

State of Storage

CdG 17 Gennaio, 2025



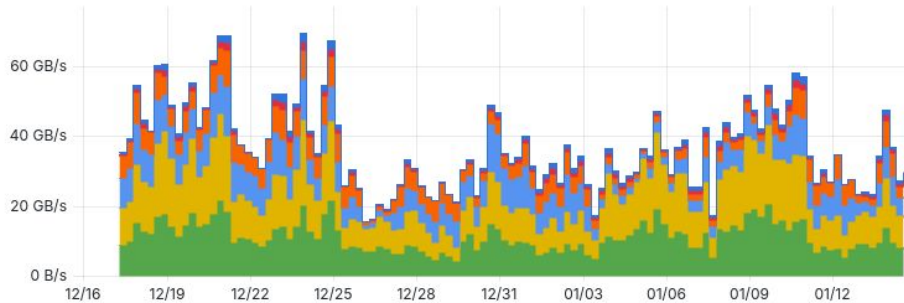
Business as usual + migration to TP



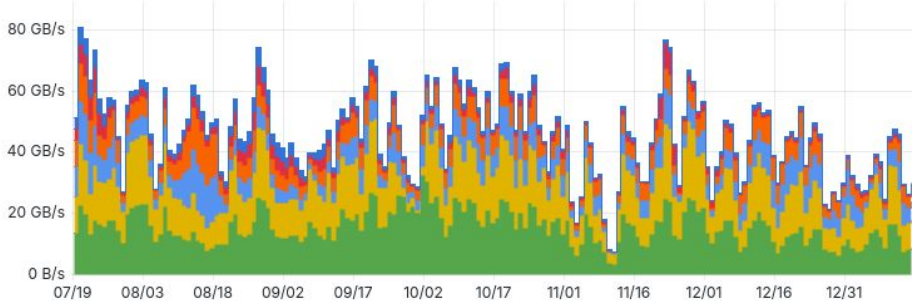
Last month

Last 6 months

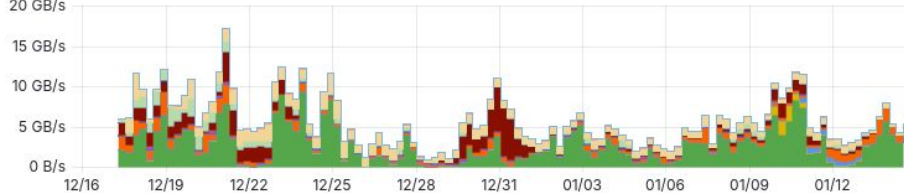
All servers network traffic out (reading)



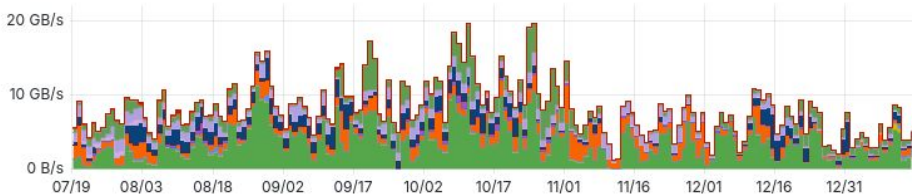
All servers network traffic out (reading)



Gateway traffic out (non POSIX reading)



Gateway traffic out (non POSIX reading)



Disk storage in produzione 77.5PB

Installed: **101.4PB**; Pledge 2025/01: **86.95PB**; Pledge 2025/04: **101PB**; Used: **60PB**

Storage system	Model	Net capacity, TB	Experiment	End of support
os5k8-1,os5k8-2	Huawei OS5800v5	8999	GR2	2027
ddn-12, ddn-13	DDN SFA 7990	5840	GR2,GR3 (commissioning)	2025
ddn-14, ddn-15	DDN SFA 2000NV (NVMe)	24	Metadata (commissioning)	2025
ddn-16	DDN SFA 2000NVX2 (NVMe)	96	LHCb metadata, hotdata	2031
ddn-17,ddn-18,ddn-19	DDN SFA 7990X	14000	LHCb	2031
od1k6-1,2,3,4,5,6	Huawei OD1600	60000(-10%!)	ALICE,ATLAS,CMS, GR2	2031
od1k6-7,8	Huawei OD1600	18000	Buffer for remake of above	2031
od1k5-1,2	Huawei OD1500 (NVMe)	400	Metadati, LHCb hotadata	2031

