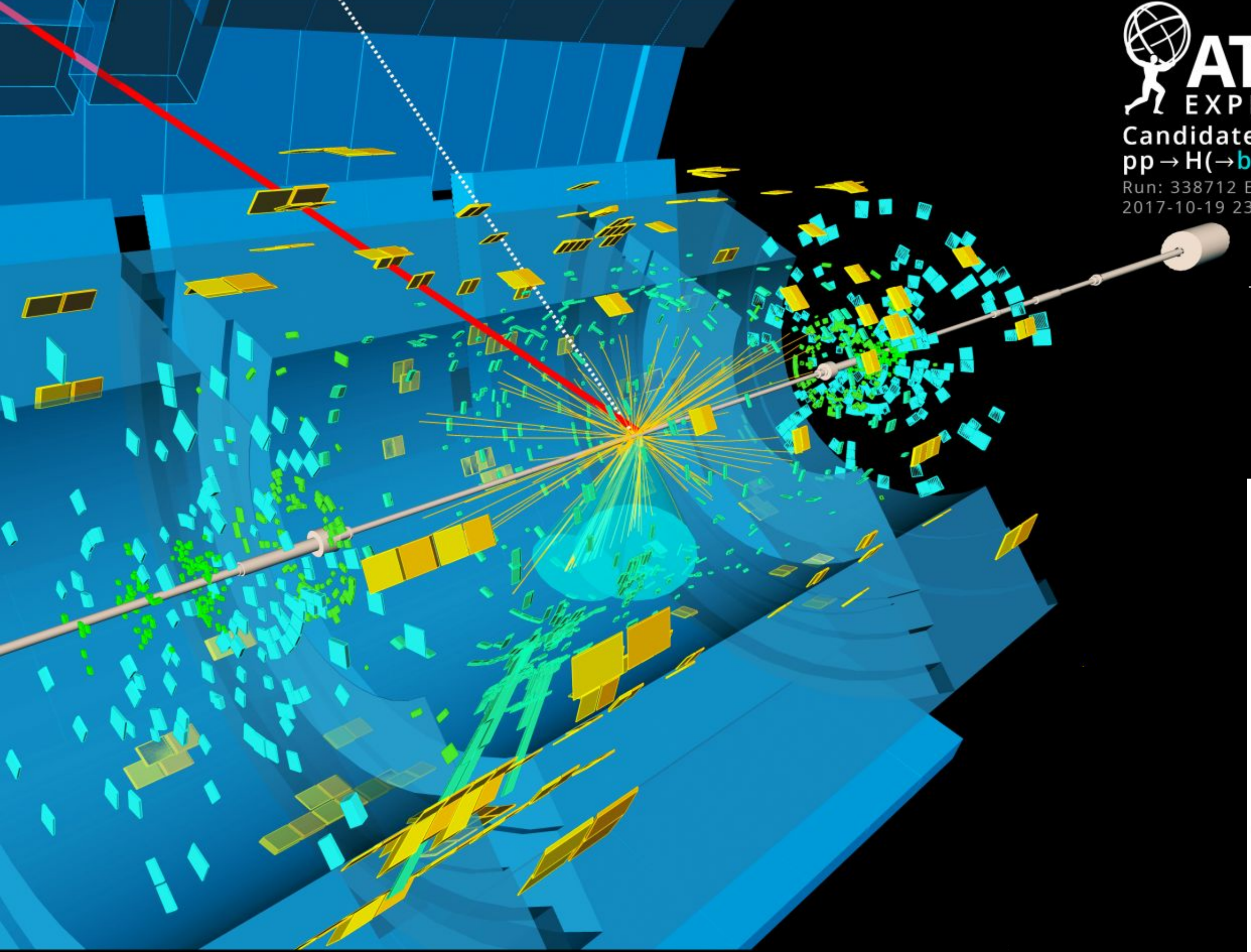
HL-LHC Data Flows for the ATLAS Experiment

or: How I learned to stop worrying and love the chaos

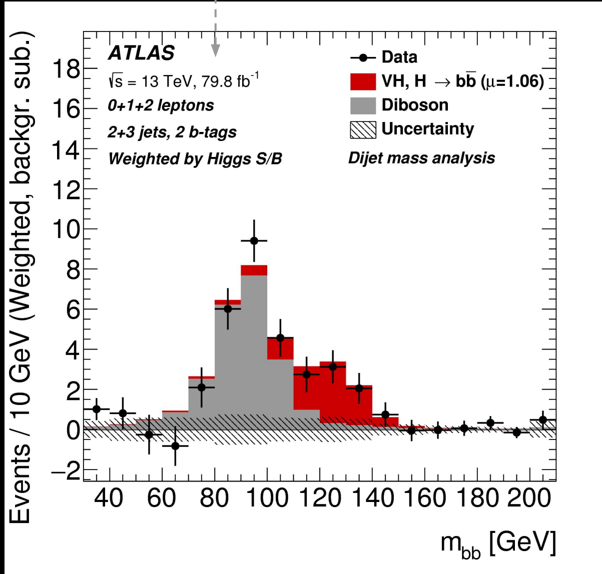
Mario.Lassnig@cern.ch

11th SIG-NGN Meeting
2023-04-20
https://wiki.geant.org/display/SIGNGN/11th+SIG-NGN+Meeting



