
CERN

CloudBank EU

Brokering cloud technology for Europe

Apostolos Theodoridis (CERN IT)

Outline

- Background
- Pilot Project Objectives
- What is CloudBank?
- Current Status
- Billing & Consumption Monitoring
- Summary

Background and Objectives

- Commercial clouds usage has increased in the recent years
  - HNSciCloud, GEANT IaaS Framework, OCRE Framework
- Financial, Procurement and Data Processing models need development
  - Allow rapid increase of available capacity without re-tendering and potential delays
  - Respond to demand of increasingly popular non-batch services (e.g. ML, HPCaaS, QCaaS) with optimized usage
  - Monitor billing accurately, individually and transparently
  - Multi-cloud, avoid vendor lock-in: rapidly select alternative cloud service suppliers

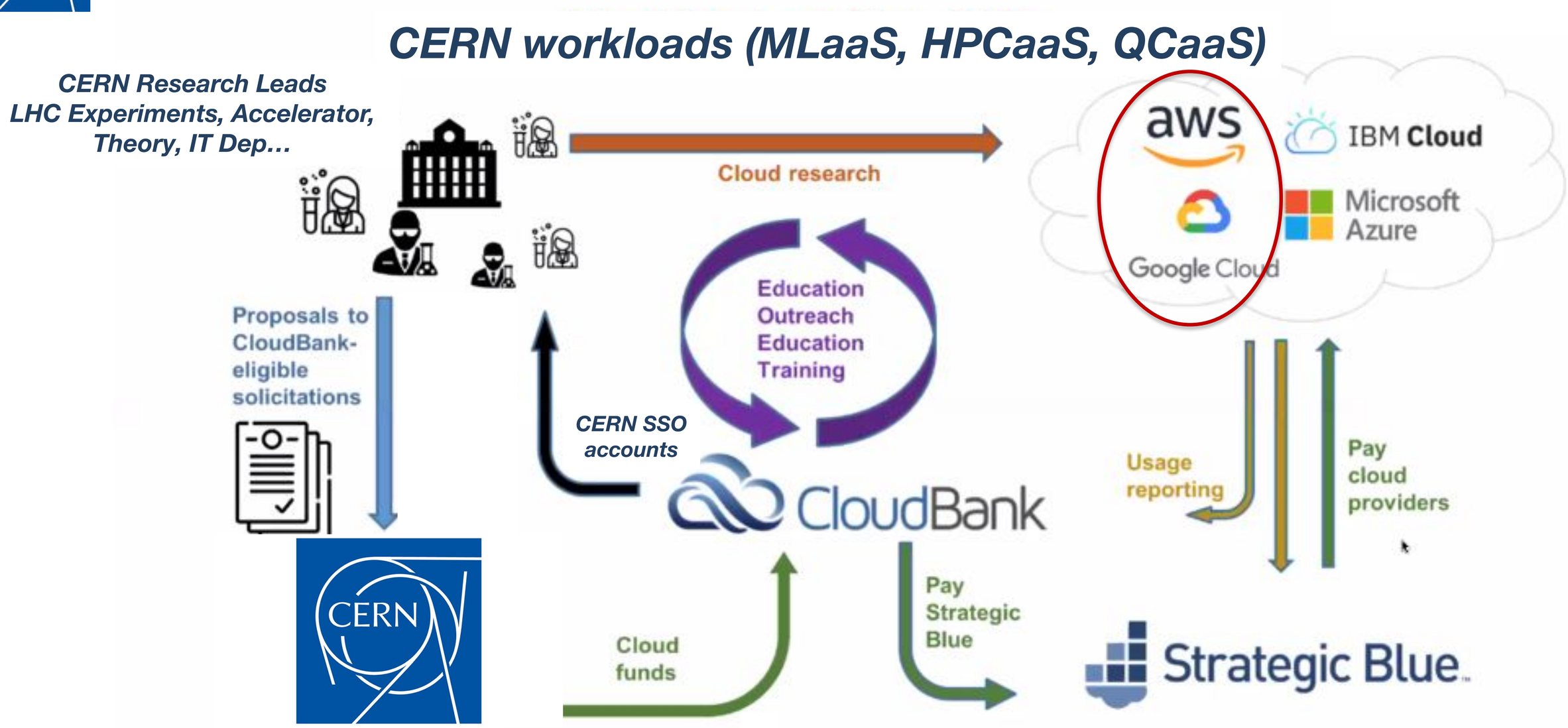
Background and Objectives

- Pilot commercial cloud usage by multiple research leads and organizational units
  - Provide structured access to multiple cloud providers with billing and data processing terms understood
  - Determine a model compatible with CERN's procurement and protect CERN financial interests
  - Determine if the model can be expanded in Europe to other organisations to monitor spending on their own contracts
- Pilot Deployments: Machine Learning (ML), HPCaaS, QCaaS use cases in AWS and GCP
- Training tutorials & support lines provided by cloud providers directly

What problem is NSF trying to solve?

