Enhancing Grid Infrastructures with
Virtualization and Cloud Technologies

Project Quarterly Report

Quarterly Report QR7 (V1.0)
20 March 2012

Grant Agreement Number	INFSO-RI-261552
Project acronym	StratusLab
Project title	Enhancing Grid Infrastructures with Virtualization and Cloud Technologies
Funding Scheme	CP/CSA
Date of latest version of Annex I against which the assessment will be made	2010-05-31
Periodic report	1st
Period covered	2010-06-01 to 2011-05-31
Name, title and organisation of the scientific representative of the project's coordinator	Dr. Charles Loomis, Research Engineer, Centre National de la Recherche Scientifique (CNRS)
Tel	+33 (0)1 64 46 89 10
Fax	+33 (0)1 69 07 94 04
E-mail	loomis@lal.in2p3.fr
Project website address	http://stratuslab.eu

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.

Contributors

Name	Partner	Sections
Marc-Elian Bégin	SixSq	WP3, WP4
Christophe Blanchet	CNRS-IBCP	WP2, WP3
Kathryn Cassidy	TCD	WP3
Vangelis Floros	GRNET	WP3, WP5, WP6
Javier Fontan	UCM	WP4
Eduardo Huedo	UCM	WP6
Stuart Kenny	TCD	WP5
Ignacio M. Llorente	UCM	WP3
Charles Loomis	CNRS-LAL	WP2, WP3, WP5, Summary, Mgt.
Louise Merifield	SixSq	WP4
Ruben S. Montero	UCM	WP6
Henar Muñoz	TID	WP2, WP4, WP6
David O'Callaghan	TCD	WP3

Document History

Version	Date	Comment
0.1	8 Mar. 2012	Initial outline.
1.0	19 Mar. 2012	Final version.

Contents

List of Figures

List of Tables

1 Publishable Summary

1.1 Project Context and Objectives

The StratusLab project is aimed at service provisioning, networking, and research of technologies that will bridge cloud and grid infrastructures to simplify and optimize the use and operation of existing distributed computing infrastructures (e.g. European Grid Infrastructure) and to provide a more flexible, dynamic computing environment for scientists.

The European production grid infrastructure has had many notable successes. It has allowed scientists from all over Europe and indeed from all over the world to federate their computing resources to advance their scientific aims. More importantly, the infrastructure allows them to federate their data and expertise to accomplish more than they would be able to do singlehandedly. Common APIs and service interfaces make it possible to take advantage of these distributed resources without having to modify applications for each site.

Despite its success, the grid also has its limitations. The uniformity of service interfaces unfortunately does not extend to the underlying computing resources, where users are exposed to significant heterogeneities in the computing environment, complicating applications and increasing failure rates. Passive calculations are handled well by the grid, but many applications require active services to coordinate the distributed analyses. Either scientists must provide their own resources for such services or negotiate with a particular site to provide them. This reduces the speed at which new calculations can be done.

Virtualization technologies provide a mechanism for offering customized, uniform environments for users with negligible performance degradation. Using grid technologies combined with virtualization allows the grid to provide users with a homogeneous computing environment, simplifying applications and reducing failures. Emerging cloud technologies allow users to dynamically allocate computing resources (often in less than a minute) and to specify the characteristics for the allocated resources. The fusion of cloud and grid technologies provides a more dynamic and flexible computing environment for grid application developers.

Cloud and virtualization technologies also offer other benefits to administrators of resource centers, such as the migration of live services for load balancing or the deployment of redundant servers. Reduced costs for managing resources immediately benefit users by freeing money for additional computing resources or

by having better user support from administrators.

A combined computing infrastructure that uses grid technology's strengths for federating resources, virtualization's strengths in providing custom, uniform environments, and the cloud's strengths in dynamic resource allocation, maximizes the utility of European distributed computing resources to scientists.

The StratusLab project creates an complete, coherent, open-source private cloud distribution to allow administrators of grid resources centers to take advantage of virtualization and cloud technologies. It provides new ways of using existing distributed computing resources to make the infrastructure more adaptable and more useful for scientists.

1.2 Summary of Work Performed and Achievements

In the sixth quarter, the project worked to enhance the existing production release and to update the architecture for the second major release due at the end of the project. Unfortunately, a major cooling failure at LAL seriously perturbed the build and test infrastructure and consequently also the ability of the project to make incremental releases of the StratusLab cloud distribution. Nonetheless, one incremental release was made that was used to support a number of tutorials. Progress was made on all of the quarter's objectives, although many were not fully completed because of the above problem. The points below summarize the progress made.

Incremental Releases of the StratusLab Cloud Distribution Version 1.1 of the StratusLab cloud distribution was made publicly available. This version was quickly installed on the project's reference cloud infrastructure and on the LAL's test infrastructure. Two additional planned incremental releases unfortunately did not take place because of the perturbations to the build and test infrastructure described above.

Implementation of Use Cases In the previous quarter, seven use cases have been defined on which to focus the project's porting and support efforts. Work on three of these–a bioinformatics application, a commercial 3-tier prototype service, and a software development PaaS–have significantly advanced, although not yet to the stage where they have concrete results and can be publicized. All of these are expected to provide dissemination opportunities in the next quarter.

Support for OpenSuSE Support for a second operating system by the StratusLab cloud distribution is desired for a couple of reasons. First, it demonstrates portability of the code and services. Second, support of OpenSuSE will allow us to test GPFS, which should have much better performance than NFS, as a shared file system for distributing images on the cloud infrastructure. Although modifications to the build system have been made to facilitate this, support for OpenSuSE has not yet been demonstrated.

Tutorials and Training Materials A major effort and success for this quarter has been a series of well-attended tutorials. These took place as part of the EGI Tech-

nical Forum in Lyon, France; as part of the ACGRID 3 school in Hanoi, Vietnam; and as a standalone event in Orsay, France. Overall, around 90 people participated in these events. The materials developed for these tutorials will continue to be updated for future releases of the distribution.

Restructured Communication Channels The communication channels between the project and the targeted communities have been somewhat restructured. In October an additional "user announcements" (announcement only) mailing list was created, specifically for users of the StratusLab reference infrastructure to inform them of upcoming service outages, upgrades and other important news. A new public user forum was created and went live on 16 November, and has been running on a trial basis so far. The aim of the forum is to provide a contact point for StratusLab users and to build a community support mechanism. The group already has 12 members, and has seen seven messages posted so far in two topics. It will thus be officially launched in December 2011.

Updated Architecture The quarter began with a redefinition of the global architecture of the StratusLab cloud distribution. The major changes were the additions of monitoring services, accounting services, and an "inter-cloud connector". The designs of components for advanced services and functionalities have also been updated, providing a clear roadmap for the second year developments.

Improved Image Management A major integration effort centered on the persistent storage service. The image management code was significantly refactored to take advantage of this service and to allow efficient caching of virtual machine images. With these changes, the primary delay for starting images is now the latencies associated with the scheduling procedure. The "create image" feature has also been refactored to take advantage of this service, drastically reducing the time necessary to create a new image.

Streamlined Release Procedures In order to better streamline the release and certification of new software releases, new procedures have been put in place, supported by changed in our continuous integration system. This should allow future releases to be done "with the click of a button" with nearly all of the test and certification tasks done automatically.

Reference Configuration and Performance Benchmarks Although an important topic, little time was spent on this in Q6 because it was more important to get the build and test system working correctly and refactored to support more fluid release procedures.

Implementation of Advanced Networking Services Work was done to provide advanced networking services such as the dynamic provisioning of VLANs and configuration of firewalls. These features are part of the OpenNebula 3.0 release and is in the process of being integrated into the mainline StratusLab distribution. Integration problems on both the StratusLab and OpenNebula sides are being resolved and these new features will appear in the StratusLab distribution once OpenNebula 3.0 has been integrated.

Sustainability As we reach the end of the StratusLab project, sustainability is an important question for the users of our software. It is a question that comes up at nearly every StratusLab presentation. The face-to-face meeting at the end of the quarter allowed a detailed discussion of the sustainability strategy that will serve as a good basis for the final plans that will be developed over the last six months of the project.

Despite the problems with the build and test infrastructure's availability, significant progress was made on all of the objectives for this quarter. Those that have not been completed will be moved forward into the next quarter. More emphasis will be placed on supporting a second operating system as this has been pushed forward two quarters now. The next quarter will see the beta of the 2.0 release with prototypes of all of the expected services.

1.3 Final Results and Potential Impact and Use

Most scientific and engineering research requires significant computing resources. Distributed computing infrastructures have brought unprecedented computational power to a wide range of scientific domains. Although, these architectures and the related software tools have been considerably improved over the years, they exhibit several difficulties, mainly due to limitations of physical platforms, which discourage adoption of grid technologies. StratusLab has the potential to profoundly change existing grid infrastructures.

1.3.1 Improved Interdisciplinary Scientific Collaboration

Cloud technologies are expected to have significant impact, both immediate and long-term, in the way scientific research is carried out. Grid infrastructures have provided a remarkable advantage over the past years offering access to vast amount of computing power and storage space, and most importantly by offering a sustainable platform for scientific collaboration enabling the sharing of computing resources and scientific data. Cloud computing is expected to take this one step further by facilitating the easy deployment of customized grid infrastructures. These infrastructures are expected to have further positive impact on the way interdisciplinary scientific research is taking place.

StratusLab focuses on the provision of scientific infrastructures over cloud computing, investigating in particular the provision of customized Virtual Machine images. This customization will be done on the user side, which means that the user can have more immediate influence on the infrastructure itself. In this way the infrastructure will adapt to the user requirements and not vice-versa. By easing the management of grid sites and the configuration of hosting services we expect to attract a broader number of scientific communities and further facilitate their collaboration.

Table 1.1: *StratusLab Information and Support*

Website	http://stratuslab.eu/
RSS Feed	feed://stratuslab.eu/feed.php?ns=news&linkto=page
Twitter	@StratusLab
YouTube	http://www.youtube.com/user/StratusLab
Support	support@stratuslab.eu

1.3.2 Impact on DCI Evolution

Currently, there is a big shift in all e-Infrastructure projects, and related efforts in Europe, to expand their activities in order to include cloud computing technologies. StratusLab will play a key role in this landscape by providing a focused environment for development, deployment and experimentation of cloud computing services.

The projects proposal reflects an evolutionary path from the existing large-scale monolithic grid e-Infrastructures to novel, beyond the state-of-the-art, cloud-based, grid-enabled ones. Through its expected collaborations with other projects, StratusLab will disseminate its findings and drive direct impact on the way e-Infrastructure provision is currently done.

1.3.3 Improved Usability of DCI Platforms

Virtualization is the cornerstone of cloud computing and a key for achieving optimal usability of DCI platforms. Moreover, virtualized environments have the ability to adapt to different hardware platforms enabling a quick transition from one environment to another.

StratusLab operates such a virtualized platform on a variety of hardware environments. By offering customized machine images, users will be able to set-up an environment that better suits their application requirements. This will dramatically improve the current situation where current infrastructures are forced to offer a common configuration–a common denominator–that tries to do its best to satisfy many users with different runtime requirements. Another aspect where StratusLab will contribute is on power consumption efficiency (Green Computing) and the increase reliability by incorporating failover mechanisms using virtual machine snapshots and migration.

1.4 Contact Information

More information about the StratusLab project can be obtained from the sources listed in Table 1.1. Individual partners can also be contacted to obtain more specific information about their contributions to the project. Table 1.2 contains the list of StratusLab partners and relevant contacts.

