



ALICE

LHCOPN-LHCONE meeting #49

L. Betev

Intro

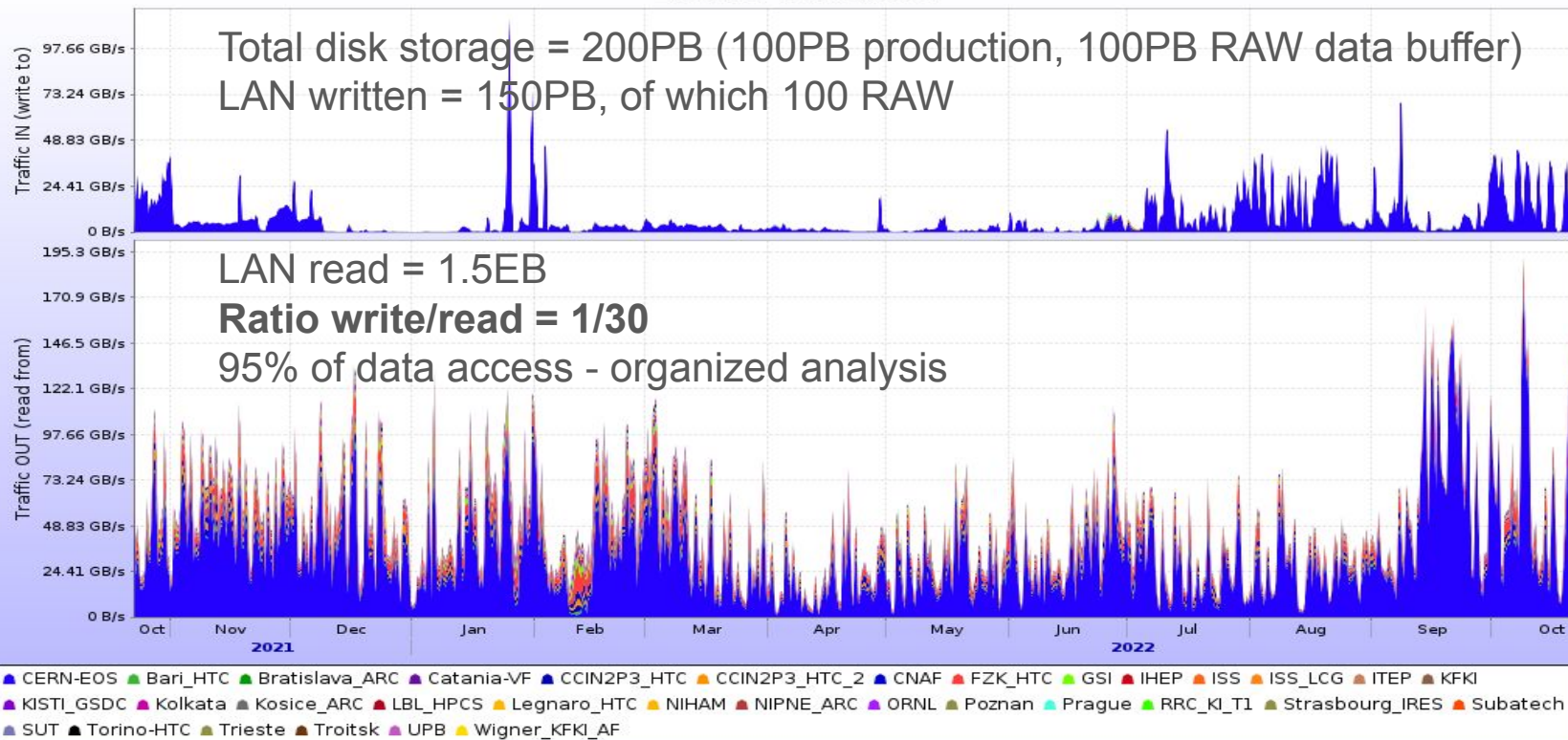
- ALICE Run3/Run4 upgrade completed, first year of operation with new detector/readout/software
 - Some uncertainties to be expected in the numbers presented
- In general, the data management model and tools remain ~same as in Run2
 - Minimize the use of WAN
 - Most of the WAN use is RAW data transfers
 - Some details to follow on
 - data processing,
 - upgrade and data transfers
 - new elements of the computing system

