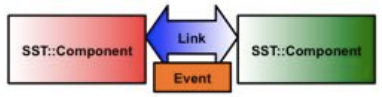
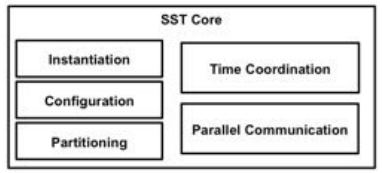
SAND2017-9713 C





Sandia National Laboratories

Exceptional service in the national interest

Sandia's ARM–centric Co-Design Strategy

Introduction to the NNSA/ASC Vanguard Project

James A. Ang, Ron Brightwell, Simon D. Hammond, K. Scott Hemmert. Robert J. Hoekstra, James H. Laros III, Kevin Pedretti and Arun F. Rodrigues
Center for Computing Research
Sandia National Laboratories
Albuquerque, NM 87185-1319 USA

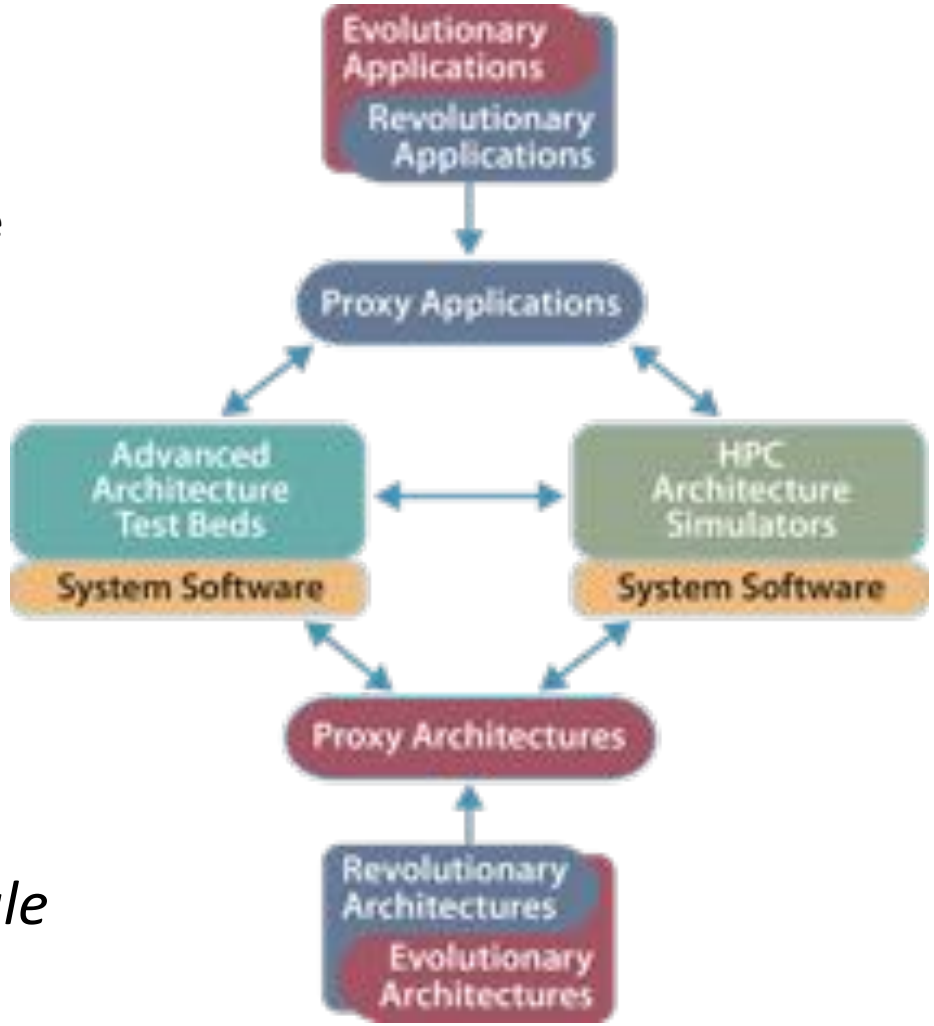
ARM Research Summit, Going ARM Workshop
Robinson College, Cambridge, UK
September 11-13, 2017

