StratusLab

Enhancing Grid Infrastructures with Virtualization and Cloud Technologies

StratusLab is developing and deploying cloud technologies with the aim of simplifying and optimizing the use and operation of distributed computing infrastructures such as the European Grid Infrastructure (EGI).

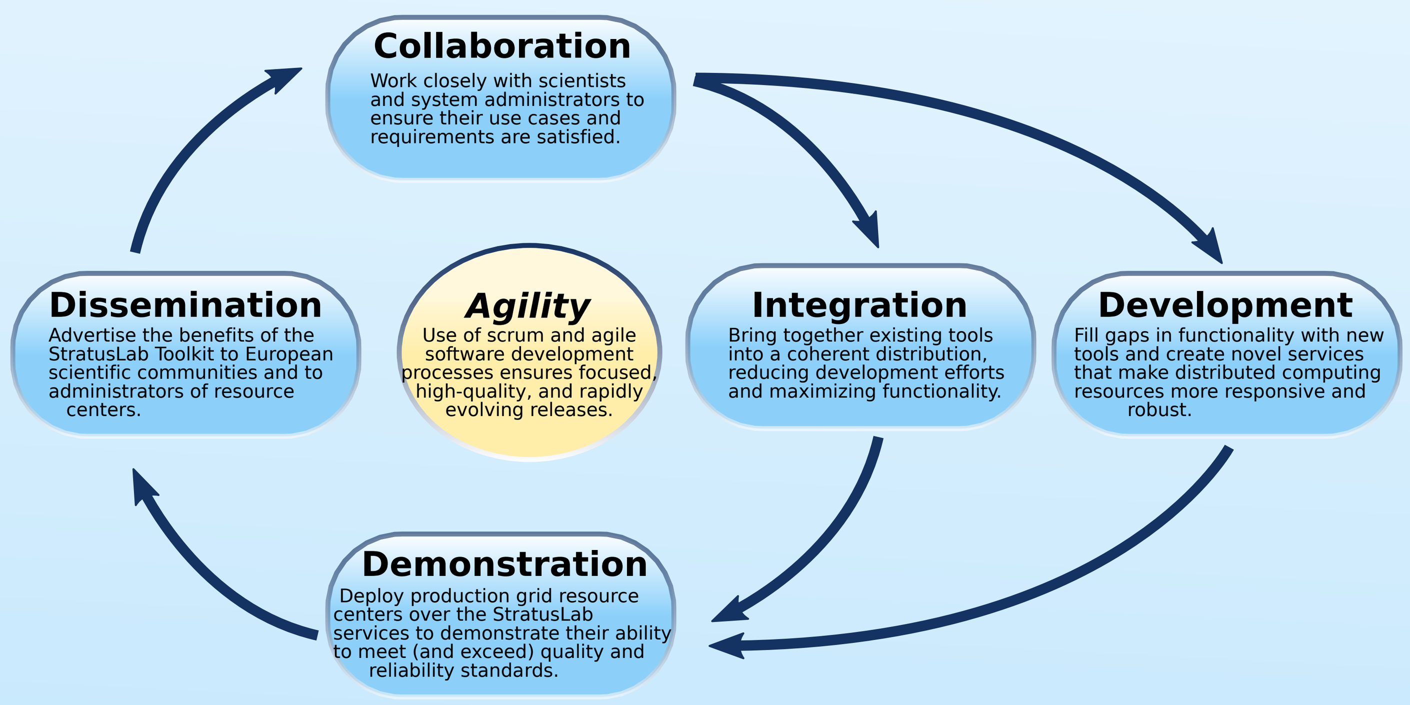
The StratusLab Toolkit will integrate cloud and virtualization technologies and services within grid sites and enrich existing computing infrastructures with “Infrastructure as a Service” (IaaS) provisioning paradigms.

StratusLab Distribution

StratusLab integrates, distributes and maintains a sustainable open-source cloud distribution to bring cloud to existing and new grid sites. The StratusLab distribution is composed of existing cutting-edge open source software and the innovative services and cloud management technologies developed in the project. It incrementally delivers a production grade distribution that is being demonstrated through the operation of production-level grid sites in the project.

Development Process

StratusLab has adopted the Scrum agile software development method, delivering releasable increments every three weeks, daily stand-up meetings, and collegial management of priorities. The result is the StratusLab Distribution, supporting public, private and hybrid clouds.



Benefits

StratusLab brings several benefits to the e-Infrastructure ecosystem, in terms of simplification, added flexibility, increased maintainability, quality, energy efficiency and resilience of computing sites. A key aim is for the cloud layer to be transparent to the layers above.

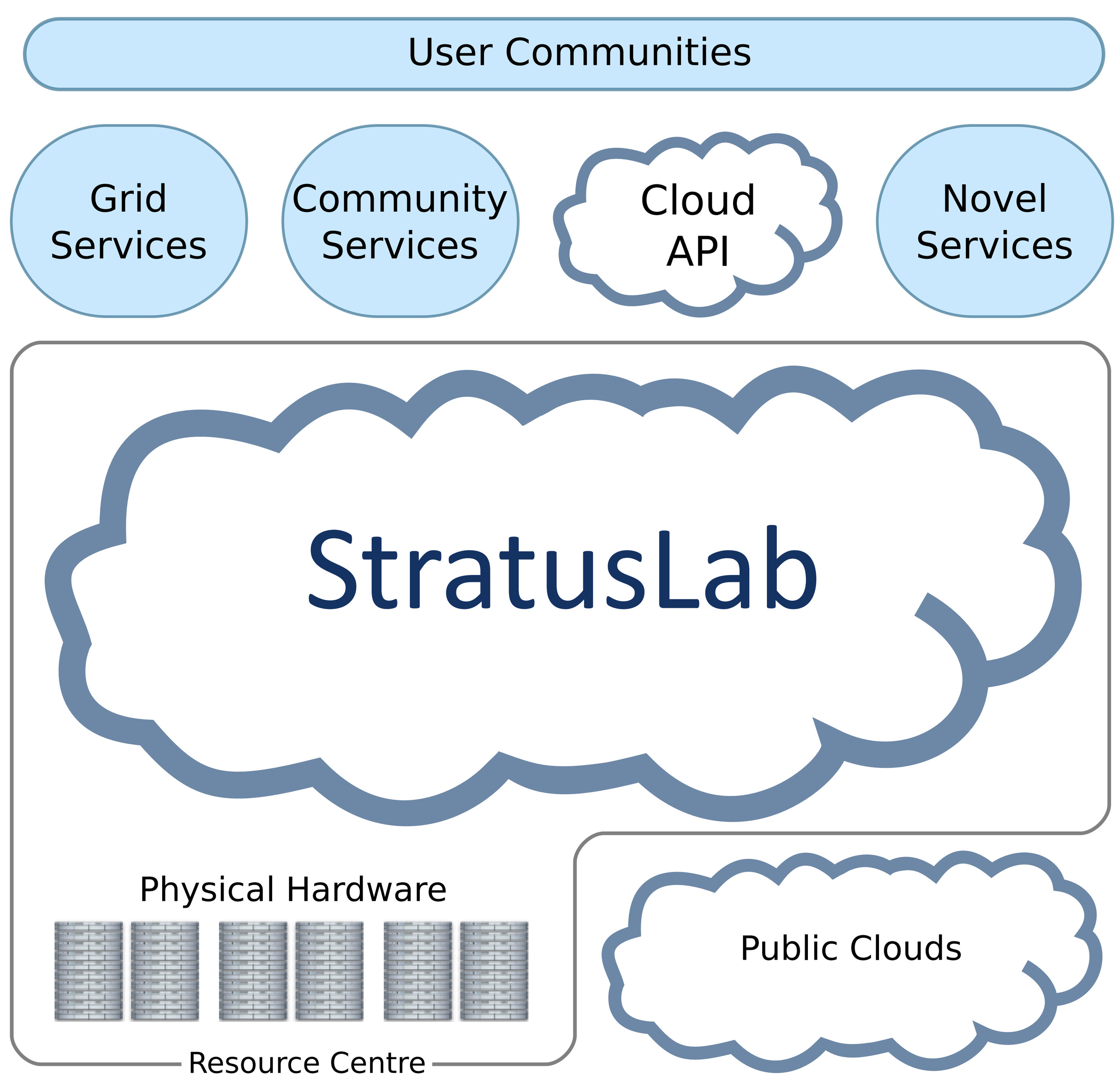
Impact

Existing grid middleware, as provided by EMI, continues to provide the glue to federate the distributed resources and the services for high-level job and data management. StratusLab improves the usability of distributed computing infrastructures; attracts scientific user communities; appeals equally to industrial users; keeps European research infrastructures at the technological forefront, and strengthens the know-how in virtualization and cloud computing of European industry.

Plan

StratusLab is a two-phase project. In the **first phase**, the project focuses on cloud computing for resource provisioning in grid sites. This entails development of the StratusLab cloud platform and creation of virtual appliances for the scientific application domains in the project.

In the **second phase** the emphasis will shift towards developing new cloud-like delivery paradigms in grid sites. This will build on the first phase, including new IaaS cloud interfaces that can be used for community services, directly by users, and to deploy novel cloud-based services.



Website: http://stratuslab.eu/
Contact: Charles Loomis
loomis@lal.in2p3.fr

Centre National de la Recherche Scientifique *France*
Universidad Complutense de Madrid *Spain*
Greek Research and Technology Network S. A. *Greece*
SixSq Sàrl *Switzerland*
Telefonica Investigacion y Desarrollo S.A. *Spain*
Trinity College Dublin *Ireland*

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

StratusLab Enhancing Grid Infrastructures with Virtualization and Cloud Technologies

Application Benchmarks

StratusLab has developed application-level benchmarks to test the stability, scalability, and performance of the StratusLab Toolkit. Equally, it will be used to validate different cloud deployment scenarios.

Running StratusLab Benchmarks

1) Define Virtual Machines
   - Define CPU and memory parameters for VM
   - VM based on SL5.5 image

2) Deploy VM in StratusLab cloud

3) Application parameters
   - Thread number
   - Timeout
   - Size of input and output files, ...

