



The Hyperion Project: Collaboration for an Advanced Technology Cluster Testbed

November 2008





Extending leadership to the HPC community



- Motivation
- Collaborations
- Hyperion Cluster
- Timeline

The word **leadership** can refer to:

1. The ability “to get people to follow voluntarily.”
2. Those entities that perform one or more acts of leading.
3. The ability to affect human behavior so as to accomplish a mission designated by the leader.

<http://en.wikipedia.org/wiki/Leadership>



Hyperion Project Goals 1



- Development and testing environment for critical enabling Linux Cluster Technologies
 - Infiniband Open Source Software - OFED
 - Lustre & other Open Source Parallel File System Testing and Scaling
 - Open Source Tri-Lab Linux software stack development & testbed
 - “Intel Cluster Ready” process to push technology out to wider HPC community



Hyperion Project Goals 2



- Evaluation testbed for new hardware & software technologies
 - Petascale I/O technology scaling for Sequoia and future capacity systems
 - Processor, memory, networking, storage, visualization, etc.
 - Designed for future technology refresh, expansion, and upgrades



Hyperion Project Goals 3

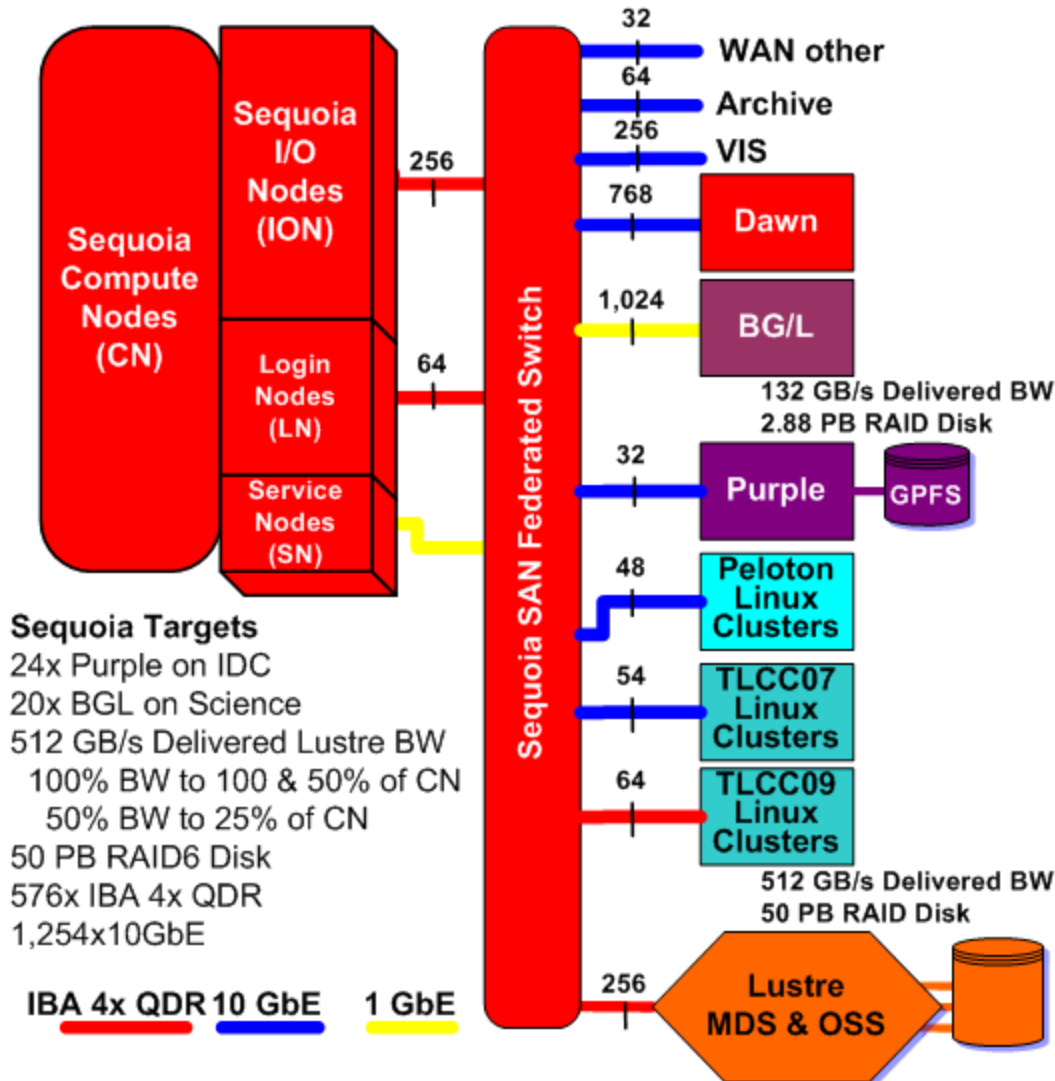


- Innovative approach for forming long term collaborations
 - OpenFabrics Alliance
 - Lustre Center of Excellence
 - Other Gov't agencies, alliances, and computing centers
 - End customers (e.g. Financial services, Oil & Gas, Pharma, etc.)

Sequoia Peta-Scale Simulation Environment Requires the Development of Critical Enabling Technologies



ASC Sequoia Simulation Environment Lawrence Livermore National Laboratory 2010/11

