktop[5]
  - 25% improvement for notebook[2]
  - AMD Turbo Core 3.0



“PILEDRIVER” CORE FLOOR PLAN



"PILEDRIVER" IMPROVEMENTS & ENHANCEMENTS VS. "BULLDOZER"

"Bulldozer" Hybrid Predictor Augmented with 2nd level predictor

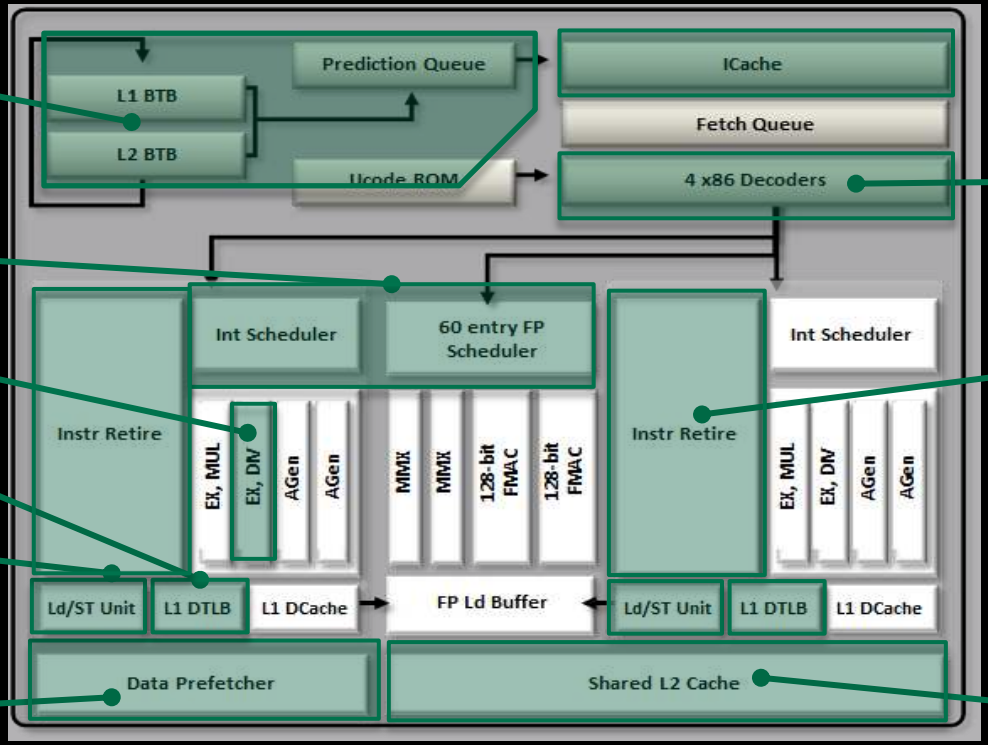
Improved scheduling FPU and INT

HW Divider

2x Larger L1 TLB

Improved Store-to-Load Forwarding

HW L1 Pre-fetcher improvements



ISA extensions FMA3, F16C

Faster Instruction exe SYSCALL & SYSRET

L2 efficiency and prefetching improvements

"PILEDRIVER" IMPROVEMENTS

30% higher CPU Freq [13]

- Design optimized for wide operational range (0.8V to 1.3V)
- 30% higher frequency at same voltage as "Stars" CPU Core in "Llano"

10% lower dynamic power vs. "Bulldozer"[13]

- Loop Predictor
- Way Predictor
- Dispatch gating based on group size
- Clock Gating
- Reduction in high power flops

Efficient operation

- 50% more base product frequency vs. "Llano" at same 35W SOC TDP

A10-4600M vs A8-3600M

Power management Latency reduction

- Intelligent L2 content tracking to speed up L2 flush
- State save/restore latency improvements to speed up power gating

MEDIA PROCESSING ACCELERATION

AMD'S UNIFIED VIDEO DECODER (UVD)

"TRINITY" ACCELERATED VIDEO CONVERTER ("AVC")

Core functionality

- Multi-stream hardware H.264 HD Encoder
- Power-efficient and faster than real-time[6] 1080p @60fps

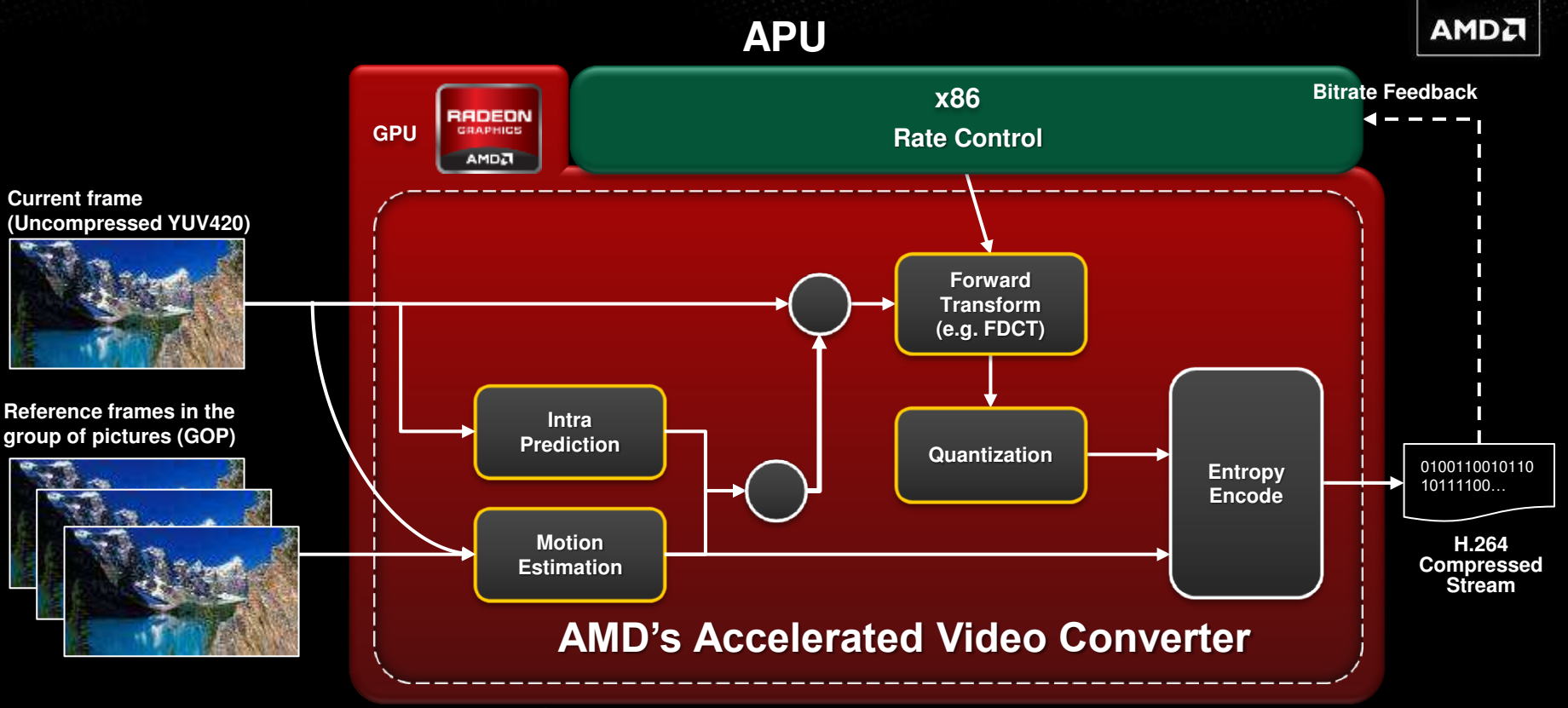
Quality features

- 4:2:0 color sampling video
- Optimizations for scene changes (games and video)
- Variable compression quality

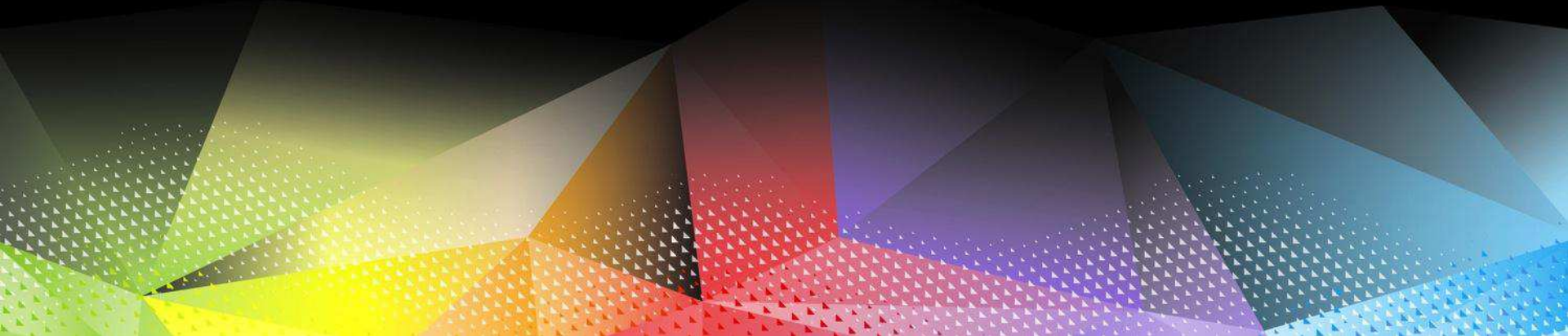
Interfacing features

- Audio / Video multiplexing
- Input from frame buffer for transcoding and video conferencing
- Input from GPU display engine for wireless display[7]

