StratusLab

Enhancing Grid Infrastructures with
Virtualization and Cloud Technologies

Infrastructure Tools and Policy Specification

Deliverable D5.2 (V1.0)
15 December 2010

Abstract

This document presents the tools, policies and procedures applied for the operation of the computing infrastructure provided by the StratusLab project. The document focuses on three services that are operated on a production-basis by the project, and in particular: the public IaaS cloud service, the virtual appliances repository and the virtualized grid sites running on top of the public cloud. For each one of the above services six aspects are considered: generic operations, maintenance and upgrades, third party access, security, monitoring and accounting, and quality of service. Many of these tools and procedures are already in place and are expected to evolve with the progress of the project, as more experience from the operation of these services is gained.

CAPACITIES

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

The information contained in this document represents the views of the copyright holders as of the date such views are published.

THE INFORMATION CONTAINED IN THIS DOCUMENT IS PROVIDED BY THE COPYRIGHT HOLDERS “AS IS” AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE MEMBERS OF THE STRATUSLAB COLLABORATION, INCLUDING THE COPYRIGHT HOLDERS, OR THE EUROPEAN COMMISSION BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THE INFORMATION CONTAINED IN THIS DOCUMENT, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Contributors

Name	Partner	Sections
Evangelos Floros	GRNET	All
Stuart Kenny	TCD	Chapter 5

Document History

Version	Date	Comment
0.1	12 July 2010	Initial draft with TOC.
0.2	25 Nov. 2010	Draft content for Cloud Services chapter
0.3	29 Nov. 2010	Details for Cloud Services. Added Appendix A (AUP)
0.4	03 Dec. 2010	Details for Grid Services and Appliance Repository. Added Appendix B (SratusLab VO)
0.5	07 Dec. 2010	Added figures and references. Fixed various typos and omissions in the rest of the chapters. Added more details about Appliance repository mirroring in Chapter 5
0.6	08 Dec. 2010	First complete draft. Added Executive Summary, Introduction and Conclusions
0.7	13 Dec. 2010	Applied various changes in response to internal review comments provided by Cal Loomis (LAL) and Nasia Assiki (GRNET)
0.8	15 Dec. 2010	Applied various changes and improvements in response to internal review comments provided by Marc-Elian Bégin (SixSQ).
1.0	15 Dec. 2010	Final version

Contents

List of Figures

List of Tables

1 Executive Summary

This document presents the tools, policies and procedures applied for the operation of the computing infrastructure provided by the StratusLab project. The document focuses on three services that are operated on a production-basis by the project: the public IaaS cloud service, the virtual appliances repository and the virtualized grid sites running on top of the public cloud. For each one of the above services six aspects are considered: generic operations, maintenance and upgrades, third party access, security, monitoring and accounting, and quality of service.

StratusLab operates a public cloud service which is build upon the software distribution integrated by the project. This service offers a technology preview of the solutions developed by the project and additionally provides the testbed for various activities that require IaaS capabilities (the deployment of virtualized grid sites being one of them). The operation of a cloud site is not a trivial activity, and although the technology is currently very hyped with various solutions being developed for the management and provision of virtual machines, sustaining a high quality of service is a challenging endeavor. Users are accustomed to cloud services providing high-availability, limited downtimes and on-demand access to VM resources at any given moment without delays. In pursue of this target we have defined procedures for seamless maintenance and upgrades of both software and hardware resources. Proper monitoring and notification of problems is required on the physical infrastructure (e.g. in order to identify hardware failures or other issues on physical nodes) and the virtualized layer (e.g. the VMs hosted on the physical nodes).

Since one of the primary usages of the cloud services will be to host grid sites many of the operational requirements of the former are imposed by the established quality of service requirements of the latter. Production grid sites are governed by the operational policies and procedures defined by EGI [1]. Grid sites operated on the cloud should adhere to the same rules and provide at least the same quality of service as their physical counterparts. Actually the exploitation of cloud technologies are expected to provide more flexibility to the grid administrators in cases of maintenance and upgrade activities as well as handling security incidents. StratusLab has a strong belief on the above that is why for example we it have set higher goals regarding grid site availability and reliability metrics.

The appliance repository is an integral service for the provision of IaaS cloud capabilities. All VMs instantiated on the cloud infrastructure or hosted on the repository and are transferred to the cloud frontend upon their instantiation. Thus our goal is to provide high availability and redundancy. For this reason apart from a main repository operated in TCD at least one more mirror will be provided from GRNET. Additional mirrors can be setup by third parties wishing to improve transfer times by reducing network proximity.

Many of the tools and procedures defined in this document are already in place. Others will be applied in the coming months as the respective services will be put into operation. For example the definitions for grid operations will be applicable once the first production grid site will be deployed on the StratusLab production cloud service. Apparently the policies defined in this document will evolve with the progress of the project as more experience from the operation of these services is gained.

2 Introduction

StratusLab WP5 is the activity responsible for the provision and operation of the computing infrastructure used for various purposes in the context of the project. Part of the computing resources are used internally to cover computational needs for the project itself (e.g. having a testbed for software development and testing) whereas a significant portion is used for hosting services offered to third parties external to the project. Moreover it is an activity that operates continuously dealing with both with scheduled procedures as well unforeseen conditions. Infrastructure operations are typically a non-trivial activity since there are many parameters that should be taken into account. Formally defined policies and properly selected tools are fundamental for the successful provision of computing services that they are able to sustain a high level Quality of Service (QoS).

This documents focuses on the operational tools and policies for three important components of the StratusLab infrastructure:

Reference cloud service which offers access to third parties wishing to try out the cloud solutions developed by the project,

One or more production grid sites operated by the project on top of the cloud infrastructure, and

Appliance repository which hosts the VM images that can be instantiated in the cloud service.

The common denominator of the above is the requirement to serve a wide range of user communities and potential applications. These services should offer an acceptable QoS in order to sustain the workloads required by real-life applications and high-level demands of actual users and scientific communities. Combined with the above services, the activity provides the testbed for verifying that the StratusLab distribution satisfies the project's primary goal: the provision of a cloud computing environment appropriate for hosting production grid services.

