

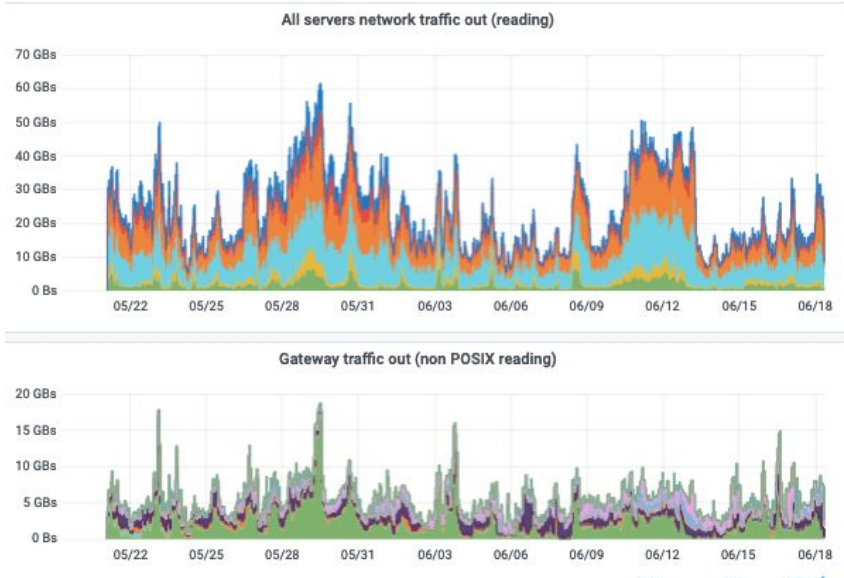
State of Storage

CdG 18th of June, 2021

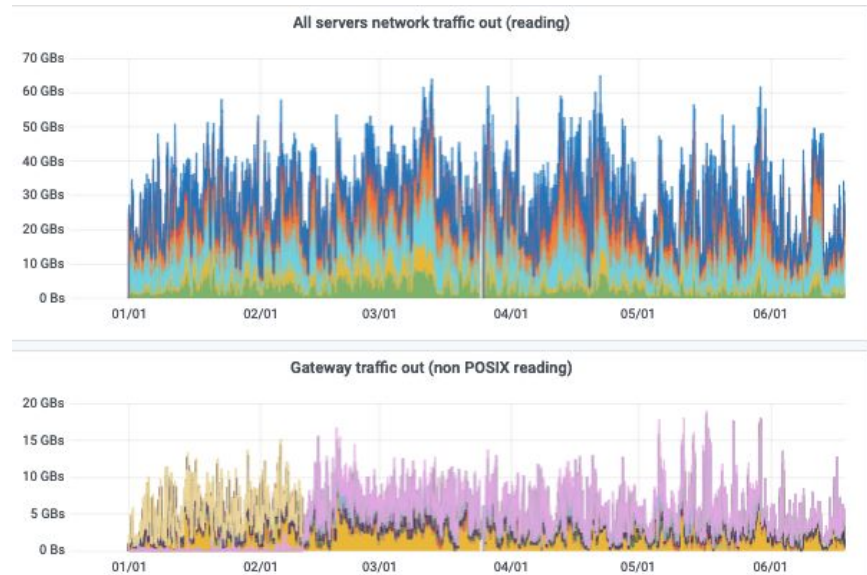


Business as usual

Last month



Last 6 months



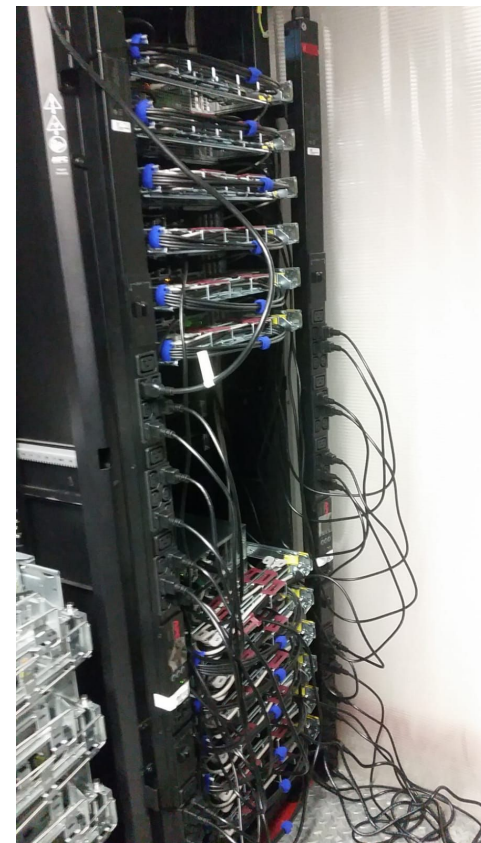
Disk storage in produzione

Installed: 41.07 PB Pledge 2020: 45 PB Used: **36.1 PB**

Sistema	modello	Capacita', TB	esperimenti	scadenza
ddn-10, ddn-11	DDN SFA12k	10752	Atlas, Alice, AMS	03/2021→ 06/2023
os6k8	Huawei OS6800v3	3400	ALICE, GR2	2022
md-1,md-2,md-3,md-4	Dell MD3860f	2308	DS, Virgo, Archive	11/2021
md-7	Dell MD3820f	20	Metadati, home, SW	04/2021
md-5, md-6	Dell MD3820f	8	metadati	06/2021
os18k1, os18k2	Huawei OS18000v5	7800	LHCb, ALICE	2023
os18k3, os18k5, os18k5	Huawei OS18000v5	11700	ATLAS, CMS	2024
ddn-12, ddn-13	DDN SFA 7990	5060	GR2,GR3	2025
ddn-14, ddn-15	DDN SFA 2000NV	24	metadati	2025

Stato installazioni storage

- Stato consegnato Lotto CEPH
- Mancano ancora le PDU 2U
- Consumi elettrici sottostimati
 - Avra' bisogno 4 PDU invece di 2



Stato installazione storage (cont.)

Lotto GPFS:

Mancano ancora PDU, switch IB e cablaggio 100Gbit

Problemi comuni:

impossibilità di montare PDU 0U nel rack con apparati di alta densità



Current SW in PROD

- GPFS 5.0.5-2 → being updated to 5.0.5-7 (to fix security vulnerability)
- StoRM BackEnd 1.11.21 (latest)
- StoRM FrontEnd 1.8.15 (latest)
- StoRM WebDAV 1.4.0 (latest)
- StoRM globus gridftp 1.2.4
- XrootD 4.11.2

Transition gridftp-> StoRM WebDAV

