

Experience in using public Cloud for Belle II experiment within HNSciCloud Project

Silvio Pardi^{1,*}, *Davide Michelino*²

¹INFN – Napoli Unit – Via Cintia, 80126, Napoli Italy

²GARR – Via dei Tizii, 6 00185, Roma Italy

Abstract. The current level of flexibility reached by Cloud service providers enables Physicists to take advantage of extra resources to extend the distributed computing infrastructure supporting High Energy Physics Experiments. However, the discussion about the optimal use of such kind of resources is currently ongoing. Moreover, because each Cloud provider offers its interfaces, API sets, and different service levels, the integration in the computing model of an experiment requires a specific work for each provider. In this paper, we summarize the experience matured in the usage of Cloud resources offered by different providers within the European project Helix Nebula Science Cloud for Belle II experiment. The goal of the user-case was to run Montecarlo Production jobs properly on Virtual Machines created dynamically. To do that, we defined the golden image, based on CENTOS; we then set up the basic environment, the internal and external network, and local DNS services. For each cloud, we implemented a cache server based on Squid, while the orchestrator has been centralized on a single server located in the INFN-Napoli infrastructure. Finally, we integrated the available resources in the Belle II production framework based on DIRAC, using VCYCLE as Virtual Machine Life Cycle Manager. After a tuning activity that required the optimization of several aspects included Image distribution and Network configuration, we were able to run production jobs over the different Cloud with success.

1 Introduction

The Helix Nebula Science Cloud (HNSciCloud) [1] is a European pre-commercial procurement (PCP) project co-funded by the European Commission Horizon 2020 which aims at creating a competitive marketplace of innovative cloud services for scientific user communities. Among them, Belle II was one of the selected use-cases, with the goal to demonstrate the possibility of successfully running standard Monte Carlo production jobs and user-analysis on Public Cloud infrastructures selected by the HNSciCloud project without bottleneck and trying to maximize the investment.

Belle II uses a distributed computing model [2] based on Grid and non-Grid resources integrated through the DIRAC framework. The DIRAC framework is responsible for managing computing and storage resource and offering the Work Load Management

* Corresponding author: spardi@na.infn.it

System for the experiment. The possibility to integrate flexible Cloud resources properly represents a topic of great significance with potentially great benefit to the community. This paper presents an experience with porting Belle II application on Helix Nebula Science Cloud infrastructure using VCYCLE [3] as Virtual Machine Manager. The paper is organized into four sections; section 2, 3, 4 and 5. Section 2 describes the HNSciCloud infrastructure and the available resources. Section 3 shows how we integrated the available Public Cloud resources in the Belle II DIRAC framework, section 4 demonstrates some utilisation statistics and finally section 5 provides a summary of the discussion and the conclusion.

2 HNSciCloud Resources

During the first stage of the HNSciCloud project, we made use of Cloud resources belonging to three different consortiums; one run by IBM, the second led by T-System and the third by RHEA. The first consortium provided a proprietary interface to interact with the IaaS system and a set of specific command line tools. Each VM in the tenant provided public and a private IP automatically which provided the ability to create their images. In the case of the second consortium, T-system provided a customized version of OpenStack through OTC cloud which created a local network and assigned a pool of public IP used to bind the single machine when needed. It provided all standard OpenStack service for image and disk volume management. The IaaS provided web access, command line, and API. The consortium led by RHEA provided an integrated system called Nulva, offering multi-cloud access. An overlay stratum from their infrastructure provided the possibility to create Virtual Machines over different IaaS. Among the accessible open cloud even OTC was available with standard Openstack interfaces.

3 Integration of HNSciCloud in Belle II Framework with VCYCLE

For each cloud, we implemented a cache service based on squid over a separate VM. We then registered it in a local DNS (for the OTC based clouds). The VCYCLE orchestrator has been centralized on a single server located in the INFN-Napoli infrastructure (ReCas-Napoli data centre). For the network setup, we had to use a specific cloud configuration. With IBM for instance, we configured the VMs using a public address, and on T-system and the RHEA solution, we used a private IP and a SNAT server on a separate VM with a 10Gbps connection to the WAN.