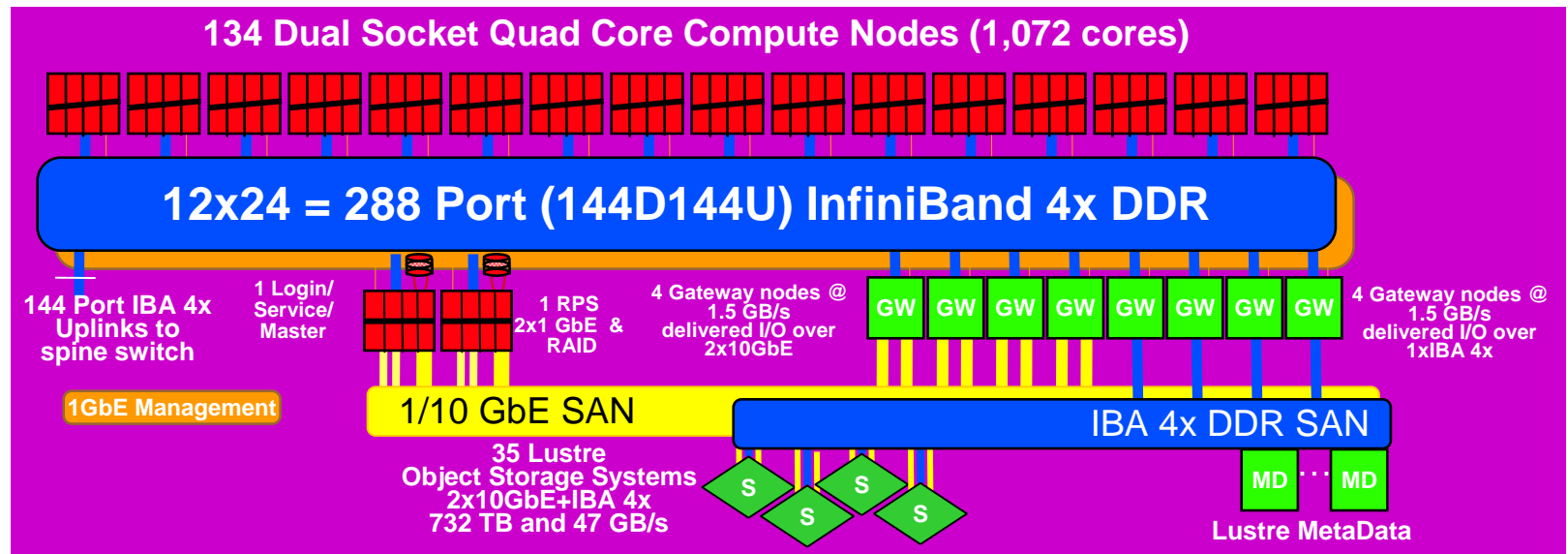


- ASC Sequoia requires a leap forward in simulation environment capability
- Deploying simulation environments for ASC Purple & BlueGene/L required developing and testing at unprecedented scale
- A new test environment is required for Sequoia and future Linux clusters
