Grant agreement no. 211714

neuGRID

A GRID-BASED e-INFRASTRUCTURE FOR DATA ARCHIVING/ COMMUNICATION AND COMPUTATIONALLY INTENSIVE APPLICATIONS IN THE MEDICAL SCIENCES

Combination of Collaborative Project and Coordination and Support Action

Objective INFRA-2007-1.2.2 - Deployment of e-Infrastructures for scientific communities

Deliverable reference number and title: **D8.3 - Phase2 & 3, Updated Deployment Test & Validation Report**
Due date of deliverable: **Month 36**

Actual submission date: **February 8th, 2011**

Start date of project: **February 1st 2008** Duration: **36 months**

Organisation name of lead contractor for this deliverable: **maat Gknowledge**

Revision: Version **1**

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Table of Contents

Glossary

Term	Definition
AMGA	Metadata Catalog Service of gLite UI
GCC	Grid Coordination Center. Core (common to all sites) services of the grid infrastructure
DCC	Data Coordination Center. Core (common to all sites) services of the database infrastructure
DACS	Data Archiving and Computational Site. neuGRID site offering and managering a set of physical resources
DCS	Data Collection Site. End-user sites acquiring data and connecting to a given DACS
Gridification	The engineering process of porting an existing application to the grid, so that it can be executed via the grid enactment environment
Pipeline	A pipeline is a set of data processing elements connected in series, so that the output of one element is the input of the next one (extracted from Wikipedia.org)
CE	Computing ElementDisk Pool Manager
DPM	Disk Pool Manager
LFC:	LCG File Catalog
SE	Storage Element
CA	Certificate Authority
sBDII	Site level Berkeley Data Base Information Index
tBDII	Top Level Berkeley Data Base Information Index
VOMS	Virtual Organization Membership Service
UI	gLite User Interface
WN	Worker Node
WMS	Workload Management System

1. Introduction

This document attempts to describe all of the tasks related with the deployment of the neuGRID infrastructure. The document covers all aspects, including the definition, profiling and hardware specification, the selection of the technologies that will be used as infrastructure, the deployment logistics preparation and the tests that will be performed over the hardware (see work package 11 – WP11 – for more information) to ensure that the site deployed will fit into the neuGRID infrastructure.

This first version of the document, covers the tasks T8.4 "Phase 2 – 2nd DACS Establishment" and T8.5 "Phase 3– 3rd DACS Establishment" This document also includes the information delivered in the D8.1 "Ground Truth & Phase1, Deployment Test & Validation Report" as it is the continuation of the work developed by WP8 in the first 24 months of the project.

This document was updated throughout the duration of all the WP8 activity. The content of the deliverable has two well distinguished parts.

The aim of the first part of the deliverable is to define all the requirements needed in order to successfully deploy a neuGRID site in the institutions (hospitals and research centers). This includes a specific hardware portfolio of specifications and the preparation of the logistics. It also includes the justification of the selected technologies by means of an analysis of the gLite technology that will be used as GRID middleware and its implications in a large project such as neuGRID.

The second part of the deliverable is the summary of the deployments performed during the reporting period. This summary includes a detailed description of the sites, including physical and logical descriptions of the services deployed. It also includes the results of the hardware tests (provided by WP11) that have been run using the deployed hardware.

2.1. Purpose of the Document

This document aims to illustrate the main objectives of all project; these objectives are detailed in the following list (extracted from the DoW):

- *To define the hardware specifications for the different sites and levels of the infrastructure.*
- *To design a series of basic hardware and software tests for validating hardware and proper connectivity performance.*
- *To deploy the infrastructure at the different institutions of the project. This includes the grid middleware and project prototype, as well as the establishment of appropriate and efficient logistic support.*

2.2. Document Positioning and Intended Audience

The deliverable D8.3 – "**Ground Truth, Phase1, 2 & 3 Large Scale Test Report**" aims to serve as a reference guide for the deployment of the sites that will be part of the neuGRID infrastructure.

By the definition of the recommended hardware portfolio, the specification of all the technologies involved in the relying infrastructure and the technological constraints, WP8 wants to help new sites to select appropriate hardware in order to successfully deploy a neuGRID Site.

3. Hardware Portfolio

This section will explain the spreadsheet that WP8 has developed (See Annex 1 – Hardware Portfolio) in order to help partners (hospitals and institutions) select the most suitable configuration for each one.

The spreadsheet presents the developed configurations in 3 different versions: Basic, Advanced and Premium. These configurations range from basic hardware specifications to the most advanced and powerful configurations.

In the spreadsheet, the following parameters for each configuration are summarized:

- Price
- Number of CPU
- Raw Disk Space
- Memory
- Power Consumption
- Heat Dissipation

WP8 has developed 2 sets of configurations. The first set is directed towards hospitals, with the main configurations having been calculated to fit within the 30.000 € available budget for this kind of institution. The second set has been designed to fit within the 10.000 € budget available to the technical partners.

First, the hospital configurations will be described, with a presentation of 5 different hardware configurations that attempt to cover all of the needs and restrictions of the IT centers at the hospitals that will be hosting the neuGRID infrastructure. In these cases the best option is to choose the Advanced configuration. Following is a description of the hardware configurations detailed in the portfolio:

- Blade Servers + Dell PowerEdge 2950: The main advantages of this configuration are moderate physical space, power consumption and heat dissipation. By using 2 DPM, running in 2 Dell PowerEdge 2950 will allow the infrastructure to manage some redundancy.
- Blade Servers + Dell PowerEdge 2900: As in the previous configuration, the use of blade servers provides moderate physical space, power consumption and heat dissipation. In this case, the use of only 1 Dell PowerEdge 2900 removes the redundancy but increases raw Disk Space.
- Blade Servers + M300 iSCSI: This is the most powerful configuration, maintaining moderate physical space, power consumption and heat dissipation. The use of an iSCSI device allows the upgrade of the DPM, should this be necessary in the future. On the other hand, this solution is the least cost effective.
- Dell PowerEdge 2950 + Dell PowerEdge 1950: This is a good solution if only taking the Power (CPU and storage) to Cost ratio into account. The main problems of this configuration are the space used, the power consumption and the heat dissipation. By using 2 DPM running in 2 Dell PowerEdge 2950, the infrastructure is able to manage some redundancy.
- Dell PowerEdge 2900 + Dell PowerEdge 1950: This is a good solution when looking at the Power (CPU and storage) to Cost ratio. The main problems of this configuration are the space used, the power consumption and the heat dissipation. In this case, the use of only 1 Dell PowerEdge 2900 removes the redundancy but increases raw Disk Space.

Secondly, the configuration for technical partners will be shown:

- Dell PowerEdge 2950 + Dell PowerEdge 1950: This is the best solution for technological partners, offering flexibility and a good ratio between price and raw disk space. In this case, the best option is to choose the premium option. This option offers the best ratio between cost and power and also fitting within the 10.000 € budget.

In the "Annex 2 – neuGRID Hardware Template Tests", we can find a set of tests that all of the servers must pass in order to ensure correct operation.

4. Technologies

One of the main challenges that WP8 has faced, and continues to face, is the shifting of needs from a development environment to a real production environment. Working within the constraints that this places on the architecture of system deployment, the solution must provide a high level of security, stability, and cost effectiveness. These requirements have therefore motivated the WP8 partners to look into solutions that enable, in a realistic way, the deployment of a neuGRID site in complex environments.

An additional important technological challenge that WP8 has been facing is the use of the gLite grid technology. gLite is a grid middleware software stack from the European project "Enabling the Grid for e-sciEnces (EGEE)."

As stated on the EGEE official website "gLite provides a bleeding-edge, best-of-breed framework for building grid applications tapping into the power of distributed computing and storage resources across the Internet." gLite was conceived as a large-scale project for providing grid capabilities to sites using a large number of computers. The architecture of gLite is complex and includes services including security, user interfaces, computing and storage elements, information systems and data and workload management.