Cloud systems have been integrated into Belle II frameworks based on DIRAC by using VCYCLE as Virtual Machine Life Cycle Manager. The original VCYCLE service was used to create and destroy Virtual Machines automatically. This was done using an internal engine that interacts with several IaaS. For the case of T-System and RHEA, we used standard OpenStack API taking advantage of all the feature of VCYCLE. Because the VCYCLE does not natively support the specific API interface in the case of IBM cloud service, a script was created inside the virtual machine that would allow automatic scratch on the user area inside the VM and restart the DIRAC-pilot script at the end of a job.

The three IaaS infrastructures were separately registered in both the central DIRAC server hosted at KEK and in the DIRAC server for Validation host at PNNL. To differentiate the three infrastructures we used the following nomenclatures: VCYCLE.HNSC01.it, VCYCLE.HNSC02.it and VCYCLE.HNSC03.it. That way, we completed the integration of the three IaaS infrastructure provided by HNSciCloud with the Belle II framework for simulation and analysis.

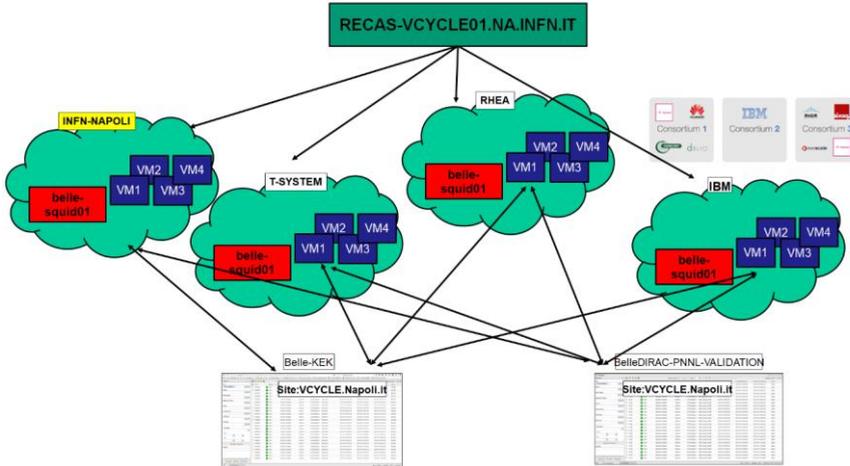


Figure 3.1 – The schema shows the three HNSciCloud integrated on the production and validation DIRAC servers of Belle II, together with the private cloud already used in Napoli.

4 Scalability test during the Belle II MC campaign

The testing activity has been divided into two steps the validation phase, and the scalability phase. The validation phase has involved the three infrastructures, while the scalability phase has been mainly performed over the OTC infrastructure. During the validation phase, a set of basic Belle II testing jobs was created and run inside several VMs of the three different infrastructures. In so doing, we demonstrated that it was possible to run the basf2 environment properly using CVMFS for software. We then tested the full production workflow over the DIRAC Validation server hosted in PNNL. This was done by sending production like jobs pushed in the VM using the pilot. Those tests completed the first phase of the project activities demonstrating the feasibility of our approach and the proper usability of the provided resources.

The scalability tests have been performed over the OTC IaaS provided by the T-SYSTEM. The goal of the activity was to proof the ability to use the Cloud resources in the production scaling up to the maximum number of available cores provided by the vendor without difficulty. To do that, we integrated the T-SYSTEM Infrastructure on the DIRAC Server for production and then included the relative site, named VCYCLE.HNSC01.it, in the official Monte Carlo Campaign.

One of the key points to take into account was the choice of the best devices for each VM (i.e., the number of CPUs, RAM, and disk size), to maximize the number of running jobs. The main factors that affected the choice were the Belle II jobs requirements and resources quota. For the job requirements, we used the following quotas: 20GB of the disk, 4GB of RAM per computational core, and 20GB of the disk for operative system. The resource quotas, instead, was changing weekly after a negotiation between INFN and other HNSC partners. Each quota limit uses different parameters, i.e., the maximum number of instantiable CPUs, the max number of VMs that we can create the maximum amount of RAM and the maximum amount of storage space that we can use, and the maximum number of disks that we can create.