Table 1.2: *StratusLab Partners*

CNRS	Centre National de la Recherche Scientifique	Charles LOOMIS loomis@lal.in2p3.fr
UCM	Universidad Complutense de Madrid	Ignacio LLORENTE llorente@dacya.ucm.es
GRNET	Greek Research and Technology Network S.A.	Evangelos FLOROS efloros@grnet.gr
SIXSQ	SixSq Sàrl	Marc-Elian BEGIN meb@sixsq.com
TID	Telefónica Investigación y Desarrollo SA	Henar MUÑOZ henar@tid.es
TCD	The Provost Fellows and Scholars of the College of the Holy and Undivided Trinity of Queen Elizabeth Near Dublin	David O'Callaghan david.ocallaghan@cs.tcd.ie

2 Project Objectives for the Period

2.1 Objectives

The primary objective of the project is to provide a software distribution that brings together cloud and grid technologies to benefit both grid resource center administrators and scientists. In order to achieve this main objective, we have defined a set of interrelated objectives to be addressed in the project. The objectives are organized, for clarity of exposition, into three groups of objectives, corresponding to networking, service and research activities (see Figure 2.1):

- The first group represents coordination and networking with users and other stakeholders in the grid and cloud ecosystems. The project will work directly with scientists using the grid to ensure that the distribution satisfies real needs; and will collaborate with related projects and contribute to standards bodies.
- The second group represents infrastructure related services to the scientific community. The project will integrate and maintain a software distribution to bring cloud to existing and new grid sites and will ensure the production quality of the distribution by running two production sites with the distribution.
- The last group represents innovation and exploration of new cloud and virtualization technologies to enhance grid infrastructures. The project will develop innovative technology for cloud-like management of grid services and resources that will be incorporated into the software distribution.

These objectives are presented by work package below. Similarly, the work program is built around these objectives. There is a one-one correspondence between objectives and activities, so facilitating an easy cross-reference between objectives and activities throughout this document, and their verification during the project execution. The activity on project coordination has not been included here.

2.1.1 WP2: Interaction with Users and Related Communities

StratusLab targets two distinct communities: resource providers and end-users. The StratusLab software will simplify grid site administration and improve the reliability of the site. Later releases in the second phase of the project will provide

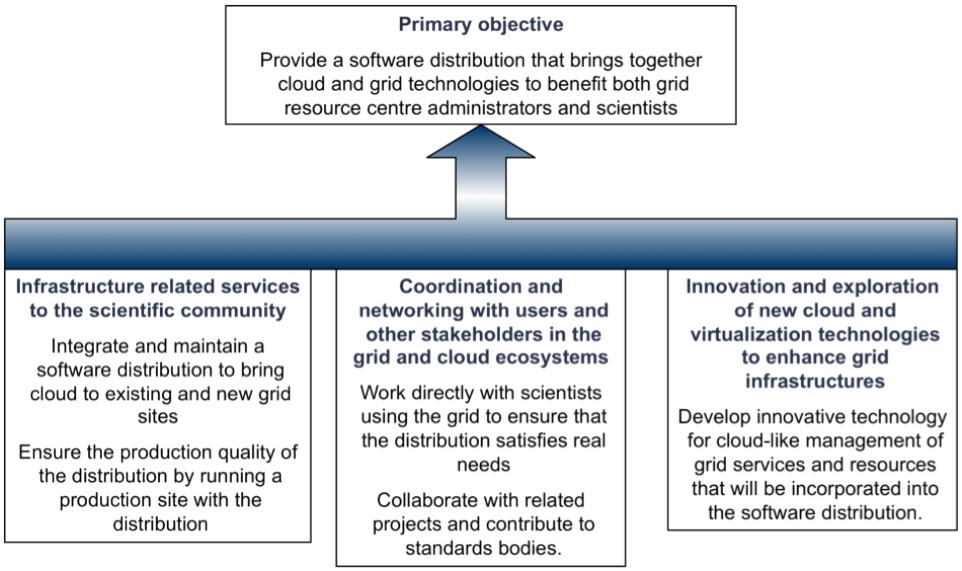

Figure 2.1: *Primary and supporting objectives*

direct cloud APIs that will be attractive for scientists porting applications to the grid. The communications between these communities and the project must be managed to ensure the project fully addresses their needs and any problems that arise. One community will work directly with the project to evaluate early releases of the software. Results of the project must be disseminated as widely as possible to those two communities as well as the general public. Scope of the objective.

- Manage communication with resource providers regarding their needs concerning virtualization and cloud technologies and their feedback on StratusLab software.
- Manage communication with end-users regarding their use of resources running StratusLab software and their needs for direct access to virtualization and cloud features.
- Training sessions will be organized to encourage dissemination of technical information and adoption of the StratusLab software.
- Evaluate early versions of StratusLab software from a users perspective with respect to utility and stability.

2.1.2 WP3: Dissemination

A large number of projects, companies, and standards bodies currently focus on cloud and virtualization technologies because of their promise and growing adoption. StratusLab must actively engage with those entities to ensure that the projects

results are well represented, that we are aware of others advances, and that we drive standardization in a direction consistent with our vision. Scope of the objective.

- Disseminate results of the project to resource providers, end-users, and the general public.
- Identify project contributions to standards bodies and standardization efforts.
- Coordinate interactions with related projects, developing Memoranda of Understanding between projects where appropriate.

2.1.3 WP4: Integration, Distribution and Support of Open-Source Cloud Distribution

StratusLab will integrate and support an open-source cloud distribution enabling grid site virtualization and dynamic scaling to remote clouds. This distribution will address the specific requirements of the grid resource providers and enable the deployment of science clouds, as well as addressing infrastructure cloud-like access demands from user communities, including industrial users. Scope of the objective. StratusLab will address the following topics:

- Selection of software components, from best of breed in open source software, to compose a robust and industry grade open source StratusLab toolkit. This distribution will integrate with typical administration tools in grid infrastructures and fabric management. This process will be driven by real needs and constraints in production infrastructures.
- Integration and management of open-source distribution, definition and maintenance of reference configurations and sustainability in the context of EGI and its official middleware distribution. The StratusLab toolkit will integrate the innovation developed in the research activity.
- Technical support for installation and configuration of the distribution, following industrial practices in term of quality, maintainability, testability and usability
- Definition of a process for automatic configuration of the virtual appliances

2.1.4 WP5: Operation of a Production Grid Site Running StratusLab

StratusLab will engage two resource centers that will be responsible for the deployment of middleware and tools developed in the project. One the main tasks of these resource centers will be the operation of two production grid sites running StratusLab toolkit. The sites should be able to pass the certification procedures imposed by EGI. The activity will demonstrate the security, performance, reliability and scalability of the distribution, and will provide support for the creation of

the virtual appliances for different user communities. The activity will also investigate the feasibility of offering a repository of reference images for cloud users, with demonstrated interoperability among the supported cloud infrastructures (including the private cloud deployed in the re-source centers, as well as a selected number of public clouds). Scope of the objective. StratusLab will address the following topics:

- Deployment and operation of virtualized grid sites
- Testbed for the StratusLab toolkit
- Support for the creation of virtual appliances for different user communities.

2.1.5 WP6: Innovative Cloud-like Management of Grid Services and Resources

StratusLab will conduct research on grid service automatic deployment and dynamic provision, including automatic elasticity mechanisms for scaling up and down to meet performance goals (typically defined by SLAs). StratusLab will also conduct research on novel infrastructure cloud-like resource provisioning paradigms, and dynamic and scalable management of virtualized infrastructures for grid services. The research will be performed to address technology gaps defined by the service activities according to user requirements collected by the networking activities. Scope of the objective. StratusLab will address the following topics:

- Framework for grid service elasticity and dynamic pro-vision of grid services
- Grid specific virtual machine management techniques
- Infrastructure cloud interfaces for grid sites and its integration with existing Grid services

2.2 Detailed Objectives by Quarter

2.2.1 Quarter 5

- Solidify the v1.0 StratusLab cloud distribution through increased testing and hardening of existing services.
- Support for a second operating system to ensure the portability of the distribution.
- Survey of the users and system administrators to see if the requirements have evolved from those already collected in Y1.
- Update and expand the target reference architecture for the distribution.
- Continued dissemination of project results.

- Continued operation of reference infrastructure and support to users and system administrators.
- Expansion of the number of users and sites using StratusLab.

2.2.2 Quarter 6

- Release incremental production versions of the StratusLab cloud distribution.
- Support for a second operating system to ensure the portability of the distribution.
- Implementation of an identified use case.
- Provision of tutorials for finding and training new users.
- Update and expand the target reference architecture for the distribution.
- Dissemination of project results with emphasis on the general public.
- Definition of reference cloud configurations and implementation of performance benchmarks.
- Improved integration of image management and caching.
- Implementation of more advanced networking services (e.g. dynamic firewalls).

2.2.3 Quarter 7

- Solidify the v1.0 StratusLab cloud distribution through increased testing and hardening of existing services.
- Support for a second operating system to ensure the portability of the distribution.
- Produce initial beta release of v2.0 of the StratusLab cloud distribution.
- Dissemination of project results of a completed use case.
- Expanding the sites using StratusLab and completion of MS4.
- Continued operation of reference infrastructure and support to users and system administrators.
- Operate a pre-production elastic Grid site elasticity, verify the applicability of the technology and move results to a production site.
- Evaluate GPFS as a backend storage solution. Prioritize and evaluate additional file systems.

- Develop additional use cases similar to MapReduce (e.g. Matlab application showcase)
- Integrate caching sub-system within the production cloud service
- Integrate NFS persistent storage service in the reference cloud service
- Integration of OpenNebula 3.0 with initial tests of new networking functionality

2.2.4 Quarter 8

- Release of v2.0 of the StratusLab distribution for multiple operating systems
- Detailed definition and initial implementation of sustainability plan
- Demonstrate of hybrid cloud computing
- Measurement of the behavior of the system in terms of scalability and IO performance, including with alternate configurations (e.g. GPFS)
- Demonstrate a pre-production elastic Grid site
- Complete migration of gLite production Grid site to UMD-1

2.3 Review Recommendations

The responses given below to the reviewer recommendations have been updated to reflect the situation at the end of Q6.

1. *Due to unscheduled availability of a physical production infrastructure experiments, testing and debugging have been affected. This was caused by problematic financial issues at the start of the project. The project should make provisions that such events are better mitigated in the future.*

   (See next response.)

2. *Several WP have shown a "delayed" start due to various hiring issues at the beginning of the project. A better process should be put in place.*

   These delays were largely related to the difficulty of hiring new personnel in the summer months and to lengthy administrative procedures. Most of the partners compensated to some extent with effort from permanent staff. Currently, all of the partners are fully staffed and no further perturbations are expected.

3. *The dissemination work is focussed towards the more technically oriented (system administrators) communities rather than the scientific user groups who could benefit from StratusLab. This is acceptable for year 1 but the focus should shift to the scientific users in year 2.*

The project agrees that this is a reasonable shift in focus for the dissemination work in year 2. The project proposed a demo for the EU Innovation Convention, which unfortunately was not accepted. However, work continues on a video demo for the general public. More effort will be put into creating general dissemination materials and demos, both of which will require effort from all activities within the project.

4. *The dissemination targets as mentioned by the related KPI metrics were not very ambitious. The project should establish more ambitious KPI metrics with respect to dissemination targets.*

   The project management along with the activity managers are in the process of redefining the metrics and the targets for the second year of the project. More ambitious dissemination targets will be proposed.

