Archive Storage Infrastructure

At the Library of Congress

Scott Rife
Digital Storage Conference
September 30, 2011
srif at loc dot gov

LIBRARY OF CONGRESS

Packard Campus for Audio Visual Conservation
http://www.loc.gov/avconservation/packard/

Formerly NAVCC

The Packard Campus

Mission

- The National Audiovisual Conservation Center develops, preserves and provides broad access to a comprehensive and valued collection of the world's audiovisual heritage for the benefit of Congress and the nation's citizens.

Goals

- **Collect, Preserve, Provide Access to Knowledge**
- The National Audiovisual Conservation Center (NAVCC) of the Library of Congress will be the first centralized facility in America especially planned and designed for the acquisition, cataloging, storage and preservation of the nation's collection of moving images and recorded sounds. This collaborative initiative is the result of a unique partnership between the Packard Humanities Institute, the United States Congress, the Library of Congress and the Architect of the Capitol.
- The NAVCC will consolidate collections now stored in four states and the District of Columbia. Once complete, the facility will boast more than 1 million film and video items and 3 million sound recordings, providing endless opportunities to peruse the sights and sounds of American creativity.

The Packard Campus – Many Formats

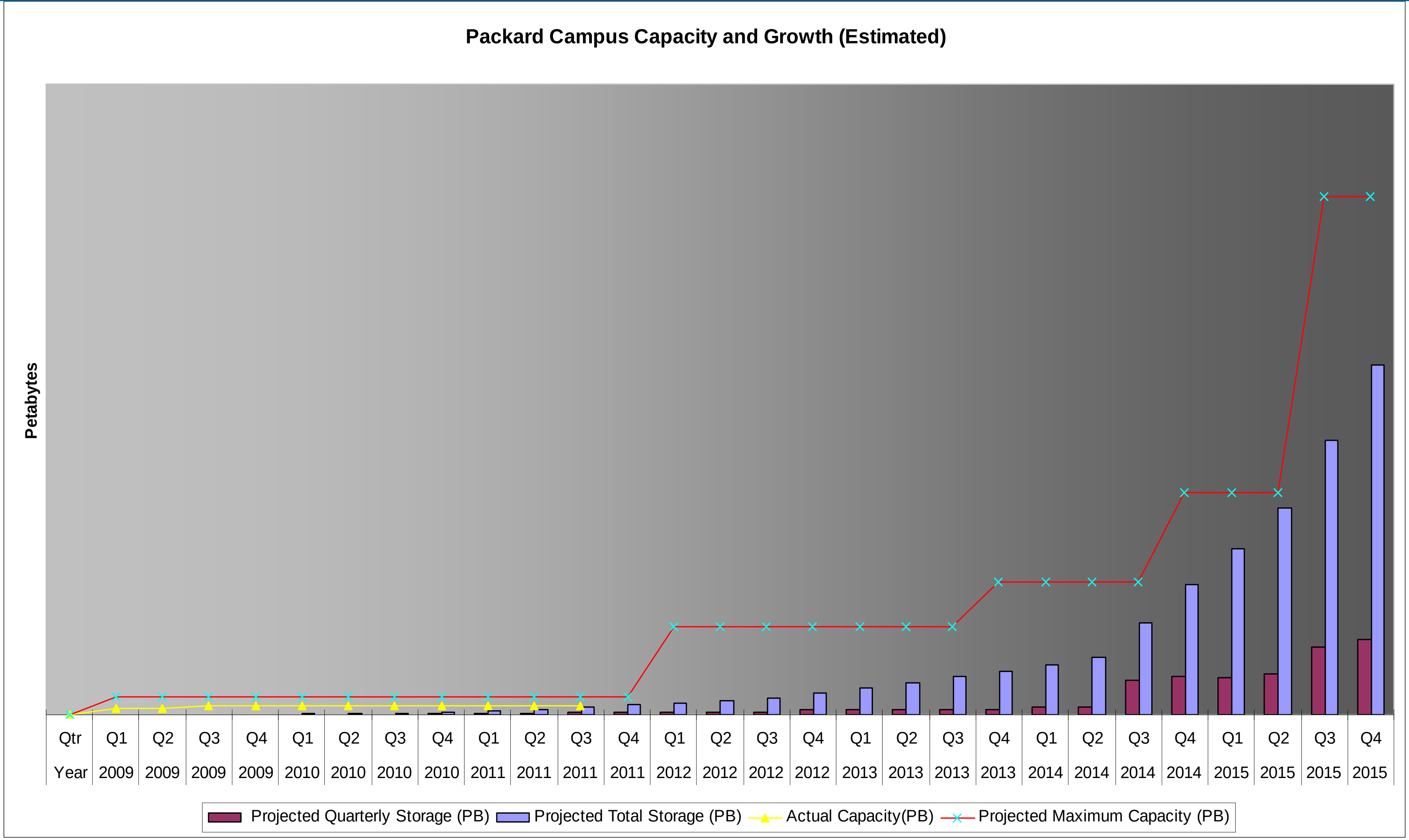
The Packard Campus – Past, Present and Future

- 38 digitization stations (PODs): 31 Solo, 6 Pyramix, 1 Quadriga
  - Daily each POD generates: 2GB-150GB for audio and 50GB-1,200GB for video
  - Additional PODs coming in the future include 2K and 4K scan for film, digital submission for Copyright and other (Live capture-264 DVRs, PBS, NBC Universal, Vanderbilt TV News, SCOLA, etc)
- Growth since production
  - February 2009: 10 TB / month
  - February 2010: 45 TB / month
  - February 2011: 91 TB / month
  - Peak in May 2011: 134 TB / month
- Current: 1.7 PB and 180,000 files in 2 locations

