CERN Computing in Commercial Clouds

C Cordeiro¹, L Field¹, B Garrido Bear¹, D Giordano¹, B Jones¹, O Keeble¹, A Manzi¹, E Martelli¹, G McCance¹, D Moreno-García¹, S Traylen¹

¹ CERN IT Department
E-mail: domenico.giordano@cern.ch, cristovao.cordeiro@cern.ch

Abstract.

By the end of 2016 more than 10 Million core-hours of computing resources have been delivered by several commercial cloud providers to the four LHC experiments to run their production workloads, from simulation to full chain processing. In this paper we describe the experience gained at CERN in procuring and exploiting commercial cloud resources for the computing needs of the LHC experiments. The mechanisms used for provisioning, monitoring, accounting, alarming and benchmarking will be discussed, as well as the involvement of the LHC collaborations in terms of managing the workflows of the experiments within a multi-cloud environment.

1. Introduction

With the planned upgrades of the Large Hadron Collider (LHC) and consequent increase in the amount and complexity of data collected by its experiments, CERN’s computing infrastructures will be facing a large and challenging demand of computing resources. Within this scope, the adoption of OpenStack cloud computing at CERN has opened the door to the evaluation of commercial cloud services, which could supply additional Infrastructure-as-a-Service (IaaS) resources to extend the current CERN computing resources for physics data processing.

In addition to demonstrating that WLCG workloads can run on public clouds, the purpose of these activities was also to probe the effectiveness of the CERN procurement rules for the cloud market, to refine the technical requirements to meet the different needs of the WLCG computing workloads, to gradually involve in the evaluation all four LHC collaborations (ALICE, ATLAS, CMS and LHCb) and last but not least to increase the complexity of the deployments by testing multiple data centres at the same time, provisioning a larger number of VMs and larger VM instances, including storage on the cloud as well as a better WAN connectivity among the cloud data centre and CERN.

2. Cloud Procurement Roadmap

The evaluation of commercial clouds started in 2011 within the context of the EU-funded Helix Nebula [1] initiative, a public-private partnership between Europe’s leading scientific research organisations and European IT cloud providers aiming to establish a cloud computing platform for data intensive science within Europe.

Building on the experience gained with Helix Nebula, three cloud procurement actions were established during 2015 and 2016 in order to deepen the understanding of commercial cloud
adoption by means of deployments of increasing size and complexity. Following the CERN procurement rules, the invitations to tender were addressed to the cloud service providers based in at least one of the CERN member states. These procurement actions generated considerable interest on the private sector, with several high-quality applications received. The adjudications were based on the rule of the lowest compliant offer. The three procurement actions were assigned to Atos, Deutsche Börse Cloud Exchange (DBCE) and T-Systems respectively. In addition, two different grant agreements gave the possibility of evaluating the IaaS offers of Microsoft Azure and IBM SoftLayer. Figure 1 summarises the procurement roadmap of the last two years.

The first cloud activity with Atos [2] was characterised by the provisioning of about 3000 single-core VMs processing simulation jobs of the ATLAS collaboration for 6 weeks. This initial demonstrator provided the ability to integrate cloud IaaS within the WLCG workflows and to run CPU intensive tasks accessing data and software libraries from remote storages over the WAN, using technologies such as CVMFS [3] and XRootD [4]. The Microsoft Azure evaluation involved simultaneous deployments in three different data centres (two in EU and one in US) and three collaborations (ATLAS, CMS and LHCb) submitting simulation jobs. A technical report about this activity is available [5]. In the second cloud procurement DBCE was the selected company to provide a total of 1 000 4-core VMs, running simulation jobs for all the LHC collaborations. Unlike the other commercial cloud providers, DBCE played the role of a broker connecting customers with the actual providers of IaaS in the DBCE marketplace where compute capacity was traded like any other commodity in an exchange. At that time DBCE, as lead procurer, settled a contract agreement among CERN and 5 different cloud providers: ClouData, Cloud&Heat, DARZ, Innovo and Ultimum.

![Figure 1: Summary of the cloud computing evaluation roadmap at CERN since 2015.](image)

The WLCG computing activity consists of distinct types of workloads classified as simulation, reconstruction and analysis. The first is solely CPU bound with low network and storage requirements while the latter two require CPU, storage capacity and network bandwidth. Having proved the sustainability of simulation workloads in commercial clouds, the investigations moved to reconstruction and analysis workloads, the so called “data-intensive” workloads. This was the objective of the third cloud procurement, where in addition to compute capacity and network, for the first time a close (RTT <5 ms) central storage of 500 TB was procured.
3. Seamless Access to External Resources

The experience acquired during the described cloud activities in the areas of provisioning, configuring and monitoring made possible the consolidation of the strategies toward a transparent extension of the CERN computing resources into commercial clouds, where the cloud resources are configured and managed using the same tools adopted to manage the CERN private cloud, namely Puppet [6] for service configuration, HTCondor [7] as batch system and the CERN-IT Unified Monitoring Architecture (UMA) [8] for the resource monitoring.

