SAND2017-12872C

Towards an Open Source Eco-System for Future HPC Designs (and the SST Simulator)

SST Research Team, Scalable Computer Architectures, Sandia National Labs, NM, USA, sdhammo@sandia.gov

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

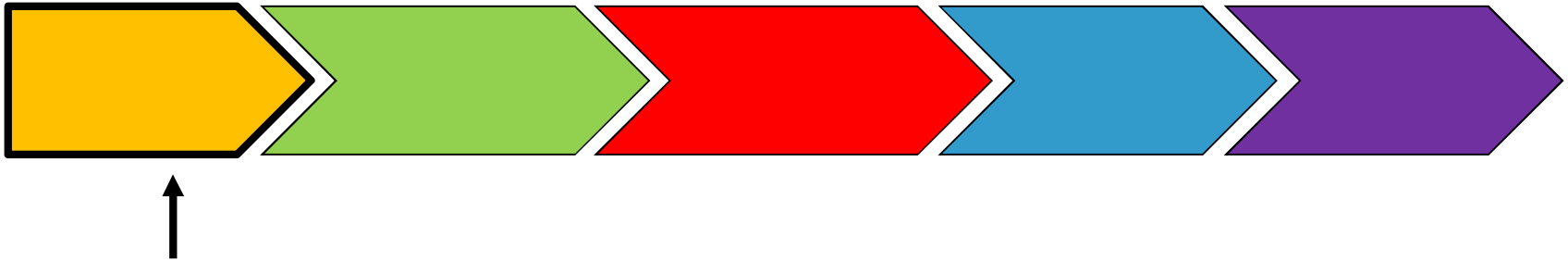
- Sandia has some of the most diverse workloads in the DOE
- Long history in HPC but also in processor design, manufacturing/fabrication, packaging *etc*
- Active computer architecture and computer science research teams
- Will become more important in the future…

Sandia CSRI Facility

Sandia MESA Facility

http://www.sandia.gov/mstc/

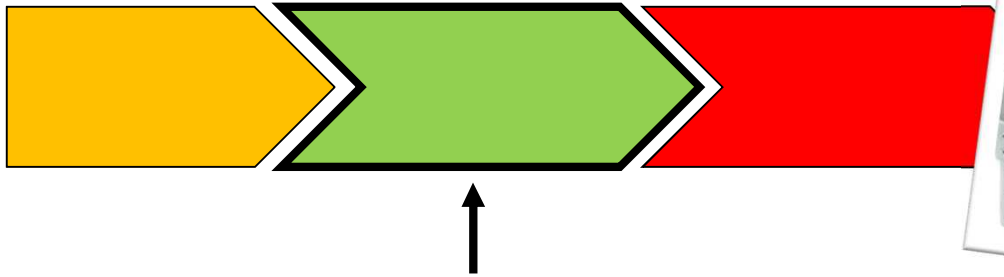
The Long Road for HPC …



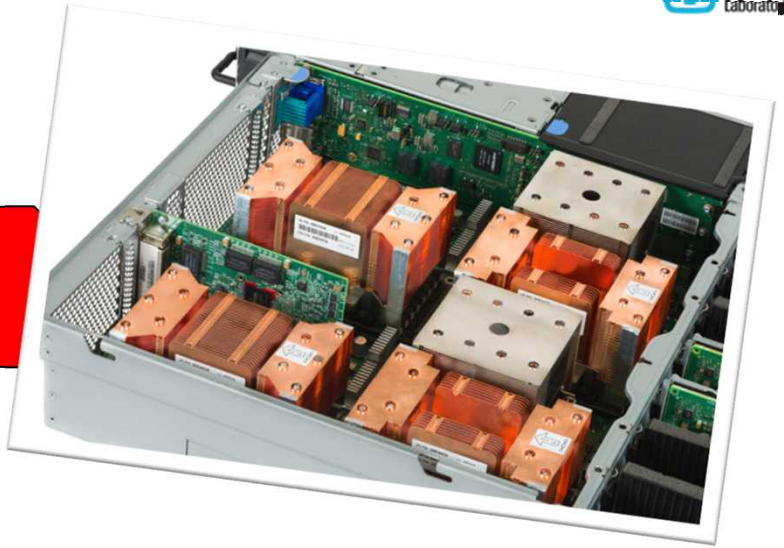
One Vendor, Single Solutions

- One vendor integrates everything, one fixed solution for every problem
- Probably in the rearview mirror for most of HPC systems now
- Little flexibility, less hardware tuning
- Take what you get given (usually get lots you don't want)