13 TeV detector data
- 8 quadrillion collision candidates
- 92 petabytes
- 130 million files

13 TeV simulation data
- 166 petabytes
- 544 million files

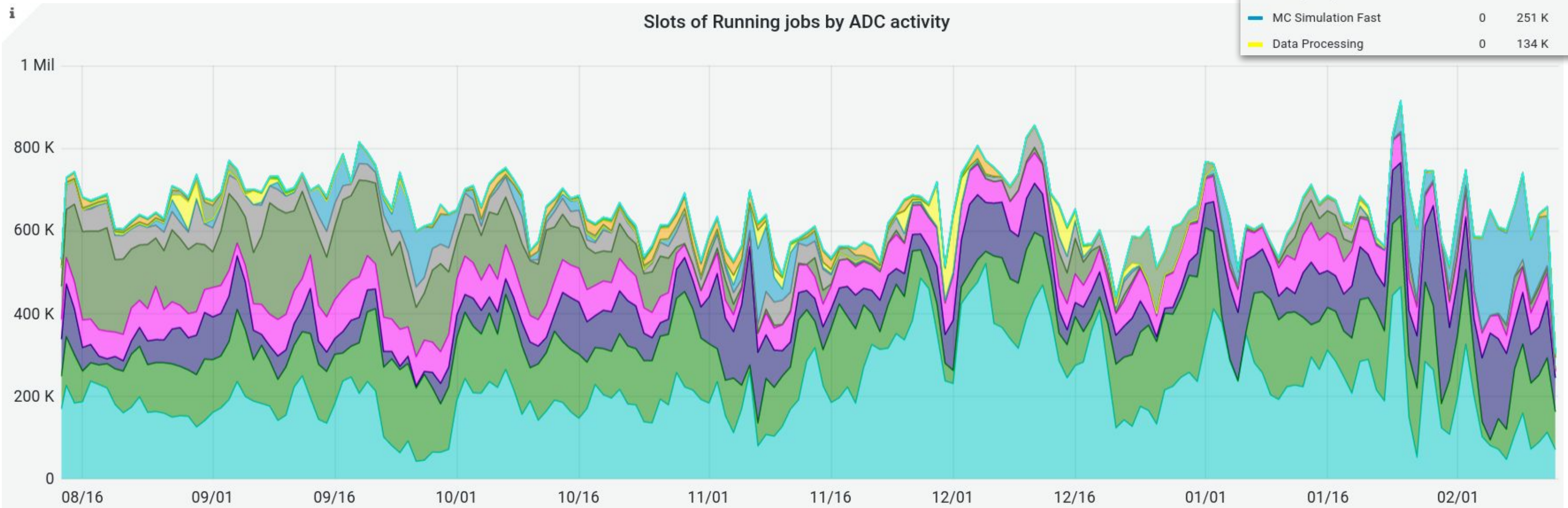


A candidate event display for the production of a Higgs boson decaying to two b-quarks (blue cones), in association with a W boson decaying to a muon (red) and a neutrino. The neutrino leaves the detector unseen, and is reconstructed through the missing transverse energy (dashed line). (Image: ATLAS Collaboration/CERN)

ATLAS computing usage

Global high-throughput computing system

- **Steady** 600k to 800k running jobs, with **full spread** of **experiment activities**
- Spread across ~250 clusters **worldwide**



ATLAS computing usage

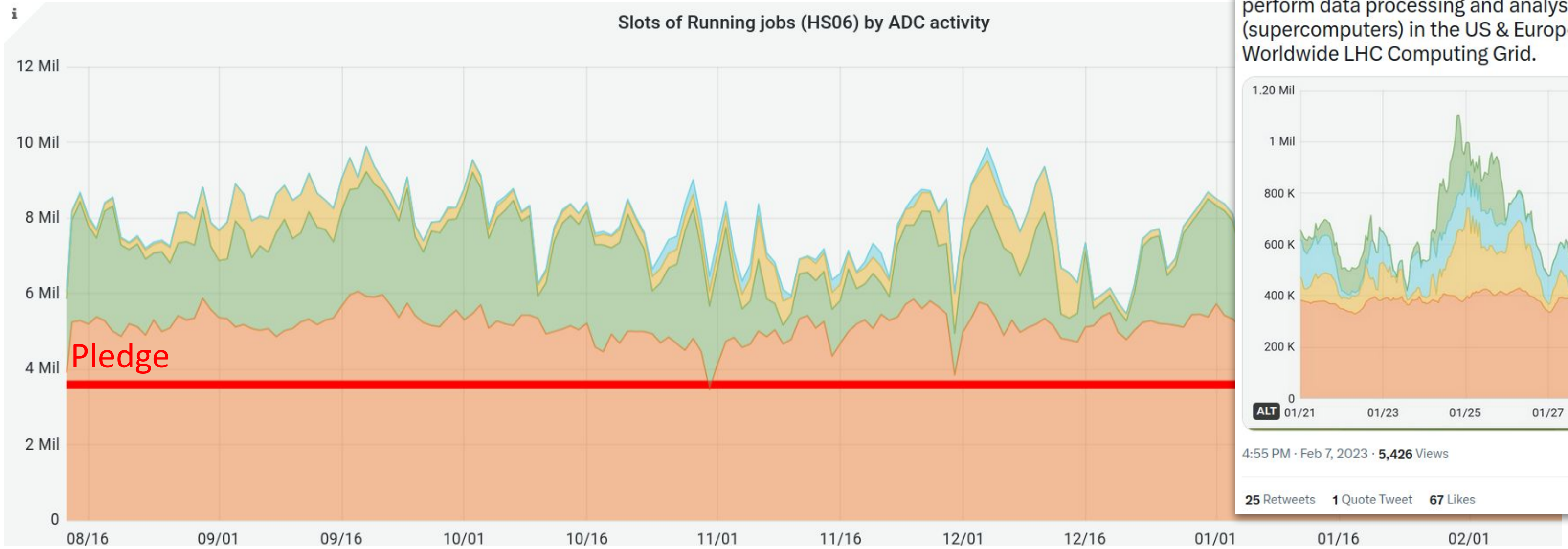
Computing power expressed in terms of HEPSPEC benchmark

