SAND2015-1224C

Exceptional service in the national interest

Sandia National Laboratories

Getting Software Ready for Future Systems

ADVANCED ARCHITECTURE TEST BEDS

James H. Laros III
Scalable Computer Architectures
Sandia National Laboratories
jhlaros@sandia.gov

U.S. DEPARTMENT OF ENERGY | NNSA

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. SAND NO. 2011-XXXXP

Overview

- Advanced Architecture Test Be
- Center of Excellence (COE) CORAL
- Trinity (ATS-1) NRE

Advanced Architecture Test Beds

- *Serve as a scout for future com[illegible]tectures*
  - **FOR OUR APPLICATIONS!!!**
- Test beds act as a con[illegible]ersation and co-design
  - With: Labs (ASC and [illegible], Universities, etc.
- Testing early hardw[illegible]ts both ATS and CTS programs
  - Early MIC/Phi and [illegible]well) -> ATS-1
  - Early Power [illegible]2
- Exploratory R[illegible]
  - Alternative R[illegible]ning Models
  - Architecture A[illegible]e Algorithms
  - Advanced Memory sub-system development, and use
  - Energy Efficient Hardware, Runtime, Systems Software
    - **Ultimately Preparing our Applications**

Test Bed Timeline

Arthur

Compton

Compton V2

Volta (XC)

Haswell 16 Core
+ PowerInsight V2

Kaveri+
PowerInsight V2

Shepard

Teller

Teller V2

Curie (XK6)

Curie (XK7) V2

Shannon

Shannon V2

SandyBridge
Kepler K20 + K40 + K80

Shannon V3

Power 8

Watson

Power 8

ESP

Power 8

White

64bit

ESP

HMC

JENGA

64bit

Hammer

Sept.
2011

Sept.
2012

Sept.
2013

June
2014

Sept.
2014

Center Of Excellence (COE)

- Partner with integrator and mor technology provider (Intel for ATS-1, IBM for ATS-2)
  - Code teams directly in tou experts
  - Start work on emulator or white box systems as available
- Focus on one product on from each of the Tri-labs
  - It is critical to include all three NW labs
- Facilitate transition eneration platforms and address ATS code migration
  - MPI+X
  - Transition fr core to many-core
  - High speed (ba dth) on-chip memory subsystem
  - Heterogeneous parallelism issues: tasks, threads, vectors, data locality, etc…
- Target efficient performance in FY17 (for Trinity)

Trinity (ATS-1) NRE

- Non-recurring engineering
- Enable new architecture for [illegible] production computing environment
- Burst Buffer
  - Tightly coupled solid [illegible] age for checkpoint restart file IO
  - Enable improved [illegible] lution
  - Data Analytics
- Advanced Power Management
  - Enable measurement and control at system, node and component levels
  - Allow exploration of application performance per watt (for example)
  - Reduce or control total cost of ownership

Backup

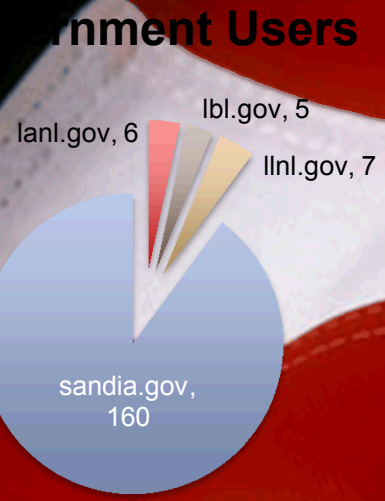
Recent Additions Detailed

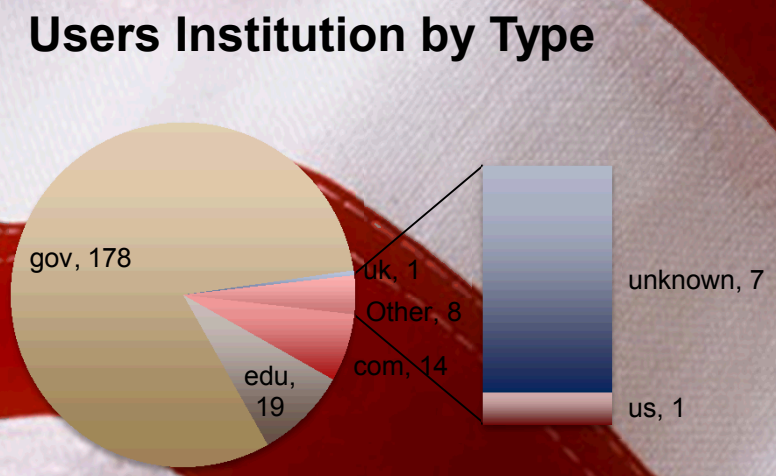
- C
  - 2 node dual socket Sandy Bridge/Nvidia GPU clu
    - 6 nodes 2 K20's per node
    - 12 nodes 4 K20's per node
    - 4 nodes 4 K40's per node
    - 9 nodes 2 K80's per node
- Shepard
  - Single Heterogeneous Cluster
  - FDR IB
  - 36 Node Haswell Xeon (16 core)
    - Instrumented with PowerIns
  - 36 Node AMD Kaveri APU
    - Instrumented with
- White
  - 10 node Power 8/
  - FDR IB (coming soo
  - 1 2U Dual Socket Po
  - 9 4U Dual Socket Pow
    - 2 K40 GPUs per nod
- Hammer
  - 45 node ARM cluster
  - 1 Xgene 1, 64bit ARM processor per node
  - 10Gb Ethernet
- Morgan
  - Evolving Heterogeneous cluster on SRN
  - As technologies mature they are introduced to the SRN to be exposed to REAL applications

