

Infrastructure

Presented by

Shane Canon

Technology Integration
National Center for Computational Sciences

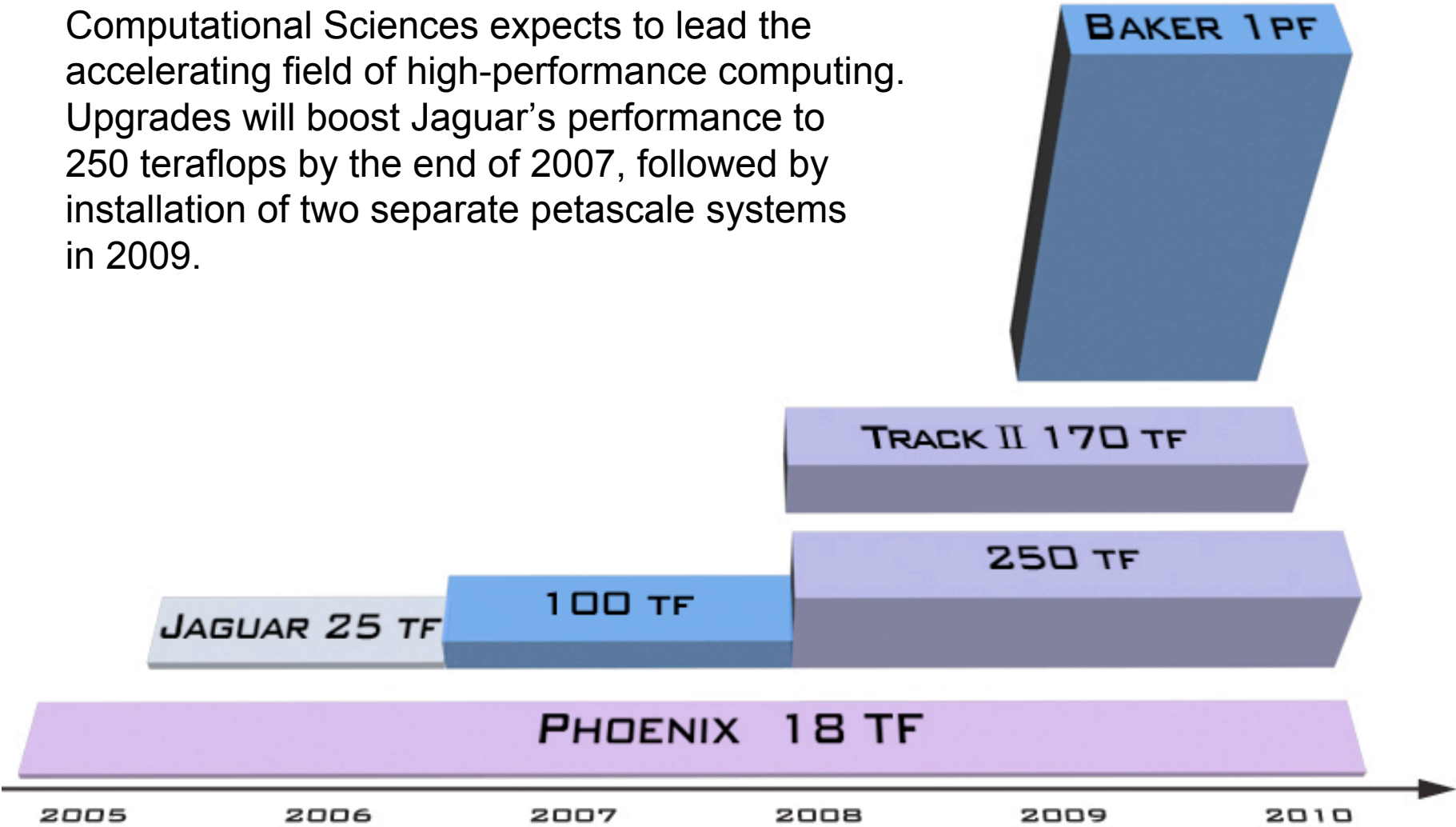


Summary

- LCF roadmap
- Infrastructure for the petascale
 - Networking
 - File systems
 - Archival storage
 - Data analytics