For the sake of exemplifying gLite's usage and scalability, it is currently used in production in the WLCG/EGEE infrastructure[1]. This infrastructure operates a grid distributed over more than 200 sites around the world, with more than 30,000 CPUs and 20 PB of data storage. These numbers can give the reader a better idea of the capabilities and sheer complexity of gLite.

While a single gLite service can be deployed in a single operating system, a site containing multiple grid services cannot be deployed in a single operating system due to its complex architecture. On the other hand, the use of one physical machine for each one of the grid services would clearly not be acceptable in the context of neuGRID.

Thus, this raised a deployment issue which had to be solved in order to provide a reasonable and successful platform deployment that could cope with the neuGRID environment. To address this situation, WP8 has been using virtualization technologies to run several instances of an operating system in one physical computer. This allowed a drastic reduction in the hardware requirements while not impacting the resulting quality of the services.

This is not the only reason to use virtualization technologies in the environments; the use of virtualization has more advantages than the example of running more than one guest OS in one host OS. These advantages are:

Independence of the hardware: The images of the guest OS can run in any compatible server. This feature of virtualization ensures the correct operation of the services deployed, without affecting the relying hardware.

[1] http://glite.web.cern.ch/glite/

Cost-effective hardware solution: It is a good solution for obtaining increased server productivity, running services that are not CPU-intensive consumers. In the case of neuGRID, this fits perfectly within the philosophy of the gLite Services.

Fast deployment of preconfigured images: Since the hardware layer at guest OS level is the same for all the servers, the building of preconfigured OS images is simplified.

In the remainder of this chapter, the virtualization paradigm and associated technologies are presented, and an evaluation is carried out to highlight the design decision(s) made in WP8.

4.1. Virtualization Technologies Introduction

Virtualization is performed on a given hardware platform using a so-called "host" software (a control program), which creates a simulated computer environment (a virtual machine) for its "guest" software. The "guest" software, which is often itself a complete operating system, runs just as though it were installed on a stand-alone hardware platform. Typically, many such virtual machines are simulated on a given physical machine. For the "guest" system to operate properly, the simulation must be robust enough and must support all the guest system's external interfaces, which (depending on the type of virtualization) may include hardware drivers.

The type of virtualization that WP8 has been and will be used is the so-called paravirtualization[2]. This is a virtualization technique that presents a software interface to virtual machines that is similar (but not identical) to that of the underlying hardware [BAR03]. This requires operating systems to be explicitly ported to run on top of the Virtual Machine Monitor (VMM), which the owner of exclusive rights in a proprietary operating system may decline to allow for strategic reasons, but which may enable the VMM itself to be simpler or virtual machines that run on it to obtain performances closer to non-virtualized hardware.

4.2. Virtualization Software Overview

This section contains a (non-exhaustive) overview of the current virtualization techniques and corresponding implementations. In accordance with existing legacy IT assets (e.g. gLite grid middleware) and other constraints that WP8 had to respect (e.g. budget limitation for hardware purchases), several criteria have been formulated for selecting an appropriate virtualization technology. These criteria are the following:

Support virtualization of Scientific Linux CERN 3 and Scientific Linux CERN 4, this constraint is imposed[3] by the gLite grid middleware, perform reasonably well when running a virtualized Operating Systems, common requirement, allow the use of SMP Kernels, all server computers use multi-processors architectures nowadays, run on i686 and x86_64 Architectures with the ability of mixing 32 and 64 bits Kernels for host Oss. 64 bits and the various possible mixed modes have to be supported for allowing future use of 64 bits-based software to be distributed under a GPL License.

[2] http://www.vmware.com/pdf/virtualization_considerations.pdf

[3] gLite is heavily reliant on the host OS. The packages for installing gLite 3.1 version are only certified to work on Scientific Linux 4.0; this is also the only operating system that is certified to properly run gLite.

4.2.1. Xen

Xen[4] is free software VMM for IA-32, x86-64, IA-64 and PowerPC architectures. It is software which runs in a host operating system and which allows several guest operating systems to be installed and used on top of, and at the same the time as, the host on the same computer. Xen originated as a research project at the University of Cambridge, led by Ian Pratt, senior lecturer at Cambridge and founder of XenSource, Inc. This company now supports the development of the open source project and also sells enterprise versions of the software. The first public release of Xen was made available in 2003.

4.2.2. Kernel based Virtual Machine

Kernel-based Virtual Machine (KVM)[5] is a Linux kernel infrastructure for supporting virtualization. KVM currently supports full virtualization using Intel VT or AMD-V. Limited support for paravirtualization is also available for Linux guests and Windows in the form of a paravirtual network driver, a balloon driver to affect the operation of the guest virtual memory manager, and CPU optimization for Linux guests. KVM is currently implemented as a loadable kernel module although future versions will likely use a system call interface and be integrated directly into the kernel. Architecture ports are currently being developed for s390, PowerPC, and IA64. The first version of KVM was included in the Linux kernel version 2.6.20.

By itself, KVM does not perform any emulation. Instead, a user-space program uses the /dev/kvm interface to set up the guest VM's address space, feeds it simulated I/O and maps its video display back onto the host. Currently, the only known program that does this is a modified version of QEMU[9]. KVM's components are available under various GNU licenses.

4.2.3. OpenVZ

OpenVZ[6] is an operating system-level virtualization technology based on the Linux kernel and operating system. OpenVZ allows a physical server to run multiple isolated operating system instances, known as Virtual Private Servers (VPS) or Virtual Environments (VE).

Compared to virtual machines such as VMware and paravirtualization technologies like Xen, OpenVZ is limited as it requires both the host and guest OS to be Linux-based (although Linux distributions can be different in different VEs). However, OpenVZ claims a performance advantage; according to its website, there is only a 1-3% performance penalty for OpenVZ as compared to using a standalone server.

OpenVZ is a basis of Virtuozzo[7], a proprietary software product provided by SWsoft[8], Inc. OpenVZ is licensed under the GPL version 2.

---

[4] http://www.xensource.com

[5] kvm.qumranet.com

[6] http://openvz.org/

[7] http://www.swsoft.com/en/products/virtuozzo

[8] http://www.swsoft.com

4.2.4. QEMU

QEMU[9] is a fast processor emulator, allowing full virtualization of a PC system. It is free software and was written by Fabrice Bellard. QEMU is a hypervisor and is similar to projects such as Bochs, VMware Workstation and PearPC, but lacks several features, including increased speed on x86 architectures, although achievable through an optional accelerator KQEMU[10], and support for multiple architectures (work in-progress at the time of writing). By using dynamic translation it performs reasonably well, and turns out to be easy to port to new host CPUs.

4.2.5. Linux-VServer

Linux-VServer[11] is a virtual private server implementation which achieves virtualization by adding operating system-level virtualization capabilities to the Linux kernel. Linux-VServer is a jail mechanism which can be used to securely partition resources on a computer system (such as the file system, CPU time, network addresses and memory) in such a way that processes cannot mount a denial-of-service attack on anything outside of their partition. Each partition is called a security context, and the virtualized system within it is the virtual private server. A chroot-like utility for entering security contexts is provided. The contexts themselves are robust enough to boot many unmodified Linux distributions, including Debian and Fedora Core.

Virtual private servers are commonly used in web hosting services, where they are useful for segregating customer accounts, pooling resources and containing any potential security breaches. Conceptually Linux-VServer is similar to the Solaris Containers (including Solaris Zones isolation technology), or FreeBSD Jail, or OpenVZ.

Linux-VServer is developed and distributed as open source software, licensed under the terms of the GNU General Public License (GPL).

4.2.6. VMware

VMware[12] is a wholly-owned subsidiary of the EMC Corporation which supplies proprietary virtualization software for x86-compatible computers, including VMware Workstation and the freeware VMware Server and VMware Player products. VMware software runs on Microsoft Windows, Linux, and Mac OS X.

