Reliable Storage by CNAF

Vladimir Sapunenko,
Head of Data Management and Storage at CNAF

Outlook

- Storage at CNAF
  - Hardware
  - Software
  - Services
- How to reach reliability targets
  - HW and SW Solutions
  - Procedures
- Experience to share
- Closing remarks and open questions

CNAF overview

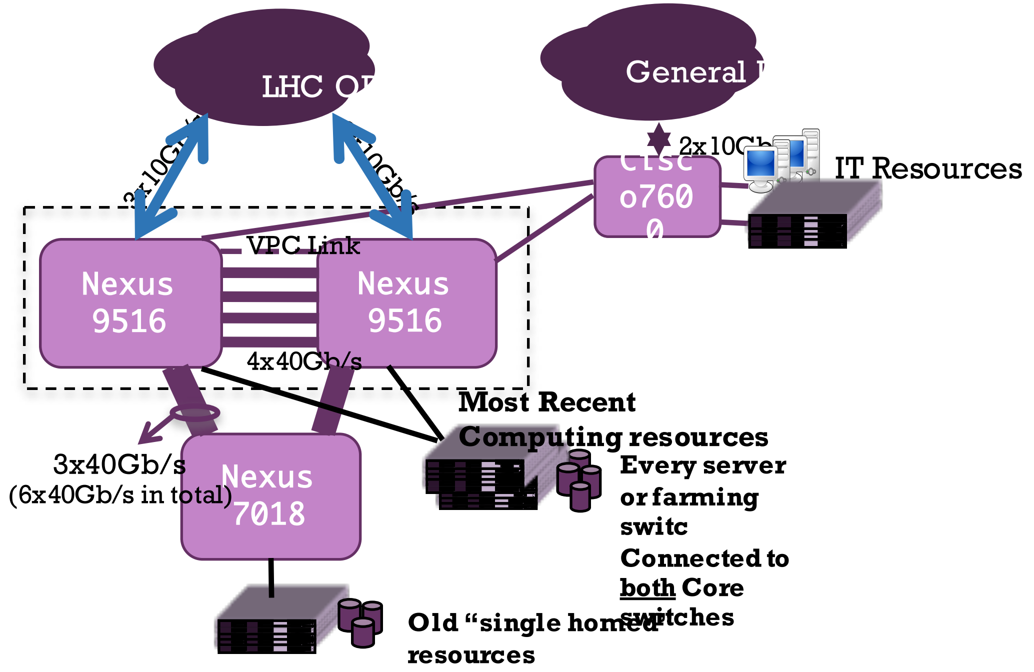
- Storage provided by CNAF (total)
  - 23 PB usable (29 PB raw) on disks
  - 56 PB on tapes
- CNAF provides
  - Tier-1 services for all 4 LHC experiments
  - Tier-1 services for several non-LHC experiments
  - General data and computing services for ~30 HEP and Astrophysics experiments

Core infrastructure

- Reliability of core services is essential
  - Power
    - Every storage component (server, controller, switch, disk enclosure) powered by two independent power lines
    - Both power lines backed-up by inertial UPS and diesel generator
  - Cooling
    - 6 chillers (4+2 redundancy)

Network

- **CORE Switches (Cisco Nexus 9516)**
  - Every **ToR** (Farming) Switch and every **Disk server** is connected to both **core switches**
  - The old Nexus 7018 is connected to both Core Switches with 3x 40 Gb/s Interfaces (6x 40 Gb/s in total) and aggregates the old single homed resources (connected with only one Ethernet interface).



Storage overview (disk)

Manuf	Model	N. units	N. Of cont/ unit	N. Encl. /unit	RAID configuratio	SSD, N*Capacity (GB)	SAS, N*Capacity (GB)	NL-SAS, N*Capacity (TB)	Year in service
DDN	SFA10K	1	2	20	6(8+2)	0	34 * 300	1166 * 3	2012
DDN	SFA12K	2	2	20	6(6+2)	0	30 * 300	1650 * 3	2013
DELL	MD3860f	4	2	3	DDP 180(8+2)	0	0	180 * 4	2014
DDN	SFA12K	2	2	10	6(8+2)	0	36 * 300	800 * 8*	2015
Huawei	OceanStor 6800	1	6	12	2.0 95(6+2)	25 * 600	0	855 * 6	2017
DELL	MD20f	2	2	1	1(1+1)	24 * 300	0	0	2015
DotHill		1	2	1	1(1+1)	4 * 300	20 * 600	0	2013

Disk space efficiency usage (usable/raw) – 80%
* - 4KN sectors

Storage overview (I/O servers)

