

Open Call 2

CloudBank EU NGI

Deliverable 3: Experiment Results and Final Report

Authors	João Fernandes (CERN), Anna Manou (CERN), Robert Jones (CERN), Shava Smallen (UCSD)
Due Date	31 st November 2021
Submission Date	31 st of January 2022
Keywords	CloudBank, UCSD, CERN, Cost Optimisation, MLaaS, HPCaaS, QCaaS, Research Community, "Bring your own Cloud Contract"

Deliverable 3: Part I

Analysis, results, and wider impact

Together with this report, the following documents are delivered:

- Annex I: Preliminary assessment from BHO legal regarding the applicability of the CloudBank model in Europe (**in revision by CERN Legal Teams**)
- Annex II: CloudBank security plan produced by UCSD
- Annex III: Full list of use cases deployed during the CloudBank EU experiment

1 Abstract <Max 100 words>

The vast amounts of data generated by scientific research pose enormous data processing challenges. CERN and UCSD (US) are experimenting the use of CloudBank in Europe. CloudBank provides a set of managed services to access public cloud for research, supporting specialised cost management and optimisation. The European NGI experiment objective is to address technical, financial, and legal challenges, determining whether CloudBank can be successfully used by Europe's research communities. **The preliminary legal analysis of CloudBank EU concludes the model being pursued does not raise any overriding legal concerns with regard to contractual arrangements and data protection law in Europe.**

2 Project Vision

The CloudBank EU NGI experiment will accelerate the adoption of public cloud services in Europe's publicly 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 analysis of the CloudBank model with respect to European legislation, builds trust among the procurement offices of public sector research organisations such as CERN, leading to a simplification of their in-house cloud service procurement processes and expanding the applicability of the model in a wider setting in Europe. The experiment will contribute to the establishment of a safe, sovereign and federated European commercial cloud infrastructure for research users, based on interoperability, open source software and open standards as measures for vendor lock-in prevention, and establish the path to find synergies with commercial EU data infrastructure initiatives. In the US, CloudBank's primary mission, known as "CloudBank Research", is to support the US National Science Foundation (NSF) in enabling US researchers to adopt the cloud services of the public cloud infrastructure providers. At the core of the adoption strategy is a streamlined, centralised access provision to multiple cloud vendors, integrating capabilities such as cloud budget management or spend alerting capabilities as well as research-specific cloud training tutorials for academic researchers. The CloudBank EU NGI experiment aimed to expand this vision to Europe beyond CERN, reaching out to Research organisations, offering and enabling a framework for efficient management of public cloud contracts from diverse funding sources, leveraging and improving some of the processes and technology built for CloudBank Research. The US partner is actively involved in key activities of CloudBank EU NGI experiment. The team of UC San Diego organised and participated in a series of webinars for Higher Education CIOs and University IT Managers illustrating how CloudBank can save university administrators time and money, discussing cloud management and optimization strategies, taking the guesswork and risk out of using for researchers who need the power and flexibility of cloud services while maintaining control of their computing environments.

CloudBank promotes a model for cost-effective use of cloud services where risks and impacts for organisations are mitigated and managed, stimulating a shift of approach for research organisations willing to bring their own cloud contracts.

In case of a future European expansion, CloudBank will foster strategic links to future procurement activities of the European Open Science Cloud (EOSC)[1] and offer cloud access and cost tracking capabilities for collaborations in several research domains. European research organisations other than CERN have already expressed interest in probing CloudBank to manage their recently signed cloud contracts.

3 Details on participants (both EU and US)

The CloudBank EU NGI experiment will accelerate the adoption of public cloud services in Europe's publicly 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 analysis of the CloudBank model with respect to European legislation, will build trust among the procurement offices of public sector research



organisations such as CERN, leading to a simplification of their in-house cloud service procurement processes and expanding the applicability of the model in a wider setting in Europe.

The experiment will contribute to the establishment of a safe, sovereign and federated European commercial cloud infrastructure for research users, based on interoperability, open source software and open standards as measures for vendor lock-in prevention, and establish the path to find synergies with commercial EU data infrastructure initiatives. This alignment will stimulate closer technological cooperation, in a “European way” as a part of a global research activity, between the public research sector and the private sector across Europe. 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 to a wider audience and proposed timeframe.

In the US, CloudBank’s primary mission, known as “CloudBank Research” is to support the US National Science Foundation (NSF) in enabling US researchers to adopt the cloud services of the public cloud infrastructure providers. At the core of the adoption strategy is a streamlined, centralised access provision to multiple cloud vendors, integrating capabilities such as cloud budget management or spend alerting capabilities as well as research-specific cloud training tutorials for academic researchers.