There are several flavors of VMware, depending on the needs of the client. VMware is offering a large variety of products, including free and non-free products licensed under a commercial license. VMware is not licensing its products under an open source license.

---

[9] http://fabrice.bellard.free.fr/qemu

[10] http://fabrice.bellard.free.fr/qemu/kqemu-tech.html

[11] http://linux-vserver.org/

[12] http://www.vmware.com/

Free products:

VMware Player: VMware makes VMware Player available, free of charge, to run guest virtual machines produced by other VMware products. VMware Player itself is not able to create new virtual machines.

VMware Server: VMware Server can create, edit, and play virtual machines. It uses a client-server model, allowing remote access to virtual machines, at the cost of some graphical performance. VMware Inc. makes VMware Server freely available in the hope that users will eventually upgrade to the VMware ESX Server.

Non-Free products:

VMware Workstation: VMware Workstation software consists of a virtual-machine suite for x86 and x86-64 computers. This software suite allows users to set up multiple x86 and x86-64 virtual computers and to use one or more of these virtual machines simultaneously with the hosting operating system. Each virtual machine instance can execute its own guest operating system. In simple terms, VMware Workstation allows one physical machine to run two or more operating systems simultaneously.

VMware ESX: This is the most professional solution. VMware ESX implements what is referred to as the "VMkernel", which is a bundle of hypervisor codes along with the device driver modules used to support a given set of hardware.

VMware ESX Server uses a Linux kernel that loads additional code. It is referred by VMware, as the "VMkernel". The VMware FAQ states 'ESX Server also incorporates a service console based on a Linux 2.4 kernel that is used to boot the ESX Server virtualization layer'. The Linux kernel runs before any other software on an ESX host, witness the console of a booting ESX machine. After the Linux kernel has loaded, there is a script that loads the VMkernel. VMware states that VMkernel is not derived from Linux. The Linux kernel continues running but under VMkernel, providing functions including the proc file system used by the ESX and an environment to run support applications.

4.3. Conclusions

These virtualization techniques and technologies have been evaluated and, when possible, tested. In the following table, their most interesting characteristics are reflected, together with the features that were retained for making further decisions regarding upon which solution to base the neuGRID deployment. As formerly stated, the following properties have been partially extracted from the virtualization technologies state-of-the-art literature.

Properties	Xen	KVM	QEMU	OpenVZ	VServer	VMWare
Full VT	Yes	Yes	Yes	No	No	Yes
Paravirtualization	Yes	Yes	No	No	No	Yes
OS Level VT	No	No	No	Yes	Yes	No
License	GPL	GPL	GPL BSD Core: LGPL	GPL	GPL	Proprietary
Architecture	i686 x86-64 IA64 PPC	i686 x86-64	i686 ARM	i686 x86-64 PPC	i686 ARM	i686 x86-64
Performance	PV: Fast FV: Medium	PV: Fast FV: Medium	FV: Slow Best if kQEMU enabled	Fast	Fast	PV: Medium FV: Medium
SMP Ghuests	Yes	Development	Development	Yes	Yes	Yes
CPU Hot Plug	Yes	No	No	Yes	Yes	Yes
Mem Hot Plug	Yes	No	No	Yes	Yes	Yes
Standalone Host	Yes	No	No	No	No	No
Standalone Hypervisor	No	No	No	No	No	Only with ESX Server

From this table, it can be noted that some of the virtualization techniques do not fulfill the main WP8 criteria, exposed in the Point 2.2. For instance, QEMU and Vserver cannot run, at the time of writing, x86_64 hosts, and VMware is distributed under a commercial license, which turns out to be quite expensive and inappropriate given the needs of our project.

WP8 has been in direct contact with the gLite development team at CERN in order to obtain additional expert advice regarding virtualization solutions and their potential current use in other projects. From this feedback, the formerly introduced evaluation and the experience of the partners using gLite, WP8 therefore decided to adopt Xen.

Xen has been extensively tested and appears to work properly with the gLite grid middleware. It allows the running of a large variety of Linux Kernels, such as i686, i686 PAE and x86_64 as host and guest kernels.

Nevertheless, WP8 is still investigating other technologies such as OpenVZ, in the spirit of comparing performance and stability aspects, for a potential future migration in case significant improvement in quality of services can be obtained.

5. Data Coordination Centre

In the DoW we can find the DCC defined as: "*DCC takes care of coordinating and maintaining the different DACS centres from level 1. TheDCC's primary functions early in the neuGRID development phase are the development, deployment and maintenance of the user-facing services, including the development of the neuGRID access portal. Towards the later phases of the project these functions will shift towards oversight and more operational responsibilities, where the DCC will coordinate operations such as standardization of acquisition protocols; development of quality control procedures; monitoring of data consistency; ontological mapping of existing databases; use, performance and validation of image analysis algorithms; use of statistical analyses procedures, etc. Many of these functions will be implemented by consensus-building among the partners as well as the user communities. The DCC will also participate in the grid as a level 1 node, providing storage space and CPU resources to the grid*"

This site was deployed during month 8 of the project, in parallel to the GCC. The current activity in the DCC is a series of developments concerning the LORIS database software, which is being integrated in the grid and in the neuGRID services.

6. Grid Coordination Centre

In the DoW we can find the GCC defined as: *"GCC is in charge of hosting, maintaining up and running the grid middleware information system services 24/7. These services are the cornerstones of the grid. They provide the inner mechanics of the distributed infrastructure. The GCC will be installed as part of the Ground-Truth phase of the neuGRID Infrastructure deployment, and once available will host its newly created virtual organization"*.

The GCC was successfully deployed at month 8 at Archamps, France within the servers provided by maat Gknowledge.

The GCC will host the entire infrastructure to run the gLite GRID middleware. This includes gLite core services, a site performing CE and SE, and the necessary infrastructure to provide servers and users with the digital certificates, which will allow operating against the neuGRID infrastructure.

- gLite Core services, including: VOMS, WMS/LB, LFC. tBDII and AMGA
- Site level services, including: UI, CE, SE and sBDII
- neuGRID UI
- CA

6.1. Services

To run a new gLite middleware infrastructure it is necessary to install some gLite services. The basic infrastructure that will allow neuGRID to provide grid services is the following:

- VOMS: Virtual Organization Membership System. This service will manage the users, the resources and the possible interactions between them.
- LFC: LCG File Catalogue, this service will store a global index of files to allow the grid users to find these files stored in the grid.
- tBDII: Top Level Berkeley Database Information Index, this service will read the configuration of the different grid sites, and will present it with an LDAP interface.

- AMGA: It is a metadata catalog for gLite.
- WMS/LB: gLite Workload Management System

These services have been installed in one DELL PowerEdge 2950; virtualization techniques have been used in this hardware in order to install 4 distinct services in one server. This technique is commonly used when deploying gLite. As explained in Section 2, gLite has significant constraints regarding the mixing of different services in one OS.

The complete gLite infrastructure security model relies on digital certificates. This means that to operate with the grid, each user and server will need a digital certificate ensuring that the actor is who he or she is supposed to be. With this purpose in mind, the neuGRID project has deployed a local Certification Authority, which will be able to issue certificates to operate with neuGRID infrastructure.

OpenCA software has been installed in a virtual machine. OpenCA provides a complete suite of Certification Authority services. This suite enables the management of the certificate life cycle of the users and servers.

An SE and a DPM were also deployed as a CE and WN.

The Computing Element is composed by 3 DELL Power Edge 1950 servers, with the specifications outlined in **Error! Reference source not found.**. These servers will be part of a gLite Computing Element, by running Worker Node software in each one. The main function of these servers is to execute the jobs that are submitted to the GRID infrastructure.

The Storage Element was deployed in a DELL PowerEdge 2950.