Hardware roadmap

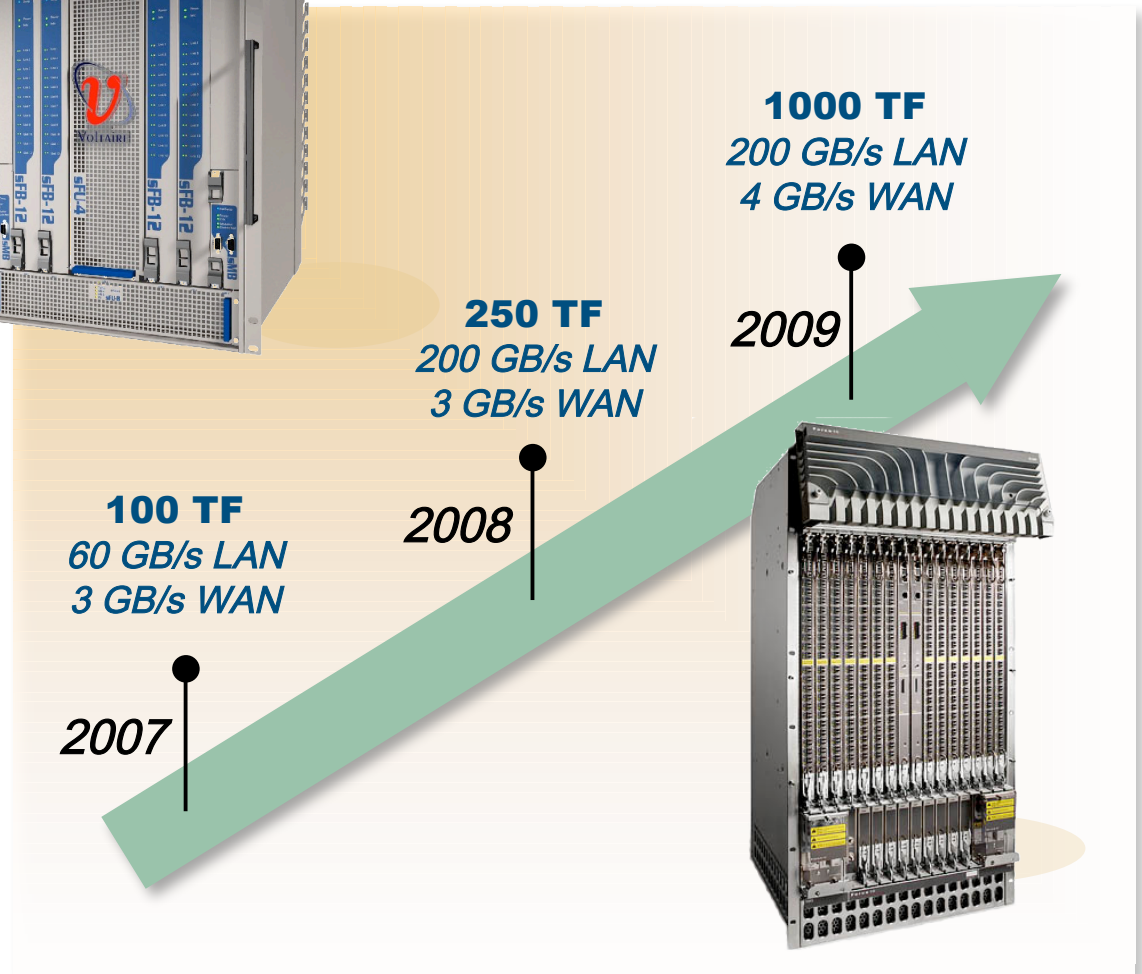
As it looks to the future, the National Center for Computational Sciences expects to lead the accelerating field of high-performance computing. Upgrades will boost Jaguar's performance to 250 teraflops by the end of 2007, followed by installation of two separate petascale systems in 2009.



