SAND2011-####C

SAND2011-4814P

Sandia's HPC Platform Strategy:
Preparation for Exascale

Presentation at:

JOWOG34: July 18-22, 2011
AWE - Aldermaston
Reading, UK

James Ang, Ph.D.
Scalable Computer Architectures Department
Computing Research Center
Sandia National Laboratories
Albuquerque, NM

Sandia is a Multiprogram Laboratory Operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy Under Contract DE-ACO4-94AL85000.

Sandia National Laboratories

Outline

- Co-design
- DARPA/UHPC X-caliber project
- Key enabling capabilities
- Leverage Sandia's MESA (Microsystems and Engineering Sciences Application) capabilities
- Industry collaborations
- Sandia's collaboration with Micron Technology
- Keys to realizing Exascale computing

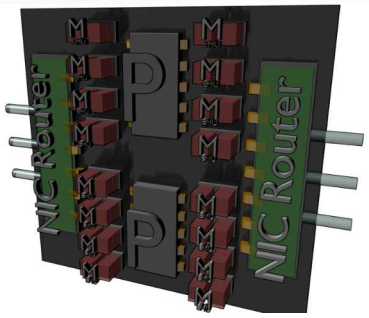
Co-design

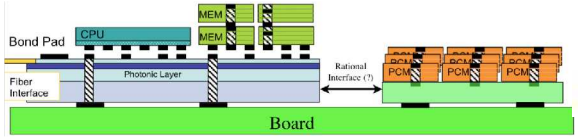
- Lessons from the embedded computing community
  - Working with Prof. Sharon Hu, University of Notre Dame, one of the pioneers of co-design for embedded computing design
  - Optimization is based on partitioning the work among different elements of the embedded system to minimize energy consumption
- The traditional focus for HPC is to minimize wall clock time
  - The objective function for Exascale HPC will be to optimize based on some weighted average between minimizing time to solution and minimizing energy per solution
- Ongoing and New Co-design efforts – see Hemmert/Rodrigues' talk
  - DARPA/UHPC X-caliber project
  - Work with LANL on the ACES – Cray Advanced Interconnection Network project – see Hemmert's talk
  - Participant in two of the recently announced DOE/ASCR Exascale Co-design Centers:
    - Combustion ECDC – Sandia led
    - ExMatEx CDC – LANL led

X-caliber – Ubiquitous HPC Proposal

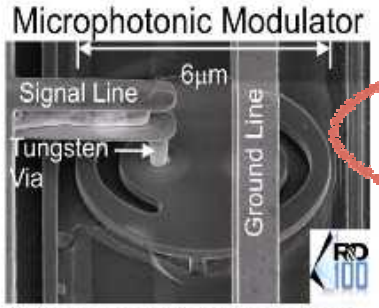
Sandia National Laboratories, LexisNexis, Micron Technologies,
LSU, USC/ISI, Notre Dame, UIUC

With support from NNSA/ASC, DOE/ASCR, DoD and Open Source Communities: e.g., SST, Trilinos





High Level Packaging Strawman



"Ultralow-Power Silicon Microphotonic Communications Platform"
2009 R&D100 Award Winner

Our Process/Approach

- Develop X-caliber architecture with multi-dimensional co-design to analyze and evaluate a rich suite of innovative technologies for HPC systems and applications
- Apply system simulators and prototype testbeds to quantify the performance of extreme-scale applications on conceptual architecture designs
- Adopt open innovation philosophy to grow and leverage ideas from the IT community and beyond

Cultivate Strategic Linkages

- DOE/ASCR, NNSA/ASC, DOE Exascale Initiative
- Other DoD, DOE/LDRD

Key Innovations

- Memory-centric approach to HPC design
- The system is memory that produces computation, not processors talking to dumb memory
- ParalleX execution model allows movement of computation to data in ways that minimize the energy consumption
- Exploit infrastructure for global control flow
- Adaptive runtime and OS to allow computation migration in response to component failures
- Adaptivity also enables "sprint mode" operations in the processor, memory and network to support transient high power operation in return for lower global energy usage

Expected Impact

- Develop a conceptual design to overcome the energy efficiency, dependability and programmability challenges of the five UHPC challenge problems
- Establish collaborative partnerships both internal and external to X-caliber/UHPC

Sandia National Laboratories

X-caliber Leadership

- **Uniqueness – deep vertical integration**
  - A technology foundation that spans – device physics, semiconductor fabrication, microelectronics packaging, HPC architecture design, system software, algorithm and application developers/users
- **While Sandia is the lead, we assembled a large diverse team in acknowledgment of Joy's Law**
  - *No matter who you are, most of the smartest people work for someone else*
- **Sandia is a *Lead User* – able to innovate in ways that are not constrained by Business Models**
  - *Lead Users are firms or individual consumers that expect to benefit from using a product or service. In contrast, manufacturers expect to benefit from selling a product or service.* – Eric von Hippel
- **Business Models are important: Sandia has first-hand understanding of the differences between:**
  - First of a Kind
  - One of a Kind