4) Execution
   - Configure networks, environments, ...
   - Install StratusLab benchmarks on VMs
   - Launch benchmarks
   - Retrieve output

Benchmarks

- Run successfully on the StratusLab cloud
- Are extensible (e.g. adding HEPSPEC, ...)
- Are important components of automated testing for StratusLab
- Are available in the StratusLab release

Benchmark Output

The benchmarks store their outputs in an XML format, which facilitates handling and analysis of the results.

The outputs contain information on: benchmark name, application input parameters, thread number, CPU time and elapsed time.

Example MPI Synchronous Benchmark Output File

```
<benchmark name='mpi_synchronous'>
 <parameters>
  <nb_threads>10</nb_threads>
 </parameters>
 <results>
  <elapsed_time unit='sec'>26.27</elapsed_time>
  <cpu_time unit='sec'>26.26</cpu_time>
</results>
</benchmark>
```

CPU Benchmarks

High-CPU usage with no input and no output data
- Fork Stress test
- CPU Stress test

I/O Benchmarks

- Writing to a file: MPI_FILE_write_at
- Reading from file: MPI_FILE_read_at
- Simultaneous: Reading from a file and writing to another

OpenMP Benchmarks

- Conjugate Gradient Method (linear system Ax=b)
- Jacobian type iteration (linear system Ax=b)
- Matrix multiplication (C=AB)

Workflow Benchmarks

- Use Kepler Platform for workflow design and management
- Workflow is built using the customizable components available in Kepler namely: directors, actors, parameters, relations and the ports.
- Workflow executes an OpenMP application

MPI Benchmarks

- MPI asynchronous non-blocking communication
- MPI synchronous blocking communication
- MPI persistent communication
- MPI standard communication

Website: http://stratuslab.eu/
Contact: Charles Loomis
loomis@lal.in2p3.fr

Centre National de la Recherche Scientifique *France*
Universidad Complutense de Madrid *Spain*
Greek Research and Technology Network S. A. *Greece*
SixSq Sàrl *Switzerland*
Telefonica Investigacion y Desarrollo S.A. *Spain*
Trinity College Dublin *Ireland*

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

StratusLab

Enhancing Grid Infrastructures with Virtualization and Cloud Technologies

StratusLab Software and Quattor Toolkit Integration

The StratusLab software must both install easily and permit management of virtual machines with automated site management tools. The Quattor Toolkit, heavily used by grid resource centers, has been chosen as the first example to demonstrate the compatibility with automated site management tools.

Automated StratusLab Installation

A complete StratusLab cloud can be installed and deployed automatically with Quattor, except for manual configuration of the OpenNebula network and node information.

Testing of Different Scenarios

Quattor is capable of managing ensembles of machines, so will be ideal for testing different deployment scenarios reliably and efficiently.

Virtual Machine Bookkeeping

Quattor keeps a complete record of a VM's configuration, allowing detailed bookkeeping, a growing problem with the poliferation of VMs and virtual appliances.

gLite Worker Nodes

These have been installed on the StratusLab cloud and run within a production grid site without any problems.

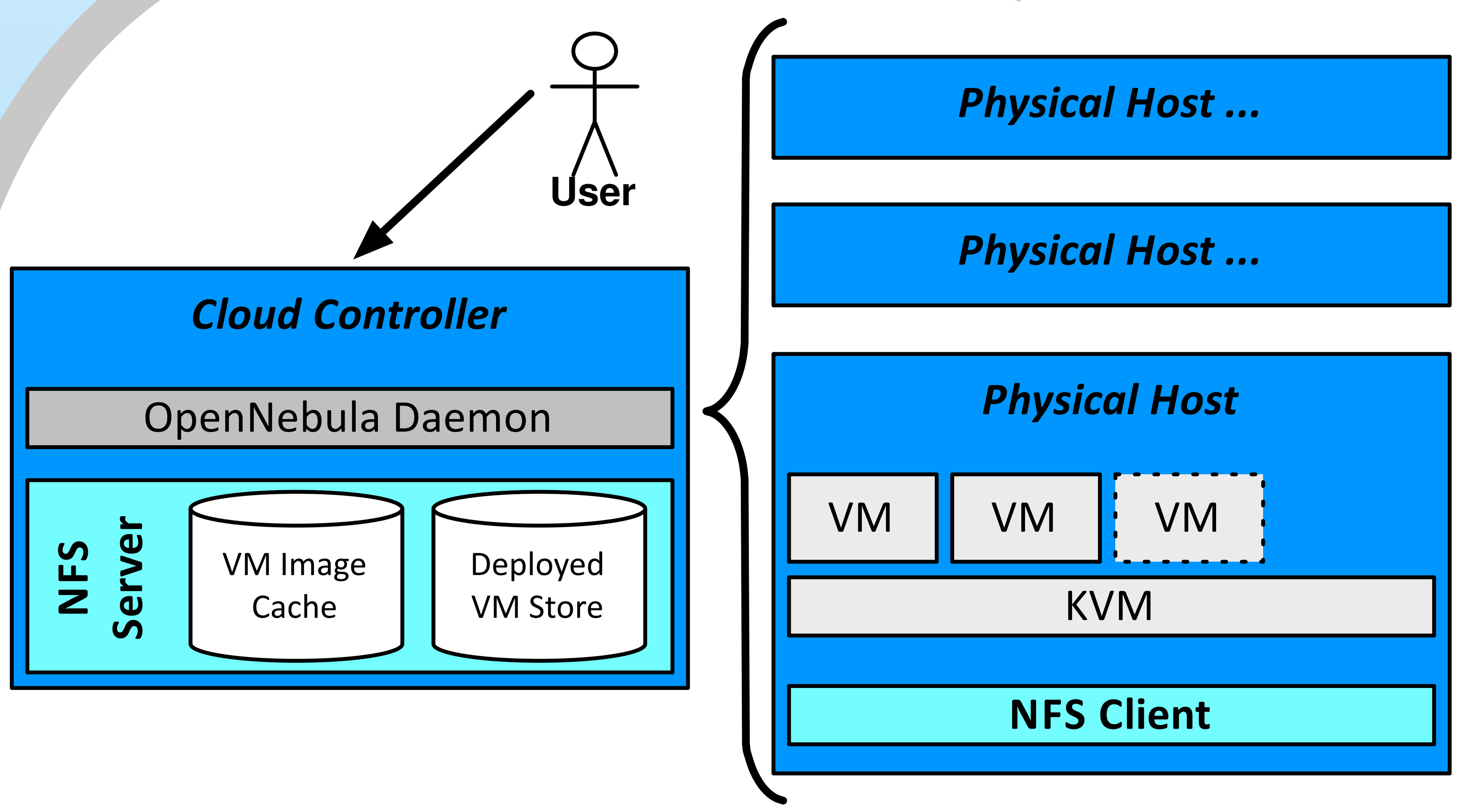
Quattor Toolkit

The Quattor Toolkit allows the automated installation, configuration, and management of an institute's machines.

System administrators describe the desired state of machines using the Pan configuration language. The language allows extensive validation of the site configuration before deploying changes to the machines.

Configuration modules running on each machine effect the desired changes to a machine's state.

More information can be found on the Quattor website: *http://quattor.org/*.



StratusLab Cloud at LAL

The CNRS Linear Accelerator Laboratory (LAL) in Orsay, France has installed its StratusLab cloud entirely with Quattor.

It consists of five 1U rack-mounted machines with dual Intel processors running at 2.33 GHz. Each machine has a total of 8 CPU cores and 16 GB of RAM.

Quattor Configuration Modules

Most configuration modules for the deployment already exist. Only two were created:

- *ncm-libvirtd*: Configuration module for virtualization daemon (libvirtd).
- *ncm-oned*: Configuration module for managing the OpenNebula daemon (oned) running on the cloud controller.

Known Limitations

Manual Configuration: Limited manual configuration of OpenNebula is required. Will be fixed with future updates.

NFS Mount Race Conditions: When installing a complete cloud, nodes sometimes start before the server is ready. Will use autoconf in the future.

Performance Issues: Server becomes unresponsive when launching several VMs simultaneously. Will use LVM and VM caching to reduce latencies.

Website: http://stratuslab.eu/
Contact: Charles Loomis
loomis@lal.in2p3.fr

Centre National de la Recherche Scientifique, France
Universidad Complutense de Madrid, Spain
Greek Research and Technology Network S. A., Greece
SixSq Sàrl, Switzerland
Telefonica Investigacion y Desarrollo S.A., Spain
Trinity College Dublin, Ireland

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

StratusLab

Enhancing Grid Infrastructures with Virtualization and Cloud Technologies

Agile Software Development

StratusLab uses the agile methodology Scrum to manage its development. This allows the project to evolve the StratusLab distribution quickly according to user and system administrator requirements while always maintaining a functioning release.

Sprints last 3 weeks on average, with a public beta release produced every other sprint. Three public releases have been made with a fourth expected at the end of April. The first production release (v1.0) is planned for the end of May 2011.

