

Backup Service at CERN

Alex Iribarren

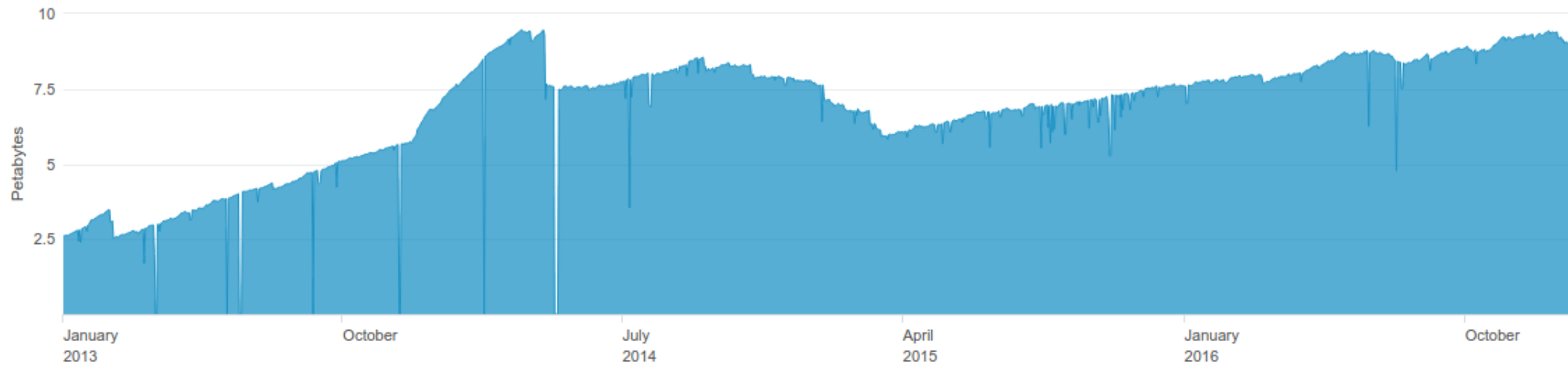
IT-DSS-TAB

What we do

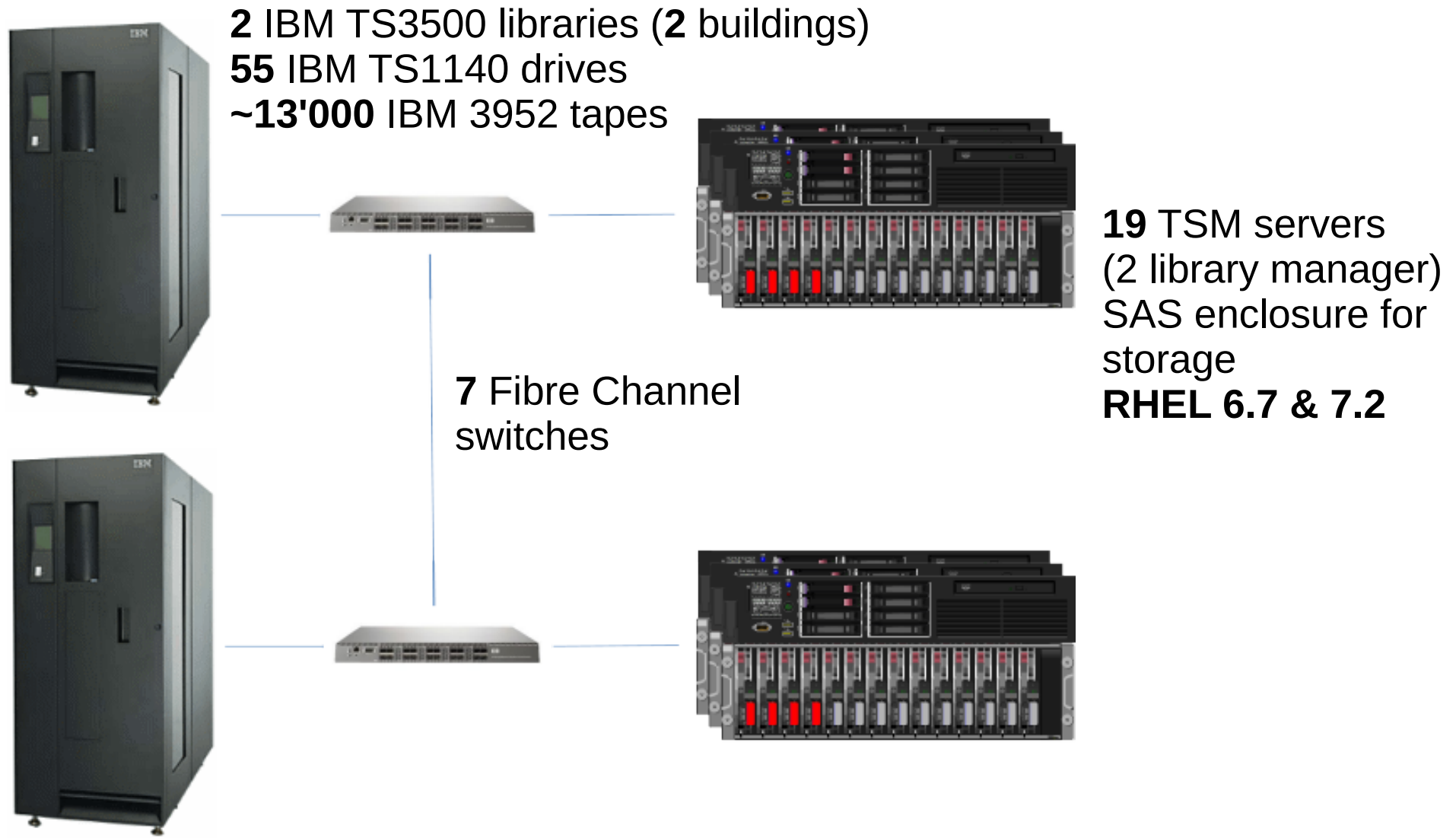
- Provide a centralized service to backup CERN-owned server machines containing critical data in the computer center
 - Network filesystems (~60000 AFS volumes, ~2500 DFS)
 - Email (~30000 mailboxes)
 - Web sites (~12000 sites)
 - Databases (~120 DB servers)
 - Servers (~1000 Linux and Windows servers)
 - Virtual machines (~200 hypervisors)
- No backup for physics data or user machines

