

LC Systems Update

LC User Meeting

August 21, 2018

David Smith, LC System Administration Group Lead



LLNL-PRES-XXXXXX

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

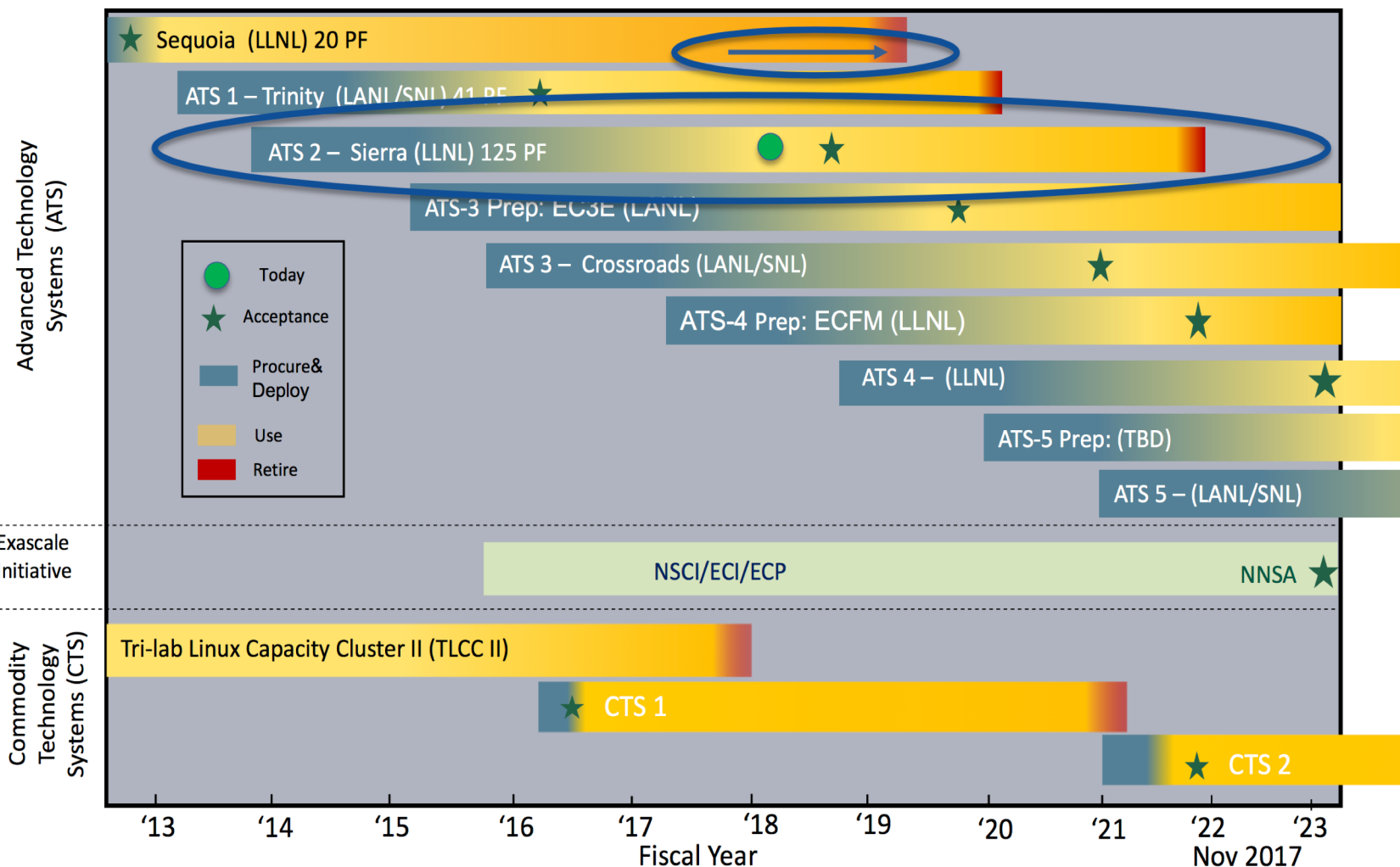


Agenda

- **Advanced Technology Systems (ATS) Hardware Review**
- **System Migrations to TOSS 3**
- **System Retirements**
- **NFS File Systems Update**
- **Other**
- **LC HPC System Summary**

Advanced Technology Systems (ATS)

Sierra



The Sierra system that will replace Sequoia features a GPU-accelerated architecture



Compute Node

2 IBM POWER9 CPUs
4 NVIDIA Volta GPUs
NVMe-compatible PCIe 1.6 TB SSD
256 GiB DDR4
16 GiB Globally addressable HBM2
associated with each GPU
Coherent Shared Memory