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

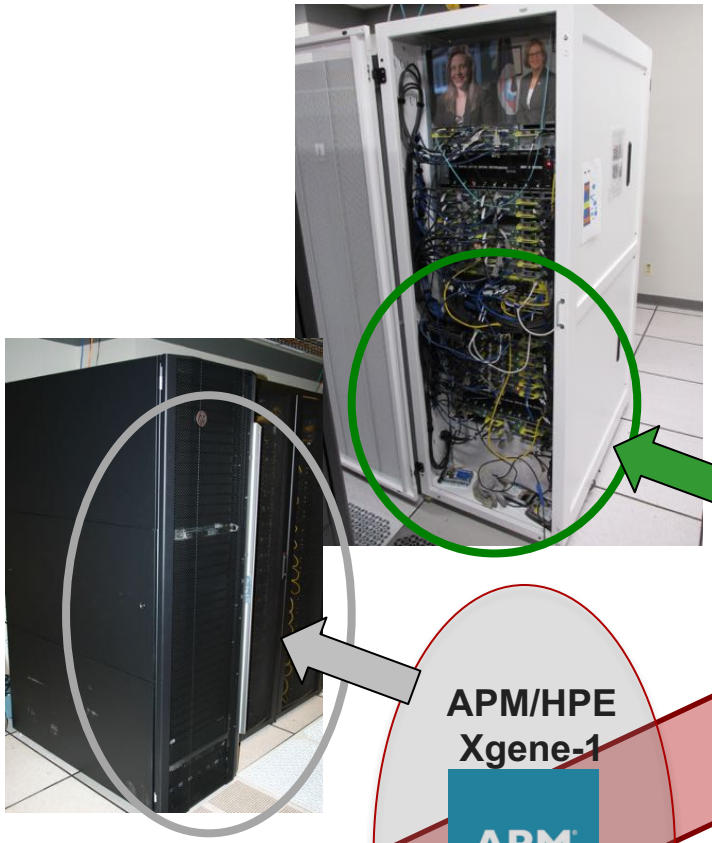
Sandia ARM R&D Overview

- Sandia's ASC Advanced Architecture Testbed project
  - For co-design early access to new architectural features is critical
  - Positive ROI for the pain associated with early hardware & immature software
  - ARM testbed experiences – with Cavium's permission to share pre-production measurements of relative performance
- Architectural Simulators
  - Structural Simulation Toolkit (SST), Open architectural simulation framework
  - Open Source, Open Development, catalyst for collaboration that can be proprietary
  - Links to Sandia/DOE's experience with application drivers
- Sandia is responsible for the NNSA/ASC's first *large scale* ARM platform
  - Overview of key requirements targeting 2019



Sandia's NNSA/ASC ARM Platforms

Sandia National Laboratories



Exascale Platform

ARM Next Gen

ARM

Vanguard "Project"

Install FY19 Target ThunderX3 (Triton)