VIDEO ENCODING SYSTEM OPERATION

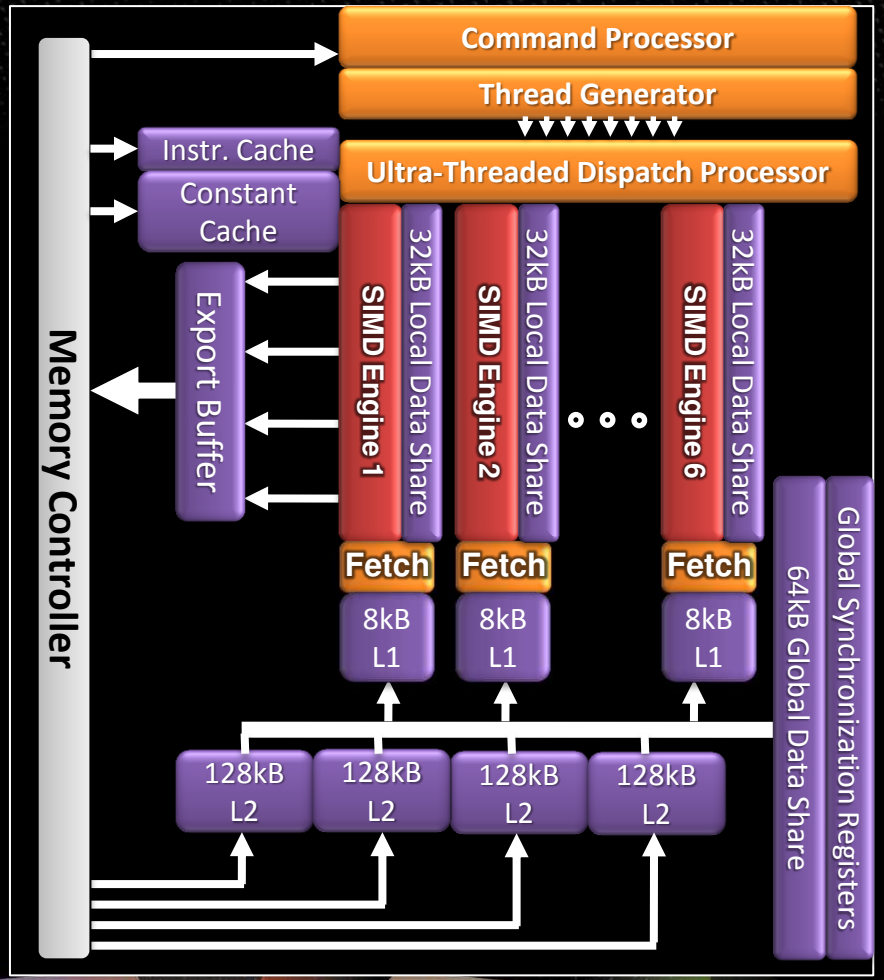


GPU DESIGN UPDATES FOR GAMING AND COMPUTE



3D ENGINE

- **DirectX® 11 – SM 5.0, OpenCL™ 1.1, DC 11**
- **GPU Core made of 384 Radeon™ Cores , each capable of 1 SP FMAC per cycle**
  - Organized as 96 stream processing units – each 4-way VLIW (vs. 5-way in Llano)
  - 6 SIMDs (each contains 16 processing units)
  - Each SIMD share 1 texture unit – achieving 4:1 ALU:Texture rate
- **32 depth / stencil per clock, 8 color per clock**
- **24x multi-sample and super sample, 16x anisotropic filtering**
- **Improved hardware tessellator vs. "Llano"**
- **Compute improvements**
  - Asynchronous dispatch: multiple compute kernels with independent address space simultaneously

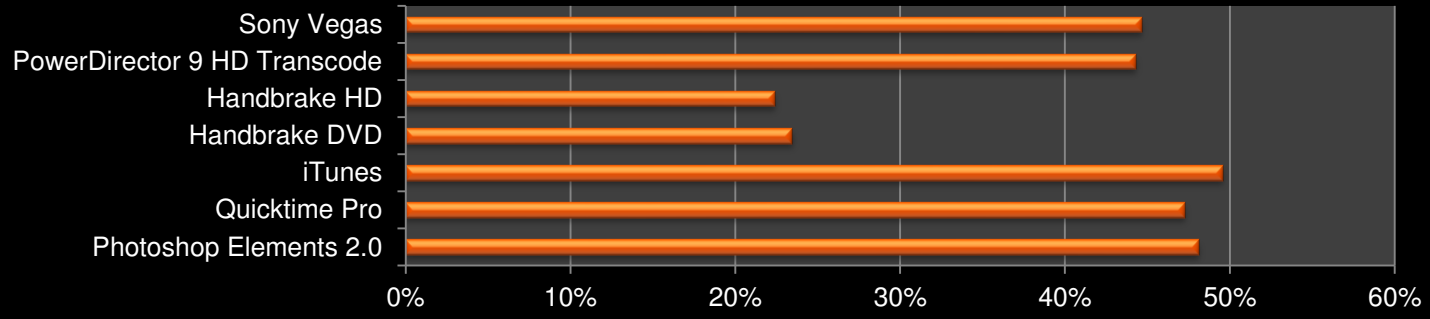


PERFORMANCE ACHIEVEMENTS

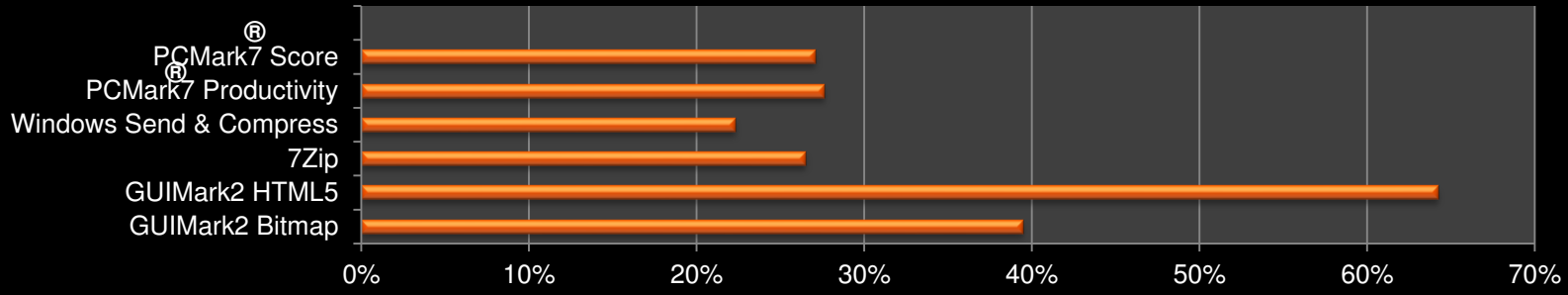
PERFORMANCE INCREASE ON CLIENT WORKLOADS (FOR 35W TDP)

"TRINITY" VS. "LLANO" CPU PERFORMANCE INCLUDING POWER MANAGEMENT, FREQ AND IPC GAINS

Digital Media

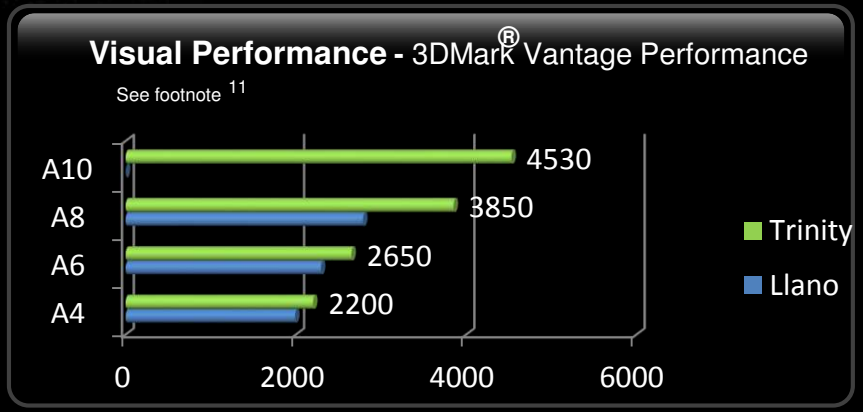


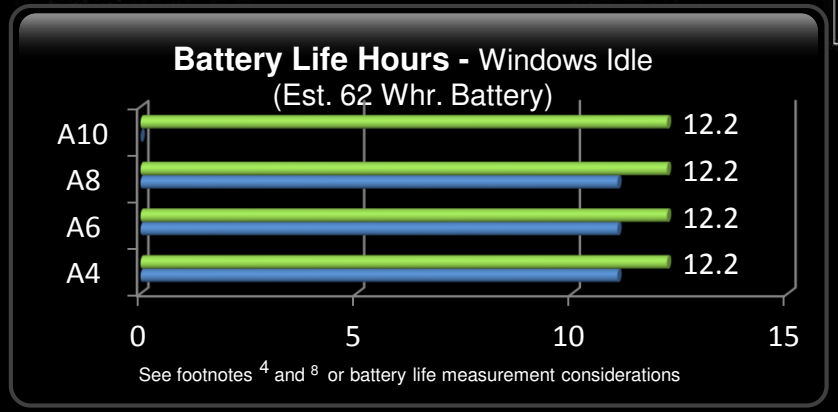
Web & Productivity: Compression & Cryptography

