ATLAS & Networking

LHCOPN LHCONE Workshop

24-25 October 2022

David Cameron (University of Oslo)
on behalf of ATLAS Computing

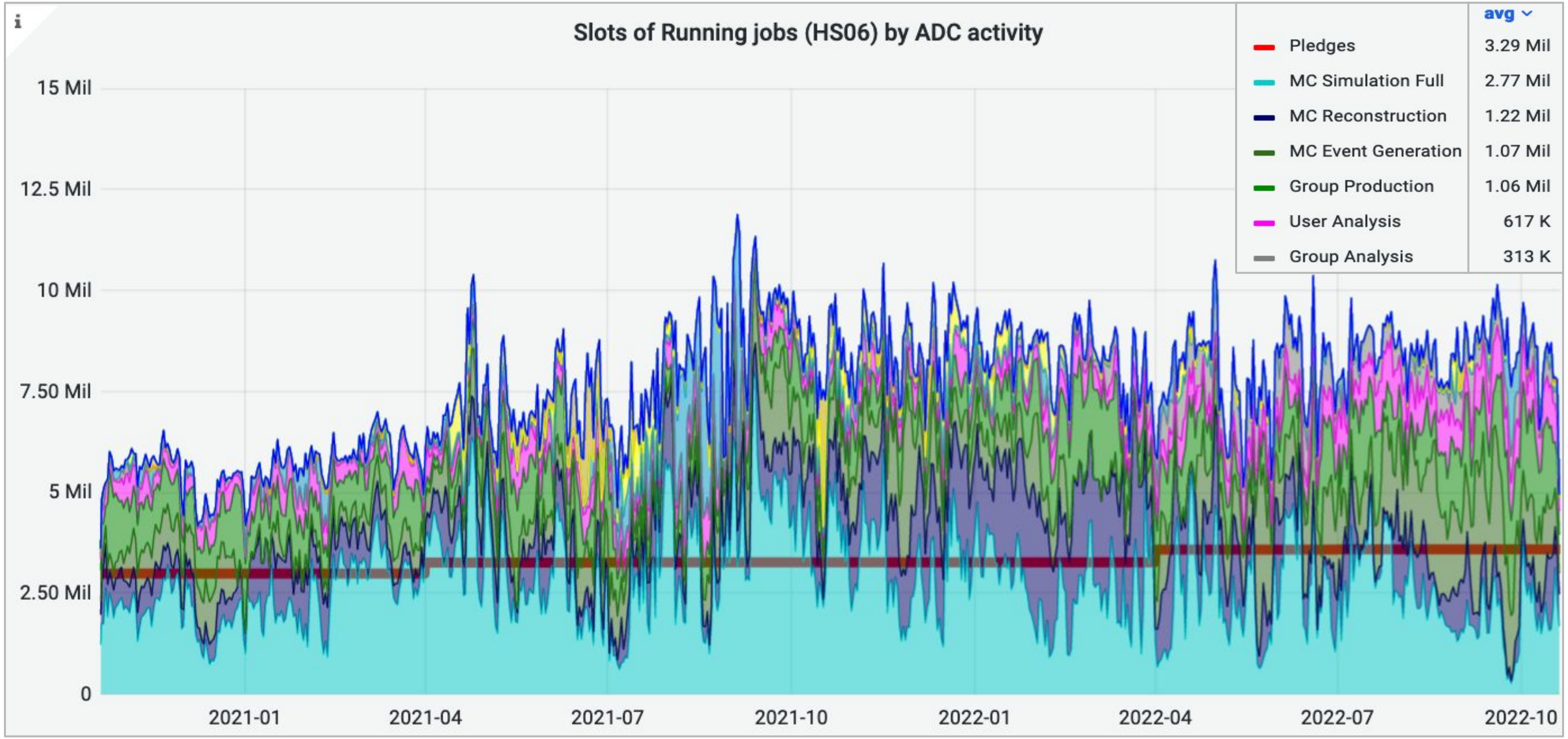
With special thanks to Rod Walker, Ivan Glushkov, Alessandra Forti, Frederic Derue, Andrej Filipcic, Zach Marshall and Ale di Girolamo for providing useful inputs and discussions