- For NSF: Create an entity that facilitates and tracks the use of cloud by many grantees from diverse organisations
- For Researchers: remove “frictions” associated with access and management and help to educate them about cloud, building the necessary skills
- For Cloud vendors: structure the research sector as a new growing market

What is CloudBank?



Use cases deployed

Usage Deployment	Cloud Provider(s) / Runs Status
Studies of TPC distortion fluctuation calibrations and secondary vertex reconstruction	GCP ✓
Statistical analysis using ML training and inference relying on Kubeflow, Knative and common ML frameworks	AWS ✓, GCP ✓
Neural ODE model, LSTM model, and Bayesian LSTM models for SPS injection kicker thermal behaviour modelling; QCaaS: Reinforcement learning for beam steering	AWS ✓, GCP ✓
Train deep learning models using GPUs, via the CERN-SWAN framework in AWS	AWS ✓
Realistic analysis physics workflows	AWS ✓, GCP ✓
ML real-time event selection based and algorithm training on GPUs/TPUs and FPGAs using HLS-based libraries	AWS ✓, GCP ✓
Commercial cloud storage in the Escape DataLake and compute to integrated with the DataLake Jupyter lab, using DIRAC as workload manager	AWS ✓, GCP ✓
Training and deployment of GANs for simulation, statistical modelling of large datasets and training of models aiming at reconstruction	AWS ✓, GCP ✓
SWAN installation on GCP	AWS ✓, GCP ✓
DBaaS, container based and serverless (function/lambda) services and NetApp evaluation	AWS ✓
Evaluate the performance of openQCD to run massively parallel MPI jobs using GPUs & AWS EFA network	AWS ✓, GCP ✓
Advanced integration of GANs architectures with Kubeflow moving into setting up an ensemble ML/DL training approach	AWS ✓, GCP ✓

