Accelerating science: the usage of commercial clouds in ATLAS Distributed Computing

Fernando Barreiro Megino on behalf of the ATLAS experiment
CHEP 2023, Norfolk, USA

ATLAS cloud projects

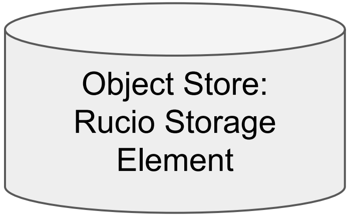
- Amazon
  - Current funding round: July 2020 - May 2023
  - Credits purchased by California State University at Fresno
- Google
  - Conversations and short R&Ds since several years. Stable relationship since ~2018
  - Latest project based on “User Subscription Agreement” for US public sector
    - July 2022 - October 2023
    - Duration, scale and cost agreed before project
    - Tracks for Cloud-site, R&Ds and Total Cost of Ownership (TCO)

What are we looking for in the cloud?

- 💡 New ideas and evolution
- 💪 Complementary sources for computing power
- Elastic usage
- Access to architectures not available on-prem
- 🏖️ Low-maintenance

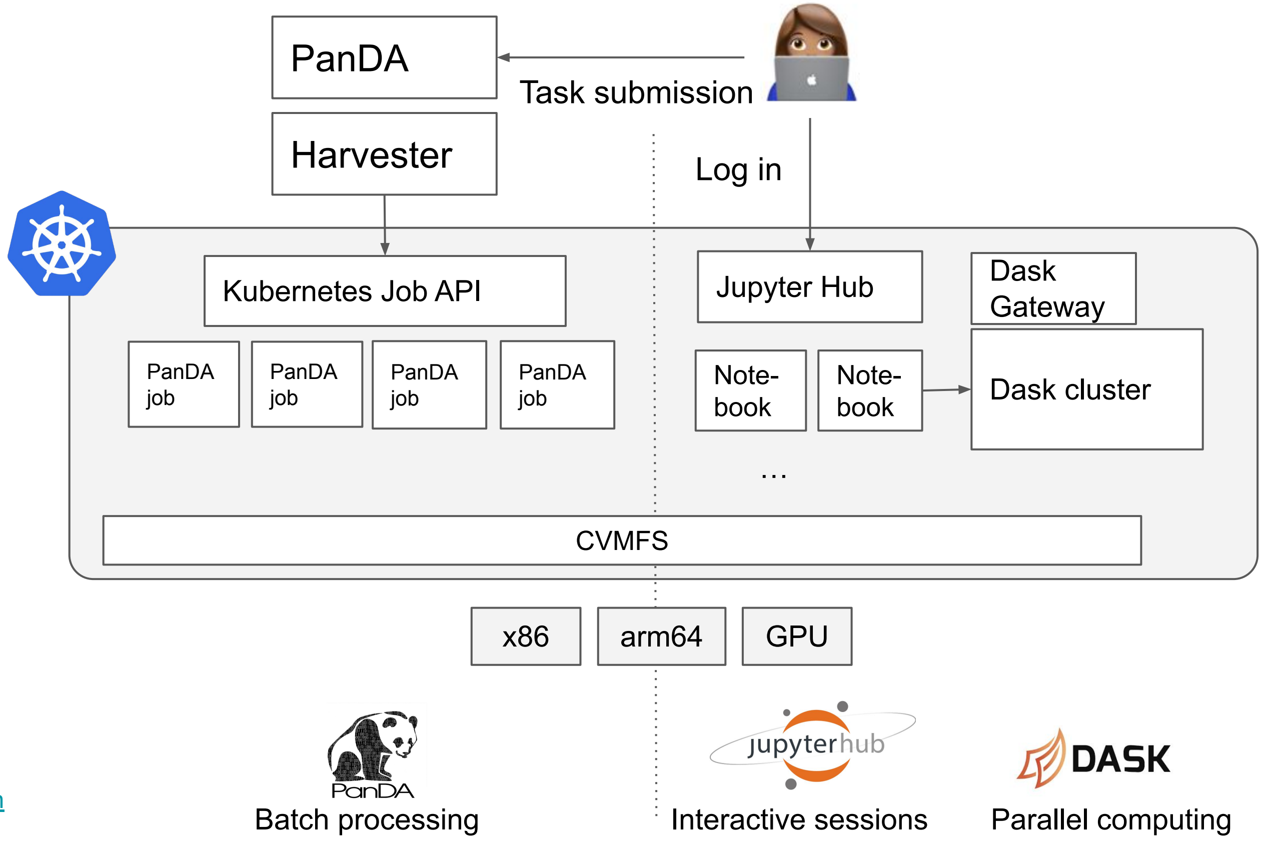
Cloud native integration

[Seamless integration of commercial Clouds with ATLAS Distributed Computing](#)



RUCIO
SCIENTIFIC DATA MANAGEMENT

FTS

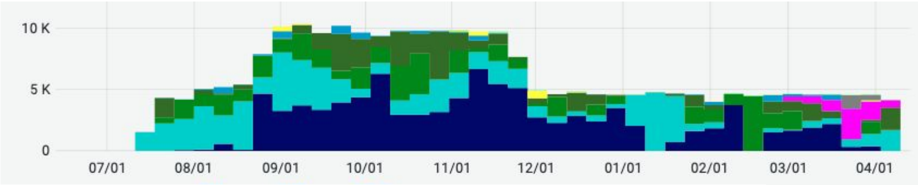


[M. Lassnig, Extending Rucio with modern cloud storage support](#)

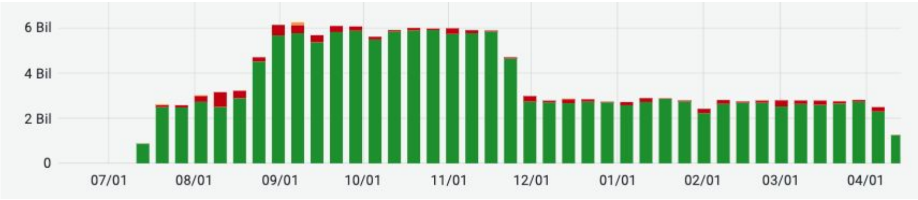
💪 Additional compute resources on Google Cloud

GOOGLE

vCPUs used by running jobs



WallClock Consumption of Successful and Failed Jobs

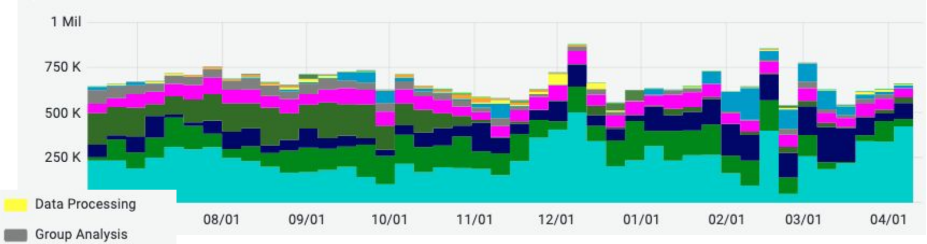


4% failed wallclock: mostly application issues, initial setup issues, Spot & DDM failures

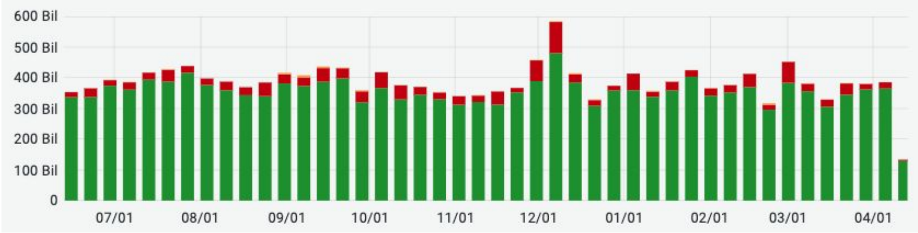
🏖️ **No interventions needed since setup**

