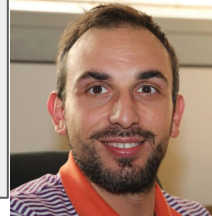
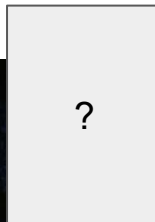
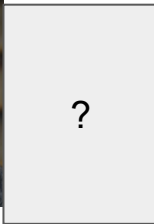


State of Storage

CdG 19 luglio, 2024



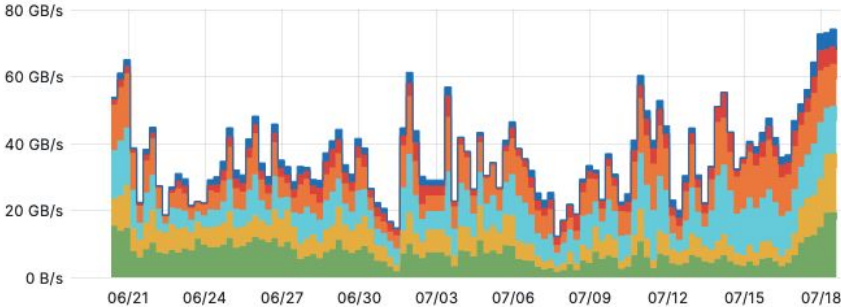
Business as usual + migration to TP



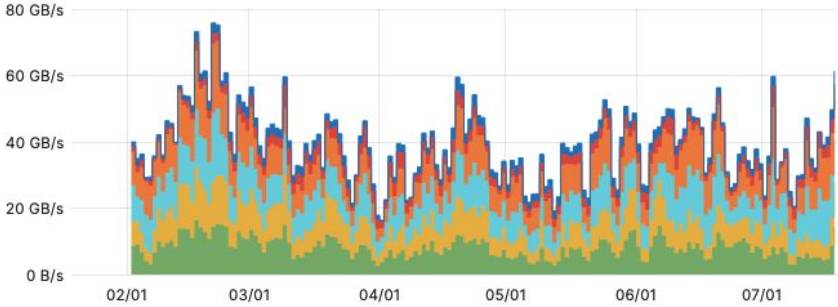
Last month

Last 6 months

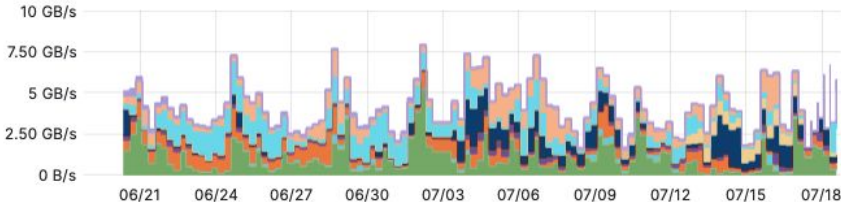
All servers network traffic out (reading)



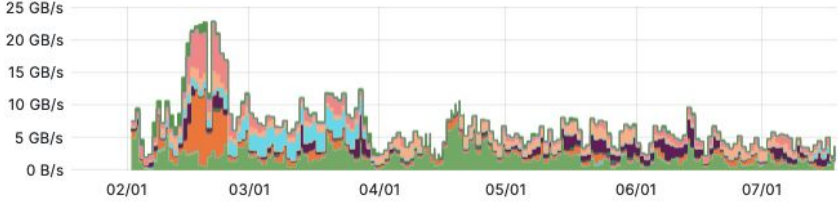
All servers network traffic out (reading)



Gateway traffic out (non POSIX reading)



Gateway traffic out (non POSIX reading)



Disk storage in produzione

Installed: 113PB - 33PB (in dismissione)=80.6PB Pledge 2024: 82.08PB, Used: 48.8PB

Storage system	Model	Net capacity, TB	Experiment	End of support
ddn-10, ddn-11	DDN SFA12k	10120	ALICE, AMS	12/2022
os6k8	Huawei OS6800v3	3400	GR2, Virgo	07/2024
md-1,md-2,md-3,md-4	Dell MD3860f	2308	DS, Virgo, Archive	12/2024
md-5, md-6 e md-7	Dell MD3820f	50	metadati, home, SW	11/2023 e 12/2024
os18k1, os18k2	Huawei OS18000v5	7800	LHCb	7/2024
os18k3, os18k5, os18k5	Huawei OS18000v5	11700	CMS	6/2024
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
od1k6-1,2,3,4,5,6	Huawei OD1600	60000	ALICE,ATLAS,LHCb, CMS	2031