- Atlas uses http for stage-out (since 11th May). 3 StoRM WebDAV servers in `xfer.cr.cnaf.infn.it`; 2 of them are also gridftp, which in principle is needed only for tape access (see however recent problems)
- 2 StoRM WebDAV servers for LHCb (alias `xfer-lhcb.cr.cnaf.infn.it`); 2 different servers are gridftp + xrootd
- 3 StoRM WebDAV servers in `xfer-cms.cr.cnaf.infn.it`; 2 of them are also gridftp

FTS Archiving feature

successfully tested by CMS (with a single file):

File ID	File State	File Size	Throughput	Remaining	Start Time	Finish Time	Staging Start	Staging End	Archiving Start	Archiving End
+ 3033924093	ARCHIVING	231.89 MiB	17.05 MB/s	-	2021-06-03T17:06:42Z	2021-06-03T17:07:01Z			2021-06-03T17:07:55Z	

📁 srm://ccsrm.in2p3.fr:8443/srm/managerv2?SFN=/pnfs/in2p3.fr/data/cms/disk/data/store/mc/RunIIFall17NanoAODv5/ttPhiPS_M-40_2mu_13TeV-madgraph/NANOADSIM/PU2017_12Apr2018_Nano1June2019_102X_mc2017_realistic_v7-v1/100000/8483D1C3-CFC9-C444-AF3A-8CE505E20689.root

📄 srm://storm-fe-cms.cr.cnaf.infn.it:8444/srm/managerv2?SFN=/cmstape/store/mc/RunIIFall17NanoAODv5/ttPhiPS_M-40_2mu_13TeV-madgraph/NANOADSIM/PU2017_12Apr2018_Nano1June2019_102X_mc2017_realistic_v7-v1/100000/8483D1C3-CFC9-C444-AF3A-8CE505E20689.root

And then 40 minutes later in the finished state (on tape):

File ID	File State	File Size	Throughput	Remaining	Start Time	Finish Time	Staging Start	Staging End	Archiving Start	Archiving End
+ 3033924093	FINISHED	231.89 MiB	17.05 MB/s	-	2021-06-03T17:06:42Z	2021-06-03T17:07:01Z			2021-06-03T17:07:55Z	2021-06-03T17:45:04Z