For each one of the above services we consider the following aspects:

- A generic overview of operations strategy,
- The procedures that should be followed for daily maintenance activities and periodic update processes,

- The process for providing access to external users,
- The security requirements or/and features offered by the service,
- The tools used for service monitoring and resource accounting, and
- The Quality of Service targeted by the service and how this is pursued.

Since this document is prepared during M6 of the project, part of these services are already available and operated by WP5 (namely the cloud service and the appliance repository). Consequently, a number of policies and operational tools have already been established. Others are under consideration and are being negotiated among the project partners. Thus this document serves both as a report of the existing established practices as well as a specifications document of the solutions planned for the coming months of the project.

3 Cloud Services

3.1 Operations Overview

Cloud services reside at the heart of the project's operations activity. These services are built upon the StratusLab cloud distribution that is developed and integrated in other project work packages, namely WP4 and WP6. StratusLab operations activity has several goals:

1. To provide a public cloud service as a showcase of the IaaS capabilities that can be enabled from the StratusLab cloud distribution.
2. To operate a private cloud in order to offer virtualized resources for various services that need to run in the context of the project.
3. To offer a testbed for beta-testing the snapshot versions of the StratusLab distribution whenever new features are ready to be released.
4. To support various other project activities that require computing resources for short- or medium-term usage.

In the rest of the chapter we focus on the production cloud services operated in the context of the project.

3.1.1 Production service

The production service is deployed using the latest release of the StratusLab distribution. This service is freely open to the public, giving third parties the opportunity to try out the capabilities of a cloud service running the StratusLab distribution. The production cloud service is operated by GRNET using the physical infrastructure that has been described in D5.1 [9].

Initially, a total of 11 nodes have been allocated to support the production service. One node acts as the frontend for the OpenNebula virtual machine manager and the remaining 10 act as hosting nodes for the Virtual Machine instances. This setup offers the ability to start a total of 160 virtual machine instances without CPU overcommitment (by considering a 1-core/1-VM allocation strategy). Depending on the popularity of the service and the number of requests from external users, additional physical resources for the production infrastructure can be allocated. An abstract architecture of the project's reference cloud service is shown in Figure 3.1.

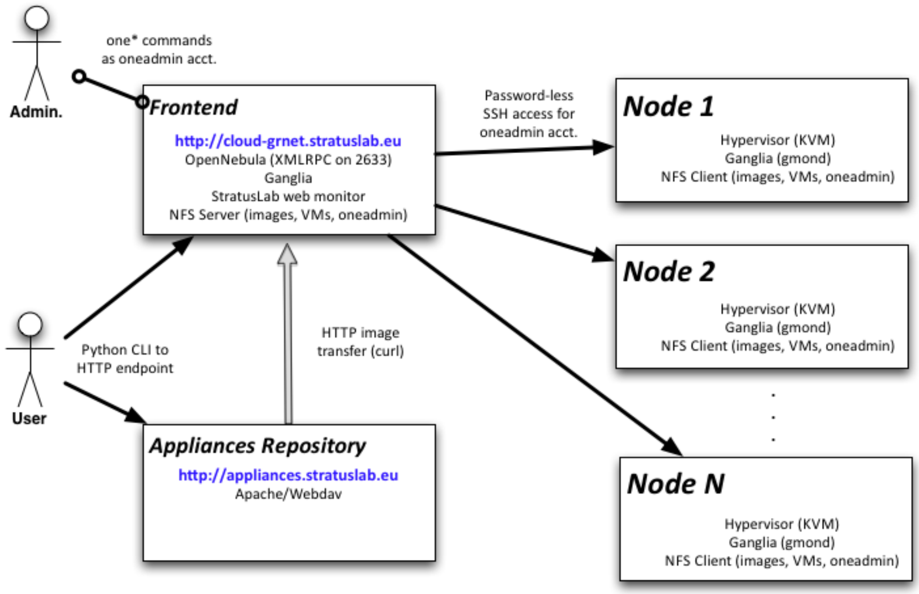


Figure 3.1: *Architecture of StratusLab production cloud service*

Currently StratusLab does not offer a cloud storage service, the ability to configure private virtual networks, or to allocate static public IPs for virtual machines. These features will be integrated as they are being implemented in the upcoming releases of the StratusLab software distribution. For the time being, if users have special requirements regarding the above resources they can contact the project support team and negotiate the possibility for a custom configuration. For example for storage, the GRNET site has allocated a total of 20 TB from the local storage server. This storage is currently used internally for various purposes (e.g. for hosting the VM images of the Appliances Repository mirror). Nevertheless, a fraction of this storage could be manually allocated per case for specific VMs. In this case, the user will have to request the amount of needed storage and indicate to which VMs this should be attached.

3.2 Maintenance and Upgrades

3.2.1 Core cloud services

The production service will stay up to date with the releases of StratusLab distribution. Since these follow a cycle of six-week development periods the reference service will pursue to follow the same pace. The upgrade process should remain transparent to the cloud users. We foresee two types of upgrades:

- Upgrades that fix bugs, improve the performance and provide additional functionality without altering the core API of the StratusLab components (e.g. OpenNebula [5] API). These upgrades are not expected to interfere with the system. An update of the system packages should be feasible without downtimes or any other actions impacting the proper operation of VMs and the other services offered to the user.

- Upgrades that alter the API or modify the integral functionality of the StratusLab core systems. For example OpenNebula implements a database schema that is used to keep information for various aspects of the cloud operation and the provision of services. In case that this schema is modified the existing database should migrate to the new one. The migration path might not be simple and straightforward depending on the number of new schema elements introduced. In this case a custom migration path will be defined to minimize the interruption to the running services. Such updates are not expected often.

