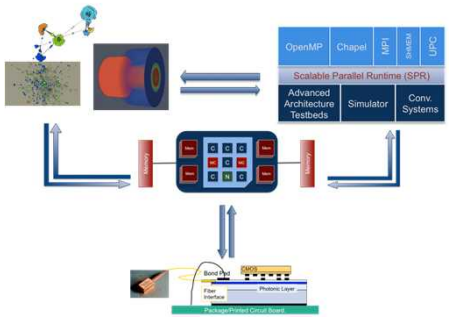
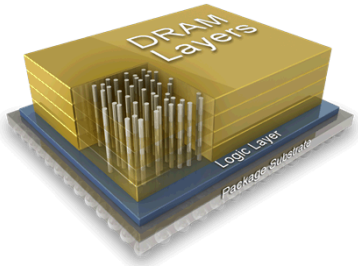
SAND2015-####

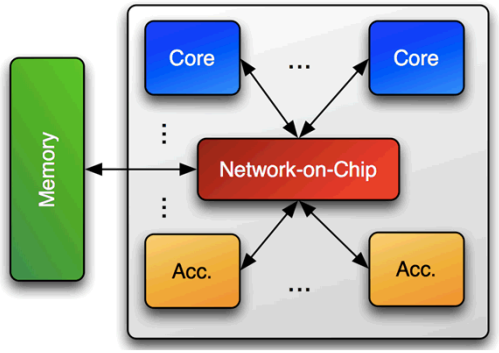
SAND2015-6269PE

Exceptional service in the national interest

Sandia National Laboratories







Sandia Exascale Architectures R&D Overview

James A. Ang, PhD, Manager
Scalable Computer Architectures Department

Centro Svizzero di Calculo Scientifico

Lugano, Switzerland

July 10, 2015

CENTER FOR COMPUTING RESEARCH

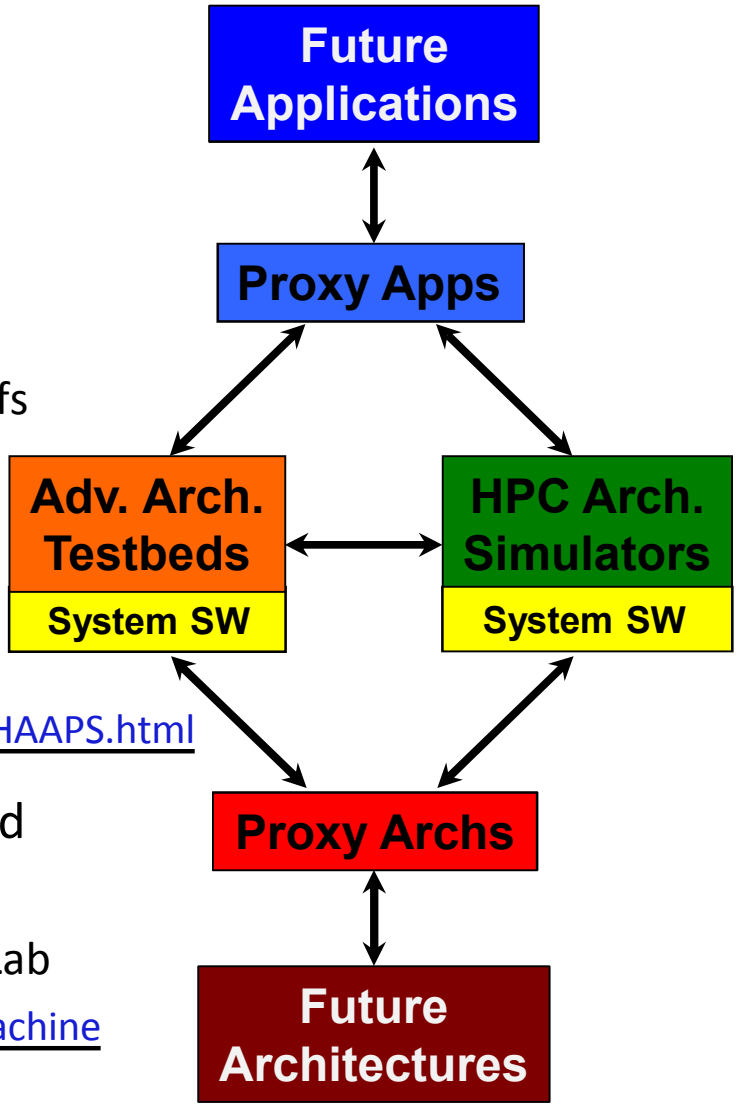
U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia Co-design Capabilities

