IRIS-HEP 200Gbps challenge

April 2024

Brian Bockelman

MORGRIDGE
INSTITUTE FOR RESEARCH

DOMA in a nutshell

- Recall the IRIS-HEP Strategic plan outlined four ‘computing gaps’ between now and the HL-LHC:
  - G1: Raw Resource Requirements
  - G2: Scalability of the Distributed Cyberinfrastructure
  - G3: Analysis at the HL-LHC Scale
  - G4: Sustainability
- Given the “D” in DOMA is “Data”, the area is relevant to all four gaps. However, the team is focusing on (G2), (G3), and (G4).

Scalability and Sustainability

- **Scalability of the Distributed Cyberinfrastructure:**
  - Can we turn “raw resources” into “effective capacity” at the HL-LHC scale?
  - If you have the CPUs in the US and the disk at CERN, do you have the CI to turn the data into science?
  - Are the networks, middleware, and services ready for the raw scale of the envisioned HL-LHC workflows?
- **Sustainability:**
  - **Can we afford to run and maintain the services in the CI?**
  - DOMA’s strategy is to commoditize parts that are not unique to our community and share the things that are.

Measuring Progress

- In the February 2020 review of IRIS-HEP, one recommendation we got was to setup a series of “grand challenges” to help focus effort to strategic items.
  - **Goal**: Have a sequence of quantifiable, increasingly realistic exercises that can be taken as a proxy for HL-LHC readiness.
- We defined the “Data Grand Challenge” which grew into the community’s “WLCG Data Challenge”.
  - In Fall 2021, DC21 was executed – 10% scale of HL-LHC – and was a success.

Year	**% of HL-LHC scale**	**Flexible** (Gbps)
2021 ✅	10% ✅	960 ✅
2024 ✅	25% ✅	2,400 ✅
2026	50%	4,800
2028	100%	9,600

DC: Scale *and* Technology Readiness

- Around the same time as DC21, we'd been working within WLCG DOMA to introduce HTTP-TPC as a transport technology.
  - We felt it was ready.
- **Problem: How do we show the community HTTP is ready?**
  - **Solution: DC21!** Use the data challenges as a staging ground for showing new ideas.

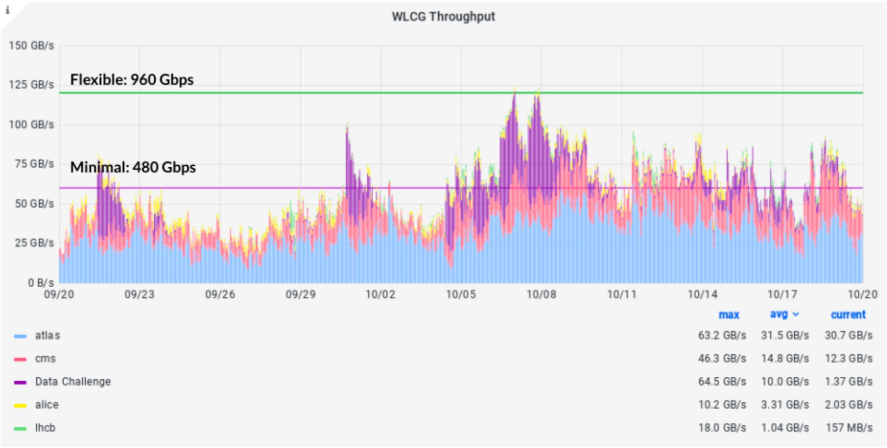


Figure 1 - Mock DC1 22/09/2021; Mock DC2 01/10/2021; Network Challenge (DC) 04-10/10/2021; Tape Challenge 11-19/10/2021.

Transfer scaling during DC21.

Figure reproduced from https://zenodo.org/record/5767913

DC: Scale *and* Technology Readiness

- **Happy ending!**
- DC21 showed that HTTP was viable for replacing GridFTP at LHC scales.
- Community adoption & uptake was rapid.
  - By the end of 2021, nearly all bulk data transfers for LHC migrated to the new protocol.

- Not all technologies will have happy endings.
- Important piece is using ‘grand challenges’ to move the community forward.