In order to ensure the stability of the production cloud service the following procedure will be followed:

1. A new StratusLab version is marked ready for release.
2. The latest version is deployed on a testing installation which is setup from scratch in the pre-production infrastructure of the project. If the testing is successful the version is marked as Release Candidate (RC).
3. The RC version is used to upgrade the pre-production site of the project
4. If the upgrade is successful the RC is marked as validated and becomes the official release of the distribution.
5. If the new release is only a bug fix or performance improvement it is immediately deployed in the production infrastructure following the clear instructions provided by WP4.
6. If the new release introduces new features and renders a major upgrade of the service functionality it is not incorporated in the production service until a clear migration path is defined based on the tests performed in the testing and pre-production infrastructure

3.2.2 Infrastructure Maintenance

The physical infrastructure where all the services are hosted will naturally go through periodic maintenance activities. Since the infrastructure is co-hosted with resources dedicated for other purposes and projects, apart from StratusLab, it is expected that these will sometimes interfere with the operations of StratusLab infrastructure. For example, so far in the GRNET datacenter it was required in many cases to change the configuration of the network setup and the cabling of the network interconnection. In other cases the router and switches firmware required to be upgraded in order to fix various bugs. Such activities introduce downtimes and require that the services will become unavailable for a short period of time. Unforeseen problems are also expected during the lifetime of the service provision.

In order to deal in the best possible way with these situations a specific window of maintenance has been negotiated with the datacenter providers. For the time being there are two maintenance windows defined, the first on Monday morning and the second on Wednesday morning. Downtime periods will be announced to the registered cloud users in order to give them time for to take preventative actions if needed (e.g. take a backup of their data from a VM or to gracefully bring a service down in order to avoid leaving it in a stale state etc.). For scheduled downtimes the event will be announced *at least 5 working days* before the downtime. Notice that this coincides with the current requirement for announcing downtimes in EGI [1] production grid sites. Unscheduled downtimes will be announced immediately the moment the decision for the bringing the service offline will be taken. Although no guarantees can be given in this case about the exact time of the downtime start, this will be done in a reasonable time before the actual start of the downtime in order to give time for the users to respond.

3.3 Third party access to cloud services

The StratusLab public cloud is available for all interested users outside the project. The service gives the opportunity to external users to try out the cloud solutions developed by the project. The purpose is also to initiate potential new collaborations with external projects. For example by trying-out the StratusLab cloud, other DCI projects will have the chance to evaluate the services developed by the project and identify potential areas of interaction and collaboration.

3.3.1 Enrollment procedure

During the first phase of service provisioning users request access to the cloud by sending an email to support@stratuslab.eu. The user should provide valid identification information and a small description of the intended usage of the StratusLab cloud service. Once the user is verified a username/password pair is generated and is send by email together with some basic information about using the service.

In the future and as the software stack is expanded a more formal and automated procedure will be followed for enrolling new users. For this purpose a web page will be developed providing an entry point where new and existing users will find information about the service status, user documentation and other support material. In particular, users will enroll through this web page providing their contact details and information about their intended use. Users will be required to possess a valid digital certificate issued by one of the IGTF (International Grid Trust Federation) accredited CAs. All authentication procedures will require for such a certificate to be presented before a user may register or access the cloud services.

In order to ensure fair access to resources, quotas will limit the usage of cloud

Table 3.1: *StratusLab cloud default usage quotas*

Maximum number of instantiated VMs	20
Total amount of memory (in GByte)	40
Maximum amount of allocated storage	–
Maximum number of assigned public IPs	–

resouces. These quotas will limit, for example, the number of instantiated Virtual Machines (VMs), the total amount of storage, and the total amount of network bandwidth. For the time being, the OpenNebula 2.0 authorization system gives us the ability to place a limit on the total number of CPU cores and the total amount of memory across all VMs for a specific user. The default quotas enforced at the moment are given in Table 3.1. Those quotas without numbers will be implemented in the future.

3.3.2 User Support Services

The main source of support is the project web site at http://www.stratuslab.eu. The web site provides all the required information for accessing and using the public cloud services as well as news about the service status and the overall project progress. Users can seek direct support using the mailing list support@stratuslab.eu. This mailing list can be used for requesting access to the infrastructure, for reporting problems, for feature requests and for any other issues related to the provision of cloud services.

3.3.3 Acceptable Use Policy

Users will have to sign an Acceptable Use Policy (AUP) upon enrollment. The purpose of this policy is to explain to users the accepted usage of the service and to make them responsible for any illegal activities they initiate using these resources. The AUP text will be visible from the enrollment web page and the user must agree to the policy in order to complete the enrollment procedure.

StratusLab has decided to adopt the AUP currently enforced by EGI. Since StratusLab and EGI will provide overlapping services it makes sense that a common policy should be followed. For this purpose StratusLab members will actively contribute in the EGI Security Policy Group that formulates a common AUP and security policies for all European DCI projects. The current version of EGI's AUP is available from https://documents.egi.eu/document/74. All candidate users of StratusLab's public cloud services will be directed to the above document and will be asked to confirm their adherence to the clauses of this policy.

3.3.4 Client tools

The StratusLab public cloud is available to the users through the public APIs implemented by the StratusLab distribution. These APIs can be consumed by user-developed applications in order to exploit the public cloud capabilities. The StratusLab distribution itself provides a set of command line tools that implement the core functionality of the cloud. For example in release 0.1 the following tools enable the remote lifecycle management of VM instances:

stratus-run-instance Instantiates a new VM from an image available from the Appliances service.

stratus-describe-instance Provides details about a running instance

stratus-kill-instance Kills a running instances releasing all the resources it currently occupies.

stratus-upload-image Uploads a new VM image file to the appliances repository

Additional tools will be released enabling access to other features of the distribution that are currently not available; for example: storage management, network management, customized VM image creation and management etc.

The client side tools are available from the StratusLab software repository (http://repo.stratuslab.eu:8081/content/repositories) in rpm, deb and zipped packaged format. Since these tools are essentially python applications that communicate with the StratusLab services through web-based APIs any platform that supports python is also supported including Linux, Mac OS X and MS Windows.

