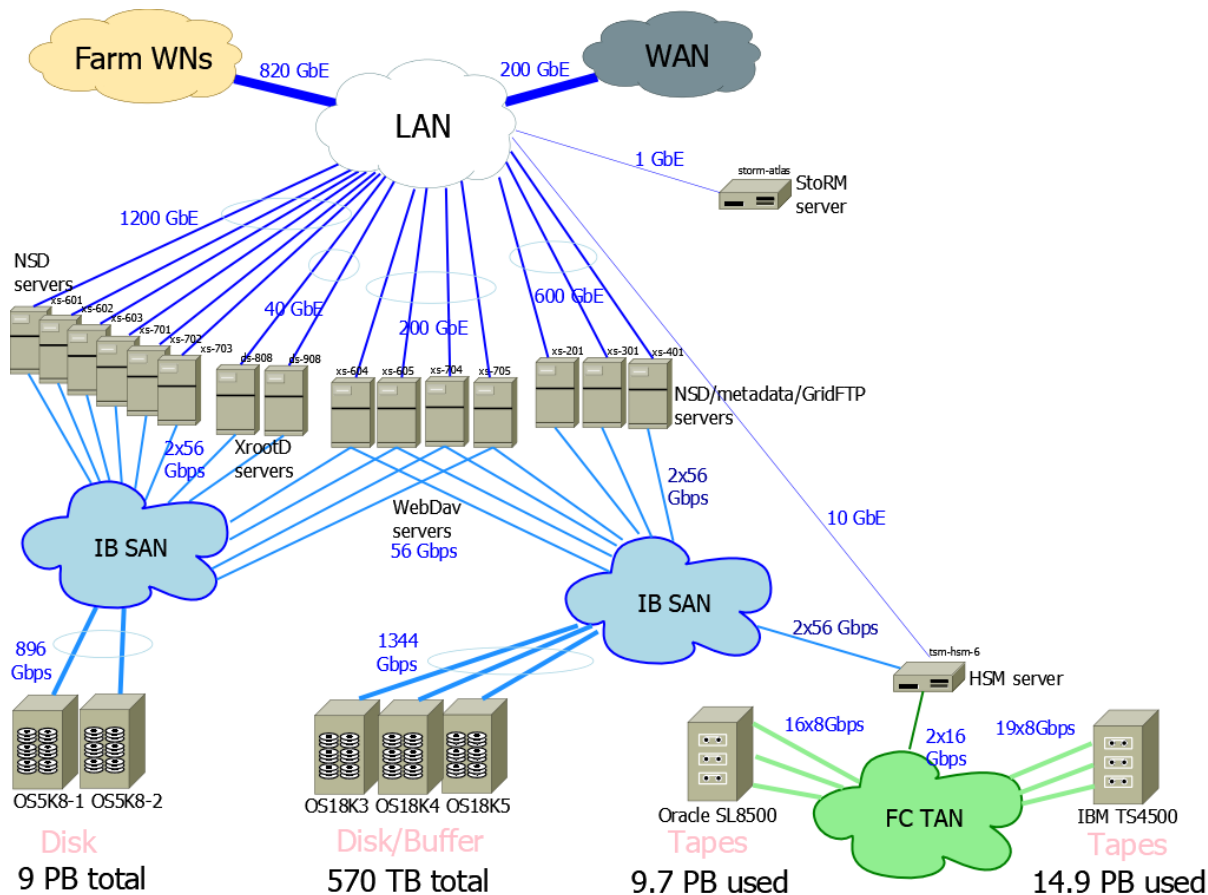


Tape system del CNAF

Enrico Fattibene

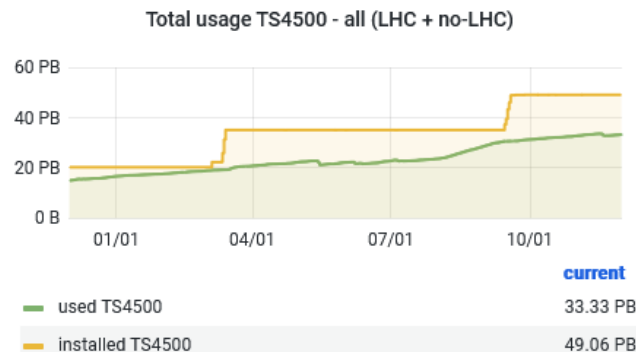
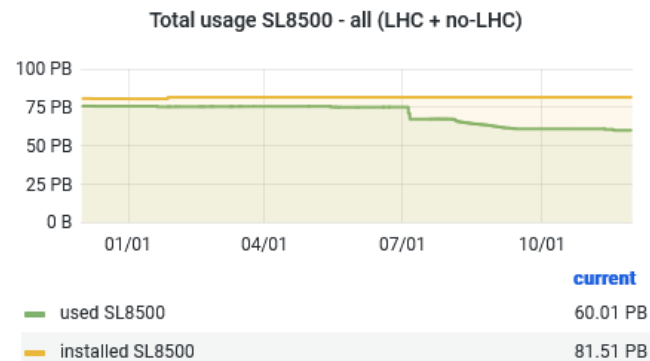
Meeting ATLAS Italia Computing
30 novembre 2022

Risorse ATLAS al CNAF



Risorse tape

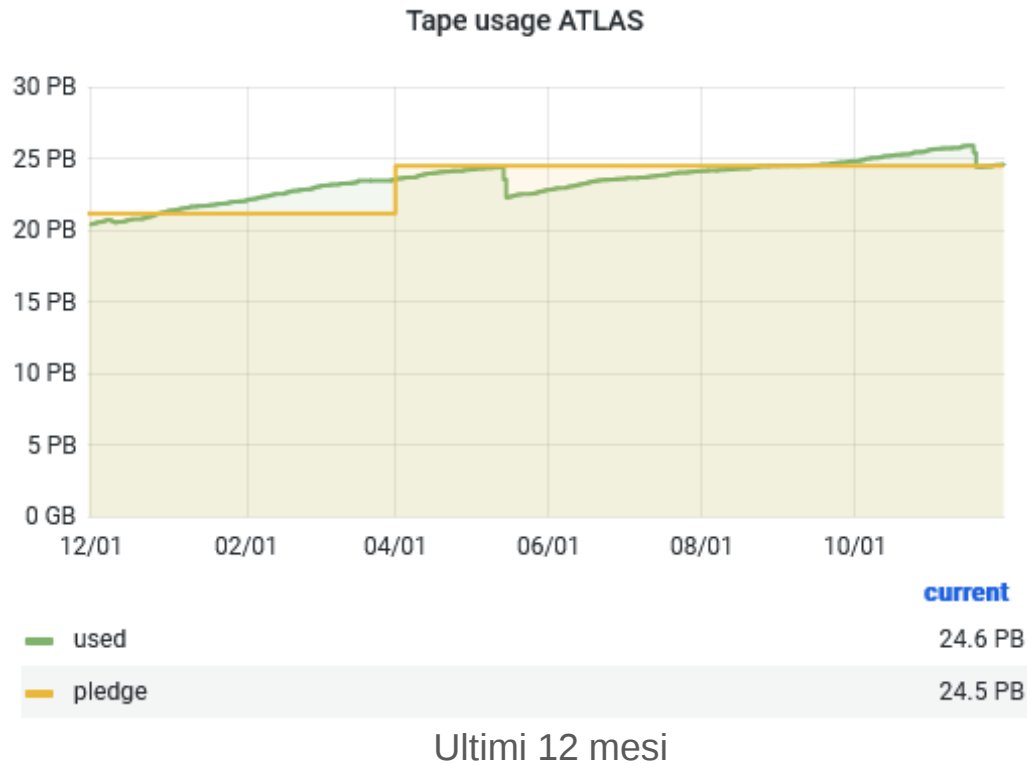
- **130.5 PB di spazio installato**, che corrisponde al pledge 2022 (**93 PB usati**)
 - **1 libreria Oracle-StorageTek SL8500**
 - 16 tape drive T10000D, condivisi tra tutti gli esperimenti
 - Cassette da 8.5 TB
 - Rate di lettura/scrittura: 250 MB/s per drive (4 GB/s complessivi)
 - Sola lettura
 - **1 libreria IBM TS4500**
 - 19 tape drive TS1160, condivisi tra tutti gli esperimenti
 - Cassette da 20 TB
 - Rate di lettura/scrittura: 400 MB/s per drive (7.6 GB/s complessivi)