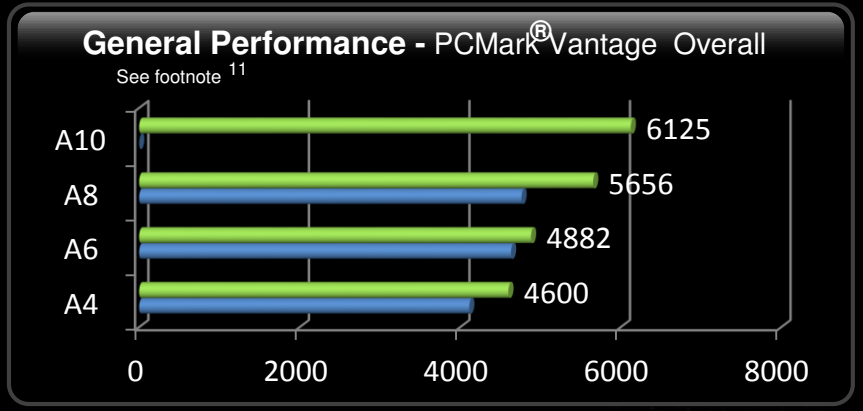


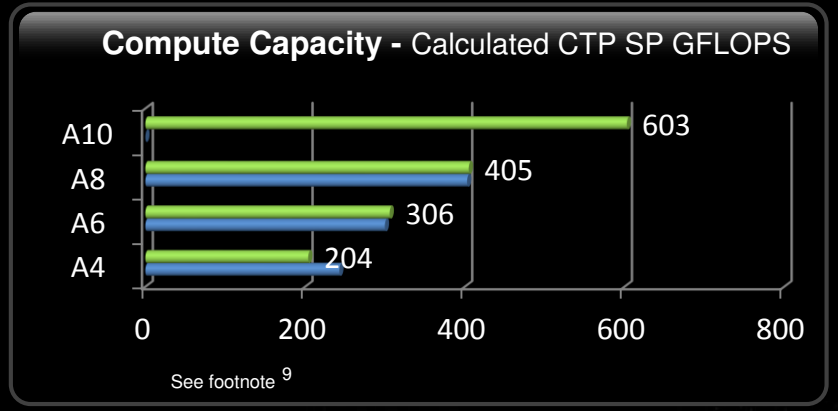
Experimental setup: see footnote 10

PERFORMANCE AND POWER COMPARISON VS PRIOR-GENERATION









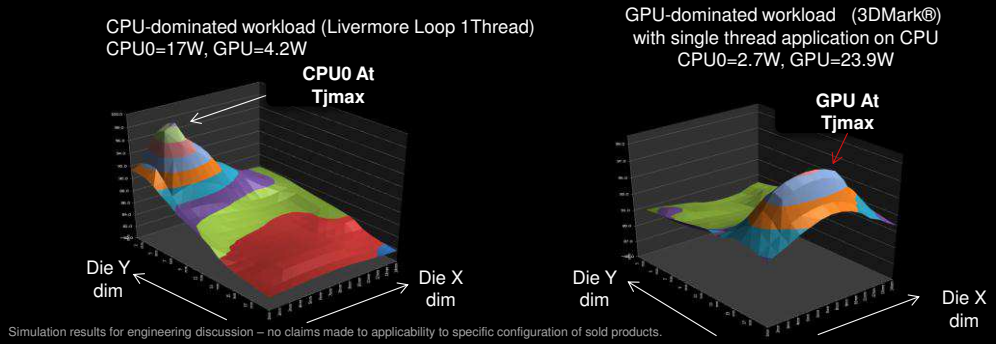
Trinity performance based on estimates and/or preliminary benchmarks and are subject to change.