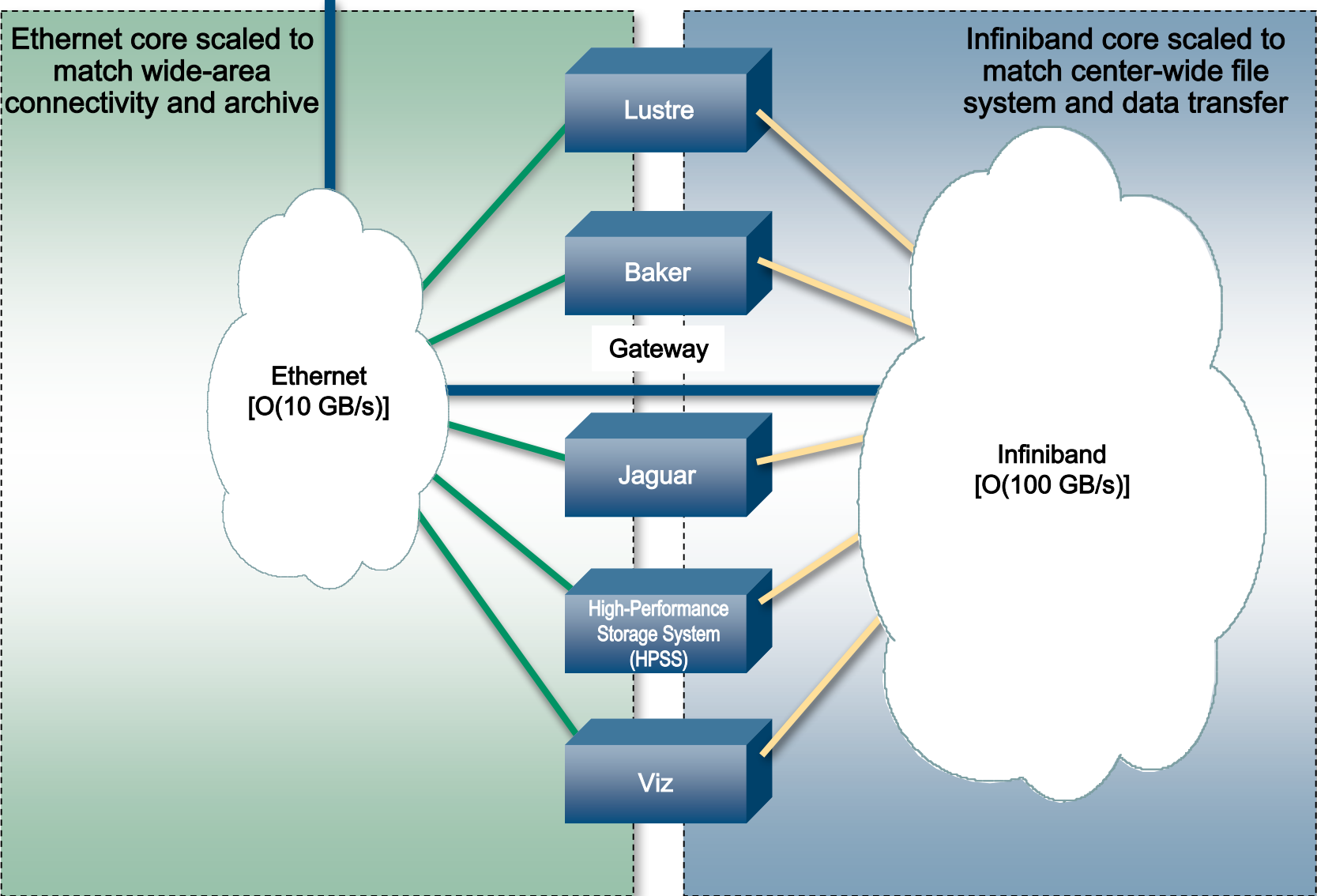
Network



- Shifting to a hybrid InfiniBand/Ethernet network.
- InfiniBand-based network helps meet the bandwidth and scaling needs for the center.
- Wide-area network will scale to meet user demand using currently deployed routers and switches.