- Hyperion will be a cluster testbed for breakthrough hardware and software technology development



Hyperion Phase 1

11.5 TF/s Scalable Unit



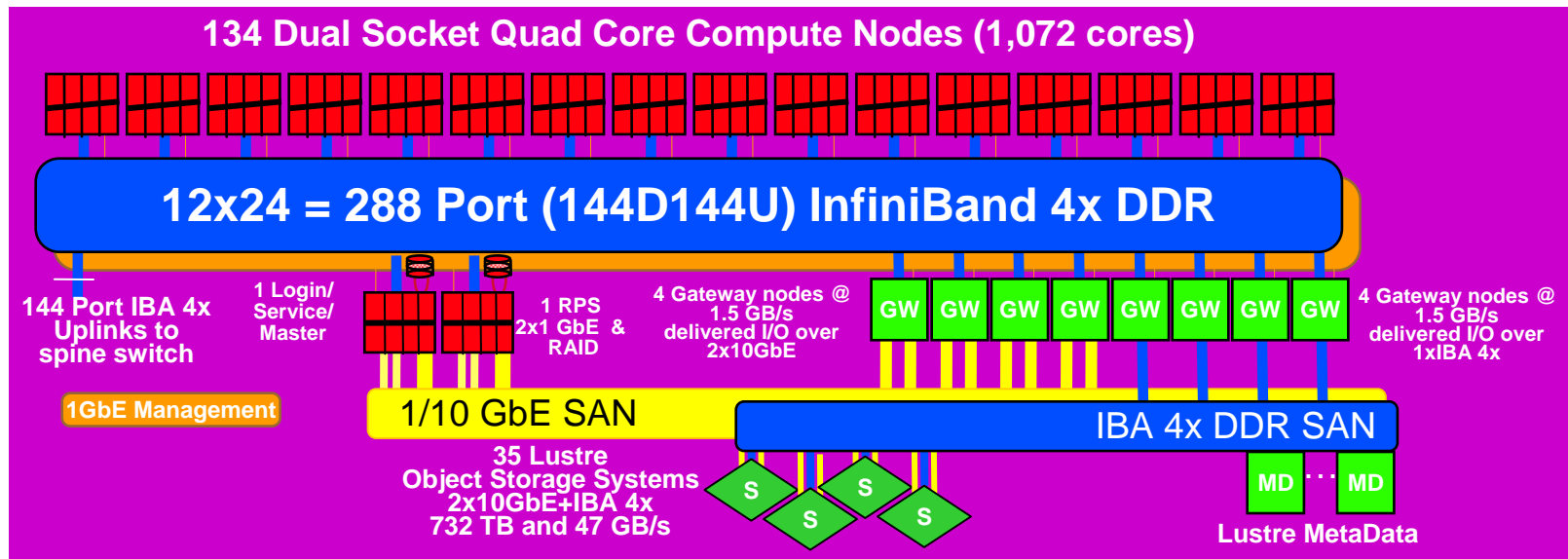
Hyperion Phase 1 Deployment is an 4 SU 46 TF/s Cluster with full IBA