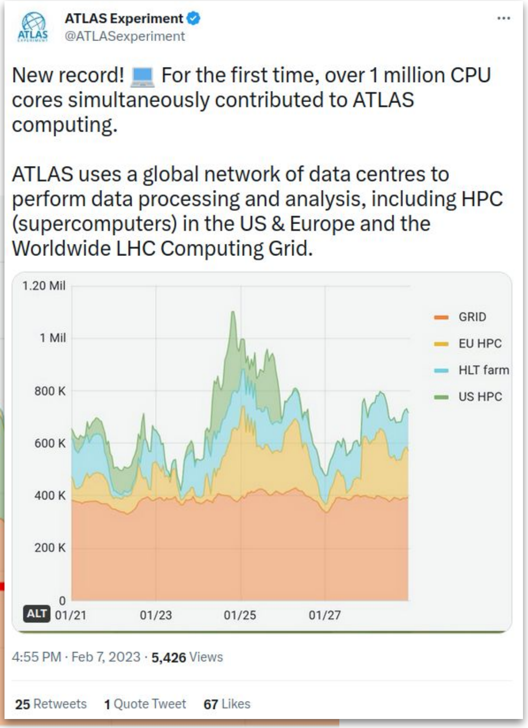
1 modern x64 core ≈ 10 HEPSPEC

Infrastructure is **consistently over pledge**

Opportunistic resources

Scale out to 1+ million jobs

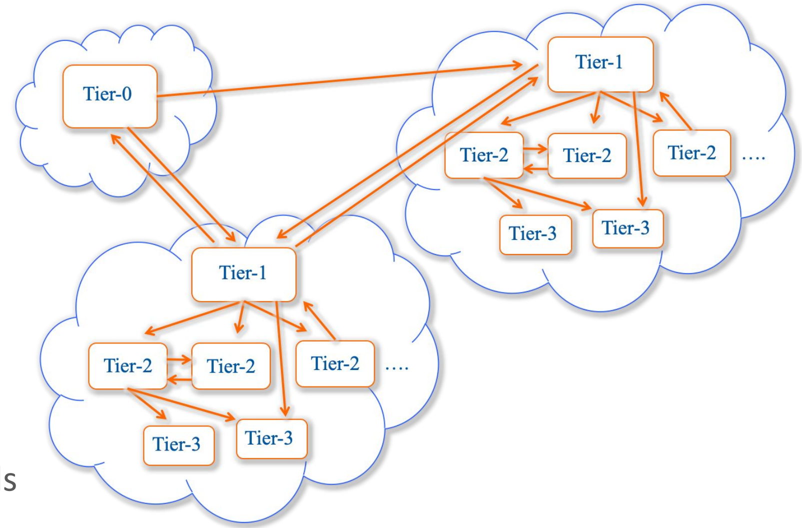




Basic experiment data flows 1/2

Original ATLAS computing model designed as static **clouds**

- ATLAS Clouds ≠ "Cloud computing"
- Mostly national or geographical **groupings of sites**
- **Common funding** agencies
- Support often using the **same language**



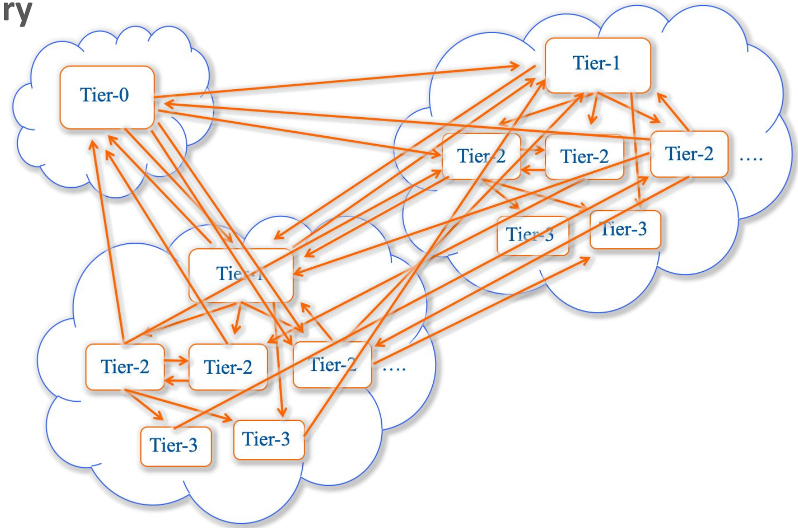
Model had a series of shortcomings

- Individual tasks **inflexibly executed** within a static cloud
- All tasks **output aggregated** at the 10 Tier-1s
- The **Tier-2 storage** was not optimally exploited
- **High priority tasks** were **occasionally stuck** at small clouds

Basic experiment data flows 2/2

WLCG networks have evolved significantly in the last decades

- **Limiting transfers** within a single cloud **no longer necessary**
- Now single **WORLD cloud** site concept



Nucleus

- **Any stable site** can aggregate the output of a task
- Site **can be manually assigned** as a nucleus

Satellites

- **Process the jobs** and **send the output** to the nucleus
- **Defined dynamically** for each task
- **No longer confined** inside the original cloud

Currently around **130 active sites** used by ATLAS

Experiment job types

ATLAS EXPERIMENT

Global shares are employed to allocate the available resources among the activities

- Done on **agreement** between the various production and physics groups
- **Hierarchical** implementation
- Related activities have the opportunity to **inherit unused resources**

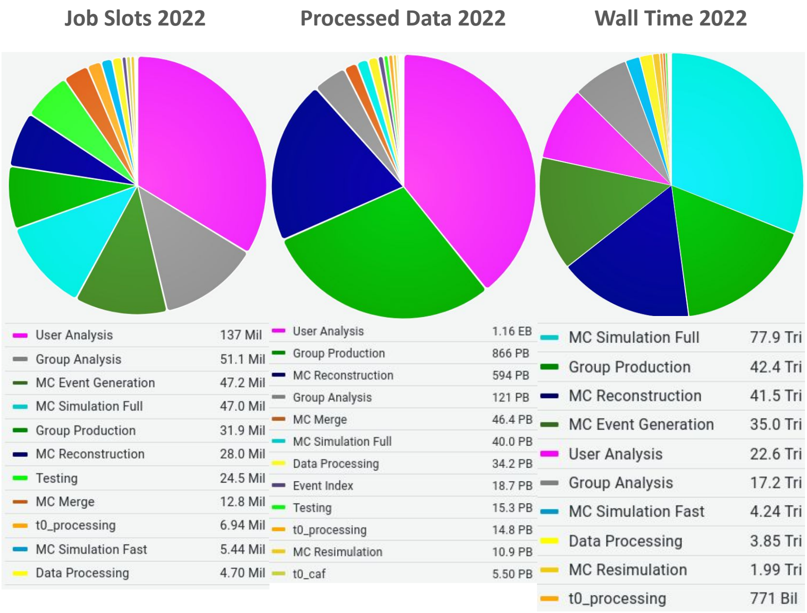
Essentially two categories of jobs

Production
- Data reprocessing
- Event generation / Simulation / Reconstruction
- Group production

Analysis
- User analysis
- Group analysis

The main activity at a given time can depend on many things

- Data **reprocessing** or Monte Carlo **production** campaigns
- **Conference** deadlines, need for an increase for user analysis
- Global **pandemics**

