# ATLAS - Google R&D Project
# Year2 Goals and Timeline

Kaushik De, Alexei Klimentov

February 17, 2021
ATLAS – Google Technical Meeting

# Timeline

- **Google – US ATLAS partnership in 2013-2015.**
  - It was focused on ATLAS WLCG sites elastic extension using Google Cloud Platform (ATLAS demonstrated how to run data processing and simulation at scale on cloud resources, a pioneering work for HEP computing at this time)
- **Google-ATLAS partnership in 2017-2018.**
  - "Data Ocean project". ATLAS and Google Collaboration R&D, ATLAS Data management and Workload management teams, use Google cloud for wordwide data analysis (publications ATLAS-SOFT-PUB-2017-002 and ATL-SOFT-PROC-2018-034)
  - Jun-Sep 2018 : series of technical discussions between US ATLAS Labs and universities and within ATLAS international about next steps led to writing a white paper submitted to HEP DOE in October 2018
- **US – ATLAS Google meeting with DOE HEP in June 2020.**
  - 3 tracks to be addressed in CY2020
    - Automated data handling between Hot/Cold storage, data analysis and virtualization
    - New data formats and I/O performance studies
    - Machine learning , new architectures and algorithms
- **September 16, 2020 : ATLAS – Google Technical Interchange Meeting (slides)**
- *October – December, 2020 : US ATLAS – Google discussions about Y2 objectives and funding*

# GCP4HEP R&D Year 1 Summary

❖ With a very modest investment at this phase of the project we have accomplished a research schedule. We have successfully proven the technical solutions for GKE/GKS integration, hot/cold storage, data formats and I/O

- Successful GKE/GCS integration for the first time with full Rucio/PanDA workflow has been demonstrated for ATLAS
- *Successful LSST-Google –PanDA story*
- Several valuable studies on data formats and ML

  o *as in any R&D project, success largely depends on the interest of researchers*

# GCP4HEP R&D impact

❖ interest from other commercial cloud providers (Amazon, Oracle) and proposals for joint projects

  ▪ *It is interesting to know LSST decision motivation*

❖ ATLAS Workflow Management SW TIM in December  discussed new Analysis Facilities concepts, workflows and analysis tools…

  ▪ A possible "common/universal" solution for analysis payloads execution on the Grid, clouds and HPCs (k8s)

  ▪ Analysis Facility concept is under discussion (ATLAS vCHEP2021 paper)

    o ATL-COM-SOFT-2021-016.  Evolution of ATLAS analysis workflows and tools for the HL-LHC era

# GCP4HEP R&D Year2

❖ **Ultimate Goal** : *demonstrate the usefulness of GCP as an Analysis Facility for ATLAS physicists*

❖ **Funding**

  ▪ No TAM for Year2

  ▪ All funds will be used for Google facilities / ATLAS users

❖ **Credit Distribution and Management**:  Allocate credit to a group of ATLAS users and developers. Developers - use credit to enable new products and features to increase productivity of ATLAS users. Users - use credit to test the usefulness of GCP in improving ATLAS physics products.

❖ **Metrics**: Demonstrate features and capabilities not available (or limited) on WLCG and Universities  resources. For example, use of ML tools, analytics and other features of GCP for physics analysis

❖ **Outreach**:  Discuss with Vera Rubin Observatory and Belle II common tracks

# GCP4HEP R&D Year2. Timeline

❖ *Feb 12th : Technical discussion about GCP cluster and data transfer to GCP*

❖ *Feb 12th - Mar 10th : prototype data analysis using CERN computing facilities with a sub-sample of data*

❖ Feb 17th : Resume ATLAS-Google (bi)weekly technical meetings

  ▪ Presentation and discussion of the in initial use-cases (NH, LH)

  ▪ Presentation of GCP infrastructure set up plan (FB)

  ▪ Presentation of DDM related topics (CS)

❖ Mar 3rd - Mar 10th : set up an initial GCP cluster and transfer data sample

❖ Mar 21st : make data sample accessible for selected users (10-15) via Jupyter hub

❖ Mar 31st : ATLAS-Google technical meeting to assess the first results and to discuss 3-6 months strategy

  *All dates TBC*