Figure 2b shows the conceptual stack of services involved in the delivery of compute capacity to the end users. In WLCG the preferred approach for the submission of workloads to the compute infrastructure is the use of the experiments’ Workload Management Systems (WMS), represented by the Pilot factories in the figure. Pilot factories interact with a batch system, via a grid Compute Element, in order to access the on-premises compute resources delivered at CERN by OpenStack IaaS, as well as the resources available through public cloud IaaS and volunteer computing. This approach was implemented during the activities with IBM and T-Systems with the deployment of a common HTCondor instance which handles the scheduling and match-making of the workloads submitted by each experiment. More details are reported in [9].

The adoption of a common HTCondor instance has implications on the VM lifetime management. The VMs do not need to be pre-configured to interact with a specific experiment WMS, as was the case in the early procurements, with separate job submission workflows as shown in Figure 2a. VMs can thus have a long lifetime and keep running for as long as they behave properly and are still needed. On the contrary in the past models the VM lifecycle management implied continuous cycles of creation and destruction to apply experiment specific contextualisation. This approach caused dead time due to the frequent bootstrap of re-cycled VMs and resulted in an increased load on the cloud provisioning system, due to the large number of calls to the provisioning APIs.

The approach finally adopted improves the utilisation of resources, with the HTCondor service in charge of the quota sharing among users. Figure 3 shows the usage of compute resources by the four LHC collaborations during activity with T-Systems. All cores available for compute (about 3.500) were fully exploited over the delivery period thanks to the dynamic allocation of the job slots to the collaborations active at any given moment.

For each one of the five cloud activities, five different APIs had to be adopted – SlipStream [10] for Atos, Azure’s ASM [11] and ARM [12], DBCE, IBM SoftLayer, OpenStack for T-Systems – and even some custom tooling had to be developed in order to manage large deployments. The implementation of connectors between the provisioning tools and those APIs often represents
an initial barrier and delays cloud adoption at large scale. For this reason the new strategy is to adopt open-source tools for the VM lifecycle management supported by large communities.

During the early cloud activities in 2015, Ganglia was the chosen monitoring framework as it provided a simple setup and the possibility of rendering clustered views of the resources, grouping metrics per collaboration or data centre [13]. In 2016, within the consolidation effort to achieve the transparent extension of resources, the CERN-IT’s Unified Monitoring Architecture [8] was adopted also to monitor the cloud resources.

In all the cloud activities the established monitoring of the VMs has also served as a reliable source of data for consumer-side accounting, used to validate the providers’ accounting reports and invoices. The same monitoring data has also been re-purposed to feed an in-memory data analytics application for the automatic detection of unhealthy and erratic resources [14].

Profiling of compute resources was established in all cloud activities to systematically measure delivered performance during production activities. The procedures introduced during the cloud activities for the resource profiling were also included in the cloud procurement process to establish the SLA conditions, to select the submitted bids and to define the compensation actions in the case of delivered performance below that defined by contract. More details about resource profiling are reported in [15].

4. Networking
One of the most important components for enabling “data intensive” workloads is the network connectivity between the cloud data centre and the CERN data centre. The connectivity must be reliable, have low latency and large bandwidth in order to support the efficient transfer of data to be processed in the cloud. In all the cloud activities the connectivity was guaranteed by the network transit across GÉANT through direct peering at an Internet Exchange Point (IXP) or via a National Research and Education Network (NREN). The largest network bandwidth tested during the described cloud activities was established during the third cloud procurement. In that case a dedicated 10 Gbps connection to GÉANT through DFN, the German NREN, was established by T-Systems. For the occasion, GÉANT set up a virtual routing instance, called GÉANT Cloud VRF, dedicated to connectivity between commercial cloud providers and Research and Education customers.

The connection through the GÉANT Cloud VRF was tested during the commissioning phase and monitored during the production activity. It proved to be able to deliver 10 Gbps of sustained traffic (Fig. 4), while other connections tested during the same activity did not prove analogous performance. For instance the connection over an existing transit peering between CERN and Deutsche Telekom could deliver only around 4.5 Gbps. Connectivity from the WLCG
Figure 4: Network traffic between T-Systems and CERN, over the GÉANT Cloud VRF. The saturation of the link was determined by the cloud VMs accessing remotely across WAN files stored at CERN. The XRootD protocol was mainly used for the file access.