Compute Rack

Standard 19"
Warm water cooling

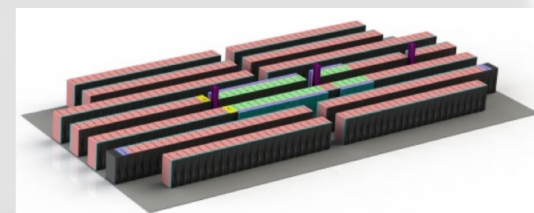
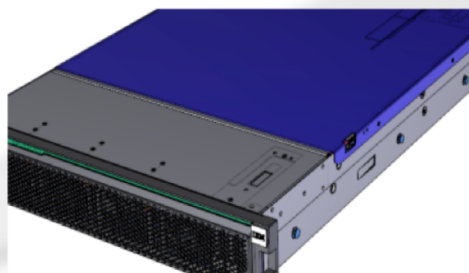
Compute System

4320 nodes
1.29 PB Memory
240 Compute Racks
125 PFLOPS
~12 MW

Components

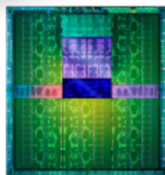
IBM POWER9

- Gen2 NVLink



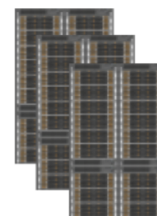
NVIDIA Volta

- 7 TFlop/s
- HBM2
- Gen2 NVLink



Mellanox Interconnect

Single Plane EDR InfiniBand
2 to 1 Tapered Fat Tree



Spectrum Scale File System

154 PB usable storage
1.54 TB/s R/W bandwidth

ATS Architecture Details

	Sierra	Lassen	RZAnsel
Nodes	4,320	684	54
POWER9 processors per node	2	2	2
GV100 (Volta) GPUs per node	4	4	4
Node Peak (TFLOP/s)	29.1	29.1	29.1
System Peak (PFLOP/s)	125	19.9	1.57
Node Memory (GiB)	320(256+64)	320(256+64)	320(256+64)
System Memory (PiB)	1.29	0.209	0.017
Interconnect	2x IB EDR	2x IB EDR	2x IB EDR
Off-Node Aggregate b/w (GB/s)	45.5	45.5	45.5
Compute racks	240	38	3
Network and Infrastructure racks	13	4	0.5
Storage Racks	24	4	0.5
Total racks	277	46	4
Peak Power (MW)	~12	~1.8	~0.14

ATS Deployment Timeline (Sierra, RZAnsel and Lassen)

- **Sierra**
 - 100% hardware integration
 - Acceptance completion scheduled by about 10/1
 - Migration to SCF by about mid January 2019
- **RZAnsel**
 - 100% hardware integration
 - Available now
 - Official acceptance after Sierra
- **Lassen**
 - 100% hardware integration
 - Integration almost complete (testing to start in Sept)
 - Official acceptance after Sierra

System Updates

TOSS 2 to TOSS 3 system migrations

- Surface (COMPLETED)
- Syrah (COMPLETED)
- Max (next week)
- Zin (November)
- RZAlastor (November)

Systems Retiring

Retiring Systems	Retirement Date	Replacement System
Vulcan	11/30/2018	Lassen RZAnsel
Zin*	TOSS 2 (11/2018) TOSS 3 (TBD)	Jade
Sequoia	10/2019 (Estimated)	Sierra

The Life of Zin

- **Extending service of Zin (delay retirement)**
 - Zin's environment has been very stable
 - Service life is typically: 5 years (Zin is 7 years in Oct)
- **Zin timeline of operation**
 - TOSS 2 until November 2018
 - TOSS 3 (2-3 days of downtime in November)
 - Continue to use system (system node count may be reduced for parts)
- **Retirement – subject to:**
 - Hardware issues
 - Floor space needed for new system deployments
- **Zin Facts**
 - First received October 2011 (Tri-Lab Linux Capacity Cluster (TLCC2))
 - 18 Scalable Units
 - 970 teraFLOPs

NFS File Systems Update

- Due to issues with current workspace servers they are being replaced earlier than planned.
- New Servers from Network Appliance
 - Purchased and delivered
 - Integration in progress
- Working on deployment and migration plan (in progress)
- Current plan is for LC to migrate users data in workspace directories from old server to new servers.
- Future phases will impact `/nfs/tmp` and `/usr/mic/post` file systems.
- CZ/RZ process should start in by early to mid September. A Tech Bulletin will be released soon. SCF systems will follow.