The CloudBank EU NGI experiment aims to expand this vision to Europe beyond CERN, reaching out to Research organisations, offering and enabling a framework for efficient management of public cloud contracts from diverse funding sources, leveraging and improving the process and technology built for CloudBank Research. The US partner is actively involved in the CloudBank EU NGI experiment. The team of UC San Diego organises a series of webinars for Higher Education CIOs and University IT Managers illustrating how CloudBank can save university administrators time and money, discussing cloud management and optimization strategies, taking the guesswork and risk out of using the Cloud for researchers who need the power and flexibility of cloud services while maintaining control of their computing environments.

In addition, the technical team of UCSD is coordinating the effort to deploy the CloudBank portal in Europe, specifically the technical requirements, Drupal[2] core modules, open-source licensing, CILogon[3] federated identity management support and overall operational effort estimation to run a CERN based CloudBank instance.

João Fernandes (male), is a senior member of staff of CERN IT Department and has more than fifteen years of experience in distributed computing infrastructure, fostering several collaborations with academia and industry. João is currently the director of H2020 EOSC ARCHIVER[4] project, an R&D initiative to establish hybrid cloud services for analysis, storage and long-term preservation of scientific data. João holds a master’s degree in Electrical and Computer Engineering at Porto University, School of Engineering. Before joining CERN, he was a systems engineer at the California Institute of Technology (Caltech) USA, with



responsibilities in distributed computing providing a hub for massive scale real-time collaboration within the global Large Hadron Collider (LHC) research community.

João is the project leader, responsible for the project technical oversight and interaction with the UCSD.

Gerhard Deiters (male), Attorney at Law Partner BHO Legal, specializes in IT, data protection and international contract law as well as the security and defence sector, especially in the areas of satellite communications and unmanned systems. His clients include international corporations, medium-sized companies, German and European authorities, and research institutes. In addition to providing legal advice, he is distinguished by numerous activities in the areas of IT and data protection law, including lectures and training courses on legal issues of future technologies. Gerhard studied law at the University of Cologne. After his first state examination, he worked as a research assistant at the Institute for Private International Law and Comparative Law and later at the law firm Oppenhoff & Partner in Cologne, where he also completed his legal clerkship. Prior to joining BHO Legal, he worked for several years as an attorney at the supra-regional law firm Heuking Kühn Lüer Wojtek in IT, telecommunications and data protection law. Gerhard is currently the data protection officer of KLEO Connect GmbH and GrapeCheck GmbH as well as the Chief Privacy Officer of the FUCHS Group.

Gerhard will be responsible for the legal and contractual analysis of the CloudBank model with respect to European legislation (such as GDPR) that includes concrete comparison of options for possible contractual relationships.

Joshua Davison (male) has eight years of experience in strategic procurement, specialising in the procurement of research and development services and software. At CERN, he manages procurement activities for the IT department, the SCOAP3[5] open access publishing initiative, and several pre-commercial procurements (PCP) including the ARCHIVER project. Joshua has a degree in Economics from University College London and previously worked for a multinational power systems company in London.

Josh will assess the contractual analysis produced by BHO Legal in view of structuring a multicloud procurement model for research-performing organisations such as CERN.

Dr. Robert Jones (male) is a senior member of the CERN IT Department management and was the coordinator for HNSciCloud[6] project. His experience in the distributed computing arena includes mandates as the technical director and then project director of the EGEE projects (2004-2010) which led to the creation of EGI[7]. He was a member of the Sustainability Working Group for the European Open Science Cloud and was appointed in December 2020 a director of the recently created European Open Science Cloud (EOSC) Association[8]. Dr. Jones will be responsible for the interaction with the EOSC governance and CERN management.



Dr. Sofia Vallecorsa (female) is a CERN physicist with extensive experience on software development in the High Energy Physics domain, in particular on Machine Learning and Deep Learning applications. Sofia obtained her PhD in High Energy Physics at the University of Geneva working at the CDF experiment at Fermilab, USA. A member of CERN openlab, Sofia works on several projects related to Deep Learning and Quantum Computing for High Energy Physics, Healthcare and Earth Observation applications, supervising the work of several doctoral students. As technical coordinator of the Quantum Computing activities in the context of the CERN Quantum Technology Initiative[9], Dr. Vallecorsa oversees the deployment of scientific workloads in the domains of Quantum Computing, Machine Learning and Quantum Machine Learning, in the context of the CloudBank EU NGI experiment.

The following personnel contributions are being exclusively funded by the EC partner (CERN):

Anna Manou (female) is a project management officer at CERN IT External Projects unit. Anna is involved in the management and administration of multiple EC funded projects in the ICT domain. She has a background in Finance and Information Systems. Anna will assist on all the project management and interactions with NGI Atlantic.eu.

