Lawrence Livermore National Laboratory

Livermore Computing Data Archives Supporting HPC Simulation

Stephen York
york12@llnl.gov

LLNL-PRES-484392

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

Topics

- LLNL HPC Archives
- Issues of Scale
- Integrity of Data
- Availability
- Efficiencies

The Data Archives at LLNL

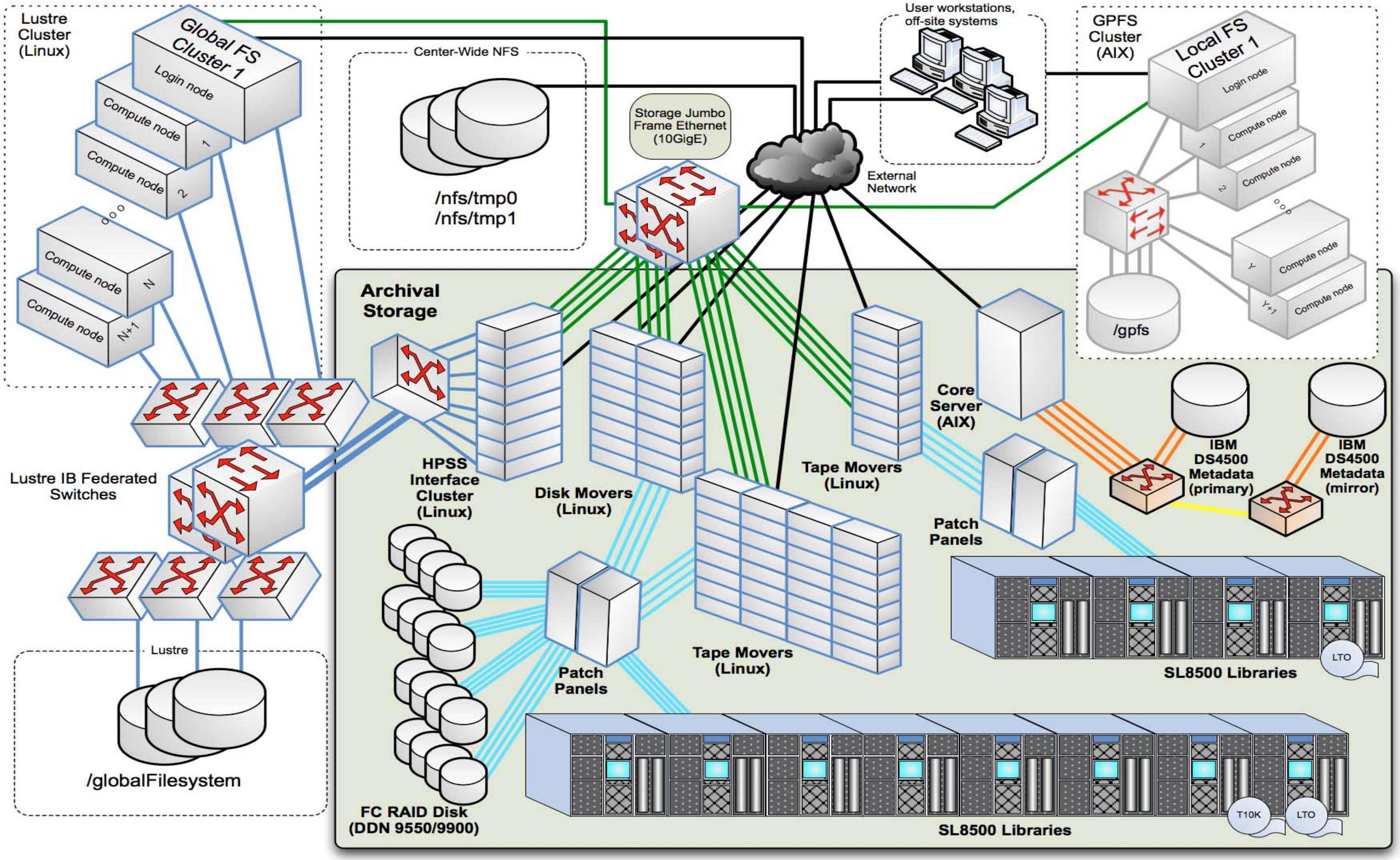
- Supporting Simulation
- Very large HPC clusters (n PF total)
- Very large Parallel File Systems (n PB total)
- Multiple production archives for different security environments
- HPSS – High Performance Storage System
  - 17+ year collaborative effort
    - IBM, LANL, LBNL, LLNL, SNL + others
  - IBM Service Offering

