Enhancing Grid Infrastructures with
Virtualization and Cloud Technologies

Infrastructure Specifications

Deliverable D5.1 (V1.0)
14 September 2010

Abstract

This document presents the computing infrastructure deployed in the context of StratusLab project, details the configuration of the computing resources used and the commitments made from the project partners to contribute with computing resources in the project. During the first months of operation this infrastructure has already been significantly exploited to deliver the first results towards the project's goals. The OpenNebula virtual management software has been used to install private clouds on different OS platforms. Two pre-production grid sites have been deployed in the private cloud and are used to test the implications of providing grid services in cloud environments. Experience from the deployment and operation of the above sites will help us identify the required tools and procedures for offering grid production sites over computing clouds. Finally the document presents related work, relevant to infrastructure operations, taking place in other projects and initiatives.

StratusLab is co-funded by the
European Community's Seventh
Framework Programme (Capacities)
Grant Agreement INFSO-RI-261552.

Contributors

Name	Partner	Sections
Evangelos Floros	GRNET	All
Ioannis Konstantinou	GRNET	Related Activities
Stuart Kenny	TCD	Preliminary results and TCD infrastructure
Evangelos Aggelou	GRNET	GRNET Infrastructure
Charles Loomis	CNRS	CNRS Infrastructure. Overall internal review
Christina Mpoumpouka	GRNET	Preliminary results

Document History

Version	Date	Comment
0.1	12 July 2010	Initial TOC version for comment.
0.2	5 Sept. 2010	Added content to introduction section. Integrated contributions regarding appliance repository. Added introduction to infrastructure section and draft content regarding GRNET infrastructure.
0.3	9 Sept. 2010	First complete draft. Still missing the executive summary and the summary in the end
0.4	13 Sept. 2010	Incorporated changes and corrections proposed from internal review (Cal Loomis)
1.0	13 Sept. 2010	Final version.

Contents

List of Figures

List of Tables

1 Executive Summary

This document presents the computing infrastructure deployed in the context of StratusLab project, details the configuration of the computing resources used and the commitments made from the project partners to contribute with computing resources in the project. This infrastructure is an important tool for the project and provides the resources required for the deployment of the project's private cloud, the provision of grid services on top of this cloud, for performing various tests and benchmarks and for the provision of auxiliary computing services needed internally by the project.

During the first months of operation this infrastructure has already been significantly exploited to deliver the first results towards the project's goals. The OpenNebula [11] virtual management software has been used to install private clouds on different OS platforms (namely Ubuntu and CentOS). Two pre-production grid sites have been deployed in the private cloud and are used to test the implications of providing grid services in cloud environments. The installations were performed either manually or using the Quattor [15] toolkit. The latter has been also used to implement various proof-of-concepts relating to the management of grid services and of the WNs (Worker Nodes) in particular. A core component of the infrastructure architecture is the appliance repository which is used to host and provide virtual machine images to be used in the context of the above private clouds. A first version of the repository has been deployed and a set of images is available for usage facilitating the fast deployment of gLite [5] based grid sites.

Taking into account this experience the project is already starting to investigate requirements stemming from the fusion of grid and cloud technologies. For example the impact on grid operations are considered. This is expected to affect the tools and procedure used currently in large scale grid computing infrastructures like EGI [4]. Aspects to be affected are for example the Service Level Agreements of cloud resources, the monitoring tools used in grid operations and the accounting policies applied to grid resource providers and end users.

Finally, the project is keeping an eye to other relevant activities taking place in the context of other national or international projects and initiatives. For instance, CERN has used OpenNebula to virtualize and manage part of its large computing infrastructure used in various experiments. Elasticity of grid sites has been also investigated by INFN in the context of their Worker Node on Demand initiative. Also, NIIF in Hungary has used OpenNebula to deploy a national cloud service, with significant contribution in the development of virtualized storage capabilities. StratusLab has already initiated contact with some of these activities and is aiming to re-use and expand their findings.

2 Introduction

The StratusLab project started in June 2010 with the goal to investigate the potential benefits and possible implications stemming from the fuse of cloud and grid technologies. This goal is twofold: on one hand grid services can be deployed and provided on top of computing clouds exploiting the inherent capabilities of virtualization, such as the on-demand provision of resources, resource elasticity, remote management and optimal usage of computing resources based on workload demands. On the other hand grids can be extended to offer cloud-like interfaces and thus advanced processing and storage capabilities to the end users. Apparently, in order to pursue the above goals the project should deploy and manage its own physical computing infrastructure.

This document presents the computing infrastructure deployed in the context of of the project's WP5 activity during the first three months of the project. The concrete goals of this infrastructure is:

- to provide a testing platform for the project,
- to demonstrate production use by running grid services with the StratusLab distribution, and
- to host various project services

The document is organized as follows: In Section 3, we detail the technical specifications of the computing resources committed by various partners of the project. During the first three months of operation already there has been significant progress towards the project goals. The operation of this infrastructure has given us a better insight on the implications of deploying and providing private cloud services and their usage for hosting grid sites. A first set of tools and procedures has been set (Section 4) which will further be elaborated in the coming months. Section 5 provides more details on the initial results achieved during the first months of cloud infrastructure setup and operation. The usage of cloud technologies for the provision of grid computing capabilities is a hot topic and is currently being pursued also in other contexts and projects. Awareness of such activities is important for StratusLab in order to avoid duplicated effort and also investigate the possibilities of collaborations. The most significant of these activities are briefly presented in Section 6.