How much data?

- 9 PB of data on tape
- 1.7 billion files
- ~600 clients, ~70 user groups
- ~55 TB average daily traffic

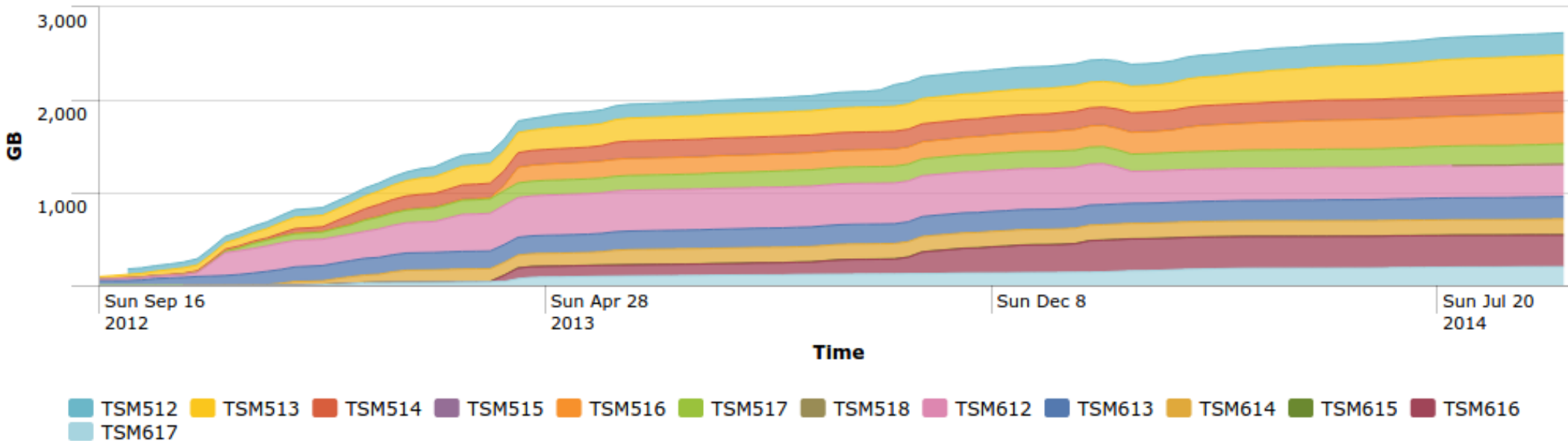


Architecture

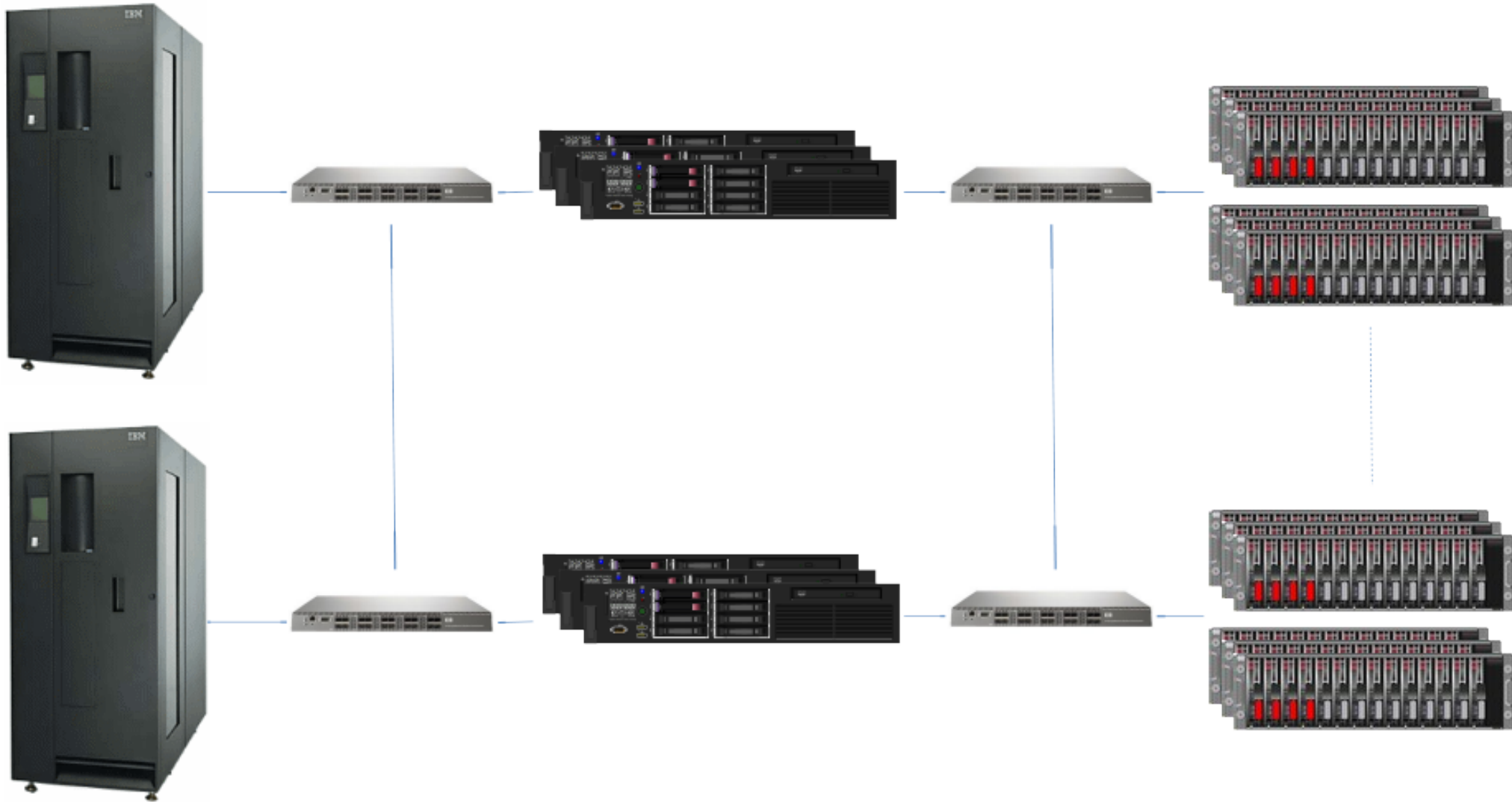


Several architectures iterations

- Pressure to reduce special cases
 - ... and cost, complexity
- Increase performance to cope with load
- Deal with TSM software evolution
 - TSM 6 DBs growing faster than expected



Pre-2013 Architecture



2013 Architecture



Linux servers
with attached
SAS expanders



- **Software RAID** was introduced
 - RAID1 (DB) and RAID1E (DATA)
 - 2 favors of disks:
 - Fast small disks for TSM DB
 - Slower large disks for DISK storage pool
- Pairs of server/SAS
 - **Redundancy in case of expander failure**



2014 Architecture



Linux servers
with attached
SAS expanders



- Software RAID
 - RAID1E everywhere (faster daily DB backups)
- 1 flavor of disks:
 - Large disks for DB and FILE storage pool
- 1 SAS expander per server
 - Less error prone during disk replacement
 - SAS expanders do not break

Our experience with RHEL

- Required by TSM for support reasons
 - More importantly, by IBM's tape device driver
- IBM's release cadence usually means we're behind the latest releases
 - We rely on Extended Update Support (EUS)
- EUS has not been without issues
 - Constant syncing issues due to CDN
 - No 7.0 EUS for some reason
 - Not released on time (EUS 7.2 released a couple of weeks after RHEL 7.3)

Questions?