Data processing

- Grid site local file access (95%), remote (5%)
 - Remote access due to local SE issues, usually temporary
- Multiple replicas sorted topologically: apps first access local replica, then the next closest
 - Sorting by network topology, availability, network quality, geo-location and other metrics
- Jobs are dispatched to the Grid sites that already have the data
 - Minimizes WAN traffic and RTT efficiency penalty
- Storing multiple replicas
 - One replica is written to the local storage element
 - The other replicas are written to the remote (but close) storage elements
 - Remote writes might go through LHCOPN / LHCONE

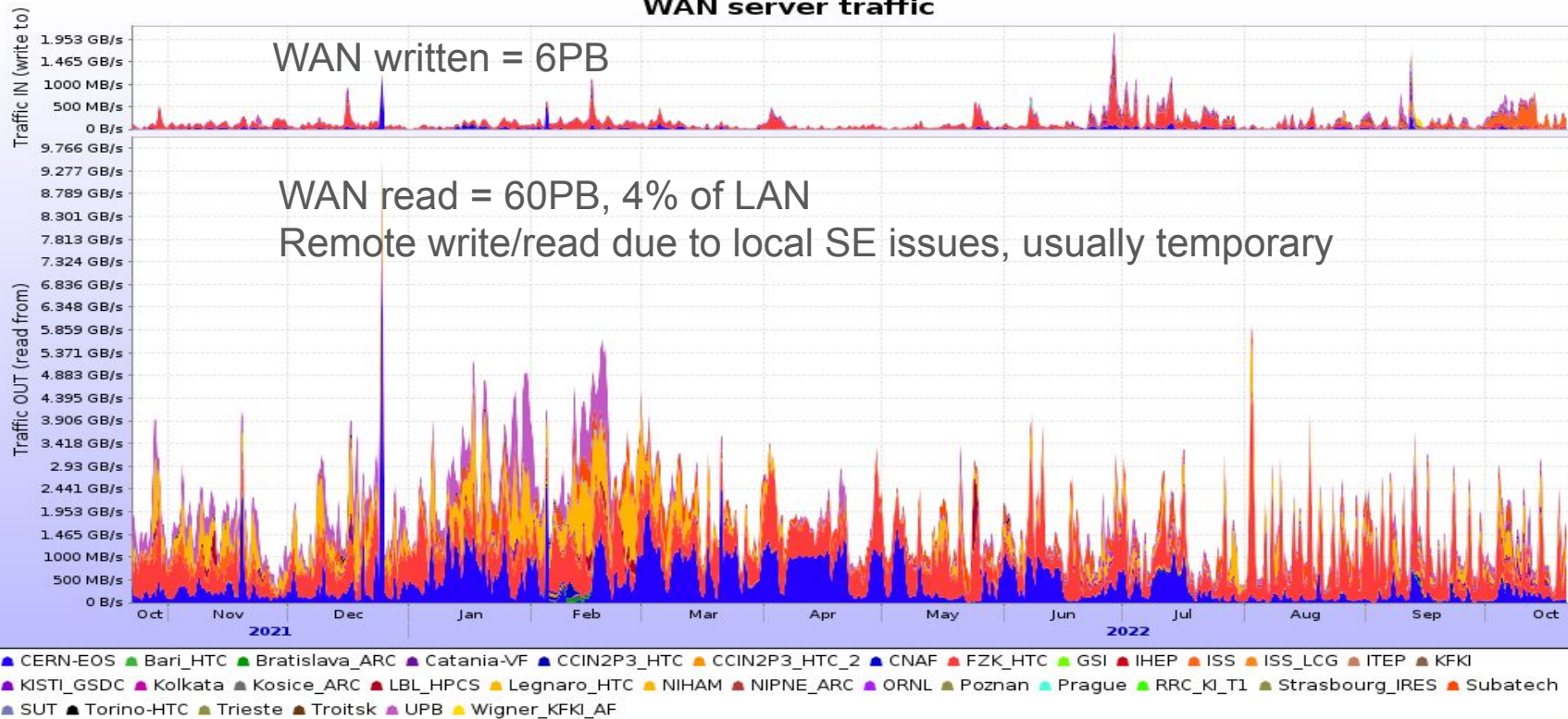
Data access - LAN

LAN server traffic



Data access - WAN (LHCONE/LHCOPN)