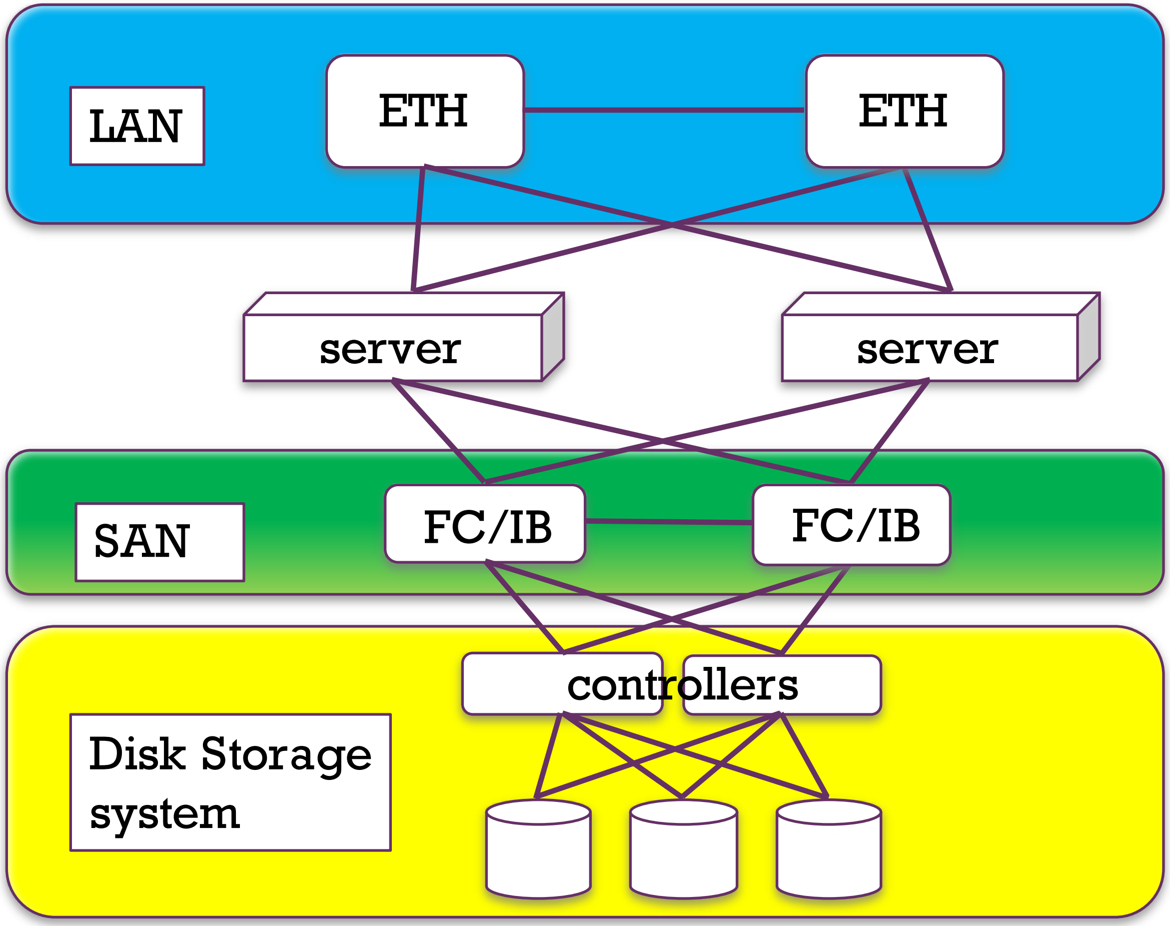
- 66 I/O servers in total (NSD, GridFTP, XrootD, HSM)
- Dual power supply
- RAM: 24 GB (1x10 GbE) - 128 GB (4x10 GbE)
- Local System disk protected by mirroring (RAID1)
- LAN: 10GbE
  - More recent servers connected to both core switches (Ethernet bonding, active-active)
  - Old servers with 1x10GbE ports connected to a "aggregating" core switch
- SAN: Every server connected to two FC or IB switches (active-active)
- Every Storage device accessible from several servers (shared storage model)
  - Any server can be taken off-line at any time without compromising data access.