Current LLNL HPSS Stats of Interest

HPSS Production Systems	
Files:	**374M**
Bytes:	**35 PB**
Disk Cache:	**1.5 PB**
Throughput:	**5.0 GB/sec**
***Availability:**	**97.52% / 99.77%**
***Mounts/month :**	**27,168**
****Byte writes/day:**	**23.2TB**
****File writes/day:**	**343,375**
****File deletes/day:**	**25,677**

**12mo avg, robotic mounts exclude tape drive cleaning*
***3mo avg*

LLNL LC Archive Layout

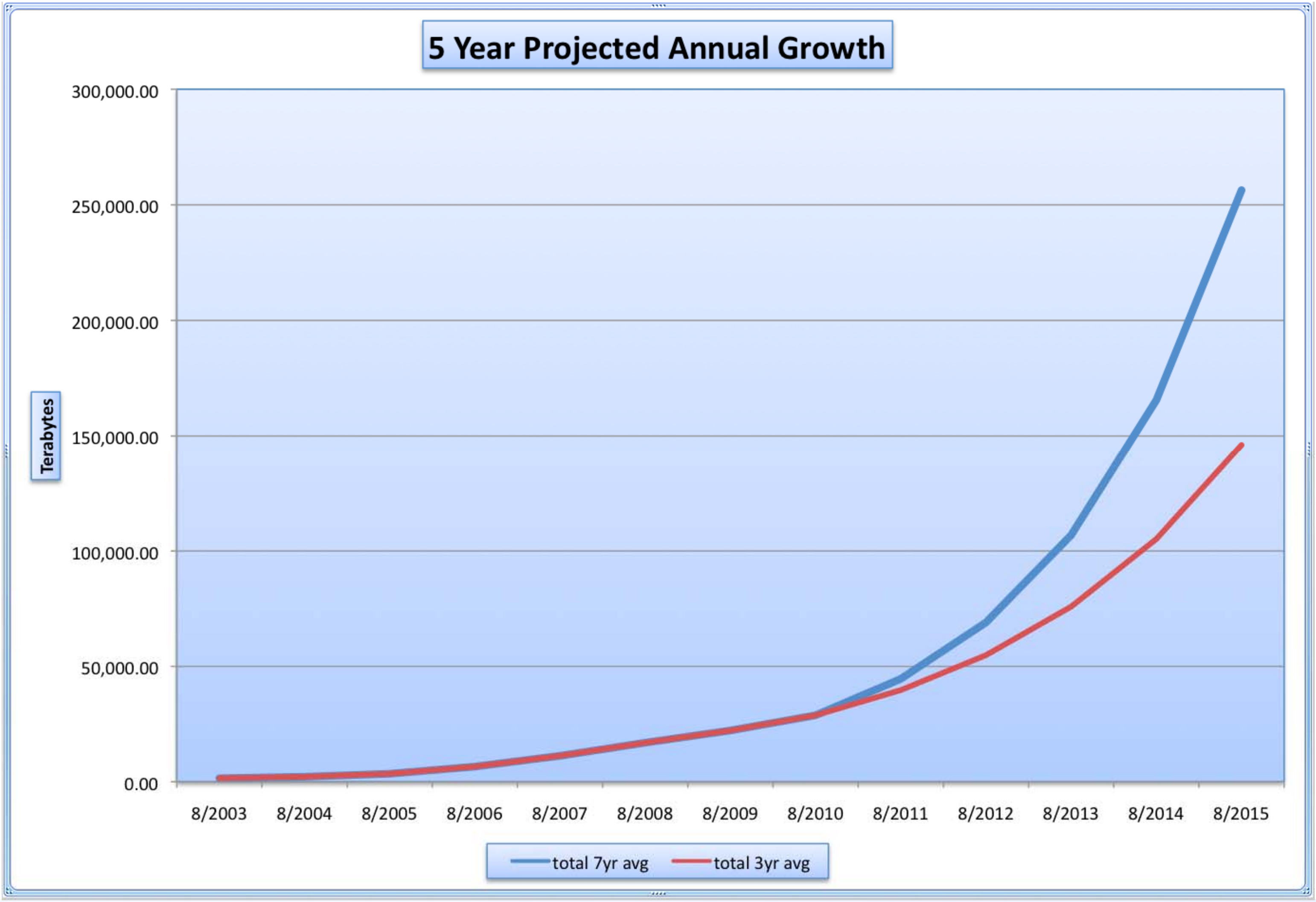


Four Raised Floor Environments

Issues of Scale

- Growth Expectations
- Managing Growth
- Archives outlast Architectures
- Rising number of components negatively affects MTBF