Advantages

- New features and improvements are released often
- Continuous feedback from users on working software informs design evolution
- Users influence requirements and their priority based on real use of earlier versions
- Users see progress as new iterative releases are made
- No 'big bang' integration required, significantly reducing risks
- End of sprint demo forces developers to integrate and show their work using a functional system
- Project advancement measured based on facts instead of subjective evaluation

Challenges

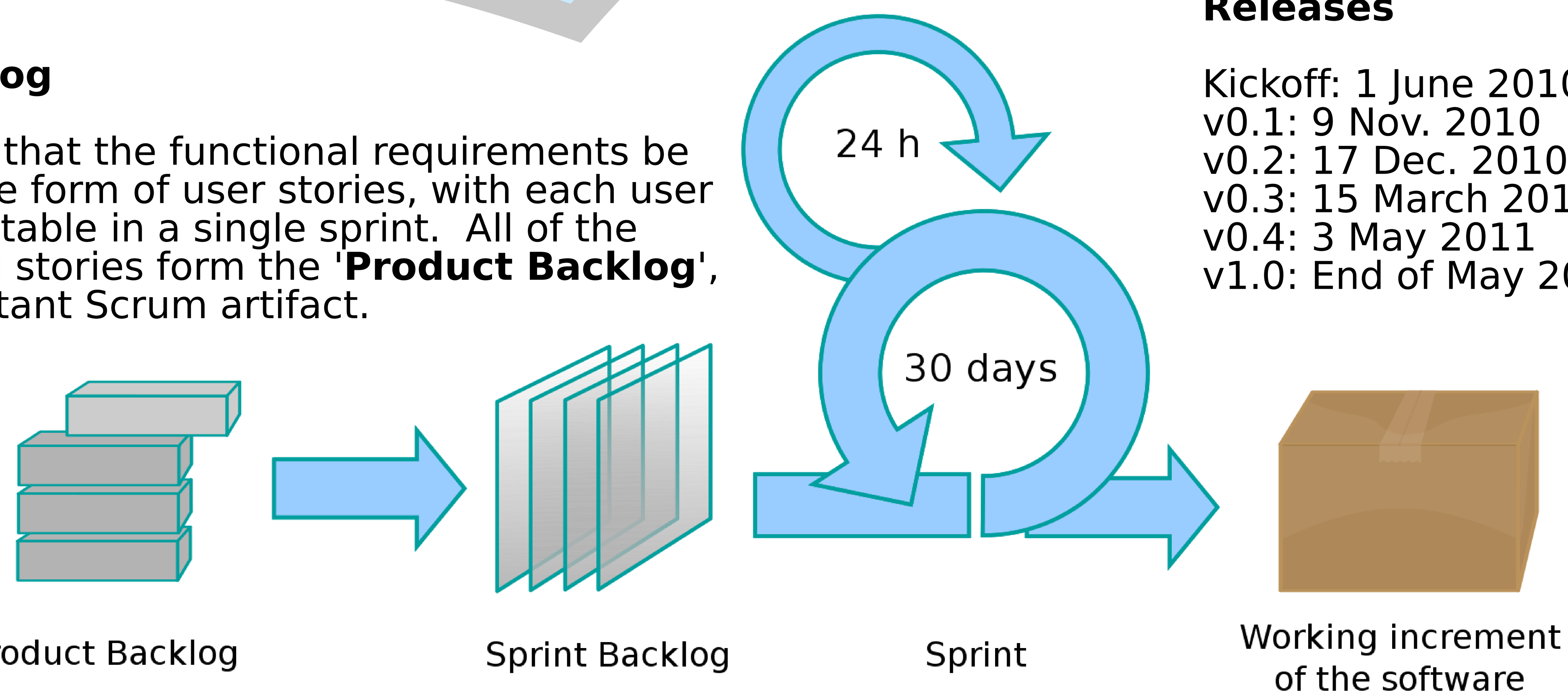
- Regular release requires higher level of automation: build procedures, test procedures, and production upgrades.
- Functional increments must be broken down into tasks that can be implemented in a single sprint
- Developer's mindset has to adapt to incremental development, rather than relying on heavyweight up-front design

Product Backlog

Scrum requires that the functional requirements be expressed in the form of user stories, with each user story implementable in a single sprint. All of the unimplemented stories form the '**Product Backlog**', the most important Scrum artifact.

Releases

Kickoff: 1 June 2010
v0.1: 9 Nov. 2010
v0.2: 17 Dec. 2010
v0.3: 15 March 2011
v0.4: 3 May 2011
v1.0: End of May 2011



Planning Meetings

Each sprint starts with a planning meeting. During this meeting, user stories from the product backlog are reviewed and selected for the sprint.

The items selected for the sprint become the '**Sprint Backlog**'.

Daily Meetings

In order to ensure a fluid communication during the sprint, we use a 'daily stand-up' meeting. This meeting, which never exceeds 15 minutes, is the place where each team reports on work completed the previous day, what it plans to work on next, and any impediments it has or foresees.

Longer topics that require further discussion are scheduled offline or just after the stand-up.

Demo Meetings

Each sprint concludes with a live demo of each user story implemented during the sprint. This is a powerful way of measuring real progress and ensuring that developments can be released in production.

Website: http://stratuslab.eu
Support: support@stratuslab.eu
Contact: contact@stratuslab.eu

Centre National de la Recherche Scientifique *France*
Universidad Complutense de Madrid *Spain*
Greek Research and Technology Network S. A. *Greece*
SixSq Sàrl *Switzerland*
Telefonica Investigacion y Desarrollo S.A. *Spain*
Trinity College Dublin *Ireland*

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

Client Installation and Cloud Installation

The project aims to make the installation of the StratusLab client and of a StratusLab cloud as straight-forward and painless as possible, minimizing dependencies and supporting multiple installation methods.

Client Installation

The StratusLab client consists of a number of command line scripts written in Python. They are portable and can be used on Windows, Mac OSX, and all versions of Linux.

Dependencies

- Python 2.6 or later (but not Python 3.x!)
- Java 1.6 or later

Tarball Installation

1) Simply follow the download link on the StratusLab website to find the latest tarball distribution and download it.

2) Unroll it into a convenient directory.

3) Set the PATH variable to point to the "bin" subdirectory. (Use "windows" subdirectory for a Windows machine.)

4) Set the PYTHONPATH variable to point to the "lib/stratuslab/python" subdirectory.

5) Use the stratus-* commands like any other command.

Packaged Client

The client is also available as an RPM package for CentOS 5.5 and compatible distributions. Just setup your machine to point to the StratusLab yum repository (v0.3).

Install the stratuslab-client-cli RPM with yum.

The commands will then be available to all users of the system.

Manual Installation

A StratusLab cloud can be installed manually with assistance from a set of scripts to automate the process.

- Define parameters in the StratusLab configuration file.
- Ensure that external dependencies are available.
- Run the system installation scripts.
- Test that the cloud is functioning correctly.

Quattor Installation

To ensure that the StratusLab distribution easily integrates with automated site management tools, a Quattor configuration for a StratusLab cloud is also maintained by the project.

- Download the Quattor configuration.
- Define the machine profiles for the cloud frontend and execution hosts.
- Trigger installation of the machines.
- Test that the cloud is functioning.

StratusLab Commands

```
stratus-run-instance
stratus-sign-metadata
stratus-verify-metadata
...
```


Centre National de la Recherche Scientifique *France*
Universidad Complutense de Madrid *Spain*
Greek Research and Technology Network S. A. *Greece*
SixSq Sàrl *Switzerland*
Telefonica Investigacion y Desarrollo S.A. *Spain*
Trinity College Dublin *Ireland*

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

StratusLab Enhancing Grid Infrastructures with Virtualization and Cloud Technologies

Models for Integrating Grid and Cloud Technologies

Both cloud and grid technologies have features that make them attractive for technical computing tasks. An improved e-Infrastructure can be provided to European researchers by combining both technologies on a single pan-European platform.

Cloud Features

- Dynamic, instantaneous provisioning of resources.
- Ability to customize execution environment and installed applications.

Grid Features

- Robust security and policy framework permitting global authentication and authorization of users.
- Mechanisms for federating distributed resources to produce a larger, more capable infrastructure for users.

A combined e-Infrastructure:

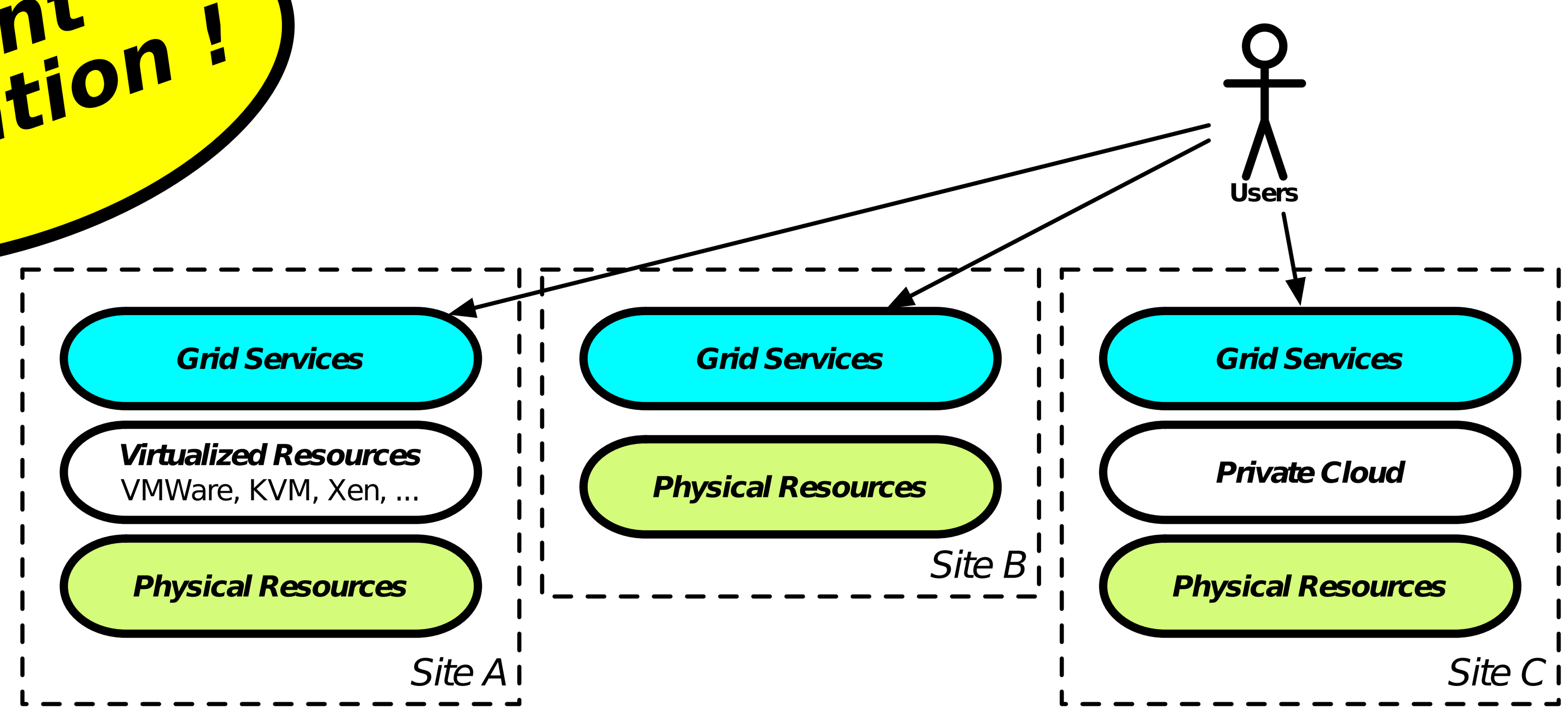
- Appeals to a larger, more diverse scientific community via customized environments and alternate operating systems.
- Allows VO and user-level services permitting the construction of complete scientific platforms and services.
- Allows sharing of same physical resources with different access modes without partitioning of resources.
- Provides more flexible management of services and resources.