Introduction

- Networking is one of the rock-solid, highly reliable building blocks of ATLAS computing successes
  - Not without considerable amount of work from many people
  - Not without issues
- Networking is not infinite
  - Saturations here and there were found (often difficult to solve or mitigate)
  - But still most of the issues/limitations come from services (i.e. storages and third-party transfer manager FTS) rather than from Network fabric
- Here today to discuss with Networking community
  - What and how we are doing (in terms of network usage)
  - R&D projects, and possible evolution of infrastructure: preparing ourselves for HL-LHC
    - Organized campaigns (e.g. Data Carousel)
    - Infrastructure (e.g. diskless sites, HPC, commercial Cloud)
  - We do not discuss about the “general” HL-LHC network needs because they will be covered by the Data Transfers Challenges presentation later

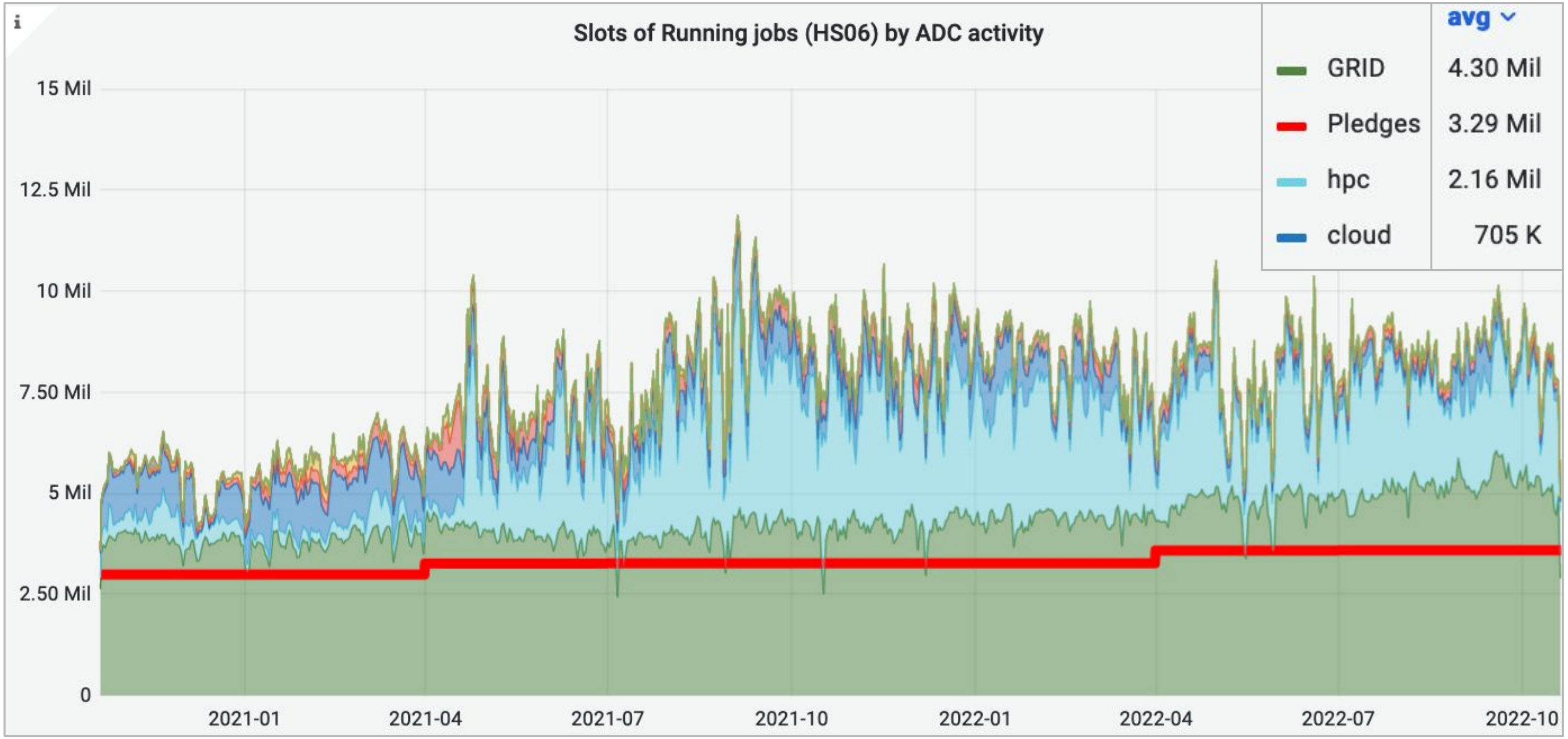
Compute Resource usage – last 2 years

- Excellent performance of our distributed computing infrastructure
  - Constant mix of activities, difficult to predict