- 576 Total nodes and 4,608 cores, 12.1 TB/s memory bandwidth, 4.6 TB capacity
- IBA is expandable to 1,728 IB ports single plane and can double to dual plane
- 80.0 GF/s dual socket 2.5 GHz quad-core Intel LV Harpertown nodes
 - 8 GB from 4 channels FB-DIMM 667 RAM @ 21.6 GB/s
- Nodes utilize PCI-Express generation 2 I/O which provides an upgrade path to IBA 4x QDR
- Storage Scalable Units (SSU) from DDN, LSI and Sun yielding >47GB/s and 732TB disk
- Full system will require 400 kW of power, 112 tons of cooling. Sited at LLNL



Hyperion Phase 2

13 TF/s Scalable Unit

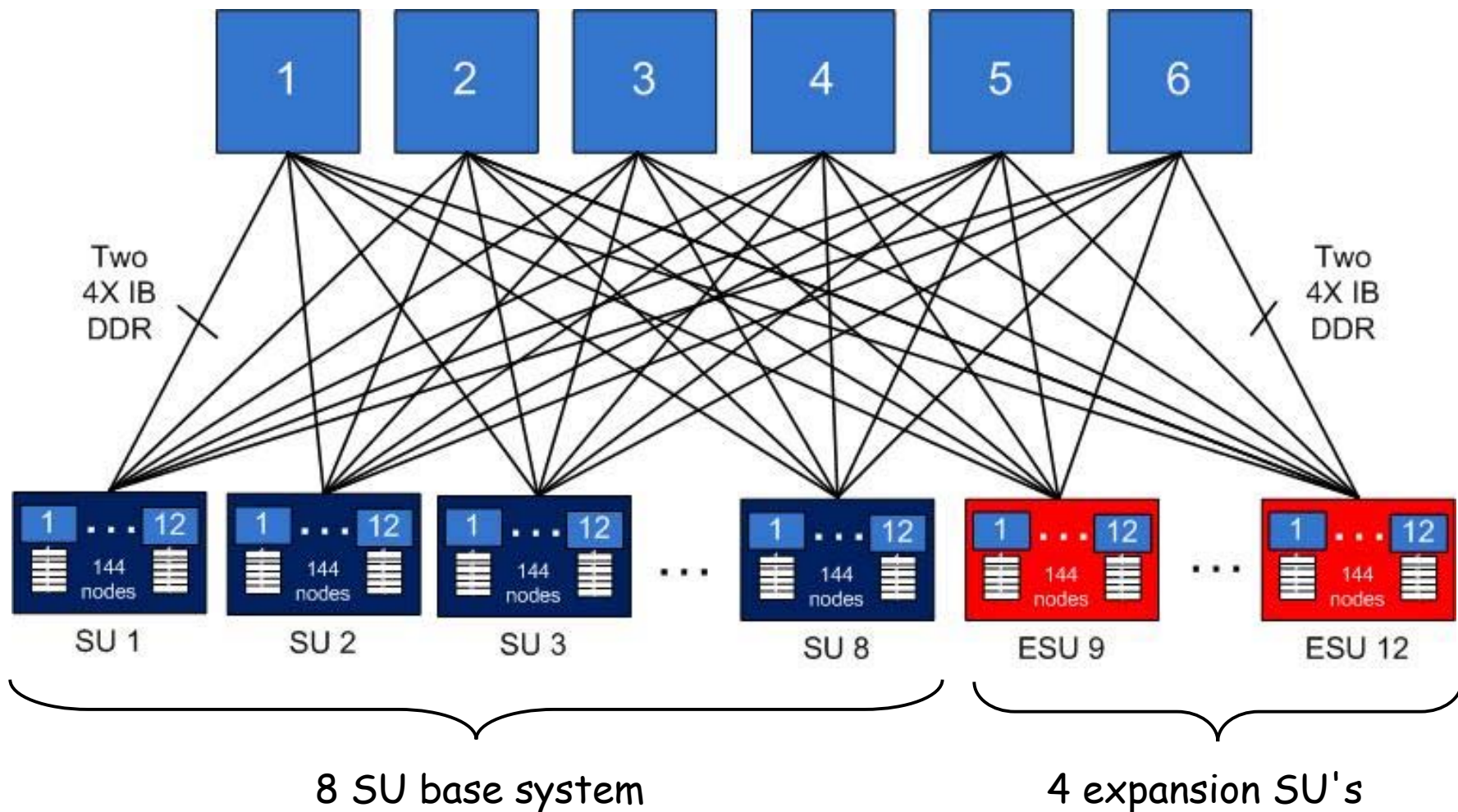


Hyperion Phase 2 Deployment is an 4 SU 52 TF/s Cluster with full IBA

- 576 Total nodes and 4,608 cores, 36.9 TB/s memory bandwidth, 4.6 TB capacity
- IBA is expandable to 1,728 IB ports single plane and can double to dual plane
- ~90.0 GF/s dual socket quad-core Intel core i7 nodes
 - ~8 GB from 6 channels DDR3 1333 DRAM @ 64 GB/s
- Nodes utilize PCI-Express generation 2 I/O which provides an upgrade path to IBA 4x QDR
- Storage Scalable Units (SSU) from DDN, LSI and Sun yielding >47GB/s and 732TB disk
- Full system will require 400 kW of power, 112 tons of cooling. Sited at LLNL