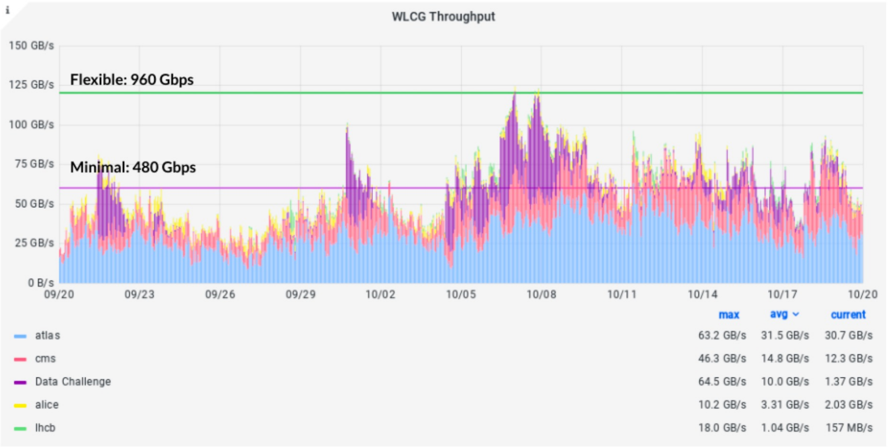
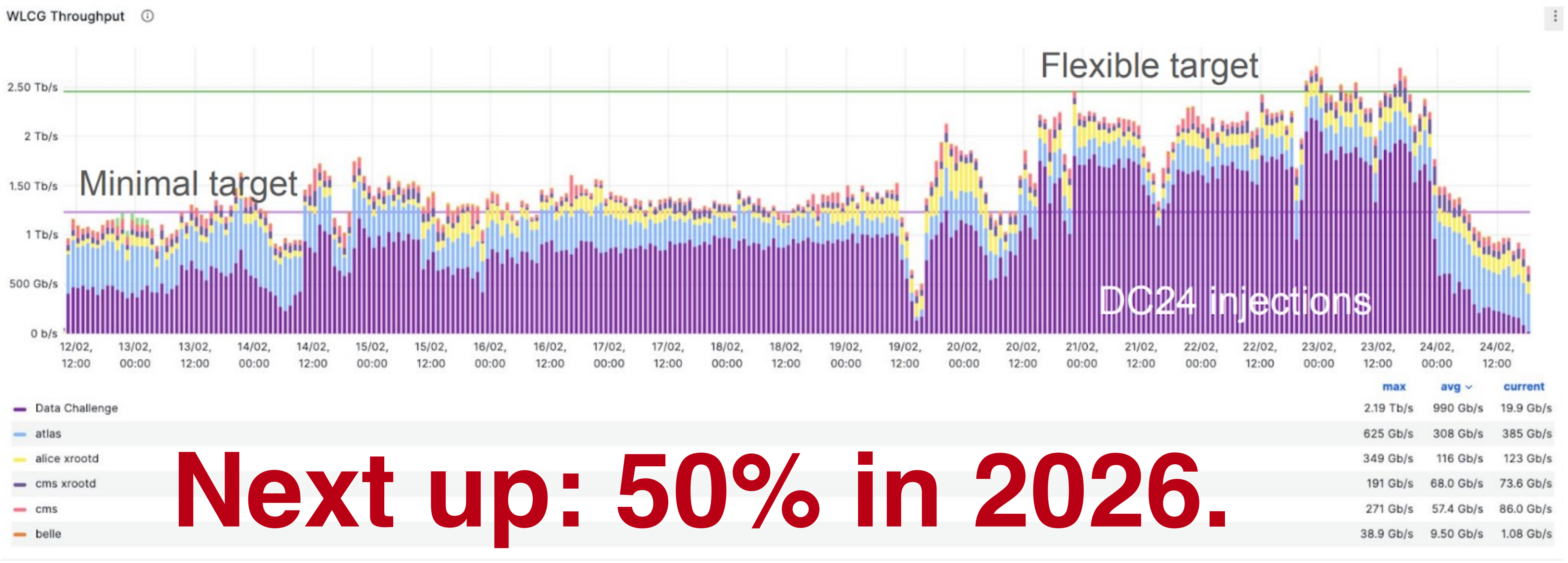


Figure 1 - Mock DC1 22/09/2021; Mock DC2 01/10/2021; Network Challenge (DC) 04-10/10/2021; Tape Challenge 11-19/10/2021.

