

CloudBank for Europe

Marion Devouassoux¹, João Fernandes², Bob Jones³, Anna Manou⁴, and Inês Pinto Pereira da Cruz⁵

^{1,2,3,4,5}CERN, Information Technology Department, Geneva, Switzerland

Abstract. The vast amounts of data generated by scientific research pose enormous challenges for capturing, managing and processing this data. Many trials have been made in different projects (such as HNSciCloud and OCRE), but today, commercial cloud services do not yet play a major role in the production computing environments of the publicly funded research sector in Europe. Funded by the Next Generation Internet programme (NGI-Atlantic) from the EC, in partnership with the University California San Diego (UCSD), CERN is piloting the use of CloudBank in Europe. CloudBank has been developed by the UCSD, University of Washington and University of California, Berkeley with NSF grant support, to provide a set of managed services simplifying access to public cloud for research and education, via a cloud procurement partnership with Strategic Blue, a financial broker SME, specialised in cost management and optimisation. The European NGI experiment is provisioning cloud services from multiple vendors and deploying a series of use-cases in the domain of Machine Learning and HPCaaS, contributing to the scientific programme of the Large Hadron Collider. The main objective is to address technical, financial and legal challenges to determine whether CloudBank can be successfully used by Europe's research community as part of its global research activity.

1 Introduction

The growing use of new methods such as those based on Machine Learning (ML), Internet of Things (IoT), HPCaaS and Quantum Computing, together with the commoditization of specialised cloud based services and frameworks have introduced the need to assess new computing models for research environments. The growing variety of hardware infrastructure platforms at scale, performance-portability frameworks and open cloud-based orchestration systems has increased the availability of scenarios to be explored, changing from basic elastic provisioning of virtual resources, evolving to a transparent and adaptive smart cloud environment continuum, available for data intensive applications, from simulation to analysis. This evolution is putting in evidence the potential to introduce new, more performant and cost-effective technologies than are currently available on-premise at many research laboratories and adapted to modern research data processing requirements.

The bulk of the cloud services provided by CERN to support its scientific programme is currently provisioned as in-house resources managed by the IT department and hosted in the computer centre on-site. Unforeseen increased demand linked to the scientific programme as well as unexpected events, such as the COVID-19, can impact the ability to meet new demands. Similarly, enterprise level risks (including major infrastructure incidents, cyber-attacks and insufficient or delayed delivery of h/w resources) can impact CERN's ability to deliver services. In a similar way as High-Energy Physics computing has made a strategic move from mainframes to personal computers 25 years ago[1], public commercial cloud services have rapidly become commodity offers and can be strategically considered in a hybrid model, to be integrated into the current computing research environments. Previous experience with public clouds has shown that it is possible to integrate commercial cloud services from different providers into the CERN cloud provisioning model[2] but a number of challenges need further investigation, such as:

- ability to rapidly increase heterogeneous capacity without the need to re-tender and associated delays;
- avoid vendor lock-in by having the possibility to rapidly select an alternative cloud service supplier;
- be able to monitor and control the consumption of the procured services assigned to multiple, independent administrative units, tracked and billed separately.

An approach based on exploring the services of a financial cloud broker has the potential to address such challenges of cloud service provisioning at scale.

1.1 Objective

At the very start of cloud adoption, the technical capability typically takes precedence over cost optimisation. However this is not a long term sustainable position. Taking this core aspect into account, University California San Diego

(UCSD), University of Washington and University of California, Berkeley, with the support of an NSF grant, have developed CloudBank[3], a set of managed services to simplify access to public clouds for research and education. CloudBank has established a cloud procurement partnership with Strategic Blue[4], a financial services SME specialising exclusively in cloud procurement, cost management and optimisation. CERN in partnership with UCSD has been funded by the NGI Atlantic programme[5] from the EC to experiment the use of CloudBank in Europe.