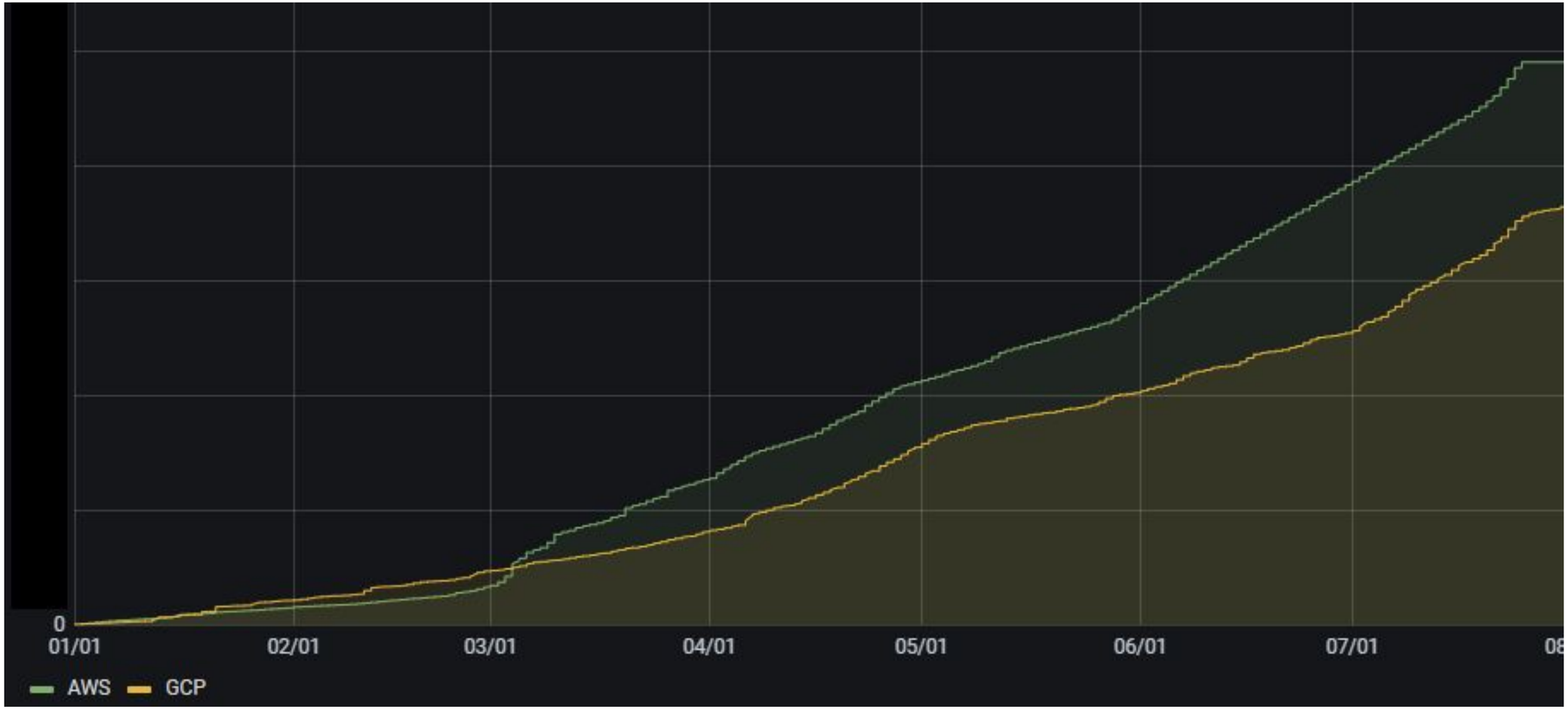
Billing Monitoring Infrastructure

Application runs in CERN Openshift and is conformed by the following core components:

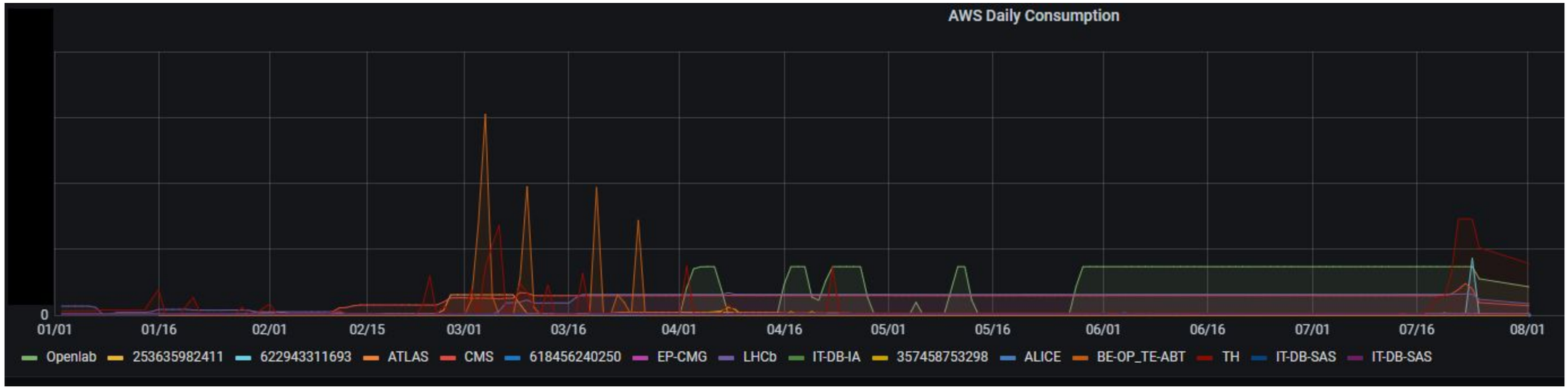
- CERN SSO: authentication for CERN users
- Grafana: dashboards displaying the data
- Prometheus: scrapes data from the exporters
  - AWS Exporter: gets billing data from AWS's reports
  - GCP Exporter: gets billing data from GCP's BigQuery
- InfluxDB: data persistence