NCCS network roadmap summary



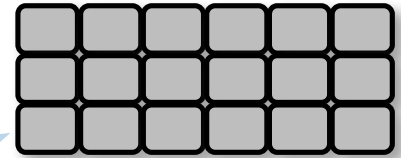
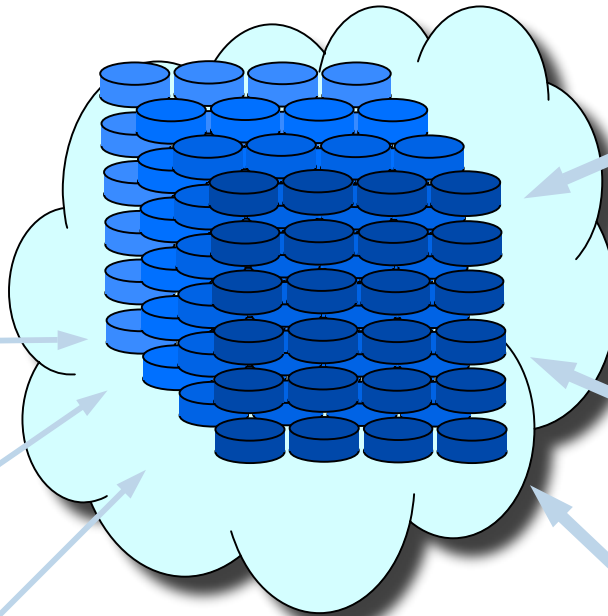
Center-wide file system (Spider)



*Phoenix
Cray X1E*



NFS servers



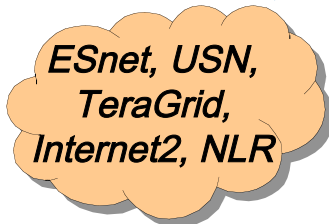
*Data analysis
and visualization*



Jaguar Cray XT4



HPSS



*ESnet, USN,
TeraGrid,
Internet2, NLR*

2007

- 1 PB
- 30 GB/s (aggregate)

2008

- 10 PB
- 200 GB/s (aggregate)