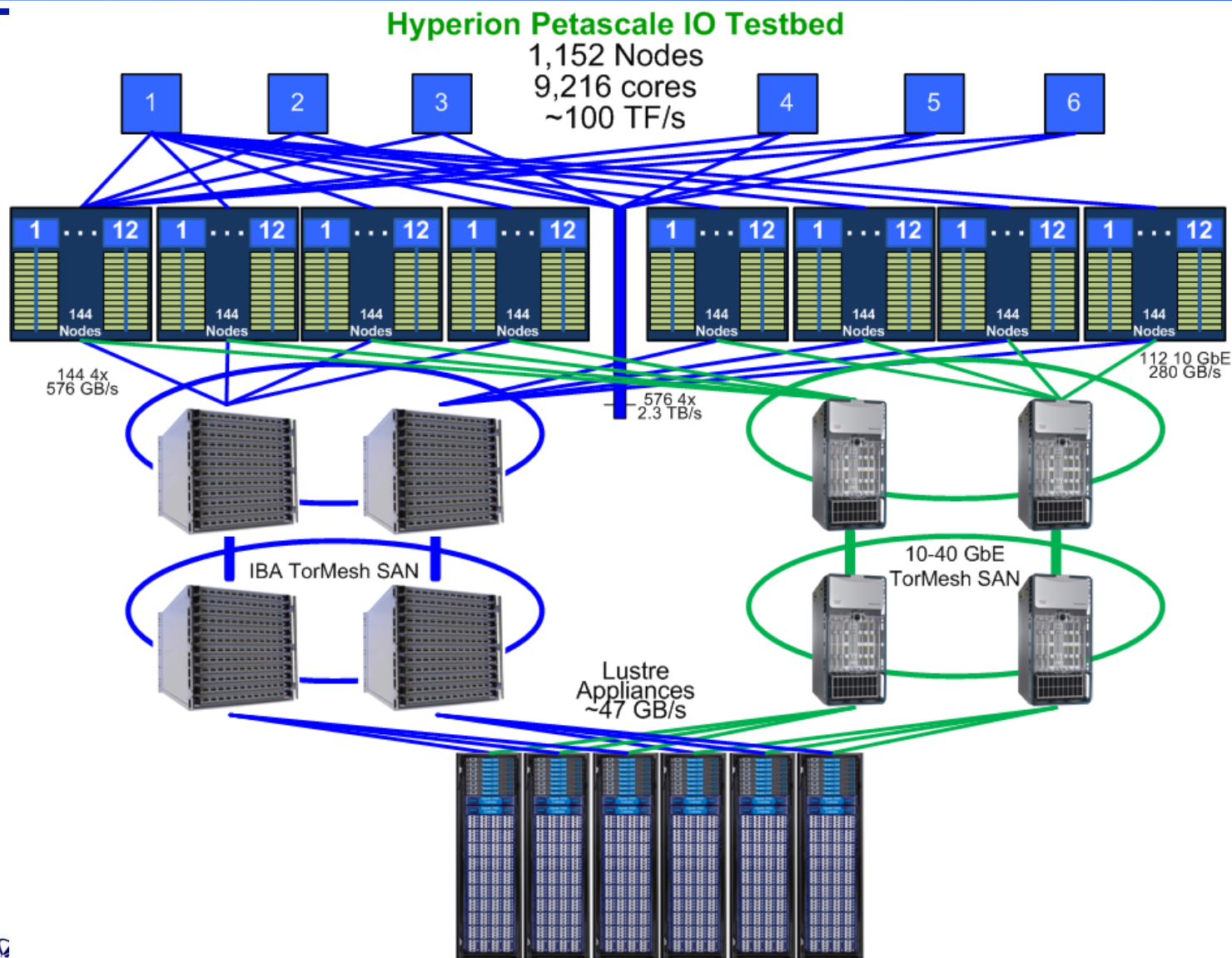


Hyperion 8 Scalable Units with Network for 50% Expansion





Hyperion petascale I/O testbed will test Sequoia I/O technologies at scale!





