Trinity
Center of Excellence

Managed by: New Mexico
Alliance for Computing at Extreme Scales (ACES)

NNSA ASC tri-lab simulation community

Los Alamos NATIONAL LABORATORY — Sandia National Laboratories — Lawrence Livermore National Laboratory

COE Leads: Hai Ah Nam — Rob Hoekstra, Mike Glass — Shawn Dawson

DOE CoE Performance Portability Workshop

April 19, 2016

UNCLASSIFIED LA-UR-16-22721

Trinity Advanced Technology System

COMPUTE NODES	
Intel "Haswell" Xeon E5-2698v3	**Intel Xeon Phi "Knights Landing"**
9436 nodes	> 9500 nodes
Dual socket, 16 cores/ socket, 2.3 GHz	1 socket, 60+ cores, > 3 Tflops/KNL
128 GB DDR4	96 GB DDR4 + 16GB HBM

#6 on Top500
November 2015
8.1 PFlops
(11 PF Peak)

Cray Sonexion Storage System
78 PB Usable, ~1.6 TB/s

Cray DataWarp
576 Burst Buffer Nodes
3.7 PB, ~3.3 TB/s

Cray Aries 'Dragonfly' Interconnect
Advanced Adaptive Routing
All-to-all backplane & between groups

Trinity - Performance (Portable) Challenges

COMPUTE NODES	
Intel "Haswell" Xeon E5-2698v3	**Intel Xeon Phi "Knights Landing"**
9436 nodes	> 9500 nodes
Dual socket, 16 cores/ socket, 2.3 GHz	1 socket, 60+ cores, > 3 Tflops/KNL
128 GB DDR4	96 GB DDR4 + 16GB HBM

- **Enabling (not hindering) Vectorization**
- **Increase parallelism, cores/threads**
- **High Bandwidth Memory**
- **Burst Buffer – reduce I/O overhead**

#6 on Top500
November 2015
8.1 PFlops
(11 PF Peak)

Cray Aries 'Dragonfly' Interconnect
Advanced Adaptive Routing
All-to-all backplane & between groups

Cray Sonexion Storage System
78 PB Usable, ~1.6 TB/s

Cray DataWarp
576 Burst Buffer Nodes
3.7 PB, ~3.3 TB/s

Trinity – Challenges/Opportunities

COMPUTE NODES	
Intel "Haswell" Xeon E5-2698v3	**Intel Xeon Phi "Knights Landing"**
9436 nodes	> 9500 nodes
Dual socket, 16 cores/ socket, 2.3 GHz	1 socket, 60+ cores, > 3 Tflops/KNL
128 GB DDR4	96 GB DDR4 + 16GB HBM

#6 on Top500
November 2015
8.1 PFlops
(11 PF Peak)

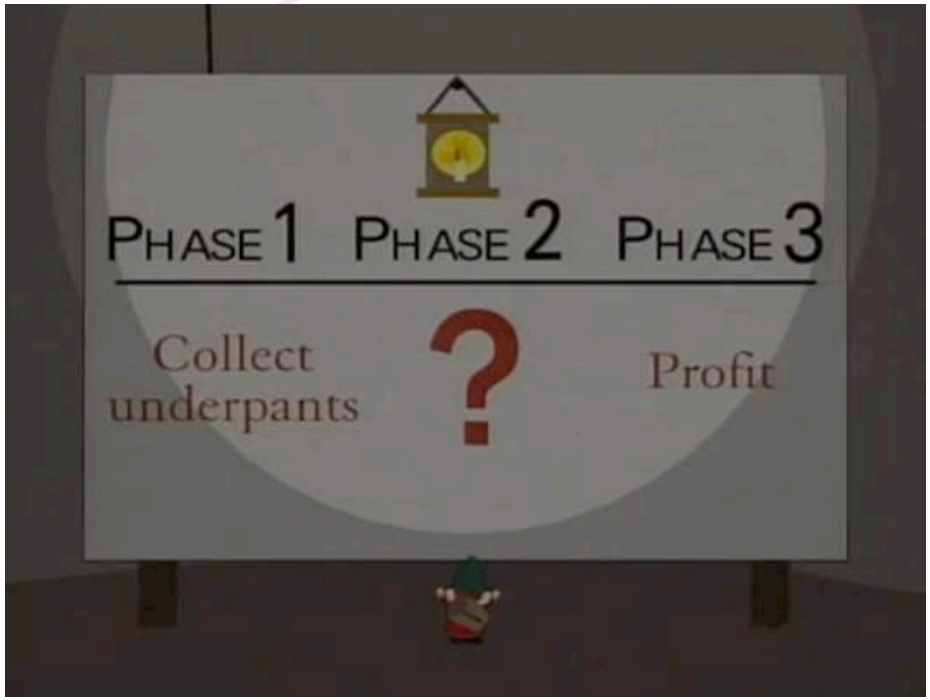
- **Scale and scaling**
- **Dual partition – new workflow & simulation capabilities**
- **Parallel FS – new Lustre DNE capabilities to improve performance**
- **BB - enable new workflow capabilities**
- **Cross compiling (impacts productivity)**