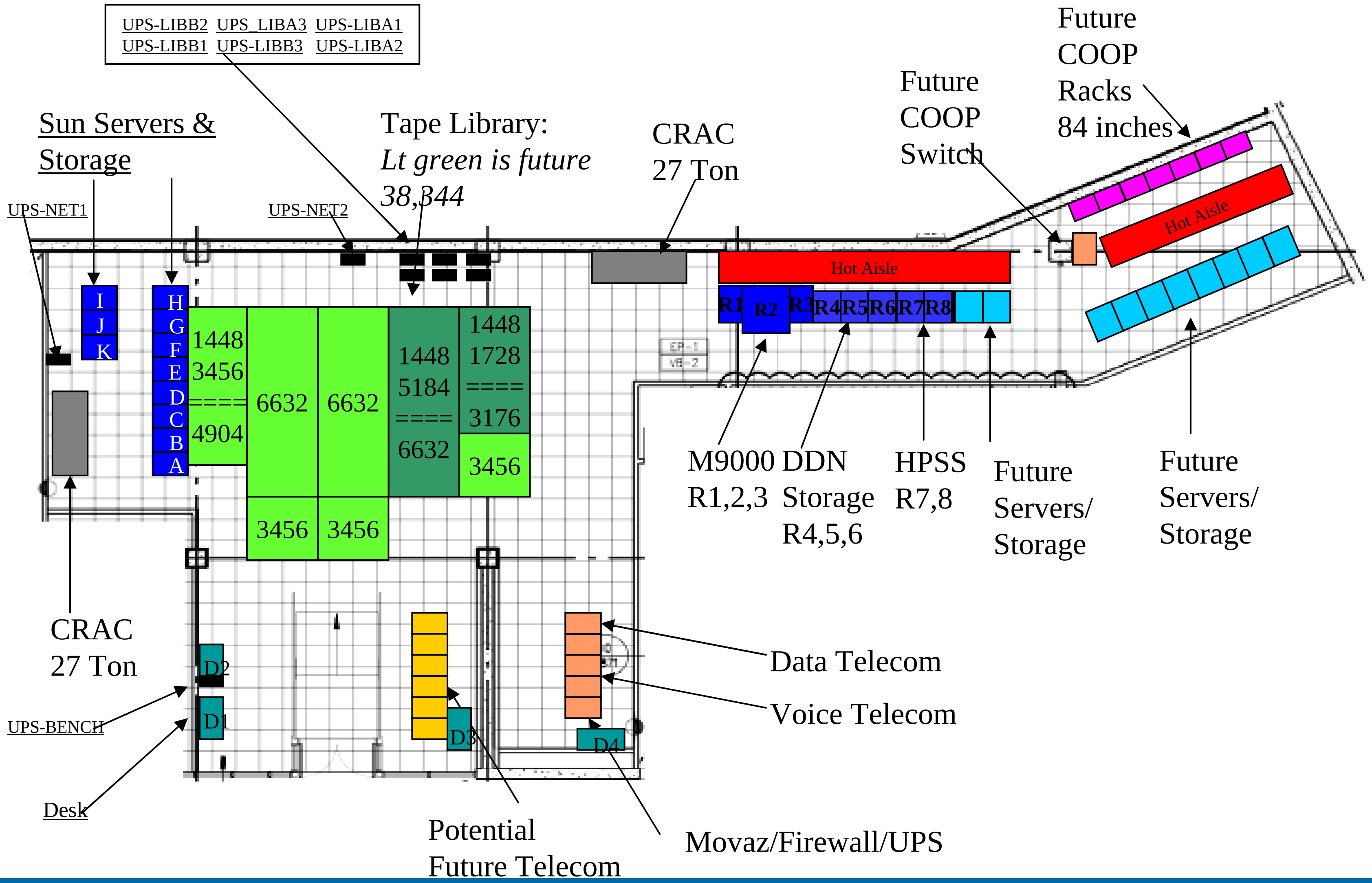
The Challenge

- **Projected: 300 TB / day or 7.5 PB / month – at least 10 years off**
- **Counting on doubling of tape density and computing power to keep us in our current 3000 Sq feet of computer room.**

