SAND2015-10955R

Trinity Acceptance
Performance Tests Summary

Mahesh Rajan

12/14/2015

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

Acknowledgements

– **LANL:** Baker, Randy S.; Lang, Michael K. ; Nam, Hai Ah

– **LLNL:** Draeger, Erik W.;

– **SNL:** Domino, Stefan P.; Agelastos, Anthony M.; Shaw, Ryan P.; Stevenson, Joel O.; Vaughan, Courtenay T.;

– **Cray:** Wichmann, Nathan; Nuss, Cindy; Davis, Mike E.; Carrier, Pierre; Olson, Ryan ; Anderson, Sarah

Trinity performance acceptance requirements from SOW

Objective: "Ensuring Real Applications perform well on Trinity is key to success"

Four components:

1. **ASC applications** (Qbox, Partisn and Nalu) at "near full scale" to achieve a Capability Improvement (CI) Metric target of 4x for Phase1 & 4x for Phase2

   CI Metric = problem-size-increase * run-time-speedup
2. **Sustained System Performance (SSP)**
   - Calculated as a geometric mean of performance of 8 apps (TFLOP/sec) times num_nodes_on Trinity (9436)
   - The 8 apps: miniFE, miniGhost, AMG, UMT, SNAP, miniDFT, GTC and MILC
3. **Extra-Large MiniApplication** Problems: To demonstrate scalability and Performance; shall scale to as large a set of compute nodes as practical

   The 5 apps: miniFE, miniGhost, AMG, UMT and SNAP
4. **Micro-benchmarks:** Helps with baselining some important performance characteristics

   The 8 micro-benchmarks: Pynamic, Ziatest, OMB, SMB, mdtest, IOR, PSNAP, mpimemu

Capability Improvement (8x) Performance Applications

Lab	Code	Fortran	Python	C	C++	MPI	OpenMP	Description
SNL	Nalu	X		X	X	X		The SIERRA Nalu is a Low Mach CFD code used to solve a wide variety of variable density acoustically incompressible flows spanning from laminar to turbulent flow regimes.
LANL	Partisn	X				X	X	PARTISN provides neutron transport solutions on orthogonal meshes in one, two, and three dimensions. Reference: "An Sn Algorithm for the Massively Parallel CM---200 Computer", Randal S. Baker and Kenneth R. Koch, Nucl. Sci. and Eng., Vol. 128, pp. 313---320 (1998),
LLNL	Qbox			X	X	X	X	Qbox is a first principles molecular dynamics code used to compute the properties of materials at the atomistic scale. The main algorithm uses a Born---Oppenheimer description of atomic cores and electrons, with valence Electrons treated quantum mechanically using Density Functional Theory and a plane wave basis.

8x Application Improvement Relative to Cielo (Split into 4x Phase 1 & 4x Phase2; WEAK SCALING)

- **Capability Improvement(CI) = (improvement in runtime) * (increased problem size)**
- **Runtime metrics are chosen to measure platform performance, not algorithmic performance**
- **E.g., A Trinity run with 8x the problem size as the Cielo baseline problem, and the improvement in runtime is 1.33x (i.e. Trinity took ¾ the time on Cielo), *CI = 8x * 1.33x = 10.6x***

