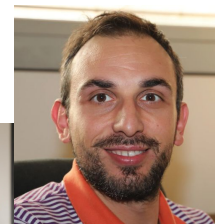


# State of Storage

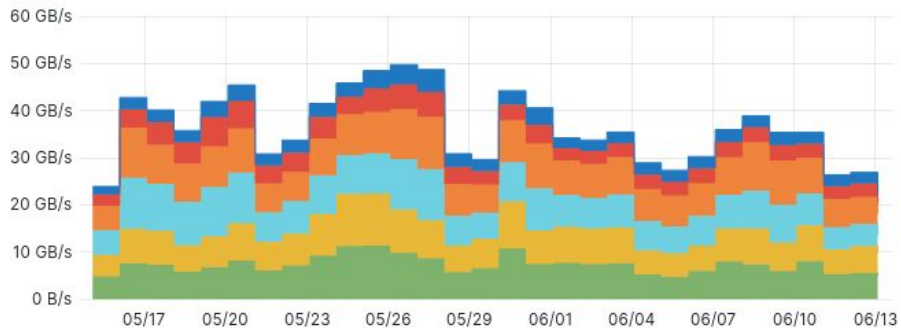
CdG 16 giugno, 2023



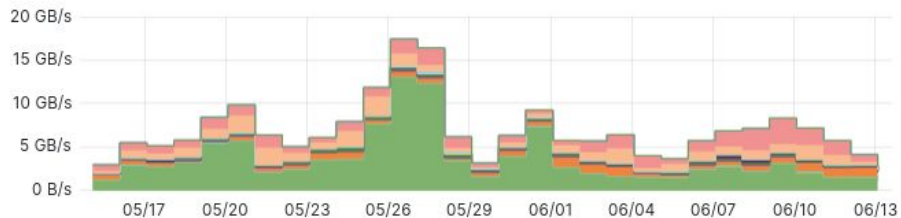
# Business as usual

## Last month

### All servers network traffic out (reading)

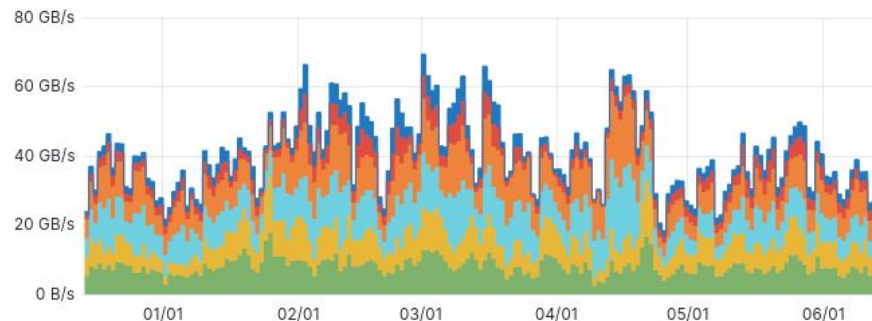


### Gateway traffic out (non POSIX reading)

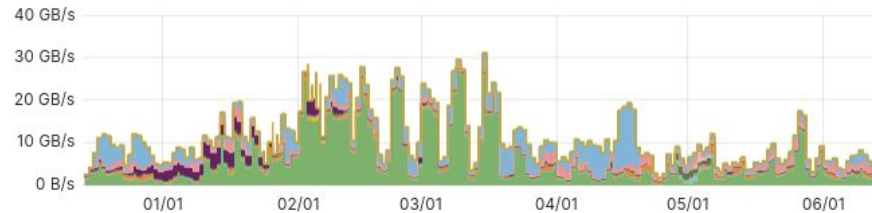


## Last 6 months

### All servers network traffic out (reading)



### Gateway traffic out (non POSIX reading)



# Disk storage in produzione

Installed: **53.64** PB, Pledge 2023: **69.6** PB, Used: **46.8** PB

Storage system	Model	Net capacity, TB	Experiment	End of support
ddn-10, ddn-11	DDN SFA12k	<b>10120</b>	ALICE, AMS	<b>12/2022</b> (20 spare hdd)
os6k8	Huawei OS6800v3	3400	GR2, Virgo	12/2023
md-1,md-2,md-3,md-4	Dell MD3860f	2308	DS, Virgo, Archive	11/2023
md-5, md-6 e md-7	Dell MD3820f	50	metadati, home, SW	11/2023 e 12/2024
os18k1, os18k2	Huawei OS18000v5	7800	LHCb	12/2023
os18k3, os18k5, os18k5	Huawei OS18000v5	11700	CMS	2024
ddn-12, ddn-13	DDN SFA 7990	5840	GR2,GR3	2025
ddn-14, ddn-15	DDN SFA 2000NV	24	metadati	2025
os5k8-1,os5k8-2	Huawei OS5800v5	8999	ATLAS	2027
Cluster CEPH	12xSupermicro SS6029	3400	ALICE, cloud, etc.	2027