Compute Resource usage – last 2 years

- Excellent performance of our distributed computing infrastructure
  - Grid, HPCs, Cloud all at comparable level, but HPC and Cloud mostly from a few big sites



ATLAS network usage: LAN and WAN

- How do we use the WAN/LAN network, how do we read/write the data:
  - **WAN**: today we (mostly) move data between sites in an orchestrated way where there are compute resources available, before the jobs start and after they end
  - **LAN**: jobs (mostly) read and write data from and to local grid storage, optimising the CPU utilization

  - Depending on the site's storage configuration there may be also significant LAN traffic (eg Ceph and erasure coding) or even WAN traffic (distributed storage) that we don't see

ATLAS network usage: synch and asynch

- Synchronous vs Asynchronous data transfers (wrt job running)
- **Synchronous**: not centrally scheduled (nor throttled) reads: data read/write from running jobs
  - **“Copy to scratch”**: copy input data from local storage to worker node before job processing starts, write output to local storage after processing ends
    - Used for vast majority of production workflows where the entire file is read
  - **“Direct I/O”**: Open and stream input data directly from local storage to compute node
    - Used for most analysis workflows where a small part of the file is read (average 15% of DAOD)

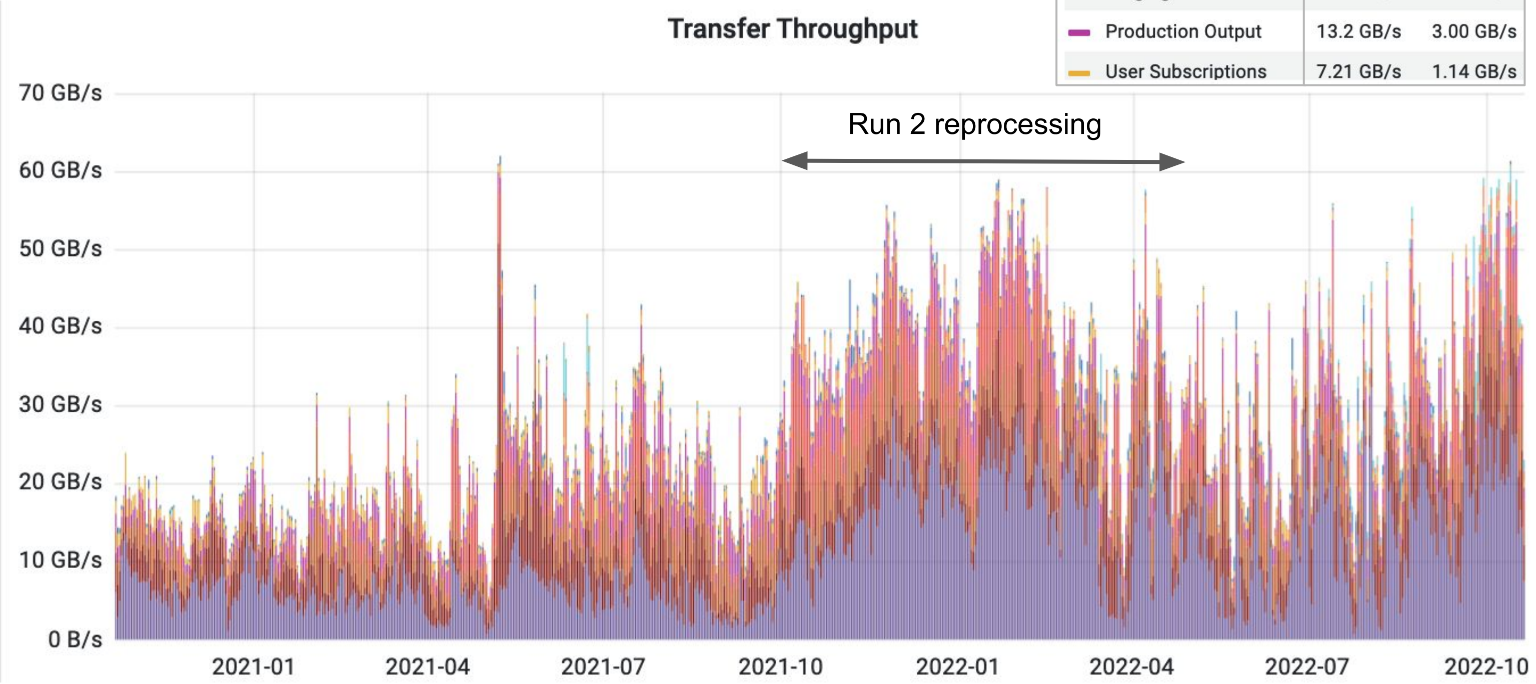
- **Asynchronous**: third-party transfers
  - FTS is the service managing the data transfers, orchestrated by Rucio
  - Rucio triggers FTS transfers from site A to site B (e.g. to consolidate datasets at a site, or move data to a site where it will be processed)
  - Once data is available locally the jobs can start