AMD TURBO CORE 3.0 TECHNOLOGY

AMD TURBO CORE 3.0 TECHNOLOGY : OVERVIEW

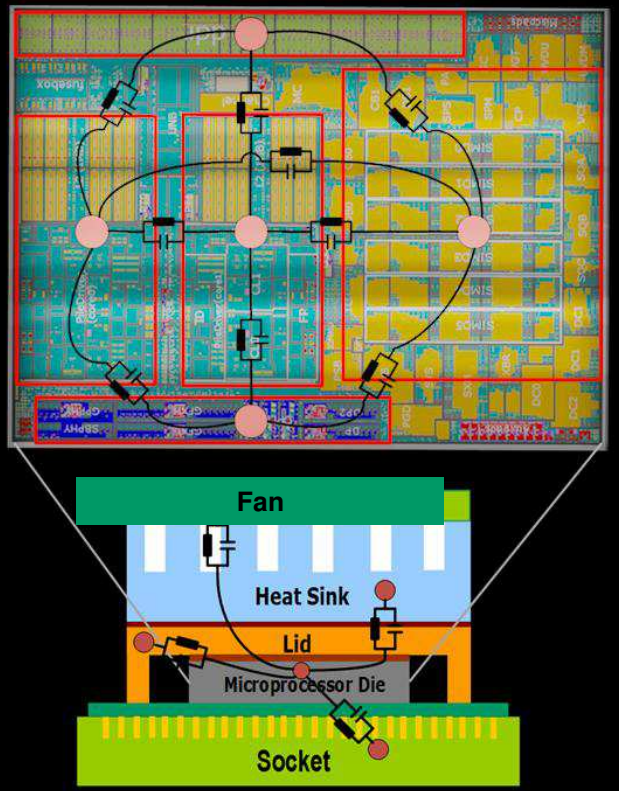
UTILIZE CALCULATED AVAILABLE DYNAMIC THERMAL HEADROOM TO IMPROVE PERFORMANCE

- **10-20 ºC variations across die during peak load**

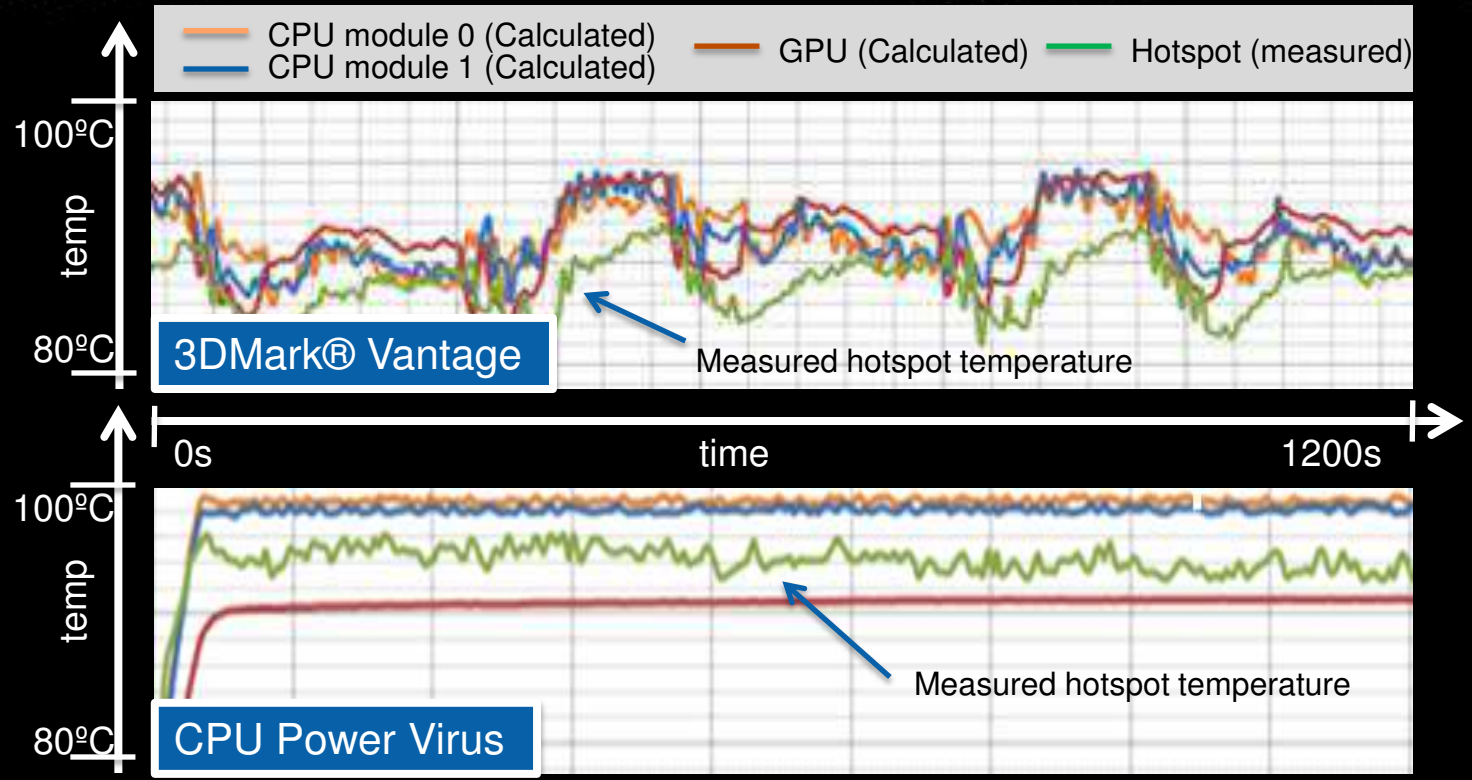


Simulation results for engineering discussion – no claims made to applicability to specific configuration of sold products.

- **Chip divided into "Thermal Entities" (TE)**
  - Thermal Entity calculate power and thermal density
- **Thermal RC network**
  - Transfer coefficients that describe thermal transfer between TEs, substrate and package are characterized
  - Numerical analysis firmware runs on the management processor which calculates per TE temperatures
  - TEs are throttled using voltage/frequency adjustments according to workload heuristics



AMD TURBO CORE 3.0 TECHNOLOGY: CALCULATED VS. MEASURED TEMPERATURE

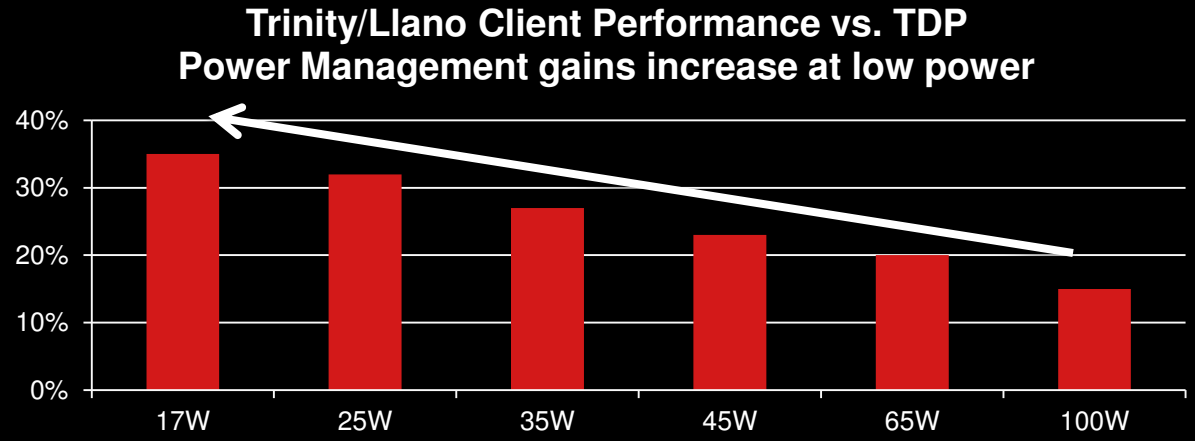


Estimated +/- 3-5C difference in calculated hotspot vs. measured hot spot temperature, at steady thermal state

Experimental results for engineering review, no observable product functional operational difference results from thermal differences. No claims made to accuracy

AMD TURBO CORE 3.0 TECHNOLOGY – PERFORMANCE

- **Workloads of moderate activity have high residency at maximum frequency**
  - Thermal headroom allows hotspot to remain below maximum control temperature
- **Higher activity workloads offer fewer opportunities to raise frequency and benefit from intelligent algorithms to bias power levels between CPU and GPU**
  - Collaborative or compute CPU/GPU applications
  - Multi-threaded workloads