# Acquisti recenti e futuri

- Gara storage 2022 (14PB netti)
  - Il vincitore è Lenovo con ThinkSystem DE6600  
(Derivato dal NetApp DE6600 con un enclosure con i dischi NVME)
  - In fase di collaudo
- AQ storage 2023-2024
  - 66PB nel 2023 + 14PB nel 2024
  - Il vincitore (provvisorio) è Huawei con sistemi OceanStore Micro 1500/1600
    - Ribasso 34% (!)
    - 4 controller x sistema, 8 sistemi capacitivi + 2 sistemi NVME
    - Backend FC32
- Gara Tape Library
- Gara Tape



# Current SW in PROD

- GPFS 5.1.2-8 (to be updated soon to 5.1.2-11)
- StoRM BackEnd 1.11.21 (to be updated soon)
- StoRM FrontEnd 1.8.15 (latest)
- StoRM WebDAV 1.4.1 (to be updated soon)
- StoRM globus gridftp 1.2.4
- XrootD 5.5.4-1
  - First updated for CMS, upon request from the experiment; recently updated for ATLAS and LHCb as well
  - ALICE CEPH updated to 5.5.5-1.el8
- Ceph 16.2.6 (Pacific)

# XrootD ALICE CEPH

- XrootD has been updated on all ALICE CEPH servers
  - In order to solve issues experienced with CEPH mount point availability
  - Updated OS (RedHat 8) and kernel
  - XrootD version 5.5.5-1
- Standard XrootD packages installation
- Puppet-based configuration
- System under observation

# Tickets and problems

- ATLAS

- Request to enable token-support for disk storage areas from Petr Vokac (24/5), allowing the following scope-based authZ schema:
  - <https://xfer.cr.cnaf.infn.it:8443/atlas/> -> storage.\*:/atlasdatadisk/
  - <https://xfer.cr.cnaf.infn.it:8443/atlas/> -> storage.\*:/atlasscratchdisk/
- Need to unify the two disk storage areas (/atlas/atlasdatadisk, /atlas/atlasscratchdisk) into one (/atlas)
- Need to redefine the access points for the tape storage areas in CRIC and Rucio (/atlas->/atlas-tape) to avoid nested storage areas
  
- TPCs in push-mode to GOOGLE\_EU fail for all StoRM sites (GGUS [158487](#))
  - Issue for StoRM developers, on hold

# Tickets and problems

- CMS
  - Gridftp still used, only for SAM tests :-(
    - Token-support enabled for disk storage area (StoRM WebDAV) and for XrootD services (GGUS [161981](#))
    - Problems with some files during staging (GGUS [162250](#))
      - Currently under investigation (mail thread, CERN jira system) together with FTS people
      - Involved both corrupt (partial) files or 0-length files; hopefully the same root cause, with the difference that 0-length files do not make it to tape
    - Confirm deletion campaign could start (GGUS [161795](#))



# Tickets and problems

- LHCb
  - 15/6: meeting at CNAF with LHCb data management team
    - Agreed upon a short, scheduled downtime to rename folders in disk so to avoid srm calls to deduce physical paths
  - FTS transfer errors between CNAF and CERN-CTA (GGUS [162189](#))
    - Waiting for CERN to check their systems for jumbo frames support
  - “No space left on tape” GGUS [161941](#)), fixed triggering garbage collector
- AMS02
  - Access to storage enabled for VOMS proxies obtained with iam-ams
  - The experiment switched from srm + gsiftp -> srm + https
    - Access point for the tape storage area changed (/ams -> /ams-cnaf)
    - On Monday 26th, we will switch off gsiftp
- Belle
  - Tests srm+https on hold

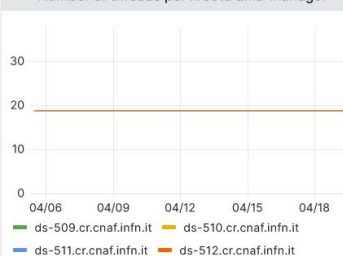
# Tickets and problems

- Dampe
  - Waiting for feedback from Bari on the issue of XrootD redirector always contacting one out of 4 available XrootD servers at CNAF
- Dune
  - StoRM WebDAV storage area for writing to tape (dedicated group 'dune' in iam-t1-computing)
- Icarus
  - Configured StoRM WebDAV storage area; issues when using root protocol for accessing files
  - Saturated buffer space during migration; need to tune the rate for writing (100 TB, then wait one day for migration and clean up)
- Virgo
  - Request for a new tape storage area, parallel to /virgod0t1
  - Waiting for confirmation to delete Runs (a folder on tape)