- Today Synchronous is *almost* all LAN. Asynch is *almost* WAN
  - *almost* for Synch: more later on diskless sites and more R&Ds
  - *almost* for Asynch: staging from tape is also LAN traffic

Network usage - last 2 years - WAN (Asynch)

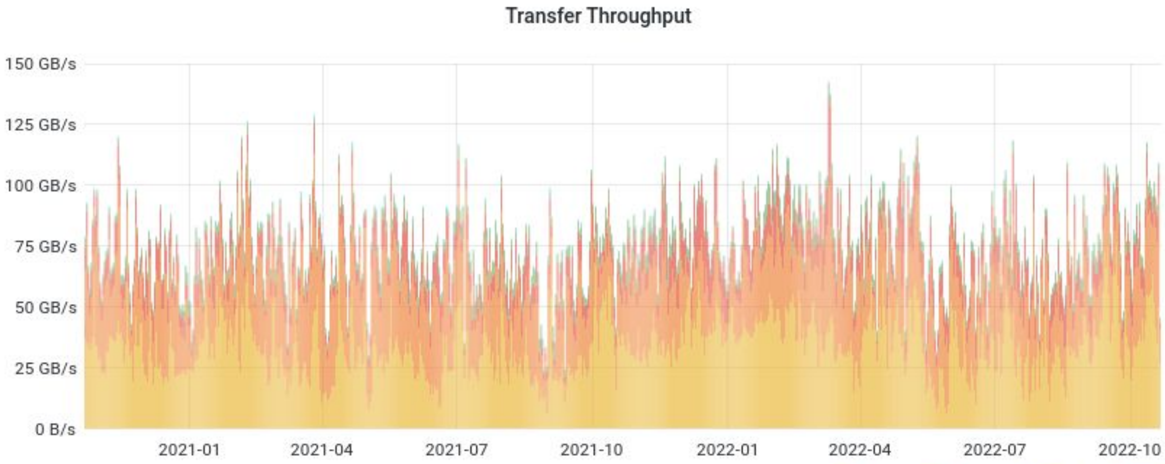
- FTS transfers orchestrated by Rucio
- Average in last months: 25GB/s
- Peaks at 60GB/s

	max	avg
Production Input	33.8 GB/s	11.1 GB/s
Analysis Input	47.4 GB/s	4.70 GB/s
Data Consolidation	20.4 GB/s	4.17 GB/s
Staging	22.3 GB/s	3.34 GB/s
Production Output	13.2 GB/s	3.00 GB/s
User Subscriptions	7.21 GB/s	1.14 GB/s



Network usage - last 2 years - LAN (Synch)

