

Trinity Architecture & Design

by the

ACES Design Team

SAND 2011-xxxxP
Unlimited Release
Printed November, 2011

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company,
for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04- 94AL85000.

Mission Requirements

- The DOE NNSA ASC Program requires an Advanced Technology computing system delivered in FY15 to support the national Stockpile Stewardship Program. The name for this platform is “Trinity”.
- The ASC Roadmap defines a path to predictive capability from today to 2020 to support the goals of the ASC Strategy. Its four focus areas are (Trinity directly supports bullets 2 and 4):
 - Address national security simulation needs
 - Establish a validated predictive capability for key physical phenomena
 - Quantify and aggregate the uncertainties in simulation tools
 - Provide mission-responsive computational environments.
- “Work in this time frame will establish the technological foundation to build toward exascale computing environments, which predictive capability may demand” (ASC Roadmap statement)

High-level Design Philosophy

- Trinity is an advanced technology platform with a schedule to be in production operation by Q3 CY2015
- Advanced technology development is assumed to be necessary to meet mission needs
 - Accelerate development of identified key technology areas
 - NRE opportunities exist
- It is highly desirable for Trinity to support programming models that ensure code teams start heading in a direction that is consistent with ASC's exascale plans
- Nominal programming model will be MPI+X
 - MPI for node level parallelism, X for on-node parallelism
 - X programming model needs to be agnostic and portable to a variety of highly threaded architectures, e.g. Multicore, GPGPU, MIC, APU, etc.
 - There may be requirements for multiple choices of X
- Programming models other than MPI+X are of value to Trinity and are expected to provide value to the ASC program
- Trinity will also need to support legacy MPI-everywhere codes, but it may be at a reduced scale and/or computational efficiency

Capability Improvement

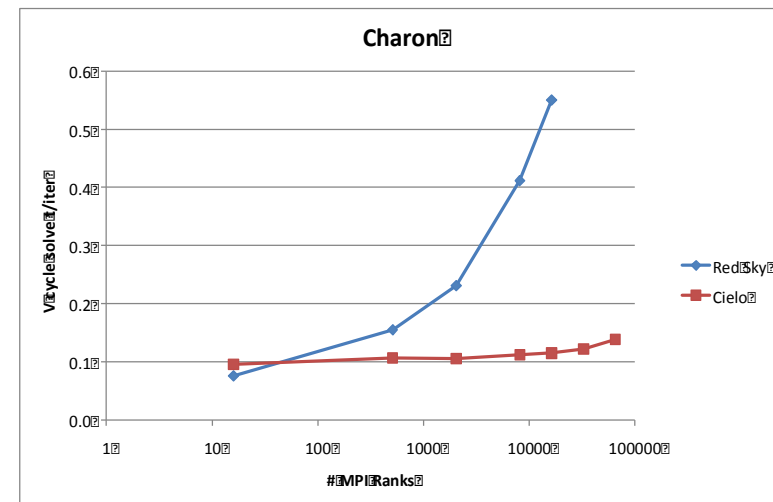
- An increase in predictive capability requires increases in the fidelity of both geometric and physics models.
- Trinity needs to demonstrate a significant capability improvement (10x) over current platforms in key areas of physics in order to meet the needs of the ASC's Predictive Capability Framework (PCF)
 - Improvement is a function of performance (total time to solution), increased geometries and increased physics capabilities
- Increased capabilities drive improvements in computational resources
 - Higher fidelity models -> increases in memory capacity per node & scale
 - While sustaining time to solution -> increased FLOPs, memory bandwidth & scale
- Advanced resilience techniques will play a major role in improving application performance (total time to solution)
- Active power management techniques within the platform will be required to meet the facility and total cost of ownership constraints

Node level characteristics

- Memory
 - An 8x to 16x increase in memory capacity per node
 - Memory bandwidth must increase to meet increasing FLOPs per node
 - Hybrid designs will be considered to meet these needs
 - E.g. High-speed embedded memory + off package DRAM
 - DRAM + Very high-density memory
 - (although non-volatile memory may have security implications)
 - ACES believes off package stacked DRAM will be required to meet capacity and bandwidth demands
- Compute
 - It is assumed that a significant increase in thread level parallelism via many-core and/or accelerator architectures is necessary
 - Nominal programming model is MPI at the node level + a second level (X) within the node
 - OpenMP, Accelerator directives, Posix threads, TBB, Cuda, OpenCL, etc.

Scalability and High-Speed Interconnect

- Scalability is a key requirement
 - 1 to 20 applications running simultaneously
- Low overhead, consistent operation more important than microbenchmark results
 - Network semantics should match the communication models used on the machine
- Scalable resource management
 - No single points of management/failure
 - Fabric management
 - Buffer management
 - Scalable flow control mechanisms
- Low overhead, low noise operating system and network stack
- Novel architectures
 - Integrated NICs
 - Collective offload
 - Matching offload



Application results (lower is better) from two platforms with similar microbenchmark results but differing scalabilities.

