

State of Storage

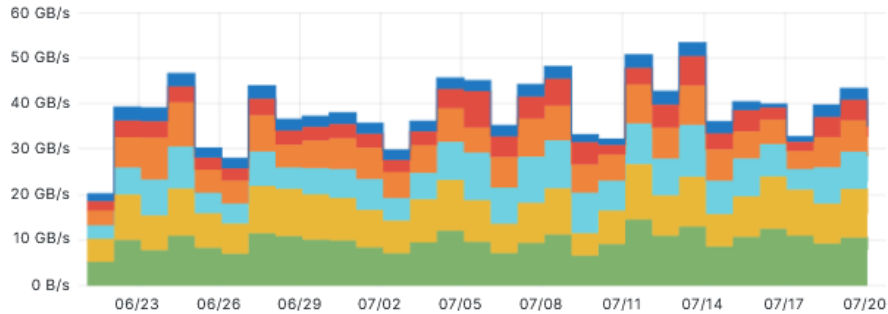
CdG 21 luglio, 2023



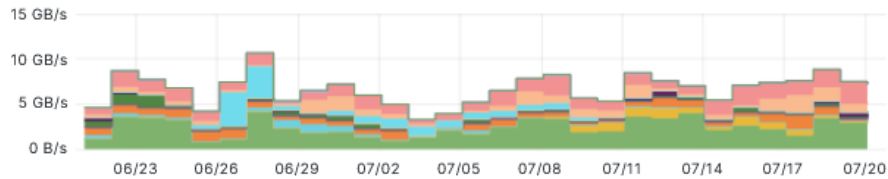
Business as usual

Last month

All servers network traffic out (reading)

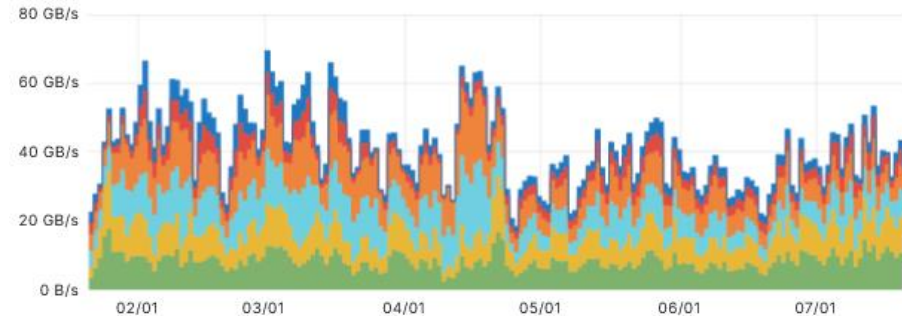


Gateway traffic out (non POSIX reading)

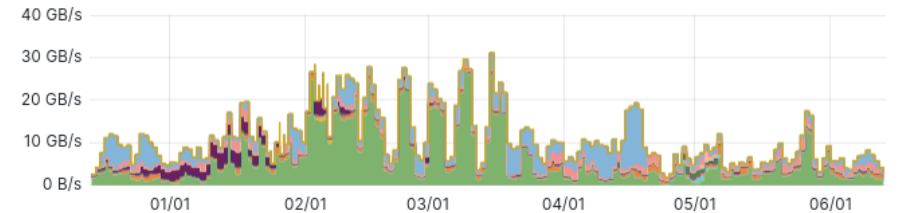


Last 6 months

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: **47.4** PB

Storage system	Model	Net capacity, TB	Experiment	End of support
ddn-10, ddn-11	DDN SFA12k	10120	ALICE, AMS	12/2022 (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
 - Collaudo non superato
 - Vista l'emergenza, tutti gli esperimenti sono invitati a cancellare i dati superflui
- **AQ storage 2023-2024**
 - 66PB nel 2023 + 14PB nel 2024
 - Il vincitore è Huawei con sistemi OceanStore Micro 1500/1600
 - Installazione direttamente al Tecnopolo
- **Gara Tape Library**
 - Apertura buste economiche oggi
- **Gare nastri**
 - 53 PB: pledge + overpledge 2023
 - consegna prevista per settembre
 - Ulteriori 50 PB di nastro per repack dei dati dalla libreria Oracle
 - da acquistare in autunno