Ultimi 12 mesi

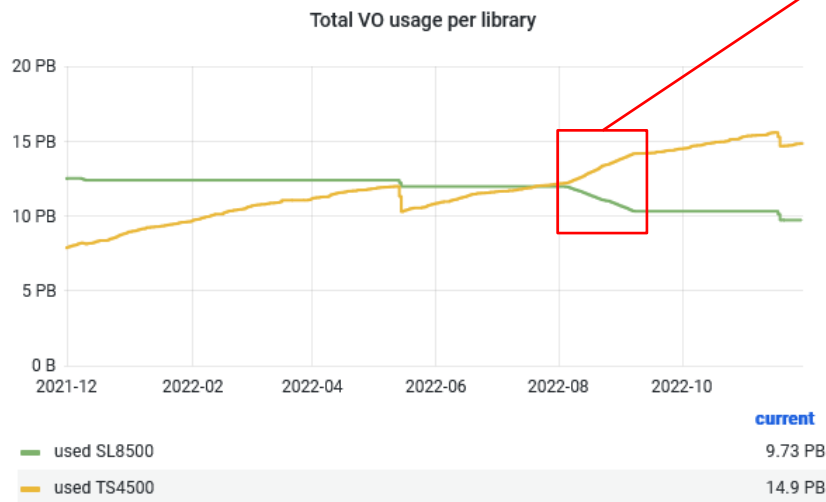
Usato tape ATLAS

- **Usato = pledge 2022**
- 2 campagne di cancellazioni nel 2022
 - 2 PB a maggio
 - 1.5 PB ad ottobre
- Richiesta di 683 TB overpledge accettata



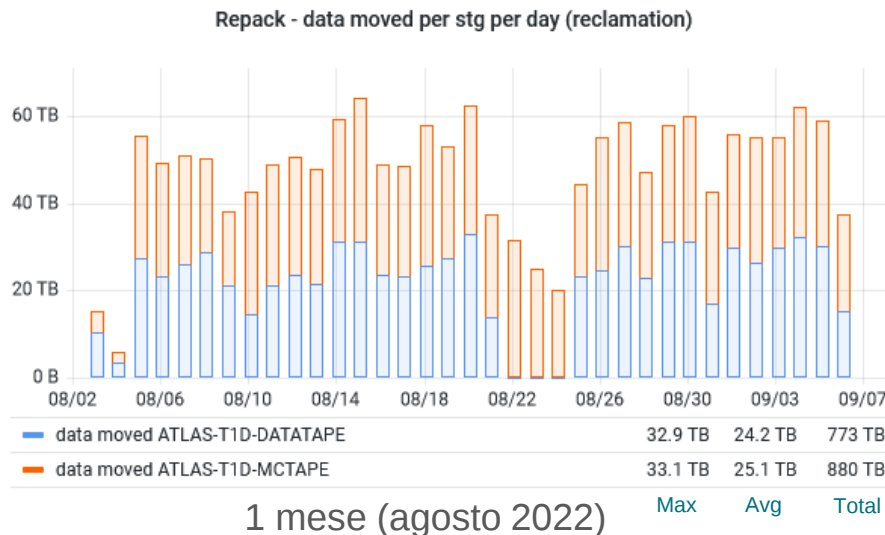
Usato tape per libreria

- Scritture su Oracle ferme da primavera 2020



Ultimi 12 mesi

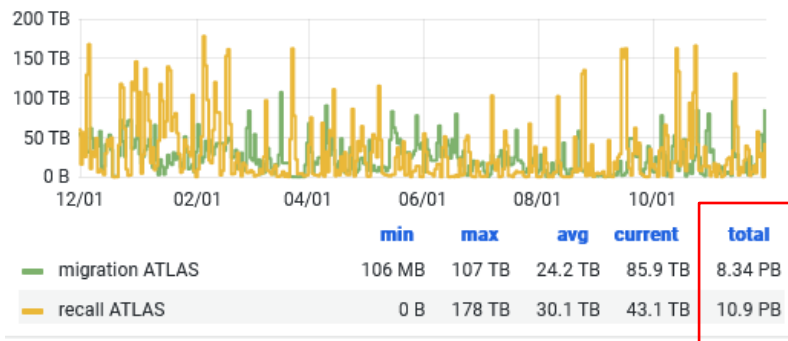
- Repack agosto 2022
 - 1.65 PB spostati da Oracle a IBM