To maximize the number of cores utilized at the same time, we created the following machine combinations:

- 4 cores 16GB RAM and 100GB Disk
- 2 cores 8GB RAM and 60GB Disk
- 1 core 4GB RAM and 40 GB Disk

The amount of VMs for each combination was defined weekly in function of the available quotas.

The period between 12 March and 1 May 2018, we had the maximum number of available resources for INFN use-case. In this period we as able to scale out up to 1275 cores, reaching even the TOP 10 among the infrastructure providing computational resources for the experiment during MC production (see figure 4.1 and 4.2). The test demonstrated that no specific challenge had prevented the proper use of a high number of resources, both the network and computational resources.

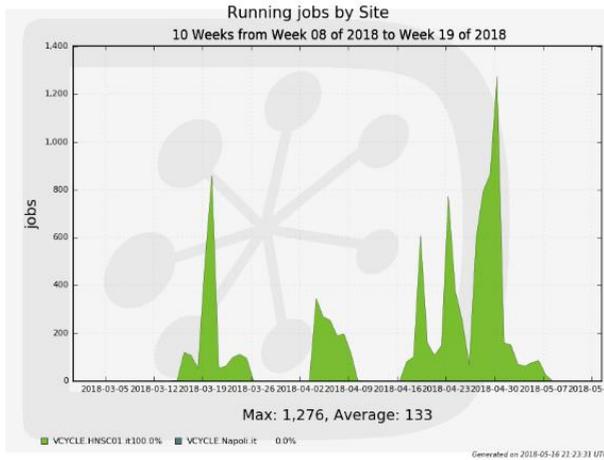


Figure 4.1 –The picture shows the number of running jobs on the OTC infrastructure

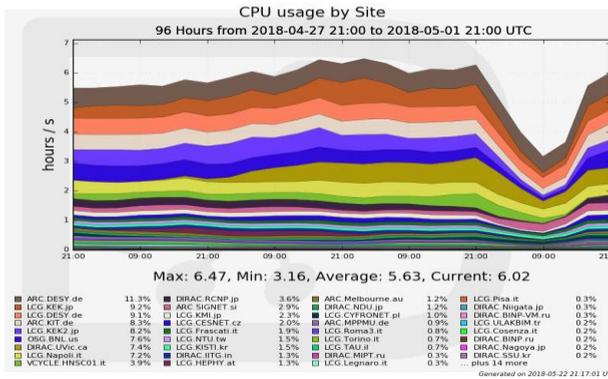


Figure 4.2 The graph in the figure shows the total amount of CPU used by all Belle II sites in the time window 27/04/2018-01/05/2018. We can see that in this period the site VCYCLE.HNSC01. It reached the top-10 of site contributors.

5 CONCLUSIONS

In conclusion, the different phases of configuring the public cloud computing for use within the HNSciCloud come with different experiences. The tuning phase allows the use of resources provided by Public Cloud providers using standard Belle II framework tools and VCYCLE as Virtual Machine Management. The scalability phase further allowed all the available resources on the cloud to be put to use growing up to the quota limits. It is, therefore, possible to use the Public Cloud services regardless of the approach. However, a certain amount of time must be taken into account when configuring the environment for the proper and efficient running of user-specific applications. Preliminary steps may be different in function of the specific platform run by the Cloud Provider, among them, the network setup represents one of the most laborious tasks. The second aspect to take into account is the balance of different quotas (i.e., #CPU, RAM, number of the machine, number of disks) that represents one of the key points to run applications in optimal conditions over a pay-per-use infrastructure. While it may, therefore, be possible to use the public cloud for Belle II experiment within the HNSciCloud project, all these aspects must be put into consideration.

References

- [1] Helix Nebula Science Cloud home page - <http://www.helix-nebula.eu/>
- [2] S.Pardi et al. "Computing at Belle II" - Belle-II computing Group - Nucl.Part.Phys.Proc. 273-275 (2016) 950-956 - DOI: 10.1016/j.nuclphysbps.2015.09.148
- [3] McNab, A., et al. "Managing virtual machines with vac and vcycle" Journal of Physics: Conference Series 664(2),022031 2015

Acknowledgments

Authors wishing to acknowledge the assistance of the staff of the ReCaS Data center, the computing group of Belle II and the site admin of the Belle II production infrastructure.