Issues of Scale: Growth Expectations



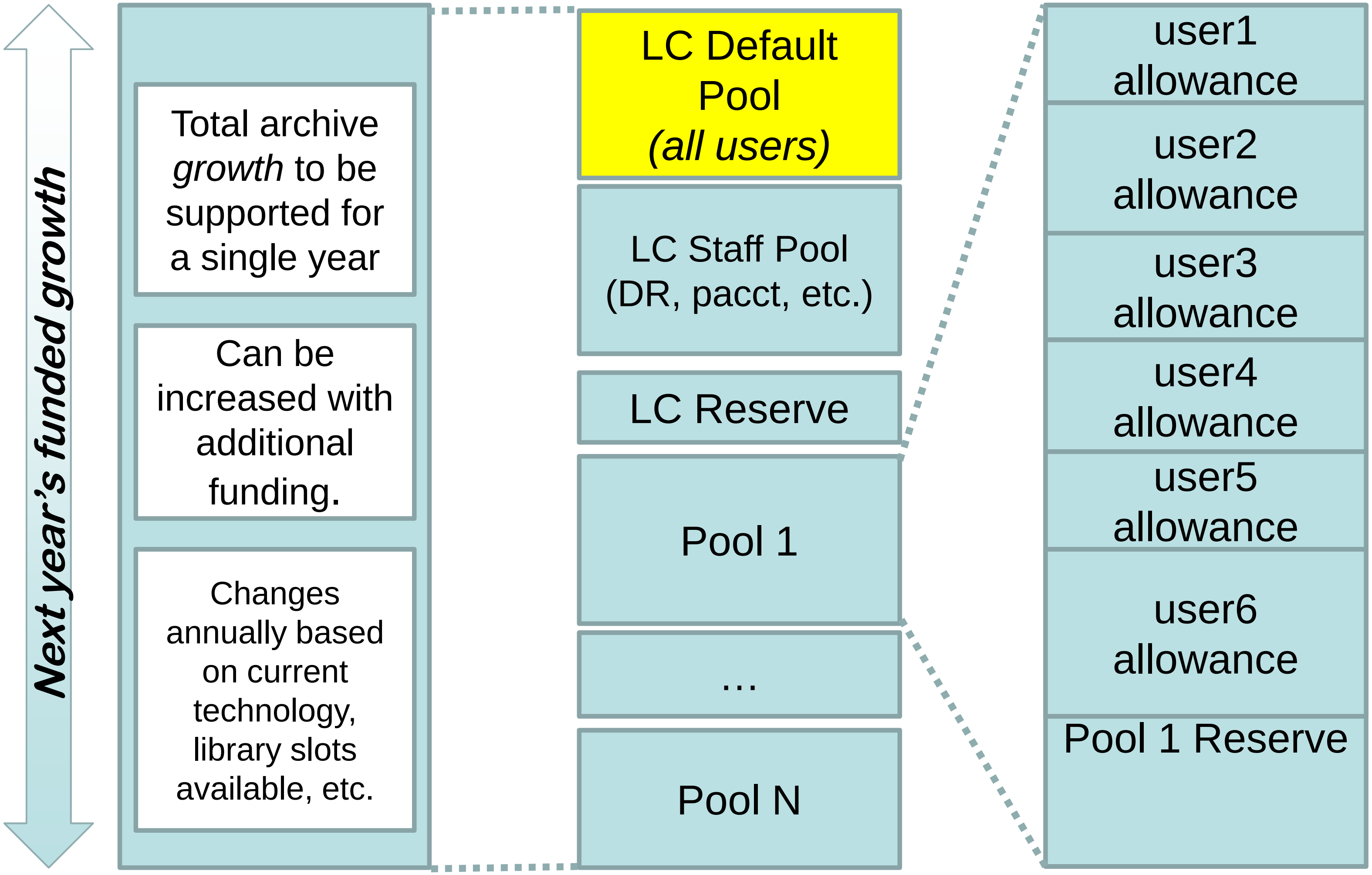
Issues of Scale: Managing Growth

Aquota:

- Tool for viewing and administering yearly archive usage allowances.
- Most users live within their default yearly allowance.
- Center allocates “pools” of space to projects.
- Individuals exceeding their base allowance need to be given space from project pools.
- “Soft” enforcement – notification only.