5. *A clear and simple demonstration showing the benefits rather than the technology itself should be produced. This benefits should be illustrated focusing on one or two clear use cases. The solutions should clearly demonstrate what the real benefits are (for Scientific Users / System administrators) of the StratusLab toolkit.*

   A major part of the D2.3 deliverable was to identify possible use cases and would serve as a good basis for demonstrating the benefits of cloud technologies and as good topics for focused dissemination efforts. Work on these use cases are advancing and we expect that in Q7 the first dissemination activities related to them will appear.

6. *Provide a clear map of the components of the toolkit. Which components are re-used, which are newly developed and which are adapted from existing components?*

   Deliverable D4.4 that provides the reference architecture for StratusLab 2.0 was structured to provide a separate description for each service. That description includes details on where the component was developed, external dependencies, and interactions with other components. This should provide a clearer overview of the development activities of the project. In addition, more effort has been made in the deliverables to describe clearly the project's work with respect to individual components.

7. *The periodic report is in draft status. Please submit a final version.*

   The final version of the periodic report for year 1 has been submitted. The delay was due the difficulty in getting the necessary financial information from the partner's institutes during the summer months. Through the PMB, the partners have been advised that this information must be provided in a more timely manner for the following periods. Forcing partners to provide the information quarterly will also help in preparing the final information at the end of period 2.

8. *Knowing year 1 budget under spending, a new forecast of the planned spending for year 2 including a recovery plan for the current under spending should be presented to the EC as soon as possible. A reasonable deadline is one month after the receipt of this review report. It should take into account the reasons of year 1 under spending and should introduce measures that allow the effective implementation of the recommendations for year 2.*

   Getting complete budget information from the partner's administrations has proven to be extremely difficult. It is now clear that several partners will significantly underspend on their allocated budget. A study is currently being done to determine what the status will be at the end of the project, but it is unlikely that changes can now be done to dramatically alter the situation.

9. *The Data Management layer should be improved. In particular, StratusLab should be able to use existing and robust parallel file-systems which have better scalability than NFS such as Panasas or GPFS.*

   This recommendation refers to the use of shared file systems to make machine images available to the various computational hosts of the cloud infrastructure. Alternate technologies such as iSCSI that do not rely on a shared file system have been investigated and used successfully. Nonetheless, shared file systems provide a convenient alternative. Consequently, tests of GPFS will be done at LAL to gauge its performance when used as a machine image cache. Furthermore, additional distributed and parallel file systems will be evaluated (e.g. PVFS) but also file systems that have already been tested (e.g. GlusterFS) will be re-visited for further investigation and scrutiny.

10. *Testing and benchmarking in WP5 should be more detailed including performance aspects.*

    A testing plan will be developed that defines various reference cloud configurations and measures of performance. The existing application benchmarks will be used as a starting point and then expanded as necessary. To the largest extent possible, these tests will be automated through the existing Hudson continuous integration infrastructure.

11. *More emphasis sould be put on the Cloud API rather than the GRID.*

    By moving work on the Cloud API from the second year of the project to the first year already shows that the project believes that cloud-like access to resources is important. The project has decided to support OCCI 1.1 and CDMI as formal standards in addition to the existing proprietary interfaces. The highest priority, however, will be a jClouds plugin as this has been the most requested feature related to service interfaces.

12. *Although security issues are taken very seriously, privacy issues should be taken seriously as well. For instance, in case when a "closed" Grid infras-*

tructure is complemented/bridged with an external public Cloud infrastructure when it is processing for instance medical sensitive information that can be relayed back to an individual person.

Our feeling is that StratusLab should remain a "neutral" carrier that allows users and administrators to implement their own mechanisms for ensuring privacy and confidentiality. Nonetheless, we see areas in which the services can improve to provide, for instance, better logging and auditing information that can complement user-level strategies for privacy and confidentiality. Those areas will be targeted for improvement.

13. *The security incident as reported in Q3 should be analysed thoroughly and measurements should be taken to prevent this to happen again on the live production system.*

    These security incidents were taken seriously and analysed thoroughly. As a result of these incidents two additional features were added to the StratusLab distribution: 1) enhanced logging and 2) image policy enforcement. The enhanced logging makes it easier to trace the characteristics, ownership, and history of a particular machine image aiding forensic analysis. The distribution now also contains a policy enforcement engine that allows system administrators to define what images are authorized to run on a given infrastructure. This policy enforcement mechanism is closely tied to the information provided in the Marketplace. Cloud administrators will still need to monitor the cloud for suspicious activity and take corrective actions as necessary. From the operations point of view (WP5) we will remain alerted for potential future security incidents and will be ready to respond quickly in coordination with other infrastructure operations teams (NGI NOC, EGI CSIRT team etc.).

14. *The project should clearly define a small number of use cases and focus the project towards delivering real value to these user communities, targeting system administrators as well as scientific users. As also indicated later in section 5 "Use and dissemination of foreground" the project should adopt an attitude of someone who tries to offer services to the market and must convince someone to spend funds for the services. This way the project can maximize impact and the work can contribute to sustainability.*

    The deliverable D2.3 defines a set of seven initial use cases. These will be scheduled for implementation over the second year of the project. A reorientation of the dissemination activities to provide better "marketing" is being planned by WP3 and will be supported by WP2 and the other activities.

15. *A person should be appointed in the Project Management Board that can help the project to move from technology towards real end user solutions and benefits. This person should also help the project to establish solid relation-*

ships with stakeholders outside the traditional high energy physics (HEP) community.

The project agrees that having a "user champion" within the project is a good idea. However, the PMB is not the proper body as it meets infrequently and isn't in direct contact with the daily technical advances of the project. In the end, a viable solution for incorporating an outside representative has not been found. Instead we have relied on input from the scientists, engineers, and administrators with whom the project already has contact. In addition, commercial contacts through the partners have proven invaluable for validating the project's software and indicating new directions for improvement.

16. *While it is clear what StratusLab could offer to the scientific community, the impact of StratusLab would be much bigger when the toolkit could also be used for users in the commercial world. Through collaboration with the Spanish TID private Cloud project, a large number of potential additional use cases Telefonica's customers) could be developed and should be taken seriously, including dissemination towards other DCI projects.*

    Better marketing of the StratusLab distribution will include both commercial and scientific communities as described above. TID will be more involved in presenting the project's work to enterprises with several possible venues being considered.

    Recently several commercial ventures have shown interest in the StratusLab software and have created demonstrators with it. This is a significant validation of the StratusLab concept for use within the commercial sector.

17. *Dissemination of the StratusLab Toolkit should become more marketing driven and should target both ICT press to reach potential industry users as well as to potential VOs beyond the current bioinformatics users. A demo centred on a use case could help.*

    As stated in other responses, we agree that the dissemination should become more marketing driven. Deliverable D2.3 has identified initial use cases for implementation. Specific dissemination activities and possible demonstrations will be considered as each use case is sucessfully implemented.

18. *Future reports should contain less "fat" and should be more cripy and to the point. Executive summaries should be self-contained and should answer: (a) why should I read the deliverable, (b) the benefits for my company/organisation, (c) aspects addressed in this deliverable, (d) summary of recommendations/findings. The report should clearly describe if components are newly developed, improved or reused and integrated by StratusLab.*

    The project will make year 2 deliverables more concise with better executive summaries following the guidelines given above.

19. *Ensure the project periodic and final reports are available at least two weeks prior to the review meeting.*

    The review for year 2 will be scheduled to ensure that the final periodic report is available at least two weeks prior to the review. All of the deliverables for year 1 were available at least two weeks prior to the review and we intend that to also be the case for year 2.

20. *On top of the internal reports, consider publishing for the general press, like a newspaper and possibly to organise a public demonstration for the less technical audience, e.g. at the European Parliament.*

    In August, a general paper was prepared that describes the goals of the project and the StratusLab distribution. Although still technical, this gives a good overview of the project. This and other project documents will be further generalized to appeal to a wider audience. A first step in this direction was a proposal for a booth at the EU Innovation Convention in Brussels, which was unfortunately not accepted. Nonetheless, work on general materials and demos for non-technical audiences continues.

3 Progress and Achievements

In the seventh quarter, the project concentrated its efforts in four areas: improving the StratusLab distribution (through bug fixing and expanded testing), demonstrating cloud use cases with real applications, expanding the functionality of the services. The improvements and new functionality have been rolled into the incremental releases of the distribution. Two such releases (v1.2 and v1.3) have been made in this quarter, acting as beta releases for the upcoming v2.0 release at the end of the project.

Implementation of Use Cases Signficant progress has been made in running prototype and real applications on StratusLab cloud infrastructures. The most advanced are the two bioinformatics use cases–one providing bioinformatic web services and the other, TOSCANI, an application from Institut Pasteur. In addition a proof-of-concept multi-tier application has been prepared to demonstrate the applicability of the cloud to commercial use cases. Initial steps to interface StratusLab with the high-energy physics analysis framework DIRAC have been taken.

Improvements of Cloud Distribution External users from the above applications have identified many areas where the user interface can be improved. During this quarter many minor bugs have been corrected in the command line interface. As part of the bioinformatics efforts, a web-based graphical interface has been developed to provide customized access to the cloud services; this complements existing service web-based interfaces. As the pdisk server has been more closely integrated into the overall virtual machine workflow, problems with scalability and performance have been noted. Significant effort has gone into solving those issues.

Integration of OpenNebula 3.0 The integration of the more recent version of OpenNebula makes many of the new features available to users of the StratusLab distribution. These features include better network management and security features as well as more options for controlling placement of virtual machines. Integration of OpenNebula 3.2 is ongoing, but should happen quickly as there are fewer fundamental changes between these releases.

Improved Testing The releases have also been hardened with more thorough testing of existing features. The certification procedure has been improved to help ensure that fewer problems appear in production. Moreover, the procedure has been automated to make the procedure quicker and more efficient. This will further improve in the future with the possibility of scalability testing of StratusLab on the

Grid'5000 infrastructure and the validation of the release on multiple operating systems.

Expanded Functionality Development of functionality intended for the final releases of the StratusLab distribution is already well advanced. Services to provide machine and application-level monitoring have been developed and integrated with Claudia, allowing more advanced autoscaling capabilities. In addition, some work has been done on standard service interfaces/libraries, including TCloud, OCCI, and DeltaCloud.

Sustainability Detailed planning for sustaining the StratusLab developments after the project started in the previous quarter and have been further refined through participation in a Sustainability Workshop between the DCI projects in January and collaboration with the ERINA+ project. This ERINA+ collaboration will, in particular, allow us to better understand the socioeconomic impact of the project and better define the stakeholders involved.

Despite these achievements, several problems have arisen in this quarter and consequently some of the quarter's objectives have not been met. The most notable problem was with the performance and scaling issues coming from the storage service; these problems appeared as the storage service was more tightly integrated into the virtual machine management workflow. Significant effort was put into understanding and resolving these issues as quickly as possible. Consequently, the upcoming 1.4 release will provide a much improved storage service, although there are still further improvements to be made. Because of this change of focus, tests of other storage services like GPFS have not been done.

On the application side, signficant progress was made in running real applications on the cloud infrastructure. Unfortunately, these successes have not been adequately publicized. Presentations at the EGI Community Forum early in Q8 will compensate somewhat for this, but a separate special effort needs to be made to maximize our benefits from these activities.