Overall ATLAS including Grid, HPC, P1, BOINC, Cloud…

vCPUs used by running jobs

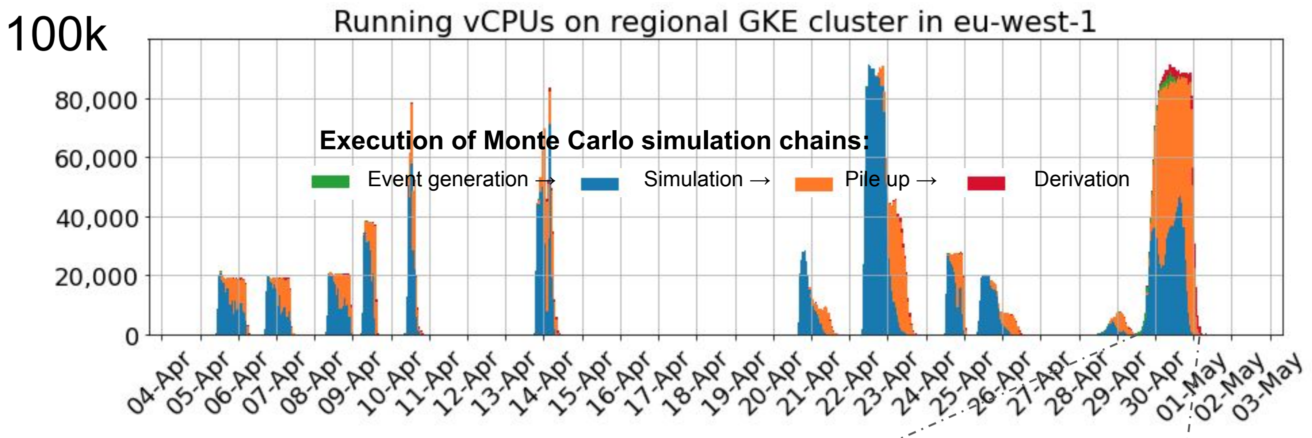


WallClock Consumption of Successful and Failed Jobs

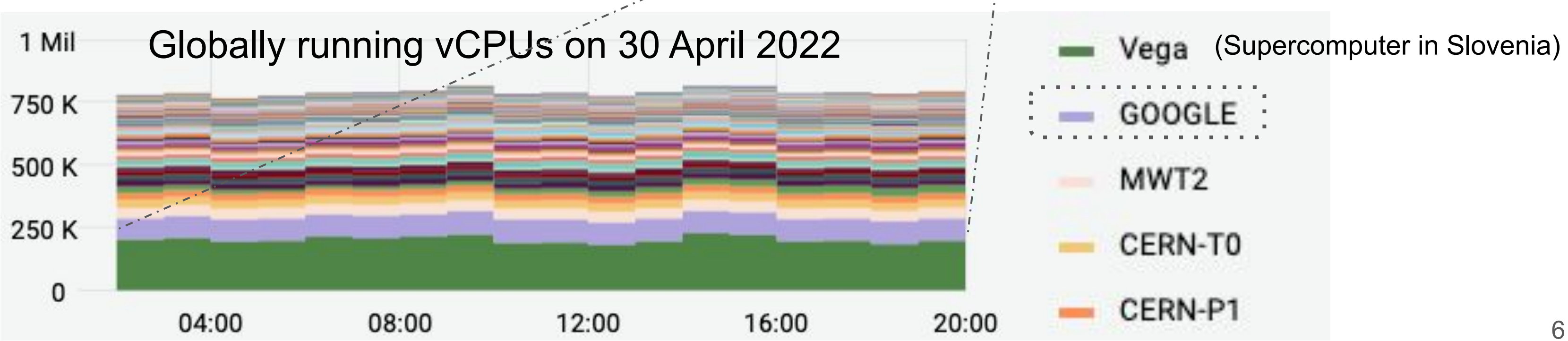


8% failed wallclock: higher failure rate concentrated at particular, unpledged sites