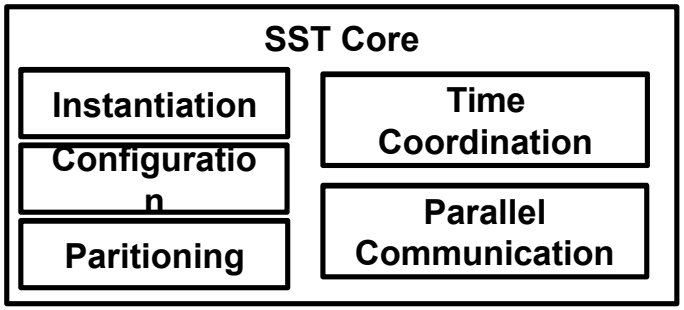
- Proxy Applications (Mantevo):
  - Application source for architecture-centric optimization and analysis
  - http://mantevo.org
- HPC Architectural Simulation Framework (SST):
  - Flexible to accommodate fidelity/speed tradeoffs
  - http://SST-simulator.org
- ASC Advanced Architecture Test Beds:
  - Evolving examples of COTS "state-of-the-art"
  - http://www.sandia.gov/asc/computational_systems/HAAPS.html
- Abstract Machine Model (AMM) Definitions and associated Proxy Architectures
  - Supported by SC/ASCR Computer Architecture Lab
  - http://crd.lbl.gov/assets/pubs_presos/CALAbstractMachineModelsv1.1.pdf