The objective of the CloudBank EU NGI experiment is to leverage on the aforementioned work and pilot the use of the service under European legislation to provision services from multiple cloud providers. These services will initially support the deployment in Amazon Web Services (AWS) and Google Cloud Platform (GCP) of a series of challenging use-cases in the domain of Machine Learning (ML) and HCPaaS that are contributing to the scientific programme of the Large Hadron Collider. The experiment will determine if CloudBank can be successfully used by Europe's research community as part of a global research activity addressing not only technical, but also financial and legal challenges to determine if such an approach can be applied in a European setting. Ultimately, the results of this analysis will lay the foundation to explore the mechanisms to extend CloudBank to European cloud providers, aligning it with the rise of European digital sovereignty, strengthening data self-determination for public European research and overall, lead to a simplification of in-house public cloud service procurement processes for the public research sector, integrating these into an hybrid research computing model, in view of a smooth transition to a heterogeneous cloud infrastructure.

2 CloudBank EU Experiment

The procurement of commercial cloud-based services has increased at CERN in recent years and the IT department has gained some experience in procuring cloud services to support its scientific programme. Procurement activities by CERN have also been supported by the EC through projects including Helix Nebula[6], PICSE[7], HNSciCloud[8] and more recently OCRE[9] and ARCHIVER[10]. In spite of very advanced initiatives, when onboarding public commercial cloud services, challenges remain in a combination of technical, financial (cost-effectiveness) and regulatory (data privacy).

Such challenges can span from delivering network traffic from cloud provider data centers to Research and Education Networks (NRENs), cloud-to-cloud provider traffic, or from public clouds to research organisations, to aspects related to cost-optimization considering that cloud environments have a substantial number of optimisation options. Depending on the workload and the real-life cloud scenario, one needs to consider a balance of performance vs. cost. Several options can be explored at different performance/cost ratios depending on the workload in question: type of resources (e.g. CPU, GPU, TPU), cloud vendor regions availability, type of instances (on-demand vs spot/preemptible) and network egress charges. Most of these elements are driven by the application. In the case of ML on cloud for example, modern workloads are used in what is usually called burst/auto-scaling mode, paying only what is used as a cost-effectiveness function, very different for example from workloads using a more traditional batch submission, that relies on predefined reserved capacity. When compared with on-premise capacity, cloud is a commodity product with increased flexibility. To use it effectively, it's not a simple case of "Lift and Shift" i. e., creating VMs and running software on top. One often needs a degree of platform optimization to get a similar level of throughput where underlying hardware infrastructure performance can be efficiently pushed to its limits. Therefore, the primary objective when using cloud is to minimise and optimise resources used for each individual application whilst when running an on-premise service, the goal is to keep the resources as busy as possible with a continuous flow of applications in order to maximise throughput and Return On Investment (ROI) for the hardware.

For successful onboarding of the potential offered by public clouds in the current research computing programmes, an analysis on models of data governance relationships with the cloud provider(s) must be performed taking into account applicable legislation. This includes assigning responsibilities for data processing and identifying financial and legal risks. In addition, the planning and practical validation of exit strategies (cloud to cloud provider or cloud to on-prem) is also required to mitigate potential vendor lock-in risks.

The CloudBank experimental service intends to address many of these combined challenges: for example, it will enable data intensive research use cases to accommodate cloud resources at scale on their data management and distribution architectures. The use cases deployed under the experiment can potentially assess the capabilities to transfer an increasing volume of research data allowing to better understand how research use cases are using cloud services and the underlying network with sustained traffic, exploring further for example the possibilities to use the transcontinental networks of large public cloud providers. Concerning the financial aspects, the CloudBank financial operations allow research performing organisations (RPOs) to understand in detail where cost is incurred, eliminate waste of unnecessary consumption and optimise procurement of the necessary.