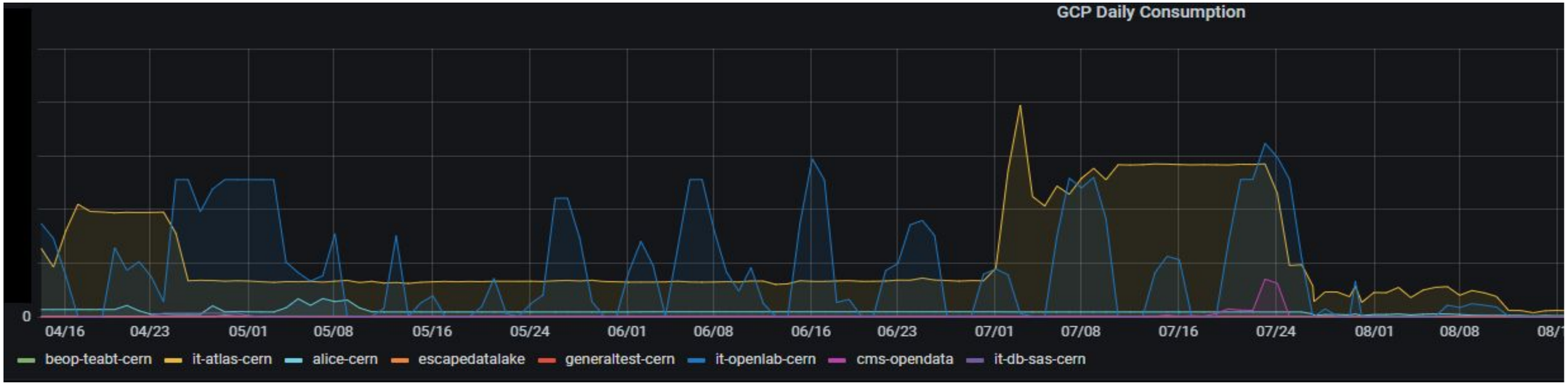
Consumption Monitoring per Cloud provider

Overall consumption per cloud provider



Billing Monitoring per Research Group





Data Privacy Analysis

Contractual analysis funded by the EC: Partners: CERN, BHO Legal & UCSD

Project Coordinator (EU) :

CERN

Organisation Type :

Research

CloudBank EU NGI

Experiment description

The CloudBank EU NGI experiment will accelerate the adoption of public cloud services in Europe's public funded research sector. The experiment's transatlantic existence would promote collaboration between the US and EU research communities in their adoption of public cloud services. The legal and contractual analysis of the CloudBank model in relation to European legislation will foster trust among procurement offices of public sector research organizations such as CERN, resulting in the simplification of their in-house cloud service procurement processes and expanding the model's applicability in a broader setting in Europe. The experiment will help to create a secure, sovereign and federated European business cloud infrastructure for research users based on Interoperability, open source software and open standards, and will help to identify synergies with EU commercial data infrastructure initiatives.

The experiment CloudBank EU NGI is to extend CloudBank through the GEANT Network and EduGAIN

Objectives: include the data privacy analysis as input for the future tender exercise; add a dashboard compliance view, so researchers can have an overview on these aspects (countries where data can be located, number and type of subprocessors, etc.)