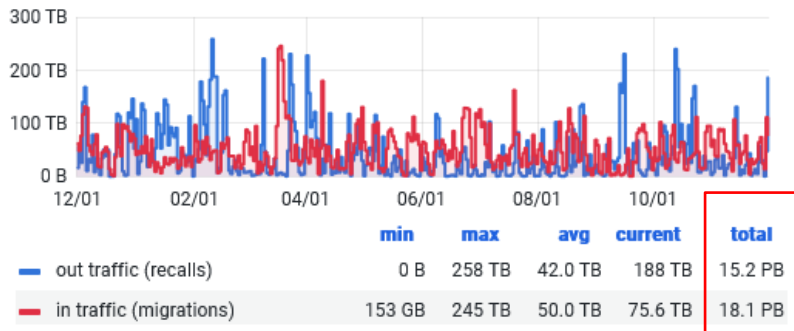
Scritture / letture tape ultimo anno

Quantità di dati

ATLAS bytes in/out per day

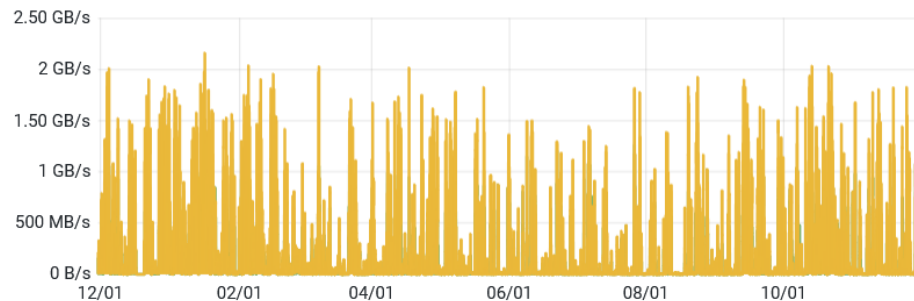


All experiments bytes in/out per day

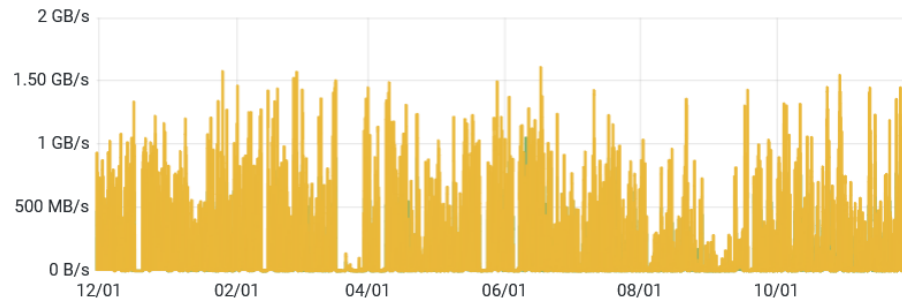


Rate

ATLAS tape read rate



ATLAS tape write rate



Media valori ogni ora

Tape challenge marzo 2022: risultati

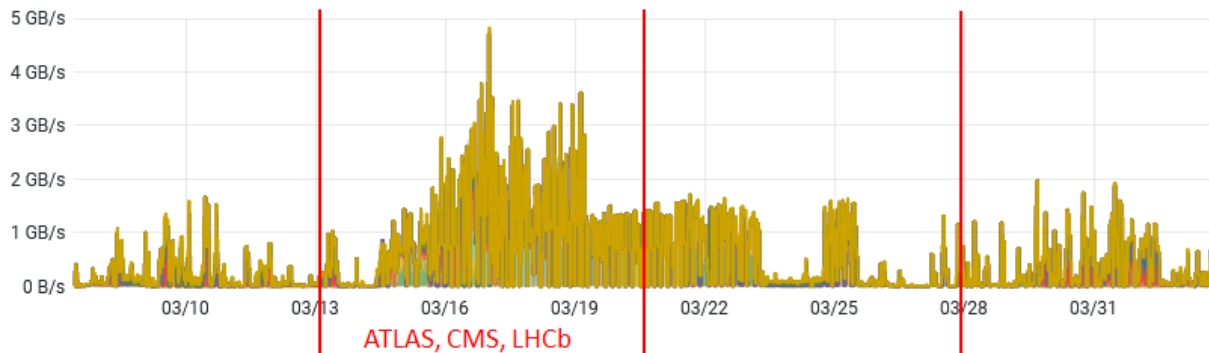
	Writes (DT)			Reads (A-DT)		
VO	Expected GB/s	Result avg GB/s	Target achieved?	Expected GB/s	Result avg GB/s	Target achieved?
ATLAS	0.9	1	YES	0.8	1.9	YES
CMS	0.37	1.1	YES	NA	1.46	YES
LHCb	1.72	1.52/1.72*	~YES*	1.35	1.8	YES