Registered users can access the services using a project-provided username and password. For the VMs that a user instantiates, she/he can access them through SSH using public keys generated by the *ssh-keygen* application. These keys are passed to the VM upon instantiation using the stratus-run-instance command line

3.4 Security

3.4.1 Firewall security

The resource providers will not impose any restrictions to the number or type of ports that an end-user can open in the VMs that he/she instantiates. By default all ports in VM images will be closed and it will be up to the user to define which will be open upon instantiation. From the point of view of the resource provider this means that all ports in the client-nodes will be open from the local firewalls.

Virtual machines are running in a separate VLAN from the physical infrastructure hosting them. One the other hand the physical infrastructure itself should apply all necessary security measures in order to ensure the integrity of the hosting system. Essentially for the frontend machine only the ports required by OpenNebula and the other components of the StratusLab distribution should be accessible from outside. Table 3.2 summarizes the ports that should be open on the frontend.

Table 3.2: *Network ports open on the frontend*

2633	OpenNebula demon over HTTP
2634	OpenNebula demon over HTTPS (secured)

Currently the public cloud service is accessible only from the non-secured HTTP port. In the imminent future the plan is to activate the certificate based authentication that will be performed on TCP port 2634. Once this is put into operation access to TCP port 2633 will also be prohibited.

On the hosting nodes all ports should remain open in order to prevent blocking of the respective user-operated ports on the VMs. Only the absolutely necessary services for hosting and monitoring of VMs should be running on the nodes. Access to the nodes must be restricted to cloud site admins through public key identification and only from inside the hosting domain.

3.4.2 Data Integrity and Privacy

The cloud layer will not provide any encryption services. The user will be responsible for securing his/her data using relevant encryption and data protection mechanisms.

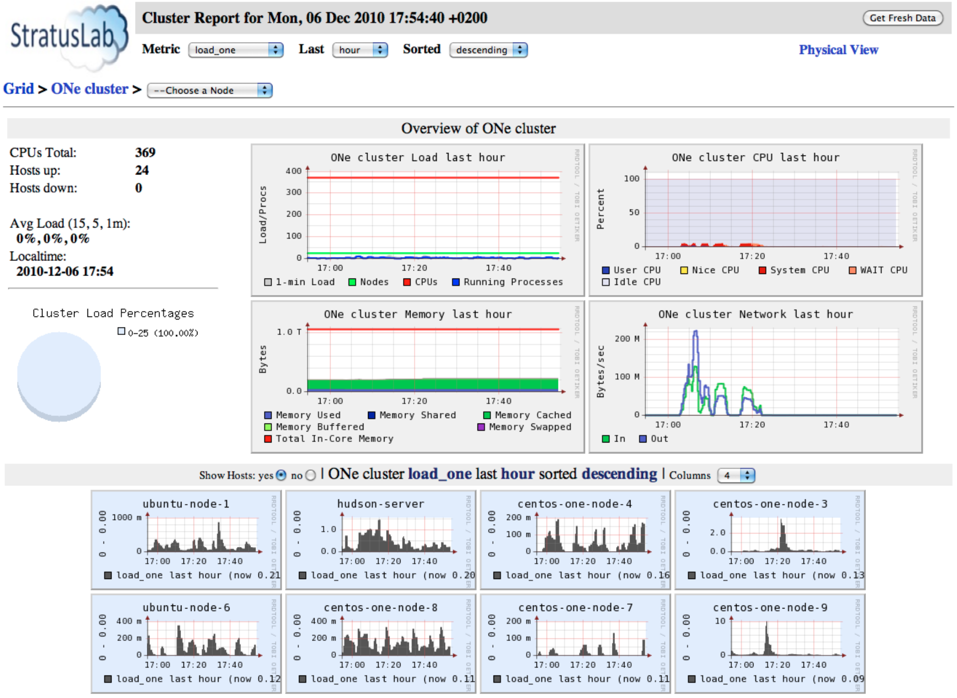
3.5 Monitoring and Accounting

System monitoring is an important aspect of cloud operations. As it has been identified in various other project documents ([9], [10] and [8]) from the resource provider point of view we consider two layers of monitoring.

Physical Infrastructure This provides detailed information about the setup and the status of the physical machines that comprise the hosting infrastructure of a resource provider.

Virtual Machines Monitoring on this layer focuses on the VM instances running on a given virtualized environment. It provides information about the configuration of the instances (number of CPU cores used, total main memory allocated, storage space etc) per cloud user. This layer is relevant not only for the resource provider but also for the cloud user itself that wishes have a clear overview of the virtual resources that she/he has allocated especially if these are provided using some kind of credit-based schema.

In the context of StratusLab one additional layer is relevant to monitoring, that of the Grid services. Monitoring of hosted grid sites is presented in more detail in Chapter 4. In brief this layer will be considered independent from the rest of the other services although a limited integration is expected to take place in the context of the project.