Transfer scaling during DC21.
Figure reproduced from
https://zenodo.org/record/5767913

DC24: ¼ of the way there

- The exercise was quite smooth. **All scale targets were hit: demonstrated we are ready for 25% of HL-LHC scale**.
- Community-wide summary is gathering inputs still (will be presented May 2024). Plenty of lessons learned from sites to middleware.

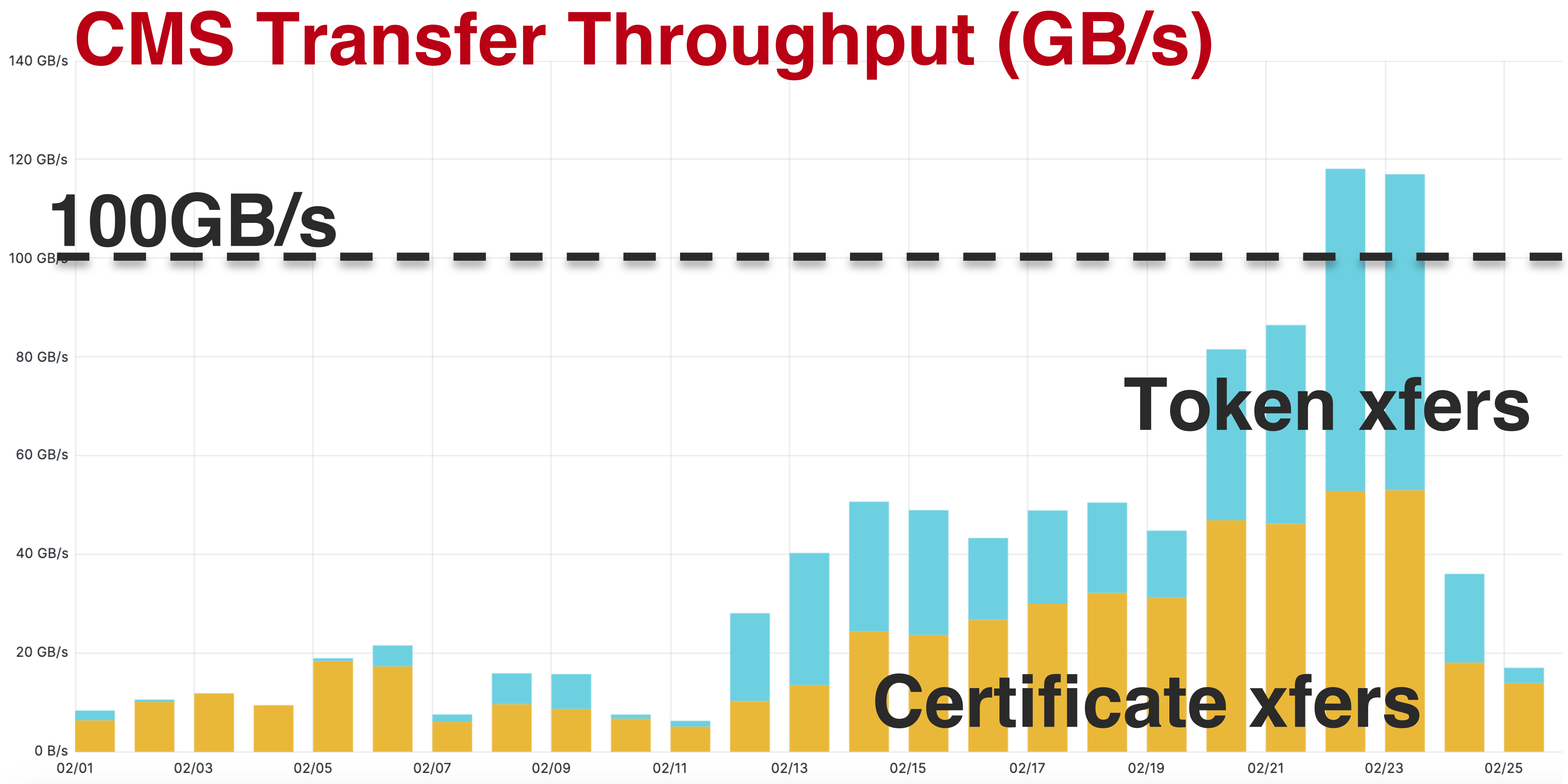


Next up: 50% in 2026.