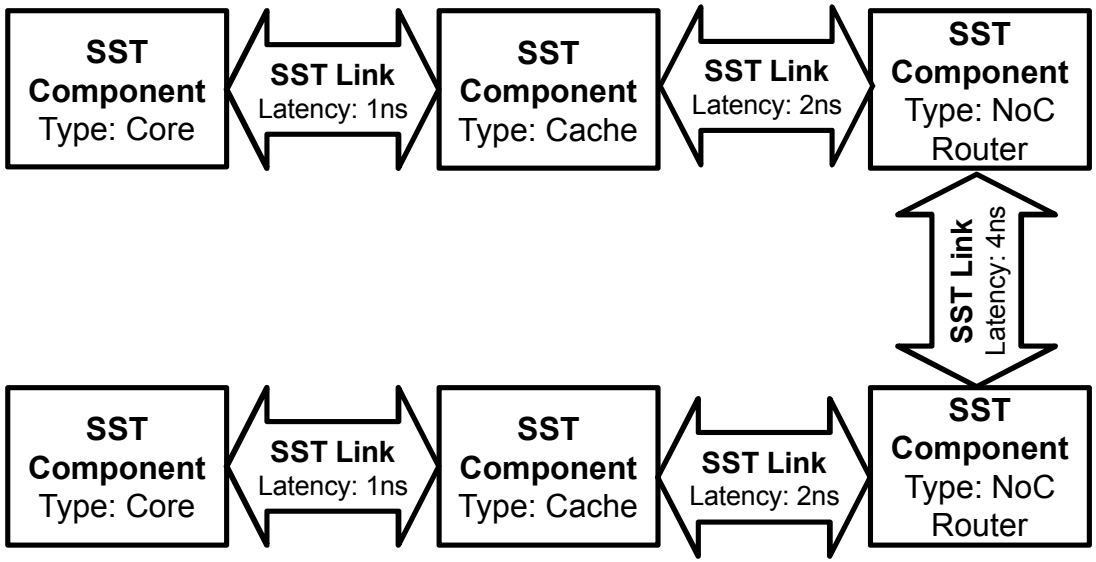


SST Capabilities

- SST Core (framework):
  - Time-scale independent: micro-, meso-, macro-scale simulations
  - Provides a number of interfaces and utilities for simulation models
- Components: SST's simulation models
  - Components perform the actual simulation
  - Many built-in models available: processors, memory, network
  - Compatible with external models: Gem5, DRAMSim2, many others
- Open API
  - Easily extensible with new models
  - Modular framework
  - Open-source core

SST's discrete-event algorithm

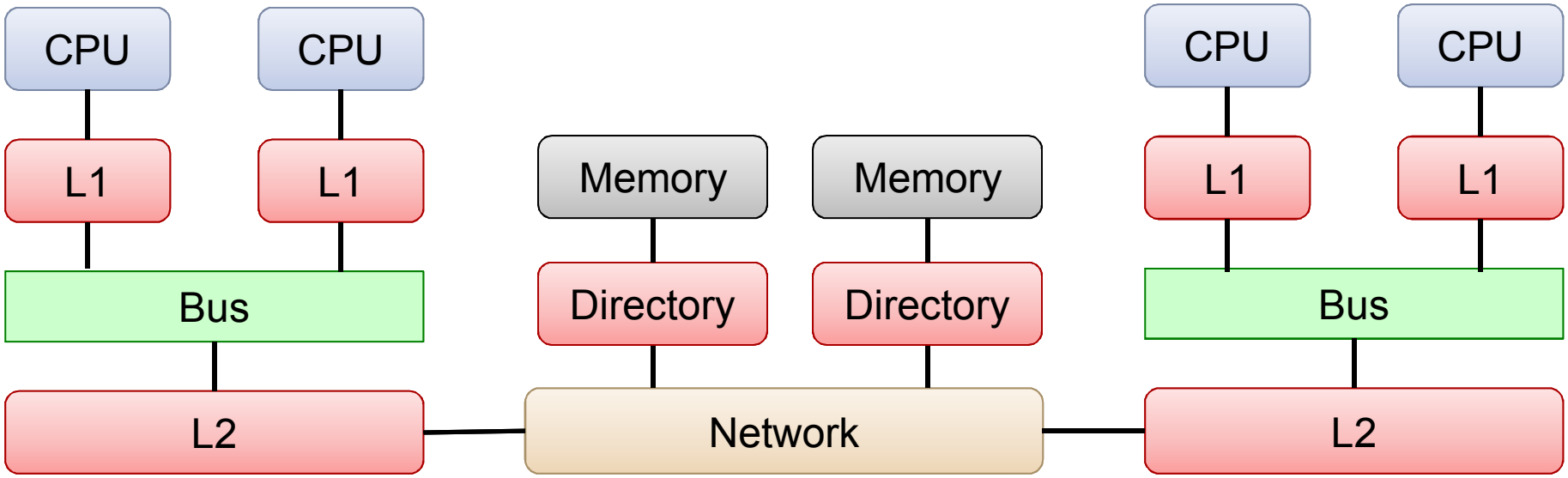




- Simulations are comprised of **components** connected by **links**
- **Components** interact by sending events over **links**
- Each **link** has a *minimum* latency
- **Components** can load **subComponents** and **modules** for additional functionality

Configuring a simulation

- SST uses a Python configuration file
  - Defines global parameters for the simulation
  - Defines and configures components
  - Specifies links and link latencies between components
- Open ‘demo_4cores.py’



Adv. Architecture Test Bed Evolution

Sandia National Laboratories

Westmere + Knights Ferry
Arthur

SandyBridge Knights Corner (B)
Compton

SandyBridge Knights Corner (C)
Compton V2

IvyBridge
Volta (XC)