TITLE	ID	TYPE	START	END	AWARD AMOUNT	BALANCE	PI	STATUS	EDIT
CERN Pilot	20200812		2020-08-22	2021-06-30	\$0.00	\$0.00 (100%)		Active	edit
CERN Test Fund	20200824	Type: Discretionary	2020-08-24	2020-12-31	\$20.00	\$19.96 (100%)		Active	edit
CERN Pilots	20201014		2020-10-14	2021-06-30	\$0.00	\$0.00 (100%)		Active	edit
CERN Pilot	20201015		2020-10-15	2021-06-30	\$0.00	\$0.00 (100%)		Active	edit
CERN Pilot	20201016		2020-10-16	2021-06-30	\$0.00	\$0.00 (100%)		Active	edit
CERN Pilot	20201017		2020-10-17	2021-06-30	\$0.00	\$0.00 (100%)		Active	edit
CERN Pilot	20201019		2020-10-19	2021-06-30	\$0.00	\$0.00 (100%)		Active	edit
CERN Pilots	20201020		2020-10-20	2021-06-30	\$0.00	\$0.00 (100%)		Active	edit
CERN Pilot	20201020		2020-10-20	2021-06-30	\$0.00	\$0.00 (100%)		Active	edit
CERN Pilot	20201027		2020-10-27	2021-06-30	\$0.00	\$0.00 (100%)		Active	edit
CERN Pilot	20201125		2020-11-25	2021-06-30	\$0.00	\$0.00 (100%)		Active	edit

Fig. 1. CloudBank Dashboard piloting the CERN ML and HPCaaS use cases.

To maximise effectiveness, cost management must include elements of cost transparency and optimisation, without losing focus on value creation. For continuous improvement, technical and financial factors need to be aligned and a framework created to evaluate the efficiency of the end-to-end cost management and optimisation process.

The CloudBank EU experiment will also contribute to the establishment of a regulatory safe, sovereign and multi-cloud commercial infrastructure for research users, based on interoperability, open source software containerisation and open standards as valid technical and organisational measures under European legislation.

2.1 Use Cases

The initial set of use cases is primarily composed of ML and HPCaaS workloads derived or evolved from deployments in early initiatives, such as HNSciCloud, expanded for the NGI Experiment by the use cases proposers themselves (Figure 1). There are ten use cases submitted and deployed so far, from different administrative structures within CERN including Experiments, Accelerator sector, Theory and IT Departments that address most phases of the research data processing workflow[11]. Examples are depicted in Figures 2 and 3.

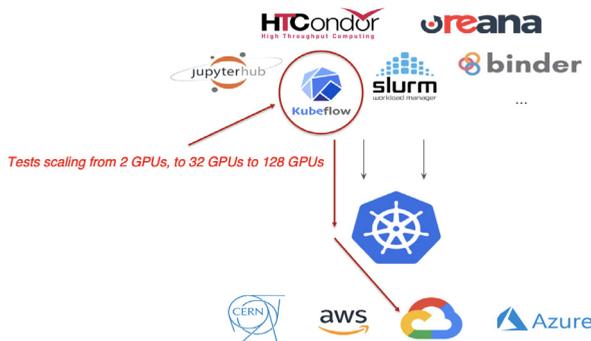


Fig. 2. Example use case CERN IT/ATLAS (courtesy Ricardo Rocha, CERN/IT).

Figure 2 depicts a cloud deployment based on Kubeflow for specialised resources across different providers

(example uses Google Cloud). Kubeflow on top of Kubernetes offers the possibility to manage the full ML workload lifecycle, from data preparation to model training (including distributed training at scale, data serving and storage).

Priority has been given to those use cases that allow the CloudBank experiment to cover a wide range of applications, experiments and departments while limiting the procurement investment and minimising any additional support effort required from IT department personnel. Another important criteria was to select use cases that would not expose personal data at this stage, as the legal analysis for personal data processing is taking place in parallel.

The deployment of ML use cases using cloud-based services is particularly relevant. While ML becomes increasingly important for the LHC computing models, the ability to scale up remains a major issue yet to be addressed. This is particularly pertinent for use cases where a large number of accelerator architectures (GPUs, FPGAs and similarly, more recently, Graphcore IPU) are required or can greatly improve the performance of training and inference algorithms.