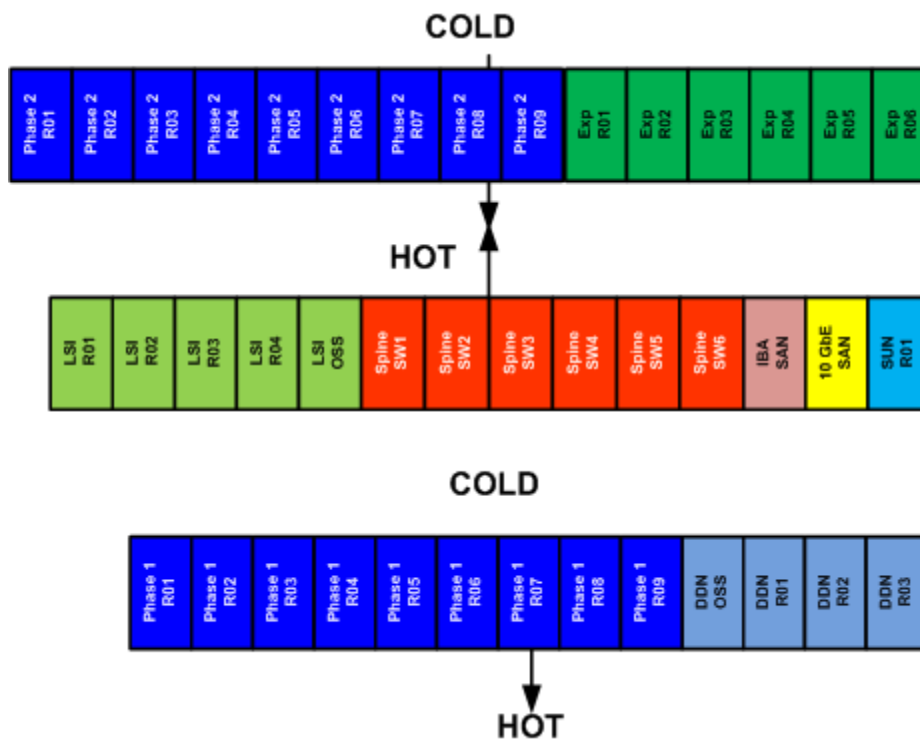
Hyperion Phase 1

46 TF/s 4SU in 9 racks

HyperionPhase1_R01	HyperionPhase1_R02	HyperionPhase1_R03	HyperionPhase1_R04	HyperionPhase1_R05	HyperionPhase1_R06	HyperionPhase1_R07	HyperionPhase1_R08	HyperionPhase1_R09
GW03 GW04 CN67 CN68	GW03 GW04 CN67 CN68	GW03 GW04 CN67 CN68			GW03 GW04 CN67 CN68	GW03 GW04 CN67 CN68	GW03 GW04 CN67 CN68	GW03 GW04 CN67 CN68
GW01 GW02 CN65 CN66	GW01 GW02 CN65 CN66	GW01 GW02 CN65 CN66			GW01 GW02 CN65 CN66	GW01 GW02 CN65 CN66	GW01 GW02 CN65 CN66	GW01 GW02 CN65 CN66
IB SW1	IB SW1	IB SW1	IB SW1		IB SW1	IB SW1	IB SW1	IB SW1
CN61 CN62 CN63 CN64	CN61 CN62 CN63 CN64	CN61 CN62 CN63 CN64	GW03 GW04 CN59 CN60		CN61 CN62 CN63 CN64	CN61 CN62 CN63 CN64	CN61 CN62 CN63 CN64	CN61 CN62 CN63 CN64
Terminal Server 1	Terminal Server 1	Terminal Server 1	Terminal Server 1		Terminal Server 1	Terminal Server 1	Terminal Server 1	Terminal Server 1
Mgmt Ethernet Switch 1	Mgmt Ethernet Switch 1	Mgmt Ethernet Switch 1	Mgmt Ethernet Switch 1		Mgmt Ethernet Switch 1	Mgmt Ethernet Switch 1	Mgmt Ethernet Switch 1	Mgmt Ethernet Switch 1
CN57 CN58 CN59 CN60	CN57 CN58 CN59 CN60	CN57 CN58 CN59 CN60	GW01 GW02 CN57 CN58		CN57 CN58 CN59 CN60	CN57 CN58 CN59 CN60	CN57 CN58 CN59 CN60	CN57 CN58 CN59 CN60
CN53 CN54 CN55 CN56	CN53 CN54 CN55 CN56	CN53 CN54 CN55 CN56	CN53 CN54 CN55 CN56	LSM01	CN53 CN54 CN55 CN56	CN53 CN54 CN55 CN56	CN53 CN54 CN55 CN56	CN53 CN54 CN55 CN56
IB SW2	IB SW2	IB SW2	IB SW2	LSM02	IB SW2	IB SW2	IB SW2	IB SW2