IRIS-HEP Technology in DC24

- For sustainability goal, IRIS-HEP work includes:
  - **New authorization technology**: Switching from X.509 ‘grid’ authorization to industry-standard JWTs.
  - **Integrating network management**: Pulling ESNet’s SENSE technology into the LHC stack, showing networks can be managed as part of the data management system.

Summary Plot



Highlight Numbers – Using tokens:

- >30PB moved
- 25 CMS sites
- On peak days, >50% by volume
- >1M xfers / day

Short version: it works

Grand Challenge as a Framework for Progress

- The “Grand Challenge” approach has been instrumental in focusing the community and the institute.
  - I feel it’s helped close (G2) scalability of the distributed CI and (G4) sustainability “HL-LHC gaps”.

Idea: Let’s do the same thing for “analysis at HL-LHC scale”

The 200Gbps Challenge

- **Observation**: IRIS-HEP innovates in
  - Facilities R&D (how do we build better compute facilities for HL-LHC; SSL area).
    - Includes pathfinder facilities that can access ATLAS, CMS, or open data.
  - Analysis systems (bringing the Python-based analysis ecosystem in production).
  - Data delivery (effective delivery of events to compute).

- **Idea (mid-March)**: Pull the three efforts together and show readiness at 25% of HL-LHC scale.
  - And present the results at the WLCG Workshop in May 2024 (7 weeks from the launch of the idea).

25% of what, exactly?

- We want to show significant, quantitative progress toward HL-LHC-scale analysis.
  - Like in DC21, use realistic proxies for HL-LHC.
- In DOMA, we were able to tap into a long history of facility planning and was able to get the community to agree to goals based on extrapolating from a decades-old system.
  - No such luck in analysis. **Very little agreement** on HL-LHC analysis models.

- We decided to put down our own axioms for the challenge:
  1. We believe a full-scale HL-LHC analysis requires high-data rates, **reading 200TB in 30 minutes**.
  2. We want to use the IRIS-HEP Data Analysis pipeline and SSL facilities.

- Longer-term, we're trying to socialize the need for the community to find common truths.

Why 200TB in 30 minutes?

- Why select X TB in Y minutes?   (X=200, Y=30)

- Experience shows we hit scaling limitations when we go up by an order of magnitude.
  - Running smoothly at 10X brings immediate benefit back to the 1X case.
  - If we fail to run smoothly at 10X then we gain valuable insight into the current limitations.
- This is ambitious-but-realistic for extrapolating today's facilities out 4 years.
  - There's nothing exotic or out of the reach of a typical US T2 in the 2028 timeframe.
- This is within reason by extrapolating today's parameters out to the HL-LHC event counts and sizes.
  - There's no first-principles derivation of the leading order. One also cannot argue that missing these targets will cause HL-LHC to fail.
    - But then again, the same is true for DC24.

Points to the need for 'common truths' in the community around HL-LHC analysis

Derived Values – Example CMS ‘napkin math’