Apostolos Theodoridis (male) is a data analytics specialist at CERN. His field of expertise lies in the intersection between data science, computer science and statistics. Apostolos obtained his master's degree in Business Analytics in University of Geneva in Switzerland and his bachelor's degree in Computer Science at Aristotle University of Thessaloniki in Greece. He has significant experience in data processing and visualization, as well as in descriptive/prescriptive/predictive analytics and data-driven methodologies to leverage data analysis and modeling techniques in order to glean insights across functional domains, such as financial management. In the CloudBank EU NGI, Apostolos is co-designing and developing an agnostic open-source billing and cost tracking solution to enhance the CloudBank portal capabilities, extracting data from different billing data sources and create metrics to be visualized in live dashboards to predict cost optimizations and the financial evolution of the contracts managed by CloudBank.

Lucie Pocha (female) member of the CERN Legal Service, provides legal support in respect of CERN's participation in EU projects and other grant schemes as well as scientific and technological collaborations, procurement and financial transactions. Prior to joining CERN, Lucie worked in the London offices Maclay Murray & Spens LLP (now Dentons) and Aberdeen Asset Management PLC, focusing on commercial law and the financial services industry. Lucie is responsible for the legal activity for the NGI contract, CloudBank EU NGI consortium agreement and a recently established CERN/UCSD MoU for collaboration and co-development of CloudBank.



The current instance of CloudBank operating in the US has received US\$ 5M funding from NSF. The effort of UCSD personnel is funded via the existing NSF award for CloudBank as follows:

Dr. Vince Kellen (male) is the Chief Information Officer (CIO) for the University of California, San Diego (UCSD), as well as a member of the Chancellor's Cabinet, and vice chancellor and chief financial officer's senior management team. Dr. Kellen is a recipient of the coveted CIO Magazine "Top 100" award in 2007 and 2014, was one of InformationWeek's "Chiefs of the Year" in 2013, received the Computerworld "Honors Laureate" award in 2013, and was one of four recipients selected globally for Dell's 2012 "Transformational CIO" award. Dr. Kellen served as a consultant in industry advising fortune 500 clients on IT strategy, customer and marketing strategies, analytics and data warehousing. He earned his Ph.D. in computer science, an M.S. degree in information systems, and a B.A. degree in communications from DePaul University in Chicago, Illinois.

Dr. Kellen holds activities such as public webinars on CloudBank for research sector organisations seeking better ways to procure cloud, provide in-depth understanding on how CloudBank Enterprise can make enable fund savings and unlock interesting financial options for research organisations using public cloud.

Shava Smallen (female) is the co-chair of the PRAGMA Steering Committee, funded by NSF, for the University of California, San Diego. Shava is also a major contributor to the Extreme Science and Engineering Discovery Environment (XSEDE), an NSF-funded virtual organization that integrates and coordinates the sharing of advanced digital services - including supercomputers and high-end visualization and data analysis resources - with researchers nationally to support science.

Shava will oversee the technical operations, specifically for the integration of the LHC use cases to be deployed under the NGI experiment, in the current instance of CloudBank.

4 Results

CloudBank EU NGI has successfully achieved the following results:

- Piloted usage of commercial clouds by multiple research teams and organizational units.
- Provided structured, central access to multiple cloud providers with billing and data processing terms understood.
- Determined a cloud contracting model compatible with European's procurement process, protecting European research organisations financial interests.
- Promoted the cost-effective use of cloud services with risks managed and increased impact to European research use cases.



- Verified the billing control capabilities of a financial cloud broker, using it as an additional source of data for cost verification and prediction.
- Assessed the quality of user support of the piloted cloud vendors.
- Benefited from training tutorials provided directly by cloud architects and support lines provided by cloud providers.
- Raised awareness about the benefits and potential of commercial cloud services for research, providing working examples of research workload deployments.
- Determined if the model can be expanded beyond a single organisation (CERN) and scaled in Europe to other research organizations either by enabling new cloud contracts or allowing organisations to *“bring their own cloud contract”*.

Specifically, during the activity from D2 to D3, in addition to the data privacy provisions analysis, a preliminary security assessment of CloudBank has been pursued via two complementary threads.

The first was an analysis of the security plan supported by SDSC in the context of the current US-based instance:

1. the **CloudBank user portal**, which comprises the software, operations, and hardware that hosts the user portal, as well as the project staff who program and support the portal and user community. This includes both on-premise systems within SDSC’s Data Center on the UC San Diego campus, as well as external services such as the federated AAI proxy service based on CILogon, which are part of the CloudBank infrastructure;
2. the **public cloud** resources that researchers use for their CloudBank-granted allocations, and the public cloud staff who collaborate with CloudBank staff and users;
3. the community of **CloudBank users**, their programming practices, their processes and tools for using both the CloudBank portal and cloud vendor resources.

This security plan and supported technical measures have been considered in the data privacy assessment performed by BHO-legal. The SDSC full security plan for CloudBank is presented in the Annex II of this document.

