

2014-2015 Xroot and NFS Plan

18 Feb 2015 **Working DRAFT**

Intro

Fermi currently supports an xrootd service consisting of 33 Sun Thumper/Thor class servers plus 12 Dell R610/620 class servers. Fermi also supports an NFS service consisting of four Sun Thor class servers. One additional Thumper class server is used by LSST. The Sun servers range in age from 4 to 8 years. The oldest Dell systems are three years old.

Fermi is looking to purchase two storage building blocks with ~640 TB **each year** with ~1/2 to accommodate new incoming data, and the remaining ~1/2 to allow for aged equipment retirement/repurposing. The following proposal resulted from a meeting held 2/18/2015 with Shirley, Yemi, Lance, Renata, Wilko, Richard, and Tom.

Two new Dell R620 systems (~640 TB) were purchased summer 2014 and are in the process of being commissioned.

Two additional Dell systems (~640 TB) are anticipated for the summer 2015. The expected use of these servers will be to increase the xrootd GPFS pool.

2014-2015 Program

- First experience with GPFS. Use new servers, fermi-gpfs01 and fermi-gpfs02 exclusively as xrootd resources running GPFS
- Retire six wains which have the highest failure rate: 053, 054, 055, 056, 069, 071 (??)
- Modernize and improve NFS system, with the special aim of increasing and improving service for the User and Group spaces currently on wain025
- Update the LSST storage server, wain006
- Donate left-over wains to Steffen Luitz for LZ use at IR-2

Proposed configuration changes for Fermi xrootd and NFS servers

- New servers: fermi-gpfs01 and fermi-gpfs02
 - Dual-connect storage between these two machines
 - Internet connectivity (2 x 10 Gbps per host?)
 - Install GPFS
 - Install xrootd
 - Balance data across xroot cluster
 - Return the 25% storage reserved for NFS back to xrootd
- New NFS/GPFS service on former fermi-xrd01 and fermi-xrd02 (and wain069, wain071)
 - fermi-xrd01 and fermi-xrd02
 - Drain xroot data (~180 TB)
 - Swap 2x R610 with SCS-owned R720 machines
 - Rename (fermi-gpfs03 and fermi-gpfs04 ?)
 - Internet connectivity (1 x 10 Gbps per host ?)
 - Decide upon storage configuration
 - Number of spindles for Users/Groups (from wain025)
 - Number of spindles for production partitions (from wain026 and wain032)
 - Number of "spare" spindles for future expansion
 - Install GPFS (total capacity will be ~160 TB)
 - Migrate wain025, wain026, wain032 to new system
 - Select two wains for CNFS service (lots of memory + fast ethernet)
 - Drain xrootd data from wain069 and wain071 (~60 TB, both are unreliable hosts due to the Seagate disks)
 - Swap the 4x1Gb interface with 10Gb nic from wain080 and wain081
 - Rename hosts (fermi-cnfs01, fermi-cnfs02 ?)
 - Install GPFS and CNFS software
 - Configure so that wain025 partitions are handled in such a way that their activities do not negatively impact access to other partitions
- Upgrade NFS service for LSST
 - Repurpose wain025 to replace wain006
- Drain xroot from other wains that are to be retired
- Retirement list:
 - wain026
 - wain032
 - wain006
 - wains that are unreliable (due to Seagate drives?)
 - wain053 (29TB to be moved)
 - wain054 (22TB)
 - wain055 (27TB)
 - wain056 (30TB, total = 108TB)

- **Retirement option:** Given that wain05x are the newest and most powerful Sun servers in the cluster, we might consider swapping physical disks from wain017/019/020/021 with wain053/054/055/056, then retiring the older machines. Advantages include: newer hardware (~2 years difference), more powerful hardware (12 cores vs 4), more memory (32 GB vs 16), more flexibility if we again repurpose the machines in the future. Wilko rightfully points out that this option requires more labor. Is it worth the extra work? Early indications suggest yes. The following disk migration is suggested:

Start	replace Seagate with Hitachi from these machines, then retire from xrootd service
wain053/54/55/56	wain017/19/20/21

Milestone Timeline

Date	Milestone
7/1/2014	new servers arrive, fermi-gpfs01 and fermi-gpfs02
7/30/2014	storage arrays arrive
9/18/2014	cables located, beginning of GPFS testing at SCS
1/13/2015	xrootd in production (readonly), and data migration/balancing begins
2/18/2015	agree upon general Fermi storage plan
3/13/2015	fermi-xrd01/02 and wain069/071 drained  GXR-71 - Jira project doesn't exist or you don't have permission to view it.
??	former fermi-xrd01/02 + wain069/071 configured for CNFS service
??	NFS data migrated to new service and ready for production
??	wain006 migrated to wainZZZ
??	all remaining retirees drained

Ref: [NFS/GPFS and Xroot Disk Assets](#)