Elastic processing on Google Cloud



$O(10^5)$ vCPUs
$O(10^4)$ Pods
$O(10^3)$ Nodes
1 managed K8S cluster
<1 Engineer



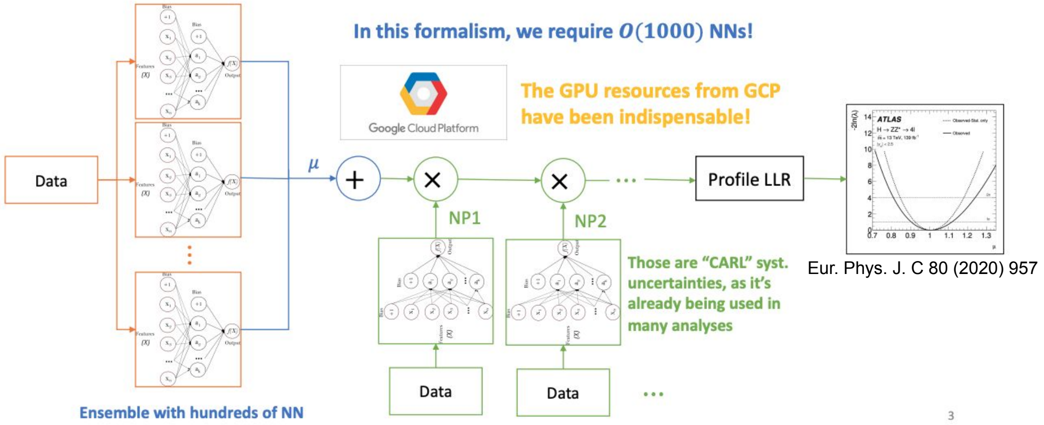
Heterogeneous architectures: GPU

[J. Sandesara, ATLAS data analysis using a parallelized workflow on distributed cloud-based services with GPUs]

1. Development on GPU notebook

2. Training at scale through PanDA GPU queue

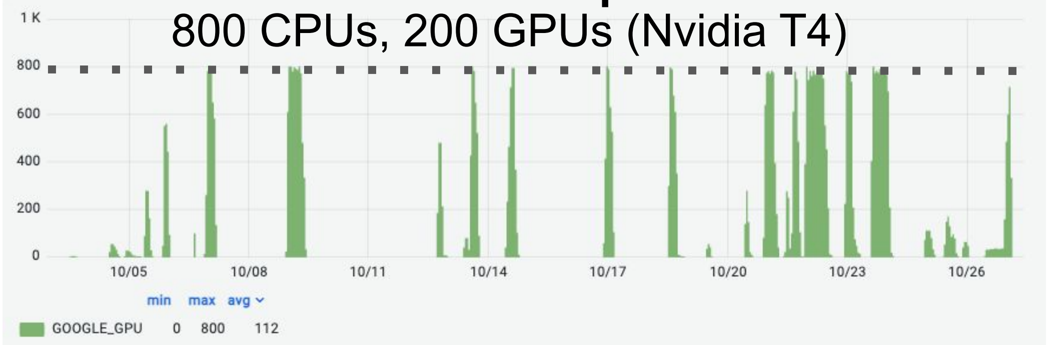
3. Fitting on mega-memory notebook (1.4TB) or PanDA queue