Cluster Report for Mon, 06 Dec 2010 17:54:40 +0200

Metric load_one Last hour Sorted descending

Grid > ONe cluster > --Choose a Node

Overview of ONe cluster

CPUs Total:	369
Hosts up:	24
Hosts down:	0

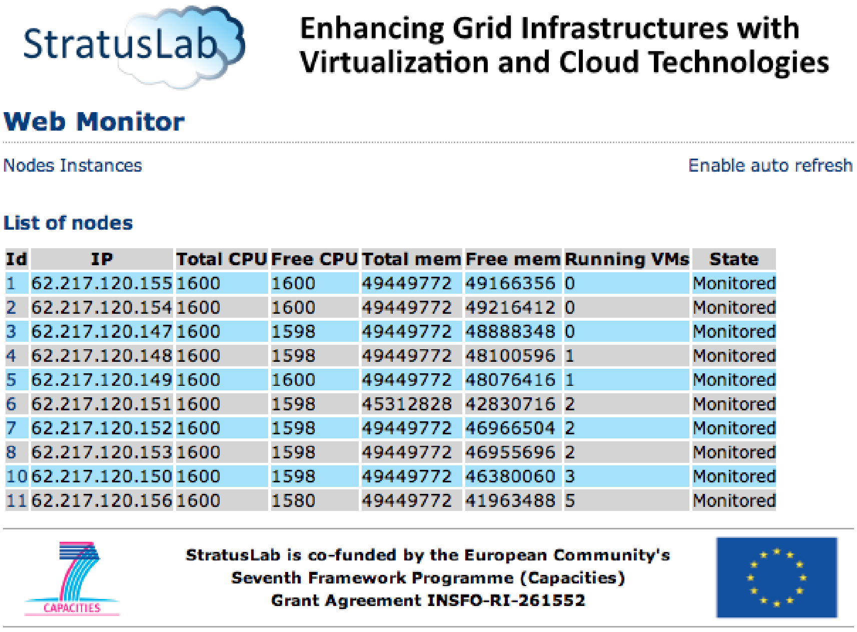
Avg Load (15, 5, 1m):
0%,0%,0%
Localtime:
2010-12-06 17:54

Show Hosts: yes ● no ○ | ONe cluster load_one last hour sorted descending | Columns 4

Figure 3.2: *Monitoring of StratusLab's physical infrastructure with Ganglia*

OpenNebula provides a set of basic monitoring capabilities that have already been integrated in StratusLab 0.1. Thus a simple monitoring web page is available on the cloud frontend. This web monitoring tools allows the real-time monitoring of all physical hosts comprising the reference cloud infrastructure. It also allows the inspection of all VMs instantiated by the *oneadmin* account (the default administrator account). A few sample screenshots of the cloud site monitoring is available in Figures 3.3 and 3.4.

StratusLab is currently working to provide an integrated view of the physical and virtual machine monitoring. This will be achieved by exploiting the monitoring capabilities of OpenNebula integrating them with a popular cluster monitoring and visualization tool, namely Ganglia [6]. Ganglia is a scalable distributed monitoring system for high-performance computing systems such as clusters and Grids. It is an extensible service that allows the incorporation of various additional monitoring capabilities. Ganglia is currently used to monitor the physical infrastructure in GRNET and in LAL. Virtual Machines can also be monitored by considering them as yet another computing node. This approach is not scalable though since it neglects the dynamic nature of cloud resources that may have shorter life span. For this reason Ganglia is currently being expanded to support, apart from the monitoring of physical nodes, also the VMs hosted on these nodes using custom developed probes that will render the system cloud-aware.



Enhancing Grid Infrastructures with Virtualization and Cloud Technologies

Web Monitor

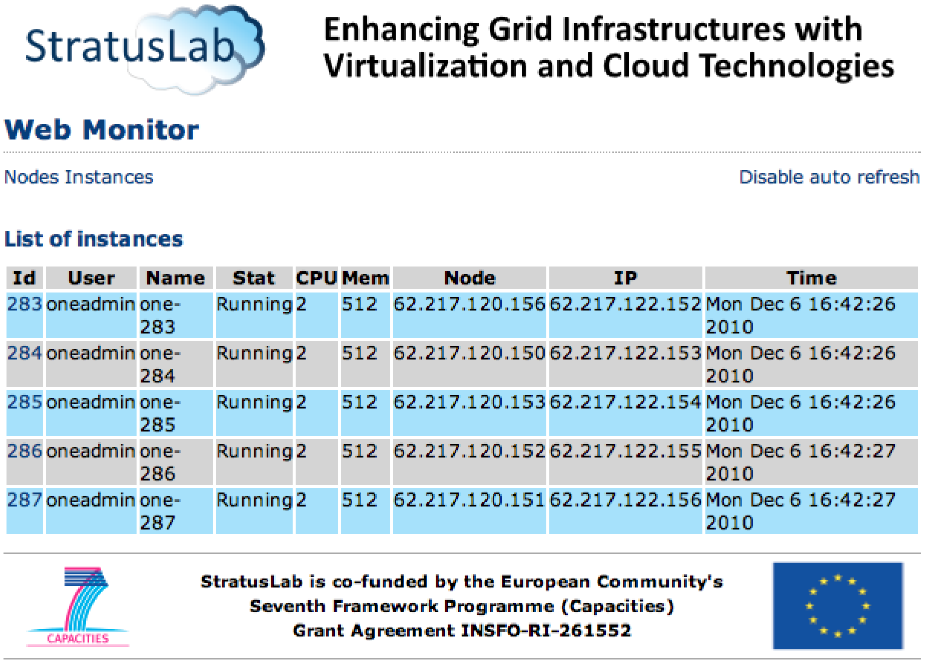
Nodes Instances Enable auto refresh

List of nodes

Id	IP	Total CPU	Free CPU	Total mem	Free mem	Running VMs	State
1	62.217.120.155	1600	1600	49449772	49166356	0	Monitored
2	62.217.120.154	1600	1600	49449772	49216412	0	Monitored
3	62.217.120.147	1600	1598	49449772	48888348	0	Monitored
4	62.217.120.148	1600	1598	49449772	48100596	1	Monitored
5	62.217.120.149	1600	1600	49449772	48076416	1	Monitored
6	62.217.120.151	1600	1598	45312828	42830716	2	Monitored
7	62.217.120.152	1600	1598	49449772	46966504	2	Monitored
8	62.217.120.153	1600	1598	49449772	46955696	2	Monitored
10	62.217.120.150	1600	1598	49449772	46380060	3	Monitored
11	62.217.120.156	1600	1580	49449772	41963488	5	Monitored

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

Figure 3.3: Web monitoring of physical hosts



Enhancing Grid Infrastructures with Virtualization and Cloud Technologies

Web Monitor

Nodes Instances Disable auto refresh

List of instances

Id	User	Name	Stat	CPU	Mem	Node	IP	Time
283	oneadmin	one-283	Running	2	512	62.217.120.156	62.217.122.152	Mon Dec 6 16:42:26 2010
284	oneadmin	one-284	Running	2	512	62.217.120.150	62.217.122.153	Mon Dec 6 16:42:26 2010
285	oneadmin	one-285	Running	2	512	62.217.120.153	62.217.122.154	Mon Dec 6 16:42:26 2010
286	oneadmin	one-286	Running	2	512	62.217.120.152	62.217.122.155	Mon Dec 6 16:42:27 2010
287	oneadmin	one-287	Running	2	512	62.217.120.151	62.217.122.156	Mon Dec 6 16:42:27 2010

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

Figure 3.4: Web monitoring of virtual machines

Regarding accounting, this has been also identified as a critical aspect of cloud services in project deliverable [D6.1]. Currently no accounting information are kept in the production service but such functionality will be integrated once the first results of WP6 activity are being released.