Disk storage in dismissione: 32.2PB

Storage system	Model	Net capacity, TB	Experiment	End of support
os6k8	Huawei OS6800v3	3400	Virgo	07/2024
ddn-10, ddn-11	DDN SFA12k	10120		12/2022
md-1,md-2,md-3,md-4	Dell MD3860f	2308	Virgo, Archive	12/2024
md-5,md6,md-7	Dell MD3820f	50	metadati, home, SW	11/2023 e 12/2024
os18k1, os18k2	Huawei OS18000v5	6520	AMS	7/2024
os18k3, os18k5, os18k5	Huawei OS18000v5	9780	ATLAS,ALICE (buffer tape)	6/2024

Traslochi recenti e futuri

- Trasloco della libreria IBM: completato
- Trasloco della libreria Oracle: in corso
- Gare nastri conclusa
 - 96PB di tape JF (50TB/cassetta) in arrivo



Problemi LHCb

- Abbiamo creato “buffer disco” per i dati “hot” con un fileset dedicato:

LHCb-Buffer Linked/storage/gpfs_lhcb/disk/lhcb/buffer

- Abbiamo definito la Placement policy:

```
RULE 'HOTDISK' SET POOL 'hotdata' FOR FILESET ('LHCb-Buffer') LIMIT(99) WHERE NAME LIKE '%.dst'
```

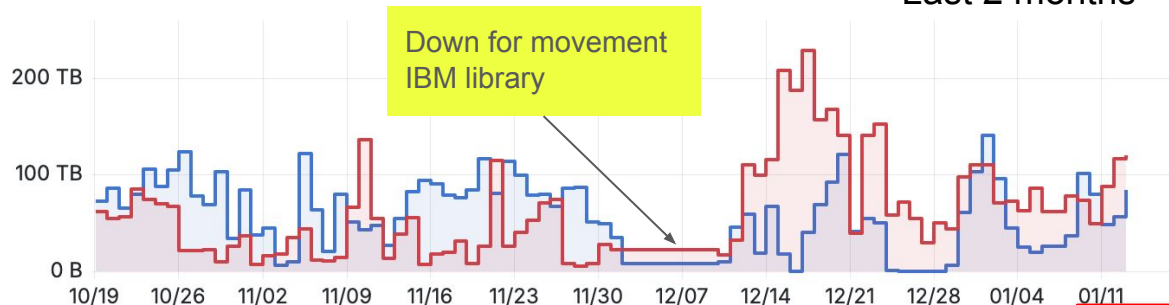
- 200 TB NVMe

- Il buffer tape è stato rimesso sull'HW condiviso con il resto dei dati.
- Dopo la migrazione sullo storage DDN, abbiamo riscontrato problemi con schede della rete dei nuovi server Lenovo

Stato tape

MSS bytes in/out (per day)

Last 2 months



Name	Mean	Last *	Max	Min	Total
— out traffic (recalls)	60.5 TB	84.2 TB	141 TB	13.2 GB	4.78 PB
— in traffic (migrations)	64.0 TB	120 TB	228 TB	5.25 TB	4.99 PB

5 PB of new data written to tapes in two month (since last CdG)

Tape system DOWN 13-23 Jan 2025
(per trasloco della libreria Oracle a Tecnopolo)



Stato tape

- Liberi ~40 PB (Scratch tape inclusi libreria Oracle).
- Usati 149 PB.
- Spazio tape sulla libreria IBM praticamente esaurito
- La nuova libreria IBM non è ancora funzionante a causa di problemi di compatibilità con la versione TSM in produzione.

