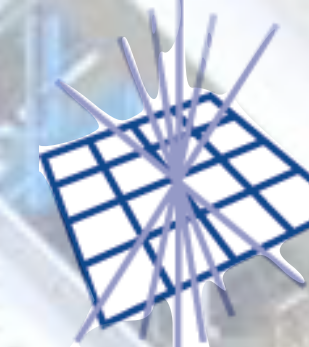




Science and
Technology
Facilities Council

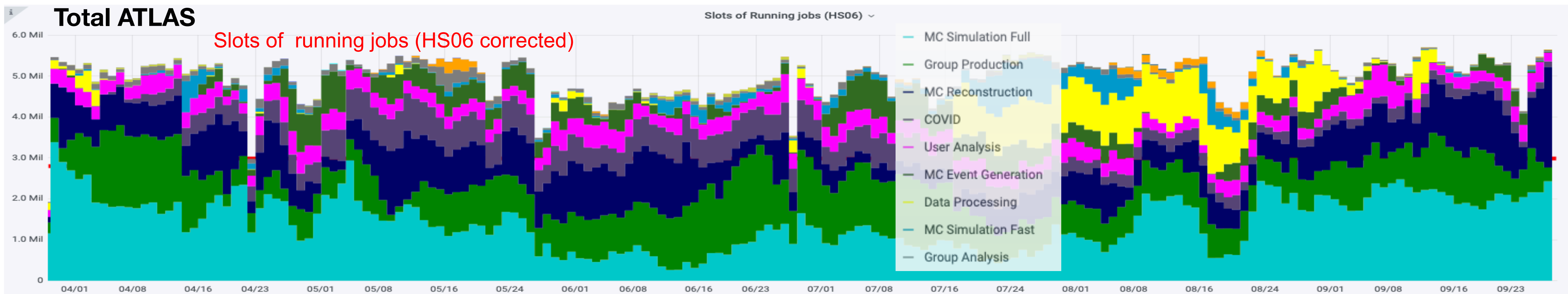


GridPP
UK Computing for Particle Physics

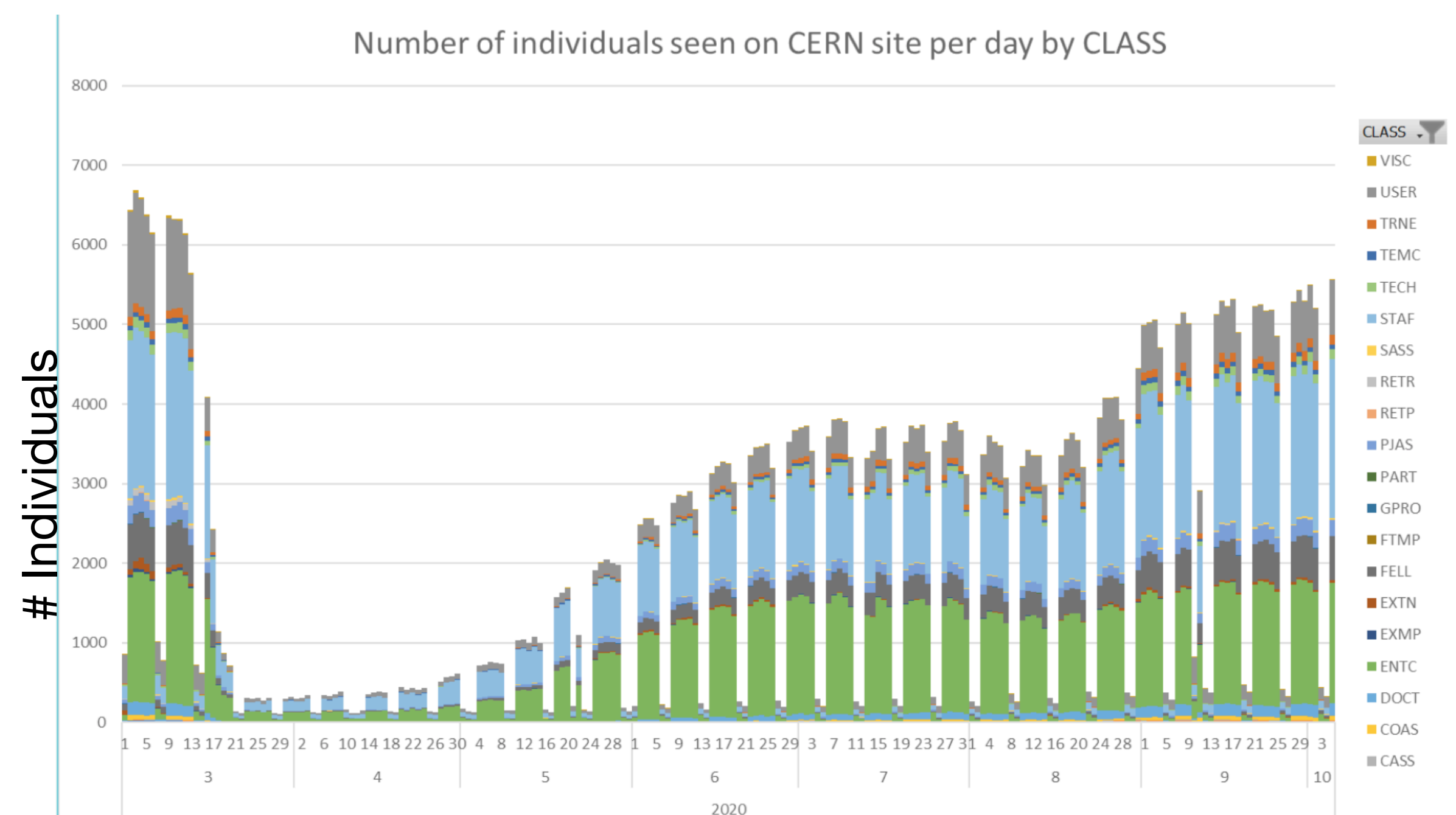
ATLAS Operations during COVID

James Walder, RAL PPD, STFC
With inputs from many others

Overview during COVID



- ATLAS computing already highly distributed:
 - ATLAS Computing success (now more than ever) relies on the ability of Tier-1/2/3 sites to continue to operate under difficult circumstances
 - CERN-based personnel move to 'remote working'; expert shifts allowed to be remote.
 - All meetings / conferences / topical weeks -> online (e.g ATLAS Software and computing weeks)
- New developments:
 - Folding@Home–COVID leveraging newly developed tools in ATLAS: Containerised jobs
- Other developments continue to progress, which are needed for Run-3: e.g.
 - Data Carousel
 - Third-party-copy (via HTTP)
- Research effort impacted;
 - fewer new-development ideas (quantified?);
 - less possibility for off-topic / social chats:
 - ensuring operational work prioritised / continues, with additional (home / covid) constraints / responsibilities.
- Access to CERN site originally highly restricted; now (currently) granted to all classes of personnel; With a number of exceptions, procedures and requirements. CERN guidelines (i.e. don't take my word for it).
(CERN account required to access links)



Communications / Remote working

- HEP contains many meetings / workshops / conferences / tutorials.
 - Community already online in many instances; however now forced to Online as only option, initially several conferences cancelled, now even larger Conferences (e.g. ICHEP) be attempted in online format.
 - More conferences / workshops move online; with reduced (or no) sign-up fees.
 - Also see move to publishing talks via (e.g.) youtube (cf. HTCondor workshops / HSF tutorials).
- ATLAS experts (CERN-based) typically had face-to-face meetings; learn to manage online
 - Democratisation in previously joint physical/online events ?
- Timezones:
 - Optimal time appears to converge around ~ 1600–1800 Cern time (Morning US, Afternoon Europe);
 - not great for Asia-pacific region (2300 Tokyo, 0100 Sydney)
- Video conferencing software:
 - Vidyo was the existing standard (for CERN);
 - Struggled in the initial period
 - Zoom quickly replacing Vidyo (despite some security concerns).
Largely stable; can support large events; more advanced feature set (e.g raise hand features) to Vidyo
 - Certain tenancies also include live-captioning.
- Online chat: Proliferation of usage: Skype, Mattermost, Slack, Discord, Google Notes, Microsoft Teams, ...
 - Appears to augment online meetings. Can it replace face-to-face chats?

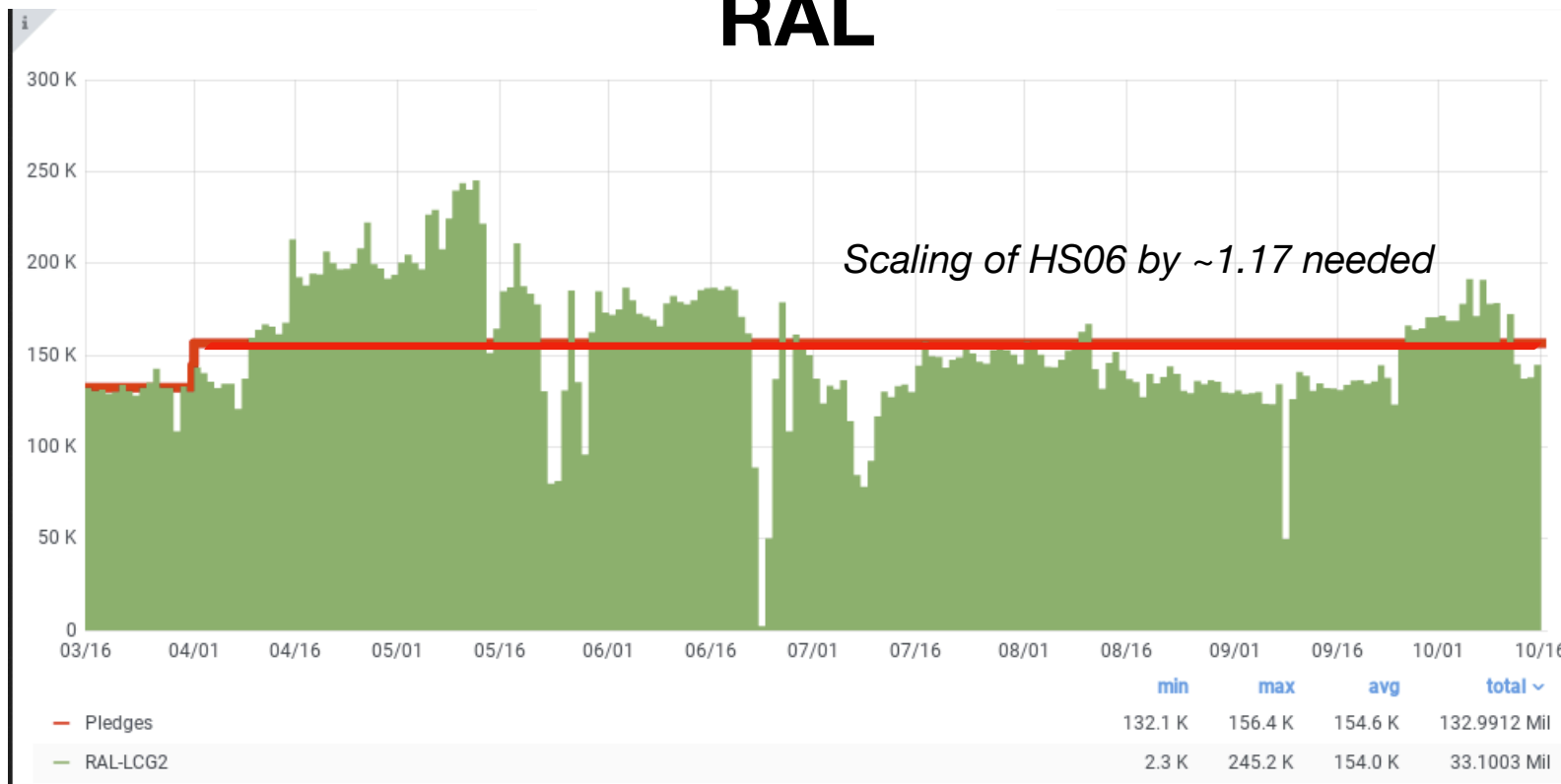
Distributed Computing: T1 & T2s

- Experiment relies on the operational ability of T1 and T2 sites (see previous talks):
 - All but 'emergency access' for significant time at most sites.
 - Each site with own unique set of circumstances / situations
 - Some delays in recovering from immediate issues.
 - e.g. RAL site-wide power cut; Only limited personal on site; Recovered well on critical services. Job slots (mostly) back within ~ 24hr.
 - Two-person operations (still?) not possible in places.
 - Number of Air Con.-related issues. Hot summer, but can COVID restrictions be attributable?
- UK Data loss:
 - DPM disk servers at: Glasgow, Manchester, Lancaster with disk server failures (some recovery possible, with significant manual effort, made more difficult with remote access)
 - Hard (for me) to estimate exact impact of COVID to these events. Restricted access certainly made situations more challenging.
- New site storage: Glasgow commissioning of CEPH (and almost farewell to DPM); not easy with restricted access.
- Is the real impact of COVID on sites still to come ?:
 - How does such limited hands-on access to machine rooms affect long-term site performance ?
 - Replacements of old disks (ability to obtain spares ?)
 - upgrades of software (e.g. delays of planned activities);
 - finalising move to centos 7.
 - Needed networking improvements / hardware infrastructure change.
- However - Should consider a such operational achievement as a significant success.
- **(Improved / simplified) access to machine rooms now essential to allow for more than just basic 'fire-fighting' activities (as much as restrictions allow).**

