

14 10

2.0





A1100 High Performance 64-bit ARM[®] CPU with Integrated Dual 10Gb Ethernet

8 Cortex[®]-A57 Cores **8**MB L3 Cache Up to **128**GB System Memory

2X10GbE Highly Integrated IO

10 Years Planned Longevity

AMD OPTERON[™] A1100 SERIES PROCESSOR ACCELERATING DATA CENTER INNOVATION



COMPUTE

- ✓ Up to 8 ARM Cortex[®]-A57 cores
- Up to 4MB shared L2 cache
- 8MB L3 cache

MEMORY

Dual channel DDR3/4 w/ECC up to 1866 MHz

- SODIMM, UDIMM, RDIMM 2DIMMs/channel
- Memory capacity 64GB/channel

INTEGRATED I/O

- PCIe[®] Gen3 | 8 lanes (x8, x4, x2 support)
- ▲ SATA3 (6Gb/s) | 14 ports
- ▲ 10GbE (KR) | 2 ports

OTHER INTEGRATION

- ARM TrustZone[®] compliant
- System co-processor (Secure boot)

AMD OPTERON[™] A1100 SERIES TARGET MARKETS ACCELERATING DATA CENTER INNOVATION



NETWORKING	STORAGE	WEB SERVING	SOFTWARE DEVELOPMENT
Multi-core ARM 64-bit SoC with dual integrated 10GE controllers	14 port SATA3 integration with balanced I/O, memory, and compute for large data sets	Power-efficient Cortex-A57 cores enabling dense form factors for scale-out workloads	Single, low cost platform for 64-bit ARM SW development

5 | PCCC WORKSHOP | FEB. 19 2016

NEXT-GENERATION, HIGH-PERFORMANCE X86 CORE







- Totally New High-performance
 Core Design
- Simultaneous Multithreading (SMT) for High Throughput
- New High-Bandwidth, Low Latency Cache System
- Energy-efficient FinFET Design Scales from Client to Enterprise-class Products
- Planned Samples Availability 2016

AMD STRATEGY FOR THE DATACENTER

THE INGREDIENTS

- ▲ High-Performance x86 and ARM[®]
- ▲ Industry-Leading & Most Efficient GPUs^{1,2}
- Scalable Designs
- Memory Innovation
- Open Approach to Heterogeneous Systems

THE OPPORTUNITY

- ▲ Offer Choice in \$10B+ Server Silicon Market
- ▲ \$4B+ Networking Market Moving to Software
- ▲ \$2B+ Storage Market Moving to Server
- Lead in HPC and Machine Learning
- Work Closely with Partners and Customers



Heterogeneous Systems Architecture

SOFTWARE ECOSYSTEM





THE UNIQUE VALUE OF AMD FIREPRO[™] S-SERIES FOR HPC AMD FIREPRO[™] S9100, S9150, S9170 GPUS

Serious compute performance

Over 2.0 TFLOPS delivered DGEMM

Performance/watt leadership

Top GPU on the Green500[™] List*

Largest GPU memory

2.7X per GPU with AMD FirePro[™] S9170 vs K40/K80

Embrace of open standards

OpenCL[™] 2.0



AMD FIREPRO[™] S-SERIES HPC FOCUS SEGMENTS ENABLED BY STREAM TECHNOLOGY

Academic & Government Clusters

Full throughput double precision performance High bandwidth GDDR5 memory with ECC support Large 16GB/32GB memory footprint OpenCL[™] 2.0

Oil & Gas

Outstanding single precision for seismic processing Leadership double precision for reservoir simulation Professional grade FirePro[™] boards, designed by AMD OpenCL 2.0

Deep Neural Networks / Machine Learning

Outstanding single precision performance Large 16GB/32GB memory footprint OpenCL 2.0, Torch 7



OPEN PROGRAMMING FRAMEWORK WITH OPENCL™



Fully featured implementation, supporting OpenCL 2.0

Supported by extensive resources on AMD Developer Central

- OpenCL Zone http://developer.amd.com/tools-and-sdks/opencl-zone/
- Tools, SDKs, libraries, programming guides, forums

Extensive network of 3rd party experts

- Can provide advanced training, hands on, classroom, and on site engineering training for OpenCL[™] development
- http://developer.amd.com/partners/training-partners/
 - ArrayFire (AccelerEyes)
 - Acceleware
 - DevelopIntelligence
 - Mindshare
 - StreamComputing
 - RunTime Computing Solutions



HPC Software Innovation from AMD Boltzmann Initiative



Enabling today. Inspiring tomorrow.

HSA FOUNDATION





Programming models for GPU <u>can</u> become equivalent to familiar CPU models

- All processors use the same addresses and access all virtual and physical memory
- Extend multicore coherency to the GPU and other accelerators
- Pass work quickly between the processors
- ▲ Enable acceleration in Java, C++ AMP and Python



LISTENING TO CUSTOMERS AND DIVING DEEP INTO THE CODE

Boltzmann Initiative

- Builds on AMD's 4+ years in research and development effort in HSA
- Dives deep into customer needs on Deep Neural Networks, molecular dynamic, bioinformatics chromodynamics, oil and gas, rendering, FEA and more
- Evolves driver/runtime platform to address latency, system level performance, scale up and scale out, ease of use and manageability
- Addresses customer request for ISO Standard Languages and support existing GPU Code