Other

- **Starting in FY19 vendor discussions will begin on next generation of systems (CTS-2).**
- **OCF VNC – deployment as a system**
 - System dedicated to VNC will be available by about mid October
- **SNSI Environment – Pinot**
 - Pinot to be replaced by CTS-1 based system (Pinot2)
 - Lustre file system recently expanded
 - NFS server updates planned for future

LC HPC System Summary – August 2018

(<https://hpc.llnl.gov/hardware/platforms>)

System	Top500 Rank	Program	Manufacture/ Model	Processor Architecture	OS	Inter-connect	Avg Power Demand (KW)	Nodes	Cores	Memory (GB)	Peak TFLOP/s
Unclassified Network (OCF)											
Vulcan	33	ASC+M&IC+HPCIC	IBM BGQ	IBM PowerPC A2	RHEL/CNK	5D Torus	TBD	24,576	393,216	393,216	5,033.2
Lassen		ASC+M&IC	IBM	IBM P9	RHEL	2x IB EDR	TBD	684	30,096	218,880	19,886.0
Quartz (CTS-1)	63	ASC+M&IC	Penguin	Intel Xeon E5-2695 v4	TOSS	Omni-Path	TBD	2,688	96,768	344,064	3251.4
Pascal		ASC+M&IC	Penguin	Intel Xeon E5-2695 v4	TOSS	IB EDR	TBD	163	5,868	41,728	1700.0
RZTopaz		ASC	Penguin	Intel Xeon E5-2695 v4	TOSS	Omni-Path	TBD	768	27,648	98,304	929.0
RZManta		ASC	IBM	IBM Power8+	RHEL	IB EDR	TBD	36	720	11,520	597.6
Ray		ASC+M&IC	IBM	IBM Power8+	RHEL	IB EDR	TBD	54	1,080	17,280	896.4
RZAnsel		ASC	IBM	IBM P9	RHEL	2x IB EDR	TBD	54	2,376	17,280	1,570.0
Catalyst		ASC+M&IC	Cray	Intel Xeon E5-2695 v2	TOSS	IB QDR	TBD	324	7,776	41,472	149.3
Syrah		ASC+M&IC	Cray	Intel Xeon E5-2670	TOSS	IB QDR	TBD	324	5,184	20,736	107.8
Surface		ASC+M&IC	Cray	Intel Xeon E5-2670	TOSS	IB FDR	TBD	162	2,592	41,500	451.9
Borax (CTS-1)		ASC+M&IC	Penguin	Intel Xeon E5-2695 v4	TOSS	N/A	TBD	48	1,728	6,144	58.1
RZTrona (CTS-1)		ASC	Penguin	Intel Xeon E5-2695 v4	TOSS	N/A	TBD	48	1,728	6,144	58.1
OCF Totals	Systems	13									34,688.8
Classified Network (SCF)											
Pinot(TLCC2, SNSI)		M&IC	Appro	Intel Xeon E5-2670	TOSS	IB QDR	TBD	162	2,592	10,368	53.9
Sequoia	8	ASC	IBM BGQ	IBM PowerPC A2	RHEL/CNK	5D Torus	TBD	98,304	1,572,864	1,572,864	20132.7
Sierra	3	ASC	IBM	IBM P9	RHEL		TBD	4,320	190,080	1,382,400	125,626.0
Zin (TLCC2)	437	ASC	Appro	Intel Xeon E5-2670	TOSS	IB QDR	TBD	2,916	46,656	93,312	961.1
Jade+Jadeita (CTS-1)	64	ASC	Penguin	Intel Xeon E5-2695 v4	TOSS	Omni-Path	TBD	2,688	96,768	344,064	3251.4
Mica		ASC	Penguin	Intel Xeon E5-2695 v4	TOSS	Omni-Path	TBD	384	13,824	49,152	464.5
Shark		ASC	IBM	IBM Power8+	RHEL	IB EDR	TBD	36	720	11,520	597.6
Max		ASC	Appro	Intel Xeon E5-2670	TOSS	IB FDR	TBD	324	5,184	82,944	107.8
Agate (CTS-1)		ASC	Penguin	Intel Xeon E5-2695 v4	TOSS	N/A	TBD	48	1,728	6,144	58.1
SCF Totals	Systems	9									151,253.1
Combined Totals		22									185,941.9

System Category	TFLOP/s	% of Total
Unclassified	34,688.8	18.7%
Capability	27,983.2	80.7%
Capacity	4,437.5	12.8%
Visualization	2,151.9	6.2%
Serial	116.2	0.3%

System Category	TFLOP/s	% of Total
Classified	151,253.1	81.3%
Capability	146,356.3	96.8%
Capacity	4,730.9	3.1%
Visualization	107.8	0.1%
Serial	58.1	0.0%

Questions?



David Smith
smith107@llnl.gov
925.422-9256

**Disclaimer**

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.