LC Systems Update

LC User Meeting

September 8, 2016

David Fox, LC System Administration Group Lead

Lawrence Livermore National Laboratory

LLNL-PRES-702944

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Agenda

- New Systems
- Systems Retiring
- HPC System Summary

CTS1

CTS Systems Highlights

- Compute Nodes
  - Dual Socket Intel E5-2695 v4 (Broadwell)
  - 18 cores @ 2.1GHz per socket (36 cores/node)
  - 128GB DDR4 @ 2.4GHz DRAM
- Intel Omni-Path 100GB/s interconnect
- 1SU = 192 total nodes, 232 TF/s, 23.5TB DRAM
  - 1 Management node
  - 1 Login node
  - 4 Network router/gateway nodes
  - 186 Compute nodes

CTS Systems

Name	Size	Network
Quartz	14 SU	CZ
RZTopaz	4 SU	RZ
Borax	48 nodes (serial)	CZ
RZTrona	20 nodes (serial)	RZ
Jade	14 SU	SCF
Mica	2 SU	SCF
Agate	48 nodes (serial)	SCF

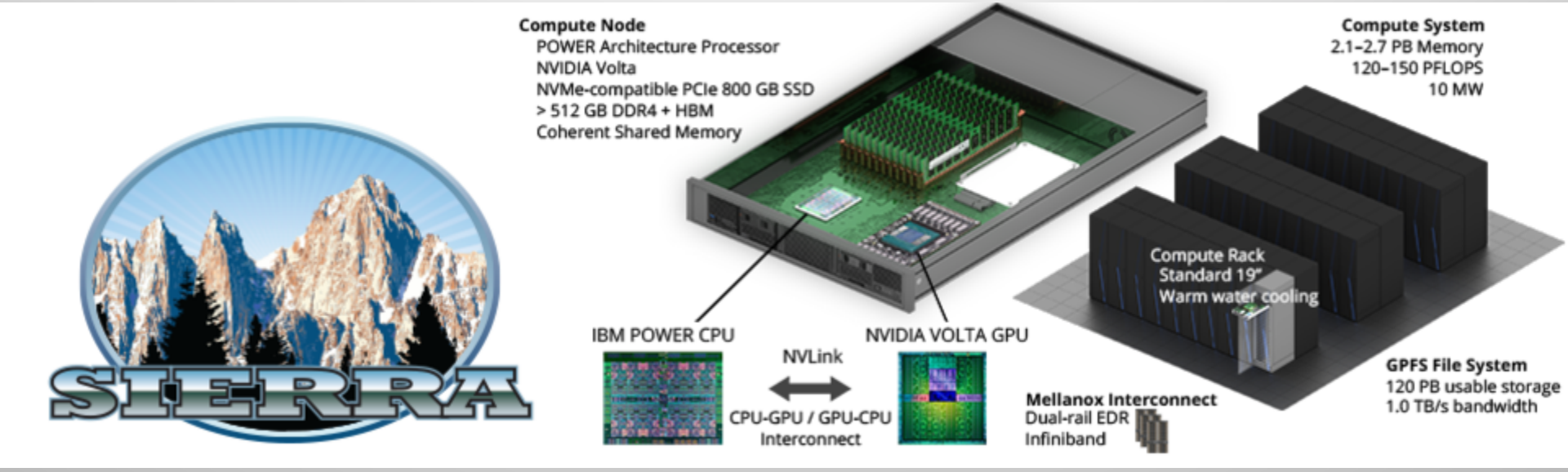
CTS Time Line

- Quartz – New building (B654), New Lustre, will require additional system testing, should be in Limited Availability (LA) mode in Oct.
- RZTopaz – New building, dependent on new water cooling system that is just being installed in B654
- Borax – Serial cluster, but will Initially bring up as a parallel cluster (tossdev) for early access to TOSS 3 in the CZ, expected mid-Sep.
- RZTrona – Serial cluster, LA late Sep.
- Jade – In testing on the open, expect to move to the SCF in late Sep, will run about 6 weeks in LA mode
- Mica – Delivered, waiting facility perperation
- Agate – Serial cluster, but will Initially bring up as a parallel cluster (ctossdev) for early access to TOSS 3 in the SCF, expected mid-Sep.

Systems Retiring

Retiring System	Replacement System
Sierra	Quartz
Ansel	Quartz
Aztec	Borax
RZZeus	RZTopaz
RZCereal	RZTrona
Inca	Agate
Muir	Jade
Juno	Jade

Sierra


SIERRA
Compute Node
POWER Architecture Processor
NVIDIA Volta
NVMe-compatible PCIe 800 GB SSD
> 512 GB DDR4 + HBM
Coherent Shared Memory
IBM POWER CPU
NVLink
NVIDIA VOLTA GPU
CPU-GPU / GPU-CPU
Interconnect
Compute System
2.1–2.7 PB Memory
120–150 PFLOPS
10 MW
Compute Rack
Standard 19"
Warm water cooling
Mellanox Interconnect
Dual-rail EDR
Infiniband
GPFS File System
120 PB usable storage
1.0 TB/s bandwidth

Sierra Early Access Highlights

- **Nodes**
  - Dual Socket Power 8+
  - 10 cores/socket @ 2.92Ghz
  - NVLINK 1.0
  - 4 nVidia GPUs/node
  - 256GB DRAM
  - 2 100Gb/s Mellanox EDR HCA
- **CZ system only**
  - 1.6TB NVMe (PCI addached solid state storage)
- **IBM Spectrum Scale (GPFS File System)**
- **IBM Spectrum LSF (batch system)**