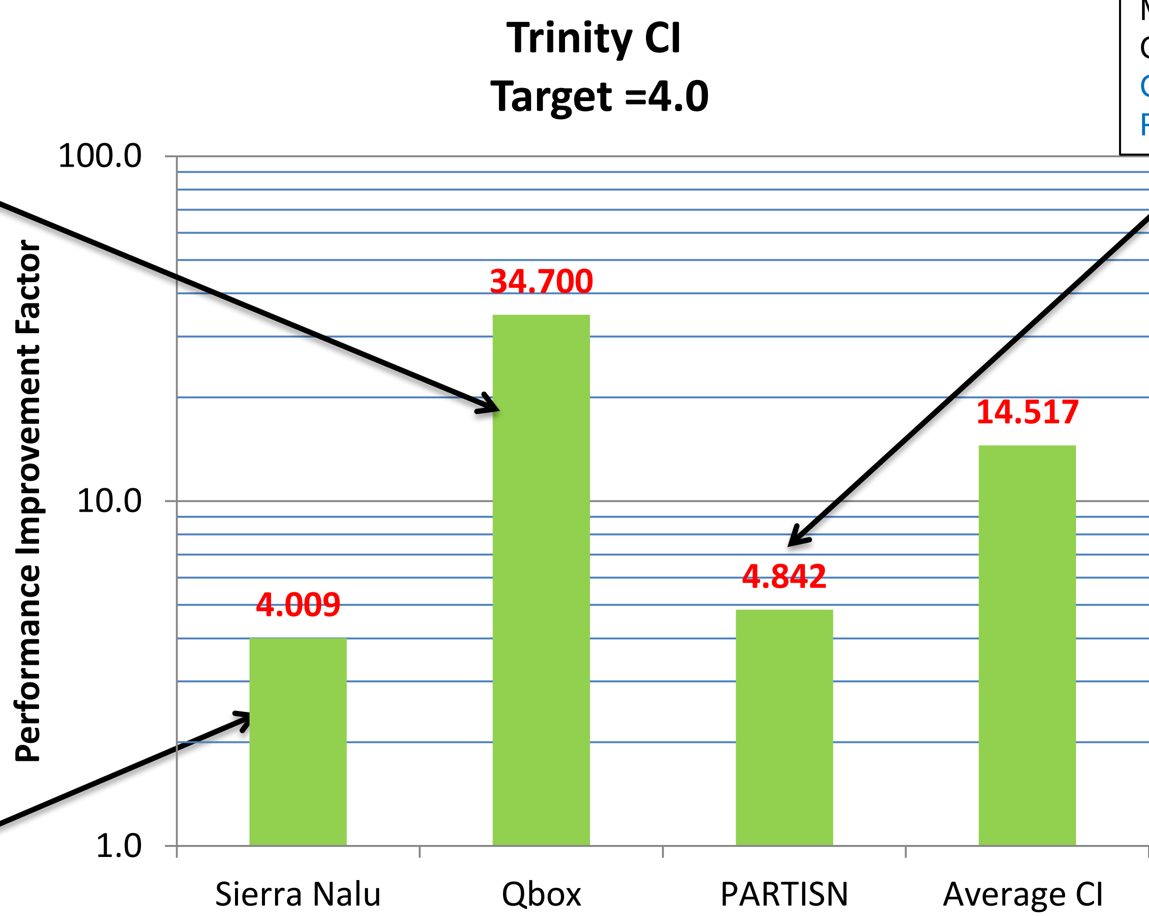
Performance Metric	Cielo	Trinity	Ratio
Application Performance	1x	Target 4x over Cielo: Phase1	4x
Number of nodes	8944	9436	1.15x
Number of cores	143, 104	301,952	2.3x
Peak FP	1.37 PF	~11 PF	8.03x
Peak Memory BW	763 TB/s	1288 TB/s	1.68x
Total Memory Capacity	298.2TB	1,207.8 TB	4x
Memory per node	16 GB	128 GB	4x
Memory per core	2 GB	4 GB	2x

Applications CI Performance Summary

(Tuning achieved with optimal OMP tasks/thread, rank_order & input mods)

PARTISN:
Nodes, = 9418
MPI Tasks = 37,672
OMP Threads/task = 4
Complexity Increase = 9.2
Run Time Ratio = 0.526

Qbox:
Nodes, = 9418
MPI Tasks= 32,762
OMP Threads/task=8
Complexity Increase = 166.3
Run Time Ratio = 0.208



Nalu:
Nodes, = 9300
MPI Tasks = 297,600
OMP Threads/task=None
Complexity Increase = 1
Run Time Ratio = 4.009

Sustained System Performance (SSP) Metric; Target Phase 1= 400; Achieved 440

Trinity SSP						Rate(TF/s per Node)
Application Name	MPI Tasks	Threads	Nodes Used	Reference Tflops	Time (seconds)	Pi
miniFE(Total CG Time)	49152	1	1536	1065.151	34.14	0.020312176
miniGhost(Total time)	49152	1	1536	3350.20032	1.77E+01	0.122949267
AMG(GMRES Solve wall Time)	49152	1	1536	1364.51	66.233779	0.013412384
UMT(cumulativeWorkTime)	49152	1	1536	18409.4	454.057	0.026395995
SNAP(Solve Time)	12288	2	768	4729.66	1.77E+02	0.034793285
miniDFT(Benchmark_time)	2016	1	63	9180.11	377.77	0.385726849
GTC(NERSC_TIME)	19200	1	600	19911.348	868.439	0.038212908
MILC(NERSC_TIME)	24576	1	768	15036.5	393.597	0.049743204
					Geom. Mean=	0.046682671
					SSP=	**440.497682**

Additional Trinity Performance acceptance tests as per SOW

1. **Extra-Large MiniApplication** Problems: To demonstrate scalability and Performance; scale to as large a set of compute nodes as practical
   a) miniFE ; 9000+ node run attempted; SegFaults; being investigated
   b) miniGhost; Global Grid of 14976^3 run completed 9216 nodes; Time (with reorder)=7.951e+01
   c) AMG; completed on 9417 nodes with 301,344 MPI tasks; Solver time=152 seconds for 77 iterations
   d) UMT; completed on 3,888 nodes with 124,416 MPI tasks; CumulativeIterationCount=119; CumulativeWorkTime=498.333 s
   e) SNAP; completed on 3072 nodes using 98,384 cores; Solve Time=257 s ; (larger SNAP problem fail: The cause of failure is that the MPI_TAG value in SNAP (greater than 3 million) exceeds CRAY MPI limit of 2^21 (2097152).
2. **Micro-benchmarks:** Helps with baselining some important performance characteristics

   The 8 micro-benchmarks: Pynamic, Ziatest, OMB, SMB, mdtest, IOR, PSNAP, mpimemu have been run on Trinity at different scales.