6.2. Services Configuration

6.2.1. Hardware Distribution

The following table shows the dom0 servers list, in addition to the management IP that will be performed over a Dell Remote Administration Card, and its Service TAG (Serial Number of the Server).

Server Name	Dom0 IP	Management IP	Service TAG
dell-1950-1	80.245.17.57	80.245.17.40	CKWZM3J
dell-1950-2	80.245.17.48	80.245.17.41	BKWZM3J
dell-1950-3	80.245.17.61	80.245.17.42	8KWZM3J
dell-2950-1	80.245.17.50	80.245.17.38	6310N3J
dell-2950-2	80.245.17.44	80.245.17.43	9310N3J

6.2.2. gLite Services Distribution

The following table presents a summary of the gLite middleware services deployed at GCC, including the reference to the physical machine that is hosting the guest virtual machine, which is running the gLite middleware service identified in the last column of the table.

Server Name	External IP	FQDN	gLite
dell-2950-2	80.245.17.45	voms.maatg.eu	VOMS
dell-2950-2	80.245.17.49	lfc.maatg.eu	LFC
dell-2950-2	80.245.17.67	bdii.maatg.eu	tBDII
dell-2950-2	80.245.17.51	ng-maat-server5.maatg.eu	amga
dell-2950-2	80.245.17.52	ng-maat-server6.maatg.eu	gLiteUI
dell-2950-1	80.245.17.47	ng-maat-server3.maatg.eu	DPM
dell-2950-1	80.245.17.68	bdii-site.maatg.eu	sBDII
dell-2950-1	80.245.17.58	ng-maat-devel1.maatg.eu	AMGA gLiteUI
dell-2950-1	80.245.17.59	myproxy. maatg.eu	Myproxy
dell-2950-1	80.245.17.61	ng-maat-server8. maatg.eu	CE
dell-2950-1	80.245.17.62	wms. maatg.eu	WMS LB
dell-1950-1	80.245.17.63	ng-maat-server10. maatg.eu	WN
dell-1950-2	80.245.17.64	ng-maat-server11. maatg.eu	WN
dell-1950-3	80.245.17.65	ng-maat-server12. maatg.eu	WN

6.2.3. Support Services Distribution

In the following table we can see all the data related to the openCA and a support web for deployments and developments that was also deployed in the GCC.

Server Name	External IP	FQDN	Services
dell-2950-2	80.245.17.46	ng-maat-server1.maatg.eu	WEB
dell-2950-2	80.245.17.46	openca.ng-maat-server1.maatg.eu	OPENCA

6.3. Infrastructure

This site will be installed over 5 servers: 2 DELL PowerEdge 2950, and 3 DELL PowerEdge 1950 in the first stage of the deployment; gLite Core Services have been deployed over one of the DELL PowerEdge2950. In the second stage, a Site including a Storage Element has been deployed in the second DELL PowerEdge 2950 and finally, a Computing Element over the 3 remaining DELL PowerEdge 1950 servers remains to be deployed.

Detailed hardware tests for each server can be found in "*Annex 3 – GCC Hardware Tests*" in which we can see that all the servers pass the hardware test suit detailed in the Annex 2.

6.3.1. Servers Configuration

DELL PowerEdge 2950

In **Error! Reference source not found.** we can see the final hardware configuration for the server DELL PowerEdge 2950, corresponding to the advanced configuration of the portfolio of solutions. This server offers a good cost to raw space ratio.

In order to fit the services, 2 of these servers have been acquired. In the first, the Core Services have been deployed, the other one will host SE and the DPM. The second one will host all core services.

	PE2950
1	PE2950 III QUAD-CORE XEON E5410 2.33GHZ/
1	RISER WITH PCI EXPRESS SUPPORT (2X PCIE
1	PE2950 SPANISH RACK POWER CORD
1	PE2950 BEZEL ASSEMBLY
1	8GB 667MHZ FBD (2X4GB DUAL RANK DIMMS)
1	PE2950 III ADDITIONAL QUAD-CORE XEON E54
2	80GB SERIAL ATA2 7.2K 3.5" HD HOT PLUG
4	750GB SERIAL ATAU 7.2K 3.5" ADDITIONAL H
1	PE2950 III - CHASSIS 3.5HDD X6 BACKPLANE
1	PERC 6/I, INTEGRATED CONTROLLER CARD X6
1	8X IDE DVD-ROM DRIVE
1	CD/DVD CABLE
1	PE2950 III - REDUNDANT PSU NO POWER CORD
1	POWER CORD, PDU (RACK)
2	BROADCOM NETXTREME 5721 SINGLE PORT GIGA
1	BROADCOM TCP/IP OFFLOAD ENGINE FUNCTIONA
1	DRAC 5 CARD
1	PE2950 OPEN MANAGE CD + DRIVERS
1	CFI EMEA SERVICE ORDER READY POWEREDGE P
1	PE2950 RAPID/VERSA RACK RAILS
1	PE2950 III - C6, MSSR0/R0, ADD IN PERC 5/
1	POWEREDGE ORDER - SPAIN
1	BASE WARRANTY
1	1YR BASIC WARRANTY - NEXT BUSINESS DAY -
1	3YR BASIC WARRANTY - NEXT BUSINESS DAY

The total power of the DPM is described in the following list:

- 8 cores running at 2.33GHz
- 8 Gigabytes of RAM
- 2 Terabytes of local space in RAID5

DELL PowerEdge 1950

In **Error! Reference source not found.** we can see the final hardware configuration for the server DELL PowerEdge 1950, which corresponds to the advanced configuration of the portfolio of solutions. This server offers a good ratio between effective cost, number of cores and RAM capacity. The full solution consists of 13 servers with these characteristics that are part of the cluster that will run the computing element.

	PE1950
1	PE1950 III QUAD-CORE XEON E5410 2.33GHZ/
1	PE1950 PCIE RISER (2 SLOTS)
1	SPANISH - DOCUMENTATION AND RACK POWER C
1	PE1950 BEZEL ASSEMBLY
1	8GB 667MHZ FBD (2X4GB DUAL RANK DIMMS)

1	PE1950 III - ADDITIONAL QUAD-CORE XEON E
2	73GB SAS (15, 000RPM) 3.5 INCH HARD DRIVE
1	PE1950 III 3.5" HDD SUPPORT CHASSIS
1	PERC 6I INTEGRATED CONTROLLER
1	8X IDE DVD-ROM DRIVE
1	PE1950 III REDUNDANT POWER SUPPLY - NO P
1	POWER CORD, PDU (RACK)
1	BROADCOM TCP/IP OFFLOAD ENGINE FUNCTIONA
1	DRAC 5 CARD
1	NO OPERATING SYSTEM
1	PE1950 OPENMANAGE KIT AND FI DRIVER
1	YOU HAVE CHOSEN NOT TO TAKE THE DELL POW
1	CFI EMEA SERVICE ORDER READY POWEREDGE P
1	SLIDING COMBINATION RAPID/VERSA RAIL WIT
1	PE1950 III - C3, MSSR1, ADD IN PERC 5I/6I
1	POWEREDGE ORDER - SPAIN
1	BASE WARRANTY
1	1YR BASIC WARRANTY - NEXT BUSINESS DAY -
1	3YR BASIC WARRANTY - NEXT BUSINESS DAY

In the following table**Error! Reference source not found.** we can see the partition table common to the 3 servers that comprise the Computing Element. Each one will run one host OS and one guest OS. All of them will run WN gLite software.

dev	**lvm**	**Size**
/dev/sda	boot	100 Mbytes
/dev/sdb	swap	2 GBytes
/dev/sdc	root	10 Gbytes
	wn	66 GBytes
	wn_swp	2 GBytes

The total power of the Computing Element is described in the following list:

- 24 cores running at 2.33GHz
- 24 Gigabytes of RAM
- 198 Gigabytes of local space in RAID1

7. DACS1 - IRCCS Fatebenefratelli (FBF, Brescia - Italy)

IRCSS will contribute to the neuGRID infrastructure with a site composed by a computational node and a data node. These nodes are part of the PROD infrastructure defined in the DoW.