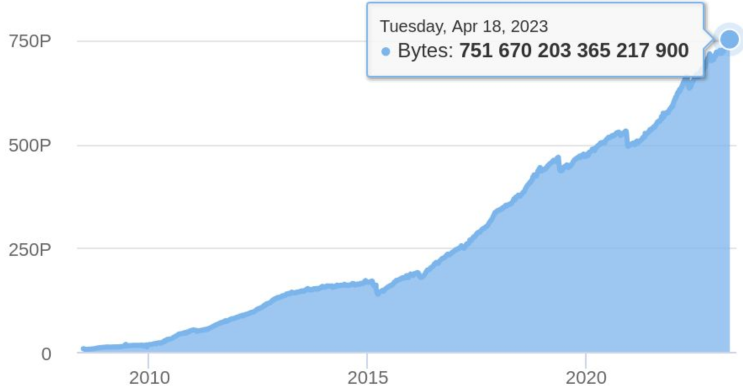


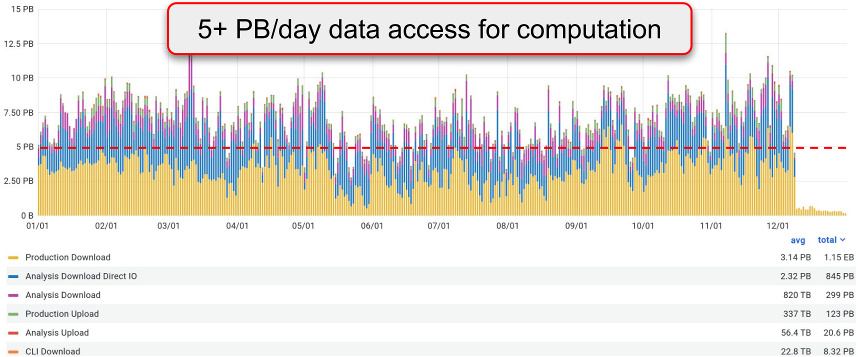
Data transfer rates

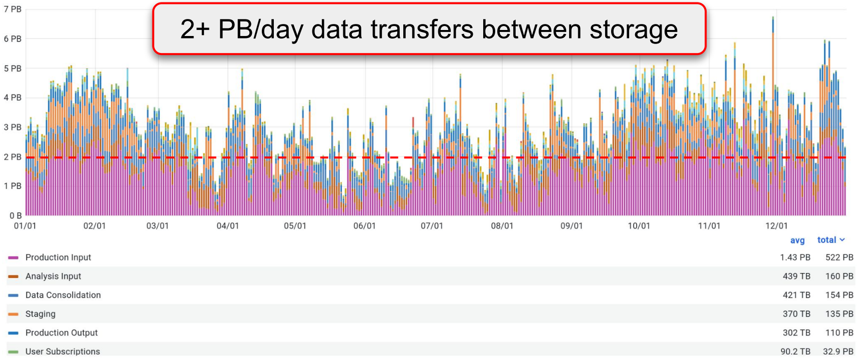
A few numbers showing the ATLAS scale

- 1B+ files, 750+ PB of data, 400+ Hz interaction
- 120 data centres, 5 HPCs, 3 clouds, 1000+ users
- 1.2 Exabytes/year transferred
- 2.7 Exabytes/year uploaded & downloaded

Increase 1+ order of magnitude for HL-LHC







Data management

Rucio handles the data management

- Creation, location, transfer, deletion, annotation, and access
- **Orchestration of dataflows** with both low-level and high-level policies
- **Coherent interface** required to allow smooth data handling for production and users
- We also have data management **internal flows** (recovery, rebalancing, …)

ATLAS sites are not homogeneous

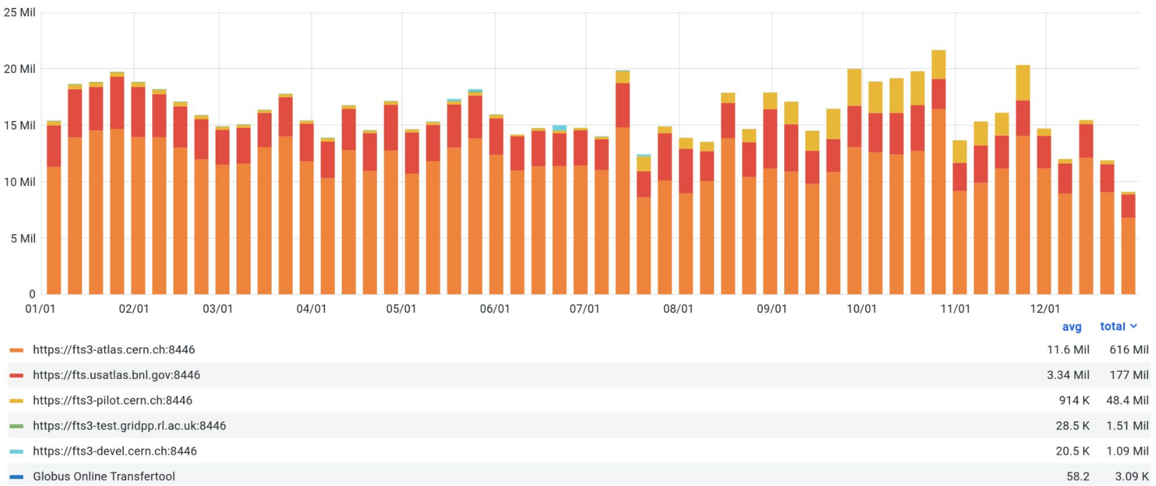
- **Different** storage, **different** protocols
- Abstracted by **FTS, GFAL** and **Davix**

ATLAS deployment

- Two FTS servers in production
- Plus regularly the pilot & test services

Average file flow rate

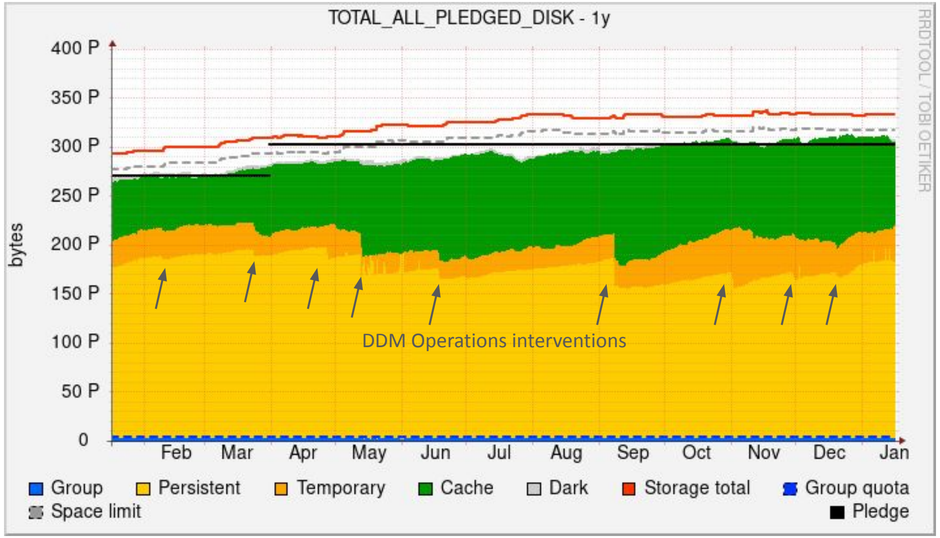
- 15 million successful transfers per day
- 2 million failed transfers per day
- Mostly site configuration problems
- Failures biased because of quick retries