Tier-1 sites to the T-Systems resources was also checked and it resulted to be via commercial Internet provider networks for almost all of them. While such connectivity is a valid option, some sites may have limited connectivity to the commercial Internet, thus connectivity to the GÉANT cloud VRF is a recommended option for at least all the European sites.

Typically VMs in the cloud need to be reachable via public IP addresses, directly assigned to the VMs or exposed through a 1:1 NAT. This was for instance the configuration adopted by T-Systems, and regardless of the NAT, the aggregated performance allowed the saturation of the 10 Gbps connection for long periods. VPNs were also tested as a possible alternative to avoid the allocation of a large number of distinct IP addresses. During the activity with IBM SoftLayer a VPN was set up. It consisted of a GRE tunnel defined between a Vyatta virtual router on a bare metal server at Softalyer, and a Brocade router in the CERN datacentre. The VMs got public addresses from the CERN datacentre address space and were accessible through the GRE tunnel. This solution provided only half the throughput of the native IPv4 option, probably due to the encapsulation overhead. In addition VPN implies higher security concerns.

Reverse DNS resolution of public IP addresses assigned to VMs is necessary for the authentication processes of certain WLCG applications and thus for the correct functioning of all the services. While direct resolution can be served by the primary DNS servers of the domain in use (cern.ch in this case), the reverse resolution is delegated to the owner of the addresses. The DNS is also used to allow the distinction between the Tier-0 and external nodes. It is achieved associating sub-domains with respect to the underlying cloud provider, e.g. *.softlayer.cern.ch for IBM and *.tsy.cern.ch for T-Systems.

5. Data Storage

With T-Systems, the procured “cluster storage” was delivered as 500 TB of block storage. In order to provide access through protocols supported by the experiments, a standard grid storage solution, the Disk Pool Manager (DPM) [16], was deployed. For this, 50 VMs were reserved to act as disk servers, in addition to a VM deployed as head node. All storage VMs mounted fast block storage devices – two times 5 TB per VM – delivering up to 20 k IOPS per volume. While deployment of the system proved relatively easy, once basic infrastructure was in place (such as Terraform, Cloud-init and Puppet), reaching a production service level required a significant time investment. This was principally through configuring for local conditions (e.g. NAT and DNS), debugging infrastructure problems, and the integration and validation by the
LHC collaborations. The latter point required particular care as the extension was neither fully transparent, nor operated as an independent site. The commissioning phase lasted deep into the contract and was characterised by iterative, tri-lateral problem solving, involving the experiments, T-Systems and CERN on infrastructure or integration issues.

As the “read once via streaming” behaviour of Tier-0 reconstruction does not benefit from local storage because data can be read directly from EOS at CERN, other workflows were sought, for example where imported data would be read multiple times within the cloud. This is generally the case for minimum-bias data, where multiple selections are made from a shared set of pregenerated events with a data volume of around 40 TB. ALICE, ATLAS and CMS successfully imported data to the system and were able to fill the 10 Gb/s bandwidth this way.

Limited use was made of the storage system, mainly for scale test purposes. Figure 5 shows the aggregated internal network traffic among worker nodes and disk nodes in a dedicated test when about 600 VMs were running data intensive workloads, doing pre-staging of input data from the DPM cloud storage. The overlap among incoming traffic in the worker nodes and outgoing traffic from storage nodes proves that the traffic was purely internal to the cloud. The design throughput of the storage VMs (100 MB/s) would have allowed an aggregated traffic of around 3 GB/s, but this was achieved only for a limited amount of time. The average throughput during the data transfer, as measured by the jobs, was around 23 MB/s. Discrepancies among the design and the actual delivered performance required deep investigation and resulted in the identification of a storage node under performing. The root cause was identified by the cloud provider to be connected to the performance of network adapter on the host machine, but couldn’t have been spotted without the support of the client-side monitoring infrastructure.

![Network Traffic](image)

Figure 5: Aggregated network traffic of worker-node VMs (incoming traffic in blue) and storage-node VMs (outgoing traffic in orange), during processing activity dominated by data access accessing from the internal storage. The orange and blue curves mainly match.

6. Cloud Operation
A common expectation in the adoption of commercial cloud IaaS is the ability to promptly scale-out the compute capacity and accommodate in the cloud the excess of job slots requested by the user community (absorbing peaks of demand). Our experience highlights on the contrary that ramp-up time can be long and is normally related to the maturity and stability of the cloud service provider. Therefore preparation and commissioning phases are often needed to test the infrastructure or simply to establish WAN connectivity with GÉANT. The refined procurement process requires that during a preparation period of a few weeks the cloud provider grants CERN experts access to a small fraction of the target capacity in order to familiarise with the infrastructure and to prepare the acceptance tests. A commissioning phase then follows, where
the capacity ramp-up takes place and acceptance tests are executed, including the verification of the API functionality and scalability, and the measurement of the compute, storage and network performance under load. If all of the acceptance tests carried out establish that the IaaS cloud resources perform properly and in every other respect are in conformity with the contract, acceptance is granted and the delivery period starts.