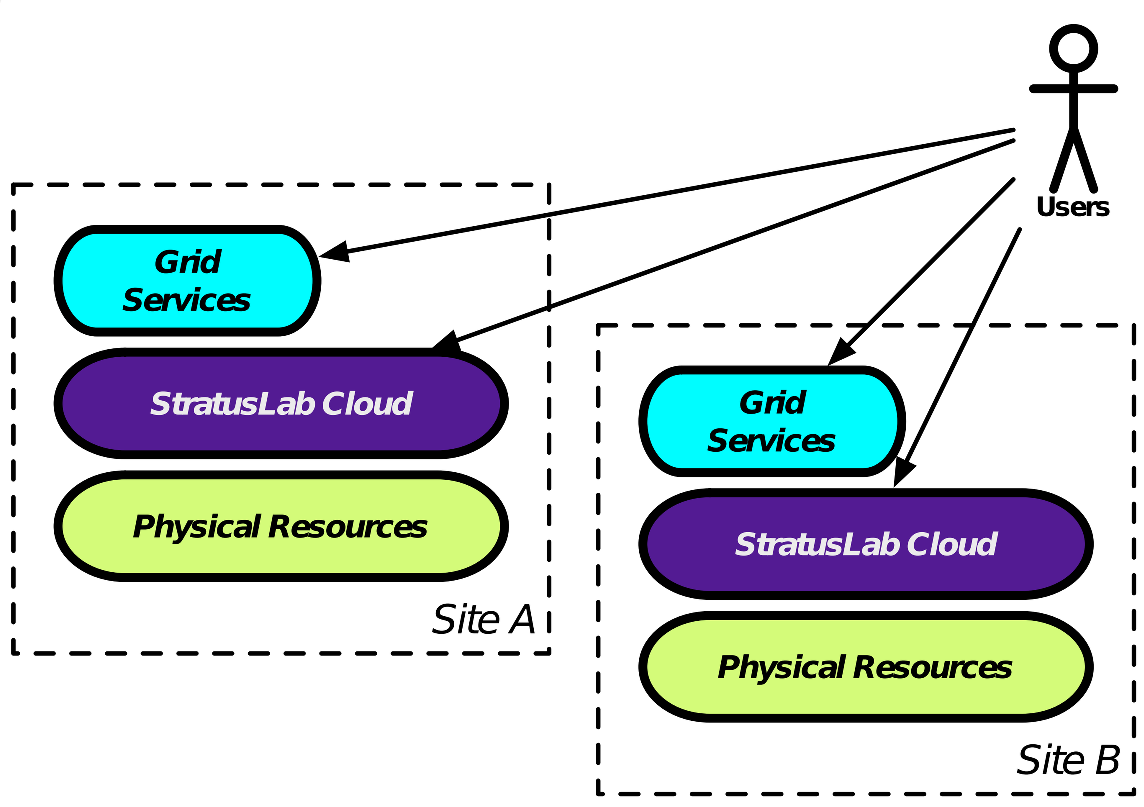
Current Situation !

Model 1: Hidden Virtualization

Site administrators use virtualization and cloud technologies but do not make this visible to grid users.

Site administrators benefit from easier system management; users do not directly benefit.

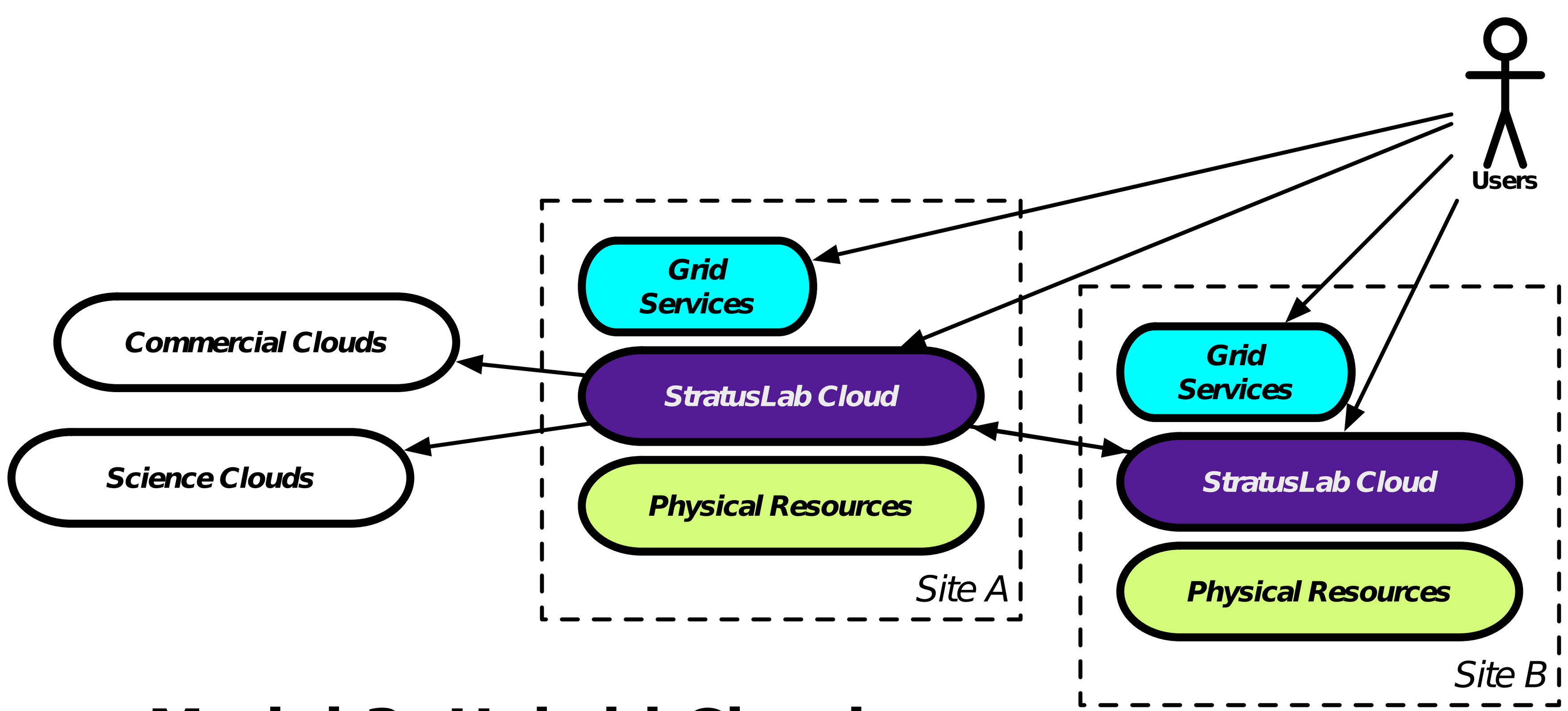




Model 2: Public Cloud and Grid Services

Each site deploys a cloud over its physical resources running the grid services within that cloud with the cloud visible to end-users.

Administrators see the benefits of Model 1 and end-users can take advantage of the customization and dynamic provisioning of clouds.



Model 3: Hybrid Clouds

Like Model 2, but clouds can now transparently use resources from other StratusLab clouds, commercial clouds or science clouds.

All the benefits of Model 2, but more resources can be made available and sharing of resources is permitted between sites, not just between users and VOs.

Website: http://stratuslab.eu
Support: support@stratuslab.eu
Contact: contact@stratuslab.eu

Centre National de la Recherche Scientifique *France*
Universidad Complutense de Madrid *Spain*
Greek Research and Technology Network S. A. *Greece*
SixSq Sàrl *Switzerland*
Telefonica Investigacion y Desarrollo S.A. *Spain*
Trinity College Dublin *Ireland*

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

StratusLab Enhancing Grid Infrastructures with Virtualization and Cloud Technologies

Marketplace: Registry for Machine & Disk Images

Making secure machine images for virtualized and cloud infrastructures remains a significant hurdle for wider adoption and use of these technologies. The StratusLab Marketplace lowers this barrier by facilitating the sharing and reuse of existing images.

Benefits for End-Users
Scientists, engineers, and other users of StratusLab cloud infrastructures can browse the Marketplace to find existing images that have the operating systems and applications to run their analyses, saving them the time needed to create their own images.