* Migration rate reaches the target if buffer has enough data to migrate

- Target raggiunti

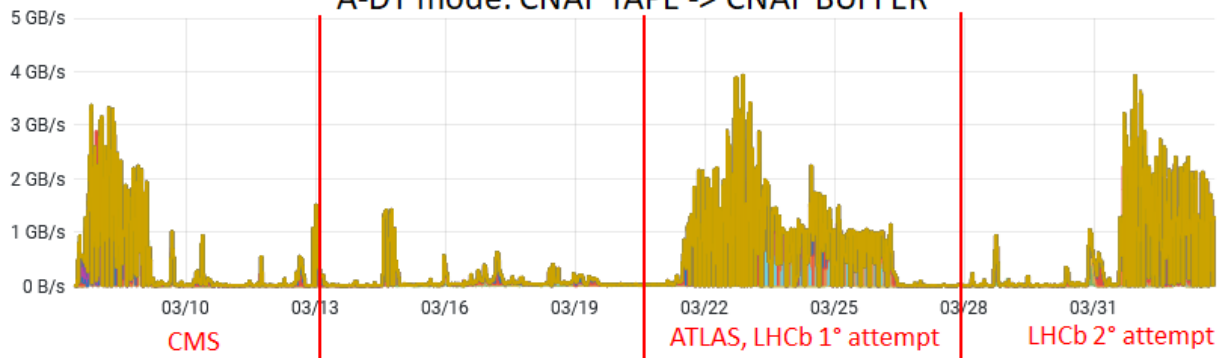
Tape challenge: rate buffer - tape

DT mode: CNAF BUFFER -> CNAF TAPE



- Media ogni 30 minuti
- Picco: **4.9 GB/s**

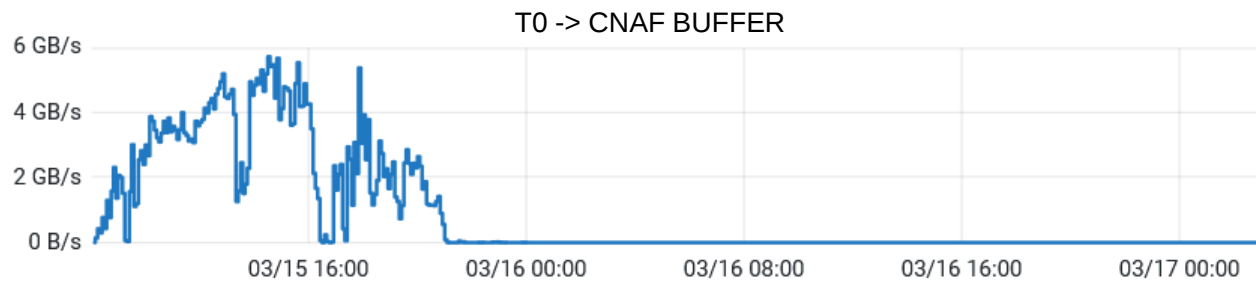
A-DT mode: CNAF TAPE -> CNAF BUFFER



- Media ogni 30 minuti
- Picco: **4 GB/s**

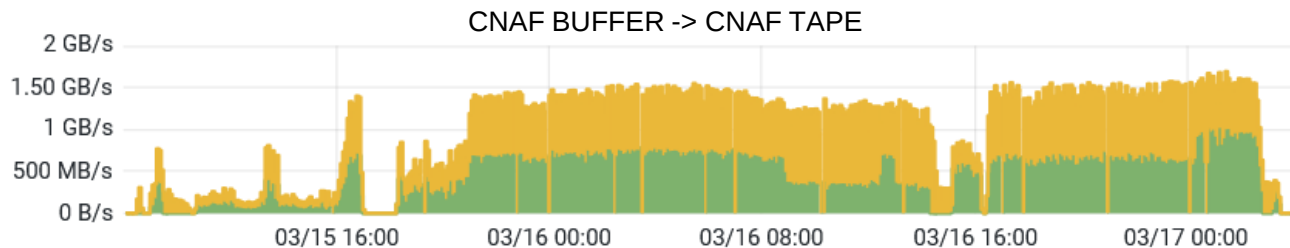
Tape challenge: ATLAS DT

- Target avg rate T0 -> CNAF buffer: 3.5 GB/s per 12 ore
- Target avg rate T0 -> CNAF tape: **0.9 GB/s**
- Performance limitata del buffer causa problemi di concorrenza tra scritture e letture



T0 -> CNAF buffer

- Rate medio: 3 GB/s
- Picco: > 4GB/s (oltre 2 ore)



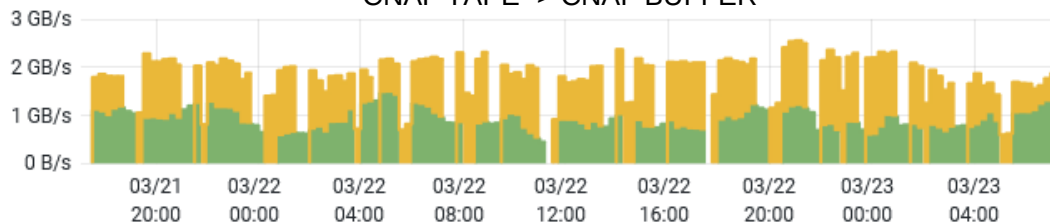
CNAF buffer -> CNAF tape

- Rate medio complessivo: **1 GB/s**
- Rate medio dalla fine della scrittura su buffer: 1.35 GB/s