The Long Road for HPC …



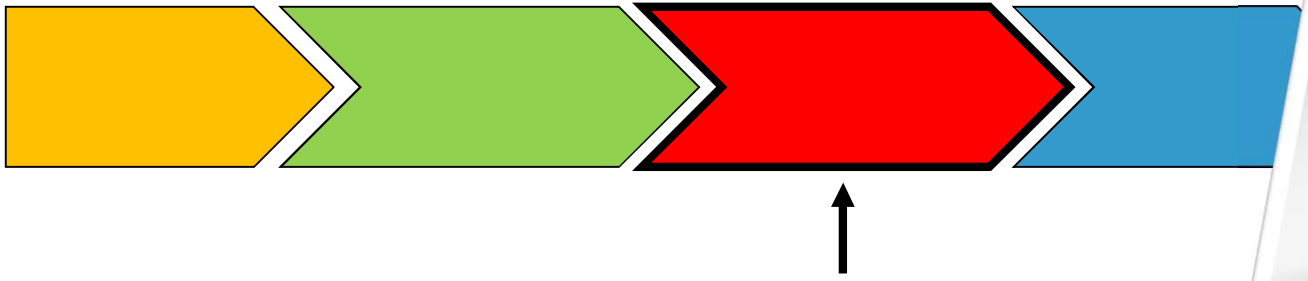
Multi Vendor, Protocol Integration



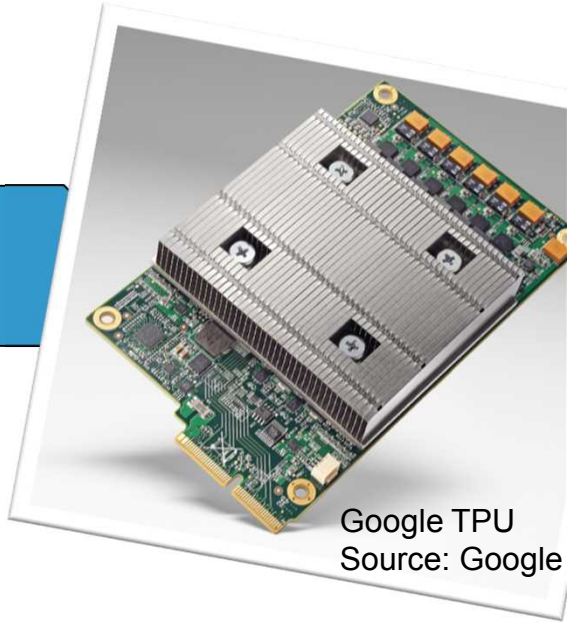
IBM NVIDIA
Source: NVIDIA

- Multiple vendors working together, not integrated hardware
- Support for a shared protocol
- Seeing this with emerging IBM + NVIDIA solutions
- CCIX, Gen-Z, OpenCAPI *etc*

The Long Road for HPC …



One Vendor, “Shopping List SoC”



Google TPU
Source: Google

- Buy a solution from one vendor from an IP catalogue
- Cherry pick components which are optimized/appropriate for your workload
- Complex task for mixed workload environments like DOE (not easy)
- Does one vendor have everything you need?

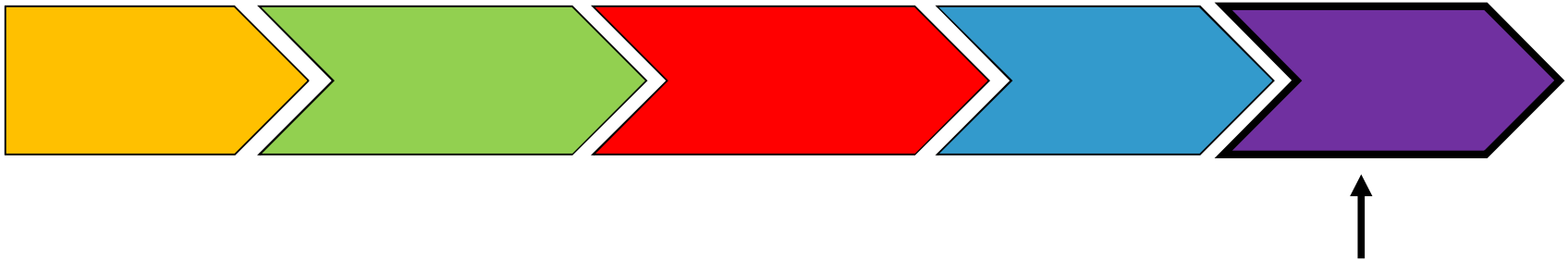
The Long Road for HPC …

Apple A12 SoC

True Customer + Silicon Provider SoC

- Custom IP included in SoC
- Becomes true plug and play for hardware, allows customization where workload permits
- Issue is validation. Open source hardware/HPC has a role to play

