Tape in WLCG and NDGF

NeIC NT1 Manager Mattias Wadenstein <maswan@ndgf.org>

> 20211201 NDGF All Hands NBI, Denmark



Overview

- Current use
- Future plans
- NDGF implications

SPEAKER | Mattias Wadenstein <maswan@ndgf.org>



Current use

- 2PB ALICE and 12PB ATLAS data at NDGF now
 - -Will increase faster than disk for the next few years, at least for ATLAS
- We read about 15% of all stored data on tape in a typical year
 - So average bandwidth is not that high, 50MB/s
- Current guidance on tape rates is about 2+GB/s aggregate effective rate
 - Especially reads will spend quite some time seeking too, which is not effective
 - Load spread over our sites





Tape Chalenge

 The bandwidths required for NDGF for reads and writes during data taking (DT) and right after data taking(A-DT), in Run3

VO	Reads (DT) GB/s	Writes (DT) GB/s	Reads (A-DT) GB/s	Writes (A-DT) GB/s
ALICE	0	0.3	0.1	0.3
ATLAS	0.1	0.5	0.5	0.3
CMS	_	_	_	_
LHCb	_	_	_	_
Total	0.1	0.8	0.6	0.6

SPEAKER | Mattias Wadenstein <maswan@ndgf.org>

4

Tape Challenge

 For ALICE, NDGF met their requirements in writing rate: 0.3GB/s, during the tape challenge, Link. -Note ALICE only tested tape write in both DT and A-DT modes. 700 MB/s 675 MB/s

σ

a

a

논







Tape Challenge

For ATLAS

- DT mode was from 10am on 10/11/2021 to 10pm 10/12/2021.
- -A-DT mode was from 10am on 10/13/2021 to 17pm 10/15/2021.
- The tape write rate for ATLAS to NDGF as seen from ATLAS (buffer \rightarrow tape rates might differ)



Tape Challenge

Below is the rate (seen by Rucio) of staging out of NDGF over the tape challenge week. It also met the target rates.

Transfer Throughput



Future LHC needs

- HL-LHC data taking from CERN
 - -140-280 Gbit/s
 - More than half of that directly to tape
 - Tens of PB per year

SPEAKER | Mattias Wadenstein <maswan@ndgf.org>





Current limitations

- dCache has a limited reordering window
 - Limited by the available space on the read pool due to reserving space for all outstanding requests to the tape system
- The dCache team has a potential solution:
 - Feed a central component all the tape locations
 - It will then issue restores for whole or partial tapes to the pools - Currently not deployable with HA dCache though
- Some sites underperforming
 - Remember, if all things align well, we should be able to do 1GByte/s to and from each tape instance





NDGF Implications

- Performance shortfalls will become hard production issues
- We might have to up the ambition level in a few years
 - At least keep up with tape drive advances
 - Big rate increase with HL-LHC in 2027 (\sim 40PB) \rightarrow 2030 (\sim 200 PB)

SPEAKER | Mattias Wadenstein <maswan@ndgf.org>