Test Bed Users

Breakdown

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





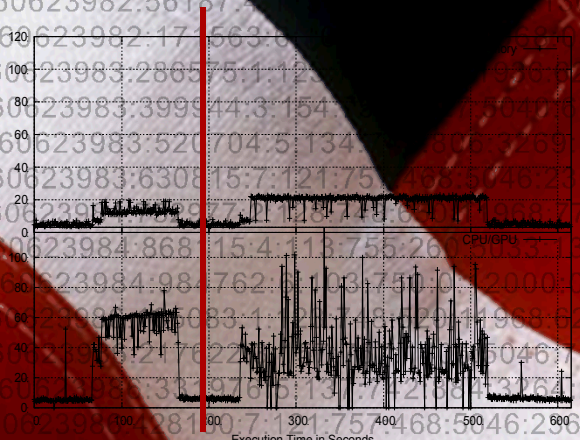
Sandia/Penguin Collaboration

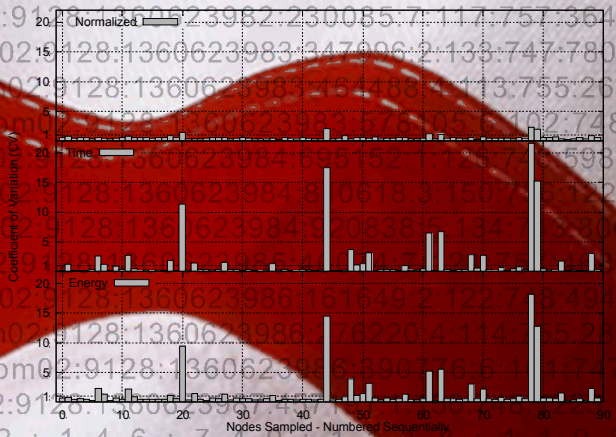
small example of co-design

Commodity

ustom PCB

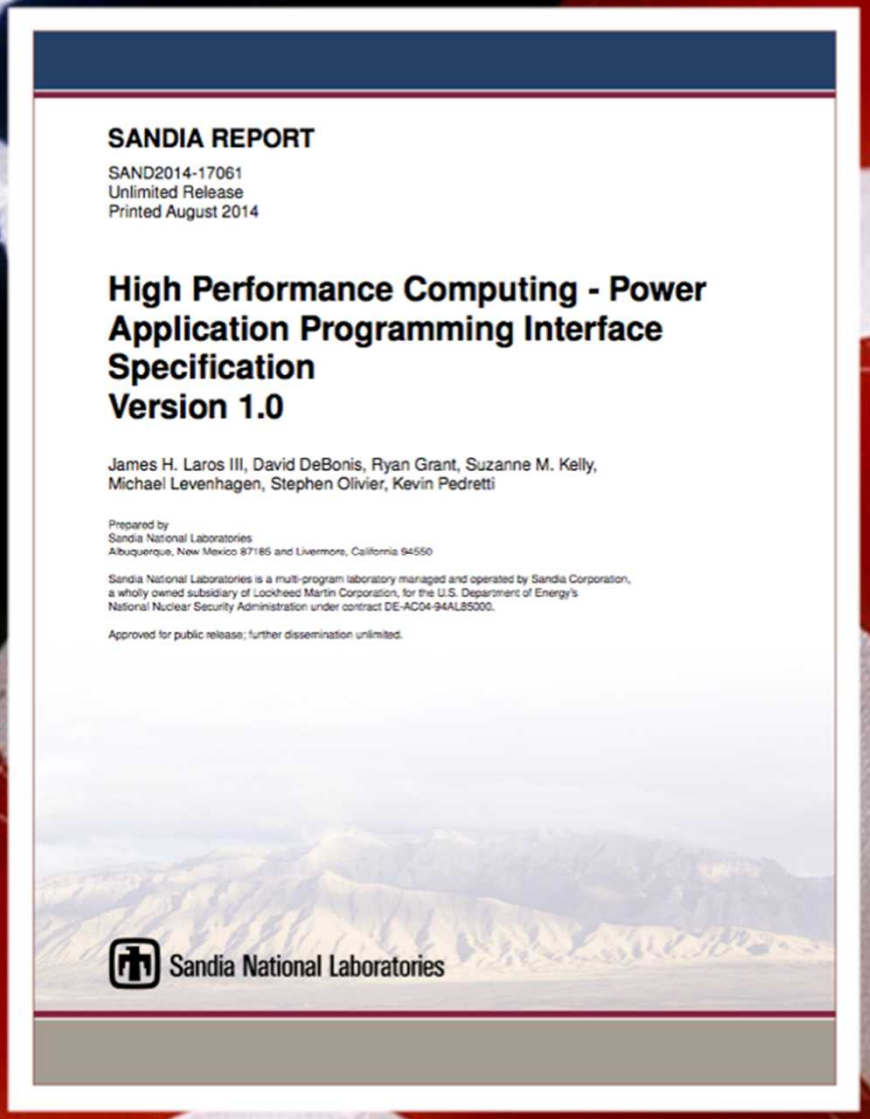
PowerInsight





Version 1.0 Delivered August 26th 2014

- Community needed a portable API for **measuring** and **controlling** power and energy
- Sandia developed PowerAPI specification to fill this gap
- Provides portable power **measurement** and **control** interfaces
- Covers full spectrum facility to compone
- First production implementation will be Trinity (ATS1)
- Continued (increasing) community involvement and influence



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

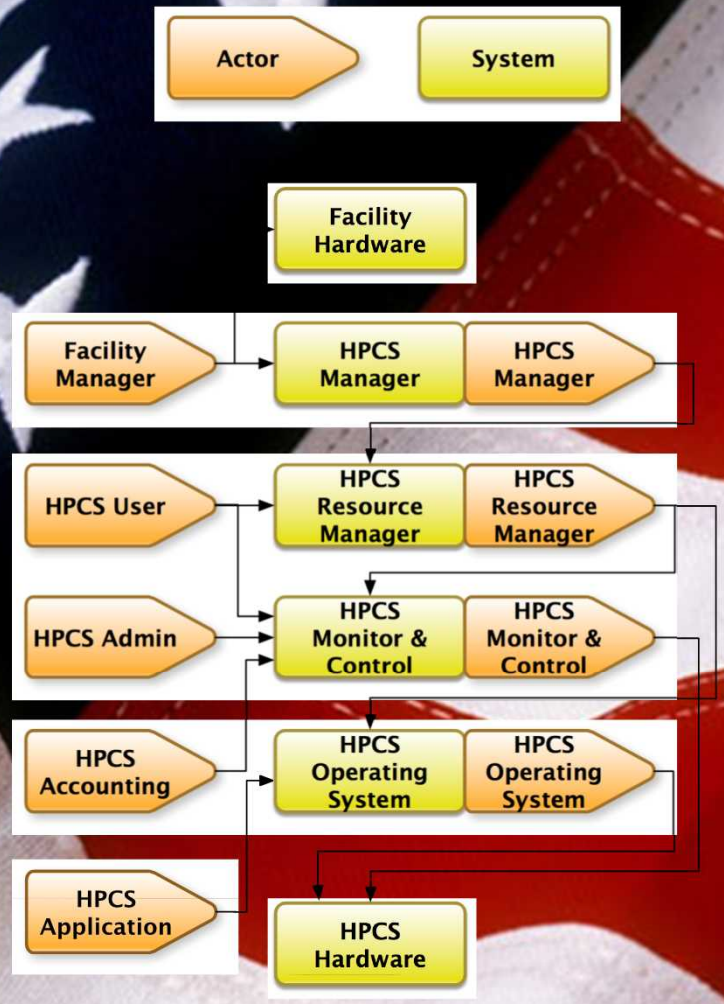
Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

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

Power API Interface Architecture



- h[illegible] efforts do you have in place for suppo[illegible]anced Technology machine and [illegible]he timeline for preparing the software e[illegible]the machine?
  - Discussed much of this in my presentation. Other efforts like SST and more targeted things like runtime and power research
- 2) What are the key issues that yo[illegible]chine? (Tools environment or systems related?)
  - Workflows (as Si mentioned), Bill Archer comment. I'd like to better understand what this means
  - Increasing the efficiency of our mission applications
  - Power, we have an opportunity for a unified standard approach, we shouldn't let it pass
- 3) What are the main areas of co[illegible]t? (e.g. Porting codes, characterizing performance, analyzing bottlenecks and lo[illegible]es, debugging at scale)
  - People, test beds would not have been successful without the people resource to use them, Rob and Si both mentioned this, ALLEN mentioned some ways this is being addressed, hard problem when we are competing with very deep pockets of industry. Expose the labs to students. My story of visiting.
- 4) How do you do in[illegible] validation of the systems environment before the machine arrives? (e.g. Simula[illegible]s, on-site at the vendor's facility)
- 5) What collaborations d[illegible]u have in place or plan to initiate to help support this effort?
  - FF, DF, test-beds, procurements all have strengthened our collaborations with the vendor community. Interlab, ACES has formed a stronger whole than the individual parts. ACES impacting CORALS NRE deliverables?
  - How to take these efforts the final yards to impact out platforms and benefit our mission
- 6) Bottom line: how will you help the users get their codes ported and help the end users run effectively on the machine?
  - Extending test-bed program into restricted networks to provide early resource for mission codes