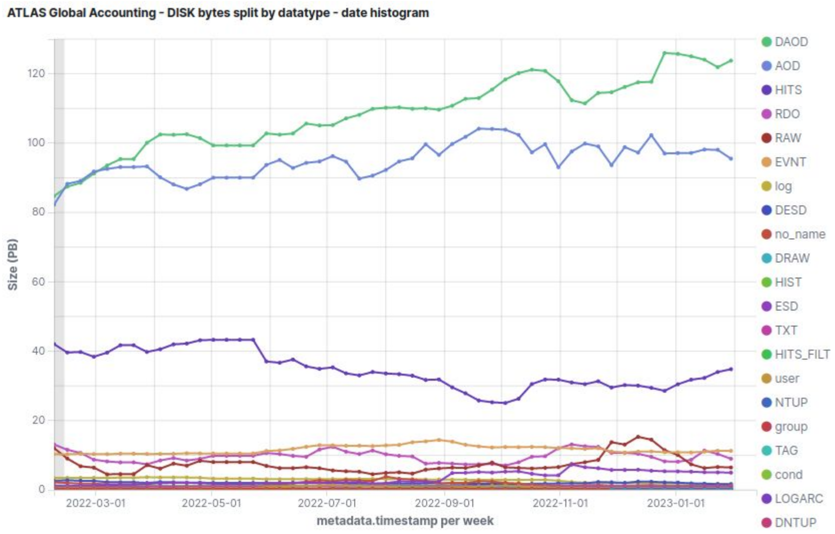


Disk resource usage

Situation improved slightly throughout the last year, however continuous intervention necessary

- Much better cached-to-persistent ratio, however we were already over the pledge
- AOD and HITS volume is stable, DAOD grows from constant production, regular obsoletion to keep it in check





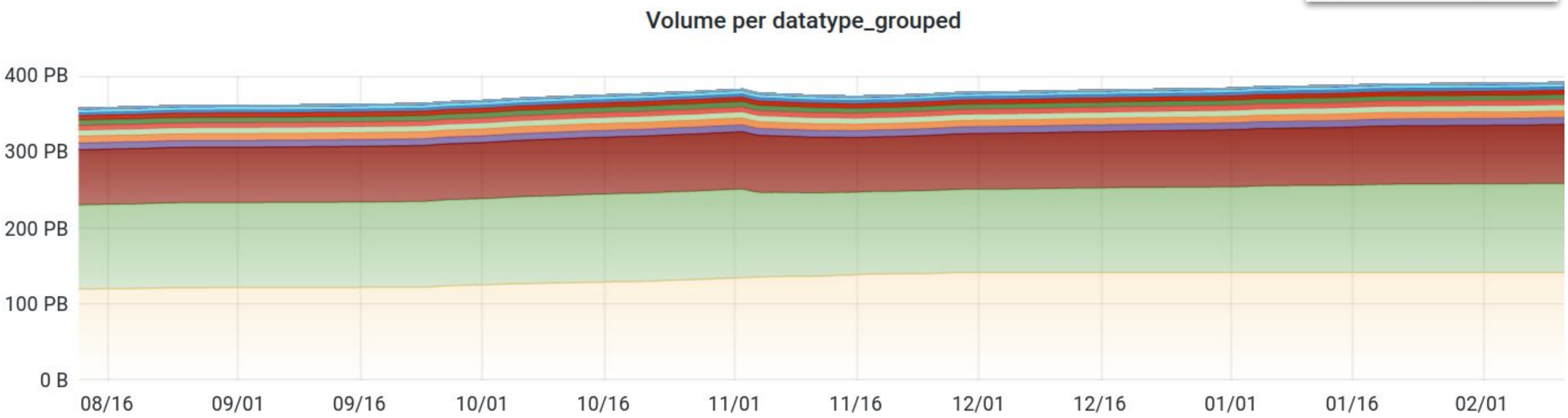
Tape resource usage

Tape situation at Tier-1s has reached the 2022 pledge

- Deletion campaign beginning of November bought us some time
- Tier-1s deploying 2023 pledge early

Tape has moved from pure archive storage towards more dynamic integration

	max	avg
RAW	142 PB	133 PB
AOD	117 PB	112 PB
HITS	77.9 PB	74.0 PB
DRAW	9.11 PB	8.93 PB
DAOD	9.26 PB	8.84 PB
NTUP	7.36 PB	7.32 PB
DESD	7.16 PB	6.85 PB
ESD	7.02 PB	6.69 PB



Network planning

Network upgrades for HL-LHC

- Planning document
- **Export of RAW data** from CERN to the T1s
- **Data processing** flows
- **Incremental steps** until HL-LHC
- **Accompanying R&D** programme

2020 estimation

4.8 Tbps of total network capacity

ATLAS & CMS	400 Gbps flat
ALICE & LHCb	100 Gbps flat
x2	to absorb expected bursts
x2	overprovisioning for operational flexibility