3 Physical Infrastructure

One of the main goals of the StratusLab project is the installation and provision of a computing infrastructure, able to efficiently host cloud computing services, capable of hosting production level grid computing sites. Apart from that many of the service and research activities of the project will require computing resources for various purposes: hosting of development and collaboration tools, testing, benchmarking etc. In this section we describe the computing resources that the partners participating in WP5 have committed to provide, for the aforementioned purposes. In particular we differentiate between the infrastructure committed for hosting of cloud services and the infrastructure allocated for the virtual appliance repository, an integral component of StratusLab toolkit architecture.

3.1 Cloud Services

Two sites have been committed to the provision of the required physical infrastructure for hosting the StratusLab cloud distribution. These sites are operated by GRNET and CNRS. The main purpose of this sites is to provide the computing resources needed for the deployment of OpenNebula [11] virtual machine manager which sits at the core of the StratusLab toolkit. At these early stages of the project this infrastructure is used for various test installations of OpenNebula in different platforms. The results of these early tests are presented in more detail in Section 5

3.1.1 GRNET infrastructure

GRNET is a key partner of WP5 being the leader of the activity as well as the provider of the largest part of the StratusLab project computing infrastructure. GRNET is interested to investigate on the emerging paradigm of cloud computing and capitalize on the novel capabilities related to the efficient infrastructure management and on-demand provision of elastic computing capabilities, which is an integral part of cloud computing. For this reason GRNET has constructed a new datacenter dedicated to cloud computing applications. The datacenter is located in the central building of the Ministry of Education (YPEPTH), is equipped with cloud-optimized hardware and is used for various activities related related to cloud services either in a research or production level. A subset of this infrastructure has been allocated for the purposes of StratusLab and is described in detail in following paragraphs. The topology of the datacenter is depicted in Figure 3.1

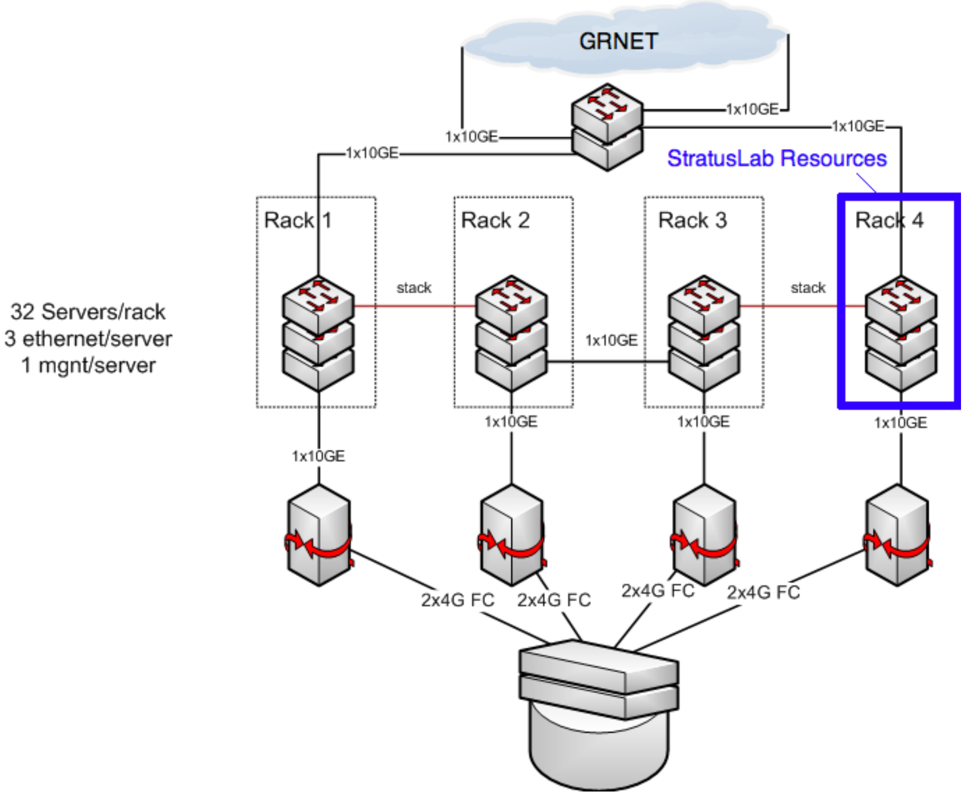

Figure 3.1: *Bird's eye view of GRNET's infrastructure*

3.1.1.1 Computing hosts

On the GRNET infrastructure, 26 physical hosts have been assigned for use by StratusLab, organized in two separate computing clusters. All nodes are Fujitsu RX200S5 1U rack servers, with 2 quad core Xeon 5500 series with the ability to run 2 threads per core through hyper-threading. The nodes reside inside the same rack ensuring optimal communication latencies. In total the above setup provides 208 physical cores, which are able to host a maximum of 416 single-core virtual machines. Each node has 48 GB of RAM, allowing for high performance VMs that can cope with memory intensive workloads. Each host is also provided with a 146 GB HD SAS 3G hard disk, for hosting the OS root and for the VM cache space.

