International Journal of Cloud Applications and Computing
Volume 9 • Issue 1 • January-March 2019

U-LITE, a Private Cloud Approach for Particle Physics Computing

Barbara Demin, INFN, CNAF, Bologna, Italy
Sandra Parlati, INFN, LNGS, Bologna, Italy
Pietro Fedele Spinnato, INFN, TIFPA, Bologna, Italy
Stefano Stalio, INFN, LNGS, Bologna, Italy

ABSTRACT

The computing infrastructure of Laboratori Nazionali del Gran Sasso (LNGS) is the primary platform for data storage, analysis, computing and simulation of the LNGS-based experiments, which are part of the research activities of the Istituto Nazionale di Fisica Nucleare (INFN). Scientific collaborations running such experiments have diverse needs and adopt different approaches in developing the computing frameworks that support their activities. The emergence of resource virtualization and the cloud paradigm has allowed a complete change to the way of operating and managing the LNGS computing infrastructure, and realize U-LITE, a private versatile environment apt at hosting such varied ecosystem and providing LNGS scientific users a familiar computing interface which hides all the complexities of a modern data center management. Since 2011, U-LITE has proven to be a valuable tool for the LNGS experiments, and provides an example of effective use of a private Cloud computing approach in a real scientific context.

KEYWORDS

Astroparticle underground physics, Batch Scheduling in a Virtualized Environment, Computing Infrastructure Case Studies, Cooperative Resource Sharing, Private Clouds, Scientific Data Management

INTRODUCTION

LNGS is a world-renowned research site,[1] where cutting-edge research in several branches of particle physics spanning from, e.g., search for dark matter[2] to neutrino physical properties,[3] or studies on nuclear reactions in stars[4] are carried on within the world-largest underground laboratory for particle physics. Data acquired from the experimental apparatuses need to be stored and analyzed. Detection systems and processes need to be modeled and simulated. Such diverse tasks require a suitable computing environment. Ever since research activities begun at LNGS in the late 1980's, the LNGS Computing and Network Service (CNS) has been committed to provide the research groups with the IT infrastructure and the professional support required for their scientific computing needs.

The way such goal has been achieved has changed over the years, often following the natural evolution of technology. The computing model of the first experimental projects run at LNGS was based on a highly centralized structure, whose main components were a VMS cluster and the DECNET network, both managed by the CNS. Later on, technological evolution brought to the rise of a more inhomogeneous computing model, based on different UNIX/Linux clusters devised by staff of the experiments built at LNGS in the late 1990's and early 2000's. The resulting variety of environments

DOI: 10.4018/IJCAC.2019010101

was too large to be managed by the CNS. Each project managed its computing cluster, while the CNS was in charge of providing resources for general use, including batch systems (Condor, NQS, LSF) for those who had not a cluster of their own, disk storage and tape backup management for experimental data, as well as the basic IT services.

This arrangement, though convenient and effective in some cases, was obviously inefficient in terms of resource and staff utilization. The rise of virtualization and the consequent advent of cloud computing as a paradigm to abstract resources and provide multiple user-defined environments on a common infrastructure, was a turning point in IT resource provisioning. CNS recognized the high potential of this approach to overcome the fragmented situation that LNGS was experiencing. Re-unifying the IT management substrate under the sole responsibility of CNS, in a multi-tenant environment (Krebs et al., 2012) that could provide each research group the most suitable and personalized computing environment for its scientific needs was the design goal for a new LNGS computing platform that virtualization tools would allow to achieve. The outcome of such effort was the Unified Lngs IT Environment (U-LITE). Characterizing features of U-LITE are a user-transparent coupling of a well-known batch system, Torque/Maui (Adaptive Computing, 2018) with a hypervisor, Proxmox, (Proxmox Server Solutions, 2018) for a seamless provisioning of virtualized resources dynamically adapting to user requests, and an overall management of the whole data stream, from the data transmitted by the data acquisition hardware in the underground Labs up to the data storage and backup systems in the Data Center which hosts U-LITE, located in the external part of the Labs.

When U-LITE started, cloud computing was not yet so widespread as it is today and open source cloud platforms were not as well-known and advanced as they are today. That is why U-LITE implemented a cloud-like platform instead of adopting an existing one.

The choice of the aforementioned scheduling tools, which only run on Unix-like Operating Systems, apparently limits U-LITE versatility. In fact, in our environment the vast majority of users develops its computational tools on Linux platforms, and being limited to Linux has not been experienced as a problem by our users.

The technical features of U-LITE will be described in the next sections, together with an analysis of the usage statistics collected over the seven years the system has been in use, demonstrating how the adoption of a cloud paradigm in a research environment can be profitable in terms of resource optimization.

MAIN FEATURES

U-LITE has been developed with a bottom-up approach, having in mind the needs and requirements of its final users, the scientific community of LNGS. In this perspective, its features are naturally adhering to those of a private cloud, (Mell & Grance, 2011) as the infrastructure is provisioned for being primarily used by the LNGS-based researchers. In fact, access to U-LITE has also been granted to a number of external users. Yet, access procedures for external users are based on scientific collaboration agreements, and do not involve financial transactions, therefore they are quite dissimilar to those commonly adopted for public clouds access.

A characterizing feature of U-LITE is its data-flow model. The *raison-d'être* of LNGS is to a large extent the production of experimental data in the underground laboratories, which may contain evidence for yet undiscovered physical phenomena. The process leading to a scientific discovery in this context passes through a careful transport of such data from source (the experimental apparatuses underground) to destination (the U-LITE SAN within the data center in the outside labs and possibly data-centers belonging to other, world-spread institutions), ensuring no data loss, and its scrupulous analysis with the software resources developed by LNGS scientists and running on the U-LITE computing platform. It is therefore clear how precious a resource are scientific data in our context, and the extreme care the CNS devotes to protect experimental data from loss and corruption.