Overall, the StratusLab distribution is advancing well with the project correcting any problems that arise and adding new functionality at it moves toward our v2.0 release in the next and final quarter of the project.

3.1 WP2: Interaction with Targeted Communities

This activity manages the relationships with the communities targeted by the project, notably scientists from diverse fields and system administrators interested in deploying a cloud infrastructure. Through those interactions, it provides requirements and feedback to the other activities within the project while at the same time evaluating the StratusLab distribution from the points-of-view of users and system administrators. The activity also helps provide support to the targeted communities.

3.1.1 Summary

The work in this quarter has been concentrated in three areas: porting of applications, usability improvements, and expanding the tests and validation of the software. These applications include two from the bioinformatics domain (TOSCANI and Bioinformatics Web Service) that allow people from outside the project evaluate the cloud paradigm and the StratusLab software for their scientific analyses. A prototype application using the standard 3-tier architecture and showing the benefits of autoscaling technique available in Claudia has been created and will be used as a concrete demonstrator of the project's software. Based on feedback from users, the command line interface has been improved; in addition, a web-based interface has been created for the bioinformatics users to simplify access. Also to improve the software, people in WP2 have concentrated on improving the functionality testing, laying the foundations for scalability testing (via the Grid'5000 platform), and providing packages for multiple operating systems to permit GPFS testing. These achievements will be presented at the EGI Community Forum at the end of March and then followed by more formal press releases.

3.1.2 Task 2.1: Interactions with Resource Providers and End-users

E-business Application Use Case The e-business application use case defined in Q6 has been developed during this quarter. The four images required: a front-end, a back-end, a balancer and a data store, have been uploaded to the appliance repository[1] and made available through the Marketplace. In addition, a set of monitoring probes are been developed to monitor the number of users and response time in the application, Key Performance Indicators (KPIs) that will drive the scalability. Finally, an OVF description has been defined for the deployment of the use case, its monitoring and scalability.

Bioinformatics Web Service Use Case New releases of the 'biodata' and 'biocompute' appliances have been built and registered in the StratusLab Marketplace. The 'biodata' appliance now contains the SwissProt and PROSITE databases; bioinformaticians can use the appliance to set up their own cloud biological databases repository. They can add the databases they want with the help of the BioMaj

[1] http://appliances.stratuslab.eu/images/comertial

system. The configuration of the databases is done through the BioMaj web interface or through the command line. This appliance can be used in conjunction with the 'biocompute' appliance to establish a biological repository for future bioinformatics analyses. The 'biocompute' appliance contains a range of tools (BLAST, ClustalW2, FastA, HMM, etc.). And the tool names can be used to apply a filter on the bioinformatics appliances at the "launch instance" step. The access to this appliance is also provided to users through SSH or through a web portal, easily accessible in the main window of the cloud web interface. Biocompute instances can mount the biological database repository of the local cloud (if available) through contextualization options. This appliance can also mount your persistent disk, at boot time if users choose via "launch instance" form.

TOSCANI Use Case We have released two versions of the ARIA appliance related to the TOSCANI use case. ARIA is developed by the group of Dr. M. Nilges at the Institut Pasteur Paris. Version 2.3 of the tool has been installed and configured, and the appliance has been registered in the Marketplace. The appliance has been evaluated by colleagues from Institut Pasteur group. They had run experiments with real data and reported that everything went fine, with good performances of the CNSsolve part. A second version of the appliance (1.1) has created with more computing (BLAS, Lapack) and OS (screen) add-ons. The user connects to the ARIA VM with SSH, uploads its data and can directly run the tool from the current directory. All the tools and libraries have been put in the environment variables. Only one instance, containing both the ARIA master and the CNSsolve tools, needs to be run by the user. The computing resources are tuned via the instance used to launch the machine, up to the maximum allowed by the "HTC" instance on the CNRS IBCP infrastructure with 24 CPUs and 16 GB RAM. Coming releases are planned which will allow an instance for the ARIA master and many instances for the CNSsolve nodes.

Specialized Cloud Web Interface for Bioinformatics CNRS IBCP has developed a new web interface for the StratusLab framework. The main goals are to make it easy to use by non-computing and non-cloud-specialist scientists like biologists and bioinformaticians, and powerful enough to perform the main tasks of the VM lifecycle on a StratusLab infrastructure. The first release has been developed and is now being evaluated by biologists and bioinformaticians. The available features allow the creation and termination of the virtual machines, allow the management of the persistent disks, and provide assistance for bioinformatics appliance selection and instance contextualization. A screenshot of the interface can be seen in Figure 3.1.

DIRAC: HEP Use Case The high-energy physics community represents an important use case for the project because its current use of grid services offers the opportunity to understand the interactions between grid and cloud services. Modifications required to run CernVM appliances (machine images created by the community) have been completed and released. A test has been added to the standard "use case" Hudson job to ensure this continues to work with future versions. As a

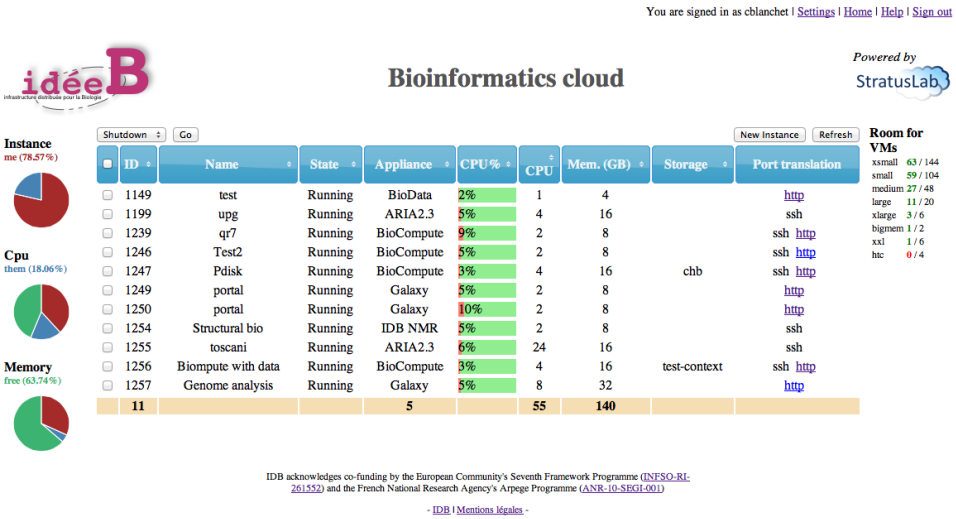

Figure 3.1: *Bioinformatics Cloud Web Interface*

follow-on to this, work has begun with the developers of DIRAC, a master-worker analysis framework used by the LHCb experiment based at CERN. The goal is the transparent integration of StratusLab resources on the cloud within the DIRAC framework, with a demonstration of production jobs being run on our infrastructure. This integration should be trivial with the availability of the CernVM images and we expect concrete results early in Q8.

Automation of Base Image Production The project has provided minimal machine images for ttylinux, Fedora, and Ubuntu for some time. However, these images have not been systematically kept up-to-date or produced in a consistent manner. Now the mechanisms for producing these base images have been automated (through Hudson) that will allow systematic production and updates. In addition, the range of supported base images has been expanded to include also OpenSuSE and CentOS. This will allow users to find high-quality images that are guaranteed to work well on StratusLab cloud infrastructures. This work was done in collaboration with WP5.

Improvements to Client Based on feedback from users at the project's tutorials a number of usability issues were identified with the StratusLab command line client. During this period a number of these issues were resolved with the client. These improvements generally provide better error messages to users and improve the overall consistency of the tools.

Multi-platform Support Support for multiple operating systems has been a long-standing request of system administrators. (The client already supports a large range of operating systems.) This has become more urgent as tests of the GPFS file system as an image sharing mechanism require support for a platform other than Fedora 14. Moreover, Fedora 14 no longer receives security updates, so system

administrators are, quite reasonably, hesitant to install that system. During this quarter, the necessary jobs and resources were put in place to provide packages for all of the StratusLab services on Fedora 14, OpenSuSE 12.1, and CentOS 6.2. This will allow tests with GPFS to proceed and will give system administrators the ability to use operating systems with long-term support. The project will move to CentOS 6.2 as the core platform for the next release and will also upgrade to use Fedora 16.

3.1.3 Task 2.2: Intensive Evaluation of StratusLab Products

StratusLab operations in a Bioinformatics Lab CNRS IBCP has continued its evaluation of the StratusLab distribution in a bioinformatics context. The PAT (Port Address Translation) implementation has been integrated with the command line interface and the standard web monitor and is now part of the official release sources of the project. The CNRS IBCP infrastructure has been advertised to some colleagues from the bioinformatics community. Of course, the first ones have been our partner from Institut Pasteur (Paris) with whom we are collaborating for the TOSCANI use case. Two accounts have been created. And they are using our cloud for the evaluation of the ARIA appliance. We have also opened our cloud to colleagues from within IBCP working in the structural biology field. They have requirements of accessing from time to time some computing resources in order to run tools for protein structure modeling. One of these tools has highlighted the need to run software under a license agreement, in that case an academic license. With these first accounts we are beginning our objectives to open our cloud to the bioinformatics community. This action will bring feedback to the project from real usage by non-computing-specialist scientists. But in our mind, this will be better achieved with a dedicated and simple web interface to the VM lifecycle integrating bioinformatics assistants (see related paragraph).

Grid'5000 Deployment Being able to test the StratusLab distribution at scale is critical for its adoption. Resources within the project have limited the level of scalability testing that can be done. As a possible solution, a StratusLab deployment inside of the Grid'5000 infrastructure has been investigated. This infrastructure, dedicated to computer science research, allows dynamic deployment of complete systems to the scale of 1500 of machines. An initial deployment was done by hand and validated, requiring only minimal changes in the standard StratusLab configuration. This was then automated, allowing easy redeployments of the system through Hudson. This system will be used as the basis for planned scalability tests in the final quarter.

Use Case Tests Comprehensive testing of the StratusLab features was scattered through a number of different test jobs in Hudson. This made it difficult to determine what features were tested and what was passing/failing. Moreover, many of the tests mixed the installation with client tests making it difficult to ascertain what a "pure" client without elevated access priviledges would experience. A set of distinct use cases covering the expected functionality has been extracted. These are

kept in a distinct code repository and run through a job that runs on a machine outside of the test cloud infrastructure. This permits better tracking of the StratusLab distribution's functionality.

3.1.4 Issues and Corrective Actions

Signficant progress was made on getting specific applications running on StratusLab cloud infrastructures; however, these successes have not been disseminated widely through a press release or public presentation. In Q8, this progress will be presented at the EGI Community Forum, but a formal press release and wider publicity needs to be planned. This will be a priority for Q8.

As seen in MS4 there are a good number of sites that have deployed StratusLab both inside and outside of the project. However, there is a need to push more sites to deploy StratusLab and to provide production cloud services to their users. The work by SixSq for commercial development will help, but must be complemented by specific actions within the academic community.

3.2 WP3: Dissemination

Work Package 3 coordinates the project's activities in dissemination, collaboration, exploitation and sustainability. Its objectives are to disseminate results of the project to resource providers, end-users, and the general public; identify project contributions to standards bodies and standardization efforts; and coordinate interactions with related projects, developing Memoranda of Understanding between projects where appropriate.

3.2.1 Summary

Dissemination efforts have continued this quarter with activities centered around the release of versions 1.2 and 1.3. The release announcements were widely circulated via the project mailing lists and were picked up by some technical press outlets. A number of talks and presentations featuring StratusLab were also given by project partners.