3.1.1.2 Storage

Storage is provided by an EMC Cellera NS-480 Cloud/NS-480 storage server. The total capacity of the system is 280 TB. From them, 20 TB have been initially been allocated for the requirements of StratusLab infrastructure, made available through NFS over the resident networking infrastructure. Alternative file access protocols will also be investigated (e.g. iSCSI).

3.1.1.3 Network

The Networking infrastructure consists of 3x10 GbE (Gigabit Ethernet) switches. Each of the physical hosts has 3 GbE interfaces plus a management interface for access through IPMI. The relevant traffic (host/VM/IPMI management) is isolated using separate VLANs. The connectivity between the servers rack and the storage server is provided by a fiber channel bridge which connects the 2x4 Gb FC storage link with the 10 Gb servers link.

The datacenter is connected through 2x10 GbE links with the GRNET network backbone. GRNET is the Greek NRN (National Research Network) offering data connections up to 10 GbE with the GEANT network.

3.1.1.4 Management software

The nodes' management is accomplished through IPMI, using either a web or command line interface. Host OS installation on the nodes is automated through PXE. A central management server provides the host OS repository (currently Debian Lenny, Ubuntu 10.04 and CentOS 5.5) for network installation and a DHCP server for host IP address allocation. Monitoring of the physical nodes is provided using Ganglia scalable distributed monitoring system.

3.1.2 CNRS infrastructure

CNRS/LAL is the second resource provider in WP5. LAL will offer a cloud test bench for testing the installation and performance of the StratusLab cloud distribution. Once the StratusLab cloud distribution is stable, LAL will gradually migrate its production grid infrastructure to it, starting with the grid Worker Nodes and progressing to other services as confidence increases. Although technically not part of WP5 activities, the Institut de Biologie et Chimie des Protéines (IBCP) will also run a test bench and production infrastructure over the StratusLab distribution.

3.1.2.1 Computing Hosts

The test bench at LAL currently consists of five 1U, rack-mounted machines. These are dual processor IBM x3550 machines with Intel E5345, 2.33 GHz, 4 core processors for 8 cores in total. Each machine has 16 GB of RAM and 150 GB of local disk space. These machines allow hardware virtualization; correspondingly, LAL has decided to use KVM [8] as our standard hypervisor technology. Although acceptable for deployment and small-scale tests, the test bench will have to grow to accommodate larger and more diverse tests. Internal funding requests have been made to double the size of the test bench.

3.1.2.2 Storage

A Storage Area Network (SAN) provides large-scale storage, accessible to all machines at LAL, including specifically all of the grid and cloud machines. A DDN 6620 connects to controllers via double Fibre Channel connections. The controllers are then connected to the LAL's primary network switch via 10 Gb/s connections.

Currently, the cloud test bench has no allocated storage; however, internal funding requests have been made for providing dedicated storage for the StratusLab activities. Regardless of the outcome of those requests, an allocation of 5–10 TB from the current storage resources can be easily and immediately provided when the need arises.

3.1.2.3 Network

All of the machines connect to the local network via 1 Gb/s network connectors. Typically each rack of machines has a dedicated network switch, operating at 10 Gb/s, which is in turn connected to the wide area network by RENATER [16]. The cloud machines have a dedicated range of IP addresses which are isolated from other nodes at LAL with a VLAN.

RENATER currently provides a 10 Gb/s link which is shared by LAL and several other surrounding laboratories. The network supports teaching and research activities in France; current policies prohibit commercial use of the network.

3.1.2.4 Management Software

LAL uses the Quattor Toolkit [15] to manage all of the machines in the LAL machine room as well as the services provided by those machines. Quattor permits centralized, shared management of the installation, configuration, and maintenance of a large number of machines. As for all other machines and services, the cloud test bench is installed, configured, and maintained via Quattor. This demonstrates the compatibility of the StratusLab distribution with automated site management tools and serves as a continuous verification of that compatibility as the StratusLab distribution evolves.

3.2 Virtual Appliances Repository

An Apache web-server, hosted by Trinity College Dublin (TCD), provides the initial virtual appliances repository. The web-server is accessed using WebDAV (Web-based Distributed Authoring and Versioning), with authentication via the StratusLab central LDAP server. A single virtual machine is used to host the web-server, with an initial storage allocation of 200 GB. As the number of appliances increases, and the average size of appliance is better known, this allocation will be reviewed and increased as required.