In addition, to address points 2) and 3) above in a European setting, the CloudBank EU pilot project following suggestions and guidance from the CERN Cybersecurity Team, has initiated a cloud security exchange in the scope of the EC funded Medina project[10] with one of the project partners, Bosch. Medina’s main objective is to create a holistic security framework for achieving a continuous audit evidence-based certification in compliance with the EU Cybersecurity Act.

Specifically with Bosch, the main objective of the exchange for CERN is to collect relevant information about Bosch’s journey to the cloud, i.e., the *“Lessons Learned”* going beyond the standard operational security practices.



Aspects covered include security controls & security governance, secure migration to cloud, shared responsibility between organizations as cloud customers vs. the cloud vendors and experience with Managed Security Service Providers (MSSPs) to support cloud users from novice to expert levels.

Bosch maintains about 300 SaaS services in its service portfolio and supports 600 AWS accounts, with no restrictions in cloud use.

Some of the highlights of the exchange that are beneficial to CloudBank EU include:

- Establishment of centralized purchase criteria and central security policies for cloud services.
- Definition of a 3-layer security defence model: 1) security governance framework (i.e., internal cybersecurity regulation, ISO controls based), 2) concrete secure best practice reference implementation as a footprint for end users and 3) automation via machine-readable policies.
- Description of the cloud service onboarding process.
- Definition of security benchmark baseline templates for cloud deployments, strongly inspired on cloud vendors accumulated experience (e.g., Azure Security Benchmark)
- Role of external experts (MSSPs) in providing support to a heterogeneous set of users (from novice to expert).
- Establishment of a cloud “Center of Excellence” with representatives of internal stakeholders across the organisations consuming public cloud.

4.1 Discussion and Analysis on Results

An analysis of the results of the CloudBank EU experiment has been collected in the tables below addressing procurement, cloud consumption control and optimization, data privacy and security. Based on the results of the experiment each topic includes possible ways forward based on that analysis, in case of future European expansion.

Procurement	Experiment Results Analysis	Proposed way forward in case of European expansion
Pricing	Prices depend on multiple commercial and technical factors that might require the development of in-house capabilities to manage them.	As a first step, a financial broker can partially help the capability building. A central contract manager role is needed for the function.
Cloud Contracts Adjudication	A contract model based on a broker has the advantage of allowing it to propose commitments and optimize the cloud spend of researchers whenever possible. Alternatively, there are advantages of a direct relationship with cloud providers that are also relevant such as to influence the	A hybrid contractual model should be explored to increase benefit, for example, maximise discounts by creating competition between cloud providers and their resellers.



	list price offered to research organisations, access to credits as a “buffer” for consumption and additional negotiation possibilities.	Adjudication can be based on forecasted consumption for 3+ years across all user groups.
Country of Origin	Some European organizations have procurement rules that prioritize suppliers from European states to balance industrial return and this has implications on where public cloud services are located. To note that cloud infrastructure can be in Europe, even if a reseller is not, or even if the original cloud platform provider has its headquarters in the US.	CloudBank can have a critical role in this aspect, by allowing possibilities to implement contractual location requirements at the CloudBank portal level via the cloud vendor supported APIs for region management.

Consumption Control & Optimisation	Experiment Results Analysis	Proposed way forward in case of European expansion
Dynamic Consumption	<p>Spikey consumption and variations on initial consumption forecasts for research workloads can be accommodated by the use of a framework contract with no pre-defined commitments.</p> <p>Invoices that include a breakdown by deployed use-case, allow easy allocation of costs to different teams (and budget codes) by the invoicing teams.</p>	<p>In order to permit rapid increases of capacity and resource types, CloudBank, acting as a central cloud vendor management framework, allows the research leads to choose one or multiple clouds without the need to re-tender (and no additional delays).</p> <p>A hybrid consumption model based on usage profiling that can combine pre-defined commitments (in the case of more traditional use-cases), with “pay as you go” capacity and even offered credits from cloud providers is potentially the most cost-effective.</p>
Billing & Invoicing	<p>Live billing monitoring can avoid surprises at the end of the billing cycle. The combination of multiple billing information sources (e.g. broker, cloud provider, cost calculations per use case, invoices, etc.) allow organisations to detect billing anomalies (e.g. small anomalies were detected for example at the end of the first phase of CloudBank, when comparing monitored data with invoicing).</p>	<p>A sustainable model would imply that monthly invoices, broken down by deployed research projects, are then allocated to each nominated budget/grant holder, after central validation.</p>
Consumption Control	<p>When using cloud services, the risk of overspending exists. In order to protect financial interests, a set of capabilities needs to be in place.</p>	<p>The creation of a programmatic “Big Red Stop Button” to be able to stop usage in an emergency remains important, exploring the programmatic budget management capabilities via</p>