WAN server traffic

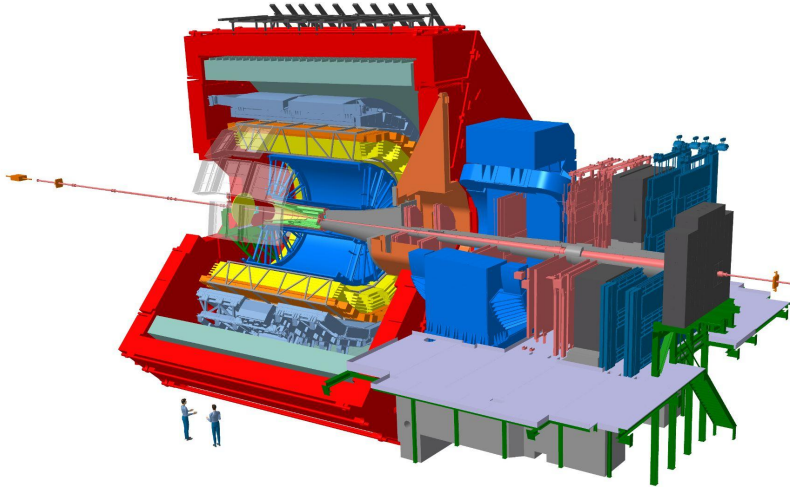


Summary of data access for past year

Description	Data volume PB
LAN write	150 (50 processing, 100 RAW data)
LAN read	1500 (95% organized analysis)
WAN write (LHCONE/LHCOPN)	6 (inaccessible local storage)
WAN read (LHCONE/LHCOPN)	60 PB (inaccessible local storage)
Other WAN transfers	20PB (file recovery/storage decommissioning, data replication)

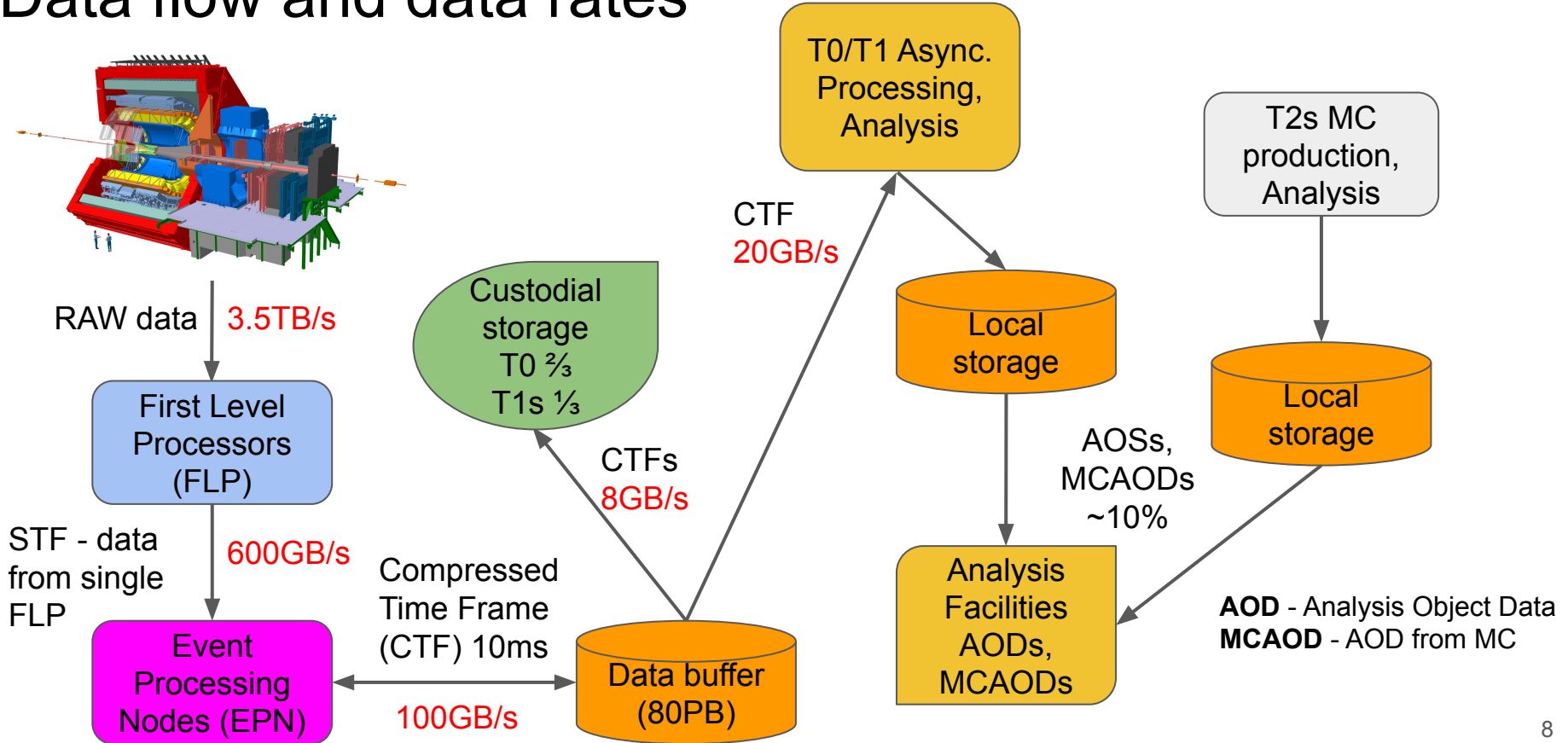
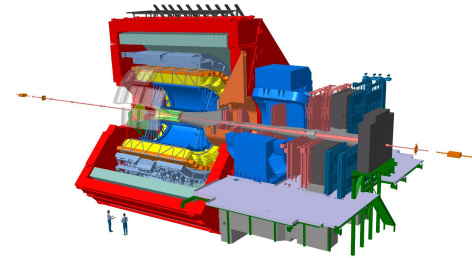
- We expect these numbers to increase $\sim 15\%$ per year
- Strongly dependent on LHC programme, maximums reached after Pb-Pb data taking