- Picco: 1.6 GB/s (1 ora)

Tape challenge: ATLAS A-DT

- Target avg rate CNAF Tape -> CNAF Disk: **0.8 GB/s**

CNAF TAPE -> CNAF BUFFER



CNAF buffer -> CNAF tape

- Rate medio: **1.9 GB/s**
- Picco: 2.5 GB/s (1 ora)

CNAF BUFFER -> CNAF DISK



CNAF buffer -> CNAF disk

- Rate medio: **1.9 GB/s**
- Picco: 3 GB/s

Tape challenge: letture

- Rate medi per drive
 - **ATLAS: 211 MB/s (85% rate nominale)**
 - CMS: 146 MB/s (59% rate nominale)
 - LHCb: 360 MB/s (90% rate nominale)
- Una certa inefficienza è dovuta al carico sui server causato dallo scan del FS

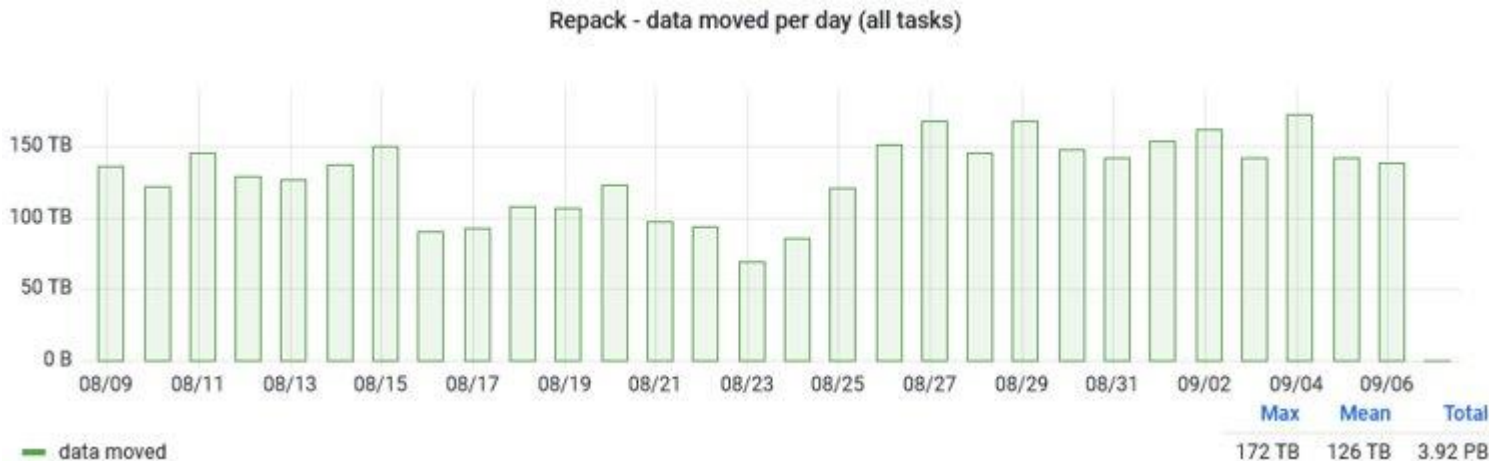
	ATLAS	CMS	LHCb
Total files	73902	10761	61378
Total tapes	52	164	18
Total GiB	157943	164397	349895
Total GB	169590	176519,9	375696,9
Avg GiB per file	2,1	15,3	5,7
Avg GB per file	2,3	16,4	6,1
Avg files per tape	1421	66	3410
Avg GiB per tape	3037	1002	19439
Avg GB per tape	3261	1076	20872