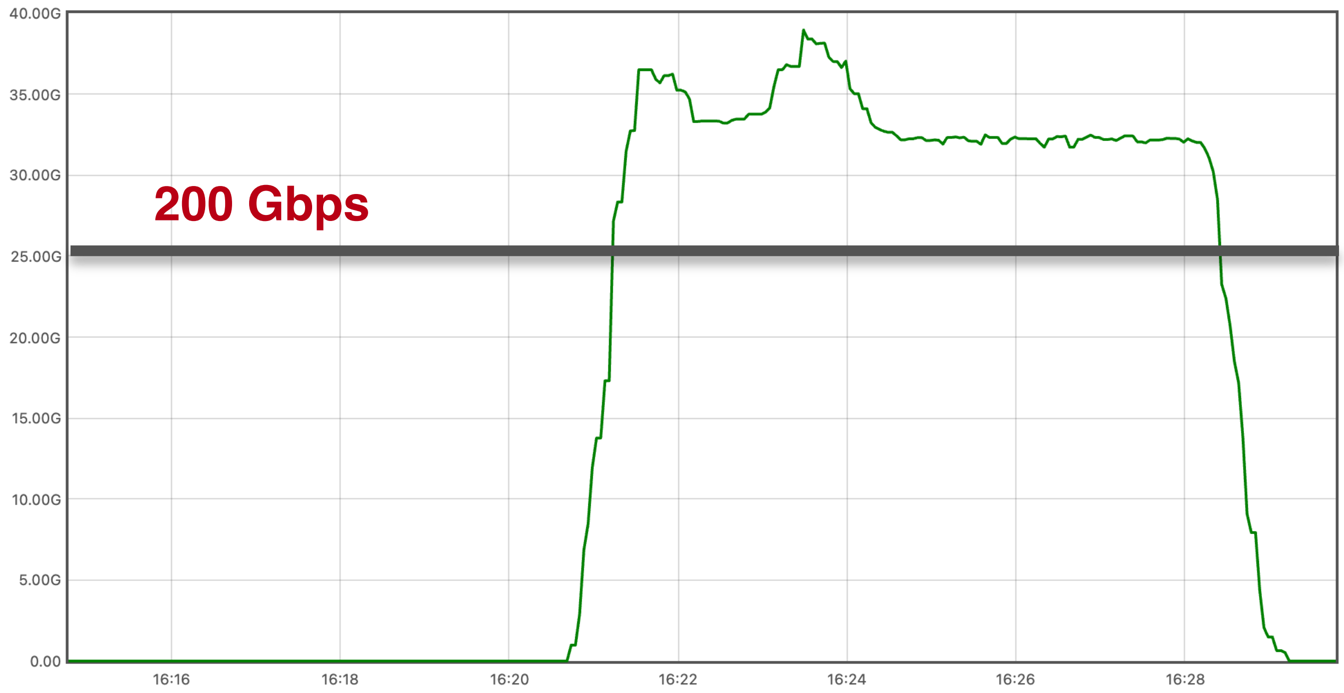
- Start with 200TB read in 30 minutes. => ~900Gbps sustained.
- 25% scale => 200Gbps sustained. Hence, **200Gbps challenge**.

- 200Gbps over 30 minutes => 45TB of data into the analysis process.
- Assume 25% of the data read from the CMS NanoAOD
  - => 180TB of NanoAOD is required to push 45TB of branches.
- At 2KB/event, 180TB of NanoAOD is **96B events**.
- 96B events in 30 minutes => sustained 55MHz event rate.

Our sample analysis runs at 25KHz per core, meaning 2,200 cores are needed to sustain the 55MHz event rate.

200Gbps Challenge

- Given we want realism (use real data, not Open Data), we split into two teams – one working with ATLAS analysis data and approaches at Chicago, the other CMS at Nebraska.
  - The "napkin math" from prior slide was repeated for ATLAS
- **Immense**, focused activities across the institute.
- We are in week 6 of 7 for the exercise.
- **Pieces are starting to come together.**
  - Plot to the right shows hitting >200Gbps for a pure data movement test (no processing).



CMS Toolset

- For CMS, we decided:
  - Start with Run2 NANOAOD.
  - Process with Coffea 2024. Read data from XCache on the Coffea-Casa facility at the Nebraska Tier-2.
  - Start with the IDAP notebook from the AGC work last year, expand work out into the site HTCondor.
  - Dask tasks processed in TaskVine & Dask.
  - Compute values from the events read in; accumulate into histograms. "Direct from NanoAOD" style analysis.
- Notes on realism:
  - Real XCache setup. Token-based auth using the IAM service at CERN.
  - LZMA decompression dominates analysis time (~70%). To hit our target 25KHz-per-core processing rate, we recompressed the NANOAOD using ZSTD. About 20% larger than the original dataset, ~2.5x faster.
    - N.b.: our strong opinion is CMS needs to make this change.
  - We scale-out to HTCondor but, for these tests, pre-create the workers.