Current SW in PROD

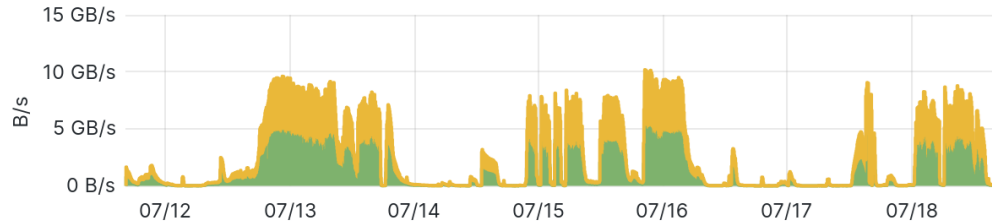
- GPFS 5.1.2-11 (updated recently)
- StoRM BackEnd 1.11.22 (updated recently, latest)
- StoRM FrontEnd 1.8.15 (latest)
- StoRM WebDAV 1.4.2 (updated recently, latest)
- StoRM globus gridftp 1.2.4
- XrootD 5.5.4-1
 - ALICE CEPH updated to 5.5.5-1.el8
- Ceph 16.2.6 (Pacific)

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
 - **6.3.5-1.el8.elrepo.x86_64**
- `mount -v -t ceph 131.154.128.248:6789:/alice-xrootd /storage/newcluster/alice-xrootd -o name=alice,recover_session=clean,secretfile=*****`
- Standard XrootD packages installation
- Puppet-based configuration

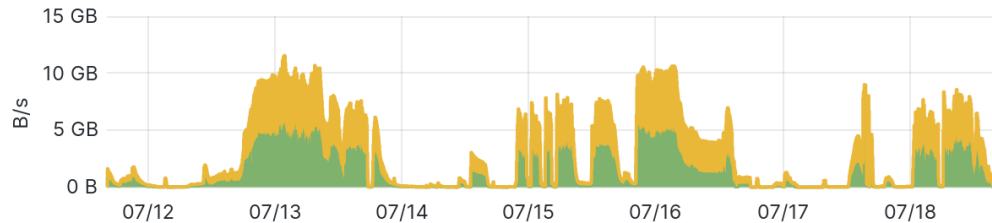
XrootD ALICE CEPH

Total network utilization (IN)



	min	max	avg	current
eth0 in	18.7 kB/s	5.25 GB/s	1.46 GB/s	19.0 MB/s
eth1 in	19.4 kB/s	4.94 GB/s	1.43 GB/s	19.1 MB/s

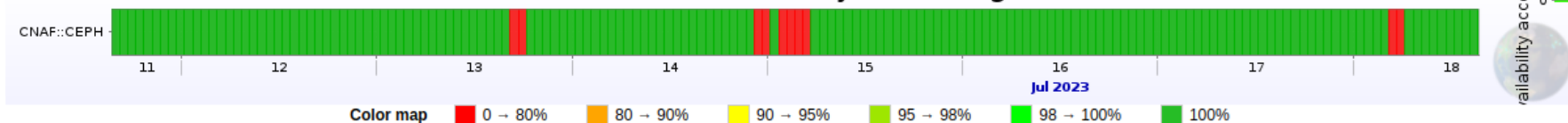
Total network utilization (OUT)



	min	max	avg	current
eth0 out	6.66 kB	5.64 GB	1.53 GB	20.0 MB
eth1 out	8.03 kB	5.93 GB	1.67 GB	11.1 MB