7.1. Services

The site at DACS1 will provide an SE and a DPM, and a CE and WN to the grid infrastructure. This infrastructure will allow users to store and access files, and to execute jobs. This fits with the specifications of the DoW.

Due to the gLite grid middleware architecture, it is necessary to set up a service sBDII that will present the site to the core services of the neuGRID infrastructure.

This site will provide 4,5 TBytes of usable space in 3 different data stores, but the most impressive numbers come from the CE, running 12 WN of 8 cores each one. This infrastructure provides a total of 96 cores running at 2,3GHz and 96 GBytes of RAM.

7.1.1. Services Configuration

In this section we will describe the distribution of the gLite services deployed in the FBF servers. In the followings tables we can see the deployed services in the set of servers. Those can be categorized in two major families: the gLite grid-site related services (UI, DPM, CE, WN) and the neuGRID Gateway. Service TAG and the IP distribution can be also viewed in the following tables.

Each server will have 3 IP addresses, the first IP address is attached to the dom0 (host OS), the second IP address is attached to the dom1 (guest OS), and finally the third IP address is attached to the remote administration card.

Server Name	Dom0 IP	Management IP	Service TAG
ng-fbf-dell-1	192.168.1.227	192.168.1.200	GPL2W3J
ng-fbf-dell-2	192.168.1.228	192.168.1.201	JFL3W3J
ng-fbf-dell-3	192.168.1.229	192.168.1.202	BFL3W3J
ng-fbf-dell-4	192.168.1.230	192.168.1.203	CFL3W3J
ng-fbf-dell-5	192.168.1.231	192.168.1.204	DFL3W3J
ng-fbf-dell-6	192.168.1.232	192.168.1.205	GFL3W3J
ng-fbf-dell-7	192.168.1.233	192.168.1.206	HFL3W3J
ng-fbf-dell-8	192.168.1.234	192.168.1.207	1GL3W3J
ng-fbf-dell-9	192.168.1.235	192.168.1.208	9FL3W3J
ng-fbf-dell-10	192.168.1.236	192.168.1.209	2GL3W3J
ng-fbf-dell-11	192.168.1.237	192.168.1.210	4GL3W3J
ng-fbf-dell-12	192.168.1.238	192.168.1.211	5GL3W3J
ng-fbf-dell-13	192.168.1.239	192.168.1.212	6GL3W3J

Server Name	External IP	FQDN	Internal IP	gLite
ng-fbf-dell-1			192.168.1.214	SE/DPM
ng-fbf-dell-2			192.168.1.215	CE sBDII
ng-fbf-dell-3			192.168.1.216	WN
ng-fbf-dell-4			192.168.1.217	WN
ng-fbf-dell-5			192.168.1.218	WN
ng-fbf-dell-6			192.168.1.219	WN
ng-fbf-dell-7			192.168.1.220	WN
ng-fbf-dell-8			192.168.1.221	WN
ng-fbf-dell-9			192.168.1.222	WN
ng-fbf-dell-10			192.168.1.223	WN
ng-fbf-dell-11			192.168.1.224	WN
ng-fbf-dell-12			192.168.1.225	WN
ng-fbf-dell-13			192.168.1.225	WN

7.2. Infrastructure

This site will be installed over 13 servers: 1 DELL PowerEdge 2900, and 12 DELL PowerEdge 1950 in the first stage of the deployment.

Detailed hardware tests for each server can be found in "Annex 4 – DACS1 Hardware Tests" in which we can see that all the servers pass the hardware test suit detailed in the Annex 2.

7.2.1. Servers Configuration

As previously explained, for the neuGRID deployment, all of the servers have been virtualized with XEN Software; the host Operating System is Scientific Linux 5.2. It runs a XEN Kernel in the top, allowing the server to run paravirtualized Virtual Machines. This technique has some advantages when compared to deploying the OS running gLite services directly in the hardware. These advantages are:

- Fast deployment of preconfigured images in the Guest OS.
- Simple maintenance and upgrades of the Guest OS.
- Simple backup of the whole system.
- Allows the running of more than one Guest OS in one physical server.

DELL PowerEdge 2900

In the next table, we can see the final hardware configuration for the server DELL PowerEdge 2900, corresponding to the advanced configuration of the portfolio of solutions. This server offers a good ratio between cost and raw space.

PE2900	
1	PE2900 III - QUAD-CORE XEON E5410 2.33GHz
1	PE2900 III SERVER RACK CHASSIS
1	RACK BEZEL ASSEMBLY
1	8GB 667MHZ FBD (4X2GB DUAL RANK DIMMS)
1	PE2900 III - ADDITIONAL QUAD-CORE XEON E
1	3.5 INCH 1.44MB FLOPPY DRIVE
8	750GB SERIAL ATAU 7.2K 3.5" HD HOT PLUG
1	PERC 6/I INTERNAL RAID CONTROLLER CARD
2	80GB SATA 7.2K 3.5INCH FRONT BAY HDD
1	16X SATA DVD
1	REDUNDANT POWER SUPPLY (2 HOT PLUG PSU)
1	POWER CORD, PDU (RACK)
1	BROADCOM TCP/IP OFFLOAD ENGINE FUNCTIONA
1	TCP/IP OFFLOAD ENGINE 2P
1	DRAC5 CONTROLLER
1	NO OPERATING SYSTEM
1	OPENMANAGE SERVER SOFTWARE, WITH CD & DO
1	PE2900 UNIVERSAL SLIDING VERSA/RAPID RAC
1	PE2900 III - C11 RAID5/1 PERC5I/6I 3-8HD
1	BASE WARRANTY
1	1YR BASIC WARRANTY - NEXT BUSINESS DAY -
1	3YR BASIC WARRANTY - NEXT BUSINESS DAY

The following table shows the partition table of the DELL PowerEdge 2900, including host OS and the partitions to be installed in the guest OS. This server will run the DPM with 3 data stores, 2 of 1.43 GBytes, and one of 1.2 GBytes, with a total usable space of 4 TBytes as data stored in the GRID. As explained before, the physical support of this space relies on a RAID5 and a spare disk in order to ensure maximum data security and stability.

Dev	lvm	Size
/dev/sda	boot	100 Mbytes
	lvswap	8 GBytes
	lvroot	10 Gbytes
	dpm	24 GBytes
	dpm_swp	2 GBytes
/dev/sdb	dpm_data	1.43 TBytes
/dev/sdc	dpm_data_2	1.43 TBytes
/dev/sdd	dpm_data_3	1.2 TBytes

The total power of the Storage Element is described in the following list:

- 8 cores running at 2.33GHz
- 8 Gigabytes of RAM
- 4500 Gigabytes of shared disk space in RAID5 with 1 spare disk to increase data integrity and security.
- 80 Gigabytes of local space in RAID1

DELL PowerEdge 1950

In the following table**Error! Reference source not found.**, we can see the final hardware configuration for the server DELL PowerEdge 1950, corresponding to the advanced configuration of the portfolio of solutions. This server offers a good ratio between effective cost, number of cores and RAM capacity. In the full solution there are 12 servers with these characteristics that are part of the cluster that will run the computing element.