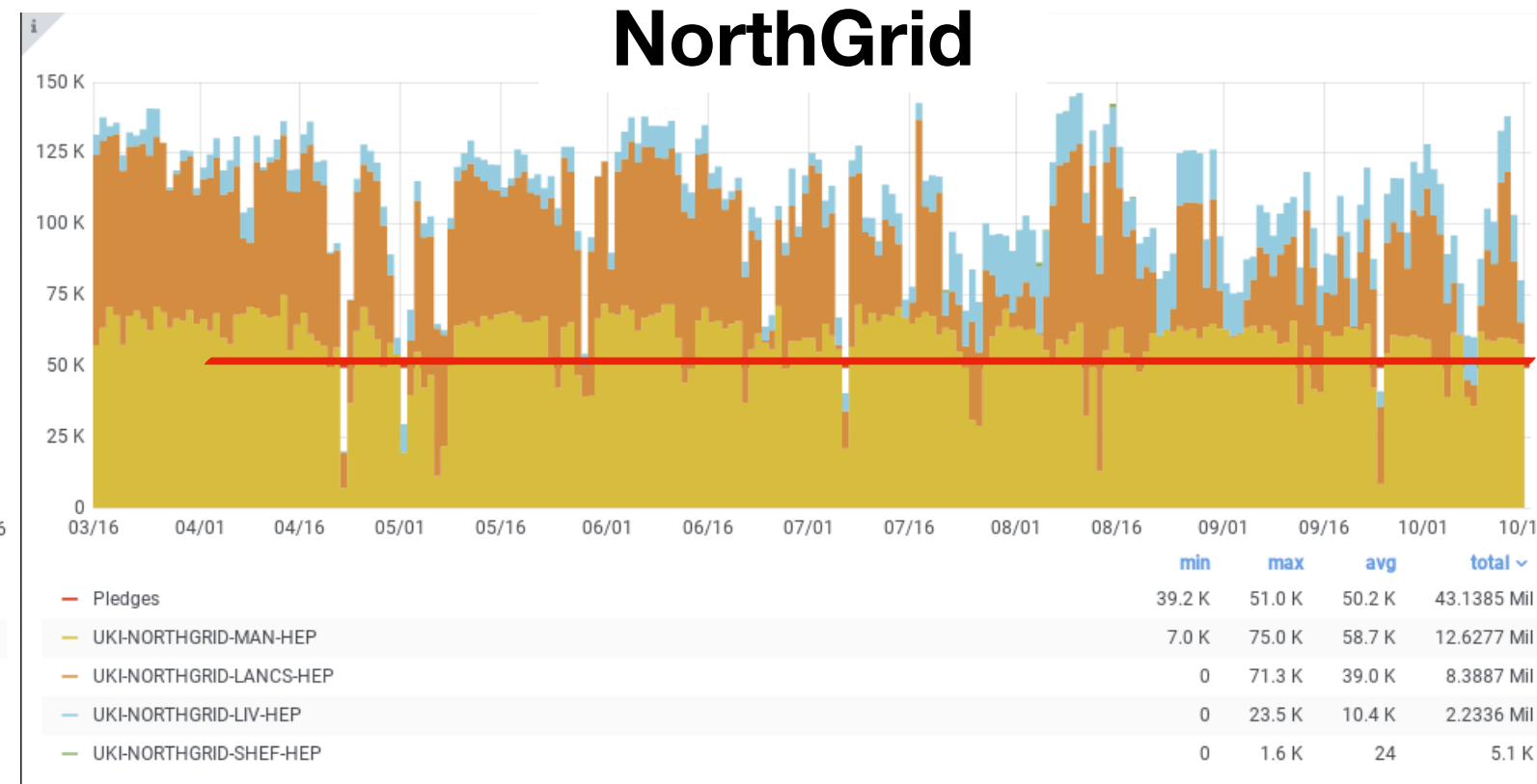
UK job accounting

- UK federations meeting (and exceeding) pledge.
- UK contributed 12% (excluding CERN) of ATLAS CPU resources