3.3 Support Services

To maintain core support services, a host similar to the computer nodes has been allocated for use by the project. Currently, this server runs the Hudson integration server, and has the same management and monitoring utilities as are available for the computing hosts (outlined in Section 3.1.1). More services can be collocated or supported on new hosts as the need arises, as GRNET's infrastructure provides the flexibility and processing power necessary for the project's purposes.

4 Tools and Procedures

Operation of a large scale computing infrastructure requires the utilization of a number of specialized tools and the establishment of a set of formal procedures for the provision of the underlying computing resources. Typically in a grid environment these tools and procedures are imposed by a centralized authority responsible for their definition, installation and proper operation. For example in the context of EGI, all sites have to adhere to a specific Operational Level Agreement (OLA) for what concerns the availability of the site's resources. Also the sites have to install a specific set of monitoring clients in order to enable centralized visual monitoring of the sites status for the various provided services. Obviously, the grid sites installed in the context of StratusLab, once scheduled for going through the official certification process defined by EGI, will adopt the set of tools and procedures that will be put in place by the organization.

On the other hand, for what concerns the cloud services, no standardized process or tools are available, and no specific SLA has to be followed. All these will be one of the main outcomes of WP5 activity. Indeed, the above will be one of the main subjects of Deliverable D5.2 scheduled for M6 of the project. Meanwhile, the activity has adopted some simple procedures for the provision of infrastructure services. For the time being such services are available only internally to the project members. Project members can apply for a certain resource (e.g. a physical node, storage, a VM running a specific image, the creation of a specific image, the setup of a grid site) by simply sending an email to WP5.

As mentioned in the previous section, a set of tools is already in place and are used for managing the infrastructure on the physical level, and the monitoring of physical hosts. These tools in some cases have been imposed by the vendor (e.g the online monitoring offered by the Fujitsu servers) and in other they have been selected due to their popularity (e.g. Ganglia[13] for host status monitoring, Nagios[9] for overall system monitoring and service availability). In the coming months the evaluation of more operational tools will be intensified in order to cover all aspects of infrastructure service provision.

For what concerns Service Level Agreements in particular, obviously it is very pre-mature to consider it this moment. A simple approach will be for cloud services to offer such a level of service that will enable the provision of grid sites that can satisfy the EGI OLAs. This rather generic requirement can serve as a good starting point. Certainly the provision of cloud services has much more implications comparing with a grid site where site availability is the main metric for assessing the site's quality.

For example in the context of EGI two are the basic metrics used to ensure the adherence of a certain site to the OLA: Site Availability and Site Reliability. The first metric refers to the percentage of time the the site is on-line and available. The former metric reflects the percentage of jobs submitted to the site and complete successfully. Certified grid site should demonstrate at least 70% availability and 75% reliability in order to retain their operational status.

In cloud environments we will consider additional metrics for SLAs such as:

- Availability and reliability of VM hosting environments and virtualized data storage
- Minimum time between requesting a VM instantiation and the time the VM is available
- Overall availability of service for accessing through the relevant cloud API

Security of data is also crucial in a cloud environment. Thus, SLAs should cover provisioning of security services, like encryption of data (with a choice of cryptographic algorithm strength used) and overall data integrity.

5 Preliminary Results

Activity has already started for the various tasks of WP5. In the paragraphs we report on the preliminary findings and the early results already available from these activities.

5.1 Installation of OpenNebula

Once the physical infrastructure was put into place it was decided that the activity should start exploiting the capabilities of offered cloud services using the current version of various StratusLab components and in particular of OpenNebula which is the core of the project's distribution. For this purpose two clusters have been installed in the GRNET infrastructure. Namely, an Ubuntu-based cluster and the CentOS-based cluster, each of them consisting of eleven physical nodes. The choice of Ubuntu Server 10.04 and CentOS Operating Systems for the clusters, amongst the ones supported by OpenNebula, was based on their popularity for use in datacenters and ease of use. There is an OpenNebula 1.4 deployment on each cluster.

The procedure of the OpenNebula installation was more or less the same for both clusters. The option of building from source seems to be the most appropriate. In both clusters the first node is the OpenNebula front-end node, where the management of the Virtual Machines (VMs) takes place, and the rest are added as host nodes in the private cloud. OpenNebula uses virtualization drivers to create, control and monitor VMs on the cluster nodes and offers support both for Xen and KVM (Kernel-based Virtual Machine) virtualization technologies. Both clouds were configured to use KVM, taking into consideration the facts that beta support is provided under Ubuntu and CentOS and that KVM is under active development. The management of the VM images is handled by transfer drivers that are used to transfer, clone, remove and create images. If the option of saving a VM image when the VM is stopped is enabled, the modified image is saved, so that in subsequent uses of the VM any configuration done or data stored in it are available. The decision for which of the transfer drivers to use is made according to the storage configuration of the cluster. The available options for the storage configuration are shared using NFS, non-shared using SSH and LVM (Logical Volume Manager) storage model. In the shared arrangement the cost of copying the images between the nodes of the cluster is minimized and allows live migrations of virtual machines between the host nodes. In the non-shared case, the images are copied to host nodes through ssh and are transferred back on the front-end node if the option of saving the VM image is enabled. The LVM (Logical Volume Manager) storage model is not fully supported and needs to be configured for the specific use case, however it provides the snapshot feature which instantly creates disk devices that can be assigned to the VM. At the moment the ssh transfer driver is used, but that will change to using NFS, as it is more effective. In addition OpenNebula provisions for the creation of virtual networks, in order for the virtual machines to acquire network access.