Cray Aries 'Dragonfly' Interconnect
Advanced Adaptive Routing
All-to-all backplane & between groups

Cray Sonexion Storage System
78 PB Usable, ~1.6 TB/s

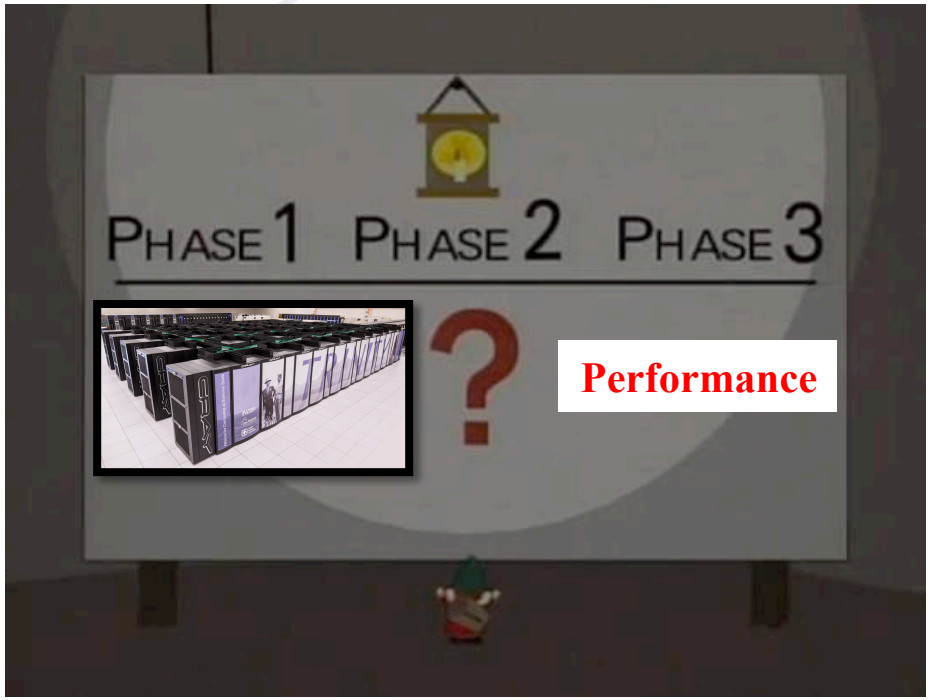
Cray DataWarp
576 Burst Buffer Nodes
3.7 PB, ~3.3 TB/s