CN49 CN50 CN51 CN52	CN49 CN50 CN51 CN52	CN49 CN50 CN51 CN52	CN49 CN50 CN51 CN52	LSM03	CN49 CN50 CN51 CN52	CN49 CN50 CN51 CN52	CN49 CN50 CN51 CN52	CN49 CN50 CN51 CN52
CN45 CN46 CN47 CN48	CN45 CN46 CN47 CN48	CN45 CN46 CN47 CN48	CN45 CN46 CN47 CN48	LSM04	CN45 CN46 CN47 CN48	CN45 CN46 CN47 CN48	CN45 CN46 CN47 CN48	CN45 CN46 CN47 CN48
CN41 CN42 CN43 CN44	CN41 CN42 CN43 CN44	CN41 CN42 CN43 CN44	CN41 CN42 CN43 CN44	RAID01	CN41 CN42 CN43 CN44	CN41 CN42 CN43 CN44	CN41 CN42 CN43 CN44	CN41 CN42 CN43 CN44
IB SW3	IB SW3	IB SW3	IB SW3		IB SW3	IB SW3	IB SW3	IB SW3
CN37 CN38 CN39 CN40	CN37 CN38 CN39 CN40	CN37 CN38 CN39 CN40	CN37 CN38 CN39 CN40		CN37 CN38 CN39 CN40	CN37 CN38 CN39 CN40	CN37 CN38 CN39 CN40	CN37 CN38 CN39 CN40
CN33 CN34 CN35 CN36	CN33 CN34 CN35 CN36	CN33 CN34 CN35 CN36	CN33 CN34 CN35 CN36	RPS01	CN33 CN34 CN35 CN36	CN33 CN34 CN35 CN36	CN33 CN34 CN35 CN36	CN33 CN34 CN35 CN36
CN29 CN30 CN31 CN32	CN29 CN30 CN31 CN32	CN29 CN30 CN31 CN32	CN29 CN30 CN31 CN32	RAID02	CN29 CN30 CN31 CN32	CN29 CN30 CN31 CN32	CN29 CN30 CN31 CN32	CN29 CN30 CN31 CN32
IB SW4	IB SW4	IB SW4	IB SW4		IB SW4	IB SW4	IB SW4	IB SW4
CN25 CN26 CN27 CN28	CN25 CN26 CN27 CN28	CN25 CN26 CN27 CN28	CN25 CN26 CN27 CN28	RAID03	CN25 CN26 CN27 CN28	CN25 CN26 CN27 CN28	CN25 CN26 CN27 CN28	CN25 CN26 CN27 CN28
Terminal Server 2	Terminal Server 2	Terminal Server 2	Terminal Server 2	RPS02	Terminal Server 2	Terminal Server 2	Terminal Server 2	Terminal Server 2