The Long Road for HPC …



- Emergence of non-conventional IP (quantum, neuromorphic, *etc*) in a single SoC
- Possible from IP catalogue but could be custom/open source components
- Potentially greater optimization across complex workloads

Thoughts on the Long Term Path

- **Exciting to see huge potential far in the future for HPC**
  - The best of HPC really is yet to come, this is just the beginning
- But .. this is a really complex path to follow

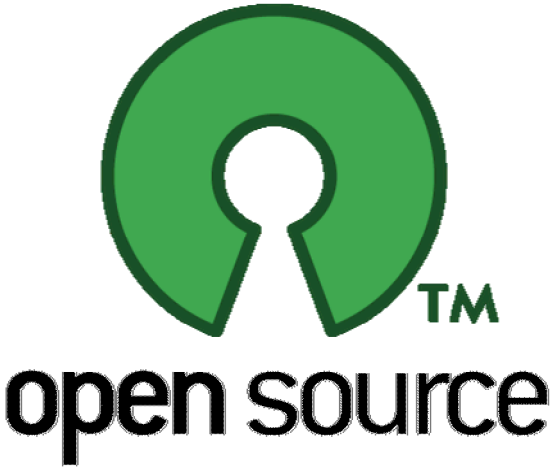
- Needs a very good understanding of workloads, not just hardware
- Diverse and flexible IP catalogue (commercial and open source)
- Validation is the challenge (particularly for *real* solutions in HPC space)
- Significant upside for some workloads but not all
- Gives life into silicon even when Moore's law finally is dead!

What's This Got to do with Open Source HPC?

- These approaches open the door to huge complexity…

- Open source hardware implementations
- Novel Programming model abstractions
- Workload and application analysis

- And .. open source hardware design tools
  - Which is the real topic of my talk …



http://opensource.org

THE STRUCTURAL SIMULATION TOOLKIT

http://sst-simulator.org

Goals

- Become the standard architectural simulation framework for HPC
- Be able to evaluate future systems on DOE/DOD workloads
- Use supercomputers to design supercomputers

Technical Approach

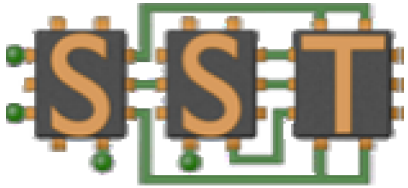
- Parallel
  - Parallel Discrete Event core with conservative optimization over MPI/Threads
- Multiscale
  - Detailed and simple models for processor, network, & memory
- Interoperability
  - DRAMSim, ,memory models
  - routers, NICs, schedulers
- Open
  - Open Core, non-viral, modular

Status

- Parallel Core, basic components
- Current Release (7.1)
  - Improved components
  - Modular core/elements
  - More Internal documentation