Storage overview (tape)

- 1 Oracle(StorageTek) SL8500 library
  - 17 T10kD (production) tape drives
  - 9 T10kC (backup) tape drives
- TSM servers
  - Production: 1 active, 1 stand-by, 1 devel
  - Backup service: 1 active, 1 stand-by
- HSM servers
  - Production: 6 active, 6 stand-by
  - LTDP: 1 active, 1 stand-by
  - Testbed: 2 active

Storage interconnect infrastructure

- Fiber Channel
  - FC4, FC8, FC16 (4-8-16 Gb/s) - 620 ports
  - I/O Server to storage
  - HSM nodes to storage
  - HSM nodes to tape drives
- InfiniBand
  - FDR (56 Gb/s) – 72 ports
  - I/O Server to DDN SFA12K storage
- Ethernet
  - N * 10 GbE
  - I/O Server to LAN
  - Ethernet bonding (N=1-4)

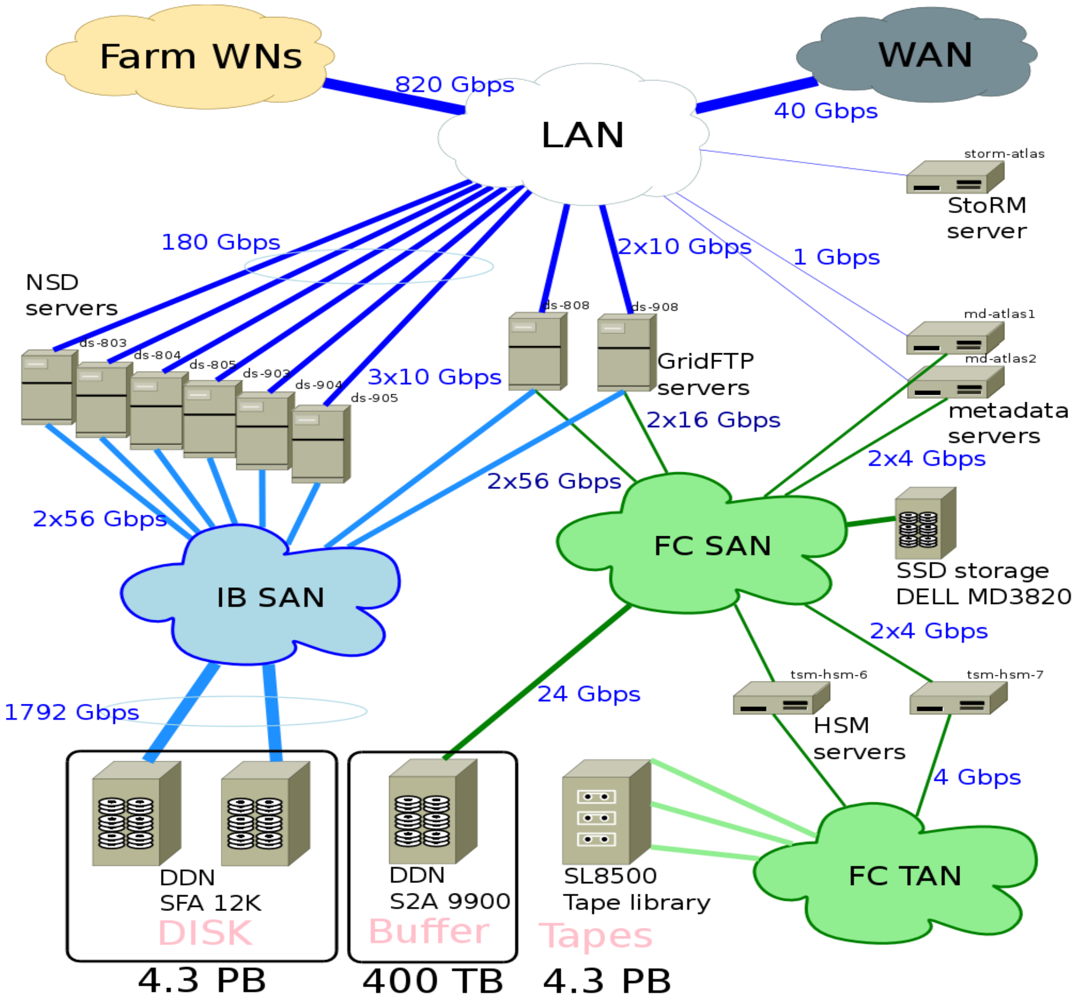


Storage Services

- Remote and local access
  - SRM (StoRM)
  - GridFTP
  - XrootD
  - WebDAV
- Local access only
  - POSIX via GPFS (~1000 clients, mainly WNs)

ATLAS example

- Servers (dedicated):
  - 2 as GridFTP/XrootD
  - 6 as NSD (POSIX)
  - 2 as metadata servers
  - 1 (VM) as StoRM FE/BE
  - 2 as HSM (1 active, 1 stand-by)
- Storage (shared)
  - 2 DDN SFA12k as main storage
  - 2 Dell MD3820f as metadata storage
  - 1 SL8500 Tape library



Software

- Proprietary
  - GPFS (IBM, site license)
  - TSM (IBM, per socket license)
- Open Source
  - GridFTP
  - XrootD
  - StoRM (locally developed)
  - GEMSS (locally developed)

Why use GPFS?