	<p>CloudBank’s ability to provide live spend monitoring and dashboard capabilities, possibilities to set triggers/warnings when spend reaches pre-defined thresholds, etc. are good examples of those added value capabilities.</p>	<p>APIs of some cloud providers, notably the hyperscalers.</p> <p>These capabilities allow controlled consumption by combining technical aspects (e.g. cloud administrators setting up technical programmatic cut offs) with terms in contractual offers.</p> <p>Contractual possibilities can be explored when they exist (e.g. GCP offers an “all you can eat” fixed price model, with overage risk on GCP’s side).</p> <p>The current strategy of cloud vendors offering “free” credits can be leveraged to “buffer” potential overspend and help to mitigate risks and allow applications to continue to run. The additional benefit is to have “free” and paid usage under the same negotiated terms, instead of scattered with no agreed terms, exposed to data processing risks.</p>
--	---	---

Data Privacy & Security	Experiment Results Analysis	Proposed way forward in case of European expansion
<p>Data Privacy</p>	<p>The analysis by BHO (Annex 1 of this deliverable) indicates the CloudBank model offers no significant legal concerns with regard to contractual arrangements and data protection law. Therefore, it’s concluded that the exposed hypothesis in D1 and D2 of the CloudBank experiment has been achieved.</p>	<p>In order to consolidate the findings, as a resulting step, it would be beneficial to have expert consultants in IT law such as BHO-Legal to analyze the contractual terms with cloud providers in any future expansion of CloudBank to Europe. As seen in the OCRE H2020 project[11], VAT handling across Europe is an issue that needs to be resolved in order to enable cloud adoption. As such, a Pan-European VAT analysis must be part of the consolidation.</p> <p>The applicability of GDPR and the possibility to use Standard Contractual Clauses (SCCs) for data transfers needs a more in-depth analysis.</p>



Security	<p>Together with the establishment of centralized purchasing criteria it's important to define a central security policy for cloud services, following a uniform footprint. It's important also to define a security benchmark baseline template for cloud deployments that can be inspired by the cloud vendors accumulated experience, resulting in a reference security deployment for research leads.</p> <p>Role of external experts such as managed security service providers (MSSPs) can be important in providing support to harmonize the heterogeneous security awareness of research leads, from novice to expert.</p>	<p>Onboard the relevant lessons learned in the context of the discussions with Bosch and the Medina project and establish a security model with a clear set of delineated responsibilities between in-house security teams of public funded research organisations and cloud vendors.</p>
-----------------	--	---

Training & Tutorials	Experiment Results Analysis	Proposed way forward in case of European expansion
User Skills	<p>The skills of research leads deploying use cases vary from novice to expert users. The tutorials and training sessions organized during the Experiment were extremely popular and oversubscribed with very positive feedback received from the researchers.</p>	<p>Build a model to allow systematic training offers to be available for European researchers and IT personnel, as a means to increasingly build the set of necessary skills.</p>

Based on the CloudBank EU results a service operation model can be expanded to European research that includes the aspects described in the table below.

Proposed Operational Model for CloudBank EU	
Multi Cloud Access	<p>Researcher Access to the CloudBank portal via Federated AAI (Edugain[12]) to multiple cloud providers, able to compare services and use the best possible service for its research.</p>
Billing	<p>Sovereign vendor neutral billing/tracking capabilities to multiple cloud providers allowing cost estimate generation, and combination of billing sources (Fig 1.). Monitoring of cloud providers regions for data sovereignty reasons. Test suite used to validate and benchmark cloud vendors and provide working examples of deployments to researchers for the supported services.</p>



<p>Support Roles</p>	<p>Clear breakdown of responsibilities between in-house teams and cloud vendors. In cases where research use cases do not need further integration or where customisation is minimum, cloud providers provide direct support lines covering multiple time zones, and are contractually responsible for ensuring services remain manageable, secure and the latest functionalities are available.</p>
<p>Skill Building</p>	<p>Skill building and training provision via cloud vendor architects that can potentially be structured in a central catalogue.</p>
<p>Invoicing & Volumes Estimation</p>	<p>Contracts must allow for dynamic consumption to accommodate peaks in demand and support switching between cloud providers.</p> <p>Establish as possible, a “hybrid” contractual model (including a direct contract with Cloud provider and indirect via a reseller) to stimulate competition between cloud vendors and resellers, maximizing the discount benefits.</p> <p>Optionally, purchase financial brokerage services/consultancy in order to develop this competence on behalf of the research community adopting commercial cloud.</p>
<p>Data Privacy & Security</p>	<p>Legal feasibility of the model was confirmed for CERN and for Europe.</p> <p>Establish a central security policy for cloud services, following a uniform footprint.</p> <p>Define a security benchmark baseline template for cloud deployments that can be inspired by the material provided by the pilot phase cloud vendors.</p> <p>Create a clear set of delineated responsibilities between Cloud vendors security teams and in-house public sector security teams.</p>