Haswell 16 Core + PowerInsight V2
Kaveri+ PowerInsight V2
Shepard

Llano+ PowerInsight
Teller

Trinity + PowerInsight
Teller V2

Interlagos + Fermi 2090x
Curie (XK6)

Interlagos + Kepler K20X
Curie (XK7) V2

SandyBridge + Kepler K20
Shannon

SandyBridge + Kepler K20 + K40
Shannon V2

SandyBridge Kepler K20 + K40 + K80
Shannon V3

Power 7 + FPGA
Watson

Power 8
ESP

Power 8
White

64bit
ESP

HMC
JENGA

64bit
Hammer

Sept. 2011 | Sept. 2012 | Sept. 2013 | June 2014 | Sept. 2014

ASC Advanced Architecture Test Beds

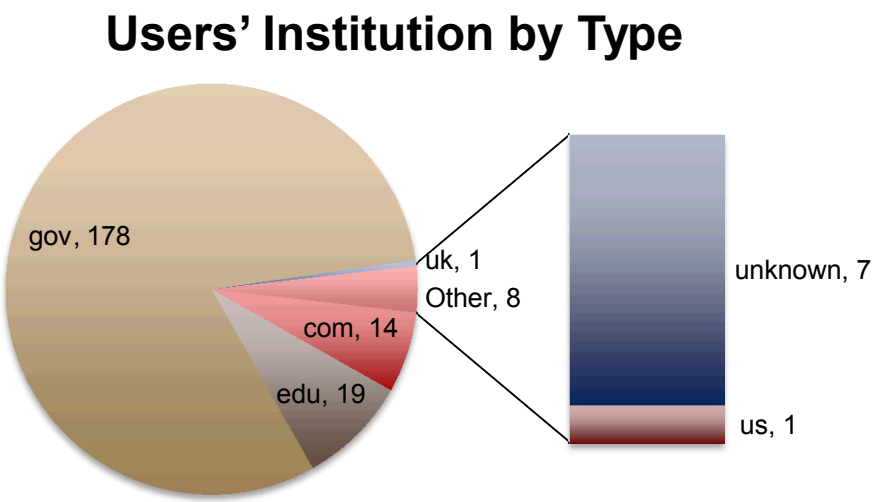
October 2014

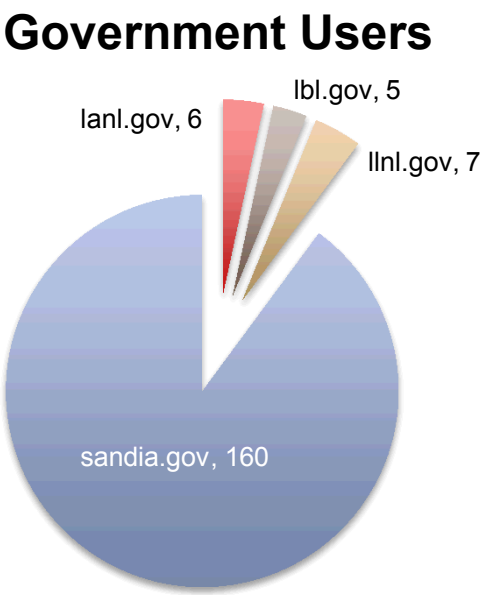
- *Shepard*
  - Single Heterogeneous Cluster
  - FDR Infiniband
  - 36 Node Haswell Xeon (16 core)
    - Instrumented with PowerInsight V2
  - 36 Node AMD Kaveri APU (hardware HSA)
    - Instrumented with PowerInsight V2
- *Shannon*
  - 32 node dual socket Sandy Bridge/Nvidia GPU cluster
    - 6 nodes 2 K20's per node
    - 12 nodes 4 K20's per node
    - 4 nodes 4 K40's per node
    - 9 nodes 2 K80's per node
- *White*
  - 10 node Power 8/Nvidia GPU cluster
  - FDR IB (coming soon)
  - 1 2U Dual Socket Power 8 node
  - 9 4U Dual Socket Power 8
    - 2 K40 GPUs per node
- *Hammer*
  - 45 node ARM cluster
  - 1 Xgene 1, 64bit ARMv8 (eight core) processor per node
  - 10Gb Ethernet
- *Jenga*
  - Micron HMC Gen2 Testbed
  - 2GB Capacity, 4 links: each with Altera Stratix® V FPGA