There are several reasons for this

- Historically, lack of personal forced us to look for a less demanding (in sense of management efforts) system
- GPFS = Software Defined Storage
  - Easy to re-size, replace HW, move data between different devices, add or remove I/O servers – everything with no service interruption
- GPFS – parallel file system
  - Always the same (max!) performance with no need of replication
- It gives POSIX access
- At the end, as somebody said "it just works" and I do agree with this ☺

Procedures and human power

- All Storage resources and services managed by the Data Management and Storage group of 6 FTE
  - Permanent part: 3 technologists + 2 technicians
  - Variable part: 2 fellowships (half-time)
- Intensive pre-production testing of new HW
- Installation and configuration of servers and services via Foreman/Puppet
- Extensive monitoring and automated actions via NAGIOS
- Redundancy everywhere (no single point of failure)
  - Active-standby (with manual fail-over) if no redundancy possible
- All the work done during working hours
  - Only one intervention per year during off-hours
- HW support contracts – 4h on site (but only during working hours)

Experience to share

- 4K sector size
- Capacity per I/O server ratio
- Dealing with TSM support costs
- On-line and off-line data movement
- Ethernet bonding
- Distributed RAID (DDP, RAID2.0, etc)
- Use of SSD for metadata

4KB sector size

- New high capacity disks often come with sector size of 4KB
- There are two flavors of such disks:
  - Native 4KN – 4KB sectors exposed to SCSI layer
  - Advanced Format – 4KB sectors translated into 512B by disk's interface
- Readiness of the support for 4 KB logical sectors within operating systems differs among their types, vendors and versions.
  - Example: GPFS v.3.5 does not support 4KB sectors

Capacity per I/O server ratio

In absolute numbers varying from **170** to **925** TB/server

- Every **GridFTP** or **XrootD** servers directly (via FC o IB) accessing the whole data of the experiment

Exp	Total capacity	N. servers	SAN	LAN, Gbps	LAN bandwdth, GB/s	TB/server
ALICE	3.7	4	IB (FDR)	4x10	20	925
ATLAS	4.3	2	IB (FDR)	2x10	5	2150
CMS	4.3	4	FC8	1x10	5	1075
LHCb	3.1	2	FC8	1x10	2.5	1550

- **POSIX** access provided by NSD servers

Exp	Total capacity	N. servers	SAN	LAN, Gbps	LAN bandwdth, GB/s	TB/server
ATLAS	4.3	6	IB (FDR)	4x10	30	1075
CMS	4.3	20	FC8	1x10	25	215
LHCb	3.1	12	FC8	1x10	15	258

Dealing with TSM support costs

- Support service for TSM is essential
- IBM requires support contract for every ACTIVE TSM node
  - The cost depends on
    - CPU family
    - N of sockets
  - Reducing N of sockets doesn't work
    - Every PCI slot hard wired to a CPU slot
    - Cannot use all needed PCI slots
  - Passed from active-active configuration to active-stand-by for all HSM nodes reducing by a half cost of the support

Ethernet aggregation (bonding)

- Generally works well
- Some problems observed in connection between bonded 1GbE and 10 GbE interfaces
- Needs great attention with HW selection
  - may have compatibility issues
- Currently using 2x, 3x and 4x 10GbE
- Helps reduce number of I/O servers
  - Helps reduce number of GPFS licenses

Distributed RAID

- Usually implemented as "floating" RAID6 (8+2) over bigger (>10) disk pool
- Using "reserved capacity" to restore missing blocks in case of disk failure
- With a disk pool big enough recovery time becomes significantly reduced
  - Failure of 4TB disk: Disk pool of 180 disks – 3.5 hours to restore redundancy (under heavy I/O load) against 20-22 hours in traditional RAID6
  - Failure of 6TB disk: Disk pool of 95 disks – 3 hours to restore redundancy
- Drawbacks
  - I/O performance may be affected by ~20%
  - Limits on LUN size may require creating more LUNs on the same disk pool reducing to zero I/O optimization made by file system software

Problems to solve

XrootD:

Small block reads vs. Large block size on FS

I/O server sizing/optimization:

How many connection one server able to manage?