Eur. Phys. J. C 80 (2020) 957

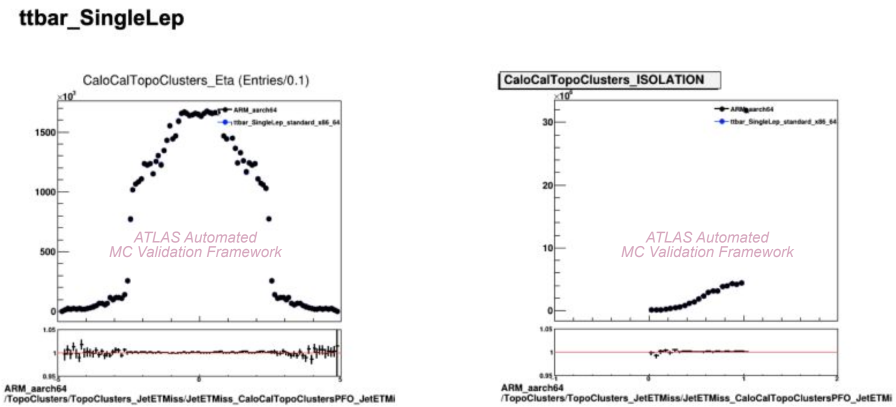
On demand GPU queue

800 CPUs, 200 GPUs (Nvidia T4)



Heterogeneous architectures: ARM

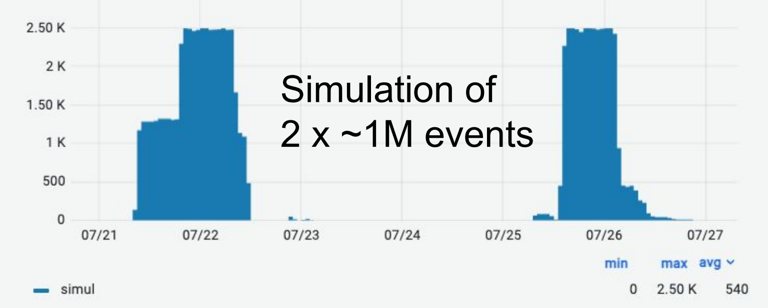
Physics Validation for Athena Simulation and Reconstruction on ARM signed off