ARM Next Gen

ARM

Mayer

Pre-production Cavium ThunderX2

Small Pre-Vanguard early access system – FY18

Cavium/HP Labs ThunderX1

ARM

Sullivan

Cavium ThunderX1

APM/HPE Xgene-1

ARM

Hammer

= Retired

= 2015

= 2017

= 2019

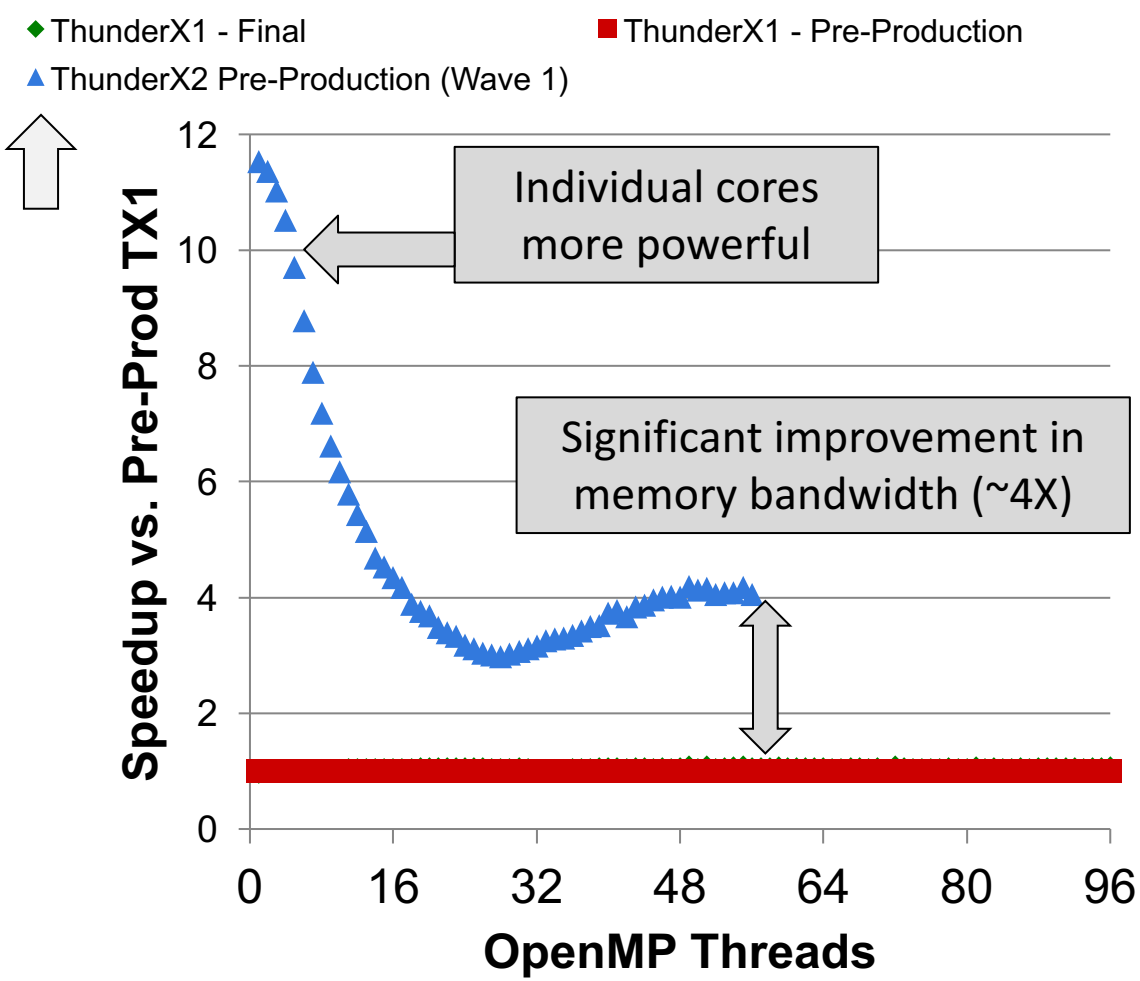
Sept 2011

TODAY

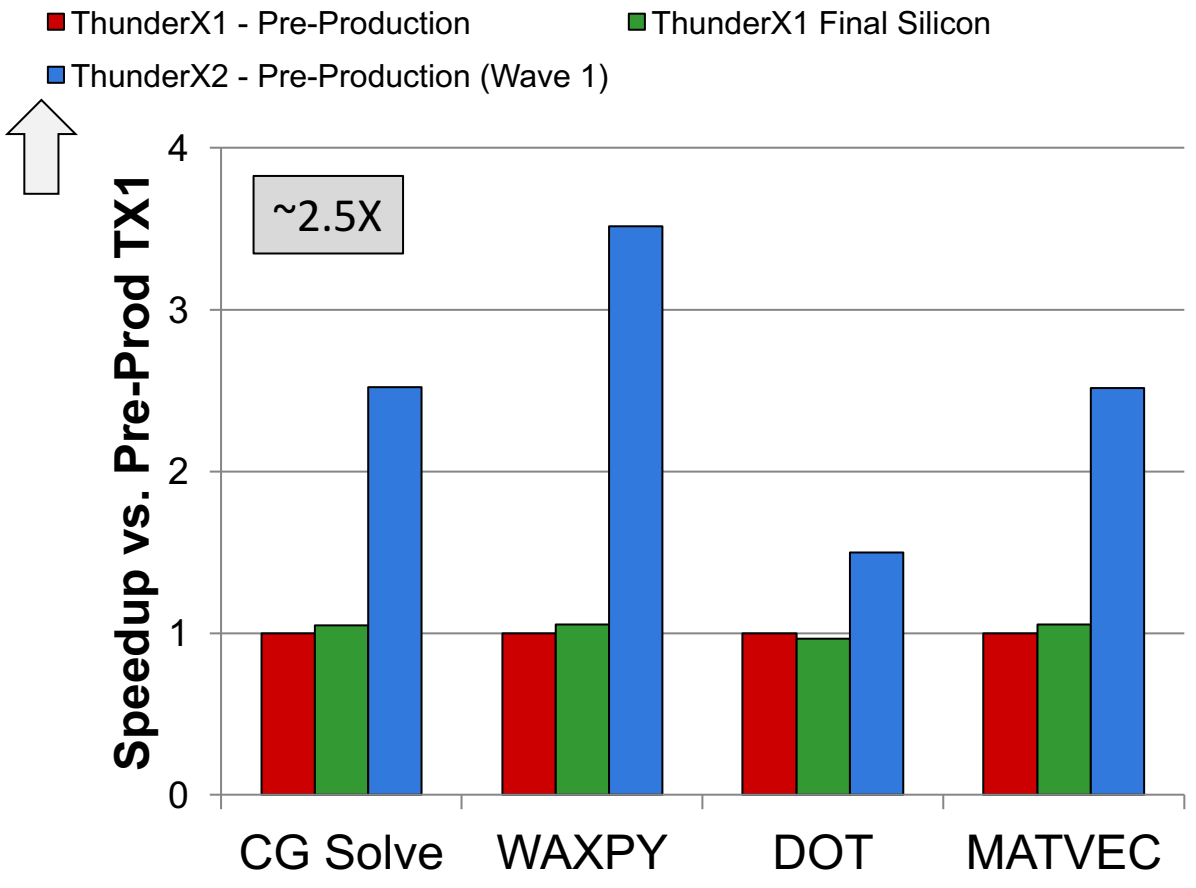