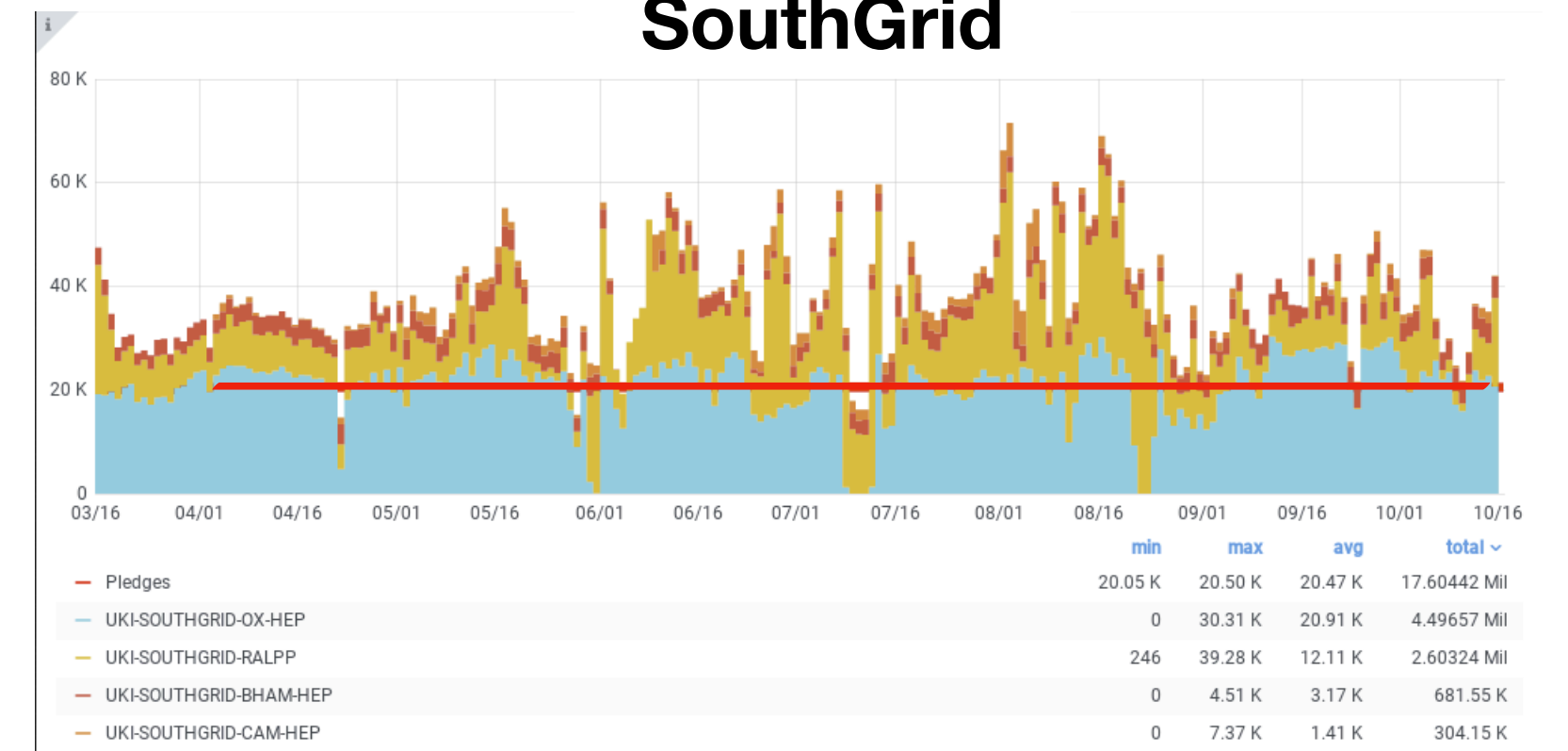
RAL



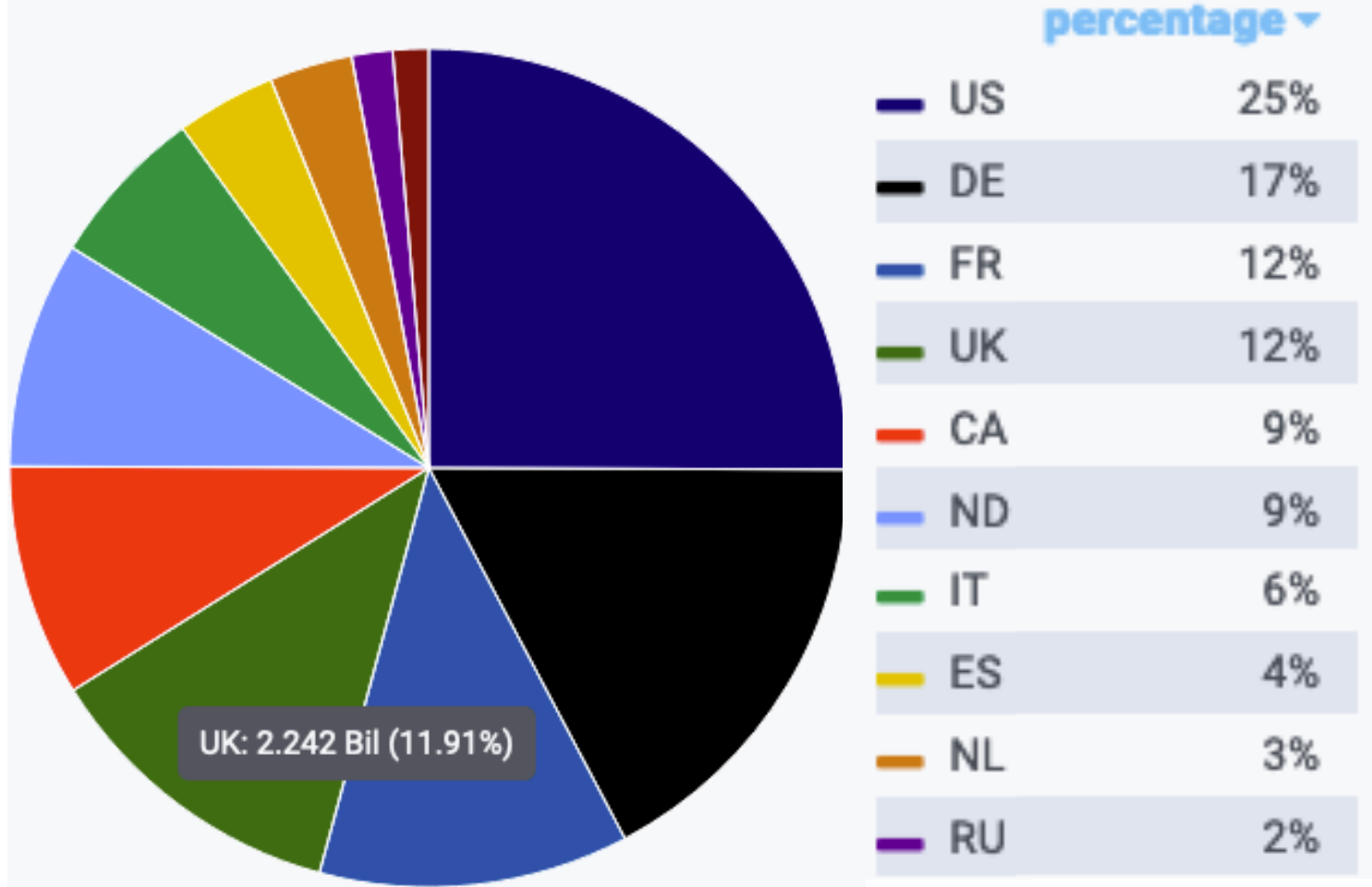
NorthGrid



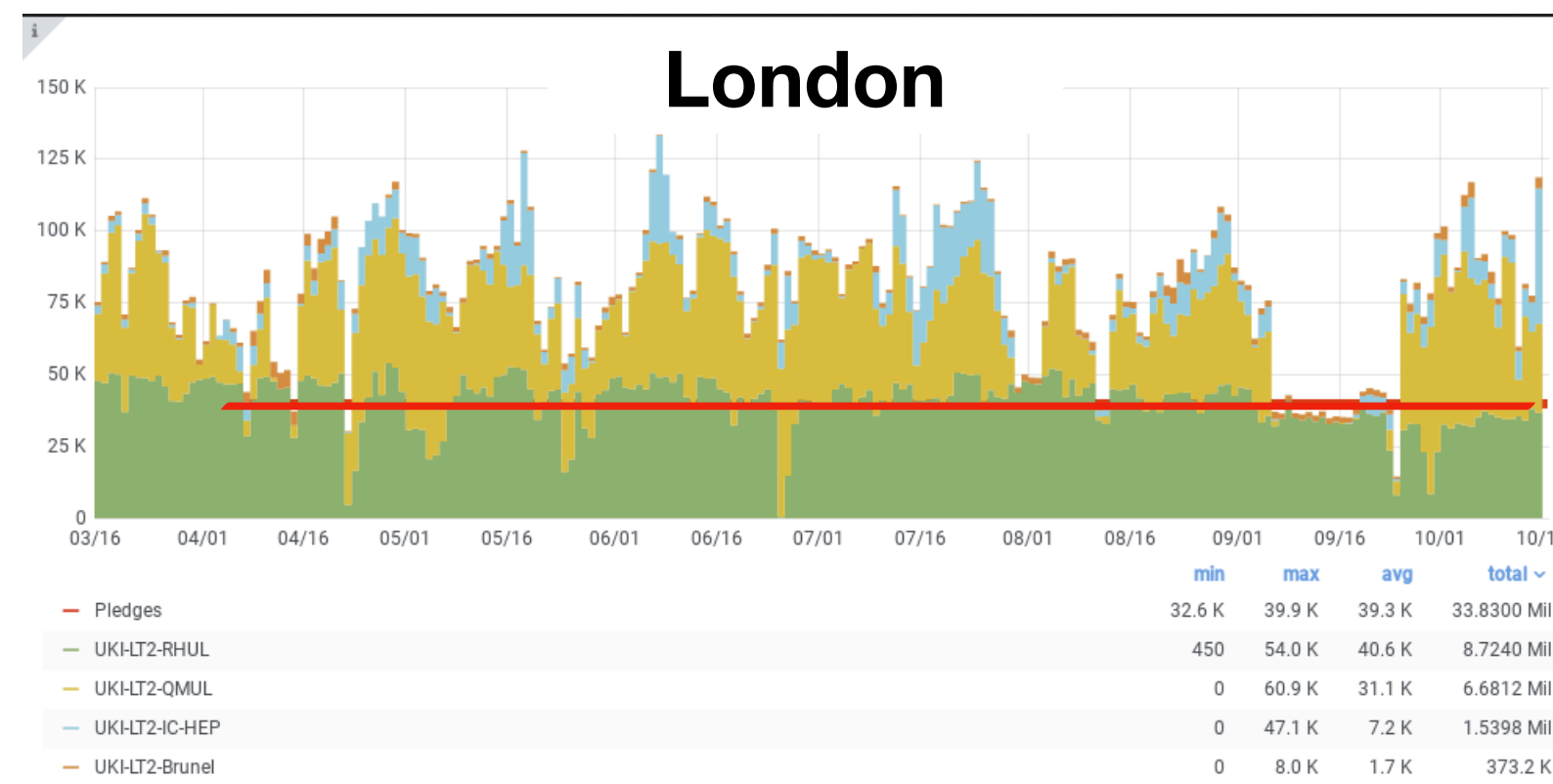
SouthGrid



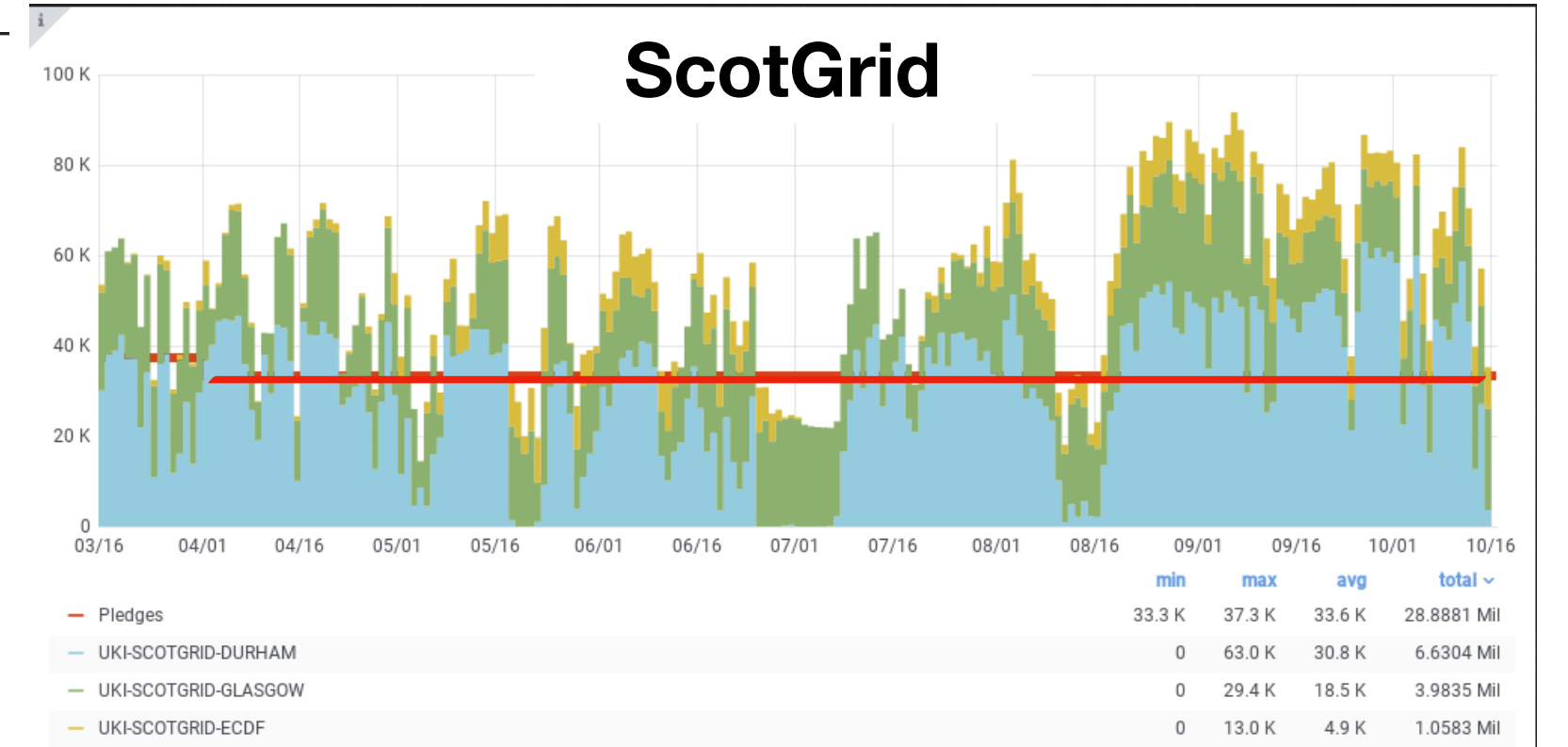
Slots of Running jobs (HS06)