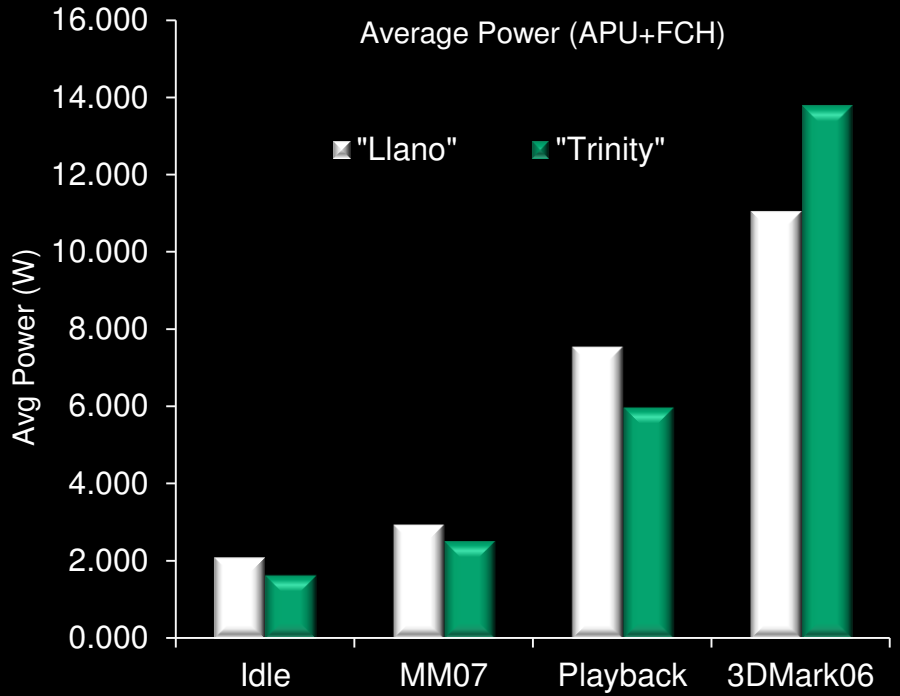
Setup information: see footnote 12

HARDWARE IMPROVEMENTS FOR LOW POWER

SYSTEM OPTIMIZATION POINTS

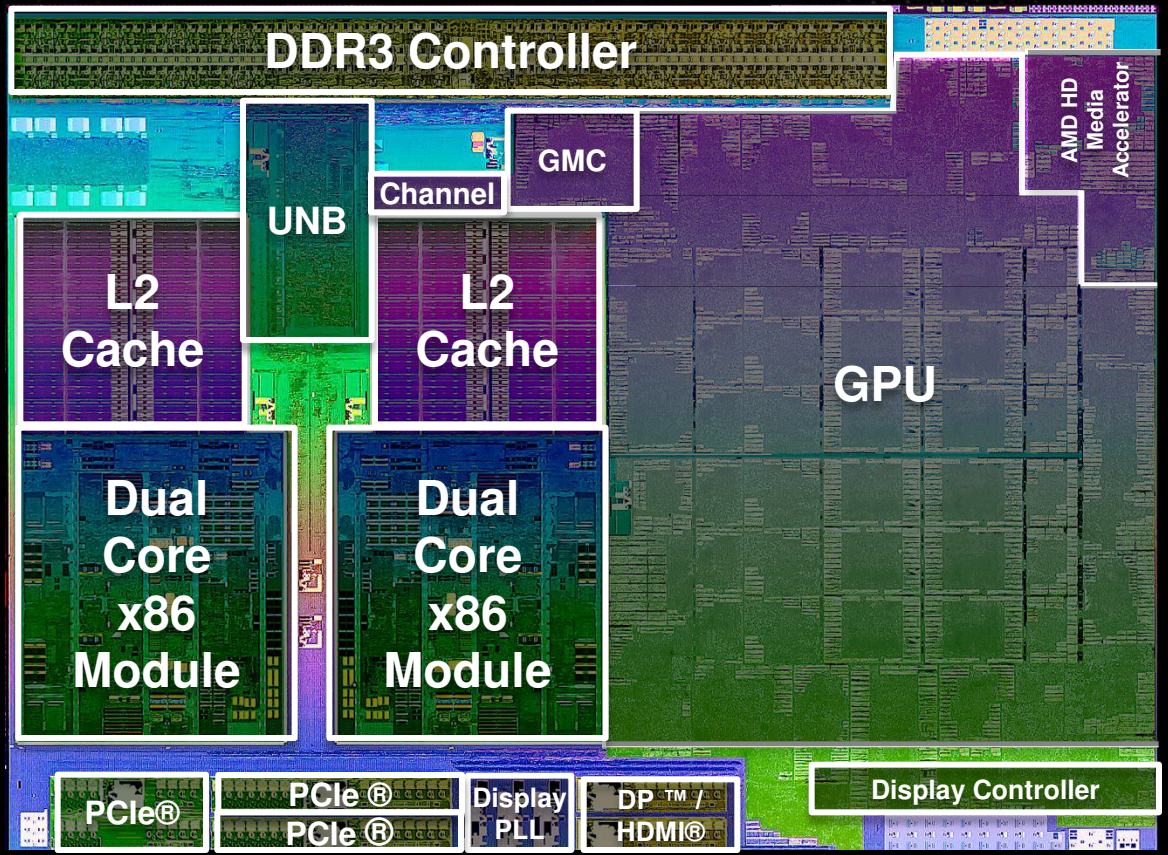
LEAP IN LOW POWER DESIGN

- **Idle — blank screen — system on**
- **MM07 — Mobile Mark 07**
- **Media playback — user experience**
- **Performance computing / gaming**
  - "Trinity" increases performance within fixed cooling solution
  - Trinity's significantly higher performance results in lower energy consumed for fixed amount of work or frames rendered, but higher power consumption during work

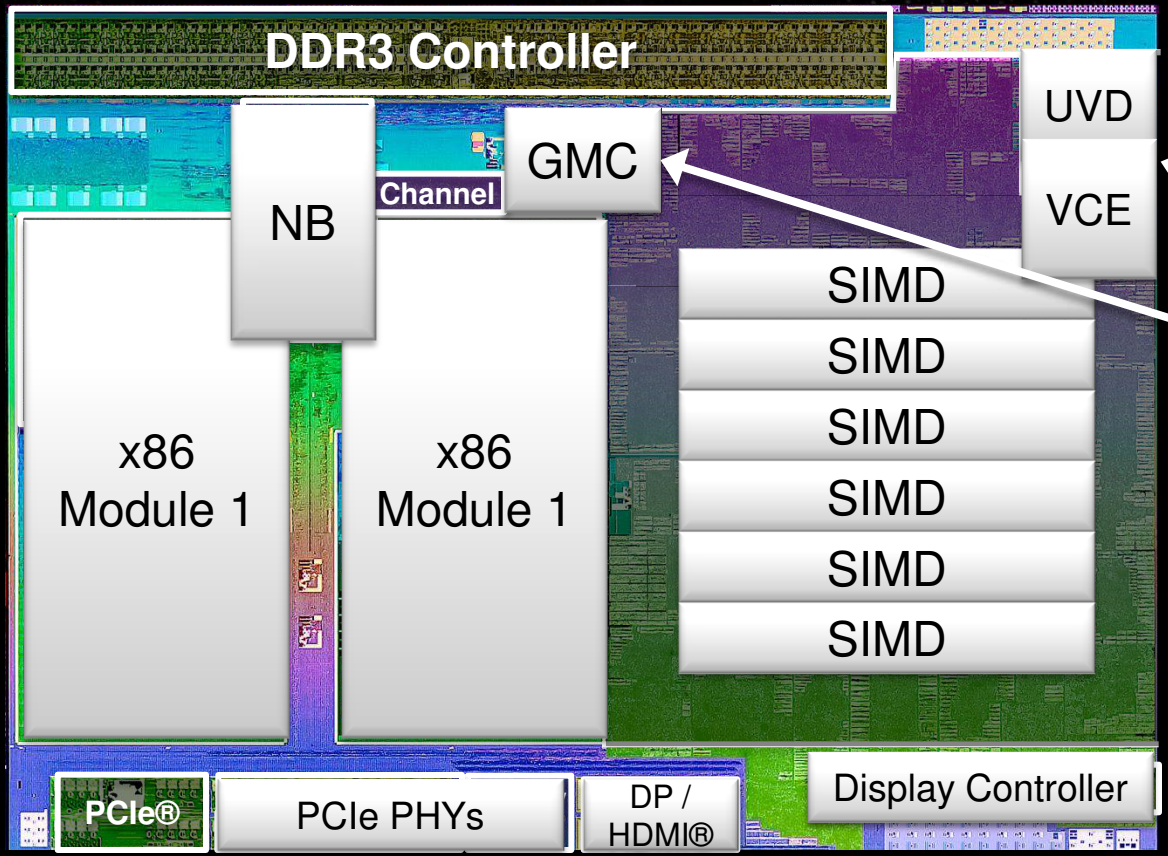


AMD A10-4600M APU on AMD "Pumori" reference board, 2x2GB DDR31600, SSD C300, Windows ® 7 64bit. Catalyst™ 8.941 vs A8-3600M,. 2x2GB DDR31600, SSD C300, Windows 7 64bit. Driver 8.941.Testing done at 1366x768. See footnotes [4] and [8] or battery life measurement considerations.

"TRINITY" APU FINE-GRAIN POWER GATING ISLANDS



"TRINITY" FINE-GRAIN POWER GATING ISLANDS (2)

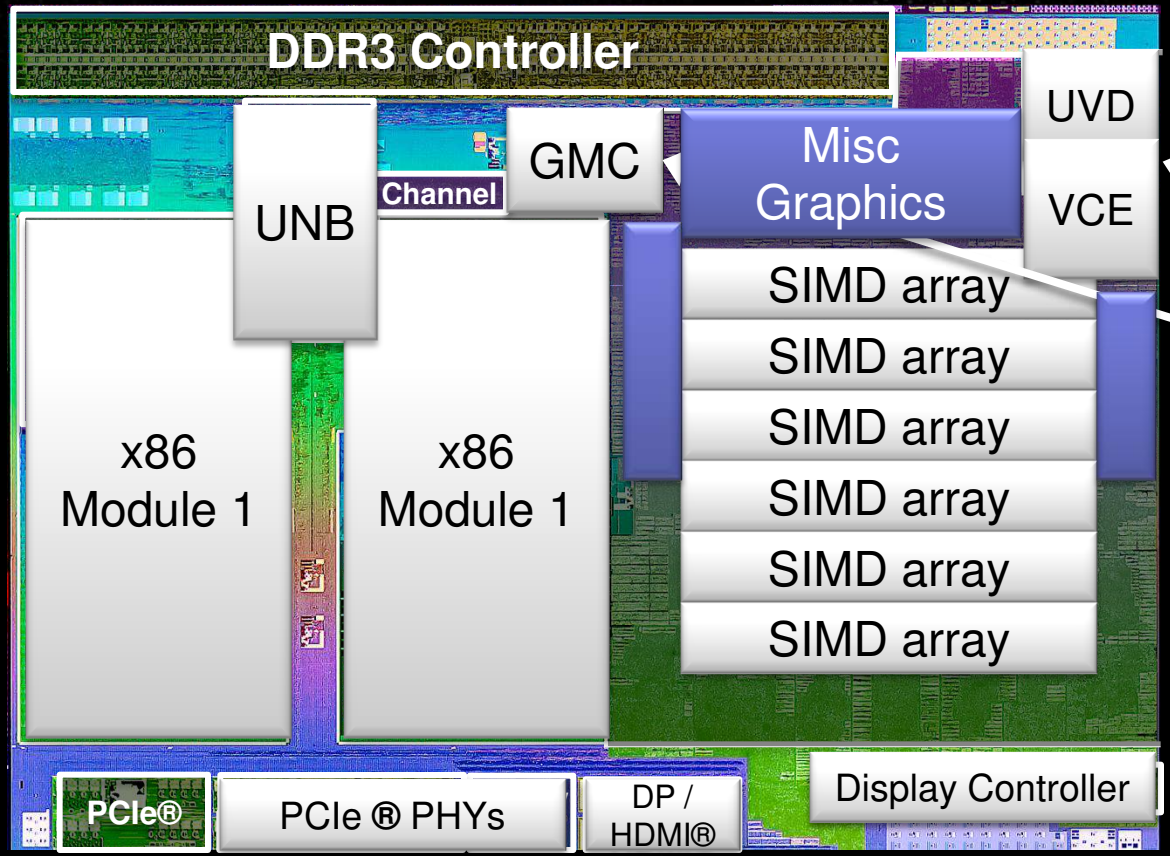


AMD HD Media Accelerator (UVD, AMD Accelerated Video Converter)

Graphics Memory Controller (Graphics optimized memory request scheduler)

Independent on-die power gated islands

"TRINITY" FINE-GRAIN POWER GATING ISLANDS (3)



AMD HD Media Accelerator (UVD, AMD Accelerated Video Converter)