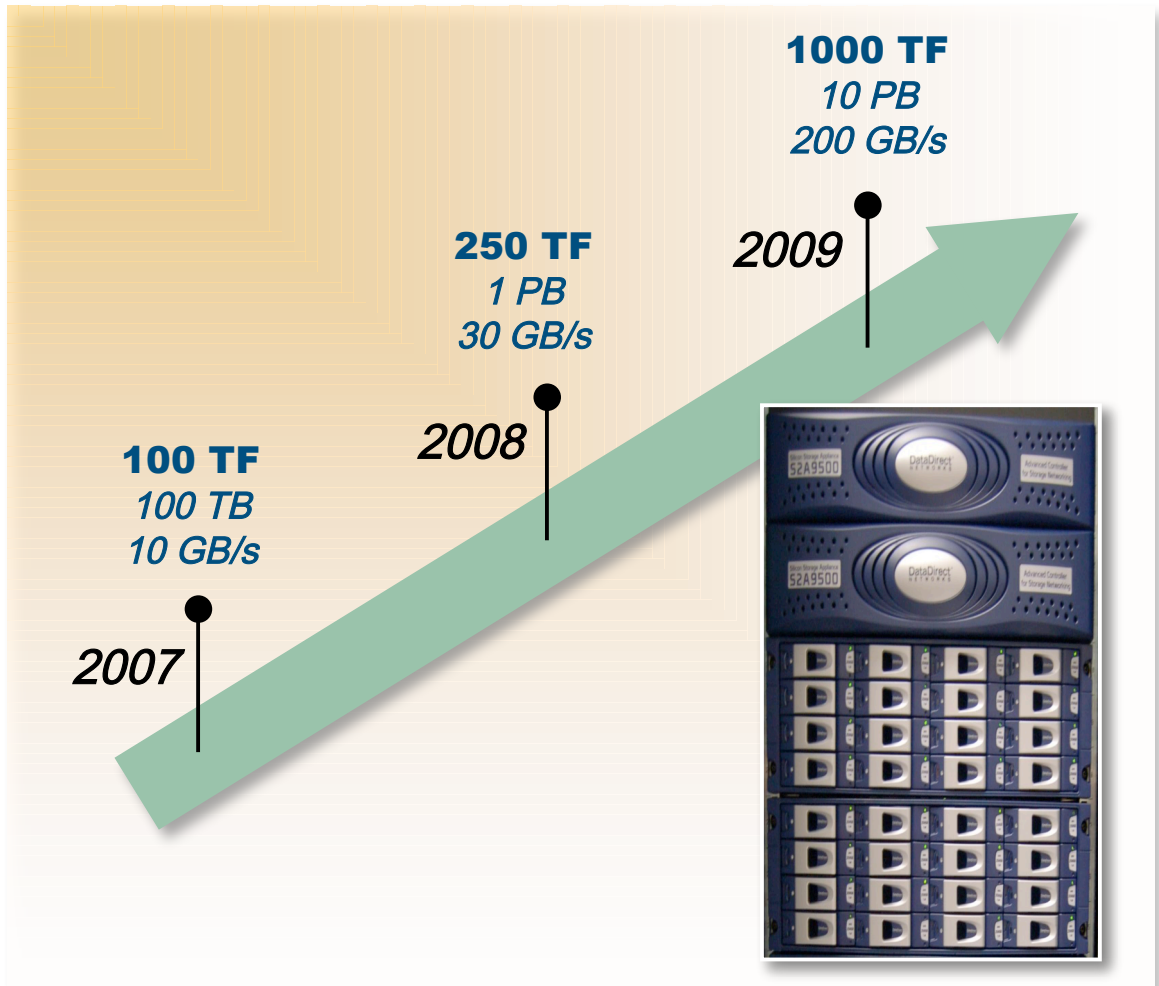


Baker

Center-wide file system



- Increase scientific productivity by providing single repository for simulation data
- Connect to all major LCF resources
- Connected to both InfiniBand and Ethernet networks
- Potentially becomes *the* file system for the 1000 TF system



Center-wide file system



- Lustre-based file system.
- Can natively utilize the InfiniBand network.
- Already running on today's XT4 at 10k+ clients.