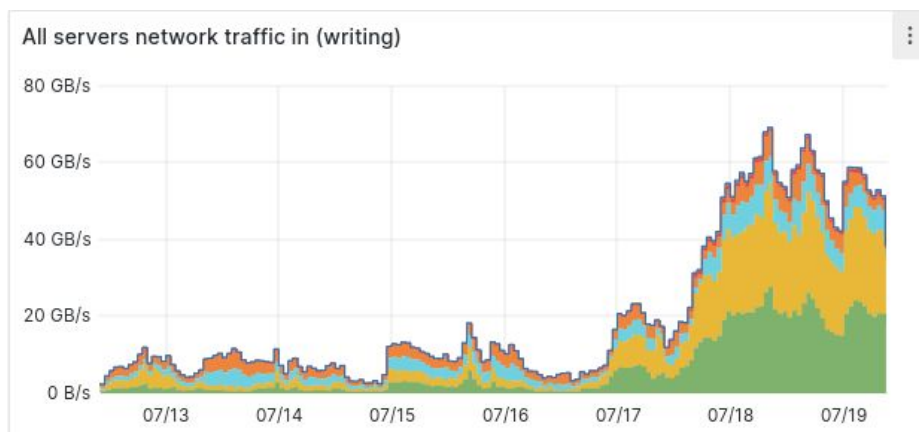
Acquisti recenti e futuri

- Gara storage 2022 (14PB netti)
 - Nuova proposta con apparati DDN SFA7990X
 - In attesa per la consegna entro ~~giugno~~ settembre
- AQ storage 2023-2024
 - Huawei OceanStore Micro 1500/1600
 - 8 sistemi di 10PB + 40 server
 - Installato e collaudato 60PB +32 server
- Tape Library
 - Installata, collaudo in corso
 - Le cassette JF da 50TB stati messi nella libreria (7.8PB)
- Gare nastri
 - Nuova gara di acquisto tape JF in preparazione



Migrazione dati sul nuovo storage a TP

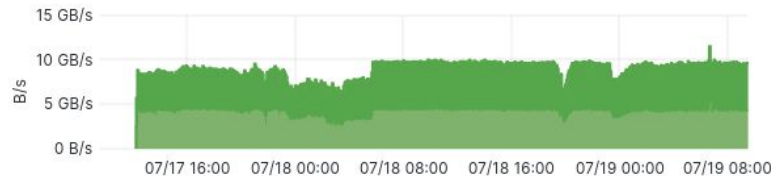
- Con il trasloco a TP dobbiamo spostare in totale ~50PB di dati;
- Pledge complessiva 2024 è 82.08PB;
- Pledge 2024 per 4 exp LHC (56.2PB) va a coprire quasi tutto il nuovo storage appena installato (60PB);
- Dal pomeriggio del 17/07 abbiamo cominciato migrare dati di ALICE, ATLAS, LHCb.
- Average rate di migrazione ~10GB/s per filesystem
- L'attività di produzione ha un impatto significativo sulle migrazioni dei dati.



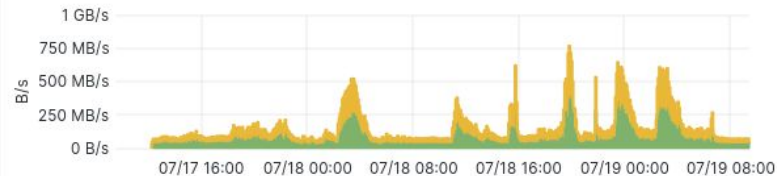
Migration vs. prod (Good, Bad and Ugly)

ALICE

Total network utilization (IN)

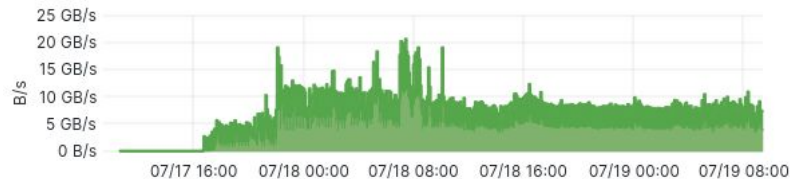


Total network utilization (OUT)