Mgmt Ethernet Switch 2	Mgmt Ethernet Switch 2	Mgmt Ethernet Switch 2	Mgmt Ethernet Switch 2		Mgmt Ethernet Switch 2	Mgmt Ethernet Switch 2	Mgmt Ethernet Switch 2	Mgmt Ethernet Switch 2
CN21 CN22 CN23 CN24	CN21 CN22 CN23 CN24	CN21 CN22 CN23 CN24	CN21 CN22 CN23 CN24	RAID04	CN21 CN22 CN23 CN24	CN21 CN22 CN23 CN24	CN21 CN22 CN23 CN24	CN21 CN22 CN23 CN24
CN17 CN18 CN19 CN20	CN17 CN18 CN19 CN20	CN17 CN18 CN19 CN20	CN17 CN18 CN19 CN20		CN17 CN18 CN19 CN20	CN17 CN18 CN19 CN20	CN17 CN18 CN19 CN20	CN17 CN18 CN19 CN20
IB SW5	IB SW5	IB SW5	IB SW5	RPS03	IB SW5	IB SW5	IB SW5	IB SW5
CN13 CN14 CN15 CN16	CN13 CN14 CN15 CN16	CN13 CN14 CN15 CN16	CN13 CN14 CN15 CN16	RAID04	CN13 CN14 CN15 CN16	CN13 CN14 CN15 CN16	CN13 CN14 CN15 CN16	CN13 CN14 CN15 CN16
CN09 CN10 CN11 CN12	CN09 CN10 CN11 CN12	CN09 CN10 CN11 CN12	CN09 CN10 CN11 CN12		CN09 CN10 CN11 CN12	CN09 CN10 CN11 CN12	CN09 CN10 CN11 CN12	CN09 CN10 CN11 CN12
CN05 CN06 CN07 CN08	CN05 CN06 CN07 CN08	CN05 CN06 CN07 CN08	CN05 CN06 CN07 CN08	RAID04	CN05 CN06 CN07 CN08	CN05 CN06 CN07 CN08	CN05 CN06 CN07 CN08	CN05 CN06 CN07 CN08
IB SW6	IB SW6	IB SW6	IB SW6		IB SW6	IB SW6	IB SW6	IB SW6
CN01 CN02 CN03 CN04	CN01 CN02 CN03 CN04	CN01 CN02 CN03 CN04	CN01 CN02 CN03 CN04	RPS04	CN01 CN02 CN03 CN04	CN01 CN02 CN03 CN04	CN01 CN02 CN03 CN04	CN01 CN02 CN03 CN04

Twice the density of standard 1U nodes with 72 nodes per rack

Layout for Hyperion Phase 1 and 2 with IBA Switches, Storage, and Expansion



Siting at LLNL on the open collaboration network



Hyperion Timeline