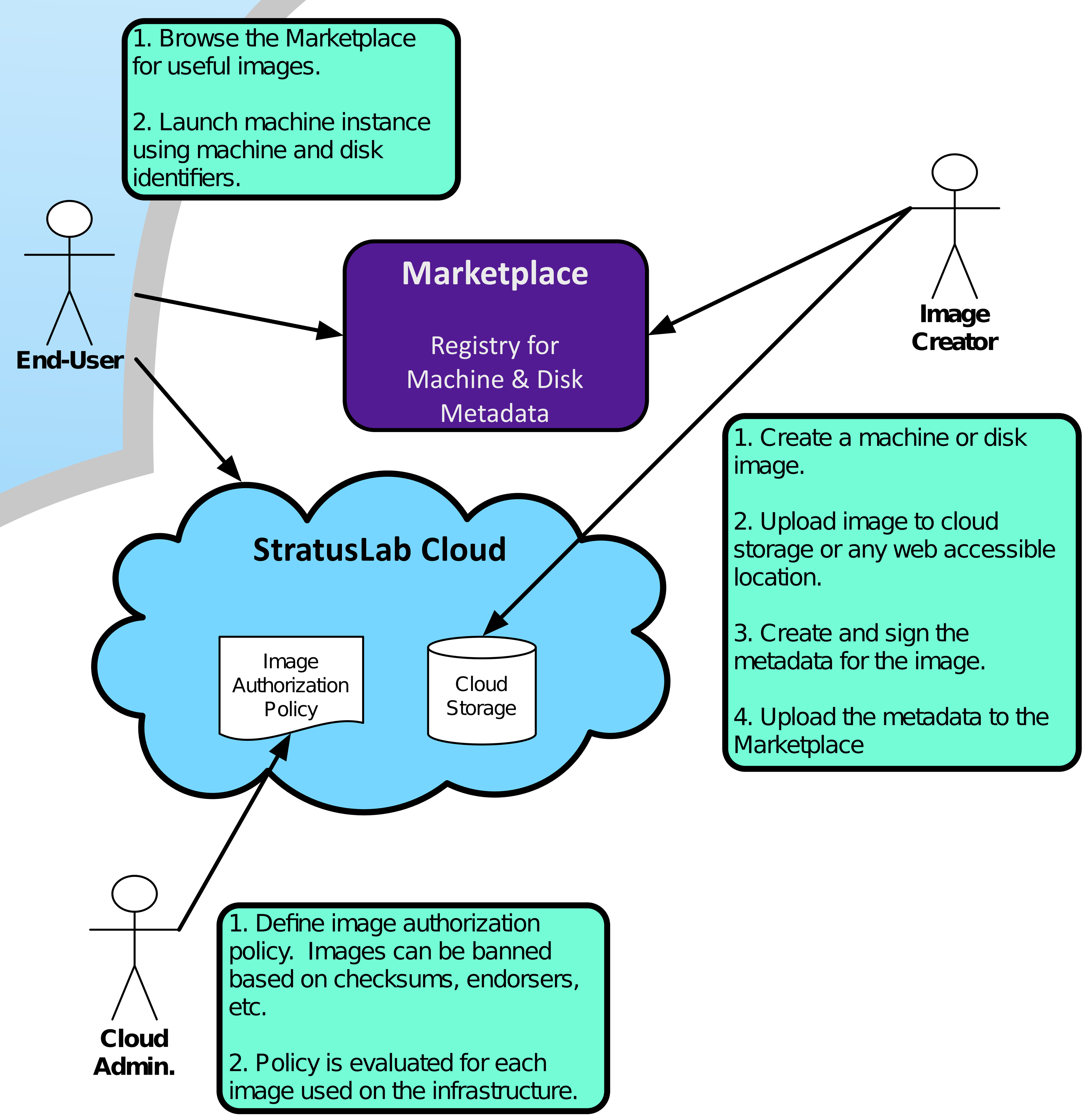
Benefits for Image Creators
People creating machine and disk images can publish metadata about them in the Marketplace, making them accessible to a wider audience and soliciting feedback to further improve those images.

Benefits for Cloud Administrators
Those running cloud infrastructures can use image metadata from the Marketplace to evaluate whether they can trust a particular machine or disk image, banning if necessary, images which do not meet the administrator's requirements.

Image Contents
Contents are uploaded into cloud storage or any other web accessible location; the Marketplace does not store the image contents. The link between the metadata and the image is based on the SHA-1 checksum.

Image Metadata
Existing standards have been used to provide a lightweight, extensible metadata format that allows searching and selection based on a variety of attributes. (See example below.)

Standards
XML (W3C)
XML Signature (W3C)
RDF-XML (W3C)
Dublin Core Metadata (DCMI)



XML Namespaces / Image Identifiers / Image Endorser / General Image Information / Additional Information / Cryptographic Signature


```
<rdf:RDF
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:ex="http://example.org/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:slreq="http://mp.stratuslab.eu/slreq#"
  xmlns:slterms="http://mp.stratuslab.eu/slterms#"
  xml:base="http://mp.stratuslab.eu/">

  <rdf:Description rdf:about="#MMZu9WvwKIro-rtBQfDk4PsKO7_">

    <dcterms:identifier>MMZu9WvwKIro-rtBQfDk4PsKO7_</dcterms:identifier>
    <slreq:bytes>100</slreq:bytes>
    <slreq:checksum rdf:parseType="Resource">
      <slreq:algorithm>SHA-1</slreq:algorithm>
      <slreq:value>c319bbd5afc0a22ba3eaed0507c39383ec28eeff</slreq:value>
    </slreq:checksum>

    <slreq:endorsement rdf:parseType="Resource">
      <created xmlns="http://purl.org/dc/terms/">2011-03-25T21:25:07Z</create
      <endorser xmlns="http://mp.stratuslab.eu/slreq#" rdf:parseType="Resourc
        <email>loomis@lal.in2p3.fr</email>
        <subject>CN=Charles Loomis,OU=LAL,O=CNRS,C=FR,O=GRID-FR</subject>
        <issuer>CN=GRID2-FR,O=CNRS,C=FR</issuer>
      </endorser>
    </slreq:endorsement>

    <dcterms:type>machine</dcterms:type>
    <dcterms:valid>2011-07-23T10:59:42Z</dcterms:valid>
    <dcterms:publisher>StratusLab</dcterms:publisher>

    <dcterms:title>linux-with-my-apps</dcterms:title>
    <dcterms:description>A 32-bit ttylinux...</dcterms:description>
    <slterms:serial-number>0</slterms:serial-number>
    <slterms:version>1.0</slterms:version>

    <slterms:hypervisor>kvm</slterms:hypervisor>
    <slterms:inbound-port>443</slterms:inbound-port>
    <slterms:outbound-port>25</slterms:outbound-port>
    <slterms:icmp>8</slterms:icmp>

    <slterms:os>ttylinux</slterms:os>
    <slterms:os-version>9.7</slterms:os-version>
    <slterms:os-arch>i486</slterms:os-arch>

    <ex:other-info>additional metadata</ex:other-info>
    <ex:yet-more>still more info</ex:yet-more>
    <ex:relatedImages rdf:parseType="Resource">
      <dcterms:identifier>MMZu9WvwKIro-rtBQfDk4PsKO7_</dcterms:identifier>
      <dcterms:identifier>NMZu9WvwKIro-rtBQfDk4PsKO7_</dcterms:identifier>
    </ex:relatedImages>
  </rdf:Description>

  <Signature xmlns="http://www.w3.org/2000/09/xmldsig#">
    <SignedInfo>
      <CanonicalizationMethod Algorithm="http://www.w3.org/..."/>
      <SignatureMethod Algorithm="http://www.w3.org/2000/09/xmldsig#rsa-sha1"
      <Reference URI="">
        <Transforms>
          <Transform Algorithm="http://www.w3.org/2000/09/xmldsig#..."/>
        </Transforms>
        <DigestMethod Algorithm="http://www.w3.org/2000/09/xmldsig#sha1"/>
        <DigestValue>go6VR...</DigestValue>
      </Reference>
    </SignedInfo>
    <SignatureValue>NrTJFF...</SignatureValue>
    <KeyInfo>
      <X509Data>...
        <X509Certificate>...</X509Certificate>
      </X509Data>
    </KeyInfo>
  </Signature>

</rdf:RDF>
```


Website: http://stratuslab.eu
Support: support@stratuslab.eu
Contact: contact@stratuslab.eu

Centre National de la Recherche Scientifique *France*
Universidad Complutense de Madrid *Spain*
Greek Research and Technology Network S. A. *Greece*
SixSq Sàrl *Switzerland*
Telefonica Investigacion y Desarrollo S.A. *Spain*
Trinity College Dublin *Ireland*

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

StratusLab

Enhancing Grid Infrastructures with Virtualization and Cloud Technologies

Cloud Storage: Disk and File-based Abstractions

The project will provide cloud storage services with both disk- and file-based abstractions. Disk-based abstractions allow both read-only distribution of disk images through the StratusLab Marketplace and persistent read-write disks within a particular cloud instance.

The project will also provide file-based storage services (likely with CDMI interface) to allow the storage and (potential) sharing of files within and between clouds. Grid users can continue to use the standard SRM interfaces for files stored on the grid.

File-Based Storage

For an IaaS cloud like StratusLab, the most natural abstraction is a disk. However for scientific analyses, file-based abstractions are often more convenient.

After v1.0, the project will provide file-based storage, probably with the CDMI interface proposed by SNIA.

Users can continue to use standard file-based grid services.

Shared Read-Only Disks

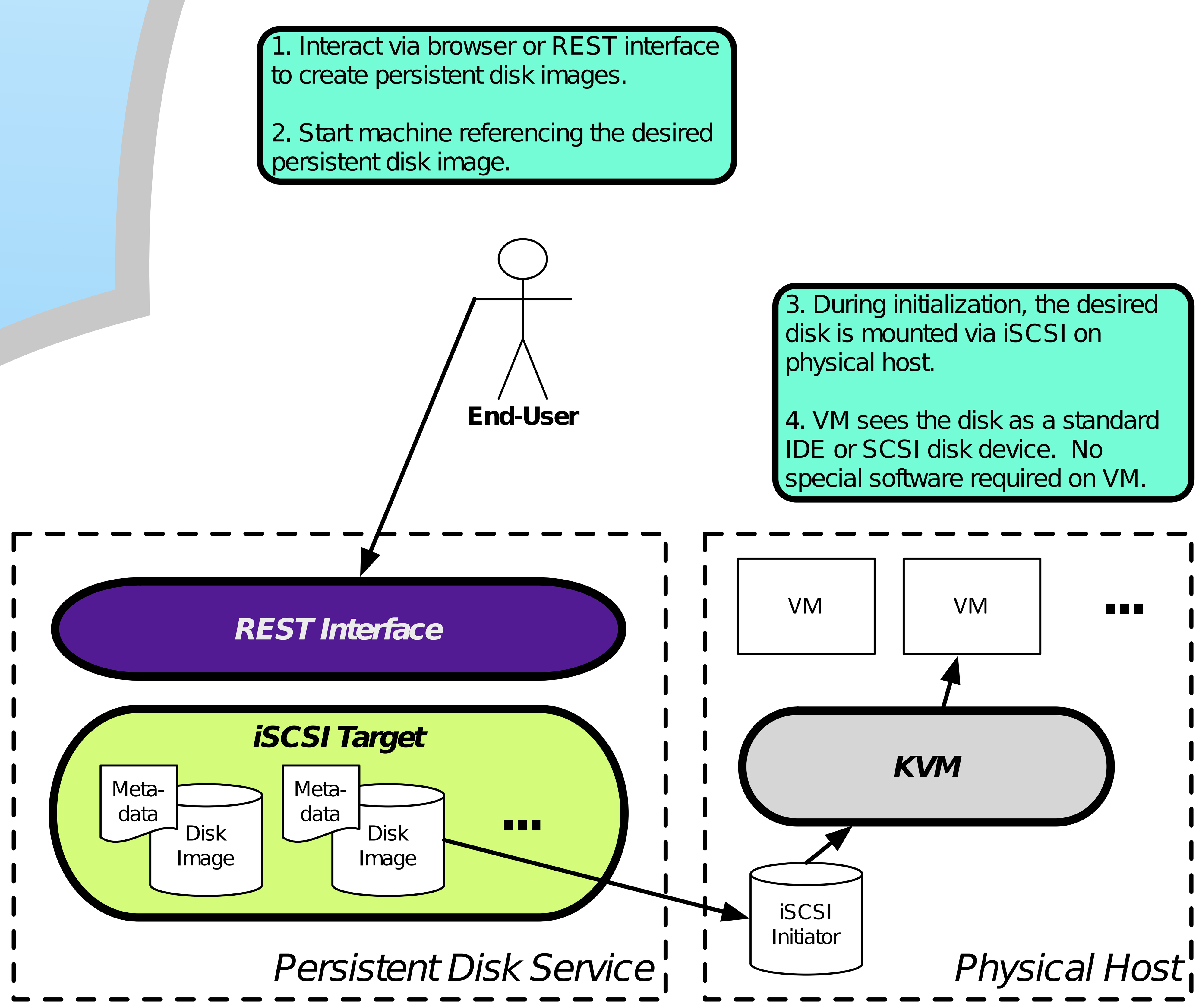
It is often useful to share datasets via read-only disks. This allows caching of these disks and sharing between different cloud instances. It also permits users to use specific versions of the database for their analyses.

