HETEROGENEOUS PARALLEL PROCESSING DIRECTIONS

Gabriel Loh
AMD Research
APU: ACCELERATED PROCESSING UNIT

- The APU provides significant benefits over previous platforms
- Combines scalar processing on CPU with parallel processing on the GPU and high bandwidth access to memory
- How do we make it even better going forward?
  - Easier to program
  - Easier to optimize
  - Easier to load balance
  - Higher performance
  - Lower power
Evolving the APU: Evolution for Better Compute, Entertainment, & PC Experiences

2011 APUs
- World’s first design with multi-core x86 and discrete-level GPU on single die
- Radeon Memory Bus and Fusion Compute Link
- AMD A-Series and AMD E-Series for PCs, G-Series for Embedded

2012 APUs
- New generations of x86 and GPU IP incorporated
- Dedicated video encode/decode
- New AMD A-Series for PCs, R-Series for Embedded

2013 APUs
- First full System-on-Chip APUs
- First tablet APUs
- New AMD G-Series Embedded SOC
- World’s first Server APUs
- World’s first Semi-Custom APUs

2014 – HSA Support
- Coherent address space
- User model kernel queuing and dispatch
- Enables efficient OpenMP
WHAT IS HSA?

Processor design that makes it easy to harness the entire computing power of an HSA-enabled APU for faster and more power-efficient devices.
HSA: REVOLUTIONARY ARCHITECTURE

UNLOCKING ALL APU GFLOPS

Access to full potential of APU compute power

ALL-PROCESSORS-EQUAL

GPU and CPU have equal flexibility to create and dispatch work items

EQUAL ACCESS TO ENTIRE MEMORY

GPU and CPU have uniform visibility into entire memory space

APU GFLOPS

GPU GFLOPS

CPU GFLOPS

hQ

CPU

GPU

hUMA

CPU

GPU
AMD OPECTRON™ X-SERIES SERVER APUs

HYPERSCALE EFFICIENCIES FOR DENSE COMPUTING CLUSTERS

2014 BERLIN APU

- Next generation Server APU
- First Heterogeneous System Architecture (HSA) Server APU

- Big Data Workloads
  Scalable platform for efficiently processing and analyzing visual and text data

- HPC Workloads
  GPU compute without the power and memory copy overhead of discrete graphics cards

- Hosted Desktop
  Provides end users with fully functional and personalized desktops

- Multimedia Services
  Cost/power efficient delivery of cloud video and entertainment services

Announced design win: HP Moonshot Server

COMBINES BEST IN CLASS AMD RADEON™ HD GRAPHICS WITH X86 TECHNOLOGY
“KAVERI” FEATURING UP TO 4 CPU + 8 GPU CORES

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

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

Visit amd.com/computecores for more detail
FastForward Program
– Started July, 2012 with a total of $12.6M in DOE funding
– Two research areas
  – Heterogeneous processors
  – Next generation memory systems

DesignForward Program
– Started November, 2013 with a total of $3.1M in DOE funding
– Research interconnect architecture to take HSA to network level
AREAS OF INTEREST

- Additional analytics research, particularly time sensitive analysis
- Predictive performance for next generation APUs (models and simulators)
- Acceleration architectures
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. 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

© 2014 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD Arrow logo and combinations thereof are trademarks of Advanced Micro Devices, Inc. in the United States and/or other jurisdictions. SPEC is a registered trademark of the Standard Performance Evaluation Corporation (SPEC). Other names are for informational purposes only and may be trademarks of their respective owners.