THE AMD FASTFORWARD PROJECT

NNSA ASC 2014 PI MEETING
FEBRUARY 27, 2014

FASTFORWARD OBJECTIVES

▲ Investigate processor and memory technologies for exascale systems
- Based on cost effective **high volume architectures** and **open standards**
- Provide significant benefits to **high-volume markets**

▲ Based on extending high volume **APU architecture**

X86
Cores

GPU

FASTFORWARD SCOPE

AMD
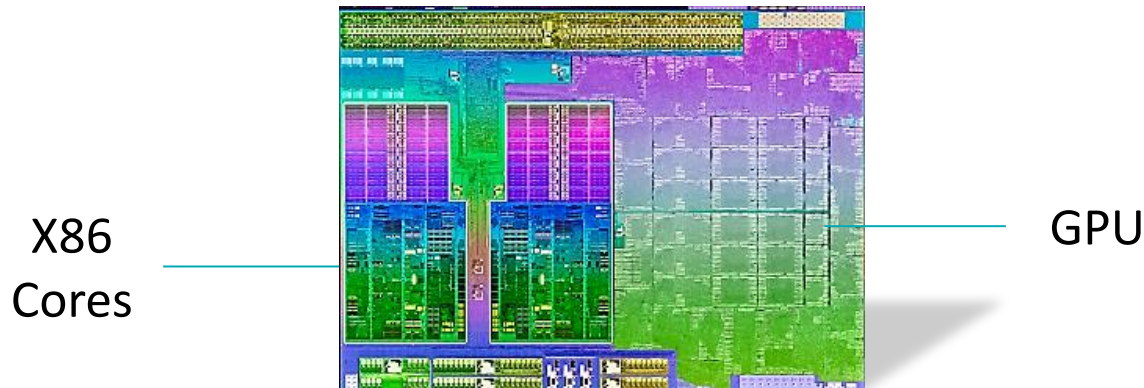
▲ **Processor research (PI: Mike Schulte)**

– Enhanced heterogeneous architectures incorporating advanced GPU features

– Improved energy utilization

– Advanced reliability and resiliency mechanisms

– Efficient communication and data movement

– Simplified programming models and tools

▲ **Memory research (PI: Mike Ignatowski)**

– Emerging technologies and new architectures

– Significant reductions in data movement and energy per access

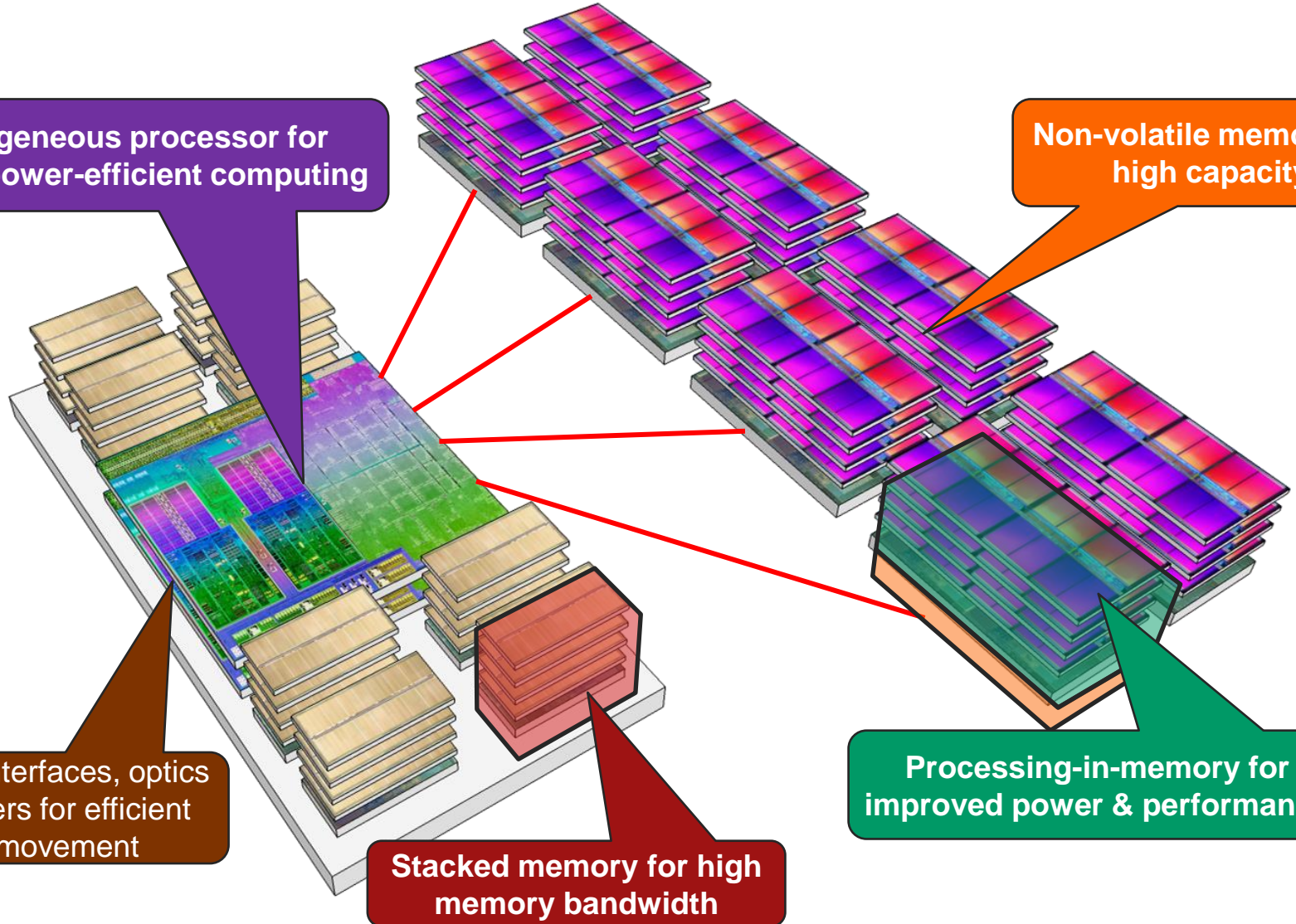– Improvements in memory latency, bandwidth, resiliency, and capacity.