Memory Bandwidth Intensive Kernels

Sandia National Laboratories

STREAM Triad



MiniFE CG Solve/Kernels

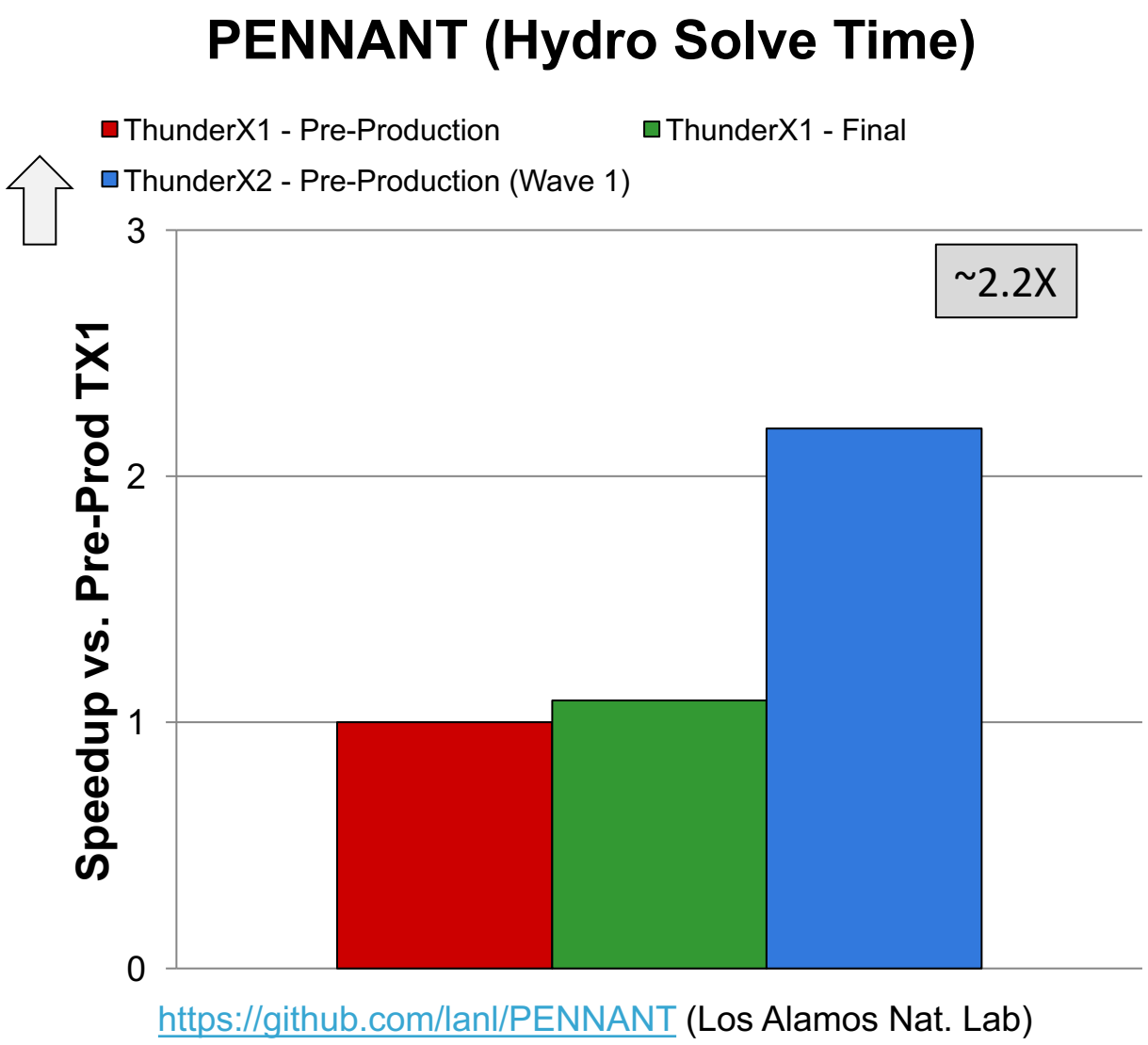


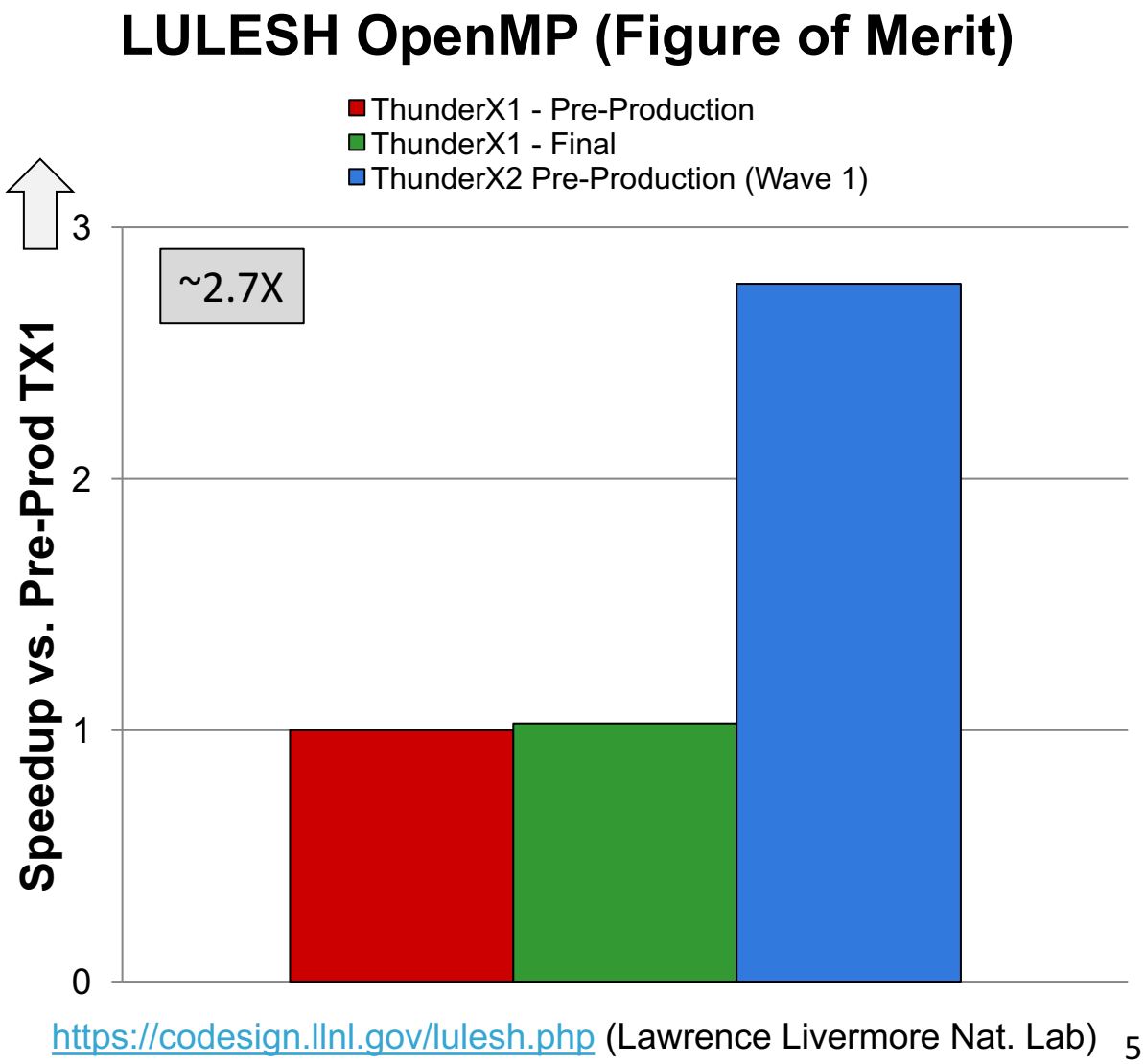
https://github.com/Mantevo/miniFE (Sandia National Labs)

Benchmarked on maximum number of cores (node-node comparison)