The website continues to be the main avenue for dissemination to the users, and the documentation section has been extensively updated during this quarter. While site visits are down again in Q6, this is to be expected due to the holiday period falling within Q7.

Significant progress has been made with collaborating projects, in particular with EGI who have deployed their own instance of the StratusLab Marketplace as part of their Federated Clouds Task Force activities, and new EGI-endorsed UMD virtual appliances available in the StratusLab Marketplace. Collaborations with other projects have also continued.

Exploitation and Sustainability efforts have increased, with participation in the EGI Sustainability Workshop in January, and progress on a number of commercial exploitation initiatives.

3.2.2 Task 3.1: Dissemination

Release Dissemination There were two incremental releases of the StratusLab distribution during this quarter, version 1.2 and 1.3. A release dissemination plan was produced for each of these releases. Notification of the release was disseminated via the website and the project's announcement email list. The release announcements for both releases were also sent to various media outlets and were carried by International Science Grid This Week (iSGTW) and HPC In the Cloud.

Media & Publications StratusLab participated in the Cloudscape IV event in February. News of the project's participation was posted on the website and was also disseminated to press contacts. The story was picked up by HPC In the Cloud.

Partner project the Initiative for Globus in Europe (IGE) ran a news item about the availability of Globus Appliances from the StratusLab Marketplace.

The release announcements and other media stories, events and general project news were also disseminated via Twitter, with several items being retweeted by related projects.

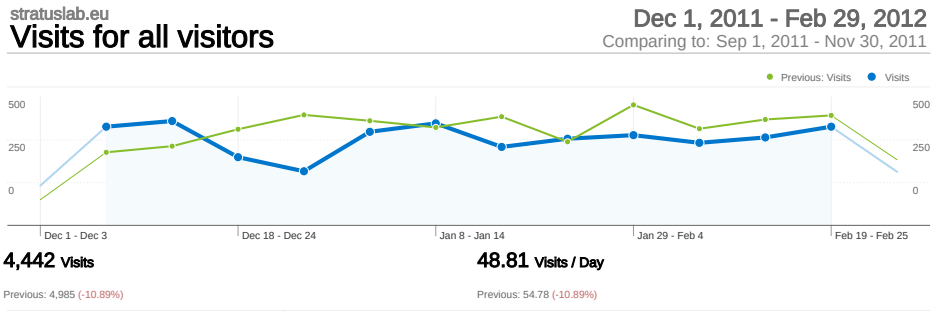

Figure 3.2: *Visits for Q7.*

Website, Mailing Lists and Fora Figure 3.2 shows the number of visits to the website. The number for Q7 (4,442 visits) is down slightly again from Q6 (4,985 visits).

The StratusLab Twitter feed now has 101 followers.

The announcements mailing list, which allows interested members of the public to keep up-to-date with news from the StratusLab project such as new releases, has 72 members.

The user announcements mailing list for users of the StratusLab reference infrastructure has 46 members, all of whom are users of the reference infrastructure.

The public user forum has 13 members, although there have been no posts on the forum during this period.

EGI Community Forum 2012 The project has booked an exhibition booth at the EGI Community Forum 2012, to be held in Munich, Germany 26-30 March 2012. The booth will be the centre of dissemination activities with a number of posters highlighting successes of the project. Promotional items such as t-shirts, brochures, pens and post-it notes will also be distributed. Video demos will be shown on a large screen at the booth, and a live demo will be available.

Project members will be involved in several of the conference sessions on Cloud computing. A StratusLab training workshop is also scheduled.

Talks A brief list of talks featuring StratusLab given by project members is given in Table 3.1.

3.2.3 Task 3.2: Collaboration with Standards Bodies and Related Projects

VENUS-C The project has continued its collaboration with VENUS-C during this quarter. Feedback was received from Venus-C users about requirements and user experience of the IaaS approach and this feedback was used to direct further development. In addition, the two projects have continued to share experiences relating to installation, performance and use of OpenNebula.

Some investigation has been done on the storage solution provided by VENUS-C and its relevance to StratusLab. Although this could probably be used to wrap

Table 3.1: *Talks*

Title / Event	Date
"Experience from the provision of IaaS cloud services using the StratusLab distribution" at eScience 2011, Stockholm, Sweden	2011-12-07
"What benefits of cloud for Bioinformatics - a user perspective" at eScience 2011, Stockholm, Sweden	2011-12-07
"What can grid and cloud computing do for you?" Seminar in School of Computer Science and Statistics, TCD, Dublin, Ireland	2012-01-27
"Grid and Cloud e-Infrastructures" at meeting with Intel Academic Research Programme representatives, Dublin, Ireland	2012-01-30
"SlipStream - Automated provisioning and continuous deployment in the cloud" at CERN, Geneva, Switzerland	2012-02-01
"OpenNebula Project" featuring StratusLab at FOSDEM, Brussels, Belgium	2012-02-04
"Open-source and Standards - Unleashing the Potential for Innovation of Cloud Computing" at CloudScape IV, Brussels, Belgium	2012-02-23
"StratusLab project status" presentation at round-table session at CloudScape IV, Brussels, Belgium	2012-02-23

the StratusLab storage services, the utlity of this is questionable especially in view of the need to improve the performance of the storage system in general.

EGI Collaboration with EGI continues in line with the MoU established between the two projects.

EGI have deployed their own instance of the StratusLab Marketplace, and evaluation of its use is ongoing. The StratusLab project undertook some development of the Marketplace in order to allow the EGI instance[2] to use its own branding, and this has been completed.

Partner TCD is participating in the EGI Federated Clouds Task Force as a resource provider with their cloud site running StratusLab software. TCD are also working with the Task Force to enable Nagios monitoring of TCD's StratusLab installation through OCCI v1.1. Minor modifications of the Nagios probe to make use of the StratusLab Marketplace have been needed.

StratusLab is planning to have a strong presence at the EGI Community Forum which will take place in March 2012. The project will take part in the Cloud and Virtualisation track at the conference and has submitted a training session and several presentations for the event.

Project partners participated in the EGI Sustainability Workshop which was held in Amsterdam, Netherlands, from 24-26 January 2012.

[2] Available at http://marketplace.egi.eu/.

StratusLab has also provided input to EGI on joint dissemination activities.

IGE Work has continued with the Initiative for Globus in Europe (IGE) project, with IGE contributing a number of fully tested and endorsed Globus Toolkit 5.0.4 appliances to the StratusLab Marketplace.

SIENA StratusLab had a strong presence at the SIENA CloudScape IV event. Project members gave two talks and participated in a number of round-table discussion sessions.

The project also provided input this quarter to the SIENA Roadmap on Distributed Computing Infrastructure for e-Science and Beyond in Europe.

ERINA+ An MoU was signed between StratusLab and ERINA+ in December 2011. As part of this MoU, StratusLab has committed to participating in the ERINA+ socio-economic impact assessment. ERINA+ has provided the project with access to their online self-assessment tool and StratusLab is currently in the process of collecting the necessary information to complete the self-assessment. This will allow StratusLab to measure the impact of the project's work in a methodologically sound way in collaboration with experts.

Commercial collaborations Partner SixSq has been involved in a collaboration with IBM and Swiss IT solutions provider Darest to develop a commercial turnkey Infrastructure as a Service (IaaS) cloud solution. The product, which will be distributed by Darest and which runs on IBM hardware, uses StratusLab 1.4 as its core, and is aimed at SMEs looking for a private cloud solution. The product will be announced in March.

3.2.4 Task 3.3: Development of Exploitation and Sustainability Plan

StratusLab participated in the EGI Sustainability Workshop which was held in Amsterdam, Netherlands, from 24-26 January 2012. The sustainability path developed during the meeting in Dublin from 30 November to 2 December provided input to this meeting.

Partner Telefónica I+D is developing the commercial use case in StratusLab to show the advantages of the StratusLab multi-tier service management, KPIs driven scalability and load balancing for an e-business application. By using this proof of concept, TID can show the advantages of using these advanced features that StratusLab provides.

As part of the StratusLab commercial exploitation strategy, partner SixSq has been involved in a collaboration with IBM and Swiss IT solutions provider Darest to develop a commercial turnkey Infrastructure as a Service (IaaS) cloud solution. The StratusLab distribution is one of the key components of this solution.

The product, distributed by Darest and running on IBM hardware, will be announced in March and is specifically tailored for at SMEs looking for a private cloud solution. This brings the StratusLab software to a new market beyond the mainly research-oriented current user base.

An instance of the StratusLab Marketplace will also be made available as a managed service for users of the new product.

3.2.5 Issues and Corrective Actions

The overall activity both within WP3, and in terms of visits to the websites, mentions in press, etc. is down this quarter when compared to the previous quarter. This is largely due, however, to the holiday period falling in this quarter and the trend can be expected to reverse in Q8.

A more concerted effort will be made in Q8 to drive traffic to the website through dissemination, news stories in the press, etc. In particular, the user forums received very little traffic and no posts during this quarter and these must be more widely disseminated in order to transform them into a useful community resource.

3.3 WP4: Software Integration and Distribution

This activity integrates and supports the StratusLab open-source cloud distribution. It integrates components required for grid site virtualization and dynamic scaling to remote clouds, addressing the specific requirements of the grid resource providers, and for the deployment of science clouds, addressing infrastructure cloud-like access demands from user communities, including industrial users.

WP4 works in close collaboration with WP5 for production deployment, WP6 for new service and component integration and all other work packages.

WP4 is also responsible to the execution of the project agile process, which includes active participation from all work packages.

3.3.1 Summary

During Q7, the integration and testing effort continued on the Persistent Disk Service, part of the StratusLab storage strategy. This service is sensitive to the local infrastructure network, which required further work in providing reliable behavior and acceptable performance. OpenNebula 3.0 was integrated in StratusLab, and integration of OpenNebula 3.2 started.

The improvements on the build and test infrastructure performed during QR6 payed-off, with significantly reduced certification time.

Three sprints were completed during this period, resulting in the release of StratusLab v1.2 and v1.3.

3.3.2 Task 4.1: Definition of Reference Architecture

The architecture of StratusLab is stable and no significant modifications were required.

3.3.3 Task 4.2: Integration of Open-source Distribution

Multiple OS Support Work to support more operating systems, in addition to the standard Fedora 14, has taken place. The target operating systems include OpenSuSE 12, Fedora 16 and CentOS 6.2. This ongoing work is significant, with important upgrades to our build and test system.

Integration of Monitoring API in StratusLab Monitoring component software was integrated in StratusLab in previous quarters. In this quarter, an implementation for a monitoring API (TCloud) has been introduced. This implementation is part of the tcloud-server-rpm.

Updated Claudia Client Tools The StratusLab client tools for Claudia have been modified to provide configuration for the monitoring systems.

Claudia Integration with OpenNebula 3.0 OpenNebula 3.0 was integrated in the StratusLab distribution, requiring several components to be adapted, including the clients and Claudia.

Port Address Translation PAT developed by IBCP was integrated; this allows virtual machines with local IP addresses to communicate with the outside world, permitting a site with a limited number of public IP addresses to still offer a viable public cloud service.

Persistent Disk Service The image creation was upgraded to take advantage of the pdisk service. The StratusLab TM uses store/snapshot capabilities of pdisk for caching of new images, fast start up of VMs, and easier, asynchronous new image creation and registration. This integration also includes a much more thorough policy check based on metadata in the Marketplace.