There are two virtual networks defined for each cloud deployment, a private and a public one. The mapping of the virtual networks on top of the physical ones is done by a network bridge set up over the physical public network of the cluster nodes. All virtual networks used are fixed, which means that a fixed set of IP-MAC pair of addresses is defined for the virtual machines.

LAL uses the Quattor Toolkit for the automated installation, configuration, and management of machines. To ensure that the cloud test bench at LAL can be similarly managed, LAL has developed Quattor configuration components for OpenNebula (ncm-oned) and for libvirt (ncm-libvirtd). In addition, a complete set of configuration templates have been developed to allow an entire OpenNebula system (front-end and hosts) to be installed automatically from bare metal. Documentation (http://stratuslab.eu/doku.php?id=quattorinstall) on the StratusLab internal wiki explains how to use the configuration as well as the current limitations.

5.2 Grid Services on Virtual Resources

Two grid sites have been deployed on top of the respective cloud sites in GRNET's infrastructure, using the above pre-configured images. The first site, named GR-42-STRATUSLAB (GR-42 for short), comprises of a CE, SE and 2 WNs. The second site, named GR-43-STRATUSLAB (GR-43) comprises of a CE, SE and 5 WNs. Both sites are in pre-production phase and are monitored by the SAM-tests (Service Availability Monitoring) running in NGI-GRNET (https://sam.athena.hellasgrid.gr). Both sites support MPI (MPICH2 and OpenMPI). Currently they don't provide shared folders among the WNs and the execution of MPI code is facilitated with the usage of MPI-START [7] tool.

For the deployment of grid sites over OpenNebula, we generated pre-configured images for each type of Grid node, i.e. CREAM (Computing Resource Execution And Management) Computing Element (CE), Storage Element (SE), Worker Node (WN) and User Interface (UI), using the gLite middleware version 3.2. The pre-configured images are running Scientific Linux 5.5 with all the appropriate packages of glite installed according to the grid node type. By creating a virtual machine with a pre-configured image for e.g. a Computing Element node, a fully functional Computing Element grid node is created. The configuration of the node using the yaim tool is done by a script that needs to be run once the virtual machine is up.

On LAL's site, the Quattor Toolkit is used to bootstrap machine installation via PXE. LAL has verified that PXE images can be started via OpenNebula and that they can be used to install grid Worker Nodes via the standard Quattor installation mechanism. This is a first indication that automated site management tools can be used with little or no modification to control grid services running on a StratusLab cloud infrastructure.

5.3 Virtual Appliances Repository

The initial prototype of the virtual appliance repository is deployed as a WebDAV-enabled Apache webserver. The repository is intended for use by both system administrators and end-users. StratusLab tools that interact with the repository have already been developed and are in active use. It is possible to customize an image and subsequently upload the resulting appliance to the repository using a single command-line tool. Appliances can be retrieved during deployment simply by specifying the URL of the required appliance contained in the repository.

Currently only a limited number of base OS images are available in the repository. This will be extended to include pre-configured appliances with core grid services, such as Computing Element and Storage Element. It is expected that the availability of these appliances will significantly simplify the task of grid site setup for system administrators.

A limitation of the initial service is that upload and download of appliances is only possible through the use of standard tools such as *cURL* [6]. The use of Maven as a front-end for uploading/downloading appliances from the repository is to be evaluated, and it is for this reason that the structure of the repository has been designed to mirror that of a Maven repository. Appliances are organized by type and operating system, for example:

```
grid
  |
   - ce
      |
       - sl-5.5-x86_64-grid.ce
                        |
                         - 1.0
                            |
                             - sl-5.5-x86_64-grid.ce-1.0.img
```

Although initial tests using Maven have proven unsuccessful with the large file sizes of the currently available appliances, use of this technology is desirable, and further testing is planned.

6 Related Activities

In this section, we describe relevant European activities which are good candidates for collaboration with StratusLab. They are also good candidates for being early adapters of the StratusLab distribution, since they also deal with grid-over-cloud provisioning.

6.1 CERN

CERN is experimenting with various VM provisioning systems in order to build its own private cloud. In the *lxcloud* [10] project, OpenNebula (ONE), along with various other commercial solutions are used to manage a large number of hardware resources. The high level architecture of *lxcloud* is presented in Figure 6.1. In the current setup, ONE is installed on 500 servers, and each server has 8 CPUs, 24GB of RAM and two 500GB hard disks. Concerning scalability, ONE was able to easily manage up to 8.000 concurrent virtual machine instances. What is more, ONE scheduler was able to dispatch virtual machines at a high rate of around 1 VM per second (machines were pre-staged in the file systems of the host machines). Currently, *lxcloud* virtual machine images are replicas of *lxbatch* worker nodes (batch cluster at CERN). Each dispatched virtual machine joins the batch cluster and accepts a part of the total workload. According to the size of the batch queue, virtual machines are initiated or destroyed.