A core objective of the experiment is to demonstrate a hybrid ML service model where the IT department can easily complement/extend in-house accelerator capacity (i.e. GPU) to hundreds of hardware accelerator instances.

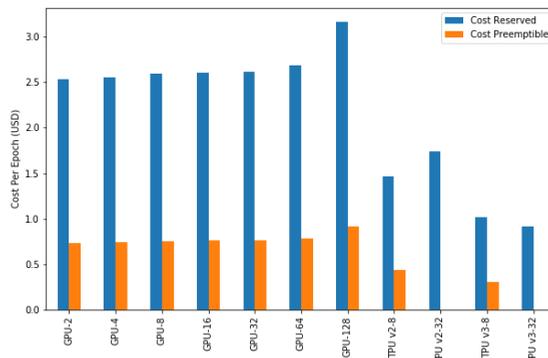


Fig. 3. An example of the studies performed for the 3DGAN use case: cost of running one epoch training on GCP. The cost per epoch remains similar for all GPU combinations, though the training time improves almost linearly. The optimisations introduced in the 3DGAN distributed training algorithm, make the use of TPUs the most cost-effective choice for this use case. (courtesy Renato Cardoso, CERN openlab).

Figure 3. depicts the results of the 3DGAN use case, a Generative Adversarial Network prototype designed to simulate electromagnetic calorimeters output. This use case explores highly parallel solutions to speedup the GAN training process. In particular, it compares TPU-based solutions to multi-GPU setups, analysing performance in terms of scaling efficiency and physics accuracy. The technical objectives when combined with the financial aspects establish a clear sustainable model for the CloudBank experiment outcomes. The results of the use case deployment will provide an assessment of the capabilities a financial broker may provide in terms of billing controls and using it as an additional source of data for cost verification and prediction without interfering with the technical offers of each of the cloud providers made available.

2.2 Brokering & Cost Tracking

In the market of cloud services, a financial broker functions in the same manner as it does in more traditional markets. The broker matches the demands of users with the suppliers, aiming to succeed in settling the best financial agreement between the two sides of the market[12]. In addition, it delivers billing services as well as finance and risk management capabilities so that clients are able to trade on their preferred terms, even when those terms mismatch with the sellers' preferred terms. A financial brokerage model for cloud, profitable for the broker, offers reduced costs for cloud users, and generates more predictable demand flow for cloud providers. The model offers access to multiple cloud providers making it contractually simpler to switch between providers without intervening in the technical delivery of cloud services.

In this context, CloudBank provides innovative financial engineering options that give researchers more flexible cloud terms tailored for their needs and contributes to the sustainability of its operations. The current US instance of CloudBank helps NSF by bundling multiple small requests from individual NSF grantees into a bulk request to cloud providers, disincentivizing more costly direct connections. Through this aggregation and innovative financial contract types, CloudBank passes along savings to NSF and researchers that would otherwise be unavailable to them.

The financial broker currently associated with CloudBank is Strategic Blue. Strategic Blue is a UK SME that offers three broad types of services, each of which is relevant and useful to CERN with regard to the use cases deployment during the CloudBank EU experiment activities:

- Consulting / Training – access to Strategic Blue's insight on cloud procurement and pricing best practice,

based on background in commodities trading.

- Cloud Options Pricing Insights – access to proprietary analysis of published historical and current public cloud pricing from vendors such as Amazon Web Services and Google Cloud Platform.
- Cloud Options – Strategic Blue’s core business involves helping clients who find they have a difference between the terms on which they would ideally like to purchase cloud computing, and the terms on which the cloud provider(s) are willing to sell at their lowest prices. Strategic Blue steps into the billing chain between cloud customer and cloud provider(s) and injects the required combination of a) billing services, b) finance and c) risk appetite in order to achieve the optimal financial result for all concerned.

CERN already has experience with Strategic Blue. In the past, it has successfully used the Cloud Options Pricing Insights services of Strategic Blue during the now completed and award winning HNSciCloud project as a means of benchmarking the bids received against market prices. This previous experience makes Strategic Blue understanding of CERN’s administrative and technical processes and collaboration better and of greater value.