Several performance and stability issues were detected during this quarter. They now seem to have being addressed but this will have to be confirmed during the last quarter of the project.

3.3.4 Task 4.3: Contextualization of Grid Services

With the more stringent policy enforcement feature integrated in StratusLab v1.3, several Marketplace entries had to be corrected and improved in order to be compliant.

WP3 provided support to WP5 in the creation of automatic deployment of Grid Sites using SlipStream on the StratusLab reference infrastructure at GRNET. This work, although not completed yet, will provide the EGI/EMI community with a simpler way of deploying and configuring a Grid Site on the StratusLab cloud.

3.3.5 Task 4.4: Technical Support

Technical support continued to be provided, coordinated with WP5, on a wide range of issues, including the Persistent Disk Service, contextualisation and command-line tools.

3.3.6 Issues and Corrective Actions

The behavior of pdisk continued to cause problems to the StratusLab reference architecture. Several improvements were investigated and implemented, which have significantly improved the situation. We believe that the latest solution provide reasonable performance and stability. This effort in testing and monitoring of the storage solution will continue during the last quarter of the project.

3.4 WP5: Infrastructure Operation

WP5 is responsible for the provision and operation of the project's computing infrastructure. It serves as a beta-tester of the software integrated by WP4 and WP6, deploying it in a production environment in order to verify its applicability for real life applications. In addition WP5 offers daily support to external users, either system administrators or those exploiting the project's public cloud services. Finally, WP5 is contributing with targeted development activities, related to the improvement of the cloud tools and services, as well as to showcase the capabilities of the StratusLab distribution to satisfy different use cases.

3.4.1 Summary

Evolution, testing and fine tuning have been the driving forces for WP5 during this quarter. The activity followed closely the evolutionary path of the StratusLab distribution by certifying and installing all the releases of the software during this quarter. During this process we intensified the effort required for testing the new releases. We also identified issues in the software and in the deployed hardware architecture, trying to alleviate them (at least those issues related to the physical infrastructure). For the hosted virtualized grid site, there is an ongoing effort to migrate to the new software supported by EGI (UMD-1) and to provide prepared appliances and support tools in order to ease the deployment and configuration of grid services. Improvements will be made in all in the coming months, correcting some unresolved issues pertaining the above tasks.

3.4.2 Task 5.1: Deployment and Operation of Virtualized Grid Sites

Reference Cloud Service The reference cloud service has followed the evolution of the StratusLab distribution, deploying first in a large scale environment the software delivered by WP4. In particular during Q7 the service has been upgraded to versions 1.2 and then 1.3 introducing in parallel the persistent disk service (pdisk) thus providing storage management and compute services. Also, the reference service was able to exploit the caching mechanisms (also a feature of pdisk) reducing significantly the amount of time it takes for a VM to instantiate improving the overall efficiency. Cached images take only a few seconds for the VM to reach the 'running" state, with the latency for access by the user dominated by the configuration of the image (number of default services and kernel modules).

The transition to pdisk was not seemless though. The introduction of the dependency on iSCSI had a significant impact on the way storage is managed from the service backend. To date, the service utilized a centralized storage server (EMC Celerra NS-480) in order to provide a shared repository between the cloud frontend and the hosting nodes for the VM images. This setup was changed when deploying StratusLab 1.2. In particular an NFS volume was allocated to the frontend in order to host a large LVM volume which was then assigned to the pdisk server in order

to offer disk devices over iSCSI for the VMs. This setup proved to be inefficient, introducing multiple levels of delay, degrading the overall performance of the service and in the end wiping out the expected benefits of pdisk. To work around this problem, we decided to drop the usage of NFS and alternatively use a relatively large local disk in order to configure the necessary volume groups for pdisk. This reconfiguration improved slightly the performance of the system with respect to storage I/O but still is underperforming comparing to the initial NFS-based setup. For example the I/O throughput recorded in the local disk iSCSI setup was around 5 MB/s whereas the throughput with the NFS is around 10 MB/s. Though both numbers are sub-optimal the difference between the two approaches is clearly evident.

Finally another problem identified was the degraded performance of the existing appliances that have been prepared using the qcow2 format. Qcow2, as the name suggests, uses a copy-on-write (cow) approach, starting from a minimum disk space and expanding to a maximum predefined size at run time based on the image disk usage. Unfortunately this behavior introduces one more performance bottleneck, since the images are used to create an LVM volume that performs poorly during copy-on-write activities.

Hosted grid site The production Grid site running on the reference cloud service, is scheduled to follow the evolution of EGI software and migrate to UMD-1 software distribution. For this purpose a set of appliances has been prepared with the required software for the CE, SE, WN and APEL grid nodes. In parallel a deployment module has been developed using Slipstream in order to automate the grid site configuration procedure. By the end of Q7 the transition has not been completed yet due to the performance issues described above. The site is still running on a backup cloud service that does not use pdisk till the issues of the reference cloud have been solved. The instantiation of the site with the UMD-1 images are expected to complete at the beginning of Q8 during Sprint 23.

Persistent Disk Service Improvements to the persistent disk service were made to make it more stable and more modular. This modularity allows different storage options (posix file system, LVM) and sharing options (iSCSI, shared file system) be used by the service. A prototype module has been written that allows a NetApp storage system to be used behind the persistent disk service. This will be throughly tested and put into production in Q8.

TERENA Certificates Work was done to ensure that all of the StratusLab services have valid certificates issued through TERENA. These certificates are signed by a standard Certificate Authority and do not require special configuration of clients or cause warnings about untrusted certificates to appear in browsers.

3.4.3 Task 5.2: Testing of the StratusLab Toolkit

Automated Certification of Release Candidates The new certification process defined in Q6 were utilized during Q7 in order to test and verify versions 1.2 and 1.3 before releasing them and deploying them in the reference cloud infras-

tructure. As expected the automated process significantly accelerated the software validation procedure. Nevertheless it has not yet reached to a level of completeness that will enable the identification of all potential issues in a release. For example the process failed to identify the performance bottlenecks that appeared in the reference deployment. The problem arises from the fact the the certification is performed on a limited infrastructure (1 frontend - 1 hosting node) which is not sufficient to perform large scale tests (e.g. bulk instantiation of 30 or more VMs in order to stress the system) nor can offer all possible hardware configuration scenarios (e.g. using centralized storage server or a distributed file system). During Q8 we will try to improve this situation by providing additional resources (even temporarily during the certification) in order at least to be able to identify any potential scaling issues.

Support for virtio_net and virtio_blk Drivers This activity was initiated from a question from one of our users noticing that the maximum bandwidth that he could achieve on the reference cloud was 100Mb/s. Initial experiments revealed that actually the maximum bandwidth that can be achieved is around 150 Mb/s. Having in mind the the hosts are using a 1 Gb/s Ethernet interface this number is significantly lower than the maximum theoretical peak. Further investigation revealed that the cause of the problem is that the VM appliances are not configured with the virtio paravirtualized driver. Therefore, the network requests from the guests are emulated, which results in poor I/O performance. By activating virtio_net on both VM and hosting node the network bandwidth increases up to 680 Mbit/s.

Similarly, the access to the local disk passes through the native IDE driver and not through the paravirtualized virtio_blk one, although the performance difference in this case does not appear to be significant.

The above issues have been forwarded to WP4 as a requirement and is expected that the next versions of StratusLab will offer native support for virtio_net and virtio_blk.

3.4.4 Task 5.3: Virtual Appliances Creation and Maintenance

Marketplace During this quarter development of the Marketplace has continued. The focus has been on improving documentation and performance. Documentation on installation and usage has been added to the project wiki.

Support for PostgreSQL as a back-end for metadata storage has been added. This provides significantly better performance than MySQL.

A requirement of the Marketplace is that email verification should be performed when metadata is uploaded. To date this has not been enabled as it would prevent automated uploads. To allow for services to upload metadata the ability to whitelist certain endorsers has been added.

Automated Production of Base Images Work was done, in collaboration with WP2, to provide a consistent procedure for creating the StratusLab base images. Moreover, this production was automated to ensure that security patches can be released regularly. See the more detailed description in the WP2 section.

3.4.5 Issues and Corrective Actions

Performance of the reference cloud service As mentioned already with the introduction of pdisk we experienced a performance degradation for what concerns activities that depend on the disk I/O. From our tests it is evident that the way pdisk works (LVM with iSCSI) requires a different hardware setup than what we currently have in the reference infrastructure. In particular the service performs on par when a dedicated iSCSI server or storage appliance is used to serve volumes. This setup for example is currently employed successfully in LAL's installation. In order to alleviate this issue in GRNET (where the reference infrastructure is located) we are currently investigating alternative setups that will resolve the bottlenecks experienced in the current configuration. Meanwhile, since this situation has impacted our work on the production grid site we are thinking to move the latter from the GRNET cloud to LAL's cloud. This will also give as a good opportunity to test and demonstrate the federation capabilities of StratusLab and also the ability to move virtualized grid sites between different cloud services, even different countries, still maintaining the quality levels required by EGI.

3.5 WP6: Innovative Cloud-like Management of Grid Services and Resources

The Joint Research Activity (JRA), carried out in WP6, develops advanced technology and features for deployment on existing Cloud infrastructures through automatic deployment and dynamic provision of grid services as well as scalable cloud-like management of grid site resources. More specifically, the objectives to be accomplished can be expressed as: i) the extension of currently available open-source service-level frameworks which provide elasticity on top of cloud infrastructures, ii) the invention of new techniques for the efficient management of virtualized resources for grid services and iii) the inclusion of novel resource provisioning models based on cloud-like interfaces.

3.5.1 Summary

The work done in Q7 has involved the continuation with the integration of the monitoring systems in the StratusLab distribution. This has implied to integrate monitoring systems with Claudia to drive with scalability and to introduce monitoring probes by contextualization. Regarding network management, the network operations are now coupled with the VM lifecycle, its definition is flexible and it is possible to reserve network leases. Placement policy can be defined in VM-basis too. Finally, the work on interoperability and APIs have carried out the introduction of the TCloud API as monitoring API, new OCCI extensions to include VM types and network definition, the introduction of the Deltacloud API and more attributes for Security and User Management. This development work is being integrated into StratusLab distribution in WP4.

3.5.2 T6.1: Dynamic Provision of Grid Services

Automatic Insertion of Monitoring Probes Monitoring software probes are inserted by the Service Manager inside the virtual machine by using the StratusLab contextualization mechanisms. This means that when the VM has been deployed, automatically it contains a set of probes that provide specific metrics for the applications to the monitoring system.

Integration of monitoring systems and Claudia for scalability Claudia is the service manager for managing the service lifecycle and the monitoring systems collects monitoring information of those services. The Key Performance Indicators (KPIs) that drive scalability are provided by probes and stored in the monitoring systems. Thus, a new process is required to obtain KPIs from the monitoring systems and inject them into Claudia. This process has been included as part as the Claudia software.

3.5.3 T6.2: Scalable and Elastic Management of Grid Site Infrastructure

Improved Network Management The network operations are now coupled with the VM lifecycle. This simplifies the management of networking (no hooks are needed) and solves previous issues with VLANs when migrating and restoring VMs. The network drivers define three actions (pre-boot, post-boot and clean) that can be easily customized if needed. Previous hooks for 802.1Q VLAN tagging, Open vSwitch and ebtables-based VLANs have been transformed to drivers, as well as the one to set simple firewalling rules.

Flexible Network Definition Networks can be now defined with an arbitrary range including an starting and ending IP address, network and network mask, or CIDR notation. It is possible also to define a network and a starting IP address to lease addresses.