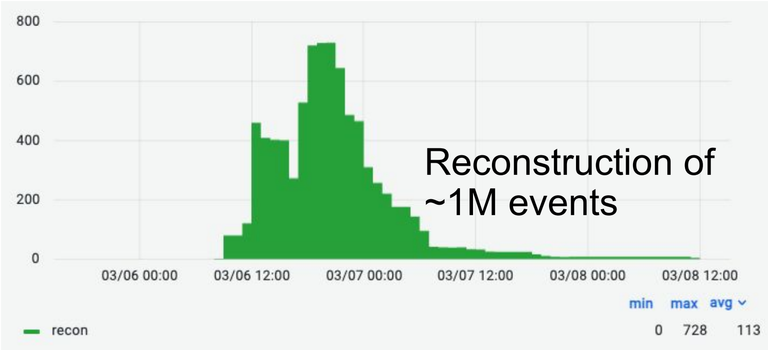


no significant difference

Example plots from calorimeter cluster validation of aarch64 vs x86_64

First ATLAS tasks on ARM: Amazon Gravitron 2 processors

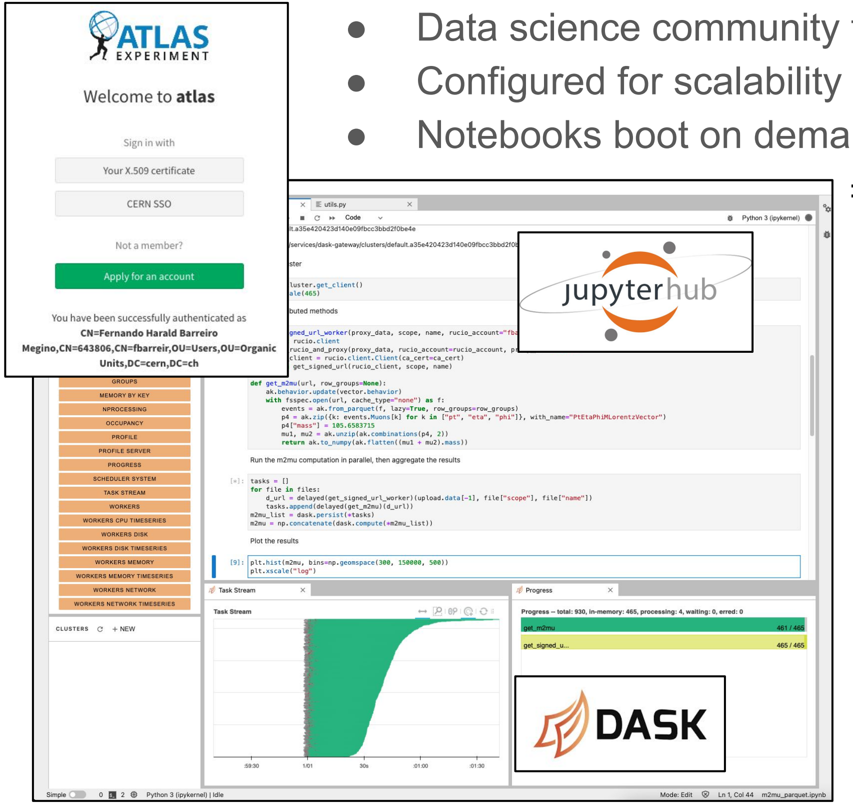




CHEP 2023 [J. Elmsheuser, The ATLAS experiment software on ARM]

Distributed, interactive analysis

- Data science community technologies
- Configured for scalability and cost effectiveness on GKE
- Notebooks boot on demand a variety of instances (GPU, XXL)



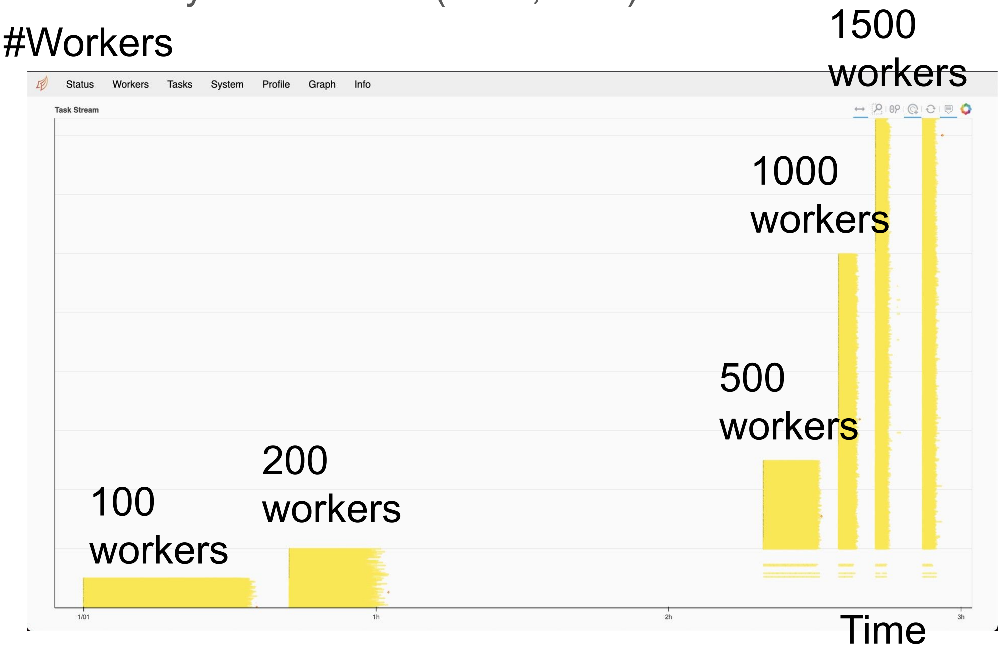
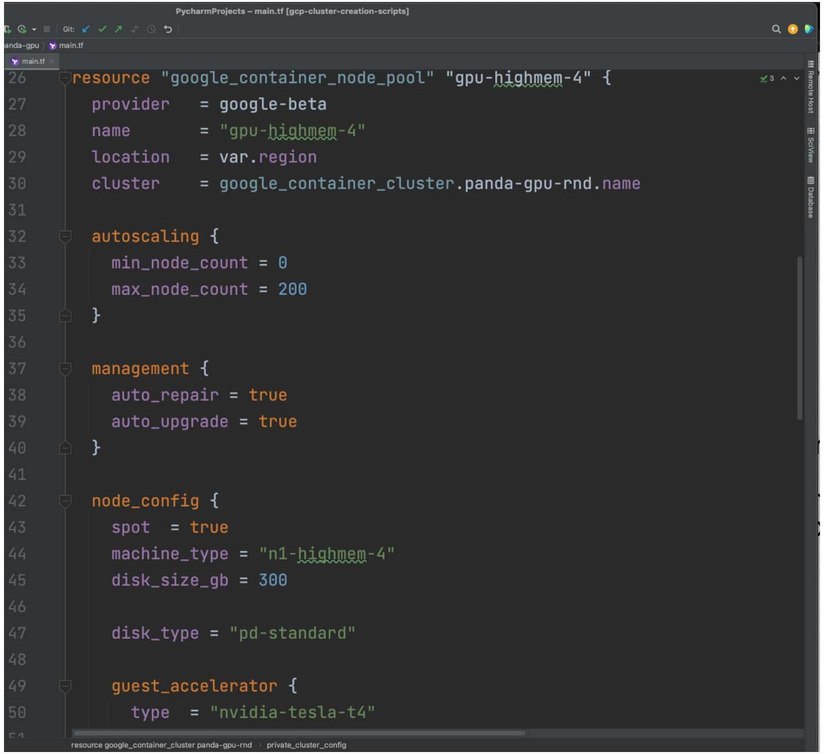