ALICE upgrade general

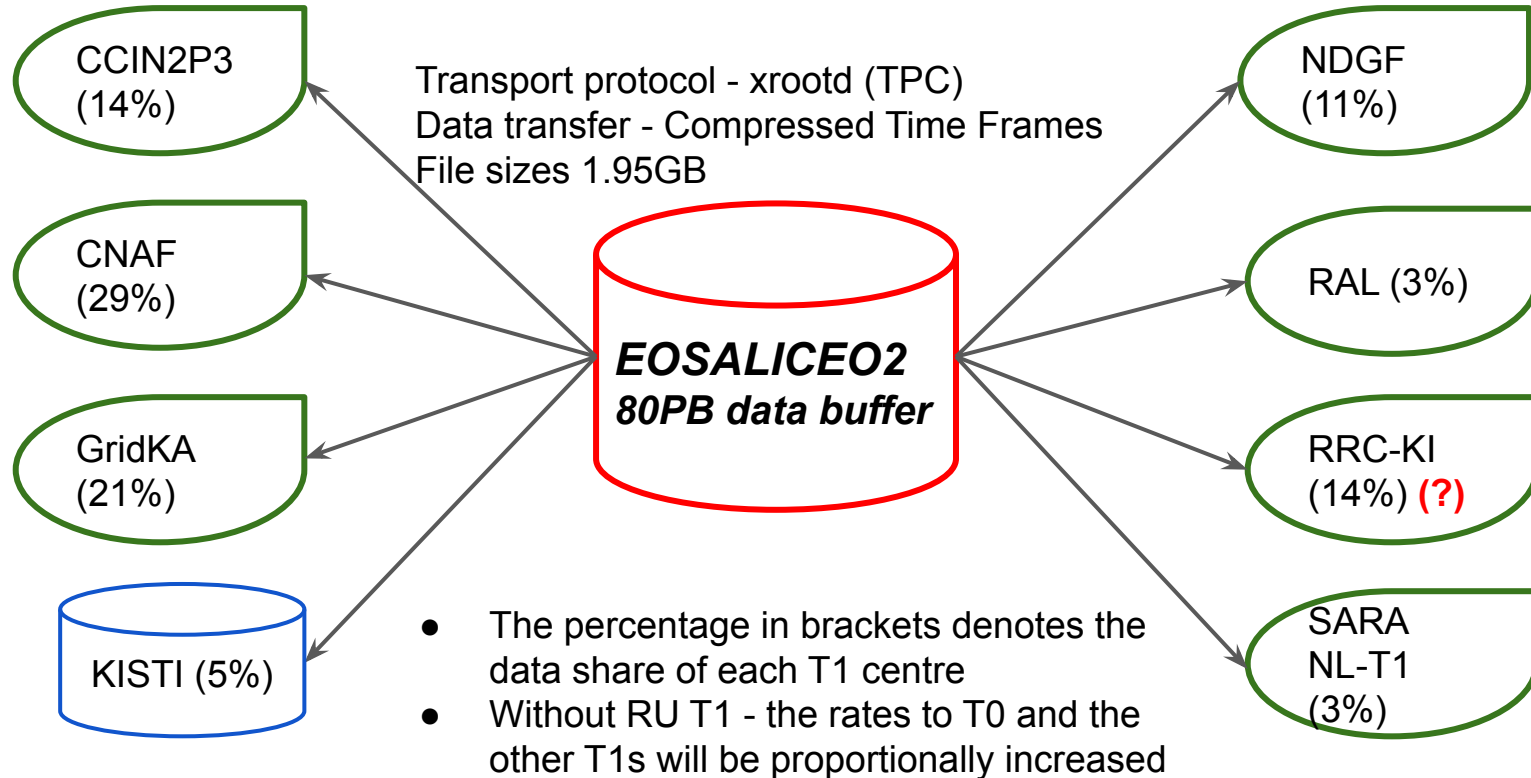


- p-p and HI physics
- 10x integrated luminosity $L \sim 10 \text{nb}^{-1}$ ($B=0.5\text{T}$)
+ 3nb^{-1} ($B=0.2\text{T}$)
- 100x event rate of Run 1/2
- Continuous readout
- Focus on data compression and real time (synchronous) data reconstruction
- => Reasonable rates and data volumes after compression to storage and secondary data formats