Know your archive budget and live within it

Soft Annual Quota System Model



Next year's funded growth

Total archive *growth* to be supported for a single year

Can be increased with additional funding.

Changes annually based on current technology, library slots available, etc.

LC Default Pool *(all users)*

LC Staff Pool (DR, pacct, etc.)

LC Reserve

Pool 1

…

Pool N

user1 allowance

user2 allowance

user3 allowance

user4 allowance

user5 allowance

user6 allowance

Pool 1 Reserve

Issues of Scale: Archives outlast Architectures

- Networks, Fibre Channel Infrastructure, Tape Drives, Tape Media, Disk Arrays, Cluster Nodes, Operating Systems, Vendors, Libraries, and Firmware/Software versions are constantly turning over.
- A new archive cannot be dropped into place. Upgrading is an evolutionary process, and occurs in a piecemeal fashion.
- Each new system component must be rigorously tested and carefully integrated causing minimal disruption to the production environment.
- Obsolete system components must be gracefully retired.
- A Pre-production system environment helps greatly in meeting the above goals.

Issues of Scale: Facility and Budget Pressures

- Exponential Data Growth with reduced headcount and flat or declining budgets
- The file systems (e.g. Lustre) and HPC clusters and memory footprints are growing and the archives are getting squeezed.
  - Leveraging center HPC platform purchases in the Archives
- Support contracts
  - For large vendor installations, permanent onsite CSE/CE support can often be negotiated for relatively modest rates, and reduce labor burden

Data Integrity

- DIVT – Data Integrity Verification Tool
- Dual Copy, Dual Technology
- Offline Tape Drive Testing
- Redundancy where required
  - Core server host components
  - Metadata (crown jewels of the archive)
  - Infrastructure switching
  - Commodity data movers and devices, so we can afford enough disk/tape across multiple hosts and tolerate losses for a short period of time