As of now, the final HL-LHC estimation has not changed

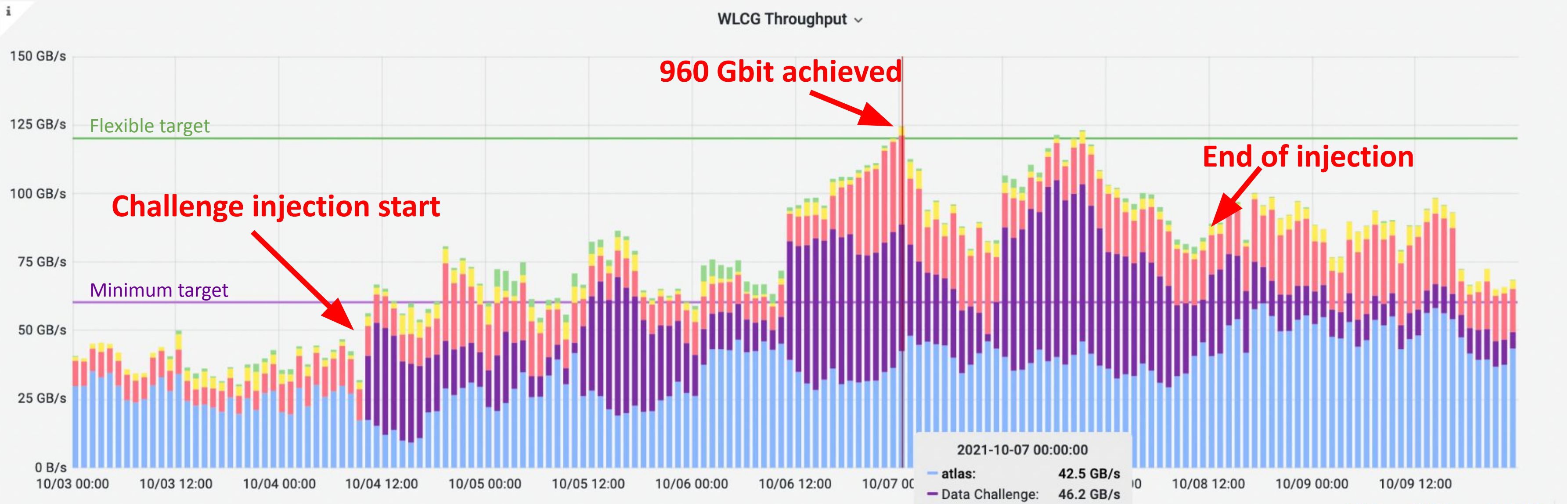
Software-defined network (SDN) developments will be crucial

T1	%ATLAS	%CMS	% Alice	% LHCb	ATLAS+CMS Network Needs (Gbps) Minimal Scenario in 2027	Alice Network Needs (Gbps) Minimal Scenario in 2027	LHCb Network Needs (Gbps) Minimal Scenario in 2027	LHC Network Needs (Gbps) Minimal Scenario in 2027	LHC Network Needs (Gbps) Flexible Scenario in 2027
CA-TRIUMF	10	0	0	0	200	0	0	200	400
DE-KIT	12	10	21	17	450	80	70	600	1200
ES-PIC	4	5	0	4	180	0	20	200	400
FR-CCIN2P3	13	10	14	15	450	60	60	570	1140
IT-INFN-CNAF	9	15	26	24	480	110	100	690	1380
KR-KISTI-GSDC	0	0	12	0	0	50	0	50	100
NDGF	6	0	8	0	110	30	0	140	280
NL-T1	7	0	3	8	140	10	30	[illegible]	360
NRC-KI-T1	3	0	13	5	50	50	20	[illegible]	240
UK-T1-RAL	15	10	3	27	490	10	110	[illegible]	1220
RU-JINR-T1	0	10	0	0	200	0	0	[illegible]	400
US-T1-BNL	23	0	0	0	450	0	[illegible]	[illegible]	900
US-FNAL-CMS	0	40	0	0	800	0	[illegible]	800	1600
(atlantic link)					1250	0	[illegible]	1250	2500
Sum	100	100	100	100	4000	[illegible]	[illegible]	4810	9620

Step estimations outdated!

T1	LHC Network Nee… (Gb… Minim…	…s …ps) …e Scenario in 2027	Data Challenge target 2027 (Gbps)	Data Challenge target 2025 (Gbps)	Data Challenge target 2023 (Gbps)	Data Challenge target 2021 (Gbps)
CA-TRIUMF	200	400	100	60	30	10
DE-KIT	600	1200	300	180	90	30
ES-PIC	200	400	100	60	30	10
FR-CCIN2P3	570	1140	290	170	90	30
IT-INFN-CNAF	690	1380	350	210	100	30
KR-KISTI-GSDC	50	100	30	20	10	0
NDGF	140	280	70	40	20	10
NL-T1	180	360	90	50	30	10
NRC-KI-T1	120	240	60	40	20	10
UK-T1-RAL	610	1220	310	180	90	30
RU-JINR-T1	200	400	100	60	30	10
US-T1-BNL	450	900	230	140	70	20
US-FNAL-CMS	800	1600	400	240	120	40
(atlantic link)	1250	2500	630	380	190	60
Sum	4810	9620	2430	1450	730	240

Data Challenge 2021



	max	avg	current
atlas	60.0 GB/s	35.0 GB/s	43.5 GB/s
Data Challenge	64.5 GB/s	25.8 GB/s	6.09 GB/s
cms	33.6 GB/s	15.4 GB/s	15.7 GB/s
alice	9.75 GB/s	3.50 GB/s	3.11 GB/s
lhcb	7.33 GB/s	1.29 GB/s	236 MB/s

HL-HLC Data Roadmap

Next data challenge jumps from 10% (960 Gbps) to 25% (2400 Gbps) of HL-LHC needs

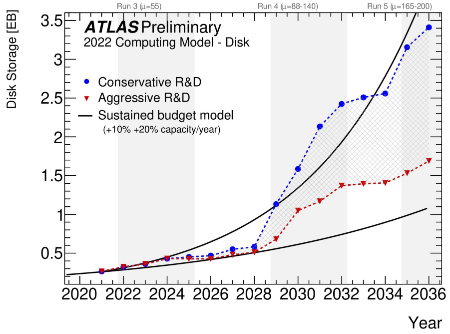
- **Large single step increase** of volume in the decade-long plan - had to reduce from 30%
- Need to reconsider due to **new HL-LHC schedule** and hardware purchasing