During the preparation phase of the experiment, an initial financial analysis was performed to estimate the total cost of the deployment of the use cases. This work was performed by the CERN IT CloudBank Office team in partnership with the lead researcher of each use case.

The process followed to estimate the use case costs is as follows:

1. The use case lead researcher provides through a questionnaire a realistic description of the expected usage and types of resources needed. The description also identifies the public commercial cloud providers (AWS or GCP) preferred by the lead researcher and the estimated timeline for deployment.
2. The CERN IT CloudBank Office calculated cost estimates for each use case according to the data provided by each lead researcher. The calculations were done based on a financial formula provided by the financial broker Strategic Blue, responsible for cost management.

The financial formula takes the following elements into account in the cost calculation of an instance:

- Quantity of compute instance (Q_c)
- On demand price for the compute instance (ODQ_c)
- Quantity of storage (Q_s)
- On demand price for the storage (ODQ_s)
- Hours of run per week (h)
- Number of weeks (n)

The equation used to determine the price (P) of an instance is the following (1):

$$P = (Q_c * ODQ_c * h * n + Q_s * ODQ_s * h * n) * 1.2 \tag{1}$$

The instance cost is multiplied by 1.2 to cover additional services and contingency. The estimated cost of a use case consists in the sum of the estimated costs of all the instances. The pricing data used for the calculation are estimated values extracted from the service catalogues of each cloud provider. Therefore, the nature of the formula is generic and its purpose is to provide a ballpark for the cost estimation. For the same reason, the term egress has been excluded from the formula, since egress costs depend on contractual agreements between the cloud provider and the institution.

These cost estimates serve as a basis for the IT CloudBank Office to allocate funds for the approved use cases during the CloudBank EU experiment, that correspond to the *Award Amount* column in Figure 1, to be managed and tracked by CloudBank, as described.

The second part of cost management implies realistic consumption tracking across organisational units, as a major indicator to improve IT service forecasting, often recurring to visualization and analytics tools. Tracking becomes therefore crucial in order to achieve transparency and trust, across both users and IT services. In general terms, a cloud service brokerage system is described as providing “a single and common interface through which consumers can provision and manage their services on multiple clouds”[13]. As the brokering is scoped at the financial rather than technical level, this creates a path to expand cloud access using transparent billing, without interference in the technical delivery. Such a model could be expanded with funds being allocated transparently and directly by multiple different funding sources to each of the corresponding organizational units.

The US instance of CloudBank is using a commercial product based on Nutanix Beam[14] to provide cost governance capabilities that is compatible with AWS and Azure. Strategic Blue also provides a service for cost tracking based on AWS Quicksight[15] (Figure 4).

As part of the activities of the CloudBank EU NGI experiment, a generic mechanism is being developed to

collect consumption and usage data from the use cases deployed over the cloud providers that are part of the experiment. The motivation behind the conception of the generic mechanism is to build a single dashboard to display aggregated data from multiple cloud providers, allowing consumption control, metrics discovery, usage patterns definition, etc. based on open source tools with a modular architecture not locked to any particular cloud vendor.

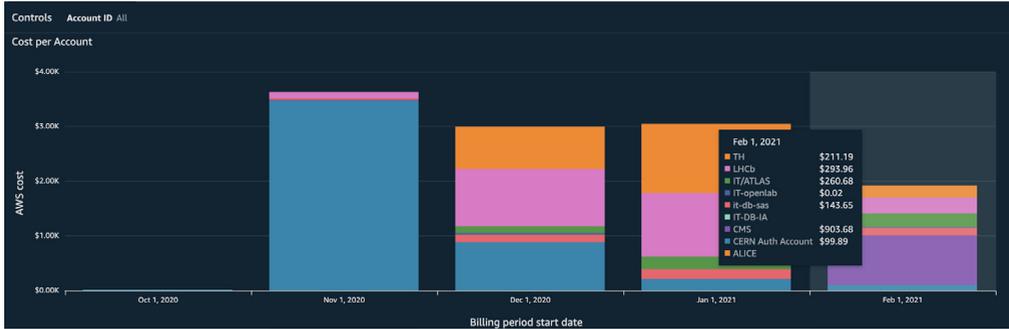


Fig. 4. Billing data for the CERN pilot use cases on Strategic Blue’s Dashboard service.