PE1950	
1	PE1950 III QUAD-CORE XEON E5410 2.33GHZ/
1	PE1950 PCIE RISER (2 SLOTS)
1	PE1950 BEZEL ASSEMBLY
1	8GB 667MHZ FBD (4X2GB DUAL RANK DIMMS)
1	PE1950 III - ADDITIONAL QUAD-CORE XEON E
2	73GB SAS 15K 3.5" HD HOT PLUG
1	SAS 6/IR INTEGRATED CONTROLLER FOR C1
1	PE1950 III 3.5" HDD SUPPORT CHASSIS
1	8X IDE DVD-ROM DRIVE
1	PE1950 III REDUNDANT POWER SUPPLY
1	POWER CORD, PDU (RACK)
1	TCP/IP OFFLOAD ENGINE 2P
1	DRAC 5 CARD
1	NO OPERATING SYSTEM
1	PE1950 OPENMANAGE KIT AND FI DRIVER
1	SLIDING COMBINATION RAPID/VERSA RAIL WIT
1	PE1950 III - C1, MOTHERBOARD SATA/SAS RAID1
1	POWEREDGE ORDER - SPAIN
1	BASE WARRANTY
1	1YR BASIC WARRANTY - NEXT BUSINESS DAY -
1	3YR BASIC WARRANTY - NEXT BUSINESS DAY

In the following table**Error! Reference source not found.** we can see the partition table common to the 12 servers that comprise the Computing Element; each one will run one host OS and one guest OS. One of them will run CE and other gLite site services; the other 11 will run WN gLite software.

dev	lvm	Size
/dev/sda	boot	100 Mbytes
/dev/sdb	swap	2 GBytes
/dev/sdc	root	10 Gbytes
	wn	66 GBytes
	wn_swp	2 GBytes

The total power of the Computing Element is described in the following list:

- 96 cores running at 2.33GHz
- 96 Gigabytes of RAM
- 876 Gigabytes of local space in RAID1

8. DACS2 – Karolinska Institutet (KI, Stockholm - Sweden)

Karolinska Institutet will contribute to the neuGRID infrastructure with a site composed by a computational node and a data node. These nodes are part of the PROD infrastructure defined in the DoW.

8.1. Services

The site at DACS2 will provide an SE and a DPM, and a CE and WN to the grid infrastructure. This infrastructure will allow users to store and access files, and to execute jobs. This fits with the specifications of the DoW.

Due to the gLite grid middleware architecture, it is necessary to set up a service sBDII that will present the site to the core services of the neuGRID infrastructure.

This site will provide 2,6 TBytes of usable space in 2 different data stores, but the most impressive numbers come from the CE, running 13 WN of 8 cores each one. This infrastructure provides a total of 96 cores running at 2,0GHz and 96 GBytes of RAM.

8.1.1. Services Configuration

In this section we will describe the distribution of the gLite services deployed in the KI servers. In the followings tables we can see the deployed services in the set of servers. Those can be categorized in two major families: the gLite grid-site related services (UI, DPM, CE, WN) and the neuGRID Gateway. Service TAG and the IP distribution can be also viewed in the following tables.

Each server will have 3 IP addresses, the first IP address is attached to the dom0 (host OS), the second IP address is attached to the dom1 (guest OS), and finally the third IP address is attached to the remote administration card.

Following it is presented a table describing the physical servers, including information about the server name, the dom0 IP address (host), the management IP (DRAC card) and finally the serial number of the server (Service TAG). This table will show us the ground infrastructure in which the grid services will be deployed.

Server Name	Dom0 IP	Management IP	Service TAG
ng-ki-node1	192.168.0.10	192.168.0.120	1CWB84J
ng-ki-node2	192.168.0.11	192.168.0.121	9CWB84J
ng-ki-node3	192.168.0.12	192.168.0.122	JBWB84J
ng-ki-node4	192.168.0.13	192.168.0.123	5CWB84J
ng-ki-node5	192.168.0.14	192.168.0.124	2CWB84J
ng-ki-node6	192.168.0.15	192.168.0.125	3CWB84J
ng-ki-node7	192.168.0.16	192.168.0.126	GBWB84J
ng-ki-node8	192.168.0.17	192.168.0.127	HBWB84J
ng-ki-node9	192.168.0.18	192.168.0.128	7CWB84J
ng-ki-node10	192.168.1.236	192.168.0.129	8CWB84J
ng-ki-node11	192.168.1.237	192.168.0.130	6CWB84J
ng-ki-node12	192.168.1.238	192.168.0.131	FBWB84J

ng-ki-node13	192.168.1.239	192.168.0.132	4CWB84J
ng-ki-dom0	130.237.143.72	130.237.143.71	9SZF84J
NA	NA	130.237.143.70	CXPB84J

In the next two tables it is represented the services hosted by each physical server, the reference comes form the previous table.

Following, in the next table, are described the 13 WN, these servers are in the hidden side of the network and builds the cluster of WN that will be accessible for the CE.

Server Name	External IP	Domain Name	Internal IP	gLite Service
ng-ki-node1	NA	ng-ki-wn-server1.neugrid	192.168.0.220	WN
ng-ki-node2	NA	ng-ki-wn-server2.neugrid	192.168.0.221	WN
ng-ki-node3	NA	ng-ki-wn-server3.neugrid	192.168.0.222	WN
ng-ki-node4	NA	ng-ki-wn-server4.neugrid	192.168.0.223	WN
ng-ki-node5	NA	ng-ki-wn-server5.neugrid	192.168.0.224	WN
ng-ki-node6	NA	ng-ki-wn-server6.neugrid	192.168.0.225	WN
ng-ki-node7	NA	ng-ki-wn-server7.neugrid	192.168.0.226	WN
ng-ki-node8	NA	ng-ki-wn-server8.neugrid	192.168.0.227	WN
ng-ki-node9	NA	ng-ki-wn-server9.neugrid	192.168.0.228	WN
ng-ki-node10	NA	ng-ki-wn-server10.neugrid	192.168.0.229	WN
ng-ki-node11	NA	ng-ki-wn-server11.neugrid	192.168.0.230	WN
ng-ki-node12	NA	ng-ki-wn-server12.neugrid	192.168.0.231	WN
ng-ki-node13	NA	ng-ki-wn-server13.neugrid	192.168.0.232	WN

Finally in this section it is presented the services exposed in the external side of the network, these services had been deployed in two virtual machines, in one physical server, and are interconnected to the hidden part of the network.

ServerName	External IP	Server Name	Internal IP	gLite Service
ng-ki-dom0	130.237.143.73	ng-ki-server4	NA	SE
ng-ki-dom0	130.237.143.74	ng-ki-server5	192.168.0.5	CE sBDII

8.2. Infrastructure

This site will be installed over 14 servers: 1 DELL PowerEdge 2950, and 13 DELL M600, containing the M600 Servers there is a M1000 Blade Enclosure, which provides M600 servers with power, connectivity, and remote management.

M10000 Blade Enclosure is a piece of advanced hardware which simplifies physical deployments, reduces power consumption and reduces the total cost of a cluster – in the case of neuGRID – bigger than 10 servers. Other benefits are the unification of the power management, redundancy, cabling, connectivity, monitoring, and remote operation of the cluster. Finally this option also offers a reduction of the space needed in the IT Center and the best factor between CPU power and Watts consumed.

Detailed hardware tests for each server can be found in "Annex 5 – DACS2 Hardware Tests" in which we can see that all the servers pass the hardware test suit detailed in the Annex 2.

8.2.1. Servers Configuration

As previously explained, for the neuGRID deployment, all of the servers have been virtualized with XEN Software; the host Operating System is Scientific Linux 5.2. It runs a XEN Kernel in the top, allowing the server to run paravirtualized Virtual Machines. This technique has some advantages when compared to deploying the OS running gLite services directly in the hardware. These advantages are:

- Fast deployment of preconfigured images in the Guest OS.
- Simple maintenance and upgrades of the Guest OS.
- Simple backup of the whole system.
- Allows the running of more than one Guest OS in one physical server.

DELL PowerEdge 2950

In the next table, we can see the final hardware configuration for the server DELL PowerEdge 2950, corresponding to the advanced configuration of the portfolio of solutions. This server offers a good ratio between cost, raw space and power consumption.