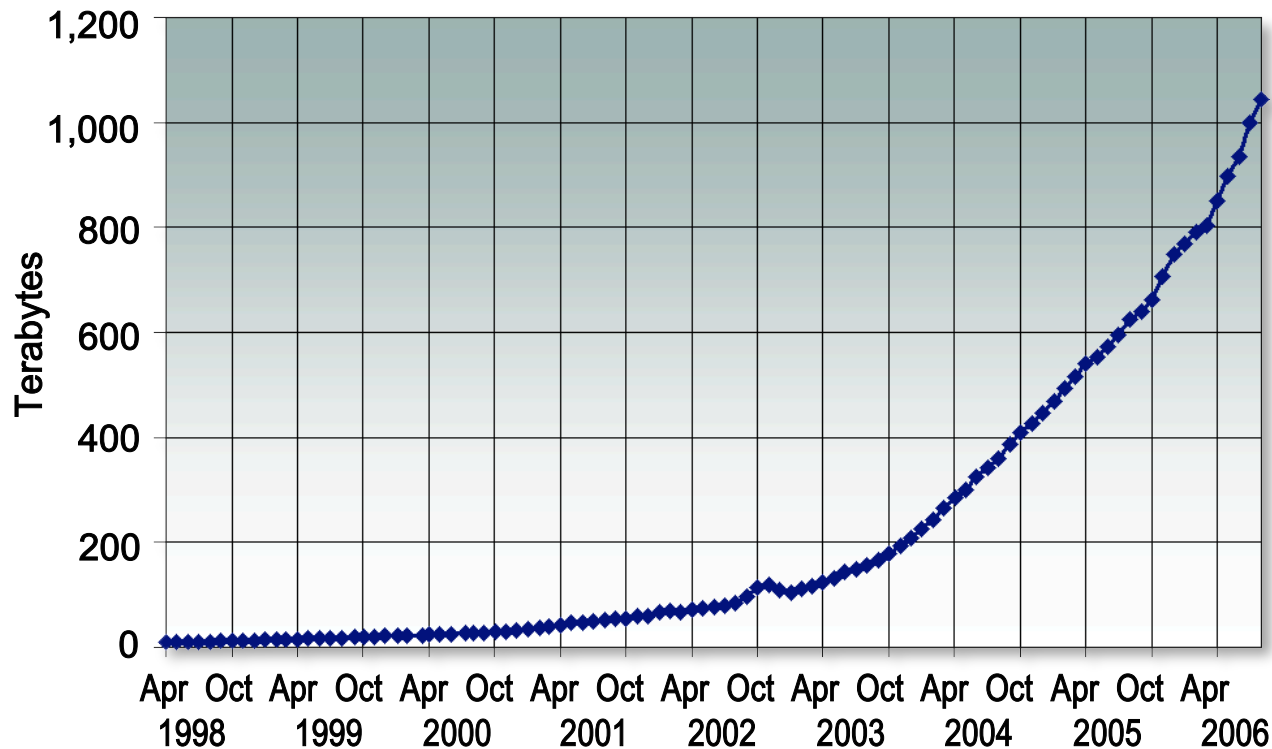
	FY 2007	FY 2008	FY 2009
OSS	20	30	160
Controllers pairs	4	8	40
Capacity (PB)	0.2	1	10
Bandwidth (GB/s)	10	30	200

End of FY, in production

- External-based system will utilize routers that are part of the transport protocol used in Lustre (route between IB and SeaStar/Gemini).
- External system already demonstrated on current XT systems.

Data storage—past usage

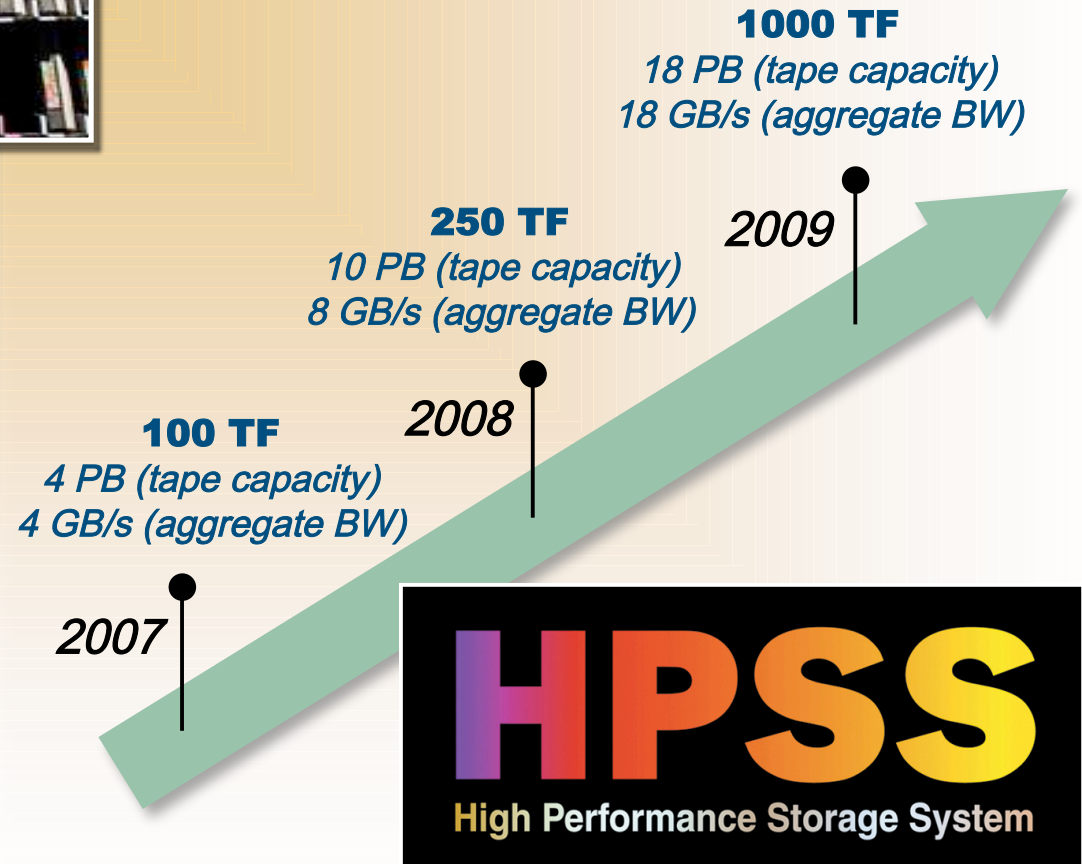
- Data growth is explosive.
- Doubling stored data every year since 1998!
- We are up to 1.6 PB stored today, and adding almost 2 TB of data per day.



