

# ATLAS Tape drives @ HL-LHC

ATLAS - September 2020

Many thanks to ATLAS colleagues

# Inputs

- Assumptions :
  - RAW reprocessing do not overlap with data taking period (RAW export)
  - HI should not require more bandwidth
  - Very rough assumption for extrapolation in 9 years : Will evolve with Data Carroussel experience
- RAW + AOD/DAOD (prompt processing) export CERN → T1s (Table 8 of CDR)
  - RAW : 53 GB/s (during stable beam)
  - AOD/DAOD : 10 GB/s
- RAW staging at T0+T1s for reprocessing
  - T0 TAPE will be used but considered as safety margin
  - 300 PB in 100 days : 35 GB/s
- Write data AOD (35 PB) at T1s (output of RAW reprocessing)
  - Scenario 1 (higher TAPE load) :
    - 2 copies → 70 PB in 100 days : 8 GB/s
  - Scenario 2 (lower TAPE load) :
    - 1 copy → 35 PB in 100 days : 4 GB/s

# Inputs : First simulation campaign

- HITS (after G4) produced over year
  - 1 copy on TAPE
  - $(50 \text{ B evts fullsim} + 150 \text{ B evts fastsim}) * 1 \text{ MB/evt} = 200 \text{ PB} \rightarrow 6 \text{ GB/s}$
- Write MC AOD (200 PB produced spread over 1 year) at T1s :
  - 100 % on TAPE : 6 GB/s

# Inputs : MC reco and derivation with existing input

- **Reprocess G4 HITS + derivation MC + derivation AOD**

- Scenario 1 :

- Process 100 % HITS in 100 days : Staging : 23 GB/s (No staging of MC AOD)
- Write 50%/50% MC AOD on TAPE/DISK : Write 12 GB/s (MC).
- Derivation 100 % data AOD in 100 days (35 PB with 50% on TAPE) : Staging 2 GB/s

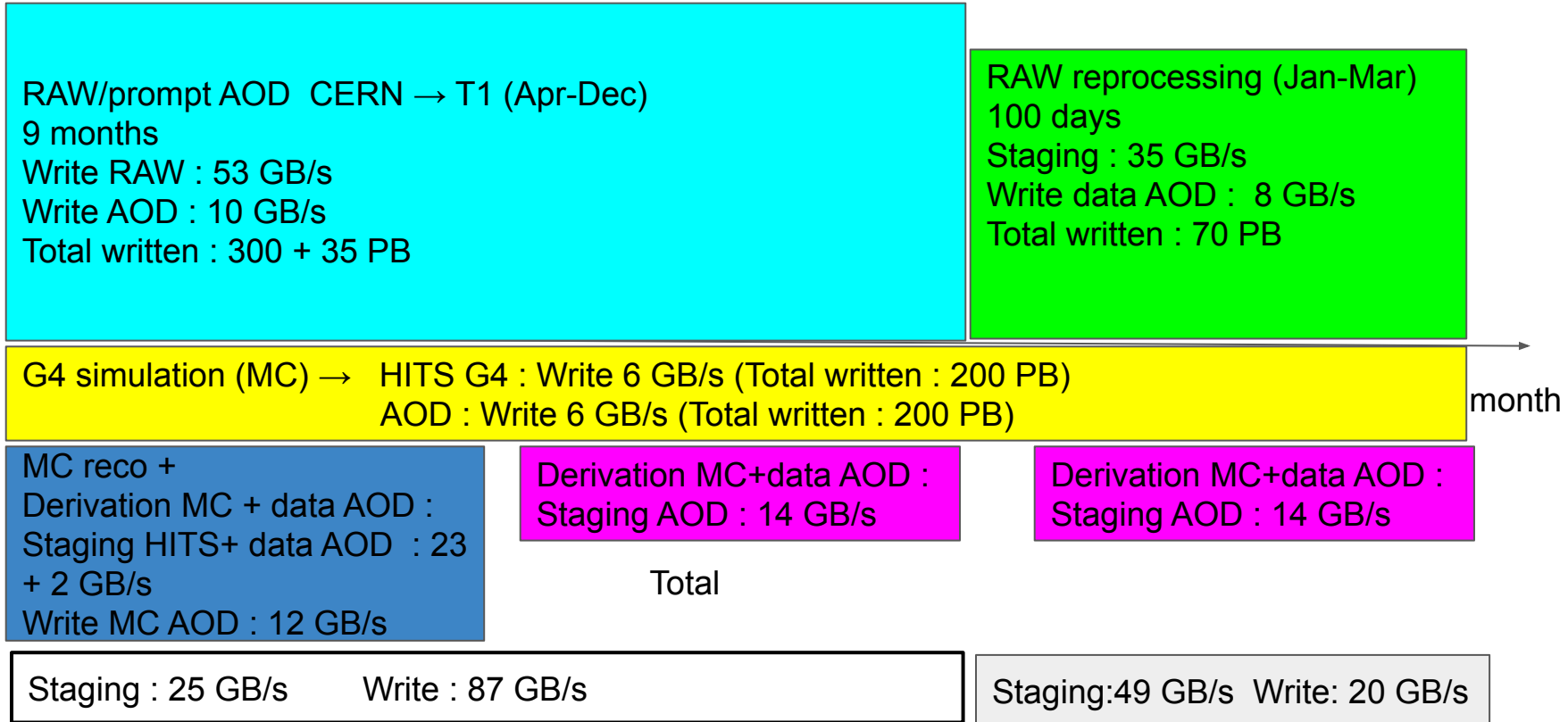
- Scenario 2 :