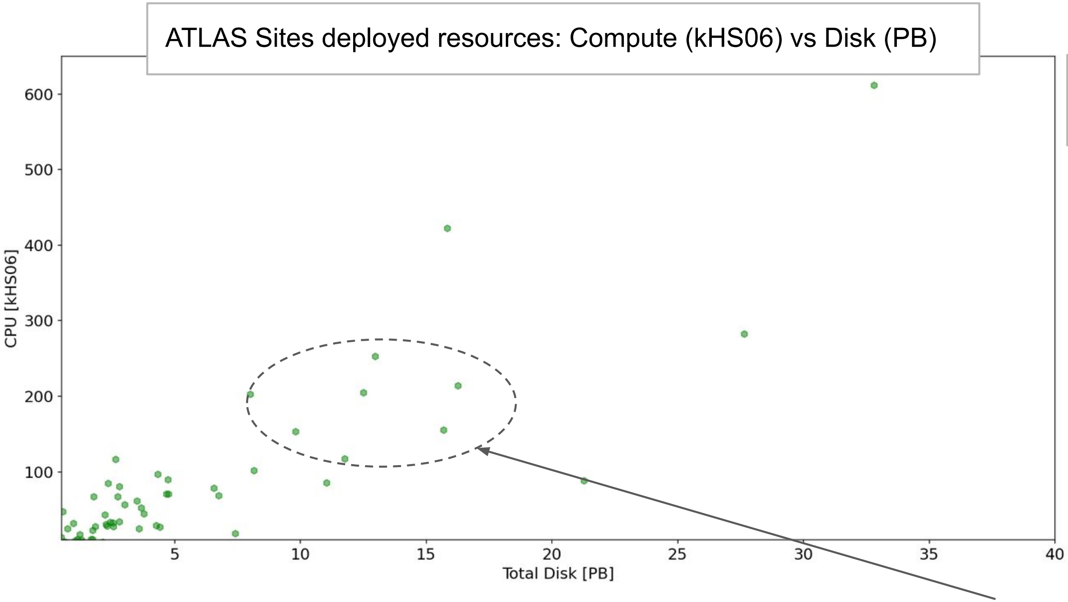
- Dominated by reads
- Average at 60GB/s, peaks at 120GB/s
- Note: Direct IO reports the whole file size, not what was actually read



	min	max	avg	current
Production Download	5.94 GB/s	76.5 GB/s	34.1 GB/s	26.6 GB/s
Analysis Download Direct IO	5.37 GB/s	67.5 GB/s	28.4 GB/s	13.9 GB/s
Analysis Download	1.84 GB/s	36.3 GB/s	9.52 GB/s	3.05 GB/s
Production Upload	0 B/s	13.4 GB/s	3.77 GB/s	1.59 GB/s
Analysis Upload	0 B/s	2.78 GB/s	705 MB/s	419 MB/s

ATLAS computing centers - compute vs storage

- There isn't a precise "ratio"
- We have many different configurations which can all be effective - our recommendations here



illustrative plot based on first 7 months of '22 data

- WAN requirements are very difficult to estimate: by experience for a 200 kHS06 site (15-20k CPU cores), at least ~50 Gbps average is needed (→ ~100 Gbps to be able to absorb peaks)
- Storageless sites are a reality, but on a small scale - see next slides

Storageless sites

- We are improving our frameworks to exploit efficiently (almost) everything that is provided to us
  - Storageless (compute only) sites are an interesting concept – they can minimise the operational needs of the site, and yet still usefully contribute to the experiment
  - Typically they are configured to read/write from/to a “close” larger site’s storage
  - Also possible to “attach” a storageless site to multiple (other sites’) storages

- We have to be careful: what if all the LAN traffic would go through WAN?
  - From the transfers numbers we know that the site internal LAN capacity is 3-4 times the WAN → if not well planned and organized, it could be inefficient, and sometimes also disruptive

- Storageless sites are an important strategic evolution of our infrastructure
  - Necessary in some situations, but not for everything
  - Caching plays a vital role here