As a backend for this mechanism, the Prometheus[16] tool was evaluated and eventually selected. Prometheus is an open-source and service monitoring system and time-series database with wide community support, part of the Cloud Native Computing Foundation[17]. It allows collecting metrics of all kinds from configured targets at given intervals, exporting them, displaying and visualising results. A Proof of Concept (PoC) deployment determined a) how data from a cloud provider could be made available to Prometheus as metrics and b) how to export and display the cost metrics.

The first step was to build a Prometheus exporter in Golang[18]. An exporter is essentially an application that exports existing metrics from third-party systems, in our case cloud providers, as Prometheus metrics. Exporters get the target’s metrics via a network protocol, so the PoC had to create a client to connect to the cloud provider via the programmatic APIs made available, request the billing data and finally expose it via the network protocol that behaves as a target. At this point, the exporter can then retrieve the required metrics. The PoC has been documented[19] and initial metrics such as hours of CPU and GPU usage, number of Virtual Machines, and total consumption costs per project are being collected.



Fig. 5. PoC Grafana dashboard showing expenditure per project for the month of February 2021.

The following step was to draft a mockup display based on Grafana[20] for the visualization of multiple metrics (Figure 5). The overall goal is to enhance dashboards to complement the billing and data governance capabilities of CloudBank. Grafana is an open-source tool tightly integrated in the Prometheus architecture providing a multi-platform analytics and interactive visualization web application. It provides charts, graphs, and other supports for the visualization of metrics. Grafana also has a built-in support for Prometheus which makes configurations faster and allows a certain level of atomization when running the solution.

Finally, in terms of infrastructure, the PoC (Figure 6) is being deployed using the central Openshift service running on-premise at CERN[21]. Openshift[22] is a platform that enables the deployment of applications using containers. The PoC is deployed as a Docker[23] container in an Openshift cluster, providing full automation to build, deploy and scale out.

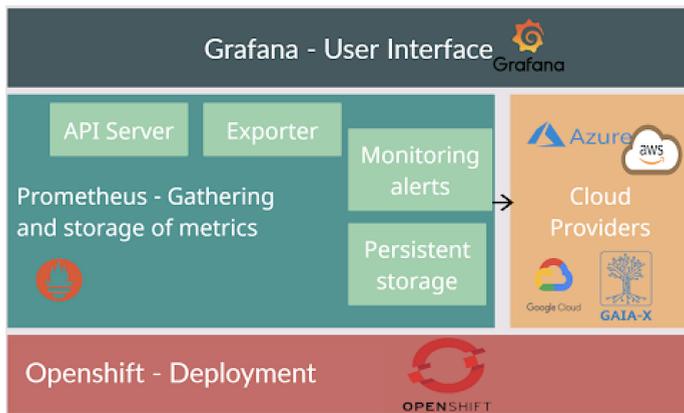


Fig. 6. High Level architecture of the PoC Dashboard based on OpenShift, Prometheus and Grafana.

Having the PoC on CERN premises will allow CERN to retain control over the billing information, ensuring confidentiality of the data collected, as best practice and obligation for contractual management with commercial partners such as financial brokers and public cloud vendors. It also allows it to benefit from “in house” deployments based on Prometheus as a widely supported technology both at CERN and in general across the HEP communities.

2.3 Data Processing

The ML and HPCaaS use cases being deployed contain no associated personal data. At the same time, one of the determining objectives of the experiment is to validate CloudBank under EU based regulatory frameworks. As part of the NGI experiment, CERN is engaging with a legal partner to carry out a data protection contractual analysis with respect to European legislation (such as GDPR). The analysis will compare possible models of contractual relationships with major cloud provider(s). The goal is to have a clear mapping and breakdown of responsibilities for processing and sub-processing of personal data on the cloud, contractual solutions that will generally lead to a simplification of in-house public cloud service procurement processes.