Test Bed Users

Breakdown as of October 31, 2014

- 220 Unique users
  - Up from 80 February 8th 2013
- By system:
  - 127 Compton
  - 136 Curie
  - 107 Shannon
  - 118 Teller
  - 100 Volta
  - 26 Morgan (SRN)
- 178 Government users
  - 160 Sandia
  - Others, LANL, LBL, LLNL
- 42 Other external
  - **.edu, .com**, net, .org, .uk, .ca, .us





Co-design: Influence Future Architectures

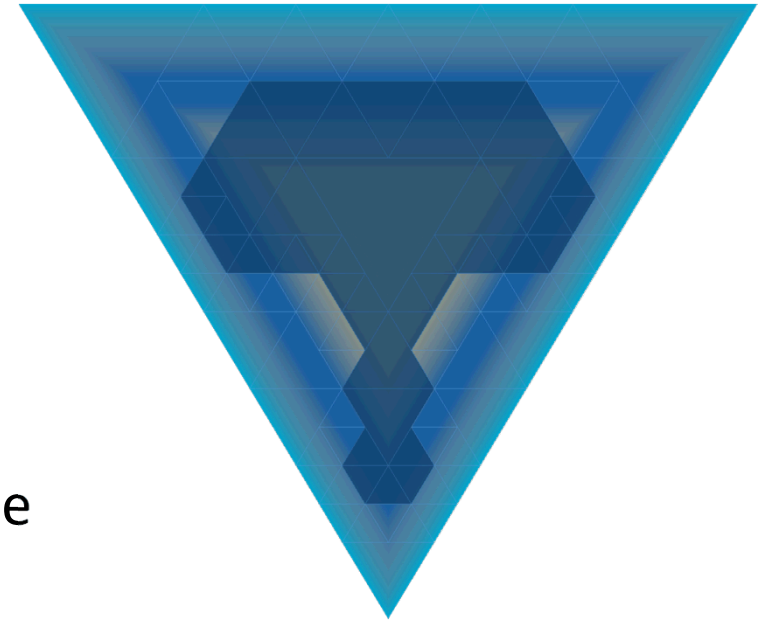
- Technical Leadership for NNSA/ASC Advanced Technology Systems
- Advanced Architecture Concepts
  - Extreme-scale computing Grand Challenge (XGC) LDRD
- Power Management API
  - Targeted NRE Investment Area for Trinity
  - http://powerapi.sandia.gov/docs/PowerAPI_SAND.pdf
- Cultivation of Industry Collaborations
  - Ongoing with FF/DF vendors: AMD, Cray, IBM, Intel, Nvidia
  - Ongoing with Micron Technology: HMC Testbeds, and R&D Contract
  - New: ARM, Applied Microcircuits, Broadcom, Qualcomm, Cavium
- SoC for HPC Workshop
  - http://sites.google.com/a/lbl.gov/socforhpc/
- Abstracted Interface for Component Integration
  - https://sst-simulator.org/downloads/abstract_interface_final.pdf

Influence Future Architectures:

NNSA/ASC ACES Collaboration on AT Systems

- Architecture, project and multi-laboratory leadership for ACES/Trinity Project
  - RFP Development
  - Vendor Negotiations
  - Architecture
- Major Impact on SC/ASCR Cori (NERSC-8) procurement
- HPC System leadership will continue with ATS3 Project