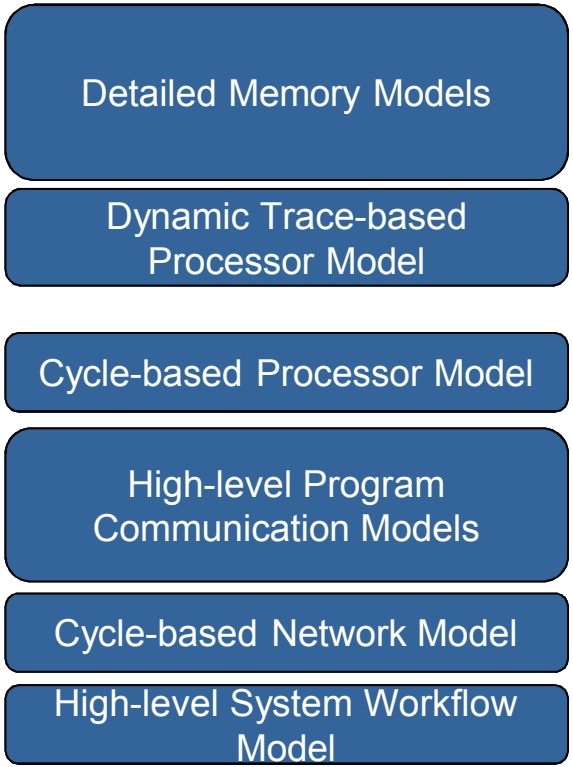
Consortium

Key Capabilities



- **Parallel**
  - Built from the ground up to be scalable
  - Conservative, Distance-based Optimization
  - MPI + Threads
- **Flexible**
  - Enables “mix and match” of simulation components
  - Custom architectures
  - Multiscale tradeoff between accuracy and simulation time
    - E.g., cycle-accurate network with trace-driven endpoints
- **Open API**
  - Easily extensible with new models, modular framework and open source

Breadth and Depth…

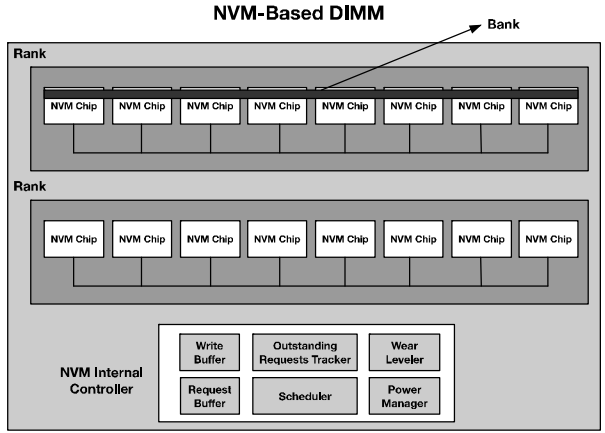


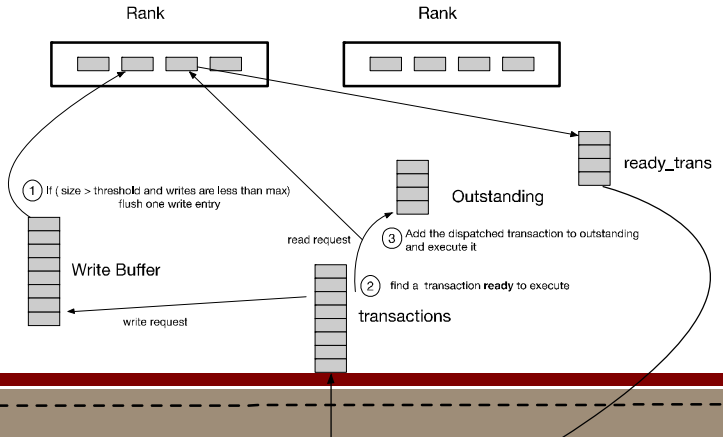
- memHierarchy - Cache and Memory
- cassini - Cache prefetchers
- DRAMSim2 - DDR
- NVDIMMSim - Emerging Memories
- Goblin - HMC
- ariel - PIN-based Tracing
- MacSim - GPGPU
- m5C - Gem5 integration layer
- ember - State-machine Message generation
- firefly - Communication Protocols
- hermes - MPI-like interface
- merlin - Network router model and NIC
- scheduler - Job-scheduler simulation models

Case Study: Non-Volatile Memory

- Messier: NV Memory model
- Focus on NV-DIMMs e.g.:
  - # Banks, Latencies
  - Row buffers, write buffers
  - policies, outstanding requests, ordering
  - Address mapping