The Master Plan



Source: http://southpark.wikia.com/wiki/Underpants_Gnomes

Phase 2 …

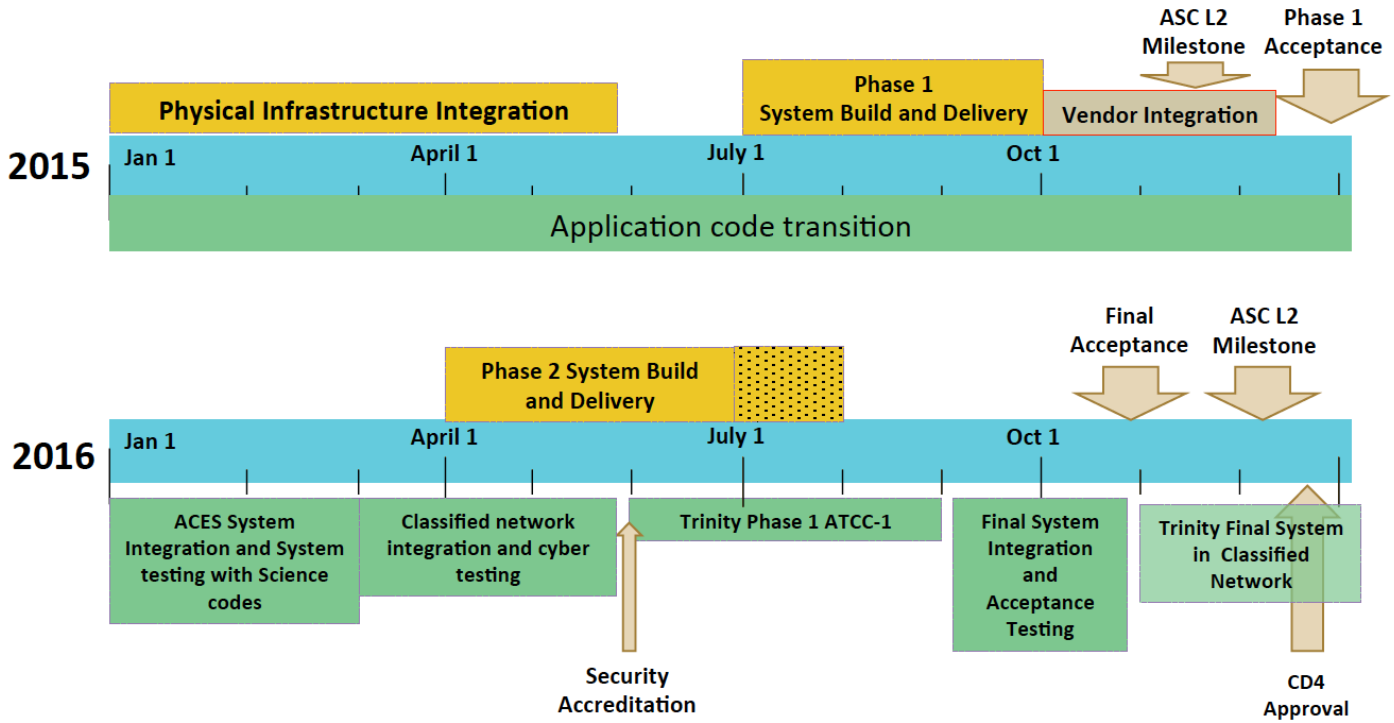


Source: http://southpark.wikia.com/wiki/Underpants_Gnomes

Phase 2

- Early access HW/SW
- Collaborating with COE vendor partners, early, often and with complete honesty
  - Kernel
  - Mini-App
  - Proxy
  - 
- Sharing our concerns
- Communicate

Access to Early HW/SW



- Application Regression Test Beds x2 (Cray) ~100 nodes (June 2015), Software Dev. Testbed < 100 nodes – Phase I, upgrades for Phase II
- White Boxes (Intel) ~ few nodes (Sept 2015/April 2016)

COE Collaborations

- Cray
  - John Levesque (50%)
  - Jim Schwarzmeier (20%)
  - Gene Wagenbreth (100%) - new
  - Mike Davis (SNL), Mike Berry (LANL) on-site analyst
  - SMEs (Performance & Tools)
  - Acceptance team
- Intel
  - Ron Green, on-site analyst (SNL/LANL)
  - Discovery Session, Dungeons - SMEs

- ASC codes are often export controlled, large and complex = a lot of paperwork
- Embedded vendor support/expertise is needed = US citizenship
- Original projects focus on a single code/lab

CoE Projects/Highlights

- SNL
  - Focused on preparing the Sierra engineering analysis suite for Trinity
  - Proxy Codes: miniAero (explicit Aerodynamics), miniFE (implicit FE), miniFENL, BDDC (Domain Decomp. Solver)
  - 'Super' Dungeon Session including
    - More realistic code/stack
      - NALU (proxy application for FEM assembly for low Mach CFD) + Trilinos multi-grid solver, Kokkos + BDDC
    - 6 weeks preparation leading up to Dungeon session
    - Expose Intel to 'real' codes & issues – long compile times, long tools analysis times, compiler issues, MKL issues.
    - Great for relationship/collaboration building
  - More embedded support from Cray (Gene Wagenbreth, March 2016)

CoE Projects/Highlights

- LLNL
  - Developed Proxy Code: Quicksilver (Monte Carlo transport)
    - Dynamic neutron transport problem (MPI or MPI+threads)
    - Use in performance portability activities
    - Proxy codes are not an example of efficient source code, rather a representation of a larger application
  - Discovery Sessions (x2) with proxy applications & performance portable abstraction layer

CoE Projects/Highlights

- LANL
  - Full application exploration – very large, multi-physics, multi-material AMR application (MPI-only)
    - Discovery session (Intel) & Deep dive (Cray) – on-site
    - Prototyping SPMD in radiation diffusion package as an option in code threading implementation
    - Addressing performance bottlenecks in solvers library (HYPRE) & code
    - Addressing technical debt
  - Broadening scope of COE projects to include deterministic Sn transport (full application and proxy)
  - Discovery sessions & deep dive activities

Sharing Best Practices… for now

- COE Tri-Lab Bi-Weekly Meetings/Mailing Lists
  - Logistics, “is anyone else seeing this?”, knlchatter
- COE (monthly) seminar – bringing the outside world in
  - March 2016 – Peter Mendrygal, Cray Performance
  - June 2016 - TBD
- KNL (monthly) working group
  - April 28, 2016 – John Levesque, Cray
- Activities (dungeon, discovery, training)
  - Observers invited