Compute Intensive Kernels

Sandia National Laboratories





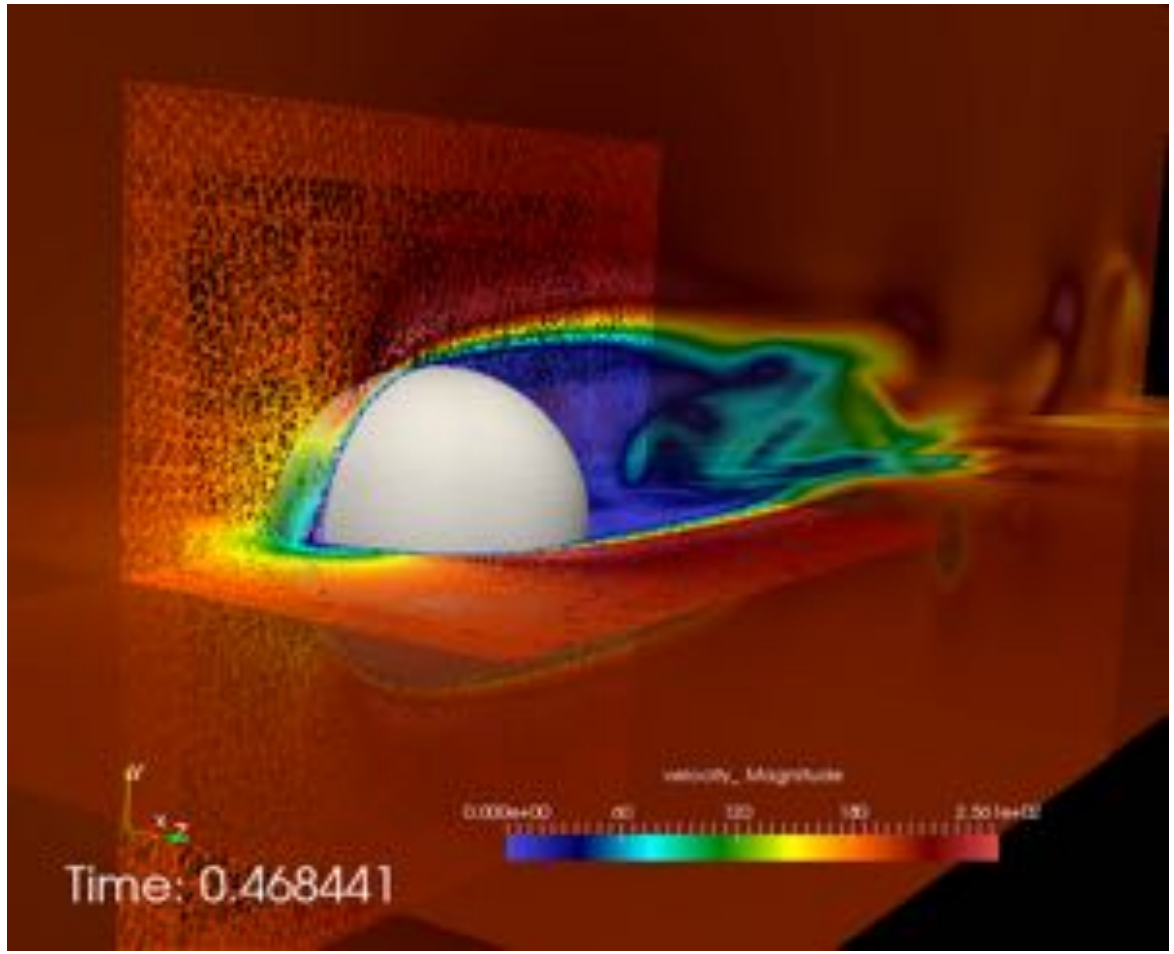
https://github.com/lanl/PENNANT (Los Alamos Nat. Lab)

https://codesign.llnl.gov/lulesh.php (Lawrence Livermore Nat. Lab)

Full Engineering Applications on ARM

Sandia National Laboratories

- On Pre-production TX2, Sandia has been actively working on libraries and packages for its ARM systems
  - Math Libraries/Kernels (Trilinos)
  - I/O Libraries (Exodus Mesh)
  - YAML
- Ported Sandia's open-source *Trinity* acceptance Nalu application to ARM
  - Representative of some SIERRA engineering applications
  - Complex mesh handling, load balancing, solvers, etc.