Building Open Platform for HPC Developer Community Innovation

THE AMD "BOLTZMANN INITIATIVE"

- Server-based GPU computing unifies computation foundation (APU and GPU) with proven heterogeneous system architecture (HSA) technologies
- Heterogeneous Compute Compiler (HCC) built on "Boltzmann Initiative" allows expression of ideas via C++ 11/14, C11 and OpenMP
- HIP Runtime (with HCC) to embrace current CUDA community, provides additional outlet for ideas and creativity with minimal effort

Efforts designed to better engage programming community

BUILDING A RICH HETEROGENEOUS PLATFORM FOR HPC INNOVATION

Announcing evolution of HSA to support dGPU and rich programming models to HPC cluster class computing

- Headless Linux[®] 64-bit Driver and HSA+ Runtime
 - ▲ Unified Address Space for CPU and GPUs
 - ▲ Optimized for lower latency dispatch and PCIe[®] data transfers
 - ▲ Large BAR
- Multi-GPU optimized for Peer-to-Peer in-node and out-of-node via RDMA via Infiniband, RoCE, and iWARP
- System Management Controls

Bringing a New Level of dGPU Platform Performance to HPC Servers

AMD ACTIVITY

▲ AMD demonstrated at SC15

- New HSA compliant C++ compiler, HIP porting tools and open source HSA programming suite
- Systems from Dell, HP and SuperMicro featuring AMD FirePro[™] GPUs running demonstrations such as AMD FireRender, Abaqus, and TUM
- Technical talks by AMD and industry experts on the future of HPC
- AMD Opteron[™] A1100 series ARM processor in SoftIron's new Enterprise Class Overdrive 3000 system for developers, and Silver Lining System's ARM-metal-as-a-Service (AMAAS) solution
- AMD FirePro[™] S9170 server GPU with Supermicro and One Stop System servers to accelerate large dataset computations

AMD demonstrations, exhibit and technical talks in booth #727

LAUNCH GPUOPEN!!

library for DLA computi

host processors. It pro-

Launched GPUOpen, which unveils the Radeon graphics drivers and software development tools.

A Tools

PROFESSIONAL COMPUTE

– http://gpuopen.com/

GPUOpen



programming framewor

Libraries

Se Infrastructure

III Applications

library provides a simple interface

for computing FFTs on GPU.

Summary

- Focused commitment to meeting customer expectations in HPC with Heterogeneous Compute
- Putting in place a system level optimized Headless Linux Driver/Runtime to create a platform for HPC innovation
- HIP Bringing hardware choice to CUDA application developers
- Heterogeneous Compute Compiler brings a foundation for developers to focus on science not device and host code integration

FOOTNOTES

- 1. AMD FirePro[™] S9150 max power is 235W and delivers up to 2.53 TFLOPS peak double and up to 5.07 peak single precision floating point performance. Nvidia's highest performing server cards in the market as of June 2014 are the Tesla K40, max power of 235W, with up to 1.43 TFLOPS peak double and up to 4.29 peak single, and the K10, max power 225W, with up to 4.58 TFLOPS peak single and 190 GFLOPS peak double precision. Visit <u>http://www.nvidia.com/object/tesla-servers.html</u> for Nvidia product specs. FP-97
- 2. AMD FirePro[™] S9150 features 16GB GDDR5 memory, and Nvidia's highest performance server GPU in the market as of June 2014 is the Tesla K40 with 12GB GDDR5 memory. Visit <u>http://www.nvidia.com/object/tesla-servers.html</u> for Nvidia product specs. FP-98
- 3. Testing conducted by AMD engineering on the AMD Radeon[™] R9 290X GPU vs. an HBM-based device. Data obtained through isolated direct measurement of GDDR5 and HBM power delivery rails at full memory utilization. Power efficiency calculated as GB/s of bandwidth delivered per watt of power consumed. AMD Radeon[™] R9 290X (10.66 GB/s bandwidth per watt) and HBM-based device (35+ GB/s bandwidth per watt), AMD FX-8350, Gigabyte GA-990FX-UD5, 8GB DDR3-1866, Windows 8.1 x64 Professional, AMD Catalyst[™] 15.20 Beta. HBM-1
- 4. Language for 4 here is: Testing conducted by AMD engineering on the AMD Radeon™ R9 290X GPU vs. the AMD Radeon™ R9 Fury X GPU. Data obtained through isolated direct measurement of GDDR5 and HBM power delivery rails at full memory utilization. AMD Radeon™ R9 290X and R9 Fury X GPU, AMD FX-8350, Gigabyte GA-990FX-UD5, 8GB DDR3-1866, Windows 8.1 x64 Professional, AMD Catalyst™ 15.20 Beta. HBM-3

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. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes.

AMD MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION.

AMD SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL AMD BE LIABLE TO ANY PERSON FOR ANY DIRECT, INDIRECT, SPECIAL OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION CONTAINED HEREIN, EVEN IF AMD IS EXPRESSLY ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

ATTRIBUTION

© 2015 Advanced Micro Devices, Inc. All rights reserved. AMD, FirePro, Opteron, the AMD Arrow logo and combinations thereof are trademarks of Advanced Micro Devices, Inc. in the United States and/or other jurisdictions. ARM and TrustZone are registered trademarks of ARM Limited in the UK and other countries. PCIe and PCI Express are registered trademarks of PCI-SIG Corporation. Other names are for informational purposes only and may be trademarks of their respective owners.