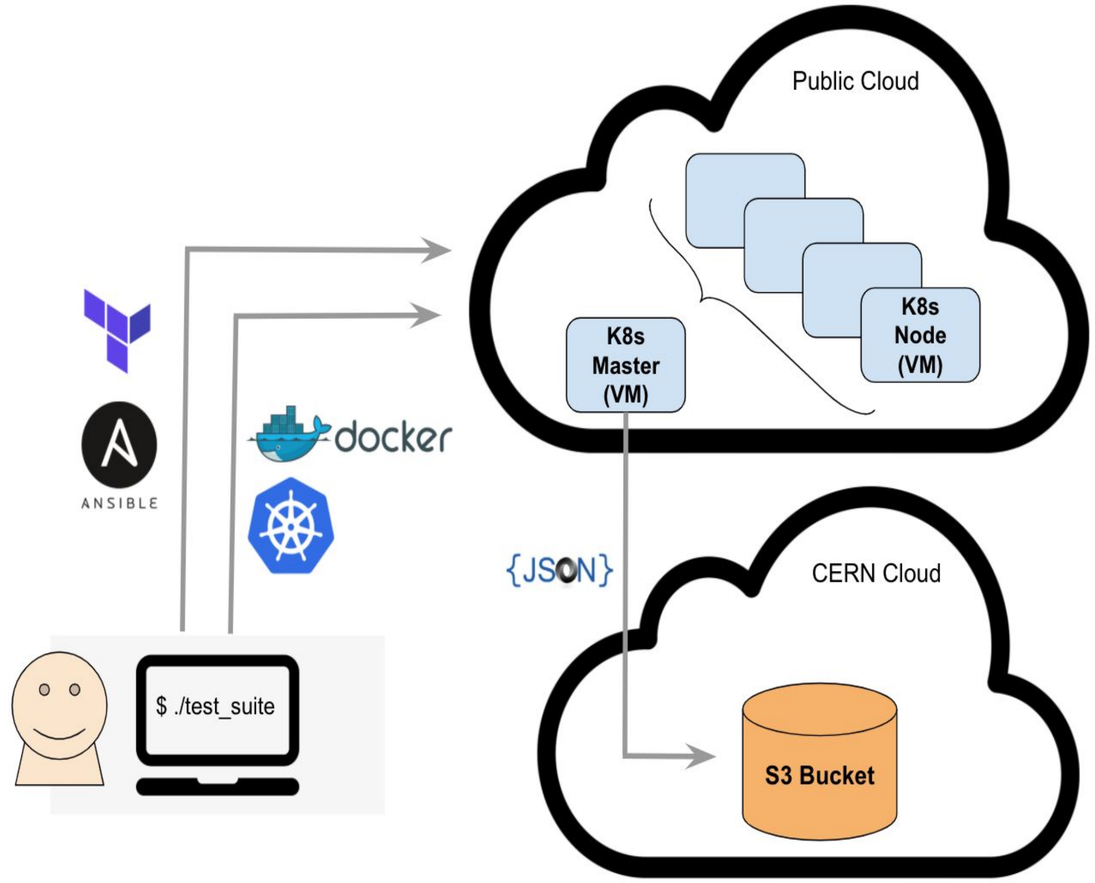
Relationships with EOSC & Gaia-X

- Determine if the model can be successfully used in the EOSC as a flexible means to access commercial cloud services at scale

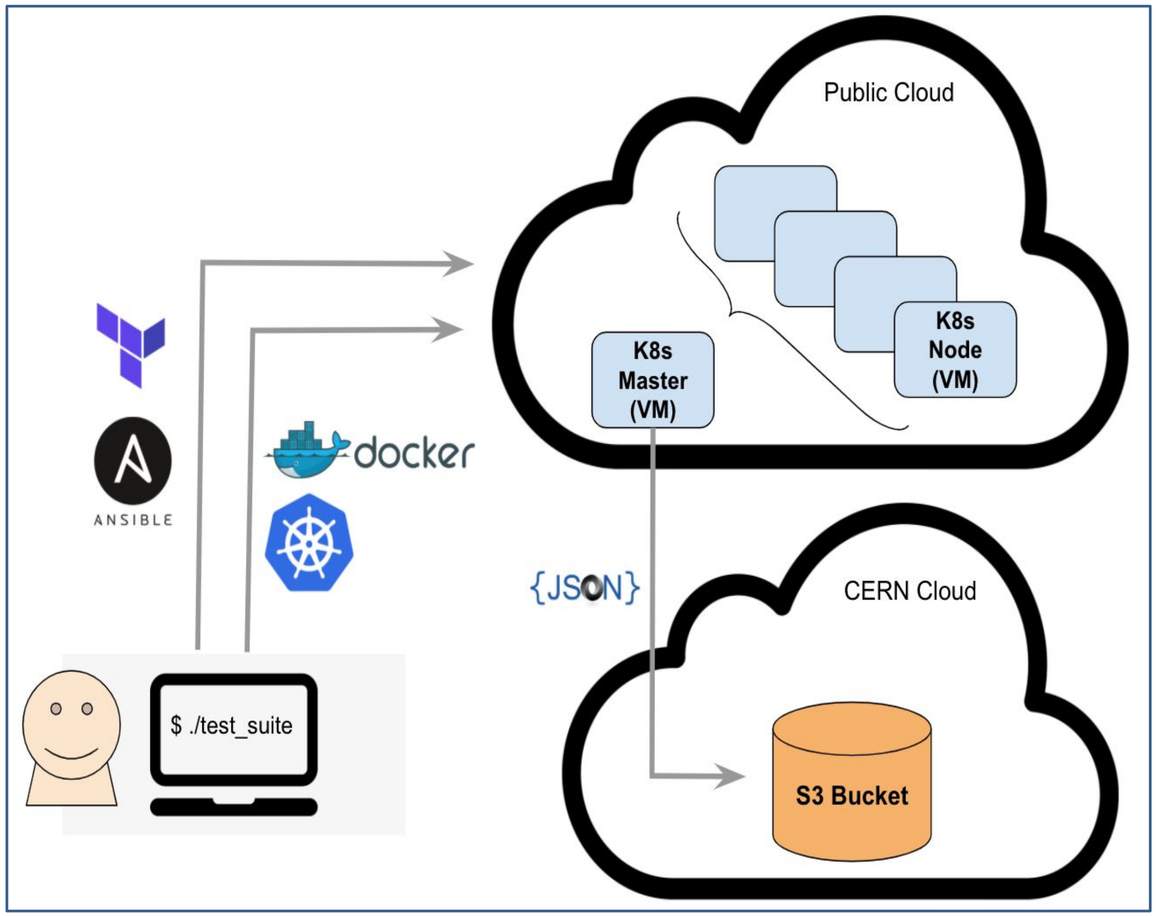
- Combine the open-source test suite framework to validate requirements as wide as possible, providing the researchers with working examples of deployments on different clouds