7. Conclusion
With over two years of evaluation of commercial cloud providers, CERN has successfully been able to transparently integrate commercial resources into the existing Tier-0 computing infrastructure. As the adoption of cloud computing becomes more transparent, network and data also becomes an integral part of the activities. Only particular workflows will benefit from in-cloud storage, thus it is crucial to understand what the intended use of the public cloud resources is in order to specify and procure them effectively. The deployment of a grid storage technology such as DPM on cloud is relatively simple, but reaching a production service level requires a significant effort in particular to sort out issues that are hidden by the cloud virtualisation layers. For this reason we consider that the use wherever possible of “native” cloud technology, such as scalable object stores, would reduce integration and operation costs.

Reliable and large-bandwidth WAN connectivity proves to be able to sustain data intensive workloads even with remote access. As data is distributed across several grid sites, the quality of the WAN connectivity should be guaranteed between the cloud provider and multiple sites, at least the Tier-1 sites. Some sites may have limited connectivity to the commercial Internet, thus connectivity to the GÉANT cloud VRF is a recommended option for at least all the European sites.

Throughout these cloud activities it was also understood that consumer-side monitoring, accounting and benchmarking are absolutely crucial during the commissioning and the production phases, for acceptance purposes as well as for the validation on the providers’ usage reports.

The procurement process has been adapted with the experience gained and preparation and commissioning phases have now been fully integrated into the tender to assure a proper delivery period. The consolidation of this work will come through the HNSciCloud project [17], aiming for an innovative hybrid cloud infrastructure to support scientific research across Europe.

References
[1] Helix nebula. http://hnx.helix-nebula.eu. Accessed: 2017-02-09.
[2] C Cordeiro, A De Salvo, A Di Girolamo, L Field, D Giordano, R Jones, and L Villazon. Accessing commercial cloud resources within the european helix nebula cloud marketplace. Journal of Physics: Conference Series, 664(2):022019, 2015.
[3] P Buncic, C Aguado Sanchez, J Blomer, L Franco, A Harutyunian, P Mato, and Y Yao. Cernvm - a virtual software appliance for lhc applications. Journal of Physics: Conference Series, 219(4):042003, 2010.
[4] S de Witt and A Lahiff. Quantifying xrootd scalability and overheads. Journal of Physics: Conference Series, 513(3):032025, 2014.
[5] Domenico Giordano, Cristovao Cordeiro, Alessandro Di Girolamo, Laurence Field, Hassen Riahi, Jaroslava Schovancova, Andrea Valassi, and Luis Villazon. Cern-it evaluation of microsoft azure cloud iaas. Technical report, March 2016.
[6] Htcondor. https://research.cs.wisc.edu/htcondor/. Accessed: 2017-02-09.
[7] Puppet. https://puppet.com/. Accessed: 2017-02-09.
[8] E Karavakis et al. Unified monitoring architecture for it and grid services. to appear in J. Phys.: Conf. Ser., 2016.
[9] B Jones et al. Future approach to tier-0 extension. to appear in J. Phys.: Conf. Ser., 2016.
[10] Slipstream smart cloud application management. http://sixsq.com/products/slipstream/index.html. Accessed: 2017-01-17.
[11] Azure service management portal and service management api. https://msdn.microsoft.com/en-us/library/jj838706.aspx. Accessed: 2016-12-07.
[12] Azure resource manager overview. https://docs.microsoft.com/en-us/azure/azure-resource-manager/resource-group-overview. Accessed: 2016-12-07.

[13] C Cordeiro, A Di Girolamo, D Giordano, L Field, D Spiga, and L Villazon. Monitoring the delivery of virtualised resources to the lhc experiments. *Journal of Physics: Conference Series*, 664(2):022013, 2015.

[14] C Cordeiro et al. Real-time complex event processing for cloud resources. *to appear in J. Phys.: Conf. Ser.*, 2016.

[15] D Giordano et al. Benchmarking cloud resources. *to appear in J. Phys.: Conf. Ser.*, 2016.

[16] A. Alvarez, A. Beche, F. Furano, M. Hellmich, O. Keeble, and R. Rocha. *J. Phys.: Conf. Ser.*, 396:032015, 2012.

[17] H Meinhard et al. The hnsicloud project. *to appear in J. Phys.: Conf. Ser.*, 2016.