These are supported though the standard Marketplace mechanisms.

Persistent Disks

These may be created and used by a single machine instance within a single cloud. They are useful for saving the persistent state of services.

Eventually snapshotting of these images will be provided.



Using a Persistent Disk

1. Create a disk with a given size via the web or REST interface.

2. Launch a machine instance referencing that disk image.

3. Partition (fdisk) and format (mkfs) the disk via the running virtual machine.

4. Store data to the disk as usual.

5. Dismount the disk or halt the machine instance.

6. Disk with the persistent data is available for use by another machine instance.

Using a Shared Read-Only Disk

1. Launch a machine giving the Marketplace reference for the image.

2. Machine is launched with the given disk image on the given device (typically /dev/hdd).

3. Mount the disk on the machine and use the data for an analysis.

4. Dismount or halt the machine instance.

The disk is read-only and no changes can be made to the disk image.

Website: http://stratuslab.eu
Support: support@stratuslab.eu
Contact: contact@stratuslab.eu

Centre National de la Recherche Scientifique *France*
Universidad Complutense de Madrid *Spain*
Greek Research and Technology Network S. A. *Greece*
SixSq Sàrl *Switzerland*
Telefonica Investigacion y Desarrollo S.A. *Spain*
Trinity College Dublin *Ireland*

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

StratusLab Enhancing Grid Infrastructures with Virtualization and Cloud Technologies

SlipStream™ Integration with StratusLab

SlipStream, by SixSq, deploys runtime environments in a cloud automatically, on-demand, and releases the resources after use. SlipStream users do not have to worry about configuring complex hardware and software to create personal runtime environments, representative of production conditions.
SlipStream first used Amazon EC2 for cloud resources, but is now been integrated with the StratusLab cloud. This allows SlipStream and the cloud resources to be deployed on entirely private infrastructures.

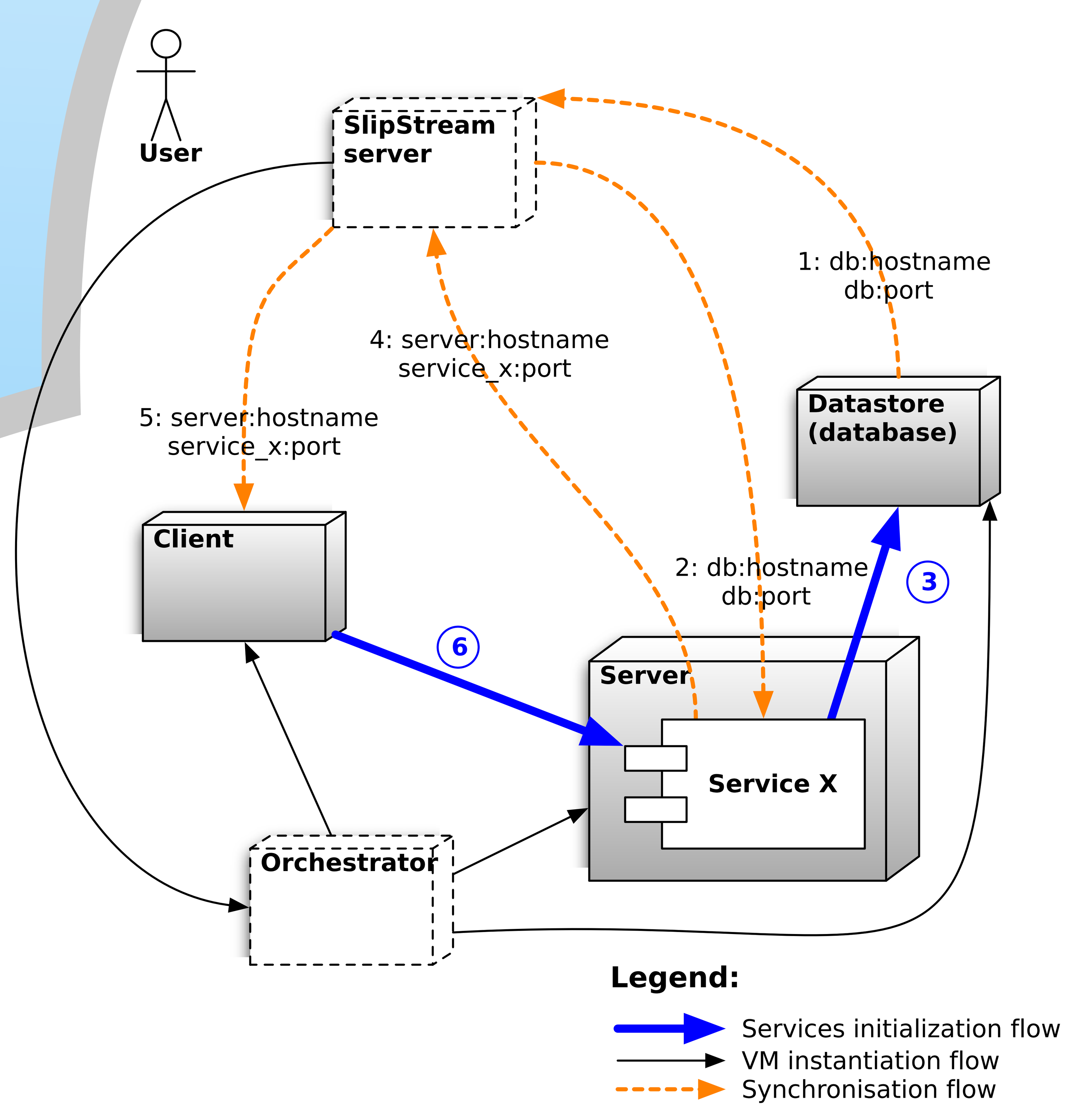
How does it work?

SlipStream is accessed via a secure Web 2.0 application. Users can organize their work in project hierarchies. Each SlipStream artifact is version controlled, providing complete history of changes. Two main workflows are available in SlipStream:

1. virtual machine (and disk) creation;
2. deployment and service synchronization.

The first flow allows a user to specify the content of each virtual machine. The second flow includes the synchronization between inter-dependent software components.

sixsq SlipStream™ Release with confidence



Deployment overview

Hybrid Cloud Solutions

SlipStream is already ready for multi-cloud usage, which is an important use-case for the second year of StratusLab. This means that users will be able to leverage their virtual machine creation recipes, as well as their deployment models, across different clouds. This feature also gives site administrators the possibility to re-generate virtual machines for different cloud services, on-the-fly.

"SlipStream is to software what the robotic assembly line is to manufacturing"

Integration lessons learned

SlipStream relies on contextualisation to control the boot sequence of VMs. StratusLab's contextualisation was improved to fulfill SlipStream's requirement resulting in a cleaner and generic contextualisation solution applicable to all types of operating systems. This feature is now available via the standard stratus-run-instance and stratus-create-image commands. All StratusLab base images have also been upgraded to support this feature. More specifically, any script running on a VM can now access contextualisation informtion driven by the user at instantiation time.
The VM creation process was also improved and optimised to shorten the machine creation process.
Finally, the disk layout of all base machines was simplified and unified to make machine creation simpler and more consistent.

"Declare war on pre-release stress. Release often and under your own terms. Improve software quality. Reduce time-to-market. These are a few reasons why we hope you'll enjoy using SlipStream!"