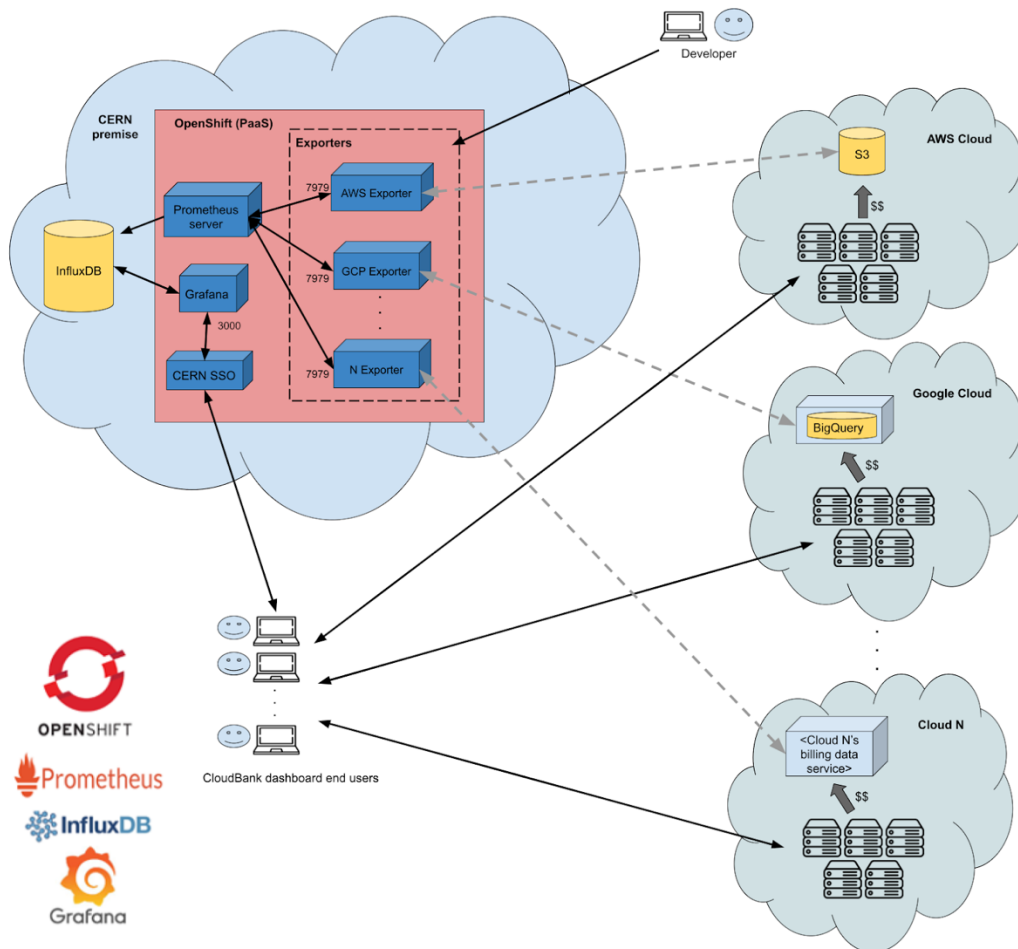


Figure 1: Vendor Neutral Cost Management PoC running at CERN.

5 Present and Foreseen TRL

The results of the CloudBank EU NGI experiment may not change the Technical Readiness Level (TRL) of the CloudBank instance deployed by UCSD but expanded the applicability of the model to a European setting.

6 Exploitation, Dissemination and Communication Status

One of the biggest challenges in the research public sector is to acquire the correct skills to use state-of-the-art public cloud resources with the best performance and cost-effectiveness for scientific workloads.

In this regard, CloudBank EU promoted a skill building cycle that engages experts (e.g., specialized cloud architects from vendors) directly with IT-savvy researchers and IT department personnel.

The acquisition of the necessary skills is a critical element for the sustainability of the CloudBank model, as skill training leads to better resource choice and consumption, allowing for improved forecasting of service needs and consequently optimization of procurement. The



resulting financial savings can then be invested in building the skills of researchers and IT personnel. This skill building cycle is depicted in Figure 2.



Figure 2: Virtuous (Sustainable) skill building cycle.

During the first stage of CloudBank EU, the following tutorials were organized:

- QCaaS tutorial organized with AWS Braket[13]
 - 60 Participants; 20 hands-on participants (using CERN SSO to access)
- ML hands-on workshop[14]
 - 133 participants
- CMS Open Data using public cloud[15]
 - 50 participants (mostly not based at CERN nor with access to CERN infrastructure)

In addition, a scheduled regular exchange between expert consultants and users has been proposed to be set up with AWS/GCP cloud architects, twice per month in virtual rooms to answer questions or help to debug problems of research leads and DevOps teams.