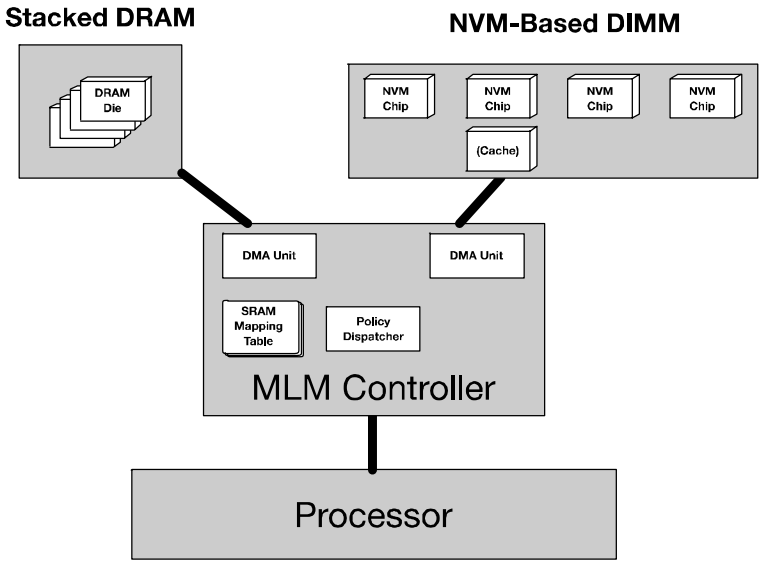
- Report: (Sandia National Labs) SAND2017-1830





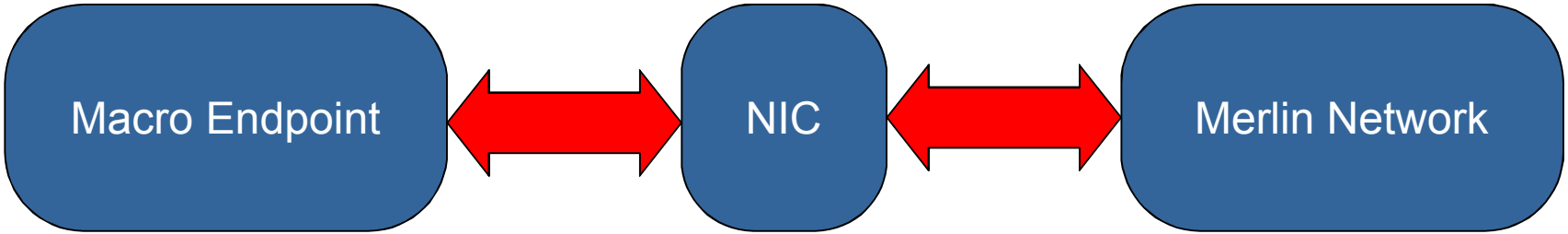
Case Study: Partnership with IBM

- Improvements to IBM CramSim to enable threaded simulations (faster analysis time)
- Improved multi-level memory models
- Performance and scaling improvements (event-driven (clock-less) memory models)
- Scratchpad support
- New TLB model



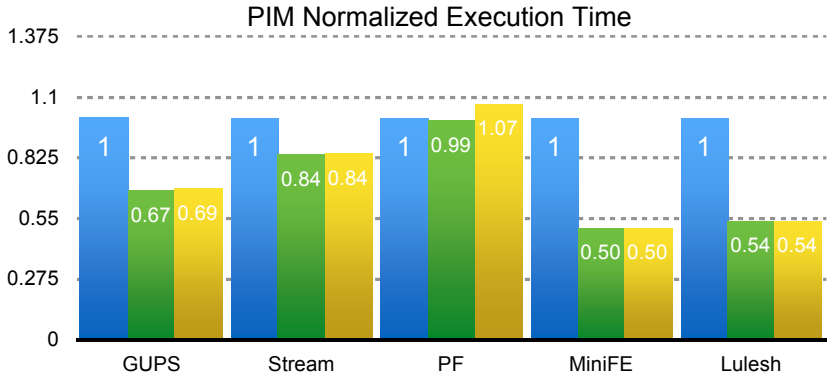
Case Study: Communication End Points

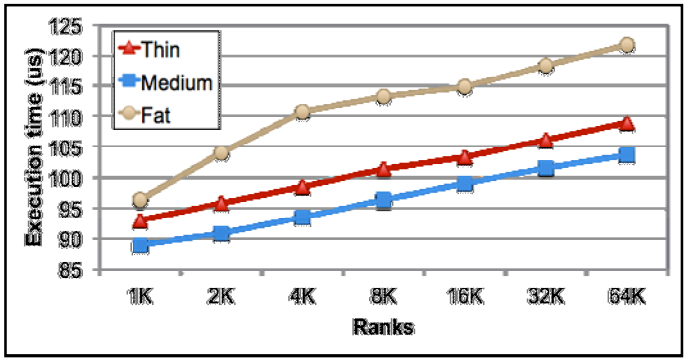
- Macro / Merlin Integration
- Beta release of OTF2 trace replay skeleton
- Beta release of Clang-based auto-skeletonization source-to-source tools
- Integrated job launcher components for simulating PBS or SLURM-like batch systems