New Network Lease Operations Network leases can now be put on hold to reserve them. This comes in handy when there are some IPs within the VLAN already assigned (e.g. .1 to the gateway). When a lease is put on hold, OpenNebula will not use it for a VM, until it is released.

Placement Policy Definition The placement of the VMs can now be defined on a VM-basis (restricted to oneadmin) or globally for the datacenter. This allows admins to set a global optimization policy to meet specific goals. The scheduling includes four predefined policies: packing, striping, load-aware, and custom.

3.5.4 T6.3: Cloud-like Interfaces Specific for the Scientific Community

Introduction of TCloud API for monitoring API Monitoring information for both VMs and services can be accessible by a TCloud API. Thus, a TCloud driver for monitoring has been included as part of the implementation.

Security and User Management Some attributes (DISK/SOURCE, CONTEXT/FILES, NIC/MAC and NIC/VLAN_ID) have been restricted in VM Templates, because they can be easily used to gain oneadmin access or to comprise VMs of any user. There are new auth drivers for LDAP, with base and group filtering, and for Cloud API and OpenNebula front-end servers (server-based drivers). Also, as some of the drivers may take some time to authenticate a request (e.g. LDAP), session tokens can now be cached by OpenNebula. Finally, a new permission set has been included to manage access control to virtual resources. The new permissions overcomes the limitations of the previous PUBLIC attribute and allow users to share resources in multiple ways. Combined with the ACL system (also simplified to match the new permissions), this allows the implementation of multiple roles.

OCCI Extensions OpenNebula's OCCI API has been extended to include VM types, that can now be defined in the server configuration file and tagged with

arbitrary information, like size, QoS parameters or price. These VM types can be programatically queried through the API. Also, to support the new VLAN features in OpenNebula, the OCCI networks can now be defined through a template, as for Virtual Machines.

Deltacloud API Updates Apache Deltacloud is a REST-based API that abstracts the differences between clouds, enabling the management of resources in different IaaS clouds using a single API. A series of back-end drivers translate the request to each cloud provider's native API. Currently, all major cloud service providers are supported.The OpenNebula back-end driver in Deltacloud has been updated to interact with OpenNebula 3.x clouds. This way, StratusLab sites could be accessed through the Deltacloud API, server and tools.

3.5.5 Issues and Corrective Actions

None.

4 Project Management

4.1 Consortium

The project consortium consisting of six partners (CNRS, UCM, GRNET, SIXSQ, TID, and TCD) has not changed since the start of the project. There have been no changes in the legal status of those partners. The representatives for TCD and TID have changed because of retirements and internal reorganization of activities.

4.2 Management Tasks

Meetings Tables 4.1–4.7 contain a list of the meetings by quarter that have been planned to foster collaboration between the project participants. Not listed are the planning meetings for each development sprint and the daily standup meetings.

Metrics Table 4.8 contains the metrics for the project. The table groups related metrics together.

Deliverables and Milestones Tables 5.1, 5.2, and 5.3 list all of the documents. In addition, these are available from the project website.

Deliverable D4.4, originally foreseen for PM15, was delivered in PM16. Deliverable D6.4 that updates the design for advanced services in WP6 has been produced in this quarter, although also with a slight delay. Deliverable D5.4 has been moved from PM18 to PM21 to allow more time to collect data from the cloud infrastructure.

Memoranda of Understanding The project has signed MoUs with VENUS-C, EMI, IGE, and ERINA+ in this quarter.

4.3 Issues

Year 2 Budget It has proven extremely difficult to collect interim budget information from the partner's administrations. It is now clear that several partners will significantly underspend on their budget. An analysis is being carried out to guage what the financial situation will be at the end of the project, although it is likely that little can be adjusted at this point.

Review Recommendations The responses to the review recommendations (included in this report) have been updated to reflect the situation as of the end of Q6. They will continue to be updated over remainder of the project.

4.4 Planning

4.4.1 Objectives for Next Quarter

- Release of v2.0 of the StratusLab distribution
- Detailed definition and initial implementation of sustainability plan
- Demonstrate of hybrid cloud computing
- Measurement of the behavior of the system in terms of scalability and IO performance, including with alternate configurations (e.g. GPFS)
- Demonstrate a pre-production elastic Grid site
- Complete migration of gLite production Grid site to UMD-1

4.4.2 Roadmap

The roadmap remains essentially the same as decided in the Lyon Face-to-Face meeting. The PMB in Q3 gave its formal approval of the following changes to the overall work program:

1. The tasks regarding having a public (user-visible) cloud and an associated cloud API have been moved from Y2 to Y1, largely because of interest from scientific communities and resource centers wanting to provide public clouds.
2. The tasks about hybrid clouds will be expanded to include also cloud federation models. This will be moved to Y2 to balance the change above. Also having a solid release will make these investigations easier.
3. As foreseen in the TA, the appliance repository consists of a single service that contains appliance metadata, appliance storage, and services for changing appliance formats. This has been split into different services. The Marketplace will handle appliance metadata. Storage will take place with normal cloud storage or outside of the cloud. Instead of providing a service for appliance format changes, client tools will be provided instead.

These changes have been made and followed at the technical level for sometime; they are now also agreed at the management level.

The architecture and roadmap have been re-evaluated in D4.4 along with D6.4. The project will continue to make incremental changes to the existing distribution. The priority for Year 2 is the demonstration of hybrid cloud functionality.

Table 4.1: *Meetings (Q1)*

Title	Date	Venue	Comments
StratusLab Kick-Off Meeting	14-15/06/2010	Orsay, FR	Kick-off of project. Detailed planning for accomplishing objectives. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1129
Technical Meeting	22/07/2010	Madrid, ES	Detailed technical discussions for StratusLab development. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1189
Sprint 1 Demo	30/07/2010	Phone/EVO	Sprint 1 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1191
Sprint 2 Demo	20/08/2010	Phone/EVO	Sprint 2 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1192

Table 4.2: *Meetings (Q2)*

Title	Date	Venue	Comments
Project Management Board	03/09/2010	Phone	PMB meeting to decide IPR policies. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1203
Sprint 3 Demo	10/09/2010	Phone/EVO	Sprint 3 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1203
Technical Meeting (TSCG)	21/09/2010	Phone/EVO	Shaping StratusLab distribution. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1213
WP6 research lines meeting	27/09/2010	Madrid, ES	Discussion about the main gaps identified in WP4 and some technologies to solve them. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1318
WP6 kickoff meeting	07/10/2010	Phone	Presentation of the lines to work on WP6 and distribution of work. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1320
Sprint 4 Demo	08/10/2010	Phone/EVO	Sprint 4 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1232
WP6 monitoring and accounting	26/10/2010	Phone	Audioconference about monitoring and accounting in StratusLab. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1321
Sprint 5 Demo	08/11/2010	Phone/EVO	Sprint 5 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1255
Face-to-Face Technical Meeting	15-16/11/2010	IBCP, Lyon, France	Discussion of StratusLab roadmap. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1243
Project Management Board	22/11/2010	Phone	Project overview; LoS policy. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1263

Table 4.3: *Meetings (Q3)*

Title	Date	Venue	Comments
Sprint 6 Demo	09/12/2010	Phone/EVO	Sprint 6 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1310
Sprint 7 Demo	17/12/2010	Phone/EVO	Sprint 7 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1323
Technical Meeting (TSCG)	27/01/2011	Phone/EVO	Feedback from EGI; priorities for distribution. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1213
Sprint 8 Demo	31/01/2011	Phone/EVO	Sprint 8 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1423
Technical Meeting (TSCG)	17/02/2011	Phone/EVO	Error reporting; priorities for next sprint. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1213
Sprint 9 Demo	18/02/2011	Phone/EVO	Sprint 9 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1442
Project Management Board	24/02/2011	Phone	Project status; MoUs; effort utilization; review planning. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1440

Table 4.4: *Meetings (Q4)*

Title	Date	Venue	Comments
Sprint 10 Demo	03/03/2011	Phone/EVO	Sprint 10 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1448
Technical Meeting (TSCG)	03/03/2011	Phone/EVO	Review of developments and priorities. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1460
Sprint 11 Demo	31/03/2011	Phone/EVO	Sprint 11 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1470
Metadata & Marketplace Demo	08/04/2011	EVO	Demo for HEPiX Virtualization Working Group. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1477
Sprint 12 Demo	29/04/2011	Phone/EVO	Sprint 12 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1492
Grid site deployment with Claudia (TID, GRNET)	09/05/2011	Phone	Discussion about how to use Claudia for the deployment of a grid site. http://indico2.lal.in2p3.fr/indico/conferenceTimeTable.py?confId=1530#20110509
Technical Meeting (TSCG)	10/05/2011	Phone	Persistent storage and cloud interfaces. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1526
Interproject Collaboration	11/05/2011	Amsterdam	StratusLab, HPC Cloud, and Mantychore discussions. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1510
Sprint 13 Demo	16/05/2011	Phone/EVO	Sprint 13 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1513
Integration Meeting	23-24/05/2011	Geneva	F2F meeting for 1.0 release. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1503
Interproject Collaboration	27/05/2011	Phone	Discussion with Contrail project. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1527
Grid site deployment and scalability (TID, GRNET)	27/05/2011	Phone	Discussion to align the work. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1529

Table 4.5: *Meetings (Q5)*

Title	Date	Venue	Comments
Sprint 14 Demo	10/06/2011	Phone/EVO	Sprint 14 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1533
StratusLab Face-to-Face Meeting	21-23/06/2011	Geneva, CH	Integration of software. Update of roadmap. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1502
Sprint 15 Demo	23/06/2011	Phone/EVO	Sprint 15 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1541
StratusLab First Periodic Review	04/07/2011	Brussels, BE	External review of project's progress. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1501
Sprint 16 Demo	29/07/2011	Phone/EVO	Sprint 16 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1557
Technical Meeting (TSCG)	25/08/2011	Phone/EVO	Priorities for upcoming sprints. Architecture for StratusLab 2.0. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1560
Sprint 2 Demo	20/08/2010	Phone/EVO	Sprint 2 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1192

Table 4.6: *Meetings (Q6)*

Title	Date	Venue	Comments
Project Management Board	12/09/2011	Phone	Project status; Reviewer Feedback; Y2 Effort and Budgets. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1561
Sprint 17 Demo	16/09/2011	Phone/EVO	Sprint 17 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1574
Technical Meeting (TSCG)	26/09/2011	Phone/EVO	Architecture review; priorities for upcoming sprints. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1575
Technical Meeting (TSCG)	17/10/2011	Phone/EVO	Use Cases for Y2 Review; priorities for upcoming sprints. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1613
Sprint 18 Demo	19/10/2011	Phone/EVO	Sprint 18 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1617
Sprint 19 Demo	16/11/2011	Phone/EVO	Sprint 19 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1638
StratusLab Cloud Training	17–18/11/2011	Orsay, France	General training for users and administrators. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1565
Technical Meeting (TSCG)	21/11/2011	Phone/EVO	F2F agenda; priorities for upcoming sprints. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1642
Project Management Board	30/11/2011	Phone	Project status; ERINA+ MoU; Reviewer Recommendations; Sustainability. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1650
StratusLab Face-to-Face Meeting	30/11/2011–02/12/2011	Dublin, Ireland	Integration of software; Update of roadmap. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1621