Graphics Memory Controller (Graphics optimized memory request scheduler)

Independent on-die power gated islands

Additional power-down region when all graphics functions are shut down

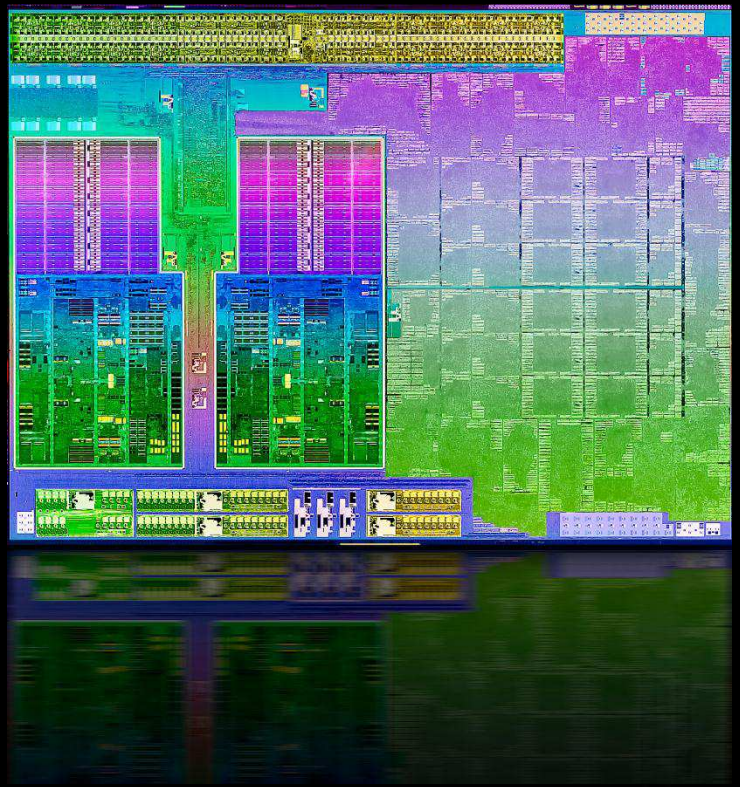
SMART NORTHBRIDGE OPERATIONAL STATES

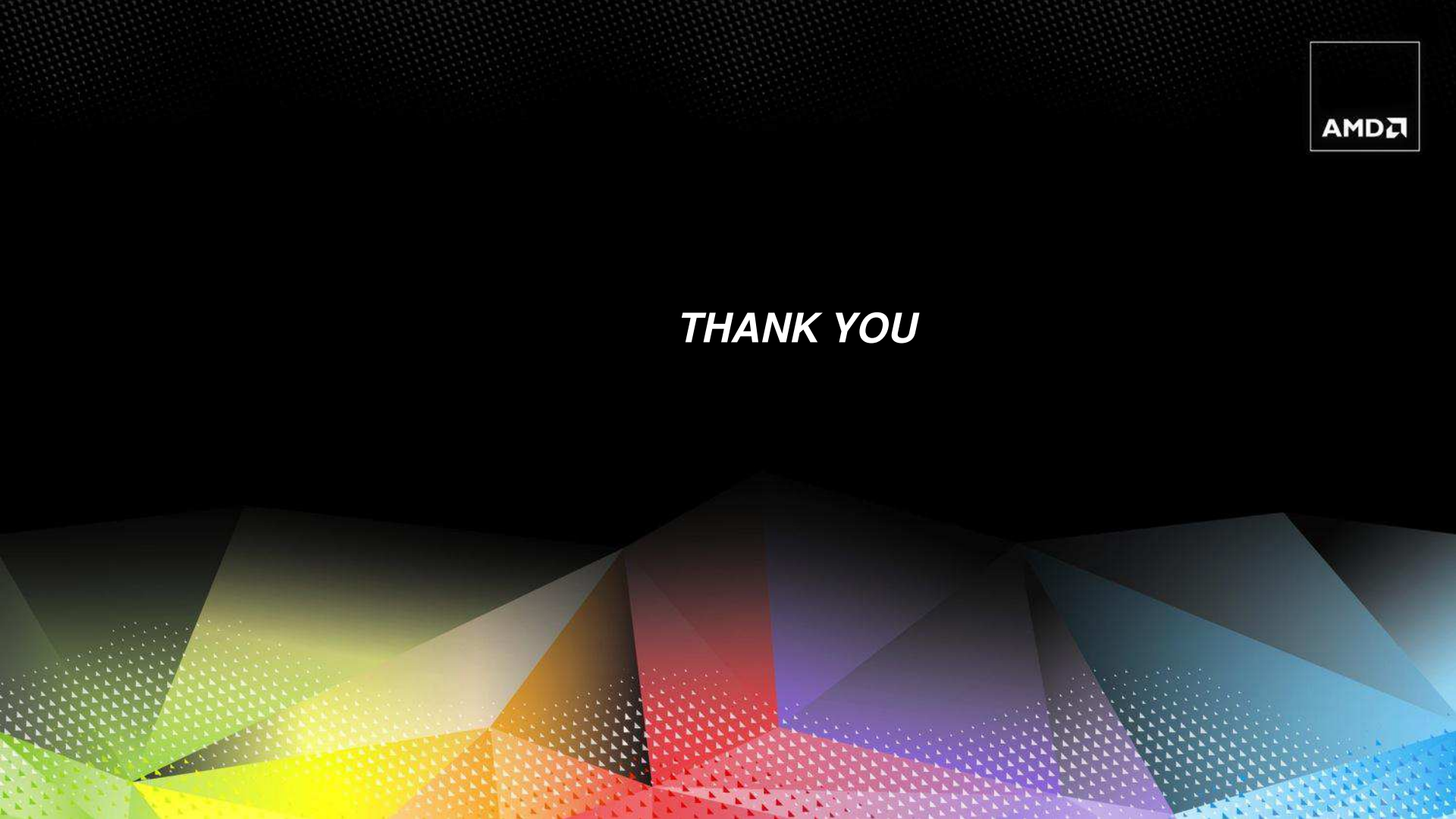
- **Goals**
  - Intelligent selection of DRAM and Northbridge frequency to meet performance and power needs
  - Additional power savings from reduced DDR termination and drive strengths at low DDR speeds
- **Design supports low-latency transitions between several operational V/F points**
  - 4 Northbridge frequencies, 2 DRAM clock rates
- **Intelligent frequency selection based on performance needs from CPU / GPU and Multi-media**
  - Memory intensive workloads and certain multi-media content types trigger switch to higher DDR speed
  - CPU intensive workloads switches to higher Northbridge frequency to improve latency to memory
  - Multi-media buffers store real-time data during low-latency switching (less than 10 µs)
- **Frequency selection is further optimized by**
  - Static user policy selection of battery or performance optimization
  - OS power management hints
  - Heuristics to ensure higher voltage and frequency will not result in additional work throttling

THANK YOU !

AMD "TRINITY" APU

- **Core redesign for greater Performance**
- **Audio and Video enhancements for the best media experience**
- **Improved GPU performance with Radeon™ Cores 2.0**
- **Low Power Leadership**




AMD
THANK YOU

FOOTNOTES

1. Testing performed by AMD Performance Labs. The score for the 2012 AMD A10-4600M on the "Pumori" reference design for PC Mark ® Vantage Productivity benchmark shows an increase of up to 29% over the 2011 AMD A8-3500M on the "Torpedo" reference design. The AMD A10-4600M APU has a score of 6125 and the 2011 AMD A8-3500M APU scored 4764.

2. Projections and testing developed by AMD Performance Labs. Projected score for the 2012 AMD Mainstream Notebook Platform "Comal" the "Pumori" reference design for 3D Mark ® Vantage Performance benchmark is projected to increase by up to 50% over actual scores from the 2011 AMD Mainstream Notebook Platform "Sabine." Projections were based on AMD A8/A6/A4 35w APUs for both platforms.

3. AMD Eyefinity technology works with games that support non-standard aspect ratios, which is required for spanning across multiple displays. To enable more than two displays, additional panels with native DisplayPort™ connectors, and/or DisplayPort™ compliant active adapters to convert your monitor's native input to your cards DisplayPort™ or Mini-DisplayPort™ connector(s), are required. AMD Eyefinity technology can support up to 6 displays using a single enabled AMD Radeon™ GPU with Windows Vista® or Windows® 7 operating systems .

4. Testing and calculations by AMD Performance Labs. Battery life calculations based on average power on multiple benchmarks and usage scenarios. These include Active metric using FutureMark® 3DMark '06 (172 min./2:54 hours), streaming YouTube video (271 min./4:30 hours), playback of a Microsoft sample clip from local HDD (303 min./5:03 hours), PowerMark ® Productivity benchmark/radio off (483 min./8:03 hours), web browsing test was average of 40 minutes via 802.11n WLAN, 2 minutes per page using the web test tool developed by AMD (570 min./9:30 hours) and Windows ® Idle (725 min./12:05 hours) as a resting metric. All battery life calculations are based on using a 6 cell Li-Ion 62.16Whr battery pack at 98% utilization for Windows ® Idle, PowerMark ® and 96% utilization for 3DMark ® 06 workload, video playback and YouTube video streaming; and 92% utilization for Blu-ray playback.