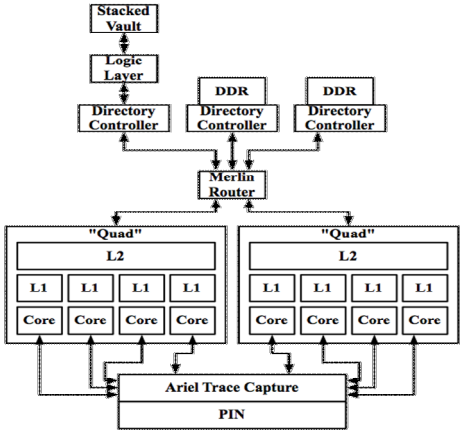


Case Study: Multi-Level Memory Analysis

- Processing-in-memory
- Multi-Level Memory
  - HW Tradeoffs: capacity ratios,
  - SW Tradeoffs: application, runtime, OS, HW control
- Scalable Network Studies
  - Network on Chip
  - Cache coherency
- Scheduling



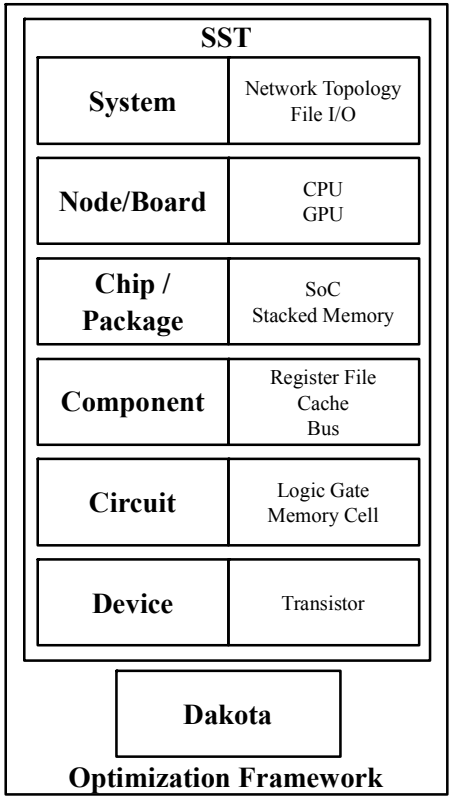




Case Study: System Optimization

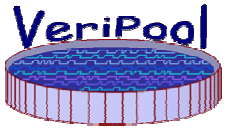
- Using Sandia's open source Dakota optimization framework
- Parameterized hardware model
  - Optimizer gradually refines input parameters
  - Re-simulates models over lots of applications/kernels
- Computationally exhaustive but potential space is so complex humans find it difficult to understand all the tradeoffs

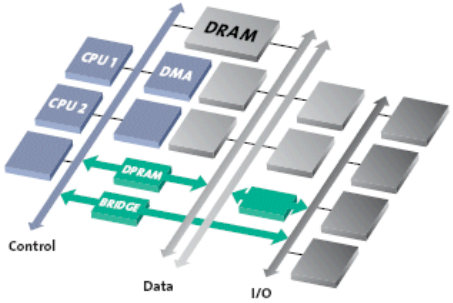
https://dakota.sandia.gov



Near Term Directions…

- HDL Simulation via Verilator & Chisel
  - Low-level hardware design
  - Path to tape-out (Chisel)
- New Processor Models
  - RISC-V
  - Juno
- Improved NoC Models
  - Faster Performance
  - NoC QoS
  - Optical Circuit Routing





THOUGHTS AND DISCUSSION…

Exciting Times to Be in HPC

- Slow down of Moore's Law is beginning to bite, roadmaps are messy
  - Challenging for supercomputing sites trying to provide capability computing

- This is opening up opportunities for *all* of computing
  - The future is more customized/specific solutions
  - Higher efficiency but less generality – customer will need to be sure

- How will you make the decisions on what to buy/put into your next supercomputer?

Flexible, Open, Hardware Toolchains

- Want to be able to virtually design hardware and nodes
  - Use medium-fidelity models from SST to do rough designs
  - Use higher-fidelity models when focusing into promising designs

- Need a common simulation framework to have a plug-and-play experience between different components

- Give us feedback, contribution source, be part of the community

- http://sst-simulator.org