Website: http://stratuslab.eu
Support: support@stratuslab.eu
Contact: contact@stratuslab.eu

Centre National de la Recherche Scientifique *France*
Universidad Complutense de Madrid *Spain*
Greek Research and Technology Network S. A. *Greece*
SixSq Sàrl *Switzerland*
Telefonica Investigacion y Desarrollo S.A. *Spain*
Trinity College Dublin *Ireland*

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

Enhancing Grid Infrastructures with Virtualization and Cloud Technologies

Publicly Available IaaS Cloud Service

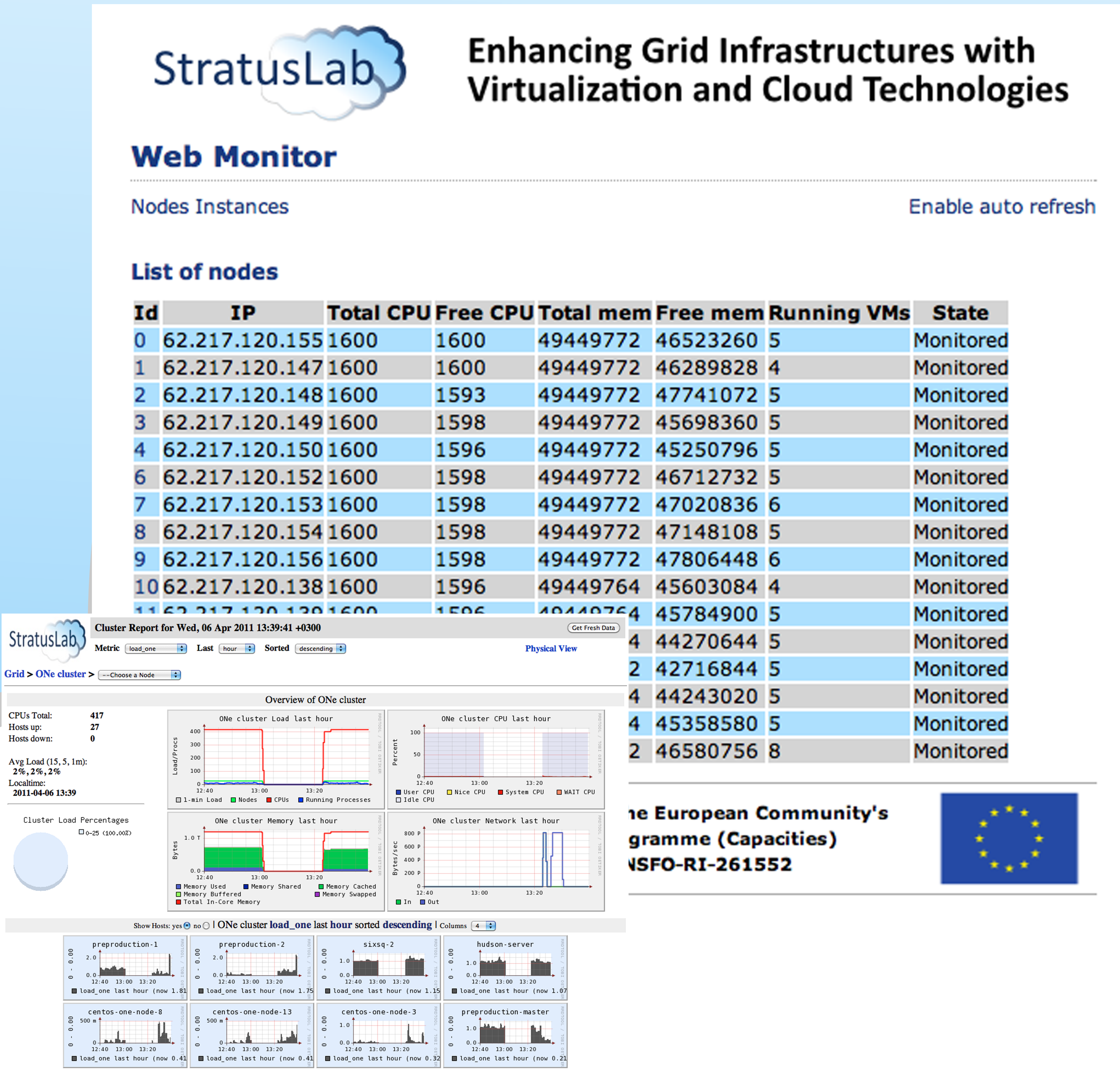
Reference Cloud service build with the latest version of StratusLab distribution. Deployed in GRNET's datacenter. Provides public access to a cloud testbed for anyone wishing to try the tools developed by the project.

Service Features

- 17 physical nodes (1 Frontend + 16 hosts)
- 256 cores
- 768 GB total memory (48 GB per node)
- 3.6 TB shared storage over NFS
- 120 public IPs available (private and local IPs to be supported soon)
- Currently running StratusLab v0.3 on CentOS 5.5.
- Uses OpenNebula 2.2 over kvm for VM management

Production grid site over cloud

- HG-07-StratusLab certified in Greek NGI
- Installed and operated on StratusLab's reference cloud service
- Offers 1 CE, 8 WNs (dual-core), 1 SE and 1 gLite-APEL node
- VM appliances pre-installed with grid middleware available from the appliance repository
- Currently working on grid site elasticity features



Id	IP	Total CPU	Free CPU	Total mem	Free mem	Running VMs	State
0	62.217.120.155	1600	1600	49449772	46523260	5	Monitored
1	62.217.120.147	1600	1600	49449772	46289828	4	Monitored
2	62.217.120.148	1600	1593	49449772	47741072	5	Monitored
3	62.217.120.149	1600	1598	49449772	45698360	5	Monitored
4	62.217.120.150	1600	1596	49449772	45250796	5	Monitored
6	62.217.120.152	1600	1598	49449772	46712732	5	Monitored
7	62.217.120.153	1600	1598	49449772	47020836	6	Monitored
8	62.217.120.154	1600	1598	49449772	47148108	5	Monitored
9	62.217.120.156	1600	1598	49449772	47806448	6	Monitored
10	62.217.120.138	1600	1596	49449764	45603084	4	Monitored
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	45784900	5	Monitored
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	44270644	5	Monitored
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	42716844	5	Monitored
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	44243020	5	Monitored
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	45358580	5	Monitored
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	46580756	8	Monitored

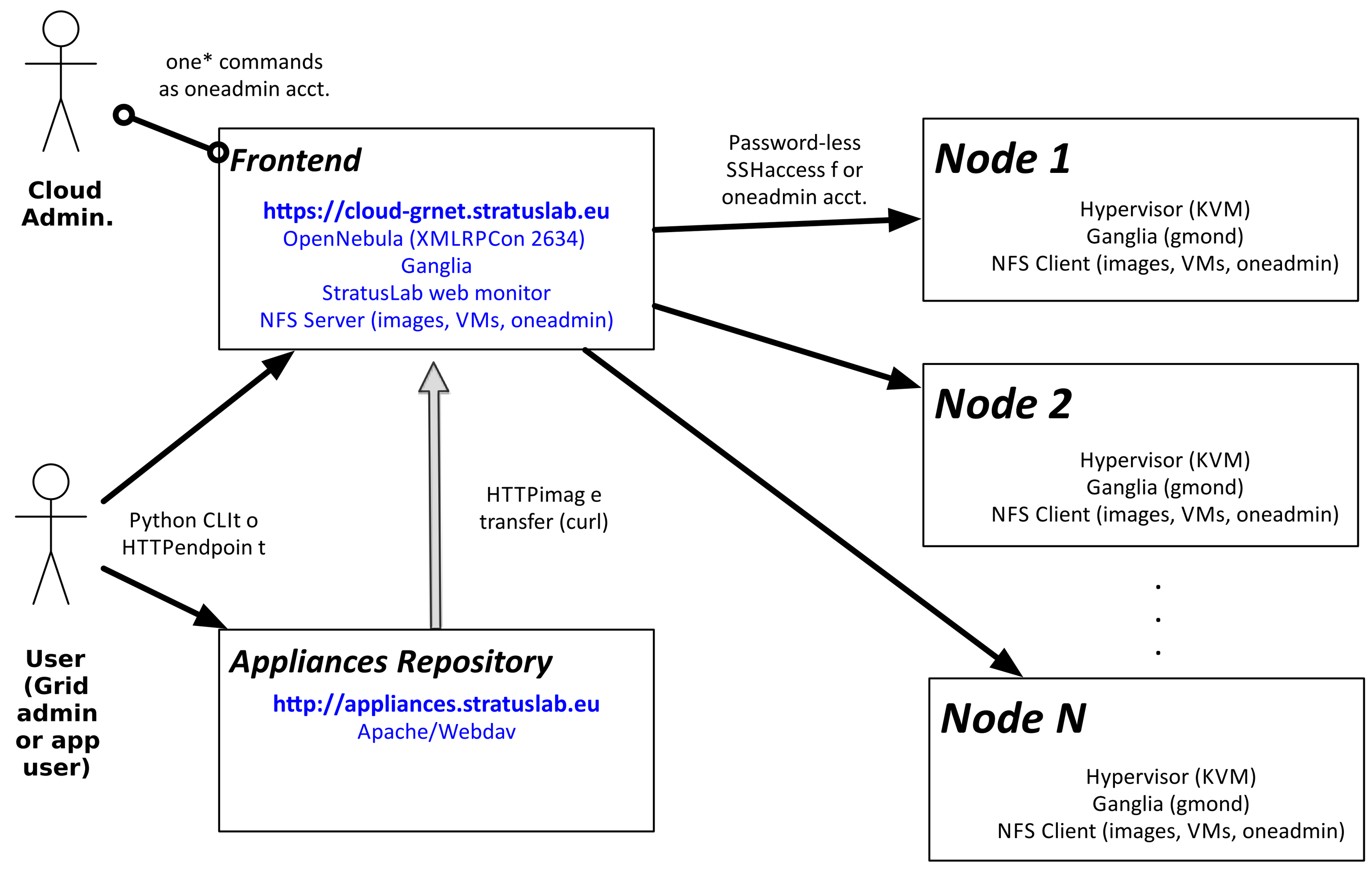
Monitoring

Web monitor provides overview of hosting nodes and VMs currently running in the cloud.

Ganglia used for monitoring the physical infrastructure.

How to access and use the service

- Send email to support@stratuslab.eu providing your details and purpose of usage
- Connect using username/password or your grid certificate (preferable)
- Manage your VMs using the StratusLab command line interface (stratuslab-* tools)
- Appliances and base images available from public repository operated by TCD (you can always use your own images!)



Website: http://stratuslab.eu
Support: support@stratuslab.eu
Contact: contact@stratuslab.eu

Centre National de la Recherche Scientifique *France*
Universidad Complutense de Madrid *Spain*
Greek Research and Technology Network S. A. *Greece*
SixSq Sàrl *Switzerland*
Telefonica Investigacion y Desarrollo S.A. *Spain*
Trinity College Dublin *Ireland*

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

StratusLab
Enhancing Grid Infrastructures with Virtualization and Cloud Technologies

StratusLab is developing and deploying cloud technologies with the aim of simplifying and optimizing the use and operation of distributed computing infrastructures such as the European Grid Infrastructure (EGI).