London

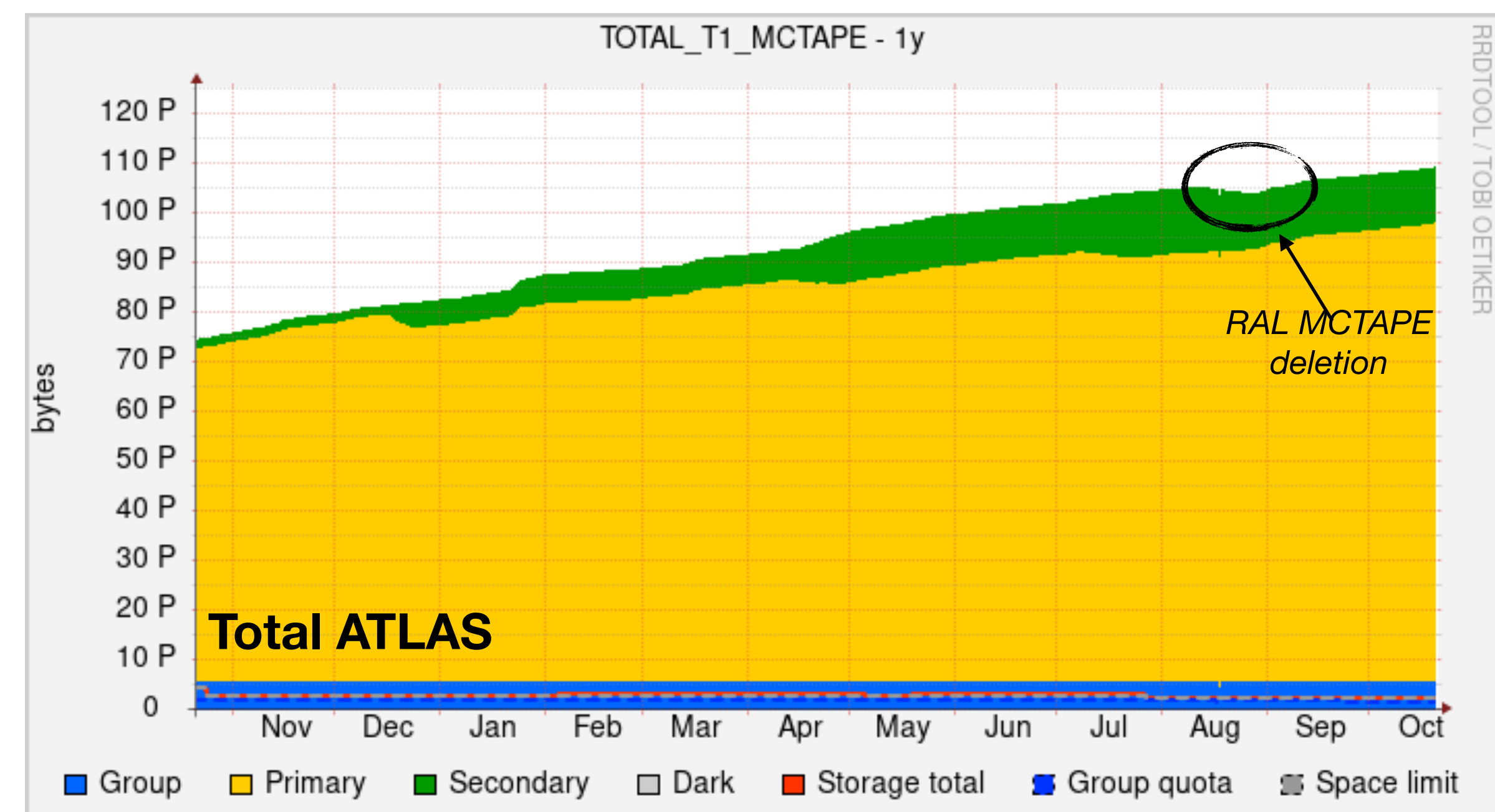
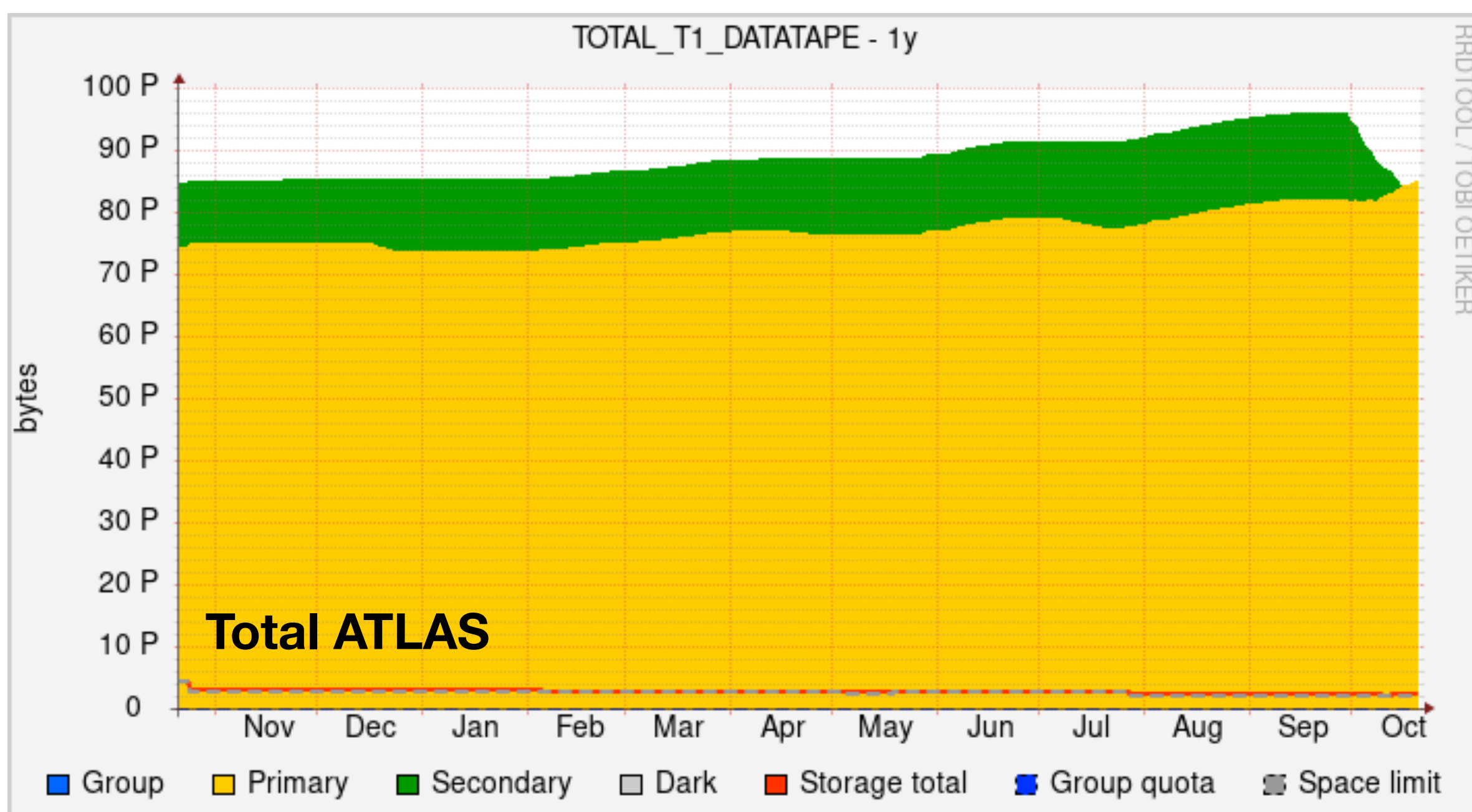
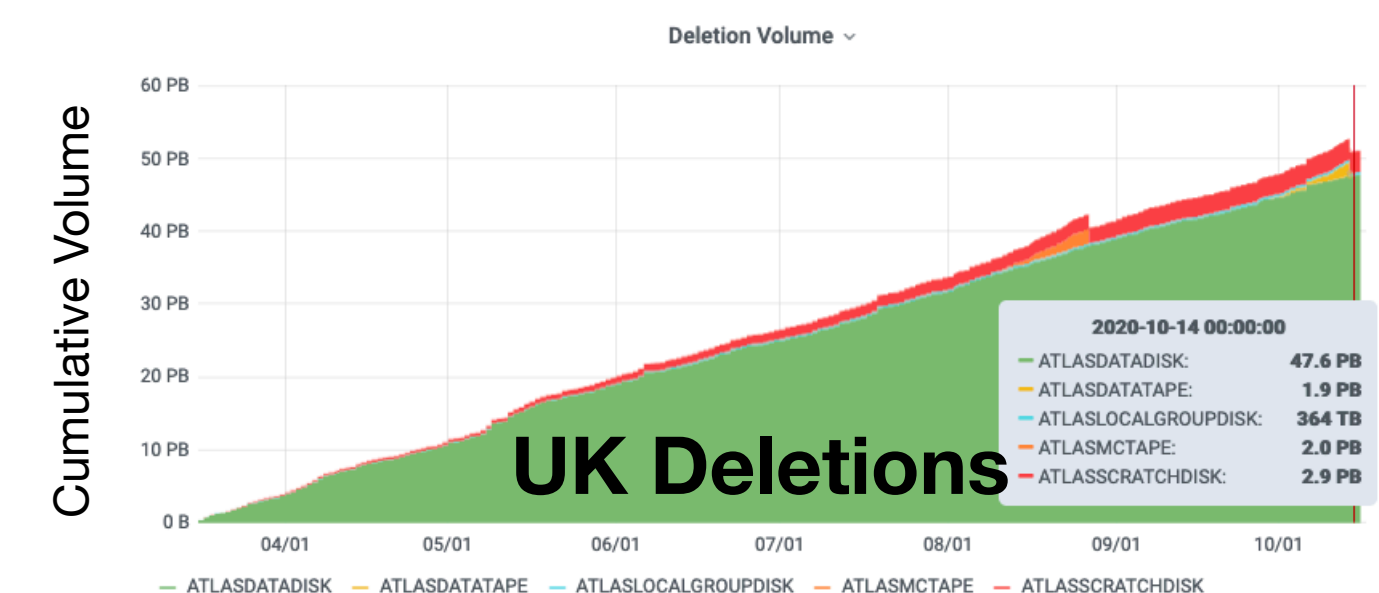
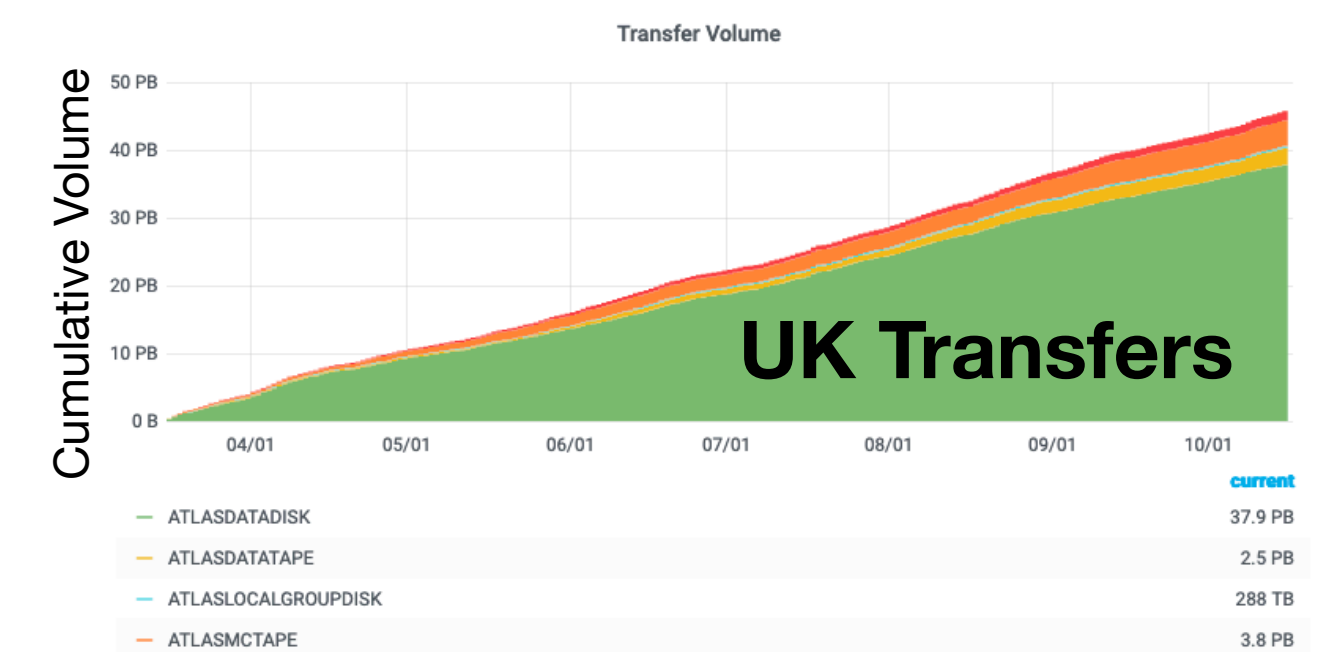
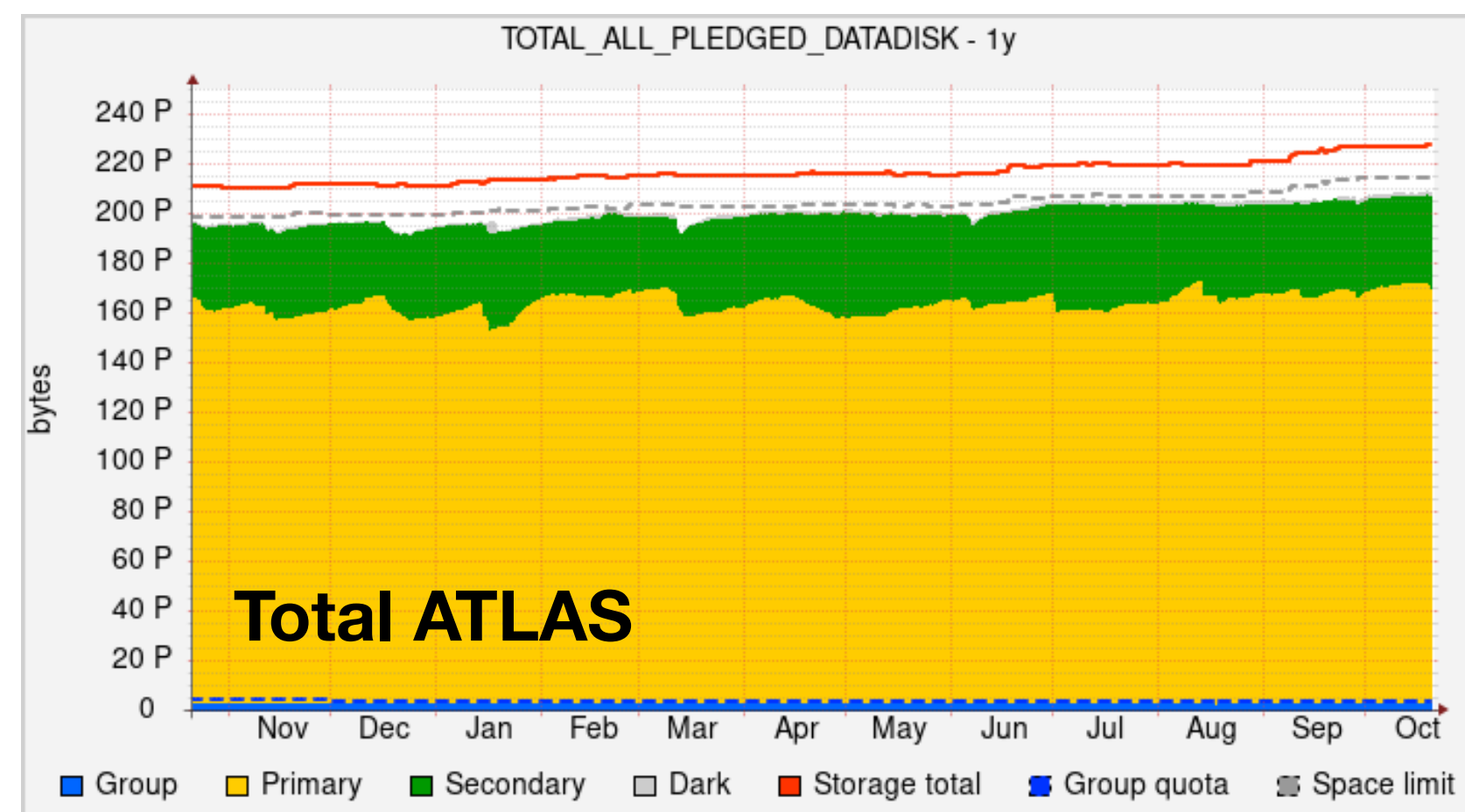


ScotGrid



Disk and Tape usage

- Disk and Tape usage will remain high
 - Regular deletion campaigns for data on disk:
 - Old or superseded datasets
 - High turnover: O(50)PB deleted, O(46) PB transferred and staged to UK in ~6months
- Removal of secondary data on DATATAPE performed recently
- Further removal from MCTAPE in Nov. (RAL first round in Aug.)
- Repack campaign at RAL for ATLAS (Tape) data started (LHCb complete).