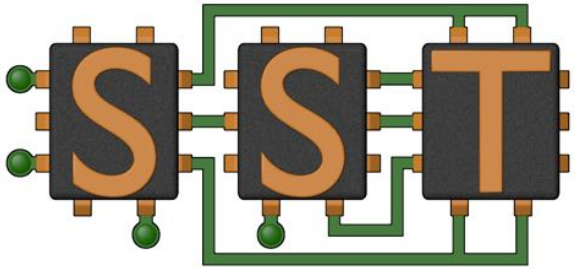


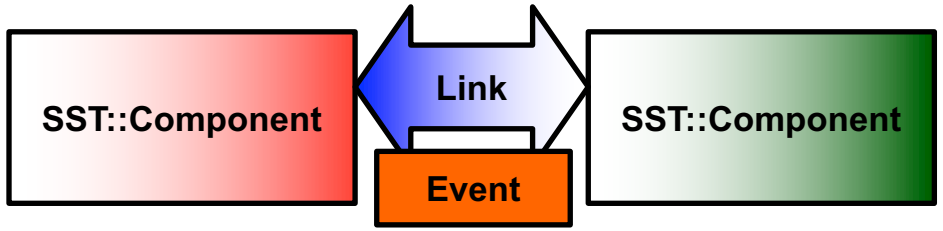
https://github.com/NaluCFD

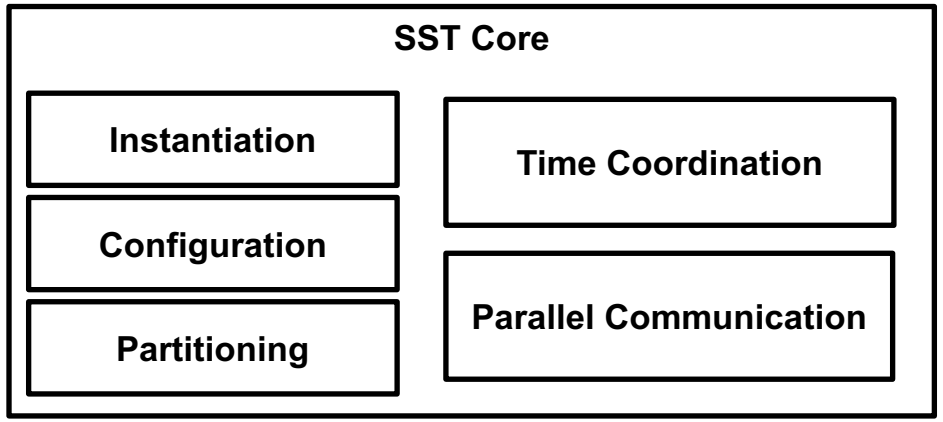
Architectural Simulation Framework: SST

Sandia National Laboratories

- Use Supercomputers to Design Supercomputers
- Parallel Discrete-Event Simulator Framework
  - Flexible framework allows multitude of custom simulators
  - Demonstrated scaling to over 512 processors
- Comes with many built-in simulation models
  - Processors, Memory, Network
- Open API
  - Easily extensible with new models
  - Modular framework
  - (Non-Viral) Open-source core
- Time-scale independent core
  - Handles Micro-, Meso-, Macro-scale simulations
- "Best of Breed" – Bring together work from Labs, Industry, Academia

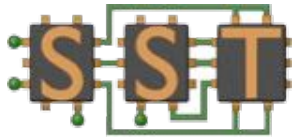






http://sst-simulator.org/

Example SST Element Libraries



Detailed Memory Models

- memHierarchy - Cache and Memory
- cassini - Cache prefetchers
- DRAMSim - DDR
- NVDIMMSim - Emerging Memories

Dynamic Trace-based Processor Model

- ariel - PIN-based Tracing

Cycle-based Processor Model

- m5C - Gem5 integration layer

High-level Program Communication Models

- ember - State-machine Message generation
- firefly - Communication Protocols
- hermes - MPI-like interface

Cycle-based Network Model

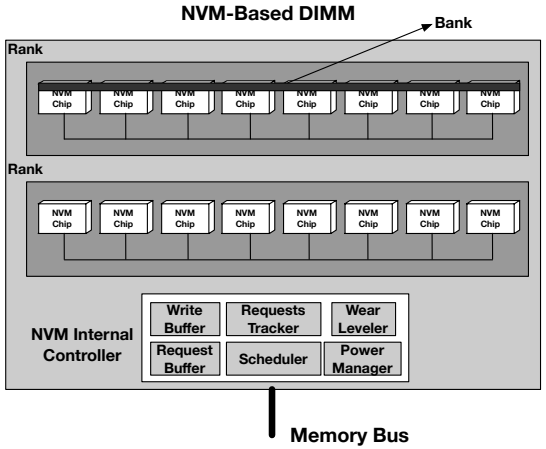
- merlin - Network router model and NIC

High-level System Workflow Model

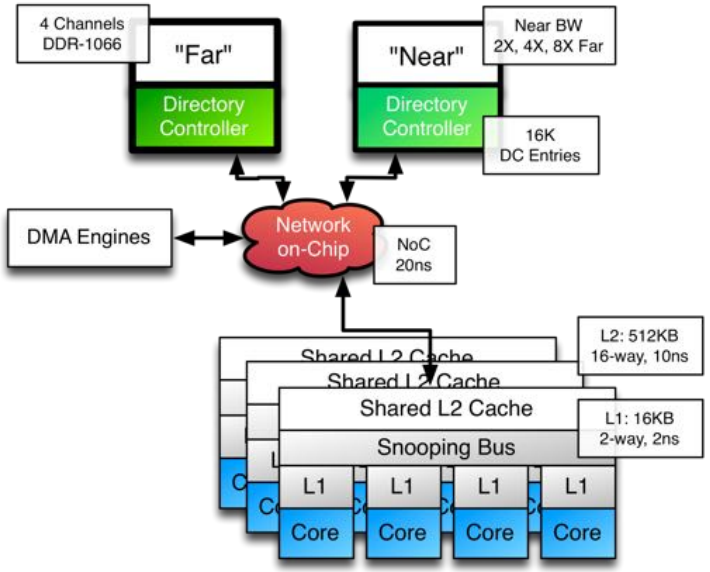
- scheduler - Job-scheduler simulation models