Archival storage



- HPSS software has already demonstrated ability to scale to many PB.
- Add 2 silos/year.
- Tape capacity and bandwidth, disk capacity and bandwidth are all scaled to maintain a balanced system.
- Utilize new methods to improve data transfer speeds between parallel file systems and archival system.



Archival storage

	FY 2007	FY 2008	FY 2009
Silos	2*	3	4
Tape drives (T10K)	16	24	32
Tape capacity (PB)	4	10	18
Disk cache (TB)	200	1000	1100
Tape BW (aggregate GB/s)	1.9	3.8	7.6
Disk BW (aggregate GB/s)	4	10	19

* Doesn't include older silos

Note: Middle of FY, in production

Data analytics

Existing resources

- Visualization cluster (64 nodes/quadrics)
- End-to-end cluster (80 nodes/IB)

Recently deployed

- Deploy 32 nodes with 4X-DDR
- Connected to center-wide file system

Data analytics—strategies

Jaguar (250 TF) (FY 2008)

- Utilize portion of system for data analysis (50 TF/20 TB)

Baker (FY 2008/2009)

- Utilize Jaguar as analysis resource (250 TF/50 TB)
- Provision fraction of Baker for analysis

Milestones—FY 2008

First half FY 2008

- Perform “bake-off” of storage for center-wide file system
- Expand IB network
- Demonstrate 1.5 GB/s sustained with single OSS node (dual socket QC)
- Deploy High Performance Storage System upgrades

Second half FY 2008

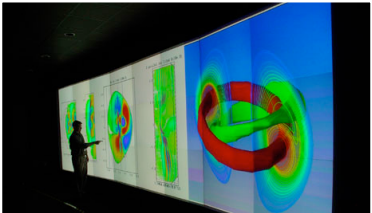
- Select storage system and procure next phase of center-wide storage (200 GB/s)
- Deploy next phase center-wide file system

Operations infrastructure systems

Now and future estimates



Archival Storage	FY 2007	FY 2008	FY 2009
Capacity (PB)	4	10	20
Bandwidth (GB/s)	4	10	19



Viz/End-to-End	FY 2007	FY 2008	FY 2009
IO B/W	10	15	60
Memory (TB)	0.5	2+20	69



Central Storage	FY 2007	FY 2008	FY 2009
Capacity (PB)	0.10	1.0	10.0
Bandwidth (GB/s)	10	30	200



Networking	FY 2007	FY 2008	FY 2009
External B/W (GB/s)	3	3	4
LAN B/W (GB/s)	60	200	200

Contact

Shane Canon

Technology Integration
National Center for Computational Sciences
(864) 574-2028
canonrs@ornl.gov