The researchers taking part on the CloudBank EU experiment participated in the tutorials mentioned in the previous section and also the first CloudBank EU workshop.

The workshop provided an overview of CloudBank and its preliminary results but also gave the opportunity to the research leads from the LHC Experiments, IT, TH Departments and the ATS sector, the opportunity to describe their research workload deployments in AWS and GCP cloud environments. The workshop included a panel discussion with representatives from all sectors to collect input on how the CloudBank EU pilot could evolve.



The research leads and their teams responded to surveys about their experience in using the cloud services available via the CloudBank EU pilot. The highlights of the feedback received from researchers can be summarized in the following points:

- CloudBank provides access to industry-leading technology solutions that have a useful impact on High Energy Physics research workflows.
- CloudBank provides access to h/w architectures and s/w packages that are not available on-premises or only available on-premises at limited scale (e.g. GPUs).
- Public cloud services provide fully configured environments (e.g. python environments, etc.) that can simplify researchers' tasks.
- Skill training (and its potential extension to cloud certification schemes) is critical to ensure that public clouds are used efficiently.
- CloudBank allows an efficient application of the concept of heterogeneous computing (i.e. optimized in terms of resources usage, energy usage and timescale) in Physics analysis.
- CloudBank provides access to mature offerings from cloud service providers (e.g. AWS has offered commodity Machine Learning services for more than a decade) and is able to test and determine their relevance on a short timescale.

In addition, the activities in CloudBank EU both from the project itself as well as the participating research leads have been disseminated via peer reviewed papers or communication actions, such as:

- Peer reviewed paper “CloudBank for Europe” in vCHEP2021 proceedings[16].
- Peer reviewed paper “Accelerating GAN training using highly parallel hardware on public cloud” from the Research Leads deploying in CloudBank, in vCHEP2021 proceedings[17].
- ISC’21 poster “Bursting DLaaS in the Cloud”[18].
- Invited talk and panel discussion on the EOSC Symposium, session “Sustaining the EOSC” on June 17th 2021[19].
- “Digital Around the World” virtual event in October 20th (US Timezone)[20].
- US partner press release “A match made in the Cloud”. [21]
- US CloudBank workshop with a presentation about ML use cases for High Energy Physics simulation.[22]

7 Impacts

Impact 1: Enhanced EU – US cooperation in Next Generation Internet, including policy cooperation.



The CloudBank EU NGI experiment is fostering transatlantic cooperation and existing US-EU cooperation on networking and AAI where policy aspects are key for global transatlantic research. Examples of these impacts are the Memorandum of Understanding (MoU) signed between UCSD and CERN on the 23rd of March 2021 to establish a collaboration for cooperative development of UCSD's existing Cloudbank software with the objective of developing future versions of CloudBank that are general-purpose, fully configurable, and open source to be used at both sites and potentially other sites.

Another example is the current already working integration of CloudBank with the CERN IdP service for Authentication and Authorization and foreseen integrations with EduGAIN Federated AAI services.

Impact 2: Reinforced collaboration and increased synergies between the Next Generation Internet and the Tomorrow's Internet programmes.

Thanks to the CloudBank EU experiment, the CloudBank model can now be proposed as an experimental platform in the NGI portfolio and made available to US and European users. This experiment represents a milestone with wide impact to launch the expansion of CloudBank usage across Europe's research community by engaging other research performing organisations. An example of this impact is the current preparation at CERN for the introduction of European cloud service providers into the CloudBank model where agreements are being discussed with European cloud providers that have expressed interest in integrating CloudBank EU. The objective is now to invite other research performing organizations to also participate as procurers (a few already expressed interest in probing CloudBank to manage their recently signed cloud contracts).

Impact 3: Developing interoperable solutions and joint demonstrators, contributions to standards.

The experiment will contribute to the establishment of a safe, sovereign and federated European commercial cloud infrastructure for research users, based on interoperability, open source software and open standards as measures for vendor lock-in prevention, in order to identify synergies with recent initiatives such as Gaia-X[23]. This alignment will stimulate closer technological cooperation, in a "European way" as a part of a global research activity, between the public research sector and the private sector across Europe. A good example of this impact is the plan to integrate the EOSC test suite[24] developed by CERN with CloudBank, to validate research requirements as wide as possible and provide researchers with working examples of deployments on different cloud environments.

Impact 4: An EU - US ecosystem of top researchers, hi-tech start-ups / SMEs and Internet-related communities collaborating on the evolution of the Internet



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 HNSciCloud Pre-Commercial Procurement action[25] has shown the appetite for researchers and public cloud service providers to work together, and this experiment established a platform where they can collaborate to develop innovative solutions.