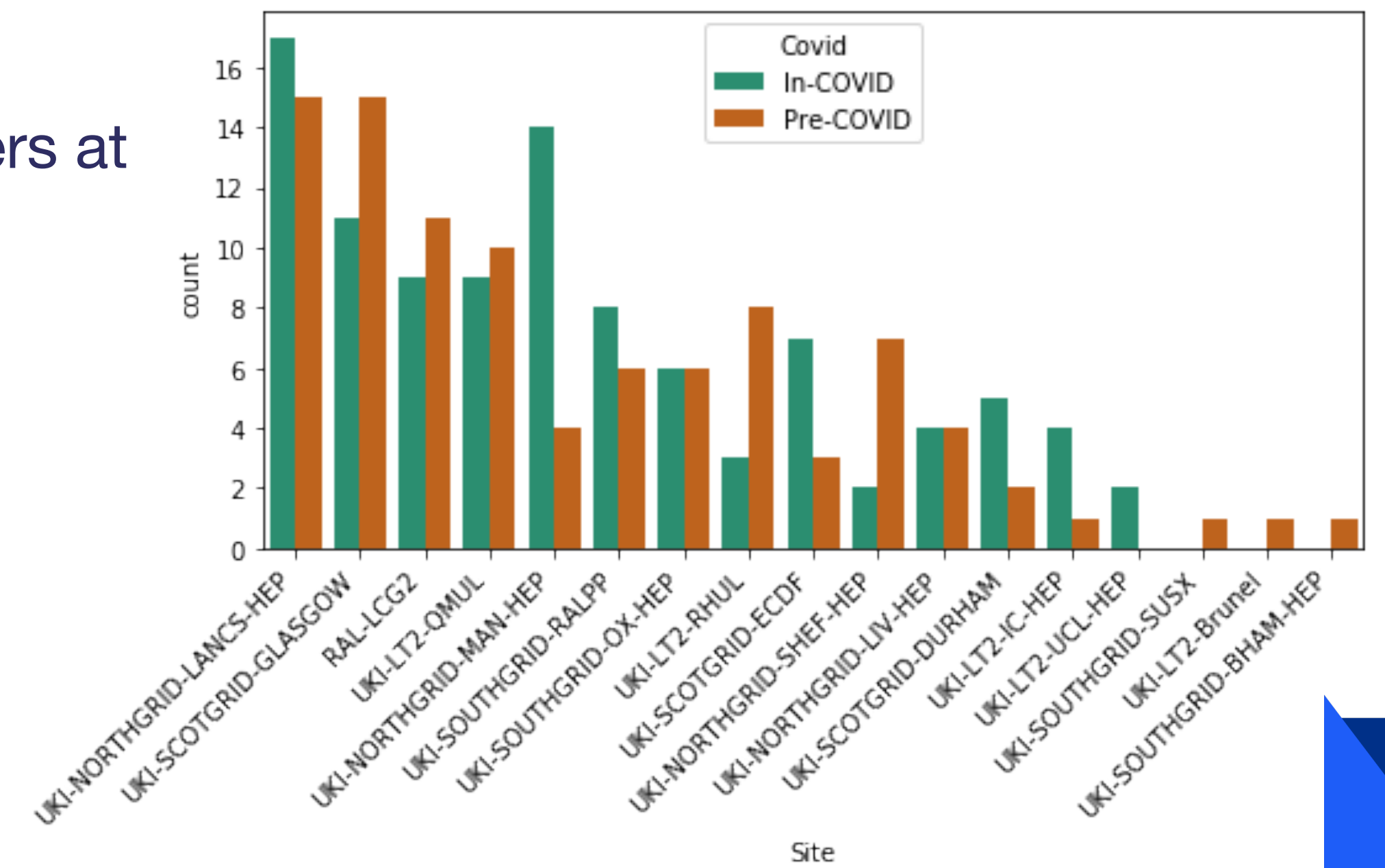
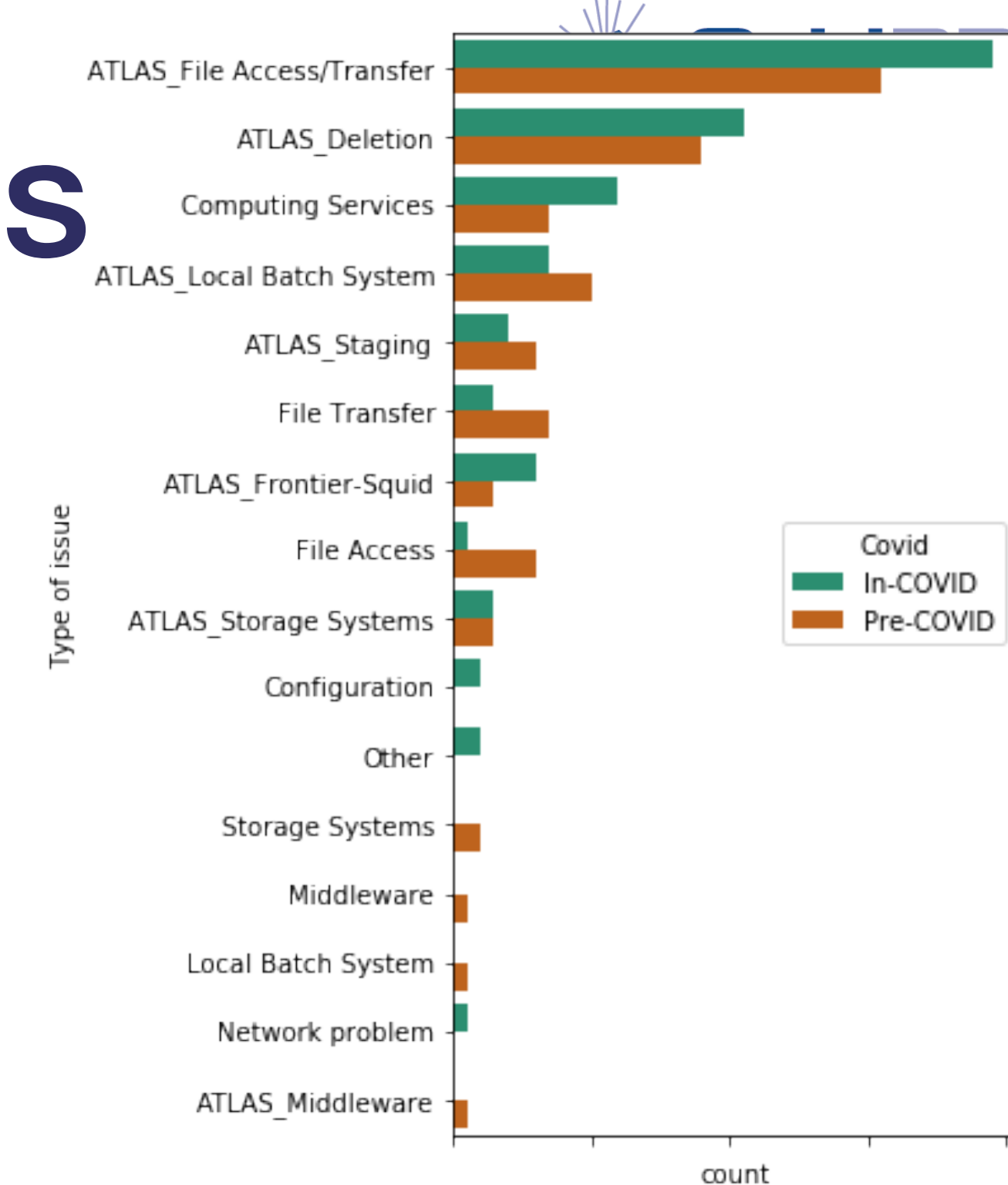
UK response to operational issues

- Quick study using GGUS tickets as metric.
- Time period: 6 Months: 16th March – 16th Sep ('Covid').
 - Compare to previous 6 months ('Pre-Covid').
- 97 (96) tickets opened within covid (pre-covid) periods,
- ~17 days median time to close (includes if ticket reopened)
 - Usually includes time to confirm that problem is solved, rather than when action taken.

	Time to Solution[d] Pre-covid	Covid
25%	15.3	15.2
50%	17.8	16.5
75%	21.4	21.4

- Broadly similar reported issue types, and (mostly) similar numbers at a given site.
- Storage, transfer (and deletion) issues dominate.
- Four long-standing tickets – impacted by COVID, or on-hold due to broader set of changes / issues.

- [GGUS Link](#)



Workflow changes: Data Carousel

- ATLAS begins to move to new analysis model format, and Data Carousel (i.e. no tape pre-staging) processing mode.
 - Run-2:
 - Primary AOD, stored on Disk + TAPE
 - O(100) Derived-AODs ~ 1% of AOD size for analysis (produced and replaced as analyses developed)
 - Run-3 (+ Run-2 reprocessing)
 - AOD, Stored on TAPE (x2 for data), short lifetime on disk.
 - Should release ~ 25PB Disk
 - 2 new derived formats DAOD_PHYS(LITE); should accommodate ~ 80% of use-cases
 - Replaces majority of existing DAODs.
 - Derivation reprocessings to become somewhat more scheduled.
 - Will take some time to fully transition across.
- Number of changes to ATLAS Workflow for TAPE usage:
 - No tape pre-staging campaigns;
 - ProdSys2 (WFM system) decides on staging requests; fine grained communication back from Rucio
 - iDDS (intelligent Data Delivery Service) Submit jobs to already staged data, without waiting for whole dataset / task to complete
 - Site staging profile: ATLAS Production decides how many files to stage at each site:
 - Site (via CRIC configuration) determines the staging profile, i.e. number of concurrent staging requests in a batch, and how batches are submitted.

*By ‘**data carousel**’, we mean an orchestration between workflow management (WFMS), data management (DDM/Rucio), and data transfer (FTS) and data archiving services whereby a bulk production campaign with its inputs resident on tape, is executed by staging and promptly processing a sliding window of X% (5%?, 10%?) of inputs onto buffer disk, such that only a small fraction of inputs are pinned on disk at any given time.*