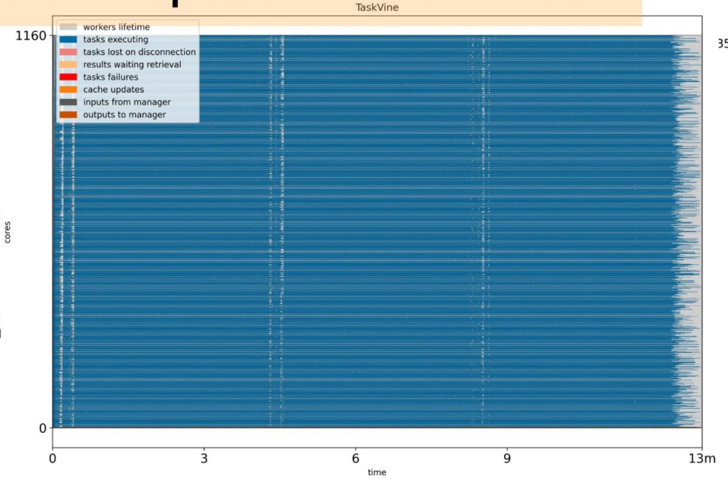
200Gbps Challenge

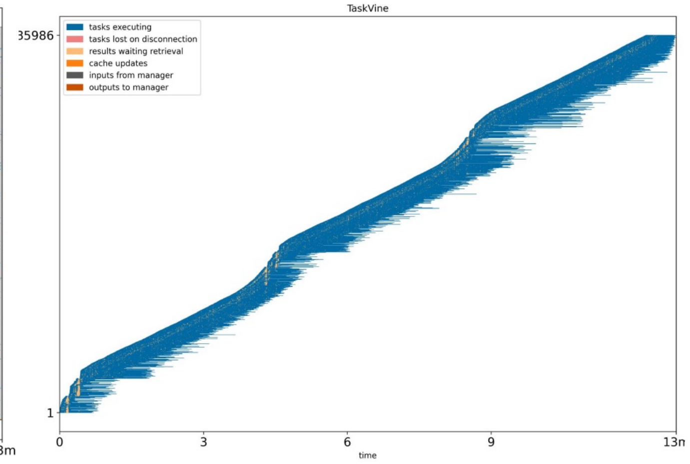
- As of last Friday, the CMS team was able to hit 178Gbps in processing data via uproot.
  - Over the weekend, test runs on a larger core count peaked at 202Gbps.
- Current obstacles:
  - With the full Coffea 2024 notebook, we see unexplained spikes in memory usage. Kills workers and causes processing “tails” (or stuck workflows).
  - Current tests hit targets using Uproot and reading via Python *but* strip out significant parts of the realism, making the work less interesting.
    - This week we’ve been bisecting the problem – adding back in the “real physics” code.

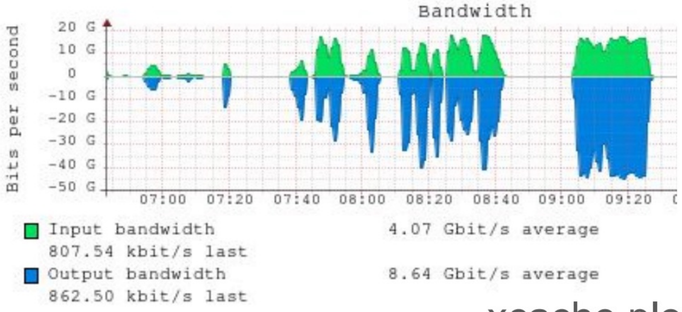
How far will we get by next Wednesday?