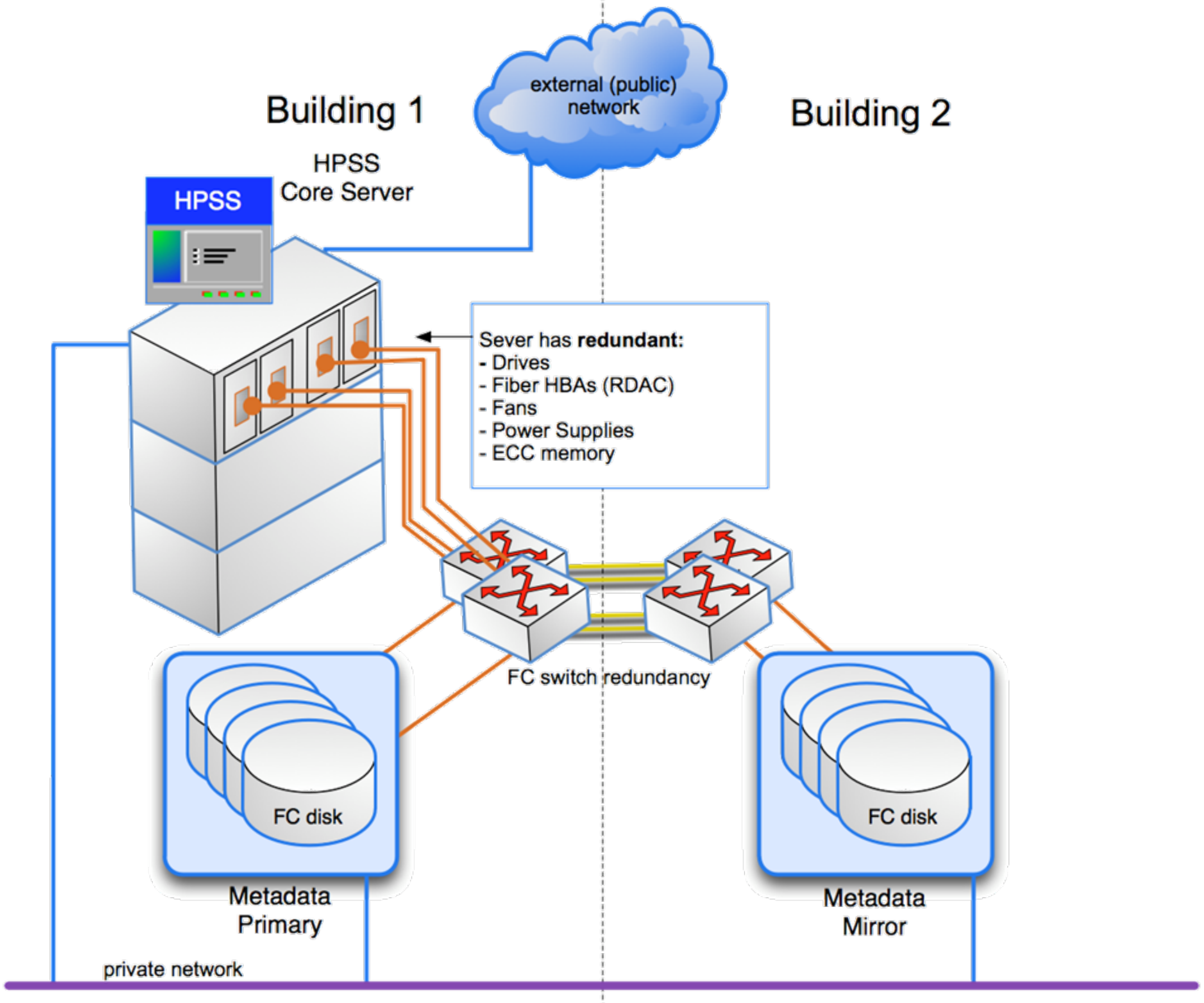
Data Integrity

- Data Integrity Verification Tool (DIVT)
  - Constantly pushing and pulling data through our entire HPC environment and checking it for validity.
  - Not all corruptions will be found, DIVT cycles through storage classes (and thru hardware).
  - We catch problems over time.
  - DIVT caught a corrupting hardware component at LLNL last year before customer data was permanently corrupted.