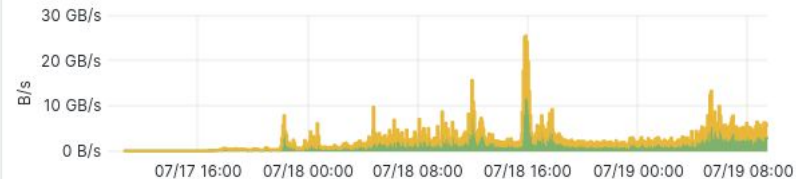


ATLAS

Total network utilization (IN)

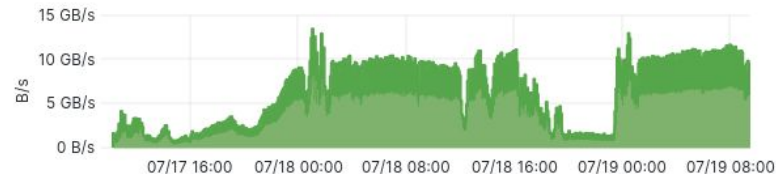


Total network utilization (OUT)

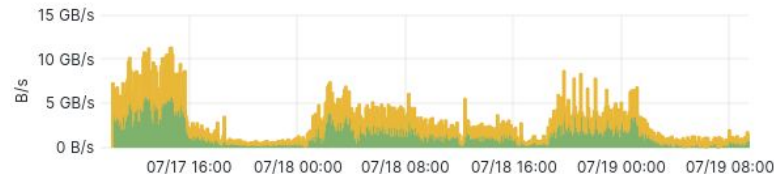


LHCb

Total network utilization (IN)



Total network utilization (OUT)



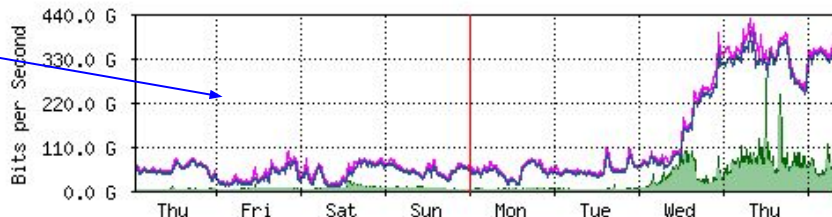
Migration to TP (few more plots)

Use of network link T1-TP

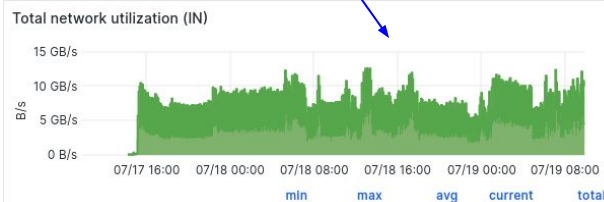
Overall rate of data migration

Migration (partial) of gpfs_data

'Weekly' Graph (30 Minute Average)



Max In: 278.4 Gb/s (34.8%) Average In: 17.6 Gb/s (2.2%) Current In: 57.9 Gb/s (7.2%)
 Max Out: 424.0 Gb/s (53.0%) Average Out: 102.1 Gb/s (12.8%) Current Out: 331.6 Gb/s (41.5%)



Current SW in PROD

- GPFS 5.1.2-13 (in preparazione migrazione alla 5.1.9-4 che supporta RHEL9)
- StoRM BackEnd 1.11.22 (latest)
- StoRM FrontEnd 1.8.15 (latest)
- StoRM WebDAV 1.4.2 (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)

Tickets and more

- ALICE
 - Finishing configuration restyling of XrootD GPFS cluster:
 - Manage configuration files with Puppet
 - Upgrade to latest version in production (5.5.4-1.el7)
 - Check on the status of the service included within sensu framework
 - Finalizing the configuration with the tape cluster (xs-204, xs-304)
 - F. Noferini has a working “tsm” RPM building procedure for EL9
 - Waiting for the migration of servers to EL9 to install and test it
 - Waiting for input from experiment on how migrate data from CephFS to GPFS
 - Rsync bulk data transfer
 - Grid transfer via an xrootd endpoint at Tecnopolo

Tickets and more

- ATLAS
 - GGUS [166882](#) (closed) and duplicate GGUS [166887](#)(closed)“INFN-T1 has staging errors”
 - 30/05: concurrent activities on two tape SA filled the disk buffer again
 - We asked again to reduce the writing rate to 1.2GB/s
 - 17/06: 645 stage requests in only one minute failed with 502 error
 - 22k file stages failed
 - Nginx closes the connection to the StoRM Tape REST unit when there is a high load on the server and a consequent timeout error occurs
 - We are setting proper Nginx parameters to avoid this behaviour
 - 19/06: Atlas reported 56k file stages failed with 403 error in the last 2 days
 - Permission denied due to trailing slash
 - Same problem reported at LHCb GGUS [164634](#)
 - 20/06: changing StoRM tape REST configuration files to accept ending slash in the path solved the issue