The Packard Campus – Storage Graph



The Packard Campus – Physical Space



Sun Servers & Storage

Tape Library: *Lt green is future* *38,344*

CRAC 27 Ton

Future COOP Switch

Future COOP Racks 84 inches

M9000 R1,2,3

DDN Storage R4,5,6

HPSS R7,8

Future Servers/ Storage

Future Servers/ Storage

CRAC 27 Ton

Data Telecom

Voice Telecom

Potential Future Telecom

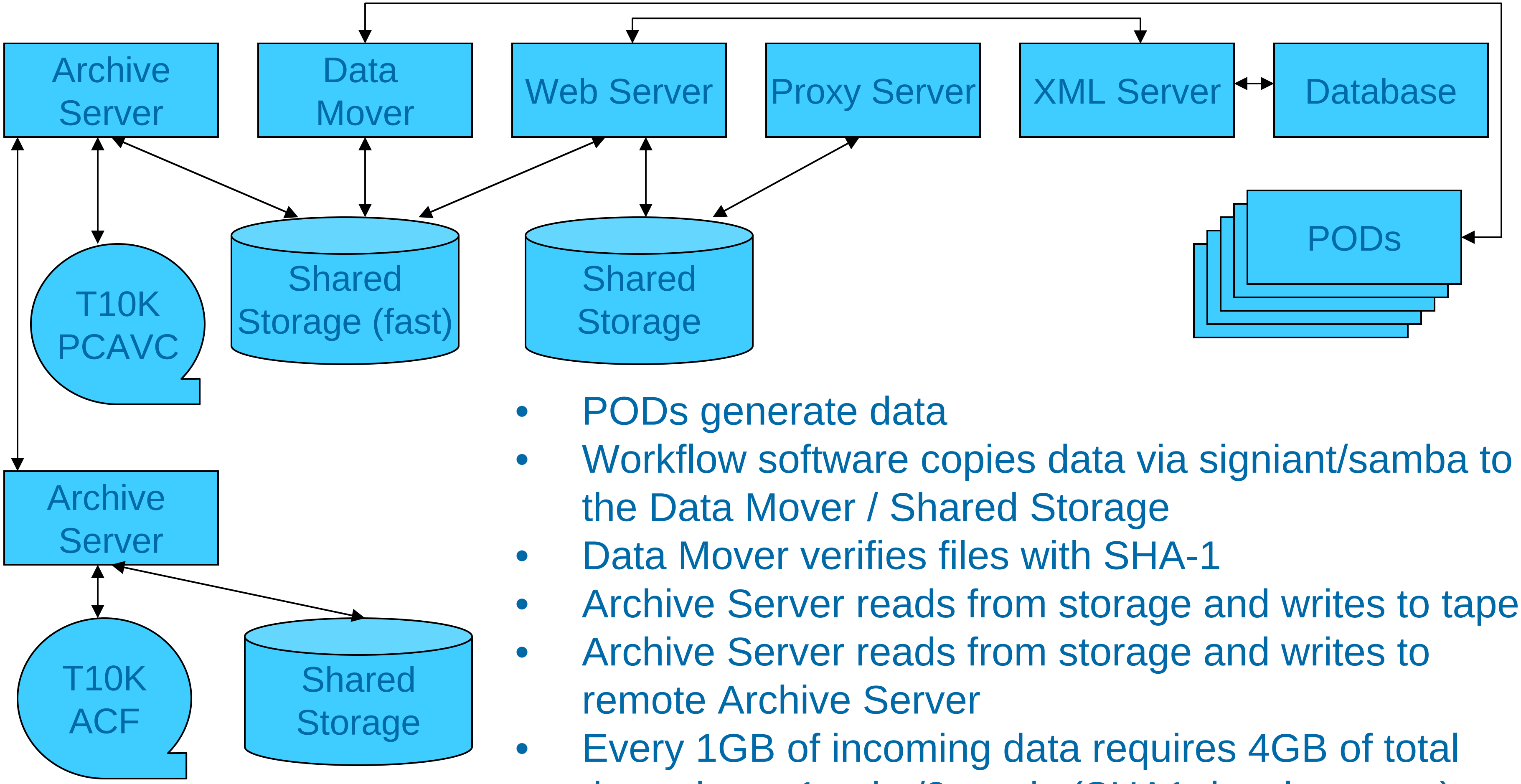
Movaz/Firewall/UPS

UPS-BENCH

Desk

Functional Architecture – Data Movement

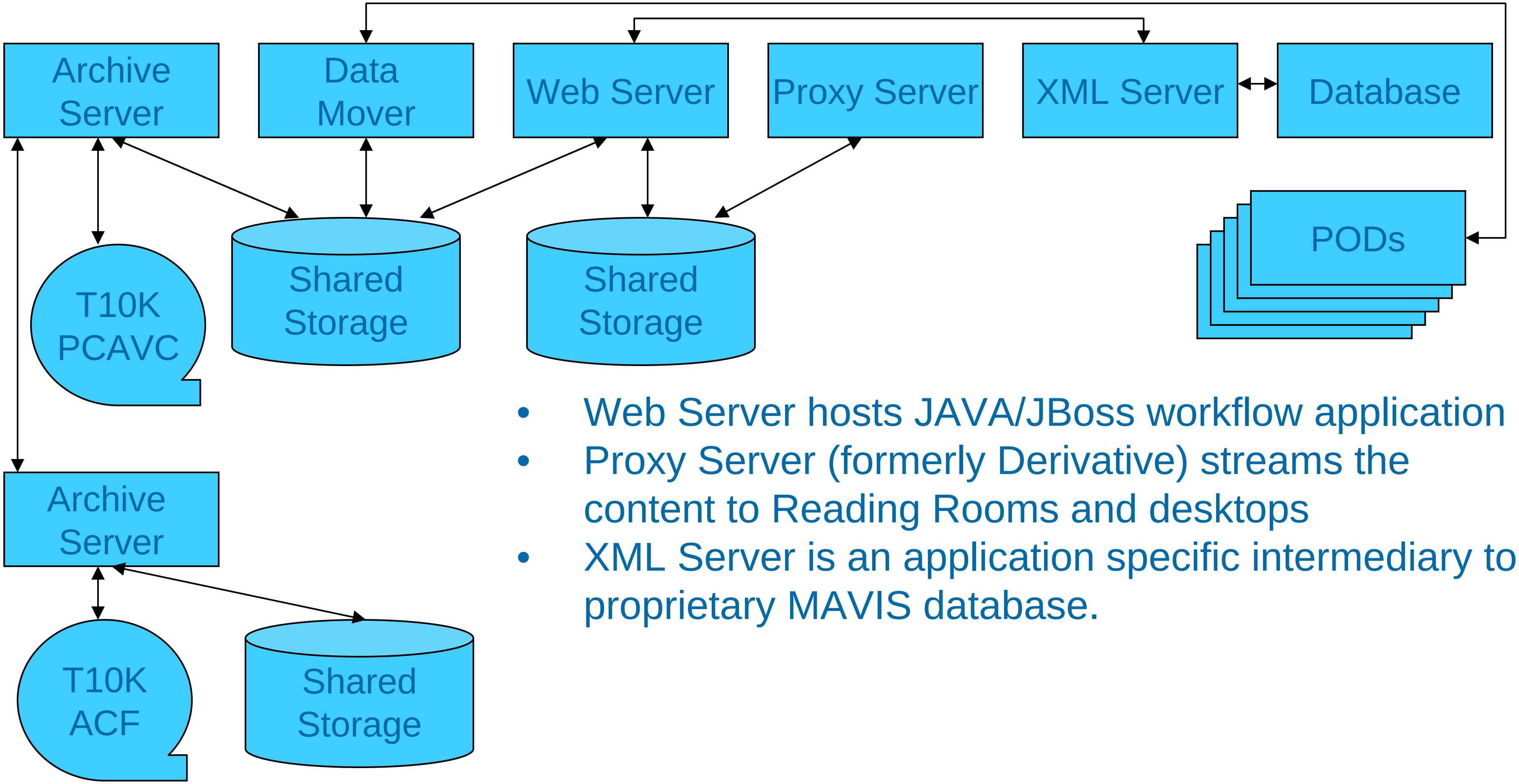
Archive Storage Infrastructure



- PODs generate data
- Workflow software copies data via signiant/samba to the Data Mover / Shared Storage