coffea.casa + uproot + taskvine

178 Gbps
1160 cores
35985 files
read: 17TB
13 min







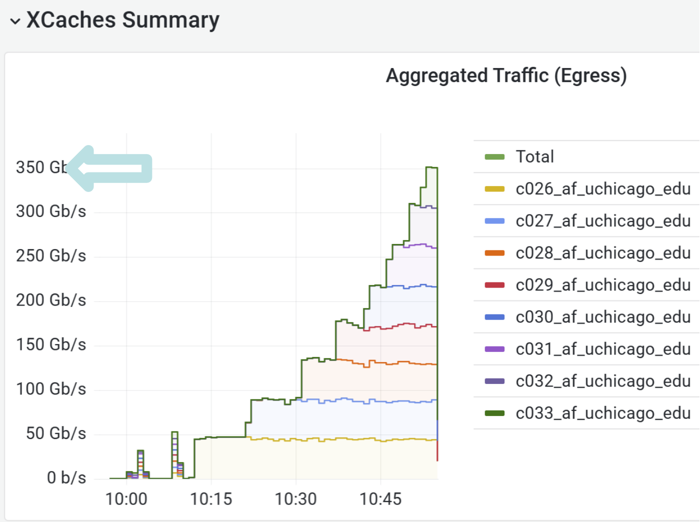
xcache plots

[Slide shown last Friday by Ben Tovar]

ATLAS – ServiceX Path

- Use ServiceX to skim PHYSLITE
  - 200 TB goal of internal Run 2 data + mc
  - All datasets are starting from the Midwest Tier 2 facility
  - Reading 25% of the data
- Internal Bandwidth Should Support 200 Gbps
- Running on 50 TB dataset with 64K files
  - Stress k8s, S3 storage of output SX fragments
  - Stable at ~40 Gbps, unstable at higher speeds
- Running on the output of ServiceX
  - 200 Dask workers works well
  - 1000 workers causes intermittent failures in S3
  - No backoff/retry in software
- 1 TB dataset in 3 minutes no problems!

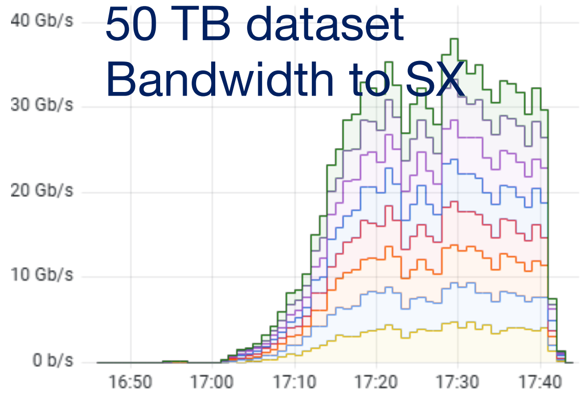
Internal UChicago AF Bandwidth



Where will we get by the workshop?

- Aiming for straight up 200 Gbps test
- Using SX for what it is good for – a prior physics motivated skim