Availability: Identify Points of Failure

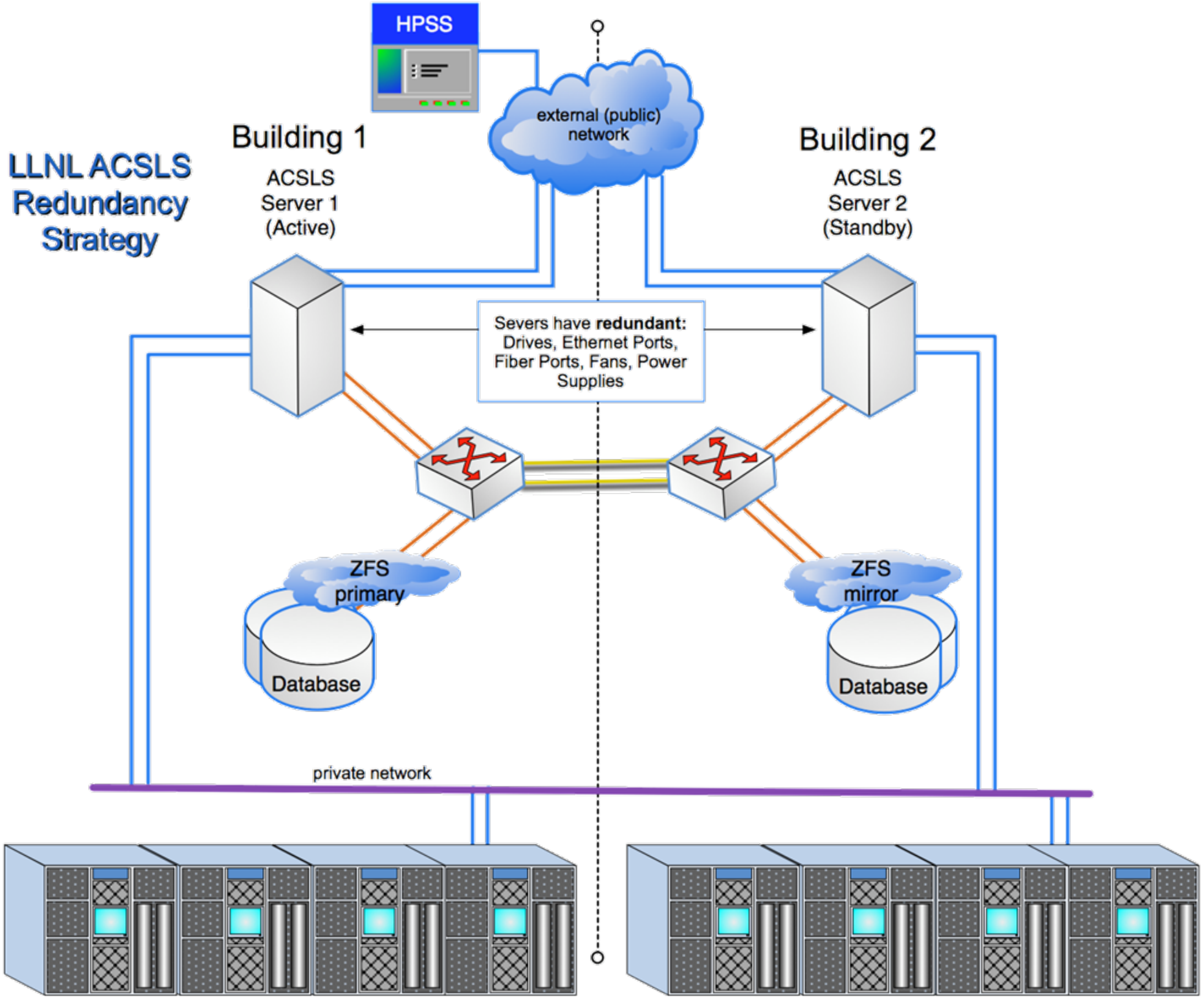
- Design for fault tolerance
  - Can you lose a disk mover host, two?
  - Can you lose a subset of tape drives?
  - What happens when a robot fails?
  - HPSS Core Server hosts mirroring and redundancy
  - Robotic software control redundancy

Availability: HPSS Core Server Architectures



Sever has **redundant:**
- Drives
- Fiber HBAs (RDAC)
- Fans
- Power Supplies
- ECC memory

Availability: Robotic Control Software



Severs have **redundant:** Drives, Ethernet Ports, Fiber Ports, Fans, Power Supplies

Availability: Archive Software Upgrades

- Stage new software on your pre-production environment.
- Pre-production environment should closely model production.
- HPSS allows devices to be used by either production or pre-production systems. Specific devices can be allocated to either side to enable testing.
- Unit test, integration test, and system test across your ENTIRE environment!
- Don’t over-patch your production system. Patches , new firmware, etc. should be justified.

Efficiencies

- Fiber Channel Doesn’t Require a Switch
- Leveraging Center HPC Procurements
  - Disk RAID Controllers/JBODs bought with file system hardware
  - Archive host servers bought with HPC clusters
  - Cost savings are realized both up-front and in ongoing self-supported maintenance activities
- Documented operational methods & strong Ops staff
- Single authoritative sources of information