PE2950	
1	PE 2950 III Quad Core Xeon E5430 (2.66GHz, 2x6MB, 1333MHz FSB)
1	PE2900 III SERVER RACK CHASSIS
1	RACK BEZEL ASSEMBLY
1	8GB (4x2GB Dual Rank DIMMs) 667MHz FBD
1	PE2950 III Additional Quad-Core Xeon E5430 (2.66GHz, 2x6MB, 1333MHz FSB)
1	3.5 INCH 1.44MB FLOPPY DRIVE
6	750GB SERIAL ATAU 7.2K 3.5" HD HOT PLUG
1	PERC 6/I INTERNAL RAID CONTROLLER CARD
2	80GB SATA 7.2K 3.5INCH FRONT BAY HDD
1	16X SATA DVD
1	REDUNDANT POWER SUPPLY (2 HOT PLUG PSU)
1	POWER CORD, PDU (RACK)
1	BROADCOM TCP/IP OFFLOAD ENGINE FUNCTIONA
1	TCP/IP OFFLOAD ENGINE 2P
1	DRAC5 CONTROLLER
1	NO OPERATING SYSTEM

1	OPENMANAGE SERVER SOFTWARE, WITH CD & DO
1	PE2950 UNIVERSAL SLIDING VERSA/RAPID RAC
1	PE2950 III - C11 RAID5/1 PERC5I/6I 3-8HD
1	BASE WARRANTY
1	1YR BASIC WARRANTY - NEXT BUSINESS DAY -
1	3YR BASIC WARRANTY - NEXT BUSINESS DAY

The following table shows the partition table of the DELL PowerEdge 2950, including host OS and the partitions to be installed in the guest OS. This server will run the CE, sBDII, TORQUE in one Virtual Machine. In another Virtual Machine will run DPM with 2 data stores, of 1.3 GBytes, with a total usable space of 2,6 TBytes as data stored in the GRID. As explained before, the physical support of this space relies on a RAID5 in order to ensure maximum data security and stability.

Dev	**lvm**	**Size**
/dev/sda	boot	100 Mbytes
	lvswap	8 GBytes
	lvroot	10 Gbytes
	dpm	24 GBytes
	dpm_swp	2 GBytes
/dev/sdb	dpm_data	1.3 TBytes
/dev/sdc	dpm_data_2	1.3 TBytes
	sbdii	20 GBytes
	sdbi_sw	2 GBytes

The total power of the Storage Element is described in the following list:

- 8 cores running at 2.33GHz
- 8 Gigabytes of RAM
- 2600 Gigabytes of shared disk space in RAID5.
- 80 Gigabytes of local space in RAID1
- 1000 Gigabytes of non shared disk space containing OS Images.

DELL PowerEdge M600

In the following table**Error! Reference source not found.**, we can see the final hardware configuration for the servers DELL PowerEdge M600, corresponding to the advanced configuration of the portfolio of solutions. This server offers a good ratio between effective cost, number of cores, RAM capacity and Power Consumption. In the full solution there are 13 servers with these characteristics that are part of the cluster that will run the computing element.

PE M600	
1	PE M600 Quad Core Xeon E5405 (2.0GHz, 2x6MB, 1333MHz FSB)

1	M600 PCIE RISER (2 SLOTS)
1	M600 BEZEL ASSEMBLY
1	8GB (4x2GB Dual Rank DIMMs) 667MHz FBD
1	M600 Additional Quad-Core Xeon E5405 (2.0GHz, 2x6MB, 1333MHz FSB)
2	73GB SAS 15k 2.5" HD Hot Plug
1	SAS 6/IR INTEGRATED CONTROLLER FOR C1
1	TCP/IP OFFLOAD ENGINE 2P
1	NO OPERATING SYSTEM
1	M600 III - C1, MOTHERBOARD SATA/SAS RAID1
1	BASE WARRANTY
1	1YR BASIC WARRANTY - NEXT BUSINESS DAY -
1	3YR BASIC WARRANTY - NEXT BUSINESS DAY
1	M600 III - C1, MOTHERBOARD SATA/SAS RAID1

In the following table**Error! Reference source not found.** we can see the partition table common to the 13 servers that comprise the Computing Element; each one will run one host OS and one guest OS. All the 13 Servers will run WN gLite software.

dev	lvm	Size
/dev/sda	boot	100 Mbytes
/dev/sdb	swap	2 GBytes
/dev/sdc	root	10 Gbytes
	wn	66 GBytes
	wn_swp	2 GBytes

The total power of the Computing Element is described in the following list:

- 96 cores running at 2.33GHz
- 96 Gigabytes of RAM
- 876 Gigabytes of local space in RAID1

9. DACS3 – Vereniging voor christelijk hoger onderwijs, wetenschappelijk onderzoeken patiëntenzorg (VUmc, Amsterdam – The Netherlands)

As the other DACS, VUmc will contribute to the neuGRID infrastructure with a site composed by a computational node and a data node and these nodes are part of the PROD infrastructure defined in the DoW.

9.1. Services

The site at DACS3 provides a SE, a DPM, a CE and WNs to the grid infrastructure. This infrastructure will allow users to store and access files, and to execute jobs. This fits with the specifications of the DoW.

Due to the gLite grid middleware architecture, it is necessary to set up a service sBDII that will present the site to the core services of the neuGRID infrastructure.

This site will provide 4,5 TBytes of usable space in 3 different data stores. The CE will run 12 WN of 8 cores each one. This infrastructure provides a total of 96 cores running at 2,3GHz and 96 GBytes of RAM.

9.1.1. Services Configuration

In this section we will describe the distribution of the gLite services deployed in the VUmc servers. In the followings tables we can see the deployed services in the set of servers. Those can be categorized in two major families: the gLite grid-site related services (UI, DPM, CE, WN) and the neuGRID Gateway. Service TAG and the IP distribution can be also viewed in the following tables.

Each server will have 3 IP addresses, the first IP address is attached to the dom0 (host OS), the second IP address is attached to the dom1 (guest OS), and finally the third IP address is attached to the remote administration card.

Server Name	Dom0 IP	Management IP	Service TAG
ng-vumc-dell-1	192.168.0.100	192.168.0.120	4FGW3J
ng-vumc-dell-2	192.168.0.101	192.168.0.121	GDR3W3J
ng-vumc-dell-3	192.168.0.102	192.168.0.122	JYT3W3J
ng-vumc-dell-4	192.168.0.103	192.168.0.123	WSA3W3J
ng-vumc-dell-5	192.168.0.104	192.168.0.124	TER3W3J
ng-vumc-dell-6	192.168.0.105	192.168.0.125	NHU3W3J
ng-vumc-dell-7	192.168.0.106	192.168.0.126	FRE3W3J
ng-vumc-dell-8	192.168.0.107	192.168.0.127	MRT3W3J
ng-vumc-dell-9	192.168.0.108	192.168.0.128	3NH3W3J
ng-vumc-dell-10	192.168.0.109	192.168.0.129	YHT3W3J
ng-vumc-dell-11	192.168.0.110	192.168.0.130	VGJ3W3J
ng-vumc-dell-12	192.168.0.111	192.168.0.131	7NJ3W3J
ng-vumc-dell-13	192.168.0.112	192.168.0.132	M3Q3W3J

Server Name	Internal IP	gLite
ng-vumc-dell-1	192.168.0.140	SE
ng-vumc-dell-2	192.168.0.141	CE sBDII
ng-vumc-dell-3	192.168.0.142	WN
ng-vumc-dell-4	192.168.0.143	WN
ng-vumc-dell-5	192.168.0.144	WN
ng-vumc-dell-6	192.168.0.145	WN
ng-vumc-dell-7	192.168.0.146	WN
ng-vumc-dell-8	192.168.0.147	WN
ng-vumc-dell-9	192.168.0.148	WN
ng-vumc-dell-10	192.168.0.149	WN
ng-vumc-dell-11	192.168.0.150	WN
ng-vumc-dell-12	192.168.0.151	WN
ng-vumc-dell-13	192.168.0.152	WN

9.2. Infrastructure

This site will be installed over 13 servers: 1 DELL PowerEdge 2900, and 12 DELL PowerEdge 1950 in the first stage of the deployment.