Table 4.7: *Meetings (Q7)*

Title	Date	Venue	Comments
Sprint 20 Demo	19/12/2011	Phone/EVO	Sprint 20 demonstration meeting. http://indico2.lal.in2p3.fr/indico/categoryDisplay.py?categId=131
Technical Meeting (TSCG)	31/01/2012	Phone/EVO	F2F agenda; priorities for upcoming sprints. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1709
Sprint 21 Demo	30/01/2012	Phone/EVO	Sprint 21 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1705
Sprint 22 Demo	28/02/2012	Phone/EVO	Sprint 22 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1749
Technical Meeting (TSCG)	01/03/2012	Phone/EVO	Priorities for upcoming sprints. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1709

Table 4.8: *Metrics*

Metric	Q2	Q3	Q4	Y1 Target	Q5	Q6	Q7	Q8	Y2 Target
No. of people trained on StratusLab software	N/A	N/A	~ 25	–	~36	~89	0		–
No. of people on StratusLab announcement list	67	67	67	25	70	72	72		75
Registered users on StratusLab discussion site	N/A	N/A	N/A	50	N/A	12	13		100
No. of views of website	2922	4623	4579	–	5472	4985	4442		–
No. of completed sprints	5	5	4	–	3	3	3		–
No. of releases	1	1	1	–	2	1	2		–
No. of open user stories	38	72	101	–	118	107	121		–
No. of implemented user stories	69	40	67	–	50	48	57		–
No. of open bugs	6	15	22	–	28	51	51		–
No. of fixed bugs	7	11	27	–	14	20	17		–
No. of prod. sites running StratusLab dist.	1	1	1	5	1	3	5		10
Availability of hosted grid sites	N/A	N/A	100%	80%	91%	74%	93%		95%
Reliability of hosted grid sites	N/A	N/A	100%	80%	92%	78%	93%		95%
No. of VOs served via StratusLab hosted grid sites	0	1	1	10	21	18	18		30
No. of sci. disciplines served via StratusLab hosted grid sites	0	0	0	3	11	9	9		15
Delivered computing resources through hosted grid services	N/A	16 cores	16 cores	–	32 cores	32 cores	32 cores		–
Delivered computing resources through hosted cloud services	N/A	256 cores	256 cores	–	256 cores	288 cores	256 cores		–
Storage provided through cloud service	N/A	N/A	N/A	–	0	0	0		3 TB
No. of jobs run in hosted grid site	N/A	N/A	N/A	–	13,960	16,916	28,345		–
Norm. CPU time consumed in the hosted grid site (hrs)	N/A	N/A	N/A	–	26,202	14,231	87,671		–
No. base machine images	5	7	8	5	8	13	–		10
No. of base machine image downloads	783	2628	7072	–	7225	6657	–		–
No. appliances	0	6	7	5	7	7	–		15
No. of appliance downloads	0	252	687	–	1010	426	–		–
No. of Marketplace metadata entries	–	–	–	–	–	–	111		–
No. of Marketplace endorsers	–	–	–	–	–	–	24		–
No. of Marketplace base images	–	–	–	–	–	–	86		–
No. of Marketplace appliances	–	–	–	–	–	–	25		–

5 Deliverables and Milestones

Tables 5.1 and 5.2 show the deliverables for the first and second years of the project. Table 5.3 lists all of the milestones. All of the deliverables and milestones for the first year of the project have been produced and submitted as foreseen in the project's roadmap. All of these are available from the project's website[1].

Two technical notes have also been produced during the first year: "StratusLab Marketplace" describing the technical specification of the Marketplace and "Installing and operating a production grid site in the StratusLab cloud: Experience and issues" providing feedback to developers and advice to administrators running grid services within a cloud. These notes are also available from the project website.

[1] http://stratuslab.eu/doku.php/deliverables

Table 5.1: *Deliverables (Year 1)*

No.	Title	Version	WP No.	Lead Beneficiary	Nature	Diss. Level	Due Date	Actual Date	Status	Contractual	Comments
D2.1	Review of the Use of Cloud and Virtualization Technologies in Grid Infrastructures	1.2	WP2	CNRS	R	PU	PM2	11/08/2010	Done	Yes	
D4.1	Reference Architecture for StratusLab Toolkit 1.0	1.0	WP4	SIXSQ	R	PU	PM3	14/09/2010	Done	Yes	
D5.1	Infrastructure Specification	1.0	WP5	GRNET	R	PU	PM3	14/09/2010	Done	Yes	
D3.1	Initial Plan for Dissemination, Collaboration and Standardization Activities	1.0	WP3	TCD	R	PU	PM4	18/10/2010	Done	Yes	
D6.1	Cloud-like Management of Grid Sites 1.0 Design Report	1.0	WP6	TID	R	PU	PM5	16/11/2010	Done	Yes	
D5.2	Infrastructure Tool and Policy Specification	1.0	WP5	GRNET	R	PU	PM6	15/12/2010	Done	Yes	
D6.2	Cloud-like Management of Grid Sites 1.0 Software	1.1	WP6	TID	P	PU	PM11	13/05/2011	Done	Yes	
D2.2	Report on Evaluation of StratusLab Products	1.0	WP2	CNRS	R	PU	PM12	15/06/2011	Done	Yes	
D3.2	Report on Dissemination, Collaboration and Standardization Activities	1.1	WP3	TCD	R	PU	PM12	16/06/2011	Done	Yes	
D3.3	Exploitation and Sustainability First Plan	1.1	WP3	TCD	R	PU	PM12	16/06/2011	Done	Yes	
D4.2	StratusLab Toolkit 1.0	1.0	WP4	SIXSQ	P	PU	PM12	15/06/2011	Done	Yes	
D4.3	First Year Software Integration Report	1.0	WP4	SIXSQ	R	PU	PM12	15/06/2011	Done	Yes	
D5.3	First Year Infrastructure Operations Report	1.1	WP5	GRNET	R	PU	PM12	16/06/2011	Done	Yes	
D6.3	First Year Cloud-like Management of Grid Sites Research Report	1.0	WP6	TID	R	PU	PM12	15/06/2011	Done	Yes	

Table 5.2: *Deliverables (Year 2)*

No.	Title	Version	WP No.	Lead Beneficiary	Nature	Diss. Level	Due Date	Actual Date	Status	Contractual	Comments
D2.3	Survey of Targeted Communities Concerning StratusLab		WP2	CNRS	R	PU	PM14	12/08/2011	Done	Yes	
D4.4	Reference Architecture for StratusLab Toolkit 2.0		WP4	SIXSQ	R	PU	PM15	03/10/2011	Done	Yes	Delayed PM15-16
D6.4	Cloud-like Management of Grid Sites 2.0 Design Report		WP6	TID	R	PU	PM17	15/12/2011	Done	Yes	Delayed PM17-18
D5.4	Economic Analysis of Infrastructure Operations		WP5	GRNET	R	PU	PM18	19/03/2012	Delayed	Yes	Delayed to PM21
D6.5	Cloud-like Management of Grid Sites 2.0 Software		WP6	TID	P	PU	PM23			Yes	
D2.4	Final Report on StratusLab Adoption		WP2	CNRS	R	PU	PM24			Yes	
D2.5	Report on Evaluation of StratusLab Products		WP2	CNRS	R	PU	PM24			Yes	
D3.4	Final Review of Dissemination, Collaboration and Standardization Activities		WP3	TCD	R	PU	PM24			Yes	
D3.5	Exploitation and Sustainability Final Plan		WP3	TCD	R	PU	PM24			Yes	
D4.5	StratusLab Toolkit 2.0		WP4	SIXSQ	P	PU	PM24			Yes	
D4.6	Software Integration Final Report		WP4	SIXSQ	R	PU	PM24			Yes	
D5.5	Infrastructure Operations Final Report		WP5	GRNET	R	PU	PM24			Yes	
D6.6	Cloud-like Management of Grid Sites Research Final Report		WP6	TID	R	PU	PM24			Yes	

Table 5.3: *Milestones*

No.	Title	WP No.	Lead Beneficiary	Due Date	Achieved	Actual Date	Comments
MS1	Establishment of Management Infrastructure and Metrics Definition	WP1	CNRS	PM3	Yes	1/09/2010	
MS6	Website Operational	WP3	TCD	PM3	Yes	6/09/2010	
MS2	Contact Procedures and Supporting Tools for Targeted Communities	WP2	CNRS	PM4	Yes	10/12/2010	
MS7	StratusLab Development, Certification and Release Procedures in Place	WP4	SIXSQ	PM6	Yes	10/12/2010	
MS3	Creation of Virtual Appliances for Bioinformatics Community	WP2	CNRS	PM9	Yes	14/03/2011	
MS10	Initial virtual appliance repository	WP5	GRNET	PM9	Yes	4/03/2011	
MS14	Release of Cloud-like Management of Grid Services and Resources 1.0 Beta	WP6	TID	PM9	Yes	14/03/2011	
MS8	Release of StratusLab 1.0 Beta	WP4	SIXSQ	PM10	Yes	05/04/2011	
MS11	Operation of Site Running StratusLab toolkit v1.0	WP5	GRNET	PM10	Yes	04/04/2011	
MS4	Adoption of StratusLab Software by External Grid Sites	WP2	CNRS	PM14	Yes	19/03/2012	
MS12	Delivery of Virtual Appliance Repository	WP5	GRNET	PM18	Yes	15/12/2011	
MS5	Opening of Virtual Appliances Repository to External Application Communities	WP2	CNRS	PM20	Yes	19/03/2012	
MS15	Release of Cloud-like Management of Grid Services and Resources 2.0 Beta	WP6	TID	PM21	Yes	19/03/2012	
MS9	Release of StratusLab 2.0 Beta	WP4	SIXSQ	PM22			
MS13	Operation of Site Running StratusLab Toolkit v2.0	WP5	GRNET	PM22			

6 Use of Resources

The effort and spending plan for year 2 is being revised. Collecting this information from the partners has been much more difficult than foreseen. Complete information is expected in January 2011. However, it is clear that the underspending continues and the project will have to develop strategy for dealing with this.

Glossary

APEL	Accounting Processor for Event Logs (EGI accounting tool)
Appliance	Virtual machine containing preconfigured software or services
CDMI	Cloud Data Management Interface (from SNIA)
CE	Computing Element in EGI
DCI	Distributed Computing Infrastructure
DMTF	Distributed Management Task Force
EGEE	Enabling Grids for E-sciencE
EGI	European Grid Infrastructure
EGI-TF	EGI Technical Forum
GPFS	General Parallel File System by IBM
Hybrid Cloud	Cloud infrastructure that federates resources between organizations
IaaS	Infrastructure as a Service
iSGTW	International Science Grid This Week
KPI	Key Performance Indicator
LB	Load Balancer
LRMS	Local Resource Management System
MoU	Memorandum of Understanding
NFS	Network File System
NGI	National Grid Initiative
OCCI	Open Cloud Computing Interface
OVF	Open Virtualization Format
Public Cloud	Cloud infrastructure accessible to people outside of the provider's organization
Private Cloud	Cloud infrastructure accessible only to the provider's users
SE	Storage Element in EGI
SGE	Sun Grid Engine
SNIA	Storage Networking Industry Association
TCloud	Cloud API based on vCloud API from VMware
VM	Virtual Machine
VO	Virtual Organization
VOBOX	Grid element that permits VO-specific service to run at a resource center
Worker Node	Grid node on which jobs are executed

XMLRPC	XML-based Remote Procedure Call
YAIM	YAIM Ain't an Installation Manager (configuration utility for EGI)