XrootD ALICE CEPH

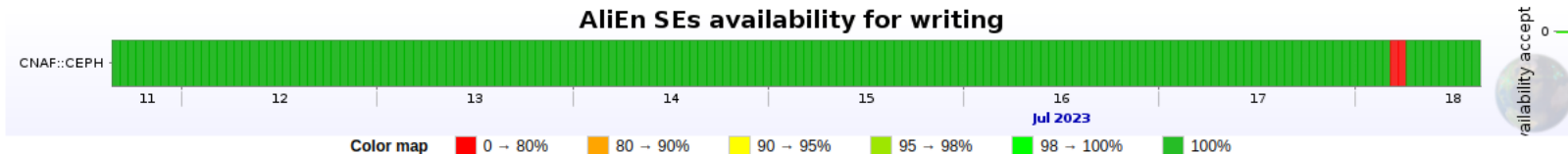
AliEn SEs availability for reading



Statistics

Link name	Data		Individual results of reading tests			Overall
	Starts	Ends	Successful	Failed	Success ratio	Availability
CNAF::CEPH	11 Jul 2023 15:06	18 Jul 2023 14:48	164	5	97.04%	97.07%

AliEn SEs availability for writing



Statistics

Link name	Data		Individual results of writing tests			Overall
	Starts	Ends	Successful	Failed	Success ratio	Availability
CNAF::CEPH	11 Jul 2023 15:06	18 Jul 2023 14:48	168	1	99.41%	99.43%

Tickets and problems

- ATLAS
 - 0% Transfer efficiency to tape (GGUS [162735](#))
 - Fixed over the weekend, triggering migrations
 - Token-support enabled for disk storage areas
 - The two disk storage areas (/atlas/atlasdatadisk, /atlas/atlasscratchdisk) have been unified into one (/atlas)
 - The access points for the tape storage area have been modified 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, under investigation

Tickets and problems

- CMS
 - Gridftp still used, only for SAM tests :-)
 - SSL connection failure for WebDAV endpoints (GGUS [162753](#))
 - Two similar tickets: GGUS [162488](#), GGUS [162436](#)
 - Thread saturation
 - Two StoRM WebDAV servers have been added in xfer-cms.cr.cnaf.infn.it (for a total of 5)
 - Problems with some files during staging (GGUS [162250](#))
 - /cmstape/store/data/%/FEVTDEBUGHLT/
 - Migrated on tape
 - Further investigation needed for the other files

Tickets and problems

- LHCb
 - Migration within filesystem (GGUS [162740](#))
 - Scheduled downtime on Wed 19/7 to rename folders in disk so to avoid srm calls to deduce physical paths
 - Jobs failed to get files from local storage, fixed configuring submission to CEs (GGUS [162386](#))
 - FTS transfer errors between CNAF and CERN-CTA (GGUS [162189](#))
 - Waiting for CERN to check their systems for jumbo frames support
- AMS02
 - The experiment has switched from srm+gsiftp to srm+https
- 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
 - GridFTP “plain” still used
 - XrootD server at IHEP to perform the transfers to CNAF
 - Waiting for a grid certificate
- DarkSide
 - Enabled StoRM WebDAV (srm+https); can we switch GridFTP off?
- Icarus
 - Enabled StoRM WebDAV (srm+https); can we switch GridFTP off?

Tickets and problems

- Virgo
 - Configured a new tape storage area /virgo-tape, parallel to /virgoD0T1
 - Deleted /Runs (a folder on tape)
 - Enabled StoRM WebDAV (srm+https); can we switch GridFTP off?
 - Call to discuss transfer inefficiency from Cascina
 - Waiting for feedback from the experiment, namely the precise time periods where the problem was spot, and tolerance to short network saturation periods

Stato tape

20 May - 18 Jul 2023



Stato tape

- Liberi 20 PB (su cassette vuote, complessivamente sulle 2 librerie).
Usati 103 PB.
 - Installato attuale: 130.5 PB
 - Spazio quasi esaurito sulla IBM -> maggior parte delle scritture 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	57.3	19.6
TS4500 (IBM)	19*TS1160	400	6198	20	2450	45.7	0.3