50 TB dataset Bandwidth to SX

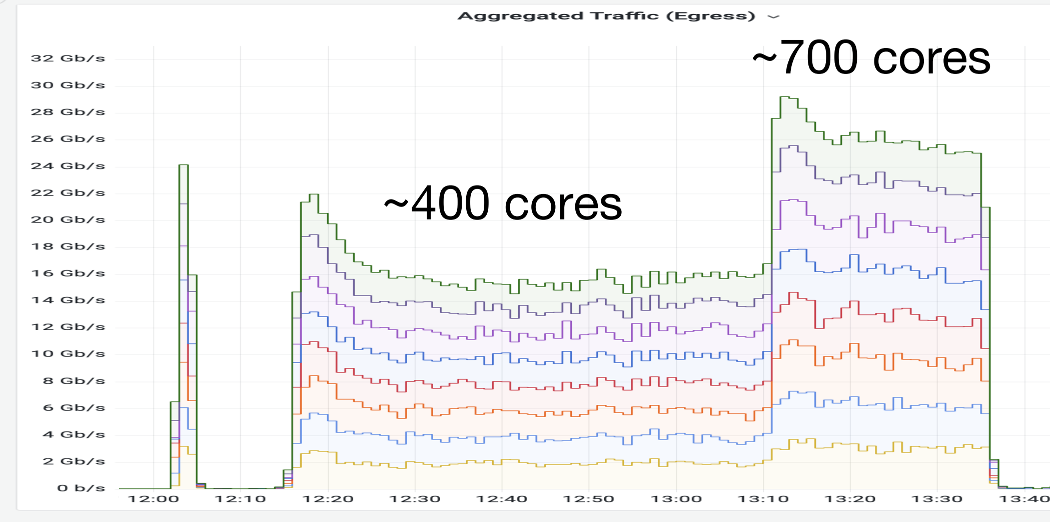


ATLAS – uproot, dask-awkward, coffea 2024

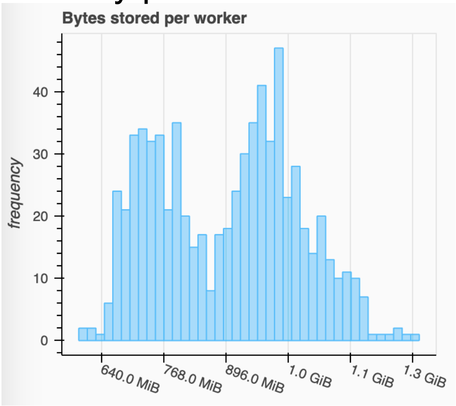
- Probe new kind of workflow
  - process PHYSLITE without intermediate steps
    - do everything "on the fly"
  - nominal setup uses coffea 2024, dask-awkward, uproot
  - same input / task as ServiceX setup

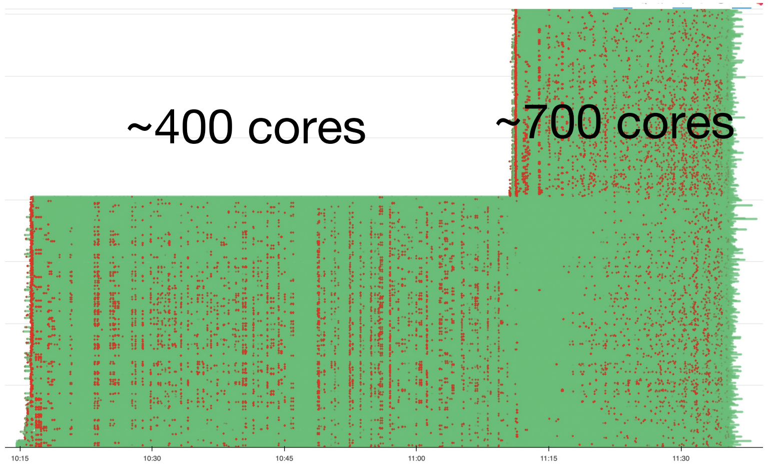
- Lots of lessons learned already, many ongoing investigations
  - scaled Dask up to around 2k cores
  - throughput up to 55 Gbps so far
  - work ongoing to go beyond

Test run with 65k files, 50 TB of data



solid memory profile across workers





Sustainability - Sharing

- A key component of the analysis facilities is XRootD.
  - IRIS-HEP funds effort toward making XRootD better.
- In fact, XRootD shows up several places:
  - Reference platform for HTTP transfers
  - Foundation of CERN’s EOS & CTA products (which manages ~1EB of data)
  - Used widely in other HEP experiments to deliver data.
  - Base of LSST’s “QServ” distributed database.
  - Transfer server for the Pelican Platform, which is the base of OSDF (used by NCAR, IceCube, LIGO, NRAO).
- At the heart of 5 different NSF Major Facilities.

Investments in IRIS-HEP for analysis have impact across the LHC, HEP, and wider communities.

Preparing for the HL-LHC

- DOMA is focusing on 3 HL-LHC “computing gaps”.
  - Demonstrated ability to move R&D into production in IRIS-HEP phase 1.
  - Reloaded with a new set of projects for phase 2.
- We have found the “grand challenge” approach to be a useful framing device for focusing effort.
  - <u>A series of increasingly-complex, cumulative exercises towards a common, quantitative goal.</u>
  - This is in addition to the “day to day” effort of bringing projects to fruition.
- Grand Challenges can be both scale and **technology readiness**.
  - Here, we’re leaning in technology readiness more.
- We’re in the middle of an intensive, time-limited exercise to show a vision of analysis at 200Gbps.
  - ~80% of the way through, it’s been a resounding success in finding weaknesses in the integration between parts of the institute.
  - Would be difficult to execute such a broad exercise outside an institute-like entity.
  - <u>Let’s see if we hit our quantitative goals as well!</u>