SST Use Cases

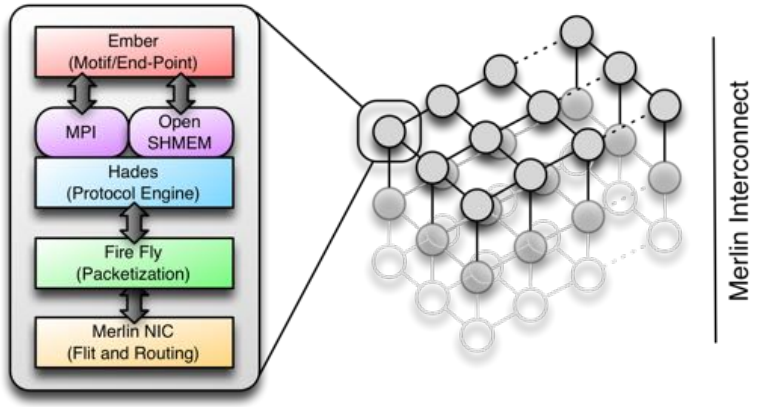
Emerging NV Memory Technologies



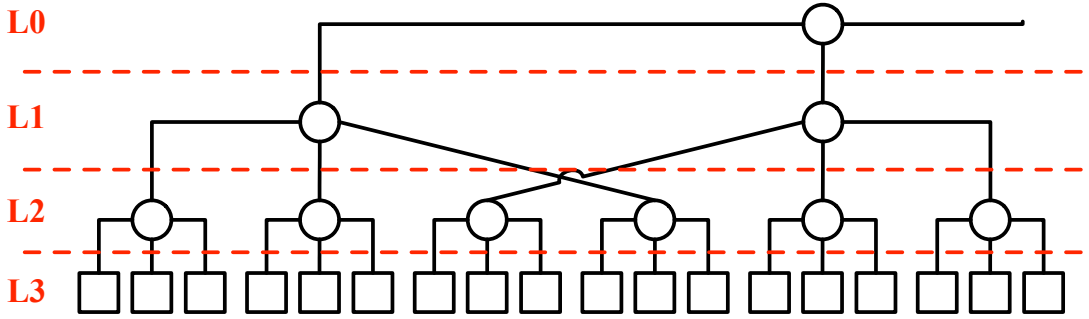
Multi-Level Memory (HBM+DDR+NV)



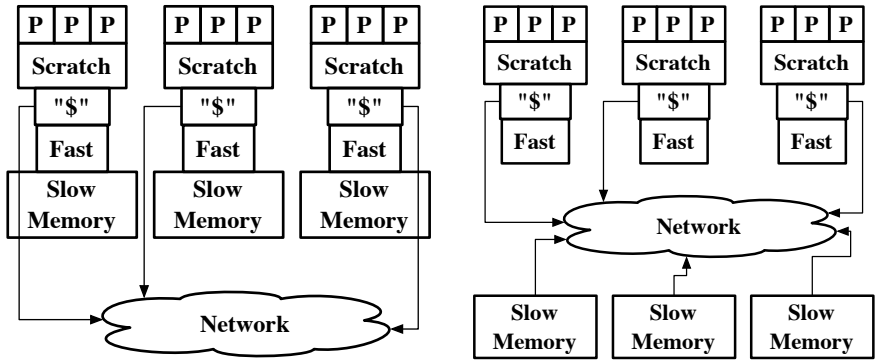
Communication/ Interconnect Modeling



Photonic Network Topology & Routing



Disaggregated Memory



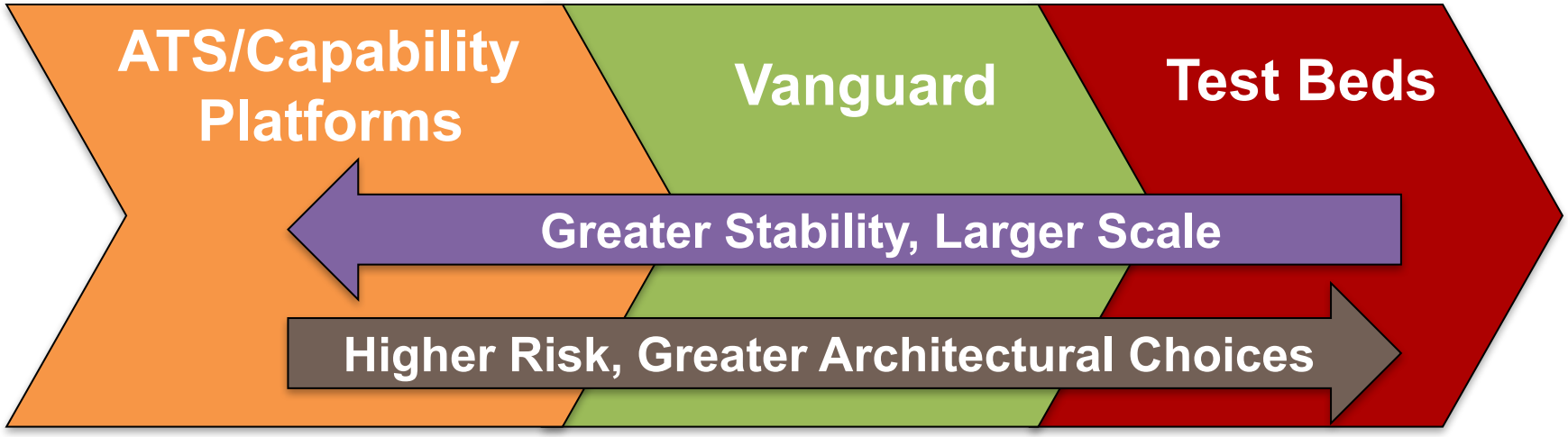
The NNSA/ASC Vanguard Project

Sandia National Laboratories

- **Expand HPC ecosystem** by developing ARM-based system necessary to enable a **credible** ARM-based Advanced Technology System offering for ASC post-2022
  - Mature ARM HPC ecosystem
    - Software stack and toolchain (OS, compilers, scalable MPI, runtime, system management, development tools, I/O, …)
    - On-package high-bandwidth memory
    - Advanced HPC interconnect
  - Influence Supplier/Integrator community to accelerate and optimize ARM technologies for leadership HPC, and foster HPC community confidence
- Leverage the open ARM architecture platform to enable innovations in processor, interconnect, memory, specialized acceleration and packaging technologies that ***focus on performance and scalability of our multi-physics production codes*** and could impact 2025 systems

VANGUARD

The Vanguard Project Approach



- Bridge the gap between what we can accomplish with Testbed project and fielding an Advanced Technology System platform
- Vanguard is a project, not a single platform
- Accelerate maturity of emerging technologies for ASC program

Vanguard Approach (cont.)