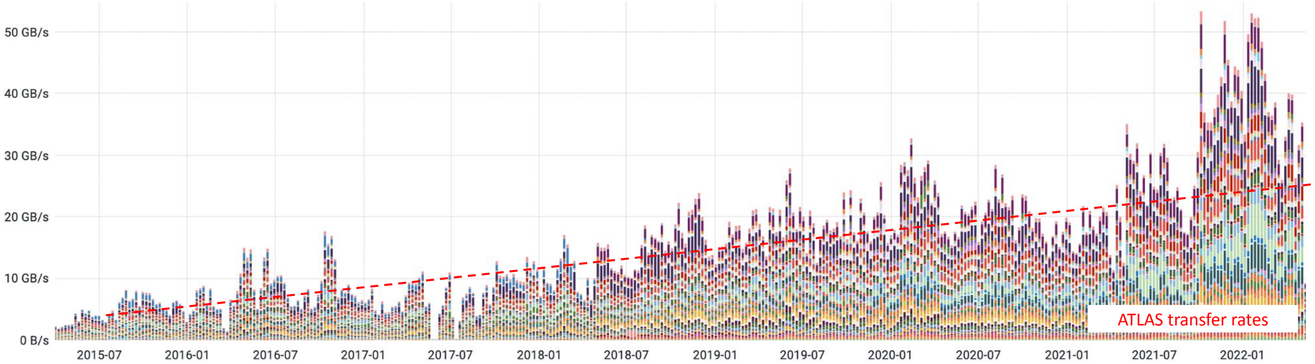
With communities beyond WLCG, such as DUNE, SKA, Belle II, JUNO, …and the NRENs

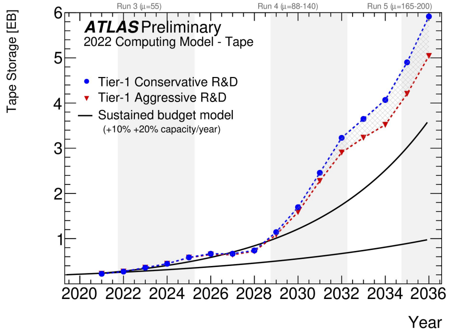
- We spend a considerable effort to **share our data management stack**
- Allows us to **work together** on these shared challenges

One interesting point: For the middleware stack, the volume is rather irrelevant

- **Number of files total**, and **number of files processed** is the key metrics
- ATLAS stance on ***big files vs. lots of files*** not yet decided







Cloud

ATLAS has **cloud R&D projects** ongoing with **Amazon**, **Google**, and **SEAL** Storage

- **Integration** into ADC systems PanDA & Rucio, and in turn FTS, GFAL, Davix
- Very **close development collaboration** across the full stack

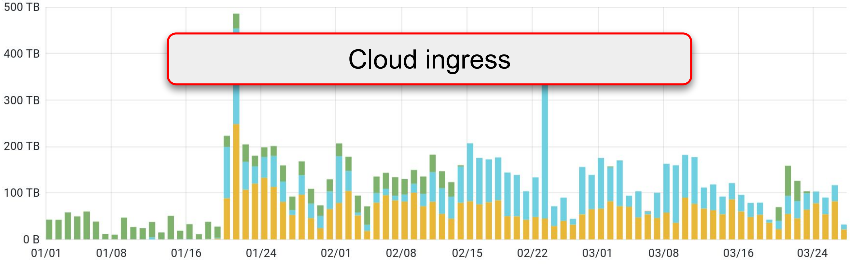
Two major angles to consider when discussing clouds

Technical	Access tools, transfer protocols, monitoring, authn/z, accounting, billing, storage, …	
Organisational	Deployed on-site or off-site	Centralised or distributed
	Public (institute, laboratory, …) or commercial	In-kind contribution or paid service

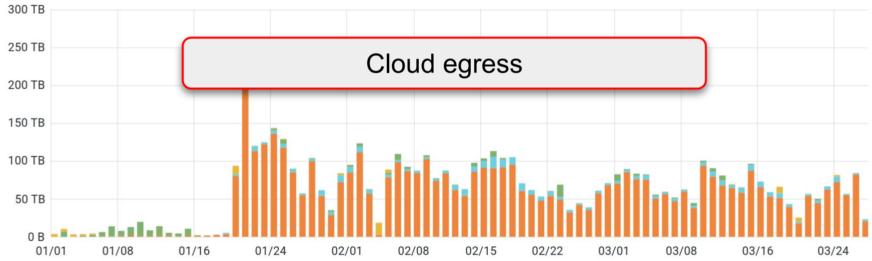
Large development programme in front of us to make cloud storage viable

- Throughput **control**, access **control**, peering **control**, cloud transfer tool **control**, lifetime **control**, cost **control**, …

WIP



	avg	total
GOOGLE-EU_SCRATCHDISK	52.5 TB	4.57 PB
GOOGLE-EU_DATADISK	44.8 TB	3.90 PB
SEAL_TEST	18.0 TB	1.56 PB
FRESNO-AWS_SCRATCHDISK	12.3 GB	1.07 TB

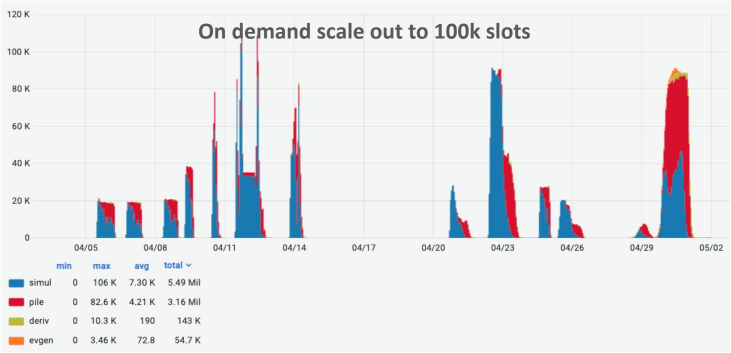


GOOGLE-EU_DATADISK	55.2 TB	4.80 PB
GOOGLE-EU_SCRATCHDISK	3.80 TB	330 TB
SEAL_TEST	2.40 TB	208 TB
GOOGLE_EU	809 GB	70.4 TB
FRESNO-AWS_SCRATCHDISK	8.64 GB	752 GB

Cloud Scale-out

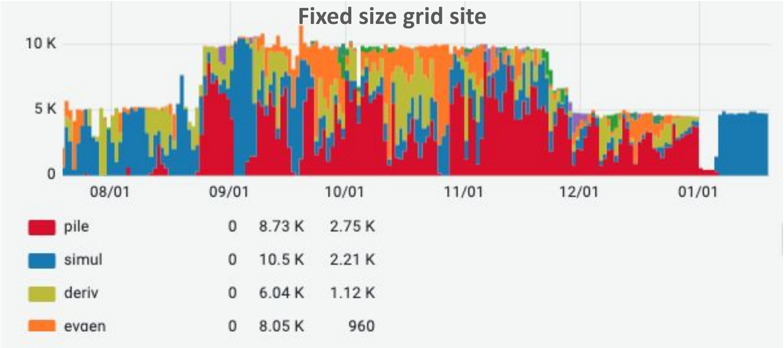
Long-term R&D collaboration with Google Cloud