Considerazioni generali su tape al CNAF

- Libreria Oracle da dismettere il prima possibile
 - Manutenzione sempre più costosa (+ 15% all'anno)
 - Tecnologia sta diventando obsoleta
- Nuova tecnologia IBM disponibile nel 2023
 - Tape drive TS1170: rate di 500 MB/s (attuali TS1160: 400 MB/s)
 - Capacità delle cassette: 50 TB ognuna (attuali: 20 TB)
 - No retro-compatibilità
 - Possibile calo del prezzo/TB
- Spazio limitato nell'attuale data center (Berti Pichat)
 - 6200 slot licenziati sulla libreria IBM
 - 2000 utilizzabili previo acquisto licenza
 - Possibile espansione con 1 frame (1000 slot)

Repack

- Da spostare **60 PB**
- Rate repack agosto 2022: 4 PB / mese
- Orchestratore drive: task di repack partono solo in caso di drive liberi
 - Viene data priorità alle scritture e letture
- Possibilità di aumento rate del 15-20% con seconda CPU al TSM server
- Tempo stimato repack di 60 PB: minimo 14 mesi
 - Considerando pledge 2023/24, lo spazio fisico non è sufficiente affinché il repack sia completato nell'attuale data center su tecnologia IBM TS1160



Scenari a medio termine

- Possibile acquisto nuova libreria nel 2023
 - Compatibile con tecnologia IBM TS1160
 - 2 possibili soluzioni
 - Libreria IBM TS4500 (come quella attualmente al CNAF) con supporto a drive IBM e LTO
 - Libreria SpectraLogic Tfinity con supporto a drive IBM, LTO e Oracle
- Repack da Oracle a IBM senza spostamento della libreria Oracle al Tecnopolo
 - Entro il 2024
 - 2 scenari principali

Scenario 1

- Acquisizione nuova libreria compatibile IBM da installare al Tecnopolo
 - Con relativa infrastruttura (compreso nuovo TSM server)
- **Repack da Oracle a libreria IBM attuale in sede Berti Pichat**
 - 1 esperimento per volta
 - Una volta finito il repack dei dati di un esperimento:
 - Spostamento fisico delle cassette nella nuova libreria al Tecnopolo
 - Migrazione dati su nuovo TSM server al Tecnopolo
 - Migrazione/installazione HSM server dell'esperimento al Tecnopolo
 - Migrazione del buffer al Tecnopolo
 - Scritture/letture dei dati dell'esperimento su nuova libreria al Tecnopolo
 - Down: per scritture qualche ora, per letture pochi giorni
- Una volta finito il repack di tutti i dati Oracle:
 - Spostamento libreria IBM attuale al Tecnopolo
 - Dismissione libreria, drive e cassette Oracle

Scenario 2

- Acquisizione libreria SpectraLogic TFinity da installare al Tecnopolo
- Spostamento libreria IBM attuale e di tutta infrastruttura tape al Tecnopolo
- Installazione tape drive e cassette Oracle nella nuova libreria al Tecnopolo
- Dismissione libreria Oracle
- **Repack da Oracle a IBM nella nuova libreria al Tecnopolo**
- Una volta finito il repack:
 - Dismissione drive e cassette Oracle
- Down: per scritture e letture 1 settimana