Storageless sites – more

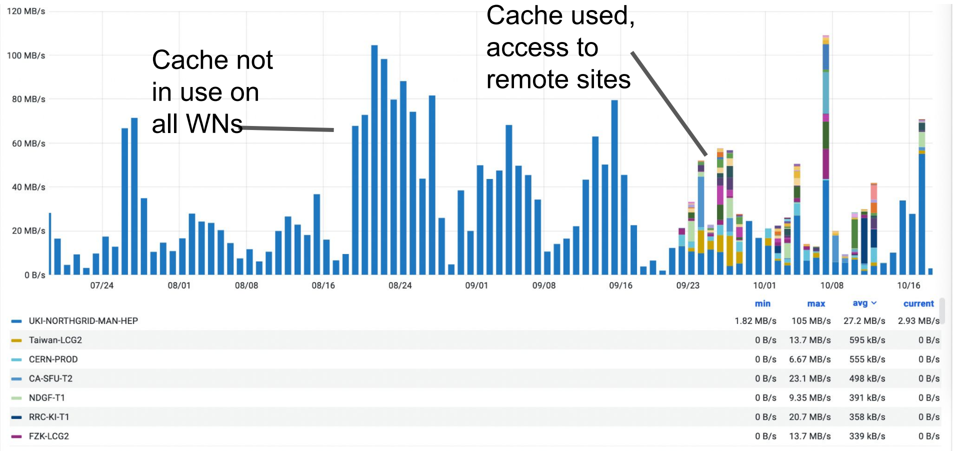
- Some funding agencies are evaluating the possibility of evolving their infrastructure by reducing the number of sites with storage, moving some sites to storageless
    - For example, in Israel, 3 separate sites with storage and CPU were consolidated in 2020 to a single storage and 2 CPU-only sites

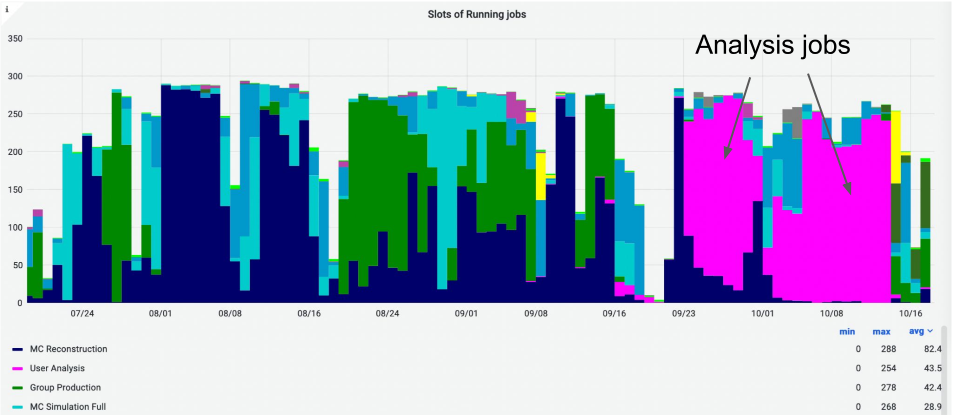
- We do have handles in case we see network limitations, e.g. limiting the workflows to the low I/O ones (simulation, generation)
    - This is only a stop-gap solution which can be used for a limited number of situations – having too many of these situations would impact all the other sites, which would have to deal with all the more I/O intensive workloads

- To cache or not to cache?
    - Caches (for now our experience is on ARC cache and XCache) are very useful: not only for data reuse, but also for reliability and robustness of data transfers.

Storageless sites – a real example – Birmingham

- 300 cores, 10 Gbps, with XCache
- Connected to storage in Manchester ~100km away
- Average transfer rate for the past 3 months ~40 MB/s
  - Before mid September XCache was not enabled on all the WNs which led to 100 MB/s peaks, a lot for 300 cores
  - After mid-September lot more analysis jobs and remote access to all sites was enabled
- Cache shows a good 35-40% of data reuse, i.e. less load on remote Manchester storage
  - Cache is filled only with bytes requested, not whole file