3.6 Quality of Service

One of the main purposes for deploying a production cloud service is to have the ability to host virtualized grid sites. Certified grid sites have specific requirements concerning Quality of Service. These requirements are presented in more detail in Chapter 4. As an immediate consequence of the above is that hosting cloud services should satisfy at least the Operation Level Agreements that a production grid service needs to support. The parameters affected are the the availability of the cloud service components that contribute to the provision of grid sites namely: computing services for hosting of VMs, storage services and network services. It should be noted that one of the initial motivations for developing the StratusLab distribution is the exploitation of cloud computing capabilities in order to improve grid site availability and other relevant QoS metrics. In the context of StratusLab, QoS is affected by the stability and maturity of the software distribution used for deploying the cloud services as well as the stability of the physical infrastructure. The project aims to provide a highly reliable cloud service. Currently all services are offered on best-effort basis. Through the operation of this services we expect to improve the various software components currently comprising the StratusLab distribution and the goal is by version 1.0 to be able to deliver a robust and stable cloud management distribution.

For what concerns the physical infrastructure, unexpected hardware problems are always part of daily operations. Prompt response to hardware failures reduces the unexpected downtimes and other critical problems (data loss, unavailability of critical services like the appliance repository).

StratusLab will also avoid single point of failures by deploying multiple instances of centralized critical services. For example for what concerns the appliances repository at least one mirror will be available in order to increase the availability of the service. Similar approaches will be followed for other components like the authorization services or the core cloud-management services.

4 Grid Services

StratusLab will deploy at least two certified grid sites that will join the pan-European grid infrastructure operated by EGI. The purpose for this will be to verify the suitability and the stability of the StratusLab distribution for hosting virtualized, production-level, grid sites. In this section we define the operational tools and policies for the provision of the these grid services.

4.1 Operations Overview

The provision of a production grid site is governed by a set of operational rules and policies that have been developed in the past years in the context of EGEE series of projects and currently are maintained and monitored by EGI [1] via the EGI-InSPIRE project. A key document that defines the expected setup and operation of a grid site is the EGI-InSPIRE Operational Level Agreement (OLA) Between NGI and Sites that is currently still in a draft status and available from the EGI-InSPIRE project web site (https://documents.egi.eu/public/ShowDocument?docid=65). According to this document in order for a grid site to be eligible for going through the certification process and be accepted as a production site in the EGI infrastructure the following conditions should be satisfied:

1. The site should provide at least one CE, one site BDII and one SE
2. The minimum storage capacity of the SE should be 1 TB.
3. The minimum number of CPU cores offered from the WNs is 8 cores
4. At least one person should be identified as system administrator and be the main contact for operations-related issues
5. The site should support at least one user-community VO

Grid sites operated by StratusLab will satisfy all of the above requirements. In particular for what concerns the last item (support for user-community VOs), StratusLab grid sites will support at least the Bioinformatics VO currently participating in the project through CNRS/IBCP. Additionally, as stated by the EGI policies, the grid sites operated by StratusLab will support the necessary OPS (operations) VOs (either EGI-wide or NGI-wide) which are used for centralized monitoring and for performing periodic Service Availability Monitoring (SAM) tests to the sites.

4.2 Maintenance and Upgrades

Maintenance administrative actions will take advantage as much as possible of the capabilities offered by cloud services. Since the grid sites are operated on the cloud the requirement for downtimes in order to perform maintenance activities on the physical infrastructure should be rare. The cloud operations will make sure that any interventions on the physical infrastructure will not impact the running VMs. For example in case a specific node needs to be brought off line, all VMs running on that node (including VMs belonging to grid sites) will be migrated to other healthy nodes. These migrations will be transparent to the grid administrators. On the other hand there might be cases that a site-wide downtime has to be introduced (e.g. in case of network shortage). In such cases snapshoting of VMs will be an important tool for administrators in order to save the state of the site and bring it back quickly and safely once the downtime conditions have been raised.

Throughout the life-cycle of a grid infrastructure the services will have to be periodically upgraded when new versions of the grid middleware are released. These new releases can be either bug fixes, improvements to the existing functionality or they may introduce new functionality and components. The urgency for applying this new versions are defined per case based on the nature of the upgrade. For example bug fixes to critical security problems in one of the core services (e.g. CE) are expected to be applied by grid sites as soon as possible otherwise the site is subject for suspension by the central grid operations. The problem with these upgrades is that sometimes they introduce incompatibilities and in some rare cases they may bring down some integral part of the grid site. Having access to a cloud service and operating a virtualized grid site is expected to improve significantly the administrative tasks in the above cases. In particular in the context of StratusLab the strategy that we will follow for grid service upgrades will be the following:

- All updates to grid middleware will be followed and the sites will upgrade to the latest version as soon as this is released.

- The updates impact both the deployed grid sites as well as the VM appliances the correspond to grid nodes (e.g. CE, WN etc.).

- For what concerns the grid sites it will be the responsibility of the grid administrator to run the required update commands on the running VMs. Before applying the updates, the administrator will make a snapshot of the VMs running the core site services (e.g. CE, SE, etc) in order to be able to rollback to the previous version in case the new version introduces bugs or any other incompatibilities.