5. Projections and testing developed by AMD Performance Labs. The AMD A-10 5800K APU with AMD Radeon™ HD 7660D graphics, versus an AMD A8-3850 APU with 14% uplift on x86 performance in measure in PCMark7 ® Productivity, and 30% planned uplift on graphics performance using 3DMark ® 11 (P). All systems using "Trinity" 100W APU, 8GB DDR3-16000 memory, Windows ®7 64 bit.

FOOTNOTES (2)

6. Based on AMD internal testing of video encoding speed of VCE of 1080p H.264 video at 47 seconds, which is faster than the 65 second size of the 480p-kid.mov video file. System configuration: OS: Windows ® 7 64-bit, CPU: AMD A10-5800K with AMD Radeon™ HD 7660D graphics, Annapurna reference board, 8GB DDR3-1600, Windows ® 7 64bit.

7. AMD Wireless Display technology provides the ability to wirelessly display local screen content onto a remote screen in real time. Compliant receiver equipment required.

8. Testing and projections conducted by AMD performance labs. Testing on the 2011 AMD Mainstream Notebook Platform show 663 minutes (11.05 hours) of Windows ® Idle as "resting" battery life. Projections for the 2012 AMD Mainstream Platform "Comal" show 748 minutes (12.47 hours) of Windows ® Idle as "resting" battery life.

9. GFLOPs calculations developed by AMD performance labs measuring compute capacity for the 2012 VISION A10-based notebook which scored 603 GFLOPS. AMD GFLOPs calculated using GFLOPs = CPU GFLOPs + GPU GFLOPs = CPU Core Freq. X Core Count X 8 FLOPS + GPU Core Freq. X DirectX® 11 capable Shader Count X 2 FLOPS.

10. Experimental results on A10-4600M with Radeon HD7660G Graphics ("Trinity") vs. A8-3500M 4GB DDR3-1600 with Radeon HD6620G Graphics ("Llano") 4GB DDR3-1333 – running under Windows ® 7 Ultimate, with Hitachi HDD 5400RPM

11. Projections and testing developed by AMD Performance Labs. Projected scores for the 2012 AMD Mainstream Notebook Platform "Comal" the "Pumori" reference design for 3D Mark ® Vantage Performance, PCMark ® Vantage over actual scores from the 2011 AMD Mainstream Notebook Platform "Sabine". Projections were based on AMD A10/A8/A6/A4 35w APUs.

12. AMD A10-4600M APU with Radeon(tm) HD Graphics, 4GB DDR3-1600, on Pumori Reference Board with Hitachi 5400 RPM HDD.

13. Power measured by AMD Perf Labs on "Trinity" A0 silicon running SpecInt ® 2006 on Pumori Reference board, and on Orochi B0 (which contains "Bulldozer" Core) at same voltage and frequency. 20% dynamic power improvement was offsetted for caching structures differences and leads to an estimate of more than 10% dynamic power reduction directly attributable to the Core ® on SpecInt 2006. Frequency improvement vs. "Stars" Core measured by AMD PEO for nominal process targeting on "Llano" Rev. B0 and "Trinity" Rev. A1

DISCLAIMER & ATTRIBUTION

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions and typographical errors.

The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. There is no obligation to update or otherwise correct or revise this information. However, we reserve the right to revise this information and to make changes from time to time to the content hereof without obligation to notify any person of such revisions or changes.

NO REPRESENTATIONS OR WARRANTIES ARE MADE WITH RESPECT TO THE CONTENTS HEREOF AND NO RESPONSIBILITY IS ASSUMED FOR ANY INACCURACIES, ERRORS OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION.

ALL IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE ARE EXPRESSLY DISCLAIMED. IN NO EVENT WILL ANY LIABILITY TO ANY PERSON BE INCURRED FOR ANY DIRECT, INDIRECT, SPECIAL OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION CONTAINED HEREIN, EVEN IF EXPRESSLY ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

AMD, the AMD arrow logo and combinations thereof are trademarks of Advanced Micro Devices, Inc. All other names used in this presentation are for informational purposes only and may be trademarks of their respective owners.

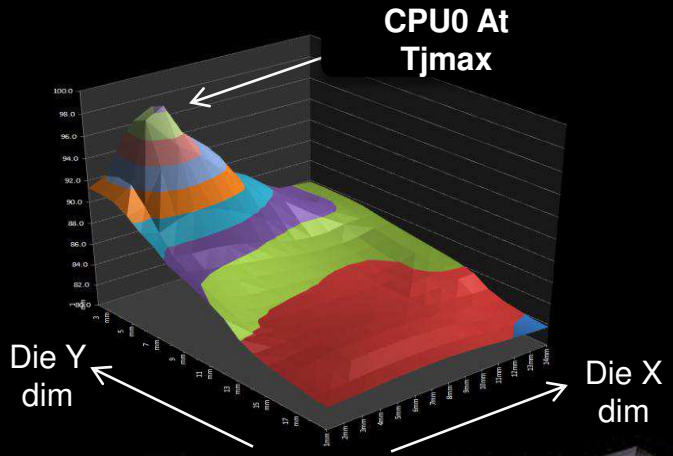
AMD, the AMD Arrow logo and combinations thereof are trademarks of Advanced Micro Devices, Inc. in the United States and/or other jurisdictions. Other names used in this presentation are for identification purposes only and may be trademarks of their respective owners.

APPENDIX

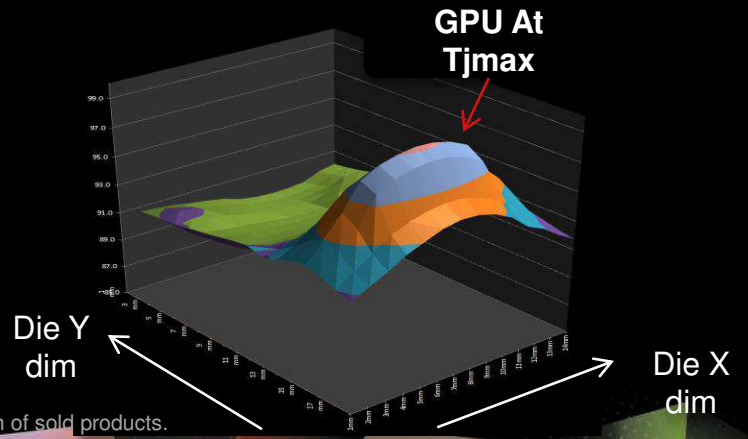
AMD TURBO CORE 3.0 TECHNOLOGY – DIE THERMAL SIMULATIONS

- **Thermal simulations for a 35W product**
  - 10-20ºC variations across the die depending on the workload, during peak activity
  - Hotspot needs to be controlled to maximum junction temperature
  - Hotspot thermal simulations are now critical part of the performance optimization flow

CPU-dominated workload (Livermore Loop 1Thread)
CPU0=17W, GPU=4.2W



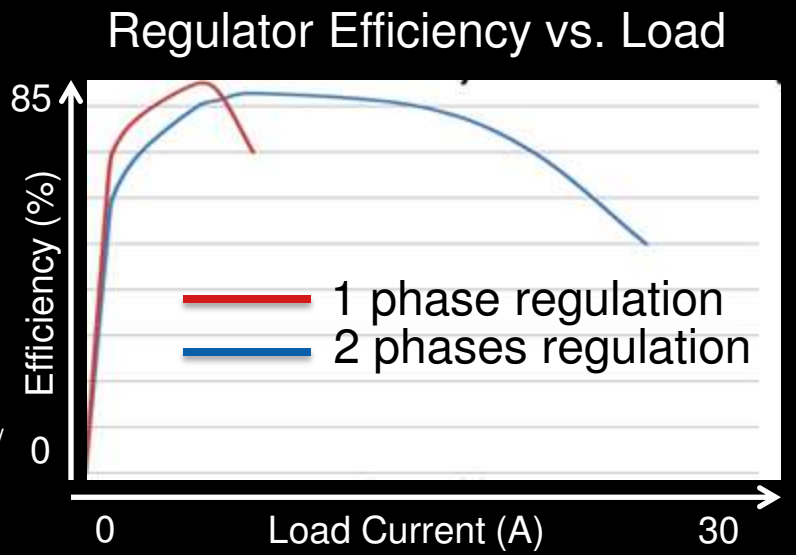
GPU-dominated workload (3DMark®) with single thread application on CPU
CPU0=2.7W, GPU=23.9W



Simulation results for engineering discussion – no claims made to applicability to specific configuration of sold products.