9.2.1. Servers Configuration

As previously explained, for the neuGRID deployment, all of the servers have been virtualized with XEN Software; the host Operating System is Scientific Linux 5.2. It runs a XEN Kernel in the top, allowing the server to run paravirtualized Virtual Machines. This technique has some advantages when compared to deploying the OS running gLite services directly in the hardware. These advantages are:

- Fast deployment of preconfigured images in the Guest OS.
- Simple maintenance and upgrades of the Guest OS.
- Simple backup of the whole system.
- Allows the running of more than one Guest OS in one physical server.

DELL PowerEdge 2900

In the next table, we can see the final hardware configuration for the server DELL PowerEdge 2900, corresponding to the advanced configuration of the portfolio of solutions. This server offers a good ratio between cost and raw space.

PE2900	
1	PE2900 III - QUAD-CORE XEON E5410 2.33GHz

1	PE2900 III SERVER RACK CHASSIS
1	RACK BEZEL ASSEMBLY
1	8GB 667MHZ FBD (4X2GB DUAL RANK DIMMS)
1	PE2900 III - ADDITIONAL QUAD-CORE XEON E
1	3.5 INCH 1.44MB FLOPPY DRIVE
8	750GB SERIAL ATAU 7.2K 3.5" HD HOT PLUG
1	PERC 6/I INTERNAL RAID CONTROLLER CARD
2	80GB SATA 7.2K 3.5INCH FRONT BAY HDD
1	16X SATA DVD
1	REDUNDANT POWER SUPPLY (2 HOT PLUG PSU)
1	POWER CORD, PDU (RACK)
1	BROADCOM TCP/IP OFFLOAD ENGINE FUNCTIONA
1	TCP/IP OFFLOAD ENGINE 2P
1	DRAC5 CONTROLLER
1	NO OPERATING SYSTEM
1	OPENMANAGE SERVER SOFTWARE, WITH CD & DO
1	PE2900 UNIVERSAL SLIDING VERSA/RAPID RAC
1	PE2900 III - C11 RAID5/1 PERC5I/6I 3-8HD

This server will run the DPM with 3 data stores, 2 of 1.43 GBytes, and one of 1.2 GBytes, with a total usable space of 4 TBytes as data stored in the GRID. As explained before, the physical support of this space relies on a RAID5 and a spare disk in order to ensure maximum data security and stability.

The total power of the Storage Element is described in the following list:

- 8 cores running at 2.33GHz
- 8 Gigabytes of RAM
- 4500 Gigabytes of shared disk space in RAID5 with 1 spare disk to increase data integrity and security.
- 80 Gigabytes of local space in RAID1

DELL PowerEdge 1950

In the following table**Error! Reference source not found.**, we can see the final hardware configuration for the server DELL PowerEdge 1950, corresponding to the advanced configuration of the portfolio of solutions. This server offers a good ratio between effective cost, number of cores and RAM capacity. In the full solution there are 12 servers with these characteristics that are part of the cluster that will run the computing element.

PE1950	
1	PE1950 III QUAD-CORE XEON E5410 2.33GHZ/
1	PE1950 PCIE RISER (2 SLOTS)
1	PE1950 BEZEL ASSEMBLY
1	8GB 667MHZ FBD (4X2GB DUAL RANK DIMMS)
1	PE1950 III - ADDITIONAL QUAD-CORE XEON E
2	73GB SAS 15K 3.5" HD HOT PLUG
1	SAS 6/IR INTEGRATED CONTROLLER FOR C1
1	PE1950 III 3.5" HDD SUPPORT CHASSIS
1	8X IDE DVD-ROM DRIVE
1	PE1950 III REDUNDANT POWER SUPPLY
1	POWER CORD, PDU (RACK)
1	TCP/IP OFFLOAD ENGINE 2P
1	DRAC 5 CARD
1	NO OPERATING SYSTEM
1	PE1950 OPENMANAGE KIT AND FI DRIVER
1	SLIDING COMBINATION RAPID/VERSA RAIL WIT
1	PE1950 III - C1, MOTHERBOARD SATA/SAS RAID1

The total power of the Computing Element is described in the following list:

- 96 cores running at 2.33GHz
- 96 Gigabytes of RAM
- 876 Gigabytes of local space in RAID1

10. Large Scale Tests

In order to test the infrastructure, two large scale tests have been done:

- AC/DC2 - Test Code Name "Highway to Hell": This test was organized and run at month 24 by WP11 on the infrastructure that was deployed by WP8. The challenge consisted in analysing the US-ADNI dataset made of 6'500 brain image scans, and occupied the neuGRID computing resources full time during 10 consecutive days. It allowed identifying some software libraries incompatibilities at the operating system level as well as an inconsistency in the pipeline itself. As a consequence, the grid Worker Nodes' operating systems were reconfigured within the entire infrastructure to better support CIVET and the latter was recently corrected and made available in the grid.
- AC/DC3 - Test Code Name "Thunderstruck": This test was just launched at the time of writing. Basically, it the same test as the previous one but with again a larger scale. Indeed, it consists in analyzing the US-ADNI dataset made approximately of 7'500 brain image scans now. The analysis is done this time by tree different pipelines (each single scan will be analyze by the 3 pipelines):
    - CIVET (take approximately 6h per scan)

- BRAINVISA (take approximately 30min per scan)
- FREESURFER (take approximately 24h per scan)

As you can see, the theorical processing time is really huge if you multiply it by the number of scans. Therefore, this very large scale test requires some external help. This was done by requesting some EGEE/EGI sites to accept the neuGRID virtual organization and to provide some computing power to the project. In order to be able to do that, a lot of work had to be done in the neuGRID infrastructure to respect the EGEE/EGI policies (VO name format etc.). This induced that most of the neuGRID sites had to be reconfigured. Once everything was properly done, we managed to get four seed sites. neuGRID is now able to expose a set of 2.000 CPU cores with an accompanying 45TB of storage capacity, over network lines performing from 100 Mbps to 1 Gbps across Europe.

More information about those 2 data challenges can be found in the deliverable D11.3 "AC/DC3 Test Suite Specification & Recommendations Report".

11. Conclusions

The present document has introduced the solution for deploying a complex set of software in heterogeneous information systems, which satisfies the security, budget and physical constraints imposed by the neuGRID project. The document also elaborated on the achievements of the work package contrasted with the initially foreseen work plan.

The first step WP8 accomplished, as part of this work, was to analyze the neuGRID platform (mainly the Gateway and gLite grid middleware) requirements in order to deliver an appropriate infrastructure.

Indeed, technologies such as gLite are quite new and resource demanding due to their scope and nature. The challenge WP8 therefore faced, was to build a solution that is capable of running such technologies, while maintaining a reasonable overall quality of services. To do so, it investigated various aspects related to hardware, software and associated costs. As a result, a deep analysis of the current virtualization techniques and technologies was undergone with the aim of drastically reducing the requirements and making it accessible while conforming to all possible institutions and respective IT configurations.

As a significant output of the work package, a complete portfolio of Hardware has been drawn up, which can be used as a reference for the partners to acquire the complete solution of hardware that will better fit in IT centres.

WP8 successfully exercised and exemplified its approach through the deployment of GCC and finally deploying the Level1 DACS. The deployment of DACS1, DACS2 and DACS3 was a big challenge – DACS1 (12 WN and 1 DPM), DACS2 (13 WN and 1 DPM) and DACS3 (13 WN and 1 DPM) - in terms of logistics and deployment.

Everything is fully operational and has been deeply tested during the two large scaled tests named AC/DC2 and AC/DC3.