- Process : 50 % HITS in 100 days during shutdown : 12 GB/s (No staging of MC AOD)
- All data AOD processed from DISK copy
- Write all on DISK

- **Read data+MC AOD for derivation (No MC reco campaign) :**

- 3 repro of 100 days each year : ~Permanent derivation activity
- Most often reprocessed (benchmark, important channel) on DISK : 50 %
- Scenario 1 : 100% of AOD data+MC
  - 50% read from TAPE → staging 14 GB/s
- Scenario 2 : 50% of AOD data+MC
  - All accessed from DISK

# Scenario 1 : Maximise TAPE usage



# Scenario 2 : Minimise TAPE usage

Data taking period (Apr-Dec)  
RAW/prompt AOD CERN → T1  
9 months  
Write RAW : 53 GB/s  
No AOD exported (1 copy already at CERN)  
Total written T1: 300 PB

RAW reprocessing (Jan-Mar)

100 days  
Staging : 35 GB/s  
Write data AOD : 4 GB/s  
Total written T1: 35 PB

G4 simulation (MC) → HITS G4 : Write 6 GB/s (Total written : 200 PB)  
AOD : Write 6 GB/s (Total written : 200 PB)

month

Derivation MC+data AOD :  
All on DISK

Derivation MC+data AOD :  
All on DISK

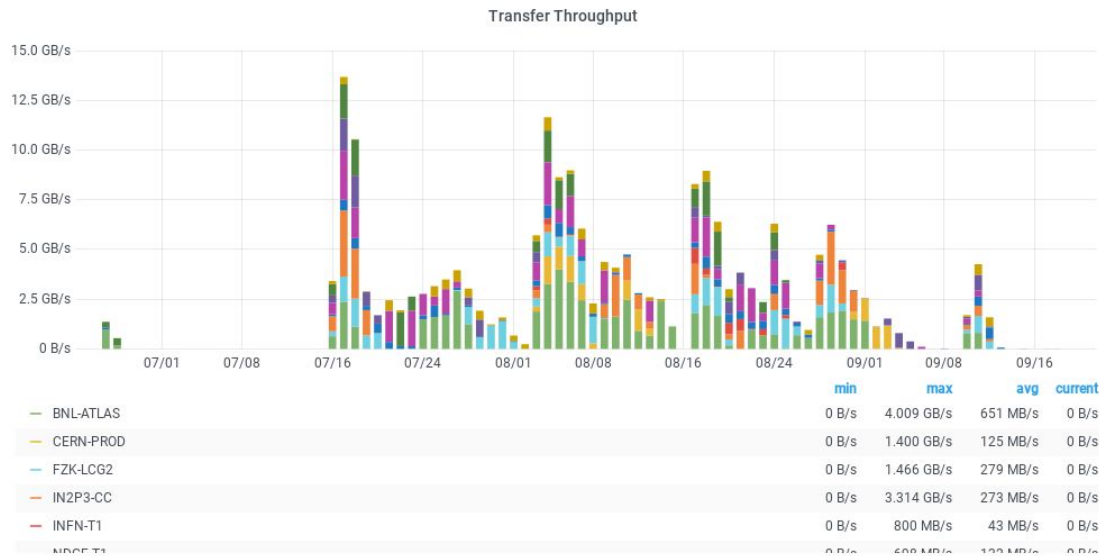
Total

MC reco +  
Derivation MC + data :  
Staging 12 GB/s

Staging : 0 GB/s      Write : 65 GB/s

Staging:47 GB/s      Write: 4 GB/s

# Staging rate in 2020



Staging activity  
(Summer 2020)