- Adherence to 'flat budget' resources funding for data processing and analysis

Data flow and data rates



Custodial data transfers over LHCOPN

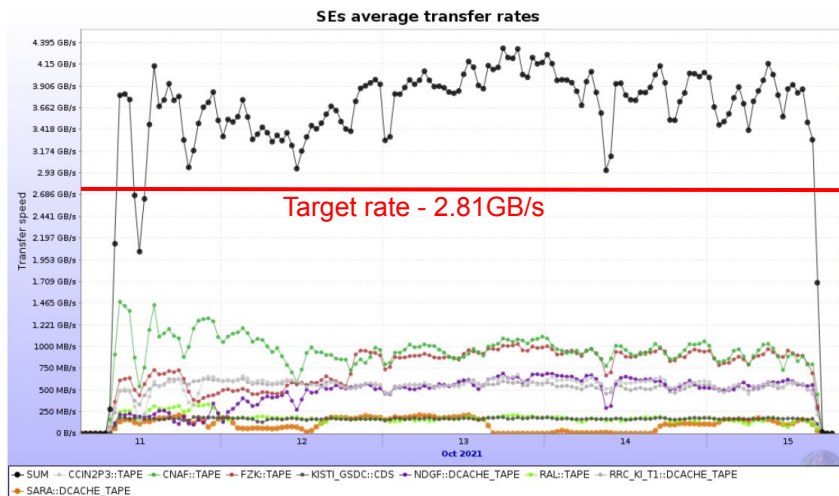


Data rates - from October 2021 data challenge

T1 Centre	Target rate GB/s	Achieved rate GB/s
CNAF	0.8	0.94 (116%)
IN2P3	0.4	0.54 (130%)
KISTI	0.15	0.16 (106%)
GridKA	0.6	0.76 (123%)
NDGF	0.3	0.47 (144%)
NL-T1	0.08	0.1 (122%)
RRC-KI	0.4	0.53 (128%)
RAL	0.08	0.17 (172%)