- Data Mover verifies files with SHA-1
- Archive Server reads from storage and writes to tape
- Archive Server reads from storage and writes to remote Archive Server
- Every 1GB of incoming data requires 4GB of total throughput: 1 write/3 reads (SHA1, local, remote)

Functional Architecture – User Interface

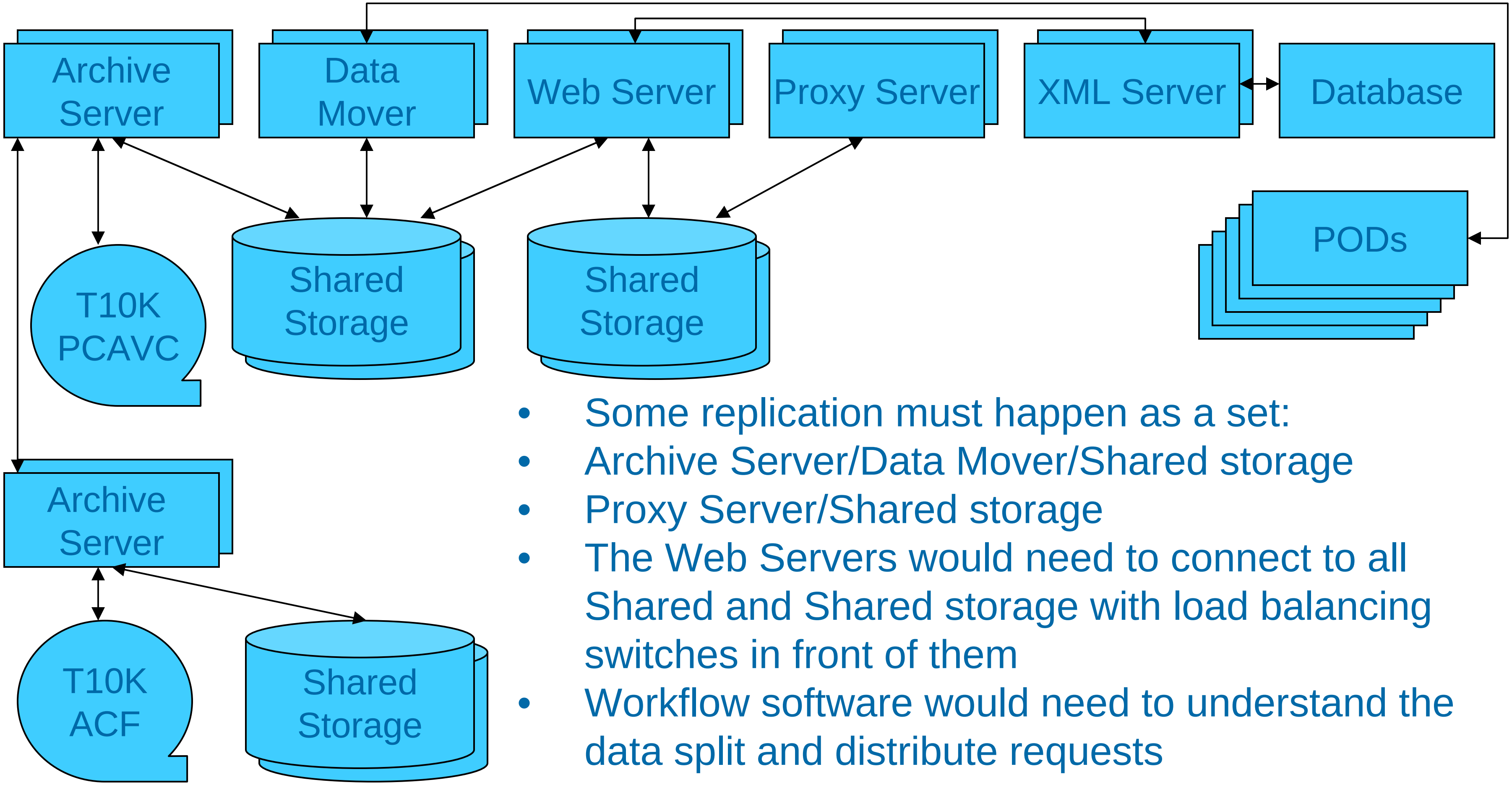
Archive Storage Infrastructure



- Web Server hosts JAVA/JBoss workflow application
- Proxy Server (formerly Derivative) streams the content to Reading Rooms and desktops
- XML Server is an application specific intermediary to proprietary MAVIS database.