Image by Nikolai Hartmann and Lukas Heinrich

Infrastructure as Code



- GCP infrastructure management through Terraform: repeatable infrastructure
- Initial development is tedious, but helps in the long run when trying to replicate a setup
  - No clicking through UI and having to remember all options
  - Share recipes in the team

Conclusions

- Cloud native integration: working at scale
- Many of the technologies are game-changers and available for on-prem
- There are challenges when not following the mainstream path
  - So far solved with creative ideas
  - Further work is needed to complete the WLCG fabric integration
- Cloud resources enabled R&Ds (GPU, mega-memory nodes, ARM, Dask) that would have been very difficult to carry out on the Grid or on-prem clusters
- Running “à la Grid” is technically possible, but not the most cost effective way
  - In particular due to very high egress costs
- Ongoing Total Cost of Ownership studies to understand subscription models and list prices

Questions?

https://www.greentechmedia.com/articles/read/inside-googles-quest-for-24-7-clean-energy-at-data-centers

ATLAS contributors in the Amazon and Google projects

Mikhail Borodin (University of Iowa)
Kaushik De (The University of Texas at Arlington)
Johannes Elmsheuser (Brookhaven National Laboratory)
Alessandro Di Girolamo (CERN)
Nikolai Hartmann (Ludwig-Maximilians-Universität München)
Lukas Heinrich (Max-Planck-Institut für Physik)
Alexei Klimentov (Brookhaven National Laboratory)
Mario Lassnig (CERN)
FaHui Lin (University of Texas at Arlington)
Tadashi Maeno (Brookhaven National Laboratory)
Zachary Marshall (Lawrence Berkeley National Laboratory)
Gonzalo Merino (Port d'Informació Científica)
Paul Nilsson (Brookhaven National Laboratory)
Cedric Serfon (Brookhaven National Laboratory)
David South (DESY)
Harinder Singh (California State University)
Jay Sandesara (University of Massachusetts, Amherst)