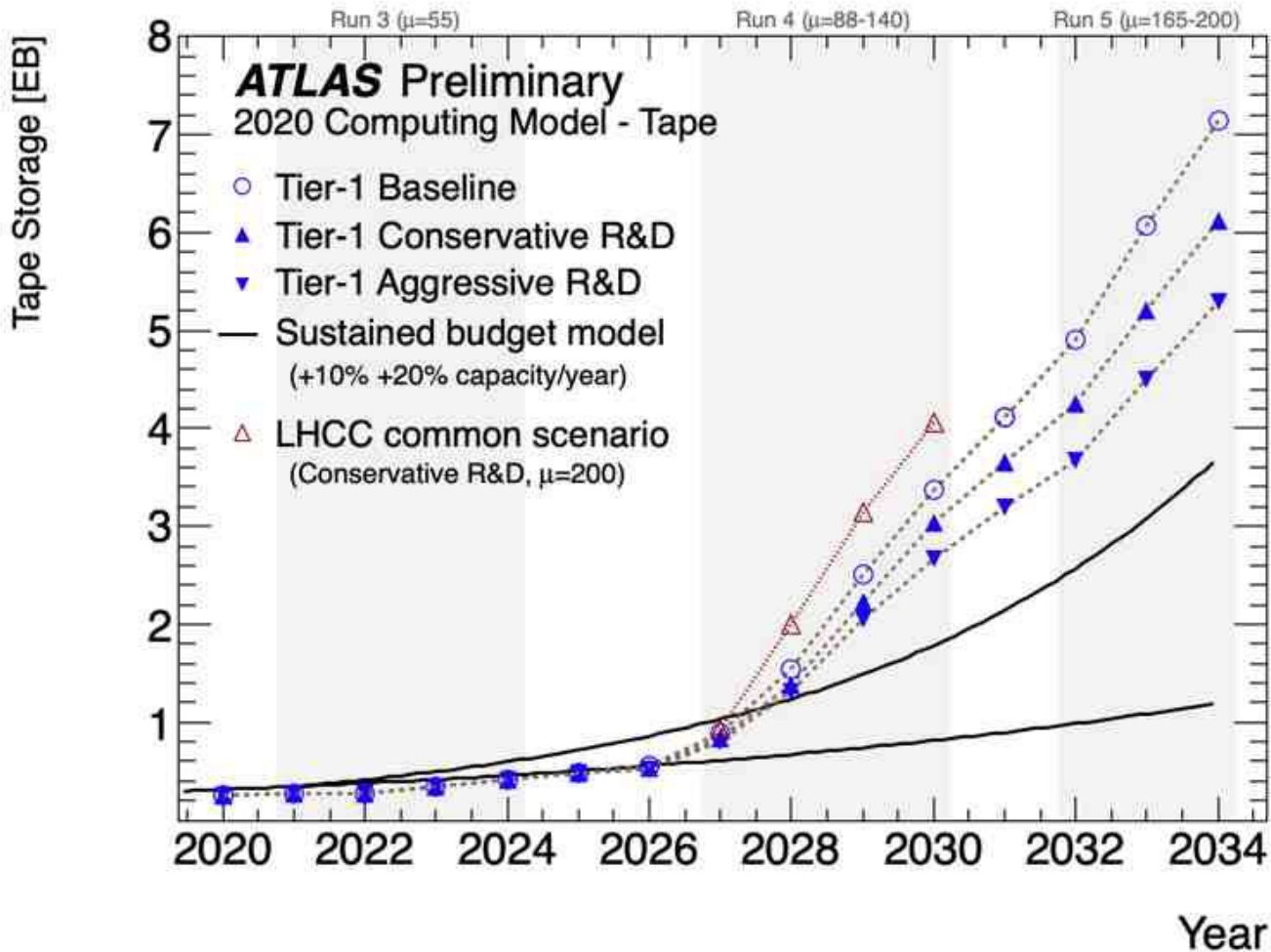
- HL-LHC read rate : 50 GB/s
  - Factor 4 compared to 2020 (+ 20 % /year until 2028)

# Typical requirements per T1

- Max requirement on tape read/write access : during data taking
  - 87 + 25 GB/s ~ 110 GB/s for all T1s
    - T1 with 25 % share : 28 GB/s
    - T1 with ~10 % share : 11-13 GB/s
    - T1 with 5% share : 6 GB/s



Backup



	Run2 repro 2020	All data (2020)	Repro year 2028
Number events (Billions)			70 <small>(10KHz * 7Msec)</small>
Event size (MB)			4.4 <small>(estimate)</small>
Total RAW size (PB)	18		300
Reprocessing time (days)	75 (not permanently full speed)		100
Reprocessing PB per day			3
File size (GB) (estimation)	2.2		10
Staging (GB/s)	15 (sum of max of each site)		40 (steady) → 60 (peak)
Fraction reprocessed outside T1	50% ?		50% ?
WAN traffic of RAW (Gb/s)	10	10-20	150
Nb transferred files / hour		100k	12 k
Processing time t <sub>bar</sub> (HS06.sec)			300 <small>(estimate)</small>
Nb core (1 core=10 HS06)	~ 150k	400k	250 k