Functional Architecture - Scaling

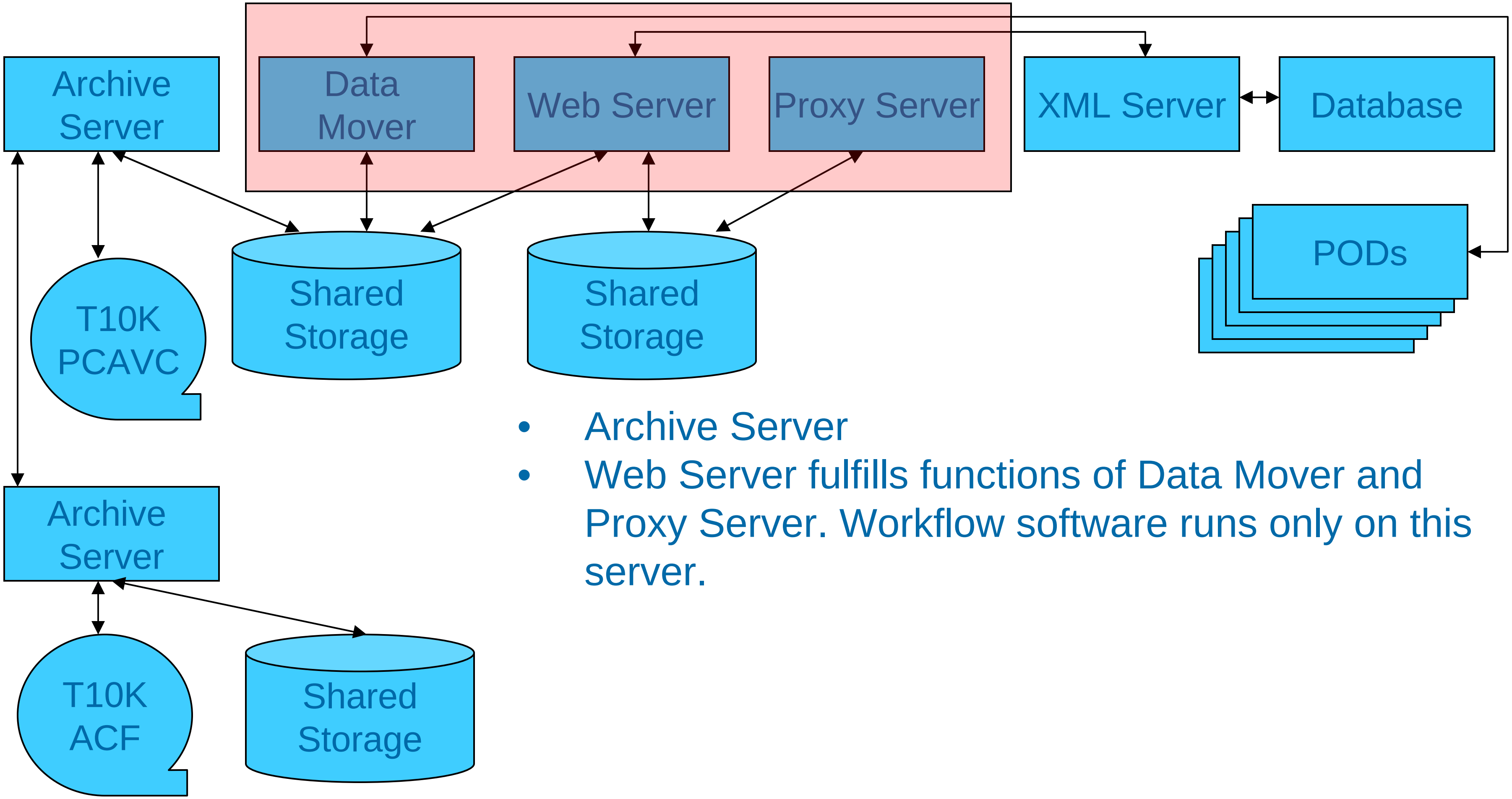
Archive Storage Infrastructure



- Some replication must happen as a set:
- Archive Server/Data Mover/Shared storage
- Proxy Server/Shared storage
- The Web Servers would need to connect to all Shared and Shared storage with load balancing switches in front of them
- Workflow software would need to understand the data split and distribute requests