📁 srm://ccsrm.in2p3.fr:8443/srm/managerv2?SFN=/pnfs/in2p3.fr/data/cms/disk/data/store/mc/RunIIFall17NanoAODv5/ttPhiPS_M-40_2mu_13TeV-madgraph/NANOADSIM/PU2017_12Apr2018_Nano1June2019_102X_mc2017_realistic_v7-v1/100000/8483D1C3-CFC9-C444-AF3A-8CE505E20689.root

📄 srm://storm-fe-cms.cr.cnaf.infn.it:8444/srm/managerv2?SFN=/cmstape/store/mc/RunIIFall17NanoAODv5/ttPhiPS_M-40_2mu_13TeV-madgraph/NANOADSIM/PU2017_12Apr2018_Nano1June2019_102X_mc2017_realistic_v7-v1/100000/8483D1C3-CFC9-C444-AF3A-8CE505E20689.root

Recent problems

- ATLAS
 - We still observe high load on gridftp servers, with 1.7 GB/s write on gpfs_atlas (disk, not tape); problem with rucio policies to be fixed in Rucio 1.26 in June.
 - SRM+https tape tests moved from DTEAM storage area to ATLAS storage area: /atlas/atlasmctape/srmhttpstests (new RSE in CRIC)
- BELLE
 - File upload failures; increased pool account size, created a SENSU check alarming when pool accounts get full (GGUS ticket [152066](#))
- CMS
 - CNAF to Florida Transfers failing in TPC push; fixed StoRM WebDAV configuration (GGUS ticket [152077](#))
 - Transfers failing having reached quota; CMS fixed Rucio quota (GGUS ticket [152274](#))
 - xrootd-access tests failing at T1_IT_CNAF; server certificate renewed but most likely the problem was caused by wn-ovirt (GGUS ticket [152318](#))
- LHCb
 - Failed transfers to CNAF-buffer from Nikhef and Manchester after scheduled downtime 16th June; tracepath shows GI is used instead of OPN; ongoing investigation

Stato tape

15 April - 15 June 2021

MSS bytes in/out (per day)



	min	max	avg	current	total
— out traffic (recalls)	1.9 TB	155.0 TB	46.4 TB	35.5 TB	2.8294 PB
— in traffic (migrations)	1.4 TB	77.4 TB	28.1 TB	26.2 TB	1.7158 PB

Stato tape

- 7.5 PB liberi (complessivamente sulle 2 librerie). Usati 88.5 PB.
 - Gran parte delle scritture su nuova libreria
 - Tutti LHC
 - Xenon, CTA, Virgo, ARGO, Juno, Icarus
 - Pledge 100.7 PB. Da acquistare poco meno di 5 PB.

Library	Tape drives	Max data rate/drive, MB/s	Max slots	Max tape capacity, TB	Installed cartridges	Used capacity, PB
SL8500 (Oracle)	16*T10KD	250	10000	8.4	~10000	78.5
TS4500 (IBM)	19*TS1160	400	6198	20	750	10

Test di repack

- In vista di una possibile dismissione della libreria Oracle
 - 250 TB, 67MB avg per file
 - A un drive Oracle che legge corrisponde un drive IBM che scrive sulla nuova libreria
 - Partito ieri test con file di dimensioni maggiori (670 MB)
- Assegnazione statica dei drive
 - Da adeguare orchestratore per ottimizzare utilizzo drive in attività di repack

	6 drive	8 drive	10 drive	12 drive
Days needed for 80 PB	820	619	521	487
Total rate (MB/s)	1128	1495	1775	1900
Avg rate per drive (MB/s)	188	187	178	158

Piano tape a medio termine

- Analizzate le roadmap tecnologiche e le previsioni di spesa
 - Libreria Oracle verrà mantenuta almeno per i prossimi 3-4 anni
 - Da traslocare al Tecnopolo
 - Repack da programmare con prossime tecnologie tape
 - Da iniziare non prima di inizio 2023
- Visti i rate e pattern di scrittura/lettura previsti, riusciremo a far fronte alle richieste almeno per tutto il 2022 con il sistema attuale
 - Non si prevede acquisto di tape drive
 - Sarebbe utile avere dati indicativi di rate e pattern previsti per ATLAS
 - Per il futuro (oltre 2022) è utile sapere il tipo di letture previste
 - Dati presenti su libreria Oracle (scritti prima di settembre 2020) oppure dati nuovi?
 - Per pianificare meglio acquisti/manutenzioni di tape drive