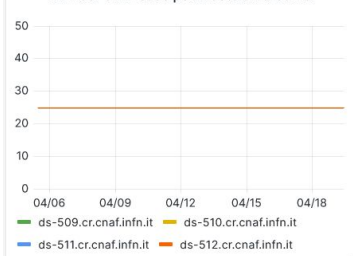
# Tickets and problems

- All the no-LHC xrootd instances have been moved to 4 servers (ds-509, ds-510, ds-511, ds-512) with the alias xrootd-archive/xrootd-ams.
  - Dampe is the main user by far
    - Unauthenticated access (!) from the farm WNs
    - Only one out of 4 servers is used (all of them subscribe to the redir in Bari)
    - Currently trying to identify a contact person for these activity within Dampe

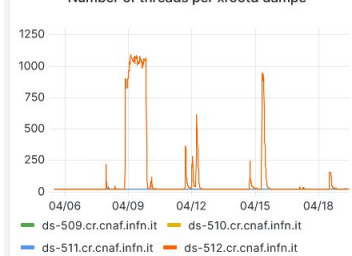
Number of threads per xrootd ams-manager



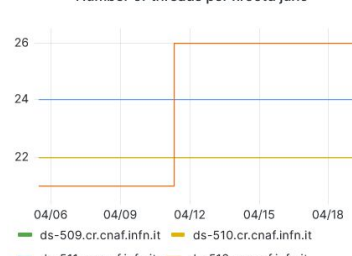
Number of threads per xrootd ams-server



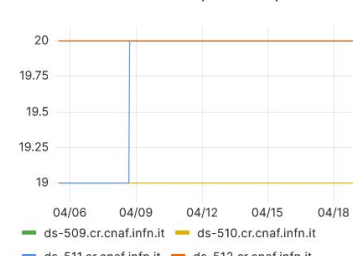
Number of threads per xrootd dampe



Number of threads per xrootd junio



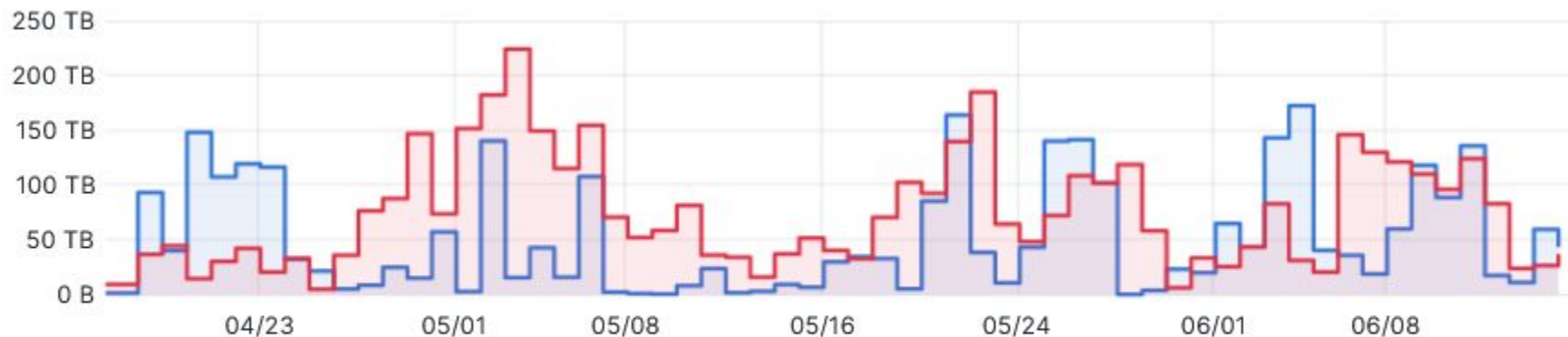
Number of threads per xrootd padme



# Stato tape

16 Apr - 15 Jun 2023

MSS bytes in/out (per day)



	min	max	avg	current	total
— out traffic (recalls)	43.8 GB	172 TB	50.8 TB	43.2 TB	3.10 PB
— in traffic (migrations)	4.79 TB	224 TB	73.0 TB	37.0 TB	4.45 PB

# Stato tape

- Liberi 23 PB (su cassette vuote, complessivamente sulle 2 librerie).  
Usati 100 PB.
  - Installato attuale: 130.5 PB
  - Gara da 53 PB in fase di assegnazione, per arrivare a 183.5 PB (pledge 2023)
  - Spazio quasi esaurito sulla IBM -> scritture di ATLAS e CMS sulla Oracle in attesa dei nuovi nastri

Library	Tape drives	Max data rate/drive, MB/s	Max slots	Max tape capacity, TB	Installed cartridges	Used space, PB	Free space, PB
SL8500 (Oracle)	16*T10KD	250	10000	8.4	~10000	54.3	22.9
TS4500 (IBM)	19*TS1160	400	6198	20	2450	45.4	0.2