OSG All Hands Meeting

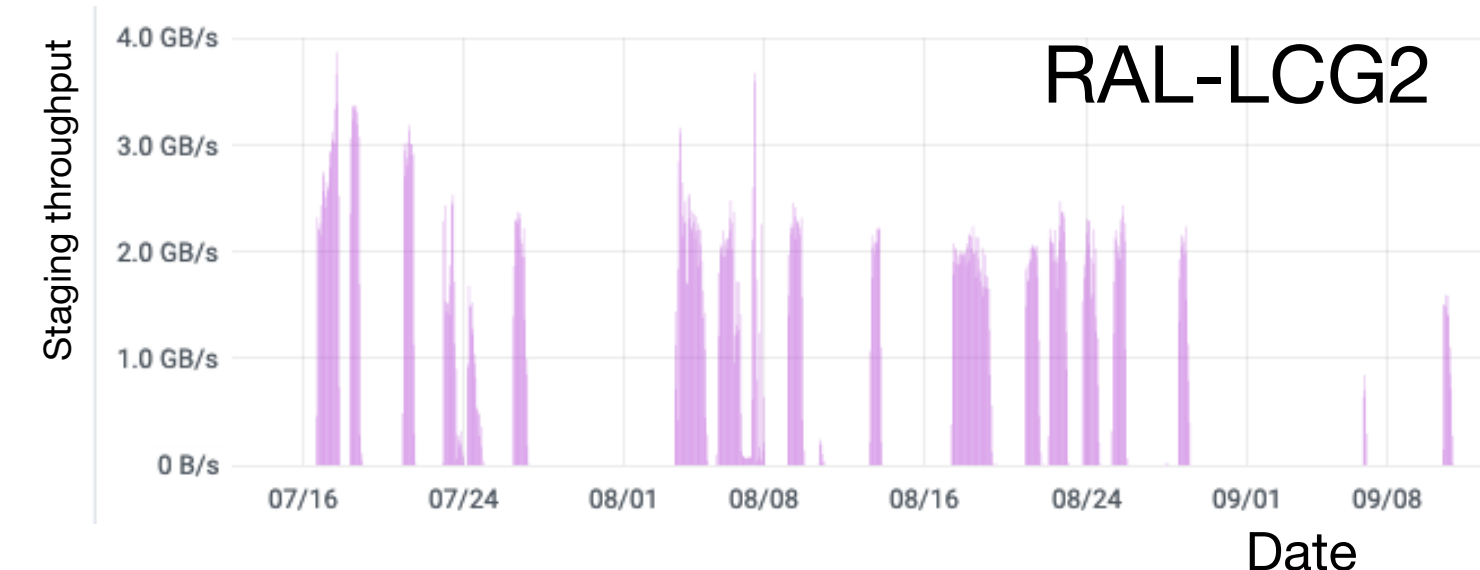
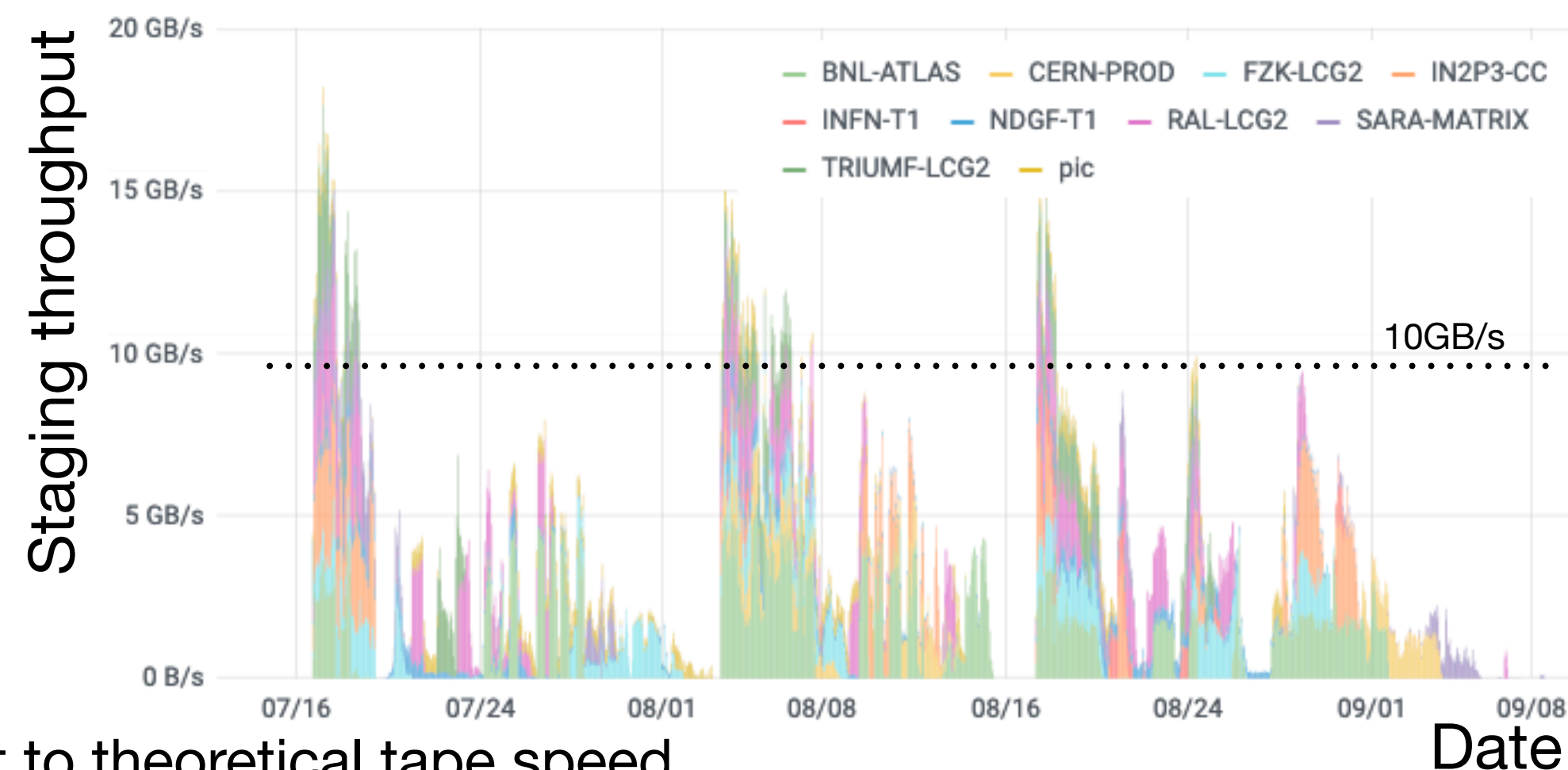
Performance: Data Carousel

- Complete Run-2 RAW (~18.5 PB) reprocessed over summer period,
 - starting ~now: AOD -> DAOD derivation production ('real world' test) using Data Carousel (for data not on Disk).

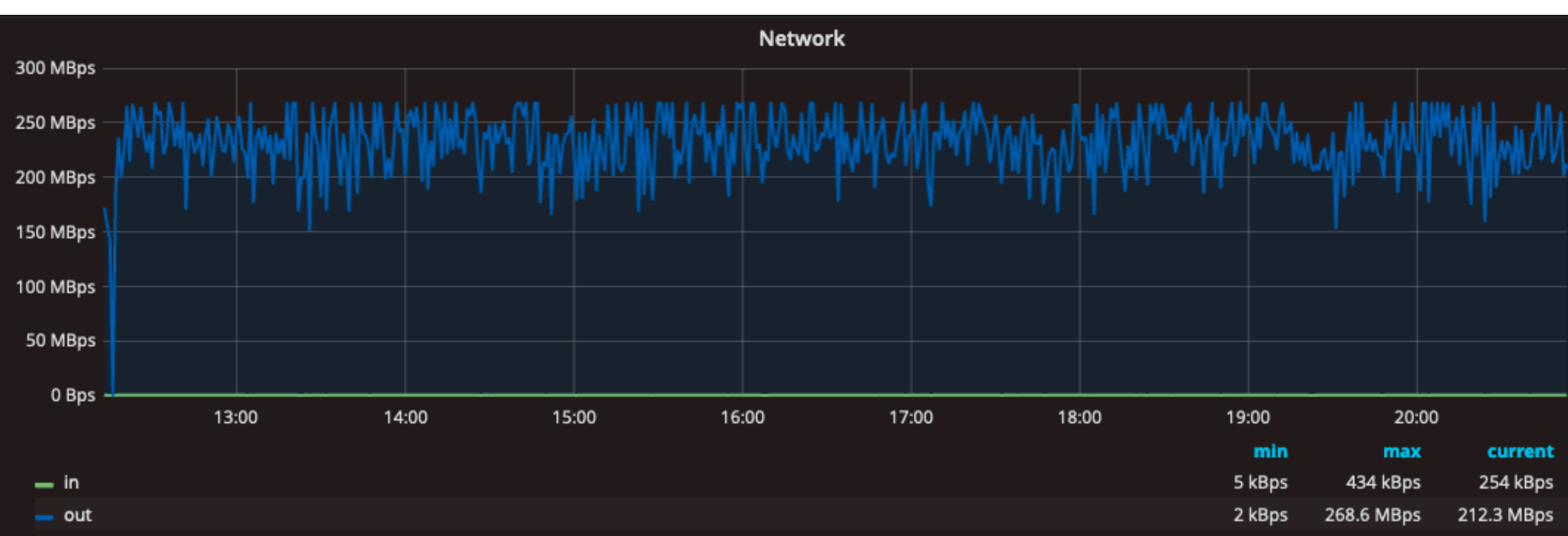
- Total T1 staging reached stable ~ 15GB/s

- Recall efficiency typically 30 – 40%

- Defined as ratio of measured tape throughput to theoretical tape speed.
- RAL managed to demonstrate near 100% (no detailed study performed) using the old robot.
 - Possibly related to initial data placement strategies across the T1s.

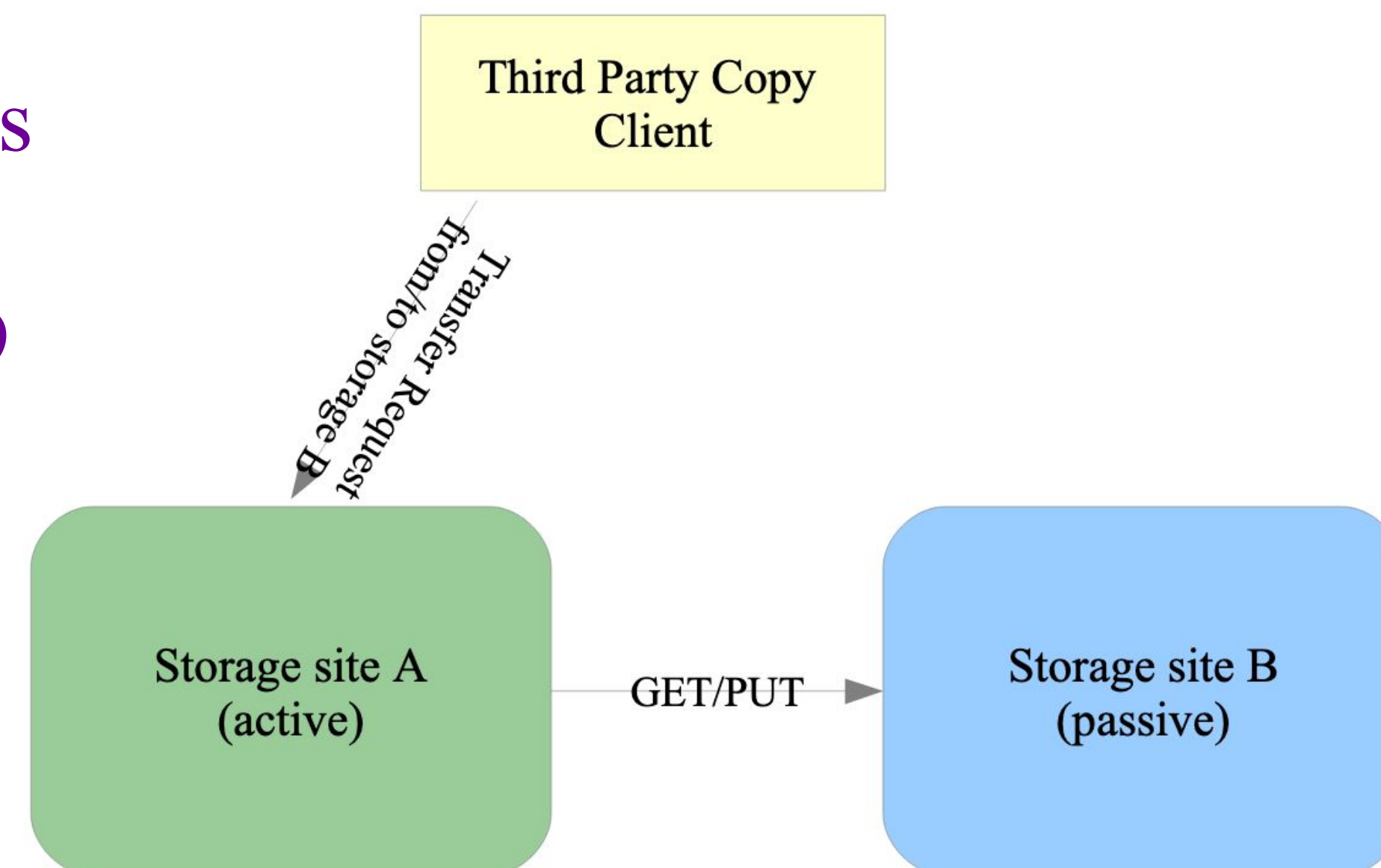


Sites	Stable Rucio throughput in Phase I test (2018)	Stable Rucio throughput in Run2 reprocessing campaign (2020)
CERN	2GB/s (CTA test instance)	4.3 GB/s (DDM link)
BNL	866MB/s	3.4 GB/s (DDM link)
FZK	300MB/s	1.6 GB/s (DDM link)
INFN	300MB/s	1.1 GB/s (DDM link)
PIC	380MB/s	540 MB/s (DDM link)
TRIUMF	1GB/s	1.6 GB/s (DDM link)
CCIN2P3	3GB/s	3 GB/s (DDM link)
SARA-NIKHEF	640MB/s	1.1 GB/s (DDM link)
RAL	2GB/s	2 GB/s (DDM link)
NDGF	500MB/s	600 MB/s (DDM link)