Efficiencies: Ongoing improvements

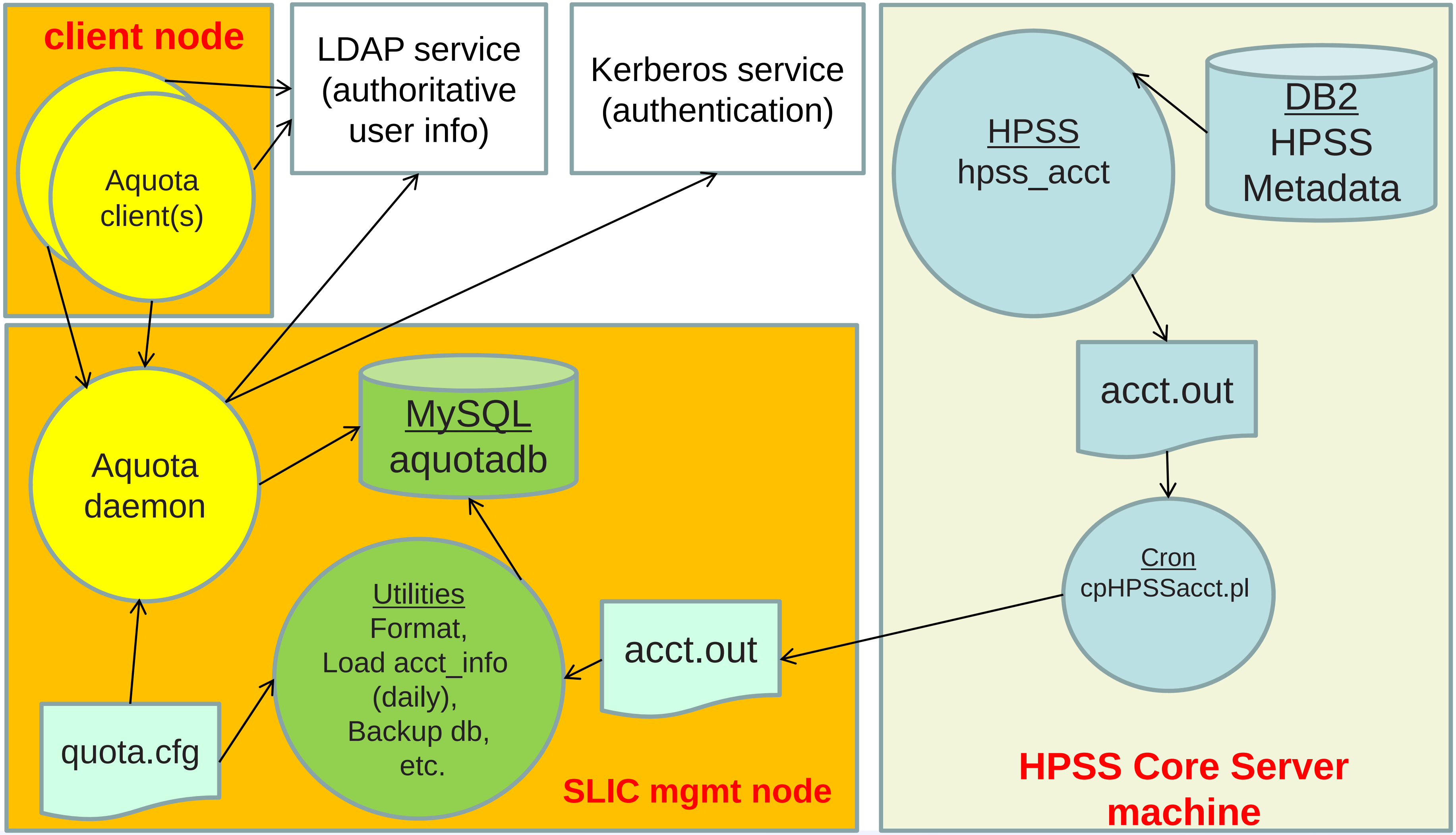
- How much of the software and hardware stack does your group support? Should it?
  - Strive to make your archive hardware easy to monitor from a single GUI.
  - Provide clear messaging on critical failures vs. degraded operation.
  - Train operations staff to replace RAID disks, mark failed drives (disk or tape) as unavailable to the system, etc.
  - Buy vendor solutions where it is practical, develop them in-house if it isn’t …

Efficiencies: A Case Study

- LTO – SCSI Log Sense data used to read tape errors and performance counters.
- T10K – Oracle Management Information Record (MIR) used to read tape errors and performance counters.
  - MIR recorded marked increase in throughput after conversion from 4x1gbE to 1x10gbE networks.
- Tape statistics are extracted before tape unload and does NOT impact HPSS performance.
- Statistics will be stored in a database.

Computation
Directorate

Aquota Architecture


client node
Aquota client(s)
LDAP service (authoritative user info)
Kerberos service (authentication)
HPSS hpss_acct
DB2 HPSS Metadata
acct.out
Cron cpHPSSacct.pl
Aquota daemon
MySQL aquotadb
Utilities Format, Load acct_info (daily), Backup db, etc.
acct.out
quota.cfg
SLIC mgmt node
HPSS Core Server machine