AMD'S FASTFORWARD PROJECT – KEY TECHNOLOGIES

AMD

Tools, language extensions, and APIs for productive programming

Heterogeneous processor for reliable, power-efficient computing

Non-volatile memory for high capacity

Network interfaces, optics and routers for efficient data movement

Processing-in-memory for improved power & performance

Stacked memory for high memory bandwidth

THE HETEROGENEOUS SYSTEM ARCHITECTURE (HSA)

▲ HSA is an industry standard with broad adoption and strong programming tools support (Industry, Universities, Nat. Labs)
  - Founders Include: AMD, ARM, Qualcomm, Samsung…
  - Many other members at different levels, including ANL, LLNL, ORNL, SNL

▲ HSA is a critical enabler for the AMD APU exascale solution.

▲ GPUs efficiently shares memory coherently with CPUs
  - Shared virtual memory, supports pointer-based structures
  - Includes user-level task queues to dispatch tasks to GPUs
  - Address more memory, avoid wasteful data copies, and much easier to program
  - Accelerate a broad range of parallel workloads

▲ Support a range of familiar languages :
  - FORTRAN, C, C++, C++ AMP, Java, OpenCL, OpenMP

▲ Initial focus on GPUs, eventually supporting many types of accelerators

SOME ACCOMPLISHMENTS

AMD

◢ Porting, characterization and optimization of selected DOE proxy applications on AMD APUs

◢ **APU Research Accomplishments**

– Updates to Accelerated Processing Unit (APU) simulation model.

– Efficient techniques for task assignment, scheduling, and synchronization

– Asynchronous communication and computation

– Fine-grained power management

– Adaptive voltage and frequency scaling

– Near threshold computing

– Data compression in GPGPU applications.

– Redundant multithreading into APU vector engines to detect and correct errors

# SOME ACCOMPLISHMENTS (CONT.)

▲ **Processor-in-Memory (PIM)**

  – Programming architecture and interfaces defined

  – Simulation model developed

  – Sample proxy applications ported

▲ **Two Level Memory**

  – Programming architecture and interfaces defined

  – Simulation model developed

  – Sample proxy applications ported

Logic / Interface

▲ **Stacked Memory**

  – Evaluations of various architectures and interface options

  – Enhancements to resiliency

  – Projections on technology and architecture trends
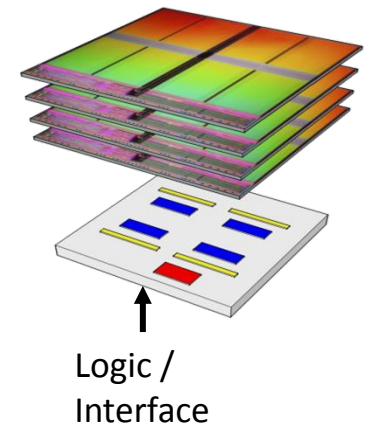
INTERACTIONS WITH THE DOE NATIONAL LABS

AMD ⤤

◢ Collaborative research efforts

◢ Working group meetings

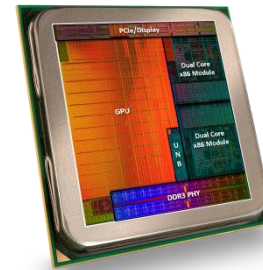◢ Technical deep dives and workshops

◢ FastForward review meetings

IMPACT ON PRODUCTS

ACCELERATING THE INTRODUCTION OF NEW TECHNOLOGIES AND DESIGNS INTO THE GENERAL COMPUTER MARKET

▲ Power-efficient circuits

▲ APU micro-architecture and architecture enhancements

▲ Resiliency mechanisms and modeling

▲ Software tools

▲ Many more in progress….

EXPECT IMPACT ON SCIENCE AT EXASCALE

AMD

- **Parallel programming** simplified, made more efficient for large-scale systems
- **Performance** and **power efficiency** greatly improved on DOE applications
- **Enhanced reliability** and improved mean time between application failure
- **Memory bandwidth** and **capacity** improvements with novel memory hierarchies
- **Efficient communication** and **data movement** throughout the system

DISCLAIMER & ATTRIBUTION

AMD

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.