TPC: How HTTP-TPC works

- Client uses x509 to request macaroon tokens for the TURLs from both storages
- Macaroon tokens are used by the storages to authorize the transfer between themselves
- In HTTP-TPC transfer can be in push or pull mode
- In production restricted to pull
 - Sites added to the production system as **active destinations** pulling data from passive sources
 - Tokens only for read access
- Now macaroons
 - In future JWT and no x509

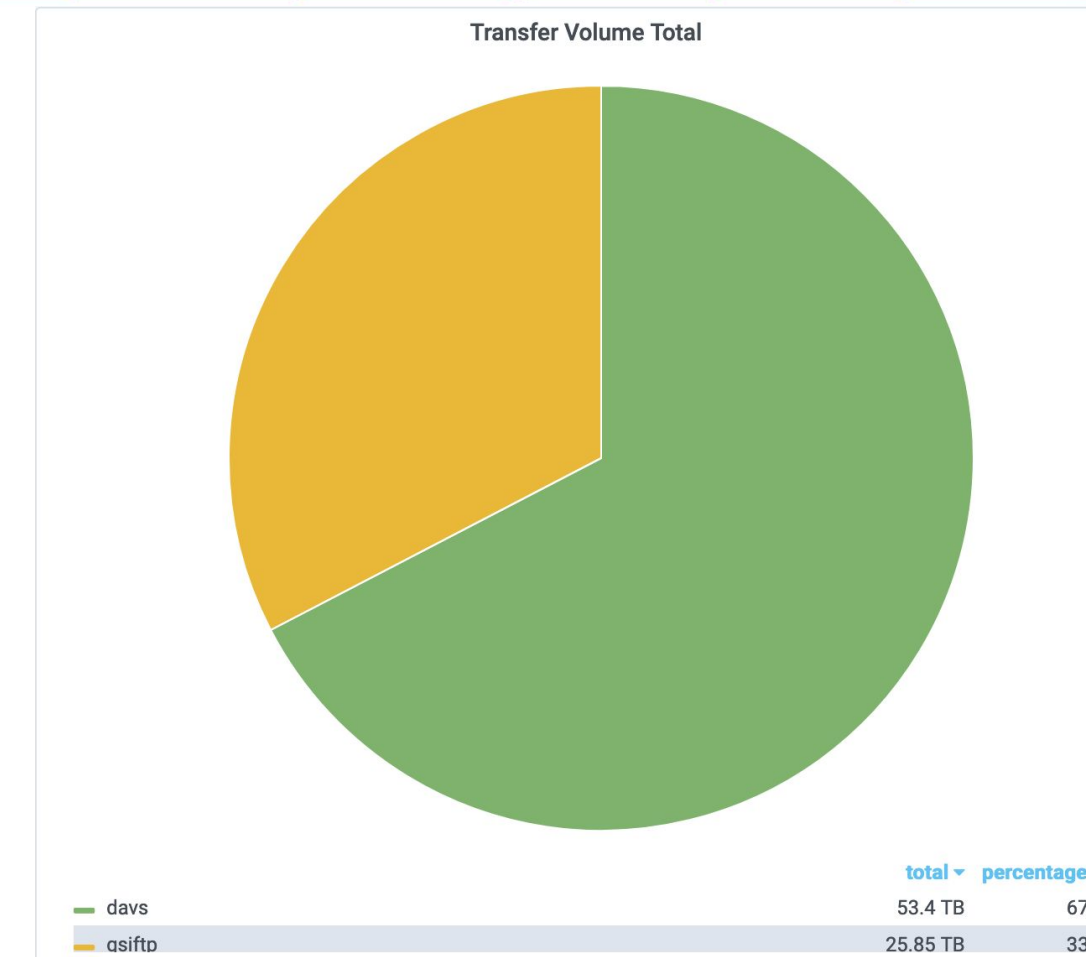


TPC in ATLAS

- ATLAS is enabling HTTP-TPC on all grid sites
 - Plan to have most grid sites enabled by **May 2021**
 - gridftp will remain as secondary protocol
 - Several months before run3 to fine tune with possible fall back
 - Started work to integrate tape end points too

Efficiency ▾												
	CA	CERN	DE	ES	FR	IT	ND	NL	RU	TW	UK	US
UKI-LT2-Brunel	88%	92%	84%	85%	81%	91%	84%	94%	84%	72%	90%	85%
UKI-NORTHGRID-LANCS-HEP	100%	100%	94%	100%	94%	100%	99%	100%	100%	98%	96%	99%
UKI-NORTHGRID-LIV-HEP	99%	99%	96%	100%	96%	96%	100%	100%	100%	97%	99%	100%
UKI-NORTHGRID-MAN-HEP	100%	99%	88%	100%	96%	92%	98%	100%	100%	96%	95%	99%
UKI-SCOTGRID-ECDF	79%	83%	93%	77%	94%	99%	67%	99%	98%	98%	97%	56%

- 21 sites enabled so far
 - 5 UK sites all DPM
 - ~65% of their transferred data over HTTP-TPC
- WLCG Ops opening tickets for **required upgrade**



Tests for TPC

- Smoke tests (atomic curl tests)
 - Smoke tests can be used by sites to do the initial tests

wget <https://raw.githubusercontent.com/paulmillar/http-tpc-utils/master/bin/smoke-test.sh>