PLATFORM POWER SAVINGS : AMD SERIAL VOLTAGE INTERFACE 2.0

- **SVI is the interface which allows the processor to communicate information to and from voltage regulator**
- **SVI2 enables quicker power state transitions**
  - Faster data transmission rates (33Mhz)
  - New regulator response when transition is complete
  - 80+% improvement in 500mV set point change latency
- **Power efficiency features**
  - Multiple Power State Indicators sent to regulator
    - PSI0 – Current low enough that regulator can shed phases
    - PSI1 – Current low enough that regulator can use pulse skipping / diode emulation
  - Load Line trim, offset
    - Ability to adjust DC offset and load line slope based on APU state



"TRINITY" DESIGN FOR LOW POWER – NEW FEATURES

Display Power Optimizations

- Static-screen display refresh from single DRAM channel
- On-die cursor caching
- Increased on-chip buffering of display memory

Power Tuning

- Voltage and frequency are automatically selected using indication from
  - GPU Power state, PCIe ® speed, Multimedia workload
- Dynamic DRAM speed – reduced power when bandwidth requirements are low
- SVI-2 Voltage Regulator interface –selection of optimal regulator power state depending on load

Power Gating, low voltage I/O

- UNB Power Gated when idle
- GPU Core support per-SIMD power gating
- PCI-Express® and Display PHY Power Gating
- Accelerated Video Converter Power down
- Support for 1.25V DDR3 Memory

Graphics and Multimedia

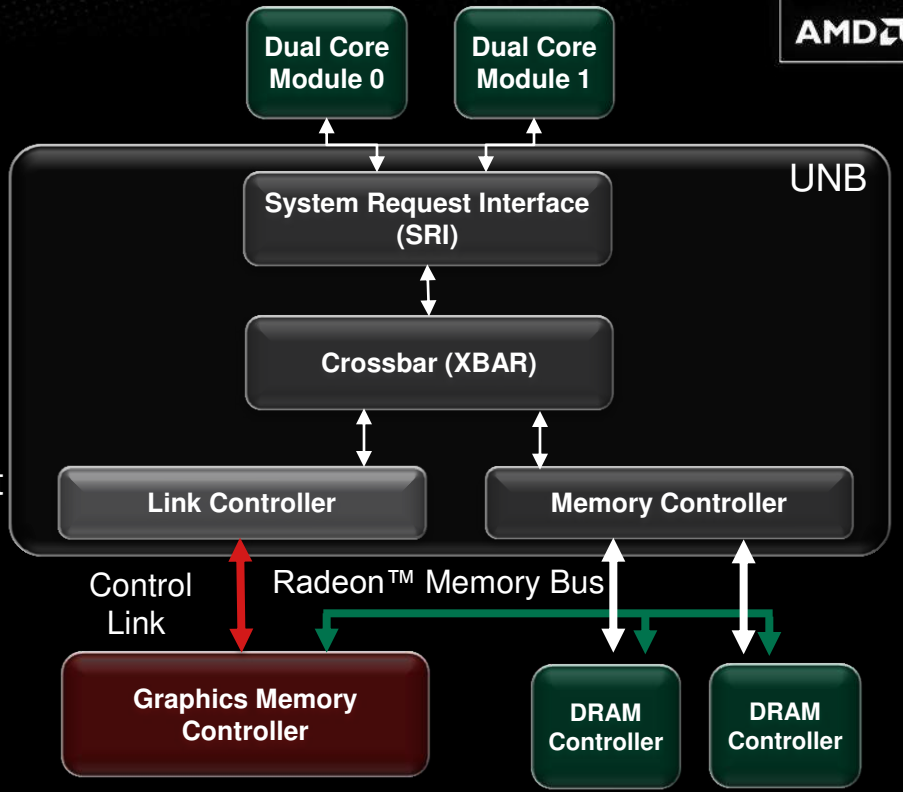
- Video Compression Engine – offload engine to save encoding power

AMD TURBO CORE 3.0 TECHNOLOGY : TEMPERATURE CALCULATION

- **CPU/GPU Temperature**
  - Firmware regularly calculates instantaneous temperature for each TE new power estimate and prior temperature
  - Uses a 5 stage thermal RC ladder
- **Other silicon contributors**
  - High Speed IO interfaces, Northbridge are modeled as power and/or temperature offsets to simplify calculations
  - This has limited impact on accuracy
- **Measured error of +/-5 ºC on 3DMark® analysis**
- **Algorithm provides deterministic operation and reproducibility of results**

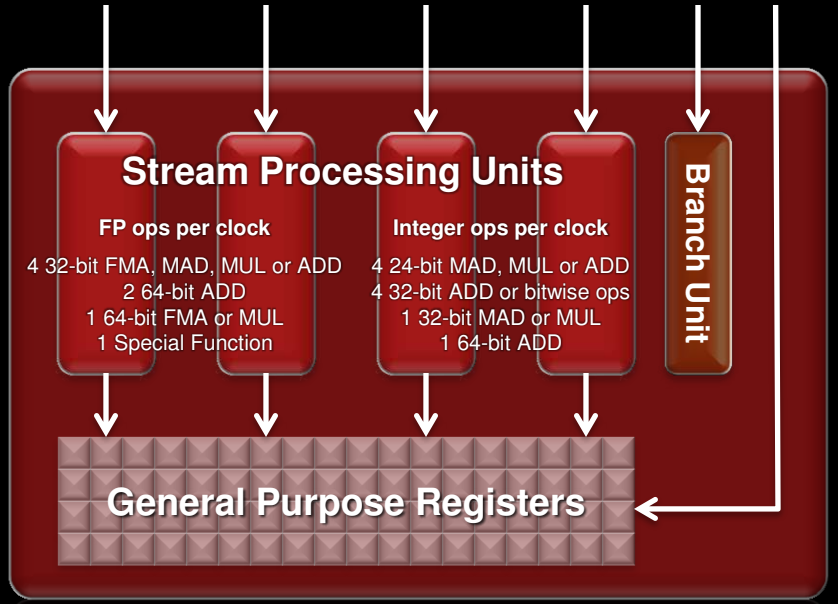
UNIFIED NORTHBRIDGE AND MEMORY

- **Unified Northbridge**
  - First UNB for APUs featured on “Trinity”
  - Supports:
    - Interface to a Graphics Memory Controller
    - Two DDR2/3 interfaces, shared with the Graphics Northbridge via Radeon Memory Bus
    - APU Power Management
- **Memory Support**
  - 128-bit interface arranged as two un-ganged 64-bit channels
  - Supports Memory P-states — with memory speed changes on the fly
  - Supports 1.25V DIMMS
  - Up to 29.8 GB/s with DDR3-1866



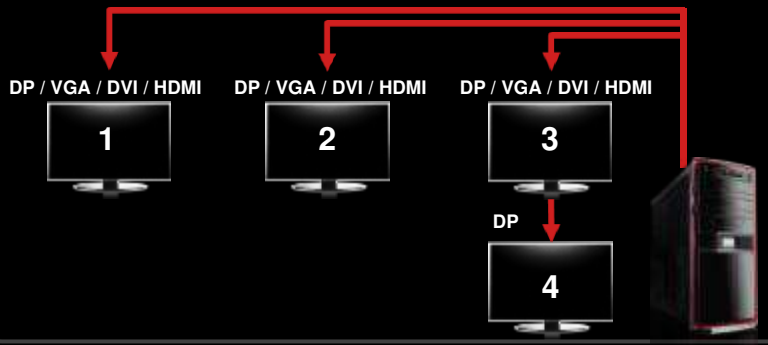
RADEON CORE S 2.0 DESIGN

- **VLIW4 thread processors**
  - 4-way co-issue
  - All stream processing units now have equal capabilities
  - Special functions (transcendentals) occupy 3 of 4 issue slots
- **Allow better utilization than previous VLIW5 design**
  - Improved performance/mm²
  - Simplified scheduling and register management
  - Extensive logic reuse

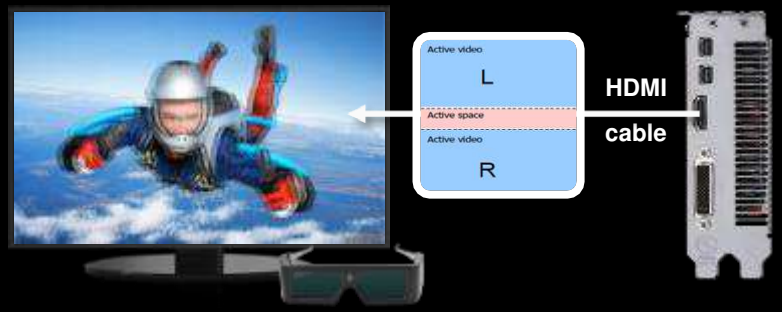


DISPLAY TECHNOLOGY LEADERSHIP

DP1.2



HDMI



Stereo 3D

Color Correction