In addition, the CloudBank EU experiment will also contribute to the definition of a data security model for research organisations using public commercial cloud providers. Aspects to be covered will include confidentiality of data, mitigation plans for DDoS attacks, antivirus and antimalware, security audits, encryption of data-at-rest and in transfer as a set of technical measures demonstrating alignment with ISO 27001 certification and implementation of the guidelines of the EU Agency of Cybersecurity (ENISA[24]), in aspects such as information security risk assessment, monitoring, data breaches and internal audits.

3 Initial Results & Future Work

The CloudBank NGI Experiment started in November 2020 with a few preliminary results at the time this paper was produced. For what concerns the use cases, 9 have started deployment both over AWS and GCP with others planned to start by the end of Q1 2021. The current CloudBank instance in the US is fully setup and can be used to pilot access to each of the billing accounts using CERN Single Sign On[25].

In what concerns the billing and tracking PoC, data was successfully retrieved from Google’s BigQuery service[26] using a generic unlocked approach, allowing the developed exporter module to receive metrics that are made available to the Prometheus backend. Metrics breakdown per project including costs could already be observed in Prometheus displayed in Grafana. As the PoC is based on Docker, there is no need to install Prometheus and Grafana as both official containers are available in the Docker Hub[27][28][29]. Besides allowing the use of containers, it also makes possible the provisioning of data sources and dashboards in Grafana[30], meaning that if the configuration files are added to the exporter, the metrics gathered by Prometheus will automatically be displayed in a dashboard already configured, eliminating the need to reinstall or reconfigure both Prometheus and Grafana components. After an initial analysis of the billing data displayed, the PoC is now requesting billing data once per day, where differences in consumption profiles can be already spotted: some projects have number of CPU hours increasing very rapidly as the use of Virtual Machines, whilst in others the number of resources is steady with no significant variation.

Some of the next development steps involve the enhancement of the Grafana dashboards display and role authentication so that the dashboard displays only billing data related to a given researcher or organisational unit. This will be followed by adding configuration options correlating type of services used, time periods and costs. Deployment on Openshift will be consolidated to make it independent of the underlying operating system, accelerating both development and deployment. Going into a pre-production scenario, high-availability will be a factor to consider, to

have service running redundantly in the event of system failures.

The modular support of data from more cloud providers will consolidate the PoC. AWS[31] is the next cloud provider to be integrated in the exporter. Once diverse data sources are supported (such as multiple cloud providers), the objective is to have the billing service producing usage reports that can help IT staff and users to make decisions to optimize usage, predict consumption for a given set of specific resources and even compare cloud provider offers in a ratio of cost vs. performance.

Other improvements are also envisioned such as the creation of monitoring alerts, support of persistent storage for metrics data, calculate trends and detect billing anomalies by comparing different sources of data (for example data retrieved from the providers vs. invoice data or direct serverless functions queries on running resources).

Resuming, the enhancements and next development stages foreseen are the following:

1. Extend Grafana dashboard configurations
2. Consolidate Deployment on Openshift
3. Support Federated Authentication and role based Authorization
4. Extend support of data from additional cloud providers
5. Create monitoring alerts
6. Support persistent storage and for metrics data and dashboard
7. Display consumption trends, predictions and detect billing anomalies

4 Conclusions

This NGI experiment will accelerate the adoption of public cloud services in Europe's public funded research sector. The transatlantic nature of the experiment will increase cooperation between US and EU research communities in their uptake of public cloud services.

The legal and contractual assessment of the CloudBank model with respect to European legislation, will build trust among the procurement offices of public sector research organisations such as CERN and lead to a simplification of their in-house cloud service procurement processes. The CloudBank EU NGI experiment will report publicly the progress made against its objectives, the lessons learned, with a recommendation for whether there is a case for expanding the model in a wider audience and timeframe.