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	47	32
TS4500 (IBM)	19*TS1160	400	6198	20	5100+380	102	0.6
TS4500-2(IBM)	18*TS1170	400	7844	50	165	0	8.2

Current SW in PROD

- GPFS 5.1.9-4, in migrazione to 5.1.9-7
- StoRM BackEnd 1.11.22 (latest, 1.12.0 in test)
- StoRM FrontEnd 1.8.15 (latest, 1.8.16 in test)
- StoRM WebDAV 1.4.3 (latest)
- StoRM Tape 0.8.0 (latest)
- StoRM globus gridftp 1.2.4
- XrootD 5.5.4-1
 - LHCb updated to 5.5.5-1
- Ceph 16.2.6 (Pacific)
- GEMSS and tape drive orchestrator updated to support X tape libraries

Tickets and more

- ALICE
 - Open action: finalize the configuration for the XrootD tape cluster (xs-204, xs-304)
 - Waiting for the migration of servers to EL9 to install and test rpm for interaction xrootd-tape
- ATLAS
 - GGUS [169504](#) (closed): Lukas Pfaffenbichler overloaded StoRM WebDAV servers
 - GGUS [169404](#), [169472](#) (closed): StoRM WebDAV threads saturation; a restart solved the issue
 - GGUS [169311](#) (closed): tape downtime
 - GGUS [169259](#) (closed): report.json no longer available on the “xfer” endpoint
 - GGUS [169250](#) (closed): the file system was left unmounted after a reboot
 - GGUS [168988](#) (closed): tape library downtime, the recovery of the services was a bit slower than expected
 - GGUS [168445](#) (closed): failed transfers due to “tape buffer full”
 - *We highly recommend not to exceed the **mean daily** writing rate limit (recalls included) of 1.0GB/s*
 - GGUS [167957](#) (on hold): StoRM WebDAV does not permit the creation of non-existent parent directory even if the scope does it,
 - Waiting for a fix from StoRM developers

Tickets and more

- CMS
 - GGUS [169589](#) (closed): tape downtime, still ongoing
 - GGUS [169260](#), [169319](#) (closed): tape downtime (on Dec); but CMS used xfer-cms (disk endpoint) for SAM tests
 - GGUS [168610](#) (closed): StoRM WebDAV threads saturation
 - GGUS [168538](#) (closed): StoRM Tape returned NONE instead of DISK for 0-size files
 - Fixed by StoRM developers
 - GGUS [167995](#) (on hold): StoRM WebDAV does not permit the creation of non-existent parent directory even if the scope does it
 - Waiting for a fix from StoRM developers

Tickets and more

- LHCb
 - GGUS [169478](#) (in progress): 19/12/24 - Massive transfer failures to/from CNAF
 - 20/12/24 - 40GB/s POSIX traffic, migration away from HUAWEI disk storage
 - Increased the number of StoRM WebDAV servers up to 8
 - Concurrent FTS/job transfers reduced by LHCb
 - 27/12/24 - Internal network issues with the new DDN disk storage
 - StoRM Tape REST API showed 502 errors (connection timeout)
 - 07/01/25 - TPCs failed with timeout expiration
 - concurrent FTS transfers were not enough to transfer all the files before that the garbage collector removed them from the tape buffer
 - 13/01/25 - Migration to DDN storage completed
 - GGUS [169555](#) (in progress): corrupted files
 - Network issues resulted in 1696 corrupted files, for a total size of 461 GB
 - For comparison, we checked 2312 files (461 GB) written between 12:40 and 13:20 of Jan 12th, and we could not find any evidence of corruption

Tickets and more

- Gsiftp protocol via StoRM backend is still available for two experiments
 - New StoRM release or new StoRM Tape REST API release should finally allow to switch GridFTP off (Xenon, CTA-LST)
- Dampe
 - GridFTP “plain” no more used
 - Rucio+FTS (https) replaced the current gsiftp transfers (WP6-DataCloud)