- Virtual Machine images stored in the appliance repository will also be kept up to date with the evolution of grid middleware. In this case responsible for applying the updates will be the grid site administrators. Images should be appropriately tagged in the XML metadata file in order to identify the version of grid middleware that they provide. Previous versions of the software will still be available from VM images in the repo but will be marked as obsolete in order to prevent users from accidentally using them.

Grid administrators are typically notified about grid middleware updates from relevant mailing lists operated by the middleware provider (EMI [2] through EGI in our case).

In some cases a major upgrade of grid middleware may impact its interoperability with the underlying cloud infrastructure, especially for those components for which a certain level of interoperability has been developed (e.g. Accounting). In these cases, grid administrators should properly notify the cloud operations by raising this issue through the appropriate support channels. Nevertheless, such a case should be considered rare and prevented by the close interaction between StratusLab and EGI-InSPIRE/EMI projects.

4.3 Third party access to grid services

Grid sites operated by StratusLab will follow the regular policies for providing access to their resources to third party users. In particular users wishing to take advantage of a site's grid resources should belong to one of the VOs supported by the site. All users are expected to adhere to the Grid AUP and security documents defined by EGI.

New VOs may be supported after negotiation between StratusLab and the managers of the interested VO. Support of new VOs will help us test the cloud infrastructure capabilities under real conditions and stress test them by putting them through periods of high load.

The sites will also support the catch-all StratusLab VO, that has been setup to accommodate grid users that do not belong to a particular user community but would like to access and try the virtualized sites operated by the project. StratusLab VO has been established since M2 of the project and currently includes many of the project members. Details about the VO and instructions for configuring a site to support it is available from EGI's Operations Portal[1]. In general, the project will provide access to arbitrary users only for a short period of time and after the user has gone through an identification procedure in order to validate his/her identity and the purpose of accessing the grid site. For longer periods of usage, users will be expected to access the resources through a discipline- or domain-specific VOs also supported by StratusLab grid sites.

Generally speaking, from the point of view of the cloud provider the grid administrator is the user of the IaaS service, thus the grid administrator should sign the AUP of the cloud service. On the other hand it is up to the grid administrator to decide which VOs his/her site will support and with what fraction of resources. The grid admin may also decide whether or not to expand the capabilities of the site by adding more virtual WNs or storage space. Since during the lifetime of the

[1] http://cic.gridops.org/index.php?section=home&page=volist&vo=2252

project it is expected that StratusLab will give access to external grid administrator in order to setup third party virtualized grid sites the above approach will be followed.

4.4 Security

Access to external users is provided only to entities registered in a supported VO, holding a valid digital certificate issued by an IGTF accredited CA. New users are enrolled with the responsibility of the VO managers. Users are responsible for the security and confidentiality of their applications and data.

Security incidents will be reported following the regular EGI incident reporting procedures. In particular we will take advantage of the cloud capabilities and in case of security breaches grid sites will be brought back to operation using grid appliances or VM snapshots taken from the site in the past. The breached VMs will be kept as separate snapshots in order to help the required investigation procedure.

4.5 Monitoring and Accounting

Currently, certified grid sites are monitored using various mechanisms. StratusLab sites will implement the monitoring mechanisms and requirements defined by EGI-InSPIRE. These include the usage of the at least following monitoring tools:

- NAGIOS [4] through the submission of hourly SAM tests.
- GStat Monitor http://gstat-prod.cern.ch through the information collected from the GIIS service running on grid sites
- GOC Database https://goc.gridops.org/ which holds relatively static information for sites, provided by the grid administrators themselves.

It should be noted that all the above tools are hosted and operated centrally on an EGI-wide or NGI-wide basis thus no special requirements are imposed to the resource centers, other than installing the necessary client part or enable the submission of SAM jobs in the case of NAGIOS.

For what concerns usage accounting, StratusLab sites will adopt the tools and procedures applied in EGI. Accounting is one of the areas that StratusLab will collaborate closely with EGI-InSPIRE and EMI in order to expand the existing mechanisms in order to take into account any special requirements imposed by the provision of virtualized grid resources over cloud computing sites. Grid sites operated by StratusLab will be one of the first adopters of any new solutions coming out from this collaboration.

4.6 Quality of Service

The Quality of Service requirements for production grid sites are defined by the EGI-InSPIRE OLA document. Grid sites operated in the context of StratusLab will comply to the above document. These requirements among others define the

Table 4.1: *Target availability and reliability values for StratusLab grid sites*

	Year 1	Year 2
Site Availability	$\geq 80\%$	$\geq 95\%$
Site Reliability	$\geq 80\%$	$\geq 95\%$

availability of grid services, the announcement of downtimes from grid site administrators and the maximum allowed time to respond to operational tickets opened in the GGUS system (https://gus.fzk.de). In particular regarding the site availability, EGI OLA requires a minimum site availability of 70% and minimum site reliability of 75%. In the previous, availability is measured over 24 hours and reliability is defined using the following equation:

$$Reliability = \frac{Availability}{(Availability + UnscheduledDowntime)} \tag{4.1}$$

Since we expect that the usage of cloud computing will significantly improve the stability of the service StratusLab has set a higher goal for the above metrics for the first and second year of the project. These goals are summarized in Table 4.1.

Lastly, regarding site downtimes, these will be announced at least 5 days before the event unless there is some urgent reason (e.g. security breach) requires to bring the site off-line immediately. The tracking of SLA conformance will be done on a monthly basis.

5 Appliance Repository

5.1 Operations Overview

The initial prototype of the virtual appliance repository is deployed as a WebDAV-enabled Apache webserver located at TCD. A backup server has been provided by GRNET using a simple mirroring scheme. The appliance repository mirror is accessible from http://appmirror-grnet.stratuslab.eu. The contents of the primary repository are currently mirrored once every day using a regular Unix cronjob. The mirroring period maybe increased in the future depending on the traffic experienced on the primary repository and the number of images registered every day.