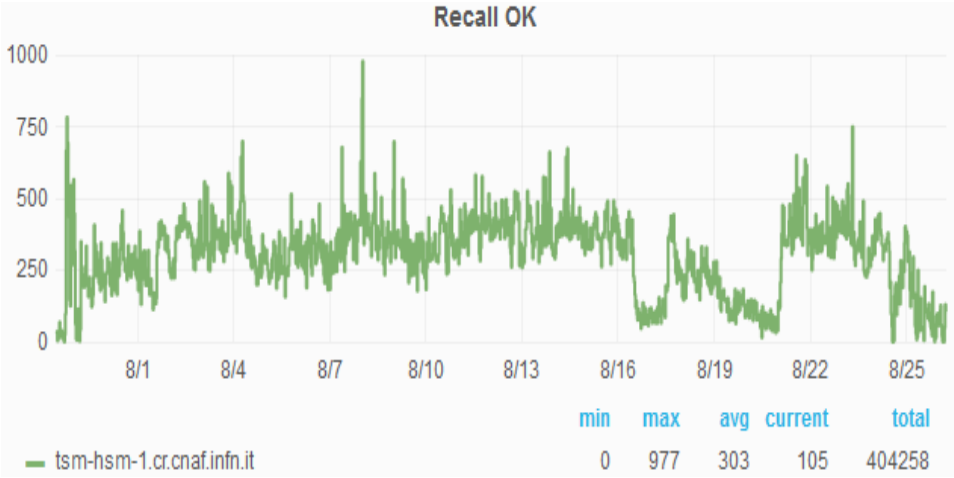
```
----total-cpu-usage---- -dsk/total- -net/total- ---paging-- ---system--
usr sys idl wai hiq siq| read  writ| recv  send|  in   out | int   csw
  2   4  93   0   0   1|1349M    0 | 232M   23M|   0     0 |  19k   24k
  2   4  94   0   0   0|1204M    0 | 196M   29M|   0     0 |  17k   21k
  2   4  94   0   0   0|1363M    0 | 191M   23M|   0     0 |  17k   22k
  2   4  94   0   0   0|1184M    0 | 223M   32M|   0     0 |  18k   21k
  2   4  94   0   0   1|1399M  76k| 188M   31M|   0     0 |  19k   23k
```

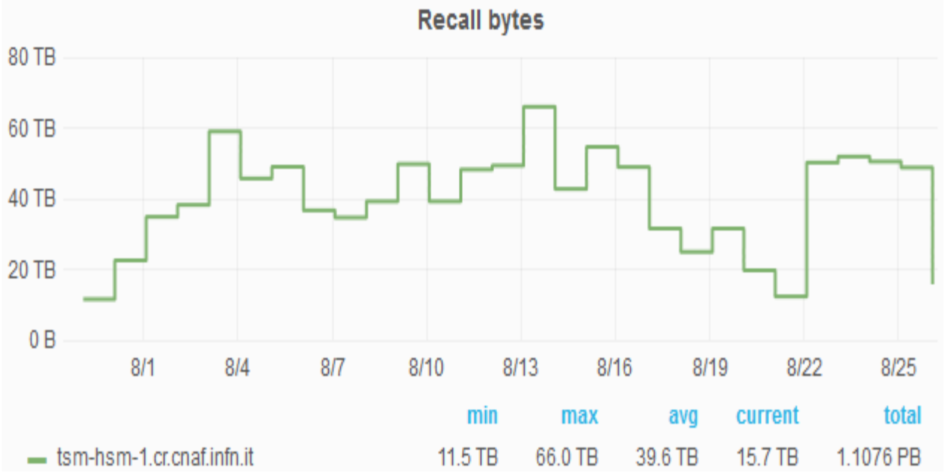
Plans for future

- Continue server consolidation
  - I/O servers: Eliminate single port 10 GbE servers
  - HSM servers: 3xFC8->4xFC16
- Increase Ethernet speed
  - With the next storage tender (to be judged this week) will get 40 or 100 GbE (server-to-LAN) connectivity

“Not intended” Stress test by CMS (400K recalls)

- In 28 days (from 30/7 to 26/8):
  - bytes recalled: 1.11PB (of 12PB total) - 9.2 %
  - tapes used in recall: 1103 (of 1544 total) - 72 %
  - files recalled 400K
  - number of mounts: 9858
  - mount/day (average) 352
  - mount/tape 9 (average, 17 mount/tape max)
  - number of tape drives used, average: 6 (max 7)





Vladimir.Sapunenko@cnaf.infn.it

Thank you for your attention!

Questions?