Resilience

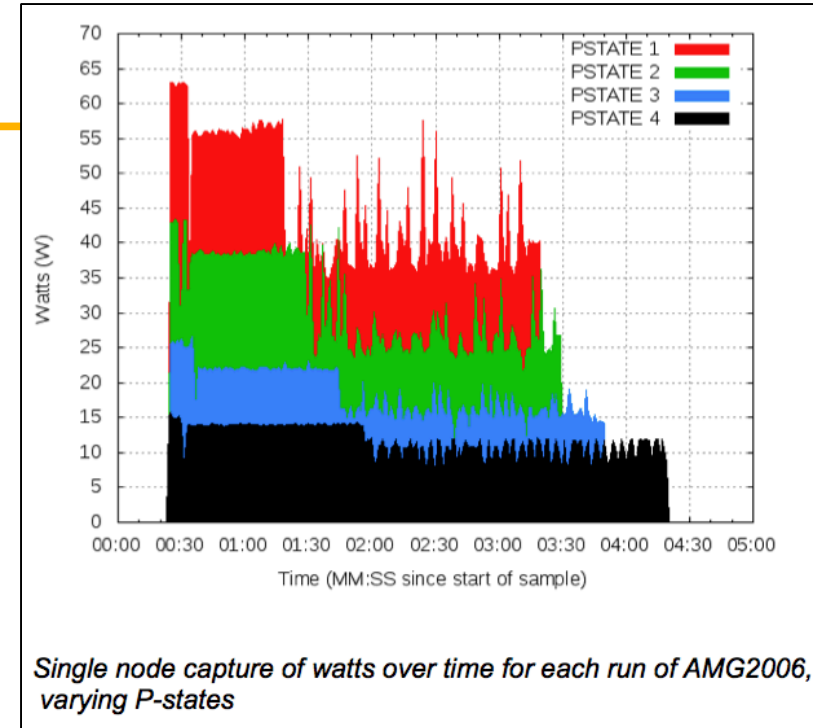
- Checkpoint with Restart will remain as the primary method of resilience
 - Job MTTI must be > 25 hours in order to meet application productivity requirements and not require excessive parallel file system BW
 - High speed, reliable, on-platform caching methods are necessary to remove dependence on excessive external bandwidth requirements (See IO slide)
- Resilience methods provided to the application in order to better allow them to recover from node failure, E.g.:
 - Fault Tolerant MPI (MPI-3)
 - Redundant MPI (rMPI)
 - Scheduling and run time techniques that allow applications to recover from failure without losing allocated resources
- Operating System techniques to recover from resource failures, E.g.:
 - Graceful memory failure and recovery
 - Transactional memory capabilities that can be used by fault tolerant algorithms
- And advanced architecture & HW techniques to improve FIT rates and provide redundancy and fault tolerance

File Systems and I/O

- High speed, reliable, on-platform caching methods are necessary to remove dependence on excessive external bandwidth requirements
 - E.g. a two tiered storage architecture:
 - Tightly coupled, high-speed, non-volatile storage providing a fast, reliable storage partition for checkpoint/restart & in-situ analysis
 - Rotating disk based storage will provide high-levels of capacity but at reduced bandwidths, complexities and cost relative to historic Byte/FLOPs ratios
 - E.g. high-speed, relatively high-capacity disk caching at the storage server or controller level
- Cielo has demonstrated the utility of a very tightly coupled, very fast file system (/udsl) for user generated input parameter files & user generated dynamic shared libraries.
 - Optimized for relative small file sizes, 1->N read access pattern
 - Read-only from the compute nodes to enable caching on IO nodes on the high-speed network, but R/W for service nodes providing access to login services
- A globally shared /home, /projects, etc via NFS

Active Power Management

- Power is a constraining factor in the operation of Trinity
- There is a need to manage power at the platform & application level
 - Policy driven
 - Weighted combination of performance & energy
 - Energy caps based on time of day, physical capacity, etc.
- Need to understand & control power
 - Cabinet & component level I & V measurements
 - Scalable collection infrastructure
 - Tunable collection fidelity: cabinets to components
 - Administrative & user accessible interface for feedback and tuning



- P-states (Frequency/Voltage States)
 - P1: 2.1 GHz, 1.25V
 - P2: 1.7 GHz, 1.1625V
 - P3: 1.4 GHz, 1.125V
 - P4: 1.1 GHz, 1.1V
- AMG demonstration on 6,144 nodes of ORNL's Jaguar shows that managing P-States allows for a 32% decrease in energy used while only increasing time to solution by 7.5%

Facility, Power & Cooling

- Trinity will be located in the Nicholas C. Metropolis center (SCC) at Los Alamos National Lab
- Facility power is the primary constraint in the design of Trinity
 - 12MW of power (becomes available in second half CY 2014)
 - 300 lbs per square foot floor loading
 - 10,000 to 12,000 square feet of floor space
- Water cooling is highly desirable
 - Direct (direct to chip or cold plate) is preferred
 - Indirect (e.g. radiator) method is acceptable
 - Tower water (directly from cooling tower) at 32° C is preferred
 - Chilled tower water at 8.5° C is available but less desirable due to additional \$
 - Under floor air at 12.5° C is available to supplement the water cooling method
- Overhead cabling is desired
- Concerns
 - Idle power efficiency
 - Ramp up / Ramp down load on power grid over 2 MW