StratusLab v0.1 was released with three fully contextualized reference images for the following operating systems:

- Ubuntu
- CentOS
- ttylinux (a minimal Linux operating system useful for testing)

The goal of the prototype repository was to quickly and simply provide centralised storage that could be used by the project to share images. It is intended in subsequent releases that the repository will evolve to become an appliance *metadata* repository. Storage of appliances will be handled by the appliance provider, and only the metadata associated with the appliance will be stored on the central repository. The metadata will conform to a standard schema, and will be signed by the provider for security. Integration with the OpenNebula image repository located at sites will also be considered.

Currently users interact with the repository through either a simple webpage, or command line tools. As part of the move from storage of appliances to metadata registry it is planned that the web interface will become an appliance 'market-place'. Users will be able to share appliances and interact through comments and appliance ratings. It is hoped that this will encourage the creation and sharing of appliances by the community.

5.2 Maintenance and Upgrades

The repository webserver is hosted on a virtual machine, managed using Quattor [7]. Security updates are applied as required.

5.3 Third party access

The images provided through the appliance repository are intended to be freely available. Users can access the images using the StratusLab command-line tools, or by using standard tools such as *cURL* [3]. A webpage available at http://appliances.stratuslab.eu shows the images currently available and also displays information about the images.

5.4 Security

Write access to the repository is currently restricted to StratusLab project members. The web server is accessed using WebDAV (Web-based Distributed Authoring and Versioning), with authentication via the StatusLab LDAP server.

5.5 Quality of Service

The appliance repository is critical to the availability of the StratusLab infrastructure. As such it is closely monitored. Also, as mentioned above, a second mirror repository has been provided at GRNET. This is intended to act as a backup should the TCD repository be unavailable for any reason. The selection of the target repository is done manually through the StratusLab user command line tools. Optimally the tools should be ”smart” enough to get information for alternative repositories and try out different mirrors in case the primary repository is not available. This is a functionality that will be investigated for implementation in the next months of the project. Additional mirrors may be deployed by anyone who wishes to increase the availability and accessibility of the images. The installation of a appliance mirror is done through the *stratus-install* command included stratuslab admin package which is available for download from the project's package repository. Instructions for setting up an appliances mirror is available from the projects web site: http://www.stratuslab.org/doku.php?id=tutorial:manualinstall:appliancerepository

6 Conclusions

In this document we presented tools, procedures and policies that aid and/or define the operation of three important services of the project, namely the public cloud service, the virtualized grid sites and the appliances repository. Some of the solutions presented are already in place and used by WP5 on daily basis. Others, and especially those related to grid services, are still planned and this document serves as a guideline and requirements definition for their implementation.

Infrastructure operations is typically a multifaceted activity that requires the coordination of many actors. Achieving a sustainable Quality of Service from the provisioned service requires proper planning and careful monitoring of the established practices.

In the past years we have witnessed the development of a significant know-how regarding the operation of large-scale multinational grid infrastructures. Many of the StratusLab partners have participated in these efforts and have gained important experience in these aspect of infrastructure provision. Nevertheless, the introduction of could computing services introduces new challenges and potential hurdles. Deployment and provision of cloud computing services either private or public still remains an open issue since it has introduced new demands and requirements from resource providers. The combination of the above technologies poses many interesting problems the solving of which is expected to have important impact on how grid computing infrastructures are currently operated.

This document presented a number of good practices and useful tools that will help the project partners, having the role of resource providers, to pursue the provision of combined cloud and grid computing services based on a production level QoS. For what concerns grid sites the project will follow the established practices which are monitored by EGI. The QoS requirements of Grid sites define to a large percent the respective requirements from the cloud services. Some of these requirements will necessitate a close collaboration between EGI, EMI and StratusLab. Others that can be isolated only at the cloud layer are the main goal of StratusLab's program of work.

The defined tools and policies are expected to evolve during the lifetime of the project as more practical experience is gained and as third-parties will start using this infrastructure regularly for real applications. The first year and final report of WP5 will provide and update and final definition respectively, of the tools and practices established and will follow the evolution of the provided infrastructure.

Glossary

Appliance	Virtual machine containing preconfigured software or services
Appliance Repository	Repository of existing appliances
API	Application Programming Interface
BDII	Berkeley Database Information Index
CA	Certification Authority
CE	Computing Element
DCI	Distributed Computing Infrastructure
EGEE	Enabling Grids for E-sciencE
EGI	European Grid Infrastructure
EGI-InSPIRE	EGI Integrated Sustainable Pan-European Infrastructure for Researchers in Europe. The European funded project aimed to establish the sustainable EGI
Front-End	OpenNebula server machine, which hosts the VM manager
IGTF	International Grid Trust Federation
Instance	see Virtual Machine / VM
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
SAM	Service Availability Monitoring
SE	Storage Element
SSH	Secure SHell
TB	Terabyte(s)
Virtual Machine / VM	Running and virtualized operating system
VLAN	Virtual Local Area Network
VM	Virtual Machine
VO	Virtual Organization
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] EGI. European Grid Initiative. http://www.egi.eu.

[2] EMI. European Middleware Initiative. http://www.eu-emi.eu.

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

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

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

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

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

[8] The StratusLab consortium. Deliverable D4.1 - Reference Architecture for StratusLab Toolkit 1.0. http://www.stratuslab.org/lib/exe/fetch.php?media=documents:stratuslab-d4.1-v1.0.pdf, 2010.

[9] The StratusLab consortium. Deliverable D5.1 - Infrastructure Operations. http://www.stratuslab.org/lib/exe/fetch.php?media=documents:stratuslab-d5.1-v1.0.pdf, 2010.

[10] The StratusLab consortium. Deliverable D6.1 - Cloud-like Management of Grid Sites 1.0 Design Report. http://www.stratuslab.org/lib/exe/fetch.php?media=documents:stratuslab-d6.1-v1.0.pdf, 2010.