- Propose it as a model to track & monitor the contracts established under EC Frameworks (e.g. OCRE)

EUROPEAN OPEN SCIENCE CLOUD



Compliance: Validation Tests 2021



CloudBank Europe

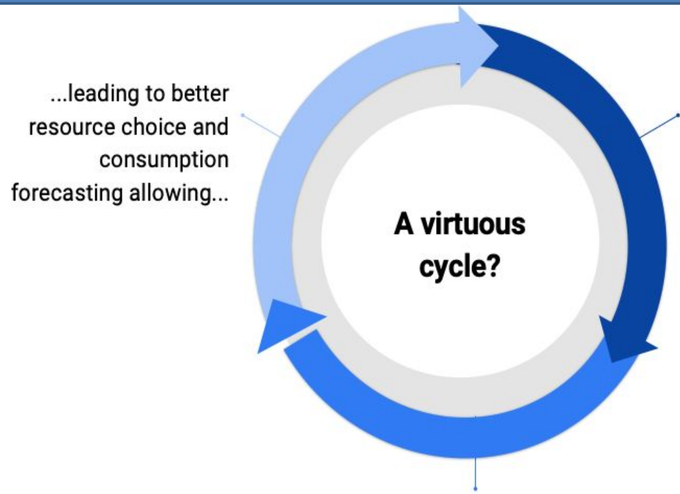
European Open Science Cloud use cases

Multi-domain research workload deployments

Access to cloud vendors via Federated IAM accounts

European Researchers

Scientific Proposals that make use of cloud services (IaaS, PaaS, SaaS)



Federated IAM accounts

Usage/Billing reporting via CloudBank EU Open Source Dashboard

Pay Cloud Services

Provision of "FinOps": Cost-Optimisation services

MoU CERN/UCSD: co-development of CloudBank EU

CloudBank

Allocate funds to be tracked

Pay "Clearing House" services

Strategic Blue

CCEX CLOUD COMMODITIES EXCHANGE

BHO LEGAL

Clearing House Services: Cost-Optimisation, Financial liability, VAT management, SLAs, TOMs for Data exchanges based on EU legislation (e.g. GDPR)

CERN, Other IGOs, EC, National Funders, etc.

"Bring your own Contract" or contracts via the "Clearing House"

Summary

- CloudBank EU is assessing a scalable organisation-wide model for procurement and use of cloud services across different Departments, Groups and Experiments
  - *”Is this cloud offering designed to the use of my research team?”, “How can I run multiple configurations for my software?”, “Is the cloud setup too time consuming?”, “Do they have GPUs at scale (multiple hundreds)?”, “I want to use it, but I need billing help”, “I have a small research team, I need access out of central job submission systems”, etc.*
- Expanding access to other organisations with transparent billing and data processing terms
- Adding a managed structure to the current diverse collection of trials and projects
- Providing a model to ***complement*** on-premise capacity ***NOT to replace***
- Building the set of necessary skills across Organizations to use modern heterogeneous services effectively and following best practices
- Determining the appropriate model for European organisations

Accelerating Science

Thank you!

cbp-team@cern.ch