Since *lxcloud* is also using the ONE platform, it could benefit from the automatic creation and maintenance of grid sites that StratusLab offers. Therefore, the *lxcloud* project is a perfect candidate for collaboration with StratusLab.

6.2 BiG Grid

The BiG Grid project [1] (led by NCF, Nikhef and NBIC) aims to set up a grid infrastructure for scientific research. This research infrastructure contains compute clusters, data storage, combined with specific middleware and software to enable research which needs more than just raw computing power or data storage.

The BiG Grid project, through its Virtual Machine working group, is investigating the use of fully virtualized grid sites. In a recently published report [2], the results of the group's feasibility study to provide virtualized worker nodes on BiG Grid resources are briefly presented. After a thorough examination through numerous use cases such as High Energy Physics or Bioinformatics applications, the

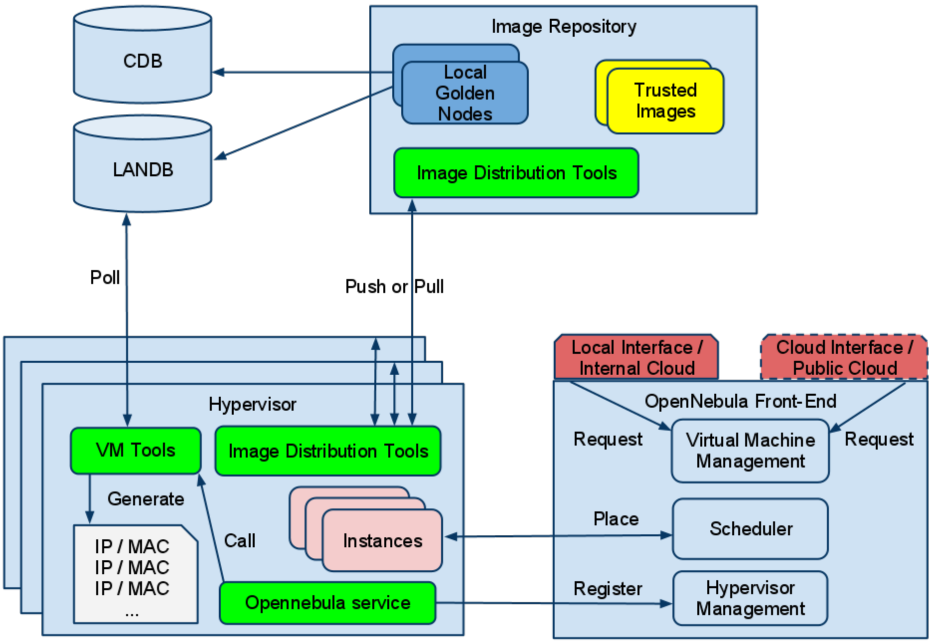

Figure 6.1: *High level architecture of lxcloud*

working group concludes that OpenNebula matches best with the BiG Grid requirements. Therefore, the group decided to utilize OpenNebula as their VM scheduler. What is more, the group has decided to create a number of "precooked" virtual machine images that could be used to automatically create a grid site. Therefore, the BiG Grid project could benefit from the StratusLab results that have to do with the automatic setup, deployment and configuration of grid sites through the use of OpenNebula technologies.

6.3 D-Grid

The D-Grid Resource Center Ruhr (DGRZR) (http://gridatdortmund.blogspot.com/) was established in 2008 at Dortmund University of Technology as part of the German Grid initiative D-Grid (http://www.d-grid.de/index.php?id=1&L=1). Different from other grid sites, the DGRZR has utilized virtualized resources from the beginning of its operation.

DGRZR consists of 256 HP blade servers with eight CPU cores (2,048 cores in total) and 1 GB RAM each. The disk space per server is about 150 Gigabytes. Half of this space is reserved for virtual machine images. Physical servers run SUSE Enterprise Linux (SLES) 10 Service Pack 3 and will be changed to SLES 11 in the near future. D-Grid users are provided with roughly 100 terabytes of central storage, mainly for home directories, experiment software and for the dCache grid