Sandia National Laboratories

- **Appropriate** scale systems
  - Must be large enough to serve as proof of concept for future Advanced Technology Systems
  - Sufficient scale to interest production multi-physics application teams
- **Appropriate** investment
  - Gain and maintain vendor and collaborator attention
- **Appropriate** level of risk
  - Goals target mission workloads but not turn-key production mission support

Software Environment Plan - Overview

Sandia National Laboratories

- Goal: Accelerate maturity of ARM ecosystem for ASC computing mission
- Need an integrated software stack for the 2019 ARM Prototype to enable application development and optimization
  - Programming environment (compilers, math libs, tools, MPI, OMP,…..)
  - Low-level OS (optimized Linux, I/O, network, containers + VMs, PowerAPI, …)
  - Job Scheduling and management (WLM(Slurm?), app launcher, user tools, ….)
  - System Management (boot, monitoring, image mgt., rapid re-provisioning, ….)
- Focus Areas
  - Integration and robustness, overall user experience
  - Address known weaknesses: compilers, libs, and tools
  - Increase modularization and openness of system software stack; seek “plugin” capability for externally developed components.

VANGUARD

Software Environment Collaboration Roles

Sandia National Laboratories

- System Vendor
  - Deliver and support core elements of the software environment necessary for a viable integrated system (part of system contract)
- Tri-lab team
  - Integrate system into our computing environment
  - Identify and resolve SW issues in collaboration with system vendor
  - Contribute tools and other capabilities to fill gaps and improve the overall computing environment
- Other Eco-system Vendors/Stakeholders
  - Linaro
  - OpenHPC
  - ARM/Allinea
  - Others

Drivers for ARM Software Stack

Sandia National Laboratories

1. Focus on NNSA/ASC's unique requirements
   - Many programs are already exploring HPC on ARM (e.g., Mont-Blanc, Post-K)
   - Need to demonstrate full multi-physics applications running at scale on ARM
2. Build an integrated team
   - Integrate work across tri-labs and vendors, managed as single project
   - Focus on strengths of each laboratory, support technology deployment
3. Provide a robust "traditional" HPC software stack
   - Meet user expectations, provide familiar environment
   - ARM is just an ISA; much of system software stack should "just work"
4. Engage with users
   - Enable low-effort issue reporting, rapidly resolve, get and provide feedback
   - Build community through talks, training, hackathons, workshops, etc.
5. Look forward
   - Improve support for real-world workflows and data-centric computing
   - Seek to reduce friction points for users when moving between platforms
   - Leverage software technology from US DOE Exascale Computing Project

Technology Development Areas

Sandia National Laboratories

- 2019 ARM platform is **not** a production system – increased latitude for experimentation and R&D activities
  - Need to address options for scheduling and types of access to the system to provide for full scale projects, alternate environments, etc.
- Plan to pursue several technology development areas:
  - On-package high-bandwidth memory
  - Advanced HPC Interconnect
  - Compilers, math libraries, tools, task-based runtimes, ...
    - Vendor framework that OpenHPC components get plugged into
    - OpenHPC framework that vendor components get plugged into
  - Experimental OS/R stacks
    - Lightweight kernels (Hobbes/Kitten, Intel mOS, RIKEN McKernel)
    - KVM support for cyber “simulate the Internet” frameworks
    - Support for HPC + Analytics (including AI and ML) workflows and compositions
    - More interactive IaaS-style usage models

Required Vendor Support

Sandia National Laboratories

- Modular system software environment architected to facilitate site modifications and additions.
- Documented APIs for integrating new software components with vendor stack (e.g. RAS APIs, Network, I/O, Health of subsystems,…)
- Ability to partition the system and manage multiple OS images
  - Carve out dedicated resources for R&D activities
  - Boot new or experimental software environments
- Buildable source – High Priority
  - Ability for labs to debug and fix system software issues
    - For example, be able to debug MPI, PMI, Lustre, Burst Buffers, app load, …
  - Ability to rebuild vendor Linux kernels with modified config options
    - Requires vendor to provide source for all “binary-only” kernel drivers and all non-standard Linux kernel patches
  - Ability to build and evaluate experimental OS/R stacks

Recap: Goals of NNSA/ASC Vanguard Project

Sandia National Laboratories

- Key Requirements for this 2019 Platform
  - System prototype for future leadership class DOE platform
  - Competitive HPC 64 Bit ARM processor technology
  - Integrated On-package Memory
  - First of a kind, Advanced Interconnect technology
  - Large focus on maturing the ARM software ecosystem
- Opportunities for *holistic co-design* of innovative test hardware
  - Logic in NIC to improve interconnection network performance
  - Logic in/near memory to support sparse linear algebra acceleration
  - Other advanced processor architectures

Acknowledgements

- **NNSA/ASC Program:** Mark Anderson, Douglas Wade, Thuc Hoang
- **Advanced Architecture Testbeds Operations Team:** John Noe, Jim Brandt, Ann Gentile, Victor Kuhns, Nate Gauntt, Mike Aguilar
- **SST Developers:** K. Scott Hemmert, Branden Moore, Gwen Voskuilen, Amro Awad, Clay Hughes, Jeremy Wilke
- **SST SQE team:** Jon Wilson, Aaron Levine, John Van Dyke
- **Tri-lab Vanguard team:** Jim Laros, Kevin Pedretti, Si Hammond, Rob Hoekstra, Ron Brightwell, Ken Alvin, Trent D'Hooge, James Foraker, Rob Neely, Becky Springmeyer, Bronis de Supinski, Mike Lang, Pat McCormick
- **Cavium:** Gopal Hegde, Rishi Chugh, Larry Wikelius, David Hass, Giri Chukkapalli
- **ARM:** Nigel Paver, Eric Van Hensbergen, Geraint North
- **HPE:** Mike Vildibill, Nic Dube, Kelly Pracht
- **Cray:** Duncan Roweth, Dan Ernst, Larry Kaplan, Heidi Poxon

Sandia
National
Laboratories

Exceptional service in the national interest

Current NNSA Architecture Strategy