8 Conclusion and Future Work

CloudBank promotes a model for cost-effective use of cloud services where risks and impacts to research organizations are mitigated and managed. The results of the CloudBankEU experiment with CERN and UCSD during 2021 have made the case for cloud services procurement and usage for the benefit of different sectors of European research stimulating a shift of approach, where both organisations and funding bodies can contribute their own funds, in a joint-funding model. Before CloudBank, individual researchers often provided their credit cards as a guarantee to get “free credits” from cloud providers with fragmented usage and terms of use that had not been reviewed for consistency with agreed policies and best practices. CloudBank does not intend to replace on-site resource provisioning. Instead, it intends to complement on-site resources with capabilities and capacities that are currently not available. It will allow research organizations to centrally cope with rapid cycles of technology evolution (major technology breakthroughs usually appear on the market every 4-5 years), centrally manage multiple cloud vendors whilst providing the ability to validate and differentiate capabilities of each cloud service provider, understanding transparently its strengths and weaknesses. Researchers can benefit from this alignment as a central cloud vendor management framework to increase optimisation of cloud service usage to their research needs, being provided with working examples of deployments and expert support from cloud architects.

The model is now ready to be expanded for multiple scientific domains and procuring organizations in the European research sector. Should the “*multiple buyer / multiple cloud*” approach be successful, it may be a candidate to be proposed as a procurement model in the context of the European Open Science Cloud.

In addition, these developments can result in laying the foundation to establish a mechanism to extend CloudBank to European cloud providers, aligning it with the rise of European digital sovereignty and strengthening of data self-determination for publicly funded European research.

Ultimately, CloudBank leads to a simplification of public cloud service procurement processes for the public research sector, integrating these into a hybrid computing model and facilitating a smooth transition to a modern and heterogeneous cloud infrastructure.

9 References (optional)

1. <https://eosc.eu>, accessed: 2022-01-22



2. <https://www.drupal.org>, accessed: 2022-01-22
3. <https://www.cilogon.org/>, accessed: 2022-01-22
4. <https://www.archiver-project.eu>, accessed: 2022-01-22
5. <https://scoap3.org>, accessed: 2022-01-22
6. <https://www.hnscicloud.eu>, accessed: 2022-01-22
7. <https://www.egi.eu>, accessed: 2022-01-22
8. <https://eosc-portal.eu>, accessed: 2022-01-22
9. <https://quantum.cern/welcome>, accessed: 2022-01-22
10. <https://ocre-project.eu>, accessed: 2022-01-22
11. <https://medina-project.eu/>, accessed: 2022-01-22
12. <https://edugain.org>, accessed: 2022-01-22
13. <https://indi.to/aws-braket>, accessed: 2022-01-22
14. <https://indico.cern.ch/event/1044836/>, accessed: 2022-01-22
15. <https://indico.cern.ch/event/1031398/>, accessed: 2022-01-22
16. <https://doi.org/10.1051/epjconf/202125102025>, accessed: 2022-01-22
17. <https://doi.org/10.1051/epjconf/202125102073>, accessed: 2022-01-22
18. <https://app.swapcard.com/event/isc-high-performance-2021-digital/planning/UGxhbm5pbmdfNTY0Mjcz>, accessed: 2022-01-22
19. <https://www.eoscsecretariat.eu/eosc-symposium-2021-programme>, accessed: 2022-01-22
20. <https://digitalaroundtheworld.org/>, accessed: 2022-01-22
21. <https://www.newswise.com/articles/a-match-made-in-the-cloud>, accessed: 2022-01-22
22. https://www.youtube.com/watch?v=2O0QfC_DF5Y&list=PLFTt6begQkmmwom-rrV8oo1nWyEkGNaaee&index=16, accessed: 2022-01-22
23. <https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html>, accessed: 2022-01-22
24. <https://github.com/cern-it-efp/EOSC-Testsuite>, accessed: 2022-01-22
25. <https://www.hnscicloud.eu>, accessed: 2022-01-22

10 Glossary (preferably in tabular form like below)

AAI	Authentication and Authorization Infrastructure
API	Application Programming Interface
ATS	Accelerators and Technology Sector
CIO	Chief Information Officer
EOSC	European Open Science Cloud
GDPR	General Data Protection Regulation
HPCaaS	High-Performance Computing as a Service
LHC	Large Hadron Collider
MoU	Memorandum of Understanding
MSSPs	Managed security service providers
NGI	Next Generation Internet
NSF	National Science Foundation



PCP	Pre-commercial procurements
PM	Person Months
R&D	Research and Development
SCCs	Standard Contractual Clauses
SDN	Software Defined Networks
SSO	Single sign-on
TH	Theoretical Physics
TRL	Technical Readiness Level
UCSD	University California San Diego
VCHEP	International Conference on Computing in High-Energy and Nuclear Physics
VNF	Virtual Network Function
WIT	Waterford Institute of Technology (Coordinating Partner)
XSEDE	Extreme Science and Engineering Discovery Environment