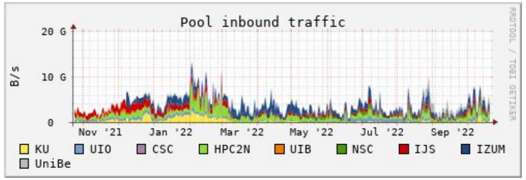




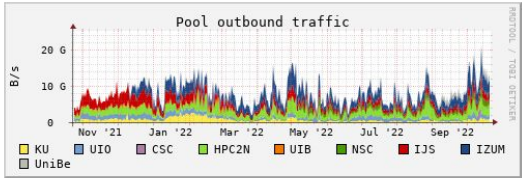
Last 3 months

HPC – Vega

- Vega EuroHPC:
  - 250k powerful cores (~14 HS06 per core)
  - ATLAS has very close contacts with the centre, allowing us to exploit it before any other users (since April 2021)
  - Data management performed by 2 ARC-CEs and 12 ARC transfer servers
    - ARC caching and controlled transfers is vital to protect the distributed dCache@NDGF



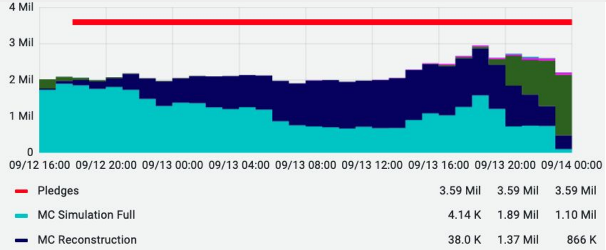
Write rate to dCache pools



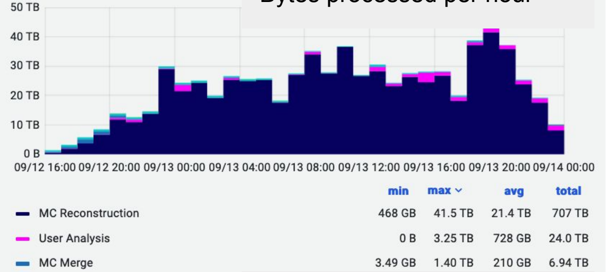
Read rate from dCache pools

In the plots: ramping up MC reconstruction on ~100k cores

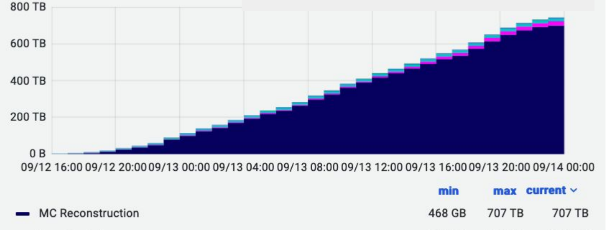
Running HS06 of Vega compared to total WLCG pledge



Bytes processed per hour



Cumulative bytes processed



HPC – not only Europe, Asia and America

- We have started using African HPCs
  - Toubkal at Mohammed VI Polytechnic University (UM6P) in Rabat, Morocco
  - 1 Gb/s connection, entering Europe in UK
  - Running simulation jobs on ~4k cores, reading and writing to RAL
  - Good network connection is critical to expand to different workflows

- We have been contacted by the University of United Arab Emirates
  - Also for them, it is not clear the optimal network path, and the possible bandwidth

- And what about the rest of Africa?
  - Providing network connectivity could help enable development for some of these countries

Network – monitoring and evolution

- All the ATLAS data transfers activities tracked through Rucio
  - Lot of work done with Rucio traces already, but still a lot more info to extract
    - Monitoring of direct I/O to know how much is really read
      - Info from xrootd/xcache
      - Info from software frameworks
    - Does anyone have spare Data Scientists?

- We are aware that we are not alone: we are engaged and we support several network-related activities:
  - traffic visibility (Packet Marking and Flow Labeling)
  - traffic pacing
  - network orchestration

- Monitoring of caches (XCache) still needs to be improved, both in terms of information and in terms of documentation for site admins.

Conclusions

- Network is one of the backbones of our distributed computing infrastructure:
  - It has been working wonderfully well
  - We need to make sure it will keep on working beautifully well
  - We know it's expensive and we will continue to try to optimise data flows

- Evolution in the infrastructure may imply more WAN network needs
  - Caches can certainly improve robustness in some cases
  - But still if there are massive "storageless" resources they will need to be coupled with adequate network

- Monitoring and more intelligence in exploiting the network is paramount
  - We (ATLAS, LHC experiments, HEP…) are not alone in heavily relying on a shared network resource, we need to evolve our tools to be able to understand and modulate our needs