

State of Storage

CdG 22-Nov-2019

Disk storage in manutenzione

Pledge 2019: 38.721 PB

Used: 35.44 PB

Sistema	modello	Capacita', TB	esperimenti	scadenza
ddn-10, ddn-11	DDN SFA12k	10752	Atlas, Alice, AMS	2021
os6k8	Huawei OS6800v3	3400	GR2, etc	2022
md-1,md-2,md-3,md-4	Dell MD3860f	2308	DS, Virgo, Archive	05/2020
md-7	Dell MD3820f	20	Metadati, home, SW	05/2020
md-5, md-6	Dell MD3820f	8	metadati	2020-2022
os18k1, os18k2	Huawei OS18000v5	7800	LHCb, CMS	2023
os18k3, os18k5, os18k5	Huawei OS18000v5	11700	LHCb, ALICE	2024

Ceph (Test) instance @CNAF

Test instance installation on the way

- HW: 4 server + 4 JBOD x 60 HDD 8TB = 1.9PB raw
 - Erasure coding 3 + 1 (1.44 PB usable)
 - Replicated pool with 2 copies (0.96 PB usable)
- Network (server side): 2x10Gbps x server => 10 GB/s theoretical limit (~10 MB/s/TB) but part of this BW will be used for replication traffic
- Ongoing tests simulating random behaviour
 - Currently limited number of client (4*num threads per client)
 - Footprint ~ Tier1 jobs (2/4GB per thread)

New data management services

- Two dedicated hosts (xfer.cr.cnaf.infn.it) with latest version of StoRM WebDAV for the DOMA TPC testbed. Currently re-installing them CentOS7 and testing StoRM WebDAV installation via puppet.
 - The ESCAPE VO is also configured on the xfer.cr.cnaf.infn.it nodes (using proxy generated by IAM).
- Xcache: configured on the same host for both Virgo (stash cache) and CMS (XCacheOnDemand).
 - Collecting monitoring metrics for our InfluxDB and sending some of them to DODAS server in Perugia every 5 minutes
- Apache: several experiments configured in the same host ds-814.cr.cnaf.infn.it: ASFIN, FAZIA, NEWSDM, NTOF. They share the same IAM instance, group-based authorization with tokens
 - Configuration with browsability to be introduced soon

Migrazione a CentOS 7

- StoRM: puppet will be used to install and configure the services on CentOS 7
 - StoRM WebDAV: currently ongoing (xfer-2.cr.cnaf.infn.it already migrated)
 - Storm GridFTP: December 2019; currently working with the developers to understand minor issues
 - StoRM frontend and backend: spring 2020
- XrootD
 - XCache currently works on CentOS7
 - We could start in December 2019. Any concern?

Nuova tape library

- IBM TS4500
 - 9 frame, 6200 slot per una capacità potenziale di 120 PB
 - 100 cartridges da 20TB (totale 2PB)
 - 19 tape drives TS1160 - data rate 400MB/s
 - Collaudo completato
- Test di funzionalità via TSM
 - Effettuati su server versione 8.1.8 - ultima disponibile e compatibile con drive TS1160
 - HSM server aggiornati alla 8.1.8 (test effettuati anche su CentOS7)
 - Nuova versione GEMSS compatibile con GPFS versione 4.2

Aggiornamento tape server

- Per poter utilizzare i nuovi tape drive di IBM occorre aggiornare
 - TSM e HSM server
 - Per l'occasione aggiorneremo i server HSM a CentOS7
- Upgrade HSM archive (gruppo II + Virgo + CDF)
 - Aggiornamento hardware, OS e software (HSM e GEMSS)
 - Da programmare un down (parziale) di 8 ore. Mercoledì 27/11?
- Upgrade del server TSM (software) e degli altri HSM (OS e software)
 - Dopo verifica corretto funzionamento HSM archive
 - Da programmare un down (parziale) di 12 ore per tutti. Potrebbe essere mercoledì 4/12?
- Entro 2019 - in modo da poter utilizzare la nuova libreria da inizio 2020

VIRGO

- Problemi:
 - Unico FS sia per dati che per il buffer tape
 - Non c'è controllo sull'utilizzo dello spazio (no filesets, no quotas)
 - Utilizzo del FS come la home (46M di files)
 - Necessario separare buffer/disk?
 - Per archiviare circa 3TB al giorno, un buffer di 25-30TB sarebbe più che sufficiente
 - Per la modifica sarebbe necessario un giorno di down
 - La modifica richiede cambiamento dei path su gpfs_virgo4 per accesso POSIX
 - Non per le storage area accedute via storm
 - La modifica dei path è problematica secondo Luca Rei prima della fine di O3b (primavera)
 - Ne discuteremo a Cascina next week

Non-LHC Experiments

- PADME over quota (expired)
- XENON dir structure in archive - example of how NOT to do

/storage/gpfs_archive/xenon/rucio	(elements in one level)	(total)
--run-id	27667	27667
-- XX	20 - 140	~2.5M
-- YY	2 - 3	~5M
-- file	1	~10M

- ATLAS example (twice as much, 20 times faster)

/storage/gpfs_atlas/atlas/atlasdatadisk/rucio		
--run-id	121	121
-- XX	256	~30K
-- YY	120-180	~5M
-- file	1-4	~20M