Tickets and more

- CMS
 - GGUS [167634](#) (open): CMS WebDAV SSL connection test fails on one server in xfer-cms
 - Issue do to overload of data transfer servers and thread limit reached.
 - Servers have been reconfigured to provide a better thread management
 - Upon the experiment request, HTTPS protocol enabled for site xrootd redirectors and servers
 - However, CMS (Bockjoo) seemed surprised that davs is actually enabled for CMS in totally different endpoints (xfer-cms)
 - Waiting for CMS input on whether (and how) to enable shoveler for xrootd monitoring
- LHCb
 - GGUS [167586](#) (solved): failed data transfers due to overload of the LHCb cluster at CNAF caused by POSIX access and FTS transfers; asked to lower submission rate of FTS jobs
 - GGUS [167045](#) (in progress): 4 servers handle both http requests, acting as StoRM WebDAV servers, and requests to the underlying file system, acting as GPFS NSD servers
 - The new servers already installed at Tecnopolo should ease the situation, separating NSD and StoRM WebDAV servers

Tickets and more

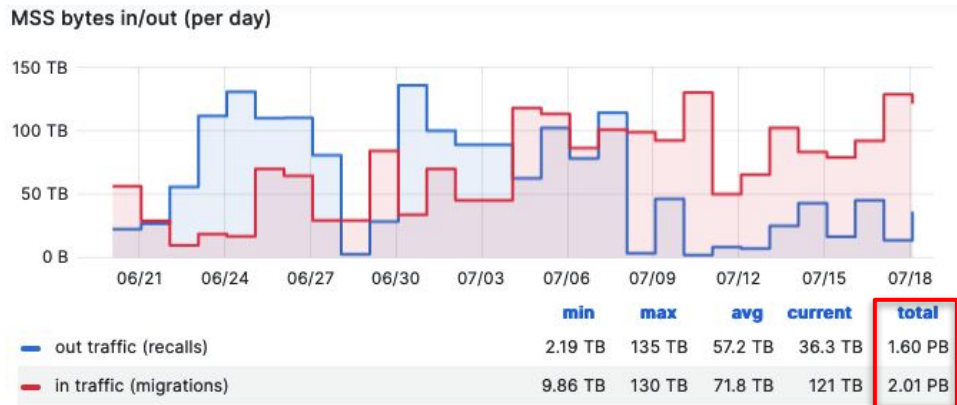
- Gsiftp protocol via StoRM backend is still available for two experiments
 - New StoRM release should allow to switch GridFTP off (Xenon, CTA-LST)
- Dampe
 - GridFTP “plain” still used
 - TPCs between XrootD server at IHEP and CNAF are working well
 - Rucio+FTS (https) should replace the current gsiftp transfers (WP6-DataCloud)
- DUNE
 - Request to expose data in read mode also via XrootD
 - 4 XrootD servers shared among other experiments
 - Waiting for the configuration files from DUNE community

Tickets and more

- HyperK
 - GGUS [167154](#) (solved): random errors contacting the StoRM WebDAV SA `davs://xfer-archive.cr.cnaf.infn.it:8443/hyperk`
 - One of the servers in the “xfer-archive” alias had been banned by `ngs.ac.uk` following a false security alarm raised one month ago from CERN (<https://www.virustotal.com/gui/ip-address/131.154.128.183/community>)
- PADME
 - E. Leonardi contacted us with an urgent request to stage-in PADME 2022 raw data due to storage problems in Frascati
 - We suggested to avoid transferring data to disk, due to space shortage in `gpfs_data`, and instead transfer data from buffer to Frascati. No reply.

Stato tape

Last month



2 PB of new data written to tapes in one month (since last CdG)
+1.6PB migrated from SL8500 (see next slide)

Tapes: Migration from Oracle to IBM library

Repack - number of free cartridges

Repack - Library Scratch Tape



Repack - Library Space Occupancy



Still have 34PB to migrate

Stato tape

- Liberi ~4.6+7 PB (Scratch tape sulla libreria IBM).
- Usati ~133 PB.

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	34	-
TS4500 (IBM)	19*TS1160	400	>8000	20	5104+380	93	4.6+7
TS4500-2(IBM)	18*TS1170	400	>8000	50	150	0	7.5