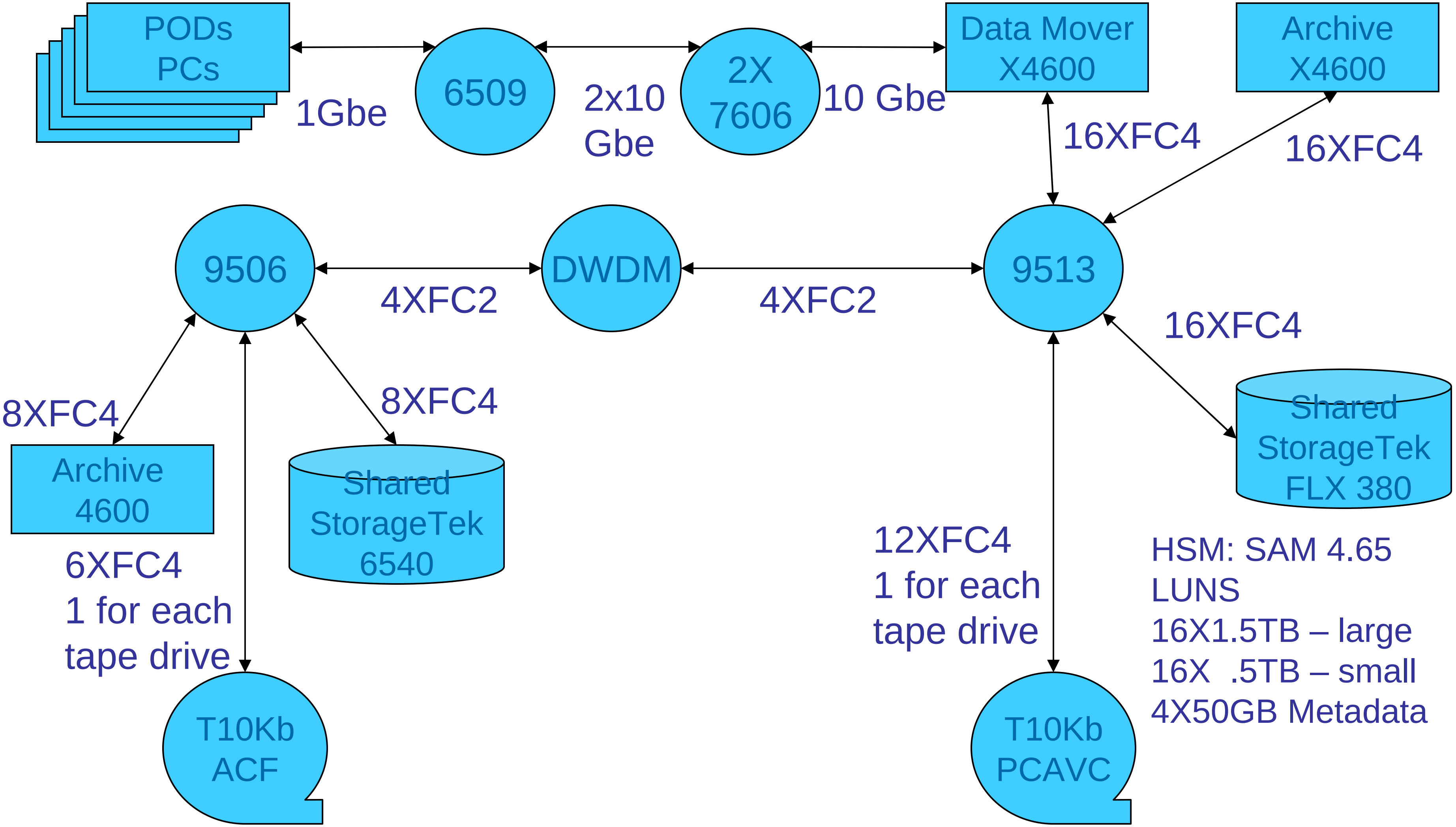
Functional Architecture – Current

Archive Storage Infrastructure



- Archive Server
- Web Server fulfills functions of Data Mover and Proxy Server. Workflow software runs only on this server.