chmod 755 ./smoke-test.sh

voms-proxy-init -voms dteam

./smoke-test.sh https://<storage_head>:<port><dpm_path>/dteam

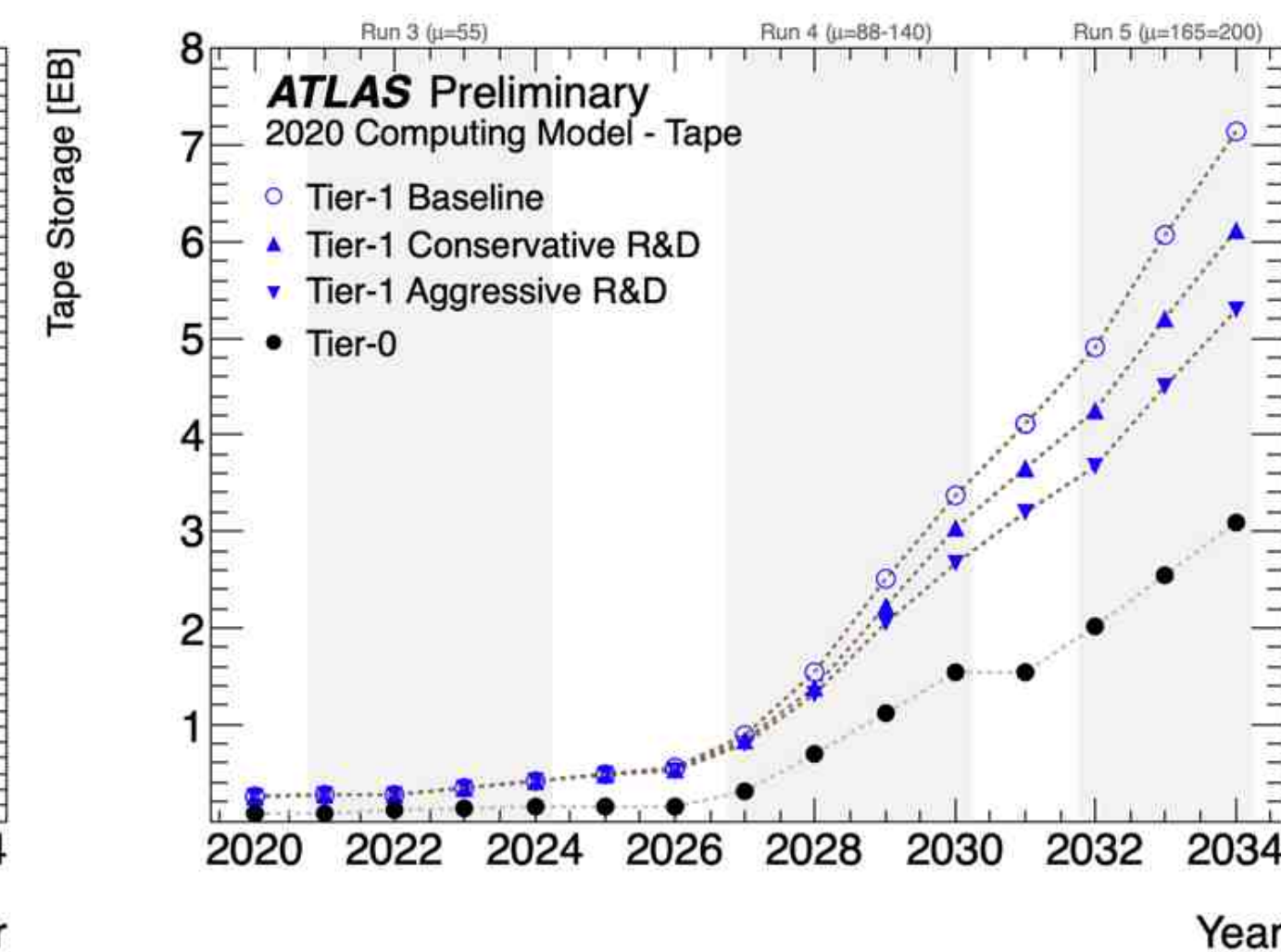
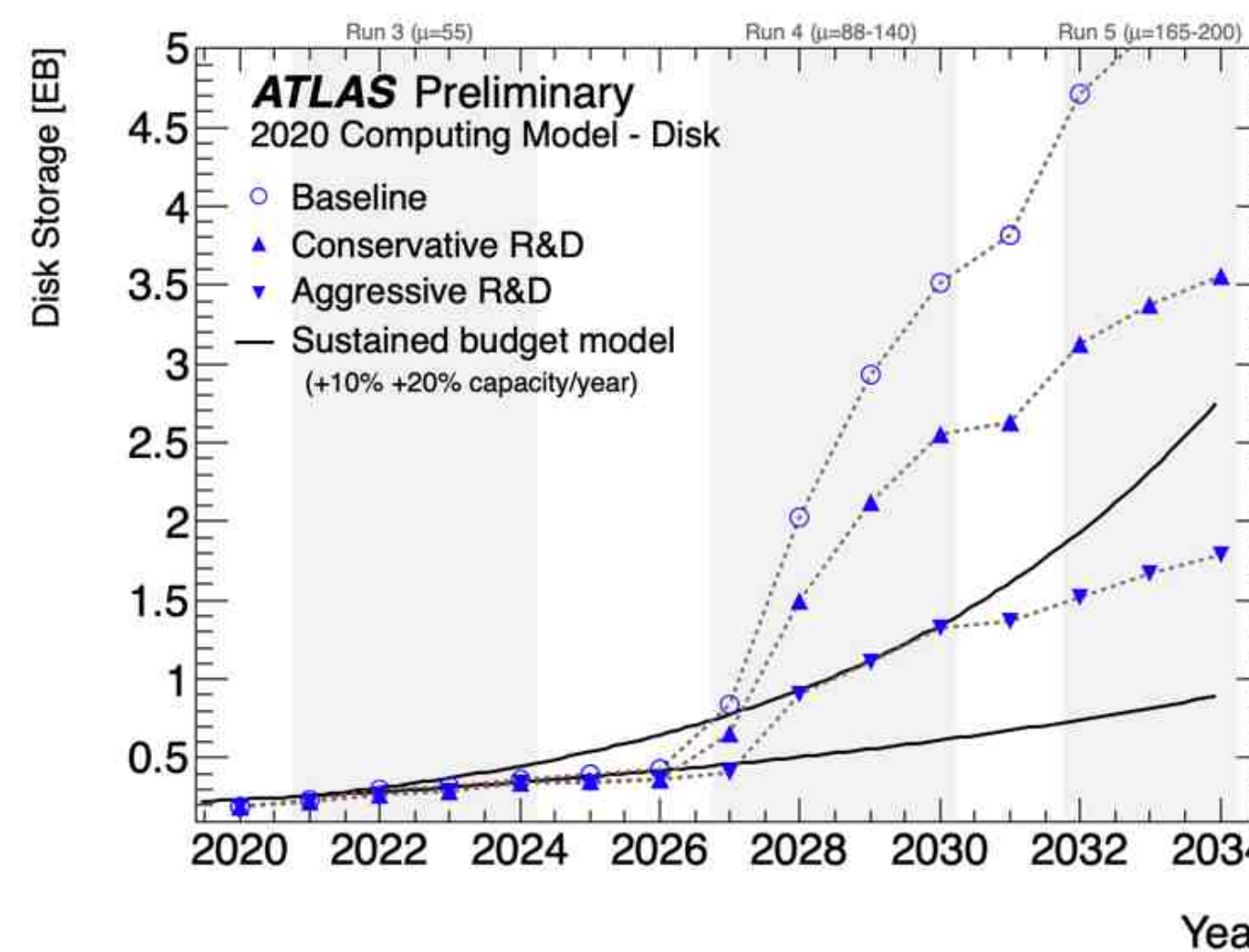
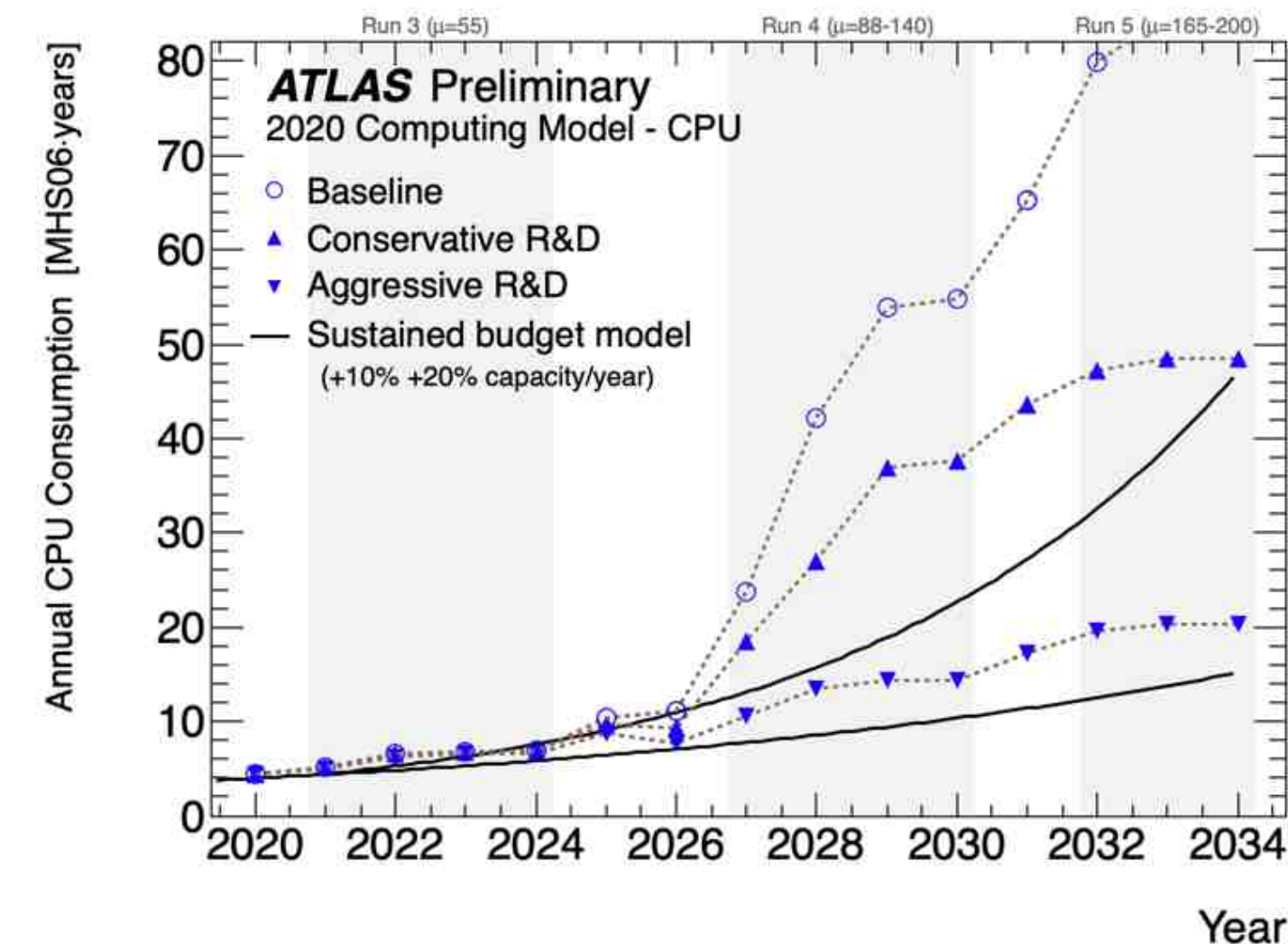
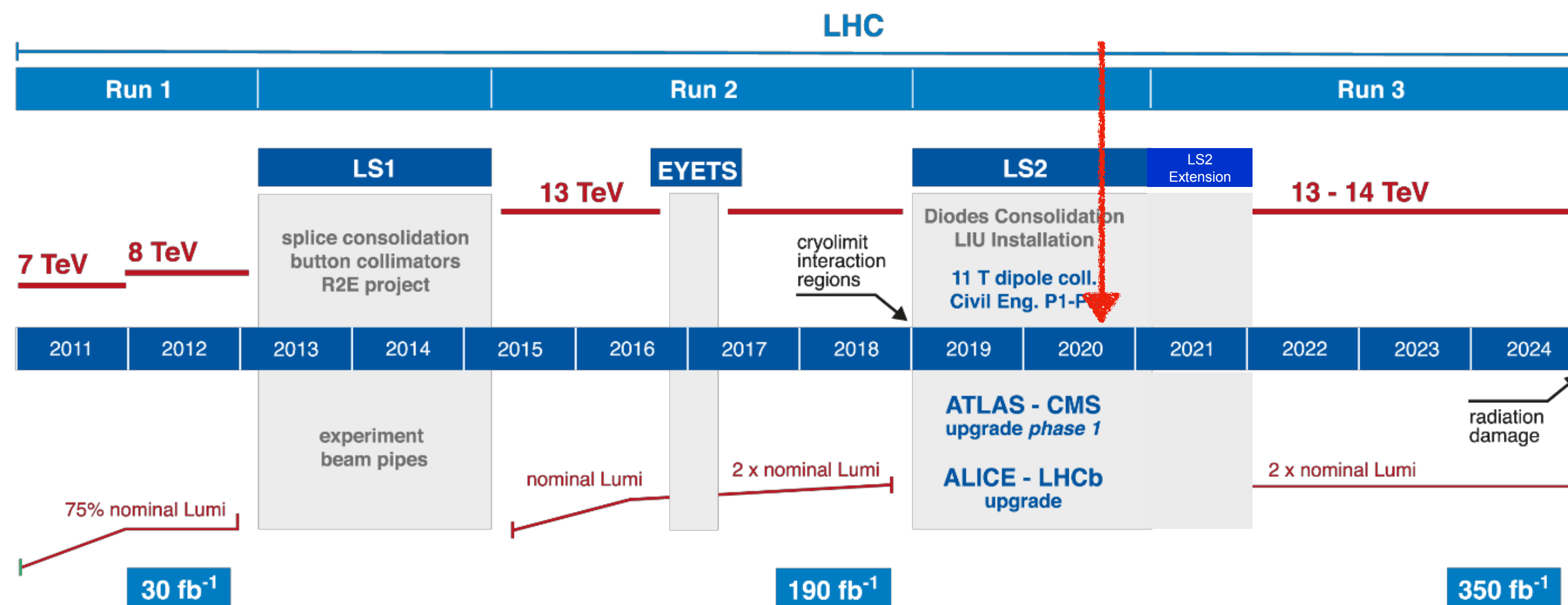
- They work also with atlas credentials
 - -v atlas and appropriate path
- ATLAS functional tests (whole FTS->gfal->davix chain)
 - Falls back on gridsite
 - Some ambiguity -> smoke tests important
 - Need to be enabled -> most sites already are but not all
 - Smoke tests first!



To Run-3, and beyond

Goal: $\Sigma(\text{Run1} + \text{Run2} + \text{Run 3}) > 350 \text{ fb}^{-1}$

Run 3 outlook



- LHC Performance workshop (“Chamonix”, November 2020) postponed in Spring 2021 to review and optimize the LHC-LS2 and Run 3 schedule.
- The largest risk / unknown [K. Jacobs]:
 - The development of the pandemic at CERN, in Europe and worldwide (working conditions at sites, travel situation (restrictions, quarantine))
- Number of detector activities remain on critical path (e.g. New-Small Wheel)
- Full reprocessing of Run-2 Data and MC expected ~ Spring 2021 (using new analysis model and release 22).
- Looking towards HL-LHC;
 - Improvements in CPU reduction / utilisation and code improvements underway
 - Meeting the Storage demands remains a considerable challenge
 - The LHCC review of Computing for run-4 (and preparation) all done online:
 - Only one in-person meeting (mid.-Feb.) to discuss the outline of the report.

Other VOs

- Brief survey of other VOs: (non-exhaustive, and thanks to Katy Ellis, Raja Nandakumar)

- CMS:

- Strong US contingent; Working time extended more into evening (UK time).
- For Tier-1 Issues requiring Inter-departmental inputs, and with increased workloads on other departments, delays / difficulties in impressing the imperative nature of problems.
 - e.g. Firewall settings, affecting off-site reads in CMS jobs.
 - i.e harder to bang on people's office doors ...
- Developments:
 - CMS to use Rucio; migration currently in progress.
 - PhEDEx also on track to be retired by end November.

CMS CPU Eff. [%]	Mar	Apr	May	Jun	Jul	Aug	Sep	Total
CERN-PROD	75.3	74.9	69.9	73.5	71.6	67.9	71.1	72.1
FZK-LCG2	78.3	80.1	76.7	77.4	67.9	80.8	79.2	77.1
IN2P3-CC	83.5	86.5	86.2	82.4	85.0	86.7	79.6	84.3
INFN-T1	83.4	79.1	75.3	79.9	76.1	83.7	81.1	80.1
JINR-T1	75.2	72.7	75.7	79.7	76.1	77.0	73.8	75.8
RAL-LCG2	64.0	59.8	61.1	57.1	56.2	46.1	37.9	55.7
USCMS-FNAL-	78.1	83.5	86.0	79.0	79.6	83.1	78.1	81.0
pic	71.1	81.3	80.2	83.3	81.4	77.3	73.2	78.1

- LHCb:

- Vidyo -> Zoom; Heavy use of Mattermost (for Ops), Slack (for Dirac development); preferred over email.
- 'Experts-on-Call': Previously CERN presence required; moved to remote shifts
- Tunneling software / techniques (vpn, socks proxy) in wide use now
- Operations have continued with some delays:
 - MC simulation continued normally (~80 - 90% of LHCb workload)
 - Operations team (non-on call expert) were regularly remote anyway
 - Applications development has had delays, with some reflection on to validations and operations

- Liaisons working within the Tier-1:

- Move to Slack for communication.
- Lost benefit of in-person discussions / meetings; Sharing screens virtually takes some learning
- More formalised meetings, and reporting.



Summary

- HEP community, and especially computing already highly distributed.
 - CERN experts with biggest change to move away from in-person meetings.
- Video-conference software (while not perfect), has help ensure continued operations.
 - Lack of face-to-face meetings and ‘corridor chats’; how significant?
- Operations continued well, with adaptations:
 - Sites survived, with occasional fire-fighting; long-term impact?
 - Enabling Covid-secure machine room access a necessity.
- Many developments ongoing during past months (did not touch on moving to multi-threaded athenaMT release 22)
 - Use of Grid infrastructure and tools to facilitate F@H:COVID.
 - Data carousel tests -> a production model for ATLAS
 - TPC-http needed across all sites
- Am looking forward to meeting all of you in person at some future event.