- **Feasibility study** about cloud resource integration
- **Total Cost of Ownership** evaluation
- **Full integration** in workflow and data management stack
- Built on **cloud-native** technologies: Kubernetes & S3v4
- **No vendor lock-in**



Gives us possibilities to try out interesting use cases

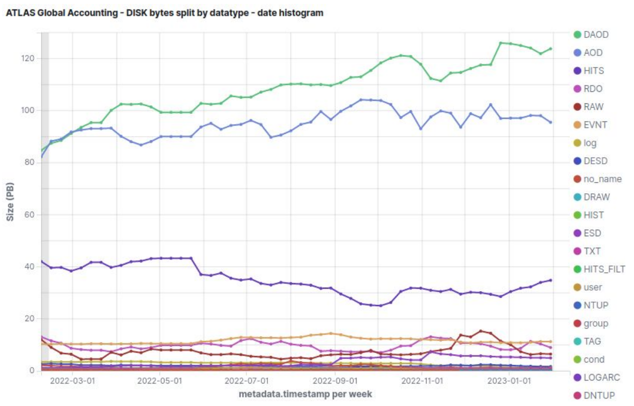
Cloud bursting	Dynamic/on-demand allocation
Network offloading	Use Google network for transfers
GPUs	On-demand GPU hardware
Special analysis	Machine Learning, Fitting



Unused data

Large volumes of unused data kept on disk due to lifetime model exceptions

- Labour-intensive procedure for ADC and physics groups
- Increase in length of publication procedures leads to data being kept on disk



Volume of DAOD on disk follows the same trend for data & MC

- Lifetime model exclusion and deletion lists are of similar volume
- Almost all DAODs are from input AODs, only 15% of input AODs are on disk
- Produced by ~10 users, then used by 580+ users

Coherent R&D of all involved mechanisms

- Lifetime model, data popularity, data placement, data caching, and Data Carousel
- Consider volume, access patterns, user requests, available resources, operational load

Demonstrator scenarios for DAOD handling underway

Status quo	Do not change anything
Delay	Keep datasets on tape/disk and delete after one year with no extension
Reproduce	Remove from lifetime exception list, delete immediately, and reproduce when needed
Archive	Archive to tape, then delete from disk

ADC-preferred solution

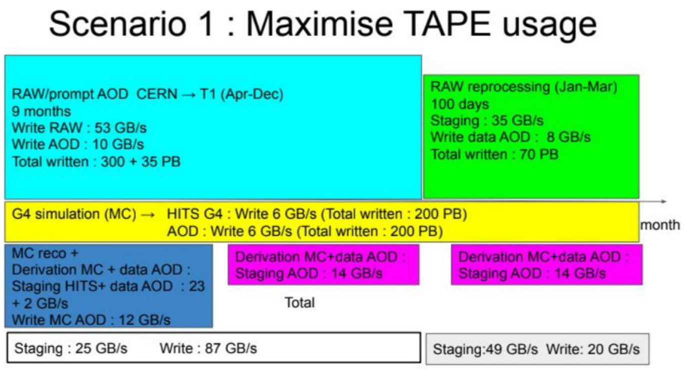
Smart archives

Smart archives: Core strategy R&D for our tape storage

- Optimise file placement on tape for efficient retrieval
- Would greatly improve Data Carousel throughput & latency

Three-phased approach

1. Definition of relevant metrics
   - Includes study of data access patterns
   - Tape IO metrics globally and individually
   - Consolidation of metadata required for efficient archival
2. Functional test to validate the full chain at FZK and BNL
   - Propagation of metadata for site to colocate data through our stack (PanDA/Rucio/FTS/dCache)
   - Manual operations and monitoring by site experts of the underlying tape system, e.g., HPSS
3. Test real use in production
   - Spawn appropriately sized tasks with data samples in the 100 TB range
   - Assess effect of automatic colocation through tasks defined by production managers

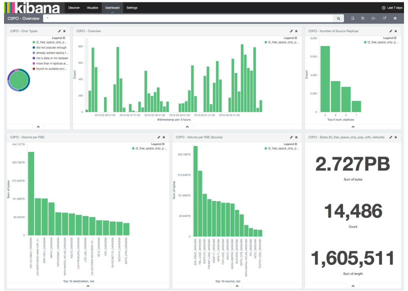


Overall ATLAS T1 tape bandwidth estimates for Run4

Dynamic data handling

Data handling in ADC is driven by two major directions

Direction 1 Physics needs, experiment agreements, processing requirements, and MoUs
Direction 2 Operational and infrastructural constraints

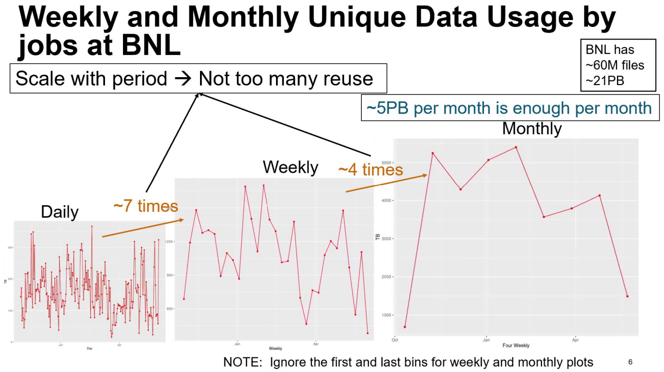


Objective

Prepare a **clear description** of the **current data flow deficiencies**

If there are any, then investigate how to

- Reduce workload execution time
- Reduce data throughput and access latency
- Make *better* use of available storage



Proposed mechanisms

- Revise **initial data placement** algorithms
- Revise **data deletion** and **lifetime models** algorithms
- Revise **data rebalancing**
- Revise data flow orchestration with **subscriptions and rules**
- Development of the **new algorithms** and software if necessary
- Compare with data balancing **strategies of cloud vendors**
- Understand the **benefits and costs of caching**

CERN

ATLAS
EXPERIMENT