References

1. 25 years ago: the strategic move to PCs in high-energy physics: <https://home.cern/news/news/computing/25-years-ago-strategic-move-pcs-high-energy-physics> (2021), accessed: 2021-02-10
2. J. Fernandes, B. Jones, S. Yakubov, A. Chierici, *HNSciCloud, a Hybrid Cloud for Science*, EPJ Web Conf. 214 09006 (2019), https://www.epj-conferences.org/articles/epjconf/abs/2019/19/epjconf_chep2018_09006/epjconf_chep2018_09006.html
3. CloudBank: <https://www.cloudbank.org/> (2021), accessed: 2021-02-10
4. Strategic Blue: <https://strategic-blue.com/> (2021), accessed: 2021-02-10
5. NGI Atlantic Programme: <https://ngiatlantic.eu> (2021), accessed: 2021-02-10
6. Helix Nebula: <https://www.helix-nebula.eu> (2021), accessed: 2021-02-10
7. PICSE <http://www.picse.eu> (2021), accessed: 2021-02-10
8. HNSciCloud Project: <https://www.hnscicloud.eu/> (2021), accessed: 2021-02-10
9. OCRE Project: <https://www.ocre-project.eu/> (2021), accessed: 2021-02-10
10. ARCHIVER Project: <https://www.archiver-project.eu/> (2021), accessed: 2021-02-10
11. Cloud Broker Pilot description and Use Cases: <http://cern.ch/go/TPm7> (2021), accessed: 2021-02-14
12. Rogers, O., Cliff, D. *A financial brokerage model for cloud computing*. J Cloud Comp **1**, 2 (2012)
13. ISO/IEC 19831:2015 Cloud Infrastructure Management Interface (CIMI) Model and RESTful HTTP-based Protocol — An Interface for Managing Cloud Infrastructure: <https://www.iso.org/standard/66296.html> (2021), accessed: 2021-02-02
14. Nutanix Beam: <https://www.nutanix.com/products/beam> (2021), accessed: 2021-02-02
15. AWS Quicksight: <https://aws.amazon.com/quicksight/> (2021), accessed: 2021-02-02
16. Prometheus: <https://prometheus.io/> (2021), accessed: 2021-02-02
17. Cloud Native Computing Foundation: <https://www.cncf.io> (2021), accessed: 2021-02-02
18. The Go Programming Language: <https://golang.org/> (2021), accessed: 2021-02-02
19. Writing exporters | Prometheus: https://prometheus.io/docs/instrumenting/writing_exporters/ (2021), accessed: 2021-02-02
20. Grafana: <https://grafana.com/> (2021), accessed: 2021-02-02

21. PaaS Web Application Hosting Service: <https://information-technology.web.cern.ch/services/paas-web-app> (2021), accessed: 2021-02-02
22. Red Hat OpenShift: <https://www.openshift.com/> (2021), accessed: 2021-02-02
23. Docker: <https://www.docker.com/> (2021), accessed: 2021-02-02
24. ENISA: <https://www.enisa.europa.eu/> (2021), accessed: 2021-02-02
25. E. Ormancey, *CERN single sign on solution*. Journal of Physics: Conference Series. 119. 082008. 10.1088/1742-6596/119/8/082008 (2008)
26. Google BigQuery: <https://cloud.google.com/bigquery> (2021), accessed: 2021-02-10
27. The official Grafana docker container Docker Hub: <https://hub.docker.com/r/grafana/grafana/> (2021), accessed: 2021-02-02
28. The official Prometheus docker container Docker Hub: <https://hub.docker.com/r/prom/prometheus/> (2021), accessed: 2021-02-02
29. Docker Hub Quickstart | Docker documentation: <https://docs.docker.com/docker-hub/> (2021), accessed: 2021-02-02
30. Provisioning | Grafana Labs: <https://grafana.com/docs/grafana/latest/administration/provisioning/> (2021), accessed: 2021-02-02
31. Amazon Web Services (AWS): <https://aws.amazon.com/> (2021), accessed: 2021-02-20