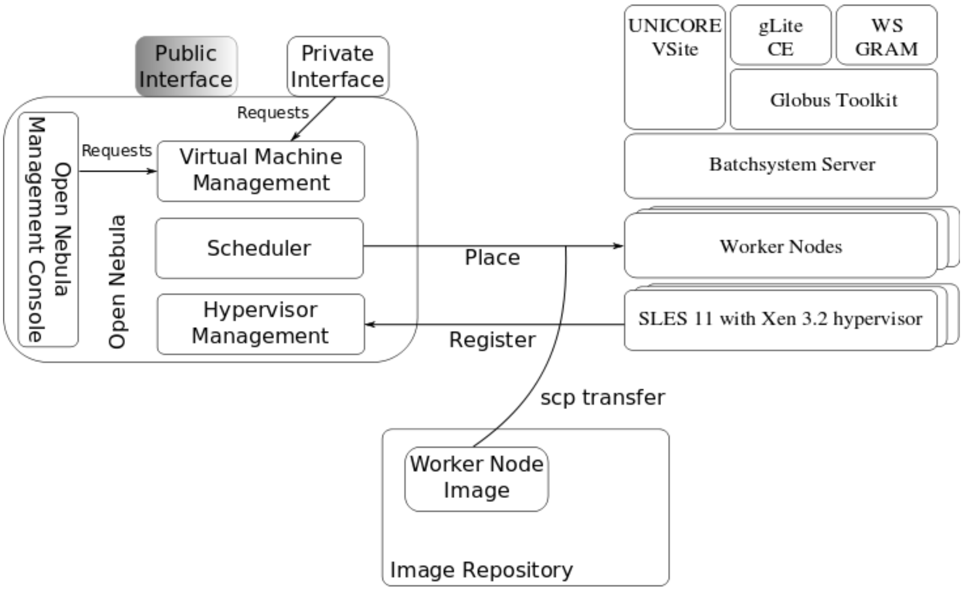

Figure 6.2: *OpenNebula at D-Grid Resource Center Ruhr*

Storage Element. In 2009, the mass storage was upgraded by adding 25 TB of HP Scalable File Share 3.1 (a Lustre-like file system). Typically, 250 of the 256 blade servers will typically be running virtual worker nodes. The remaining servers run virtual machines for the grid middleware services (gLite, Globus Toolkit and UNICORE), the batch system server, and other management services.

In 2010, DGRZR was extended by the installation of OpenNebula as the cloud middleware to manage its virtual machines as a private cloud [12]. In Figure 6.2 the DGRZR OpenNebula architecture is summarized. StratusLab could collaborate with DGRZR and provide its technical knowledge in building fully automated grid sites using virtualization technologies.

6.4 INFN Worker Nodes on Demand Service

The INFN (National Institute of Nuclear Physics - http://www.infn.it/indexen.php) is an organization dedicated to the study of the fundamental constituents of matter, and conducts theoretical and experimental research in the fields of subnuclear, nuclear, and astroparticle physics.

INFN is currently developing a service called WNoDeS (Worker Nodes on Demand Service) [3] in order to provide on demand grid worker nodes through a cloud computing management tool. The overall architecture of WNoDeS is presented in Figure 6.3. WNoDeS is build around a tight integration with LSF [14] and is running in production at the INFN Tier-1 Computing Center. It currently hosts 4,000 on-demand Virtual Machines, O(10) supported Virtual Images, serving 20 different user communities; on average, more than 20,000 jobs are executed each day through WNoDeS. Its main characteristics are:

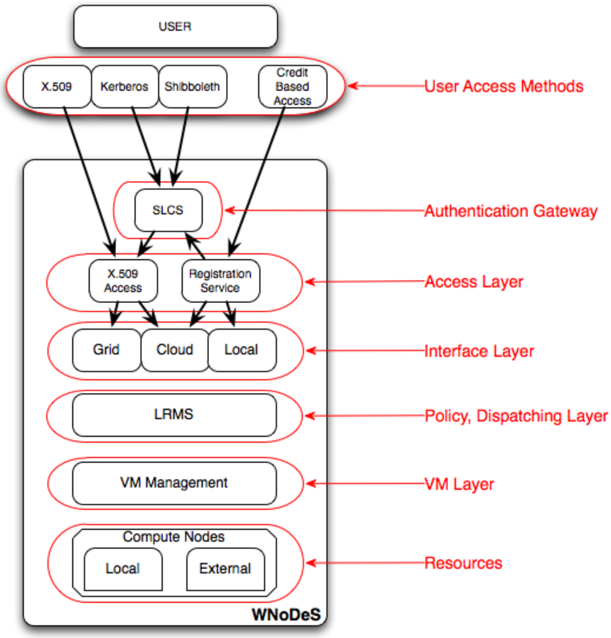

Figure 6.3: *WNoDeS overall architectural framework*

- Full integration with existing computing resource scheduling, policing, monitoring and accounting workflows.
- On-demand virtual resource provisioning and VLAN support to dynamically isolate Virtual Machines depending on service type / customer requests.
- Support for users to select and access WNoDeS-based resources through grid, cloud interfaces, or also through direct job submissions.

INFN's WNoDeS project is also a good candidate for collaboration with StratusLab, as they both deal with grid-over-cloud technologies.

6.5 Hungarian National Initiative in NIIF

NIIF (http://www.niif.hu) is the Hungarian NRN. In the context of a national cloud project, NIIF has investigated the capabilities of deploying large-scale cloud infrastructures using OpenNebula middleware. At the end of this one year project, NIIF achieved to deliver some very interesting results especially related with the storage management over clouds. During this process NIIF also contributed to the OpenNebula development in the above area of storage management. StratusLab has already made contacts with NIIF with the potential to reuse the findings for the above mentioned project and generally collaborate in various levels.