Sierra EA Systems

Name	Nodes	Compute Nodes	Network	Est. Available *
Ray	59	54	CZ	January 2017
RZManta	41	36	RZ	October
Shark	41	36	SCF	November

*** Very new hardware and software, will have lots of work to do to make this system usable and integrate into LC environment. Also expect a major software update from IBM at the end of 2016.**

HPC at LLNL – As of Nov 2014

System	Top500 Rank	Program	Manufacture/ Model	OS	Inter-connect	Nodes	Cores	Memory (GB)	Peak TFLOP/s
Unclassified Network (OCF)									
Vulcan	9	ASC+M&IC+HPCIC	IBM BGQ	RHEL/CNK	5D Torus	24,576	393,216	393,216	5,033.2
Sierra	417	M&IC	Dell	TOSS	IB QDR	1,944	23,328	46,656	243.7
Cab (TLCC2)	145	ASC+M&IC+HPCIC	Appro	TOSS	IB QDR	1,296	20,736	41,472	426.0
Ansel		M&IC	Dell	TOSS	IB QDR	324	3,888	7,776	43.5
RZMerl (TLCC2)		ASC+ICF	Appro	TOSS	IB QDR	162	2,592	5,184	53.9
RZZeus		M&IC	Appro	TOSS	IB DDR	267	2,144	6,408	20.6
Catalyst		ASC+M&IC	Cray	TOSS	IB QDR	324	7,776	41,472	149.3
Surface		ASC+M&IC	Cray	TOSS	IB QDR	162	2,592	41,500	451.9
Aztec		M&IC	Dell	TOSS	N/A	96	1,152	4,608	12.9
Herd		M&IC	Appro	TOSS	IB DDR	9	256	1,088	1.6
OCF Totals	Systems	10							6,436.6
Classified Network (SCF)									
Pinot(TLCC2, SNSI)		M&IC	Appro	TOSS	IB QDR	162	2,592	10,368	53.9
Sequoia	3	ASC	IBM BGQ	RHEL/CNK	5D Torus	98,304	1,572,864	1,572,864	20132.7
Zin (TLCC2)	66	ASC	Appro	TOSS	IB QDR	2,916	46,656	93,312	961.1
Juno (TLCC)		ASC	Appro	TOSS	IB DDR	1,152	18,432	36,864	162.2
Muir		ICF	Dell	TOSS	IB QDR	1,296	15,552	31,104	168.0
Graph		ASC	Appro	TOSS	IB DDR	576	13,824	72,960	107.5
Max		ASC	Appro	TOSS	IB QDR	324	5,184	82,944	107.8
Inca		ASC	Dell	TOSS	N/A	100	1,216	5,120	13.5
SCF Totals	Systems	8							21,706.7
Combined Totals		18							28,143.3

HPC at LLNL – As of Dec 2016

System	Top500 Rank (6/16)	Program	Manufacture/ Model	OS	Inter-connect	Nodes	Cores	Memory (GB)	Peak TFLOP/s
Unclassified Network (OCF)									
Vulcan	14	ASC+M&IC+HPCIC	IBM BGQ	RHEL/CNK	5D Torus	24,576	393,216	393,216	5,033.2
Cab (TLCC2)	390	ASC+M&IC+HPCIC	Appro	TOSS	IB QDR	1,296	20,736	41,472	426.0
Quartz (14SU)		ASC+M&IC	Penguin	TOSS	Omni-Path	2,688	96,768	344,064	3251.4
RZTopaz (4 SU)		ASC	Penguin	TOSS	Omni-Path	768	27,648	98,304	929.0
RZGenie		ASC	Penguin	TOSS	Omni-Path	48	1,728	6,144	58.1
CZSierraED		M&IC	IBM	RHEL	IB EDR	54	1,080	17,280	896.4
RZSierraED		ASC	IBM	RHEL	IB EDR	36	720	11,520	597.6
RZMerl (TLCC2)		ASC+ICF	Appro	TOSS	IB QDR	162	2,592	5,184	53.9
Catalyst		ASC+M&IC	Cray	TOSS	IB QDR	324	7,776	41,472	149.3
Syrah		ASC+HPCIC	Cray	TOSS	IB QDR	324	5,184	20,736	107.8
Surface		ASC+M&IC	Cray	TOSS	IB FDR	162	2,592	41,500	451.9
Borax		ASC+M&IC	Penguin	TOSS	N/A	48	1,728	6,144	58.1
RZTrona		ASC	Penguin	TOSS	N/A	48	1,728	6,144	58.1
Herd		M&IC	Appro	TOSS	N/A	9	256	1,088	1.6
OCF Totals	Systems	14							12,072.4
Classified Network (SCF)									
Pinot(TLCC2, SNSI)		M&IC	Appro	TOSS	IB QDR	162	2,592	10,368	53.9
Sequoia	4	ASC	IBM BGQ	RHEL/CNK	5D Torus	98,304	1,572,864	1,572,864	20132.7
Jade (14 SU)		ASC	Penguin	TOSS	Omni-Path	2,688	96,768	344,064	3251.4
Mica (2 SU)		ASC	Penguin	TOSS	Omni-Path	384	13,824	49,152	464.5
SCFSierraED		ASC	IBM	RHEL	IB EDR	36	720.0	11520.0	597.6
Zin (TLCC2)	124	ASC	Appro	TOSS	IB QDR	2,916	46,656	93,312	961.1
Max		ASC	Appro	TOSS	IB FDR	324	5,184	82,944	107.8
Agate		ASC	Penguin	TOSS	N/A	48	1,728	6,144	58.1
SCF Totals	Systems	8							25,627.1
Combined Totals		22							37,699.5

Questions?

David Fox
fox18@llnl.gov
925.422-6689