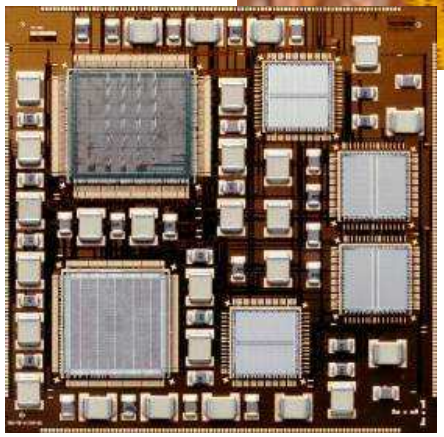
Key Enabling Capabilities

- SST Open Framework for HPC architectural simulation – see Rodrigues' talk
- Mantevo Miniapps – see Barrett's talk
- Codesign Roundtable
  - Applications
  - Algorithms
  - System Software
  - Architectures
  - Microelectronics design and fabrication capability
- Experimental architecture testbeds

Leverage MESA Capabilities

- **Key areas of collaboration:**
  - **Processor/Memory design**
  - **3D Integration and Quilt Packaging**
  - **Optical Interconnects**

- **Sandia's Microsystems Center is one of the heaviest users of IBM's trusted foundry**
- **Recent design submission history**
  - **FY06-08: 14 Design Submissions (130nm, 90nm nodes)**
  - **FY09-10: 16 Design Submissions (130nm, 90nm, 65nm, 45nm nodes)**



Sandia National Laboratories

Industry Collaborations

Cultivate strategic partnerships

- Intel – Umbrella CRADA
- IBM – Trusted Foundry Work, informatics
- Cray – CRADA for informatics & ACES D&E project
- Oracle (Sun) – See Noe's talks on Red Sky
- Micron Technology – CRADA for ECC development

How do we encourage industry to pursue the revolutionary approaches needed to address Exascale?

- Growing expertise in computer engineering and architecture – important to collaborate with industry as partners
- This expertise compliments our application, algorithm, and system software developers to be able to co-design and develop prototypes to implement and experiment with revolutionary models of computation
- We are also making a sustained investment in the open development of the SST simulation framework, and Mantevo mini-application proxies
- With our Microsystems design and fabrication capability, Sandia is also able to create proof of concept prototypes and hardware artifacts

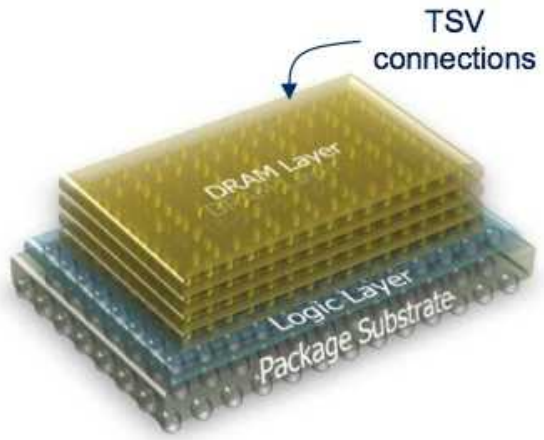
Summary of Micron-Sandia Interactions

Activity	Outputs
Kickoff Meeting	Micron-Sandia Collaboration Begins (July'06)
Foundation for Discussions	2-way NDA (Nov'07)
Advanced Memory for DOE Architectures	Simulations, Papers, PIM LDRD Effort, ASC/CSSE L2 Milestone: Evaluate Advanced Memory Subsystems – 4Q-FY10
IAA Activities	Dean Klein is a member of the IAA Advisory Board IAA Workshop on Memory Opportunities for HPC (Jan'08) IAA Workshop on HPC Architectural Simulation (Sept'09)
Collaborations with other agencies	Alignment of ASC/CSSE, DoD/ACS, & IAA support to Integrate U-MD's Memory Simulator (DRAMsim & eBOBsim) with Sandia's SST
Proposal Partnerships	DARPA/UHPC X-caliber, DOE/ASCR Data Movement Dominates
CRADA – established July'10	Micron-Sandia collaboration to analyze advanced concepts for error correction in advanced memory designs – Patent Application filed, May '11, *Automated discovery of optimal, symbol-based SECDED codes*

- **Technical Exchanges from July 2006 - Present**
  - Approximately 30 face-to-face technical meetings
  - Catalyst for Collaboration with other DOE/NNSA labs, DoD, and Universities

New Micron Technology Development and Engineering Project

- Based on our long-term collaboration, Sandia is helping NNSA/ASC establish a Cooperative Agreement with Micron Technology
- On behalf of NNSA/ASC and DOE/ASCR, Sandia is responsible for technical oversight of, and collaboration with Micron on this project
- Sandia working with Micron to define the technical scope – the project has four major parts:
  - Construction of the simulation infrastructure
  - Design explorations, focused on sets of in-memory operations to be evaluated, simulated, and analyzed
  - FPGA prototyping work of ideas that are identified as candidates for inclusion in future HMC parts
  - Research into improved low energy signaling and topologies



Functioning prototypes in silicon TODAY

Sandia National Laboratories

Keys to Realizing Exascale Computing

- Improving data movement performance / efficiency
  - Both intra-node and inter-node
- Developing a Co-design methodology for HPC
  - Including the development of the objective function that balances minimizing time to solution with minimizing energy to solution
- Providing processor, memory, and interconnection network designers with insights from application users/developers, system software developers on how best to use the additional transistors that Moore's Law will continue to provide
- Significant investment in new application development
  - Both funding and time