7 Conclusions

The StratusLab computing infrastructure is an important tool in order to successfully pursue the various goals of the project. During the first months of execution the project managed to deploy a significant number of computing resources committed by two partners (CNRS an GRNET). A third partner (TCD) has allocated storage resources and is providing a first version of the appliance repository, hosting virtual machine images to be used in the project private cloud. Table 7.1 summarizes the resources currently committed for the infrastructure.

This infrastructure has already started delivering results. Two pre-production grid sites have been deployed with support for parallel MPI applications. These sites are used for testing and benchmarking purposes. In various cases, the Quattor toolkit is used to manage the deployment of cloud sites as well as the dynamic management of Worker Nodes (WNs) in grid sites.

The operation of grid sites in production environments are expected to impose new requirements for what concerns monitoring, management and accounting tools as well as procedures for the provision of such services. The project is already considering these implications and will further elaborate on them in the context of Deliverable D5.2 planned for month M6 of the project.

StratusLab is wishing to capitalize on the work done in related projects and activities in the area of virtualized grid services and the utilization of cloud computing for hosting production level grid sites. The project has a already identified a number of significant activities in this area, and in some cases has made first contacts seeking further collaboration. More results are expected to come available towards this direction during the lifetime of the project and as the relevant requirements are slowly becoming clearer and concrete.

Table 7.1: *Committed resources per partner*

Resource	GRNET	CNRS	Total
Computing nodes	26	5	31
CPU Cores	208	40	248
Storage	20 TB	5-10 TB	25-30 TB

Glossary

Appliance	Virtual machine containing preconfigured software or services
Appliance Repository	Repository of existing appliances
DCI	Distributed Computing Infrastructure
EGEE	Enabling Grids for E-sciencE
EGI	European Grid Infrastructure
EGI-TF	EGI Technical Forum
FC	Fiber Channel
Front-End	OpenNebula server machine, which hosts the VM manager
GPFS	General Parallel File System by IBM
Gb	Gigabit
Gb/s	Gigabit per second
GB	Gigabyte(s)
GbE	Gigabit Ethernet
Hybrid Cloud	Cloud infrastructure that federates resources between organizations
Instance	see Virtual Machine / VM
IPMI	Intelligent Platform Management Interface
iSGTW	International Science Grid This Week
LAN	Local Area Network
Machine Image	Virtual machine file and metadata providing the source for Virtual Images or Instances
NFS	Network File System
NGI	National Grid Initiative
Node	Physical host on which VMs are instantiated
OS	Operating System
Private Cloud	Cloud infrastructure accessible only to the provider's users
Public Cloud	Cloud infrastructure accessible to people outside of the provider's organization
Regression	Features previously working which breaks in a new release of the software containing this feature
SAN	Storage Area Network
TB	Terabyte(s)
Virtual Machine / VM	Running and virtualized operating system
VLAN	Virtual Local Area Network
VM	Virtual Machine

VO	Virtual Organization
VOBOX	Grid element that permits VO-specific service to run at a resource center
WAN	Wide Area Network
Web Monitor	Web application providing basic monitoring of a single StratusLab installation
WebDAV	Web-based Distributed Authoring and Versioning
Worker Node	Grid node on which jobs are executed

References

[1] BiG Grid. the dutch e-science grid. http://www.biggrid.nl/.

[2] BiG Grid Virtual Machine working group progress report. Virtualization of worker nodes. https://wiki.nbic.nl/images/f/f6/ProgressReport-1.0.pdf.

[3] Davide Salomoni (INFN, Italy). Gridification of a cloud infrastructure through virtualization technologies and related security challenges. http://www.ogf.org/OGF28/materials/1994/salomoni_ogf28_100316.pdf.

[4] EGI. European Grid Initiative. http://www.egi.eu.

[5] gLite. Lightweight Middleware for Grid Computing. http://glite.web.cern.ch.

[6] Haxx. cURL. http://curl.haxx.se/.

[7] High Performance Computing Center Stuttgart (HLRS). MPI-Start. http://www.hlrs.de/organization/av/amt/research/mpi-start.

[8] KVM Community. KVM: Kernel-based Virtual Machine. http://www.linux-kvm.org/.

[9] Nagios. The Industry Standard In Open Source Monitoring. http://www.nagios.org.

[10] OpenNebula Project. Details of CERNs OpenNebula deployment. http://blog.opennebula.org/?p=620.

[11] OpenNebula Project. OpenNebula. http://opennebula.org/.

[12] OpenNebula Project. OpenNebula at the D-Grid Resource Center Ruhr. http://blog.opennebula.org/?p=664.

[13] Planet Lab. Ganglia Monitoring System. http://ganglia.info/.

[14] Platform. Load Sharing Facility (LSF). http://www.platform.com/workload-management/high-performance-computing/.

[15] Quattor Community. Quattor Toolkit. http://quattor.org/.

[16] RENATER. RENATER: Le Réseau National de télécommunications pour la Technologie l'Enseignement et la Recherche. http://www.renater.fr/.