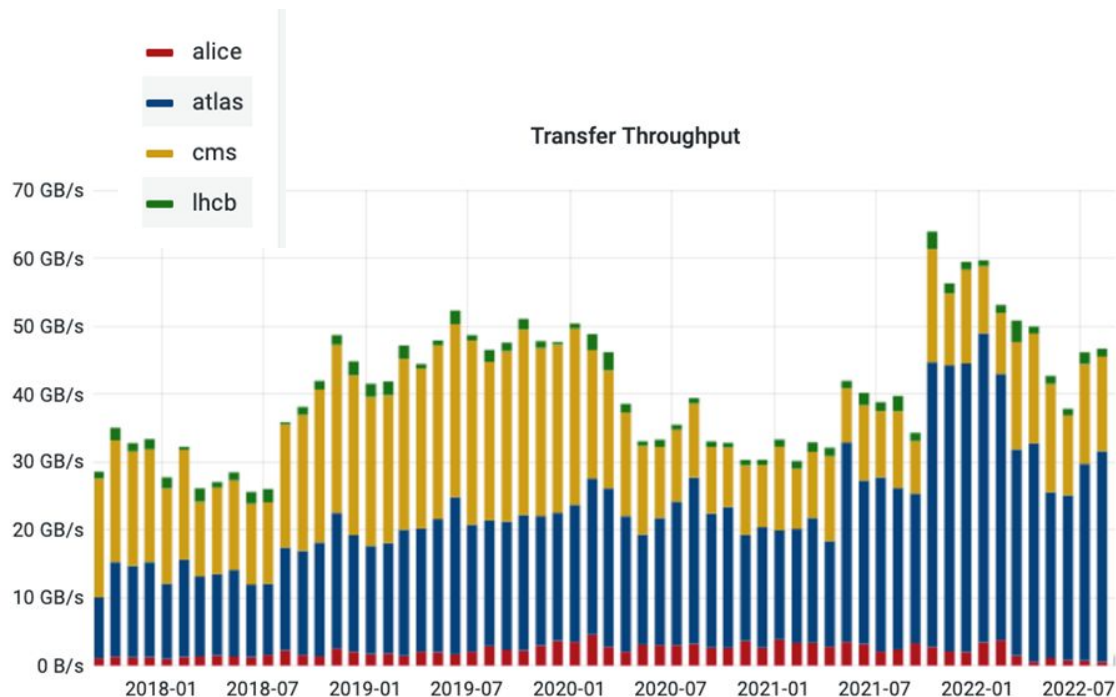
Sum 2.81GB/s

- Successful completion
- Channels tuned to slightly above the target rate, within reasonable limit
- The bulk of the bandwidth will be used after the Pb-Pb data taking period, for ~3 months
 - Since there is no Pb-Pb this year, we remain at the level of data challenges

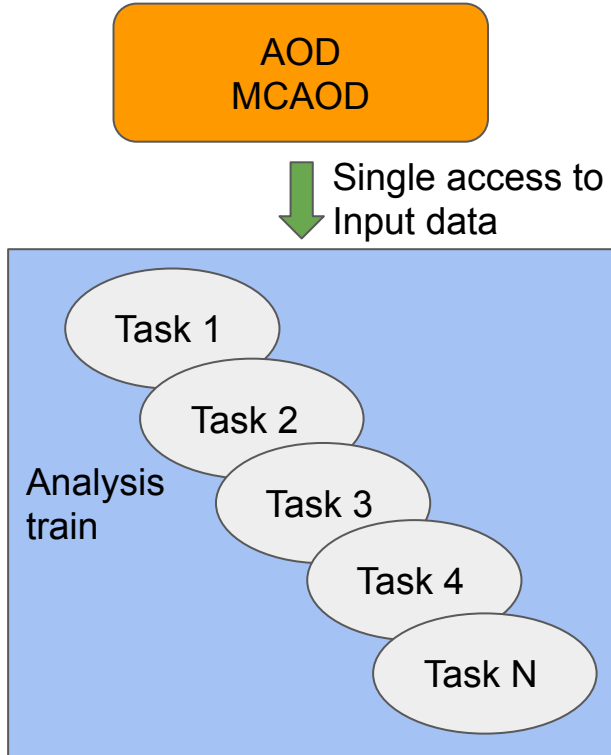


ALICE in the big picture - WLCG data transfers

- Includes RAW data distribution and other LHCONE/LHCOPN transfers



Analysis facilities (AFs)



- New element of the computing model
- Goals
 - Provide a location with comprehensive data samples from asynchronous and MC data processing at ~10% statistics
 - Fast tuning of analysis algorithms - once ready, run on full sample on the Grid
 - First data and low statistics analysis (if compatible)
- Incorporated in the Grid framework
- Sites tuned for fast I/O between storage and CPU
 - Approximate total size 6-8k cores, 10PB storage
 - ~15MB/s/core throughput
- As of today - 2 AFs (EU) possibly 1 more in US

AF data transfers

- Data is transferred to the AFs from T0/T1s/T2s
- Current AFs are co-located with T2s
- Data is transferred in blocks from the originating site
 - Can be anywhere in the world
 - Our tools take into account the network distance and copy the closest replica
- External network is not (yet) a limiting factor
 - May become one if the data turnover is greater than anticipated
- Bandwidth allocation / L3P2P service could be of interest to cover the AF use case
 - To speed up the transfers and responsiveness of the AF to analysis priorities

Summary

- ALICE is happy with LHC OPN/ONE and in general with the network performance
 - Never a limiting factor
- Our computing model favors local data access
 - WAN access for file replication and in case of issues with local storage
 - Run 3 model will continue using the same principles
- File transfers (data recovery and storage rebalancing) use will continue at the current level
- T0 to T1s data transfer of Pb-Pb data - higher LHCOPN use for 2-3 months/year
- More data from the experiment, but no significant increase on the pressure for LHC networking