Physical Implementation V1: 2 GB/s throughput



PODs PCs

1Gbe

6509

2x10 Gbe

2X 7606

10 Gbe

Data Mover X4600

Archive X4600

16XFC4

16XFC4

9506

4XFC2

DWDM

4XFC2

9513

16XFC4

8XFC4

8XFC4

Archive 4600

Shared StorageTek 6540

Shared StorageTek FLX 380

6XFC4
1 for each
tape drive

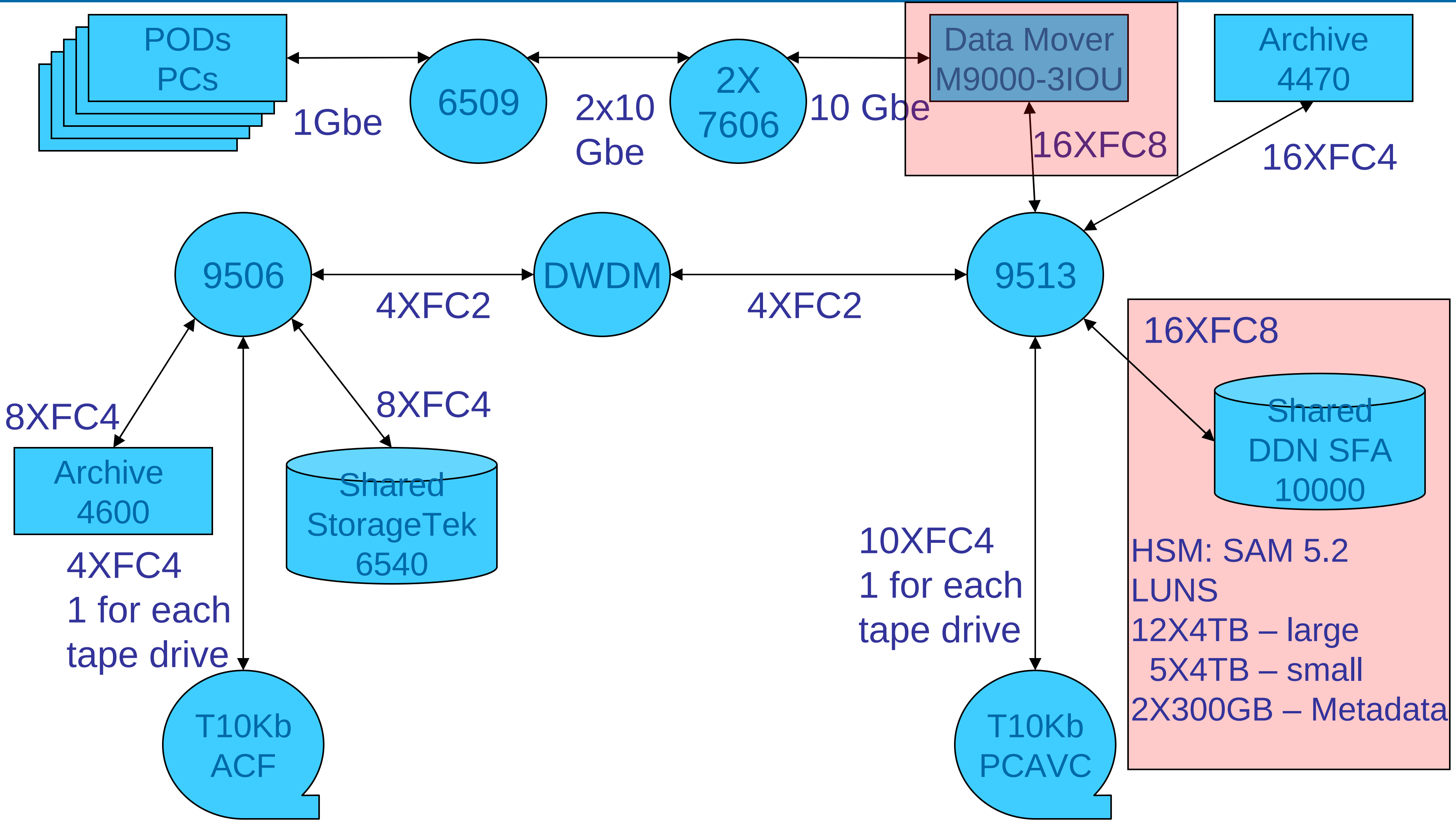
12XFC4
1 for each
tape drive

HSM: SAM 4.65
LUNS
16X1.5TB – large
16X .5TB – small
4X50GB Metadata

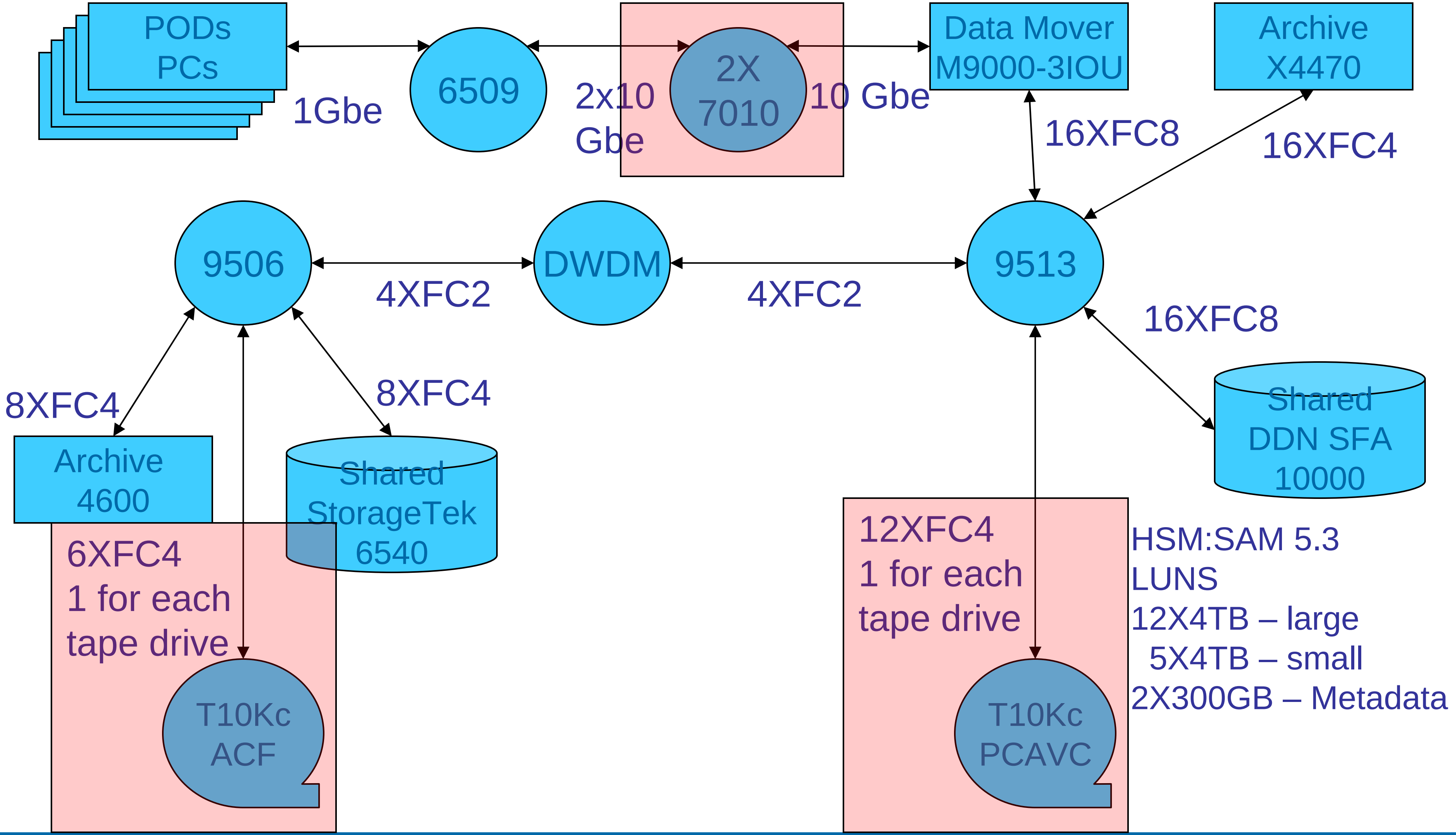
T10Kb ACF

T10Kb PCAVC

Physical Implementation V2: 6.5 GB/s throughput

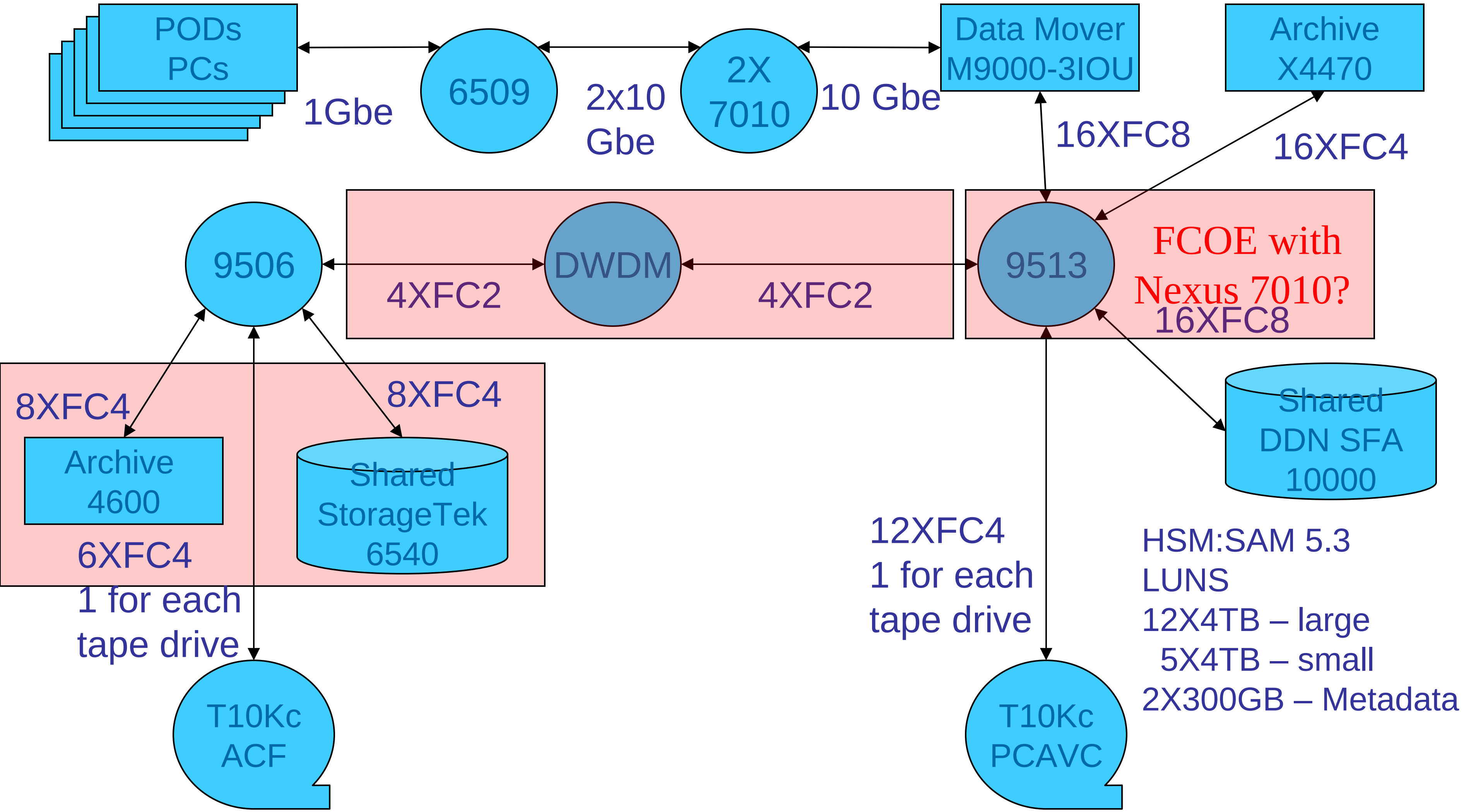


Physical Implementation V2+: 6.5 GB/s throughput Infrastructure Upgrades



HSM:SAM 5.3
LUNS
12X4TB – large
5X4TB – small
2X300GB – Metadata

Physical Implementation V3: 6.5 GB/s throughput
What's Next?



HSM:SAM 5.3
LUNS
12X4TB – large
5X4TB – small
2X300GB – Metadata

Discussion

- Market Survey of HSM solutions
- Fibre Channel over Ethernet
- Monolithic versus distributed
  - Sandy Bridge / Ivy Bridge – We can probably ride the monolithic bandwidth curve
  - Cost: compare software development and maintenance to hardware costs
  - Effective 2-5 year planning of ingest magnitude
- Contact: Scott Rife: srif at loc dot gov