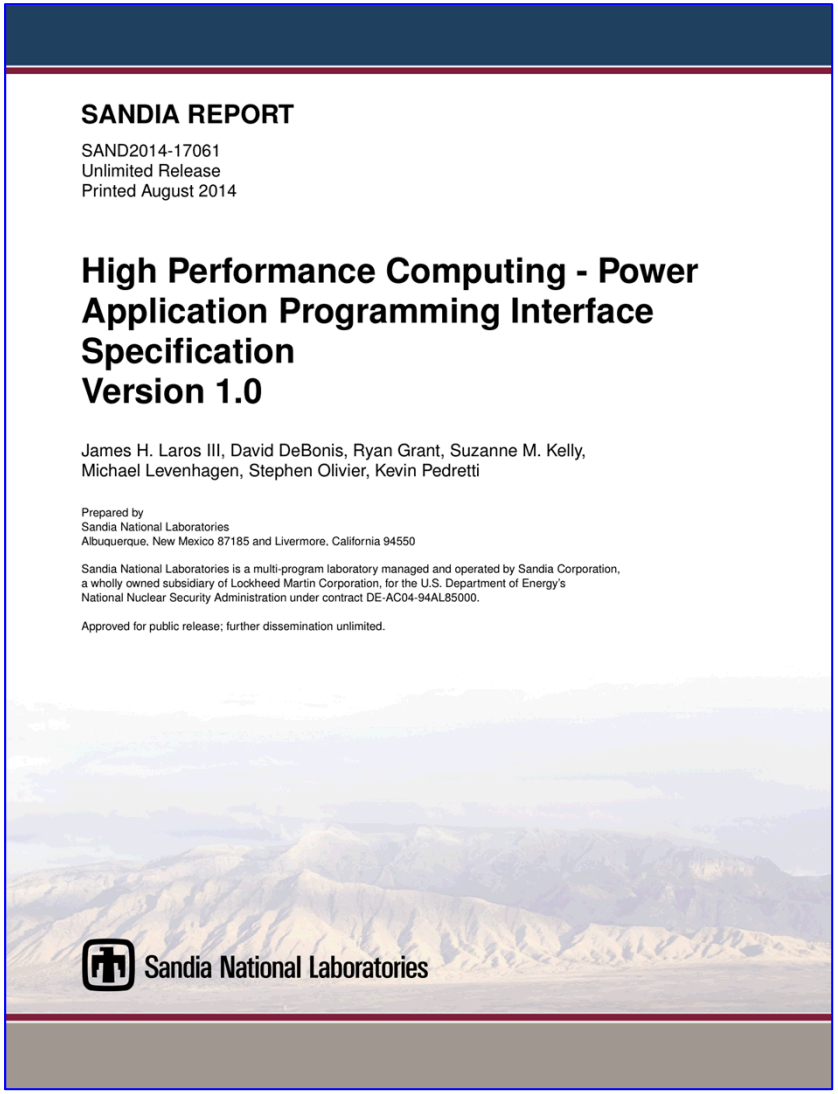
TRINITY



Influence Future Architectures:

Open HPC Power API

- Sandia HPC Power API Spec
  - 7/24/14 – Early Draft for External Review
  - 8/26/14 – First Public Release
  - 9/02/14 – ASC L2 Milestone Review
  - 9/04/14 – External Review
  - BoF and Poster Accepted for SC'14



SANDIA REPORT

SAND2014-17061
Unlimited Release
Printed August 2014

High Performance Computing - Power Application Programming Interface Specification Version 1.0

James H. Laros III, David DeBonis, Ryan Grant, Suzanne M. Kelly, Michael Levenhagen, Stephen Olivier, Kevin Pedretti

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

Sandia National Laboratories

Influence Future Architectures:

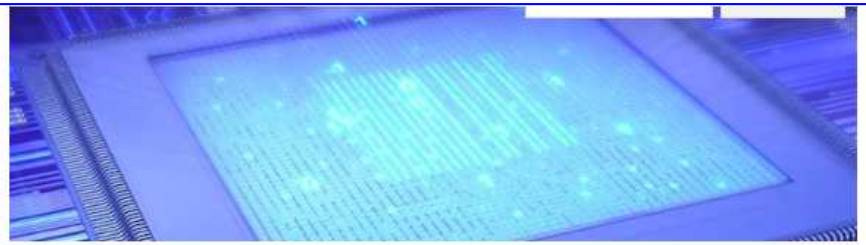
Micron Collaborations

- HMC Gen 2 T&E Boards
- DOE R&D project
  - New T&E Boards
  - Design Studies
    - Functions in memory
    - Multi-cube modules
    - Analysis of *Search* Proxy Apps

Influence Future Architectures:

Workshop on SoC for HPC

- Inter-agency workshop
- August 2014
- 35 invited participants
- Key Questions:
  - State of the Art
  - Technology Inventory and Gap Analysis
  - Software Infrastructure
  - Simulation/Modeling
  - Open Technology for SoC
- http://www.socforhpc.org



Workshop on System-on-a-Chip Design for HPC
Abstracts
Attendees
Schedule
Sitemap

Workshop on System-on-a-Chip Design for HPC

Our current HPC ecosystem relies upon Commercial off-the-Shelf (COTS) building blocks to enable cost-effective design by sharing costs across a larger ecosystem. Modern HPC nodes use commodity chipsets and processor chips integrated together on custom motherboards. We are embarking upon a new era for commodity HPC where the chip acts as the "silicon motherboard" that interconnects commodity Intellectual Property (IP) circuit building blocks to create a complete integrated System-on-a-Chip (SoC). This approach is still very much COTS, but the commodities are licensable IP for pre-verified circuit designs (the lego-blocks for SoC designs) rather than the chips. It achieves cost-competitiveness because the dominant cost of designing a chip is the cost of verifying the circuit building blocks. The cost benefits derive from the ability to leverage a commodity ecosystem of embedded IP logic components where the non-recurring expense (NRE) cost of designing and verifying a new processor or memory controller design (an IP building block) can be amortized by licensing the technology to myriad embedded applications. The market for licensed circuit IP in the embedded space is much larger marketplace (both in volume and total revenue) than for server chips and the market segment for SoC building blocks is growing at a far faster pace than the current server chip market. HPC system designers should leverage this new avenue for leveraging the cost-advantages of COTS technology.

Traditionally SoC design methods have focused on low-power consumer

SoC for HPC Workshop - Goals

- The SoC for HPC workshop will focus on semi-custom, application-targeted designs and server processing for HPC and data-centers

- The goal is to develop a strategy for a robust open ecosystem for SoC designs that serve the needs of energy efficient HPC applications for multiple government agencies

SoC for HPC Workshop - Outcomes

- Plot a roadmap for creating SoC ecosystem for HPC
  - Leverage the enormous investments taking place in the embedded consumer electronics market
  - Focus on what is effective for HPC
- Identify opportunities and weaknesses in the SoC strategy
  - What is the performance & market potential of this approach?
  - What is missing from the commodity IP component market?
  - Where will market forces NOT deliver the kinds of components required for effective Server HPC/WSC SoC designs
- Where should the U.S. government (DOE, DOD, DARPA, NASA, NSF) concentrate R&D investments for SoC technology innovation?
- Draft an *internal* report documenting our findings

Thoughts on Co-Design Paths

- Co-design paths
  - Reactive
  - Proactive
  - Holistic
- The Road Not Taken

  … Two roads diverged in a wood, and I –

  I took the one less travelled by

  And that has made all the difference

  – Robert Frost

- HPC is implicitly a road

  less travelled by

Acknowledgements

- SST ISCA Tutorial Slides – Arun Rodrigues, Gwen Voskuilen, Simon Hammond, Branden Moore, and K. Scott Hemmert
- Advanced Architecture Testbeds – James Laros
- Thank you for hosting me at CSCS