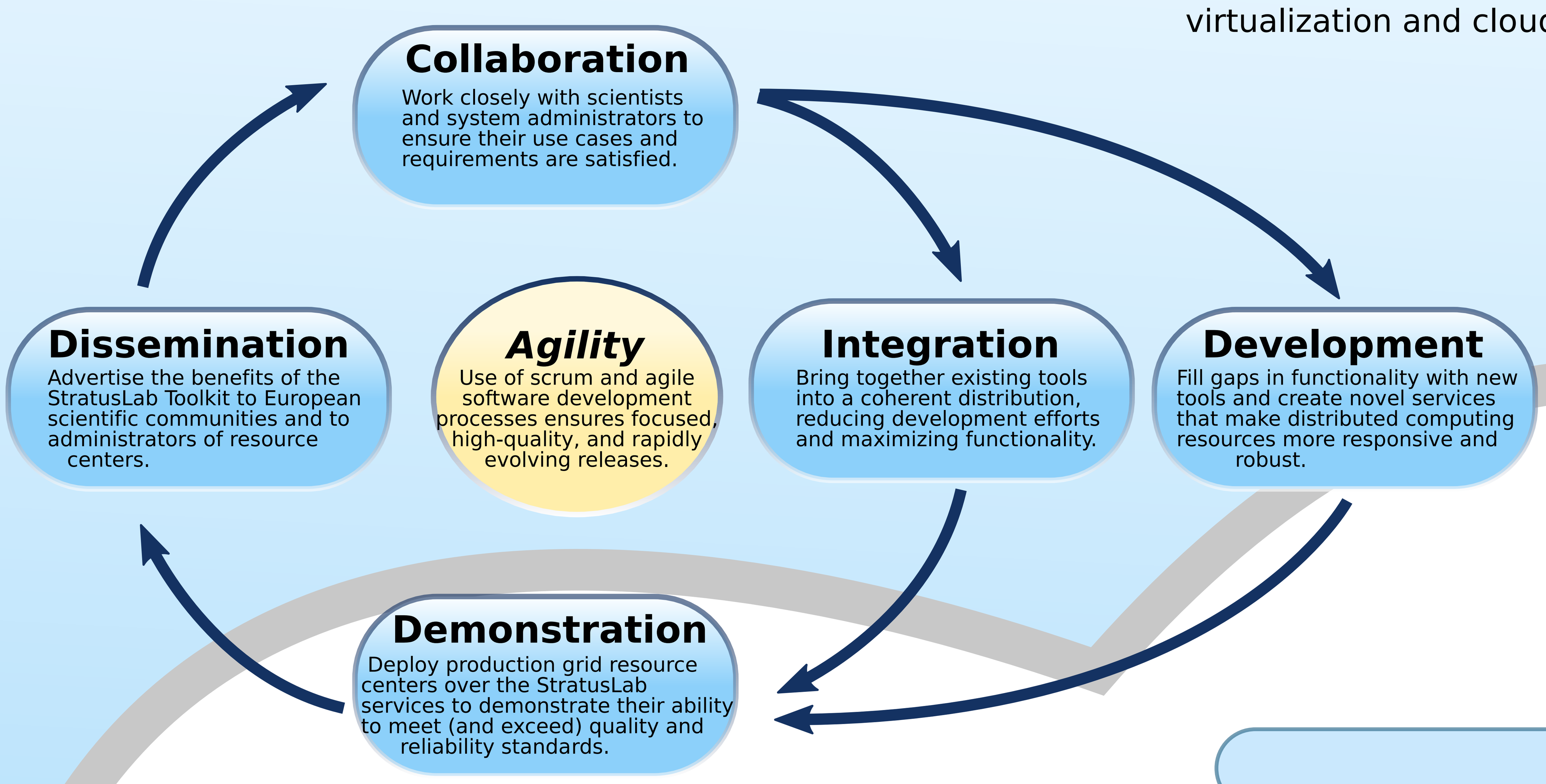
The StratusLab Toolkit will integrate cloud and virtualization technologies and services within grid sites and enrich existing computing infrastructures with “Infrastructure as a Service” (IaaS) provisioning paradigms.

StratusLab Distribution

StratusLab integrates, distributes and maintains a sustainable open-source cloud distribution to bring cloud to existing and new grid sites. The StratusLab distribution is composed of existing cutting-edge open source software and the innovative services and cloud management technologies developed in the project. It incrementally delivers a production grade distribution that is being demonstrated through the operation of production-level grid sites in the project.

Development Process

StratusLab has adopted the Scrum agile software development method, delivering releasable increments every three weeks, daily stand-up meetings, and collegial management of priorities. The result is the StratusLab Distribution, supporting public, private and hybrid clouds.

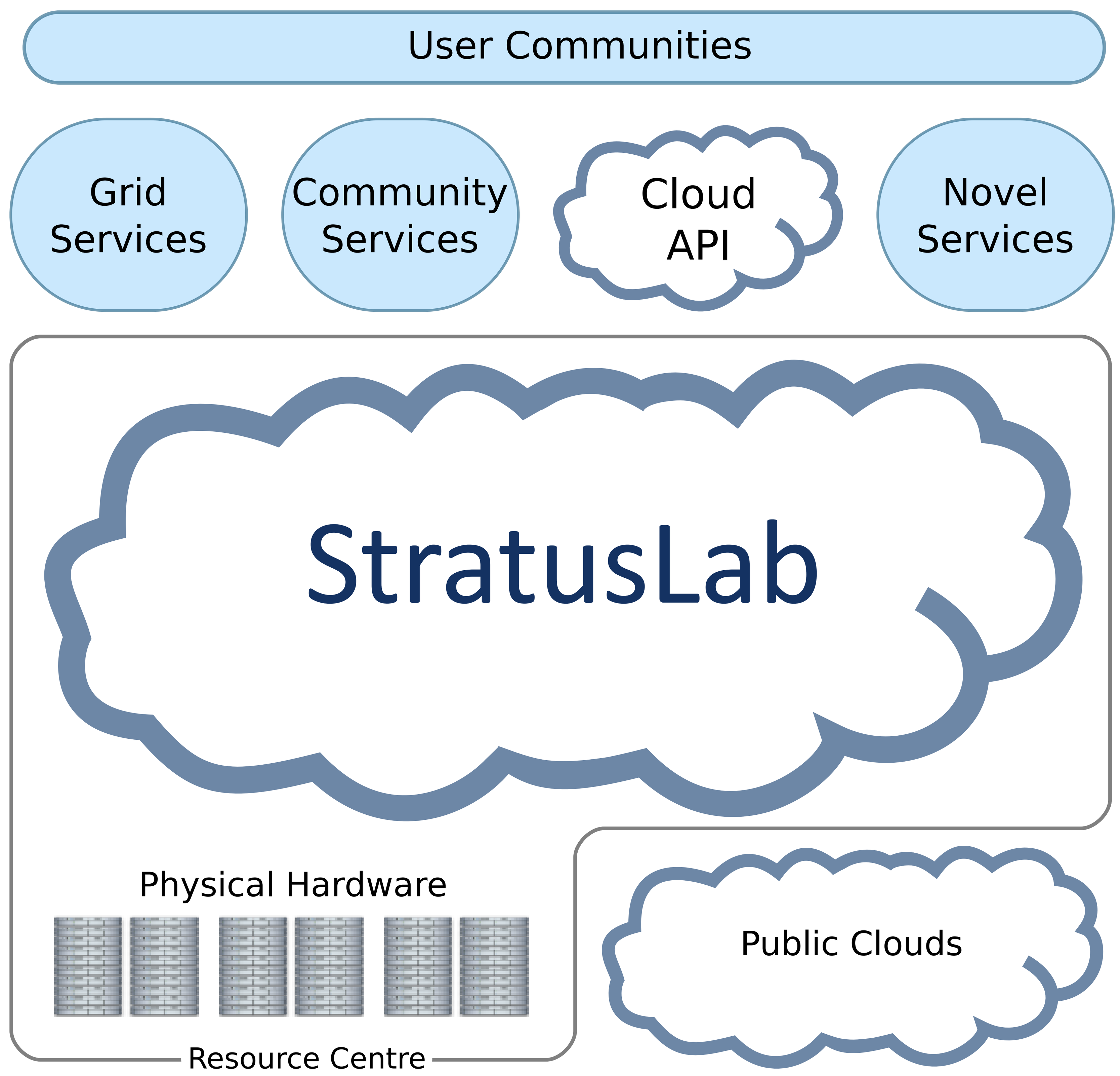


Benefits

StratusLab brings several benefits to the e-Infrastructure ecosystem, in terms of simplification, added flexibility, increased maintainability, quality, energy efficiency and resilience of computing sites. A key aim is for the cloud layer to be transparent to the layers above.

Impact

Existing grid middleware, as provided by EMI, continues to provide the glue to federate the distributed resources and the services for high-level job and data management. StratusLab improves the usability of distributed computing infrastructures; attracts scientific user communities; appeals equally to industrial users; keeps European research infrastructures at the technological forefront, and strengthens the know-how in virtualization and cloud computing of European industry.



Plan

StratusLab is a two-phase project. In the **first phase**, the project focuses on cloud computing for resource provisioning in grid sites. This entails development of the StratusLab cloud platform and creation of virtual appliances for the scientific application domains in the project.

In the **second phase** the emphasis will shift towards developing new cloud-like delivery paradigms in grid sites. This will build on the first phase, including new IaaS cloud interfaces that can be used for community services, directly by users, and to deploy novel cloud-based services.

Website: http://stratuslab.eu/
Contact: Charles Loomis
loomis@lal.in2p3.fr

Centre National de la Recherche Scientifique *France*
Universidad Complutense de Madrid *Spain*
Greek Research and Technology Network S. A. *Greece*
SixSq Sàrl *Switzerland*
Telefonica Investigacion y Desarrollo S.A. *Spain*
Trinity College Dublin *Ireland*

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

SlipStream™ Release with Confidence

On-Demand creation of runtime environments on clouds

SlipStream, by SixSq, deploys runtime environments in a cloud automatically, on-demand, and releases the resources after use. SlipStream users do not have to worry about configuring complex hardware and software to create personal runtime environments, representative of production conditions.

SlipStream first used Amazon EC2 for cloud resources, but is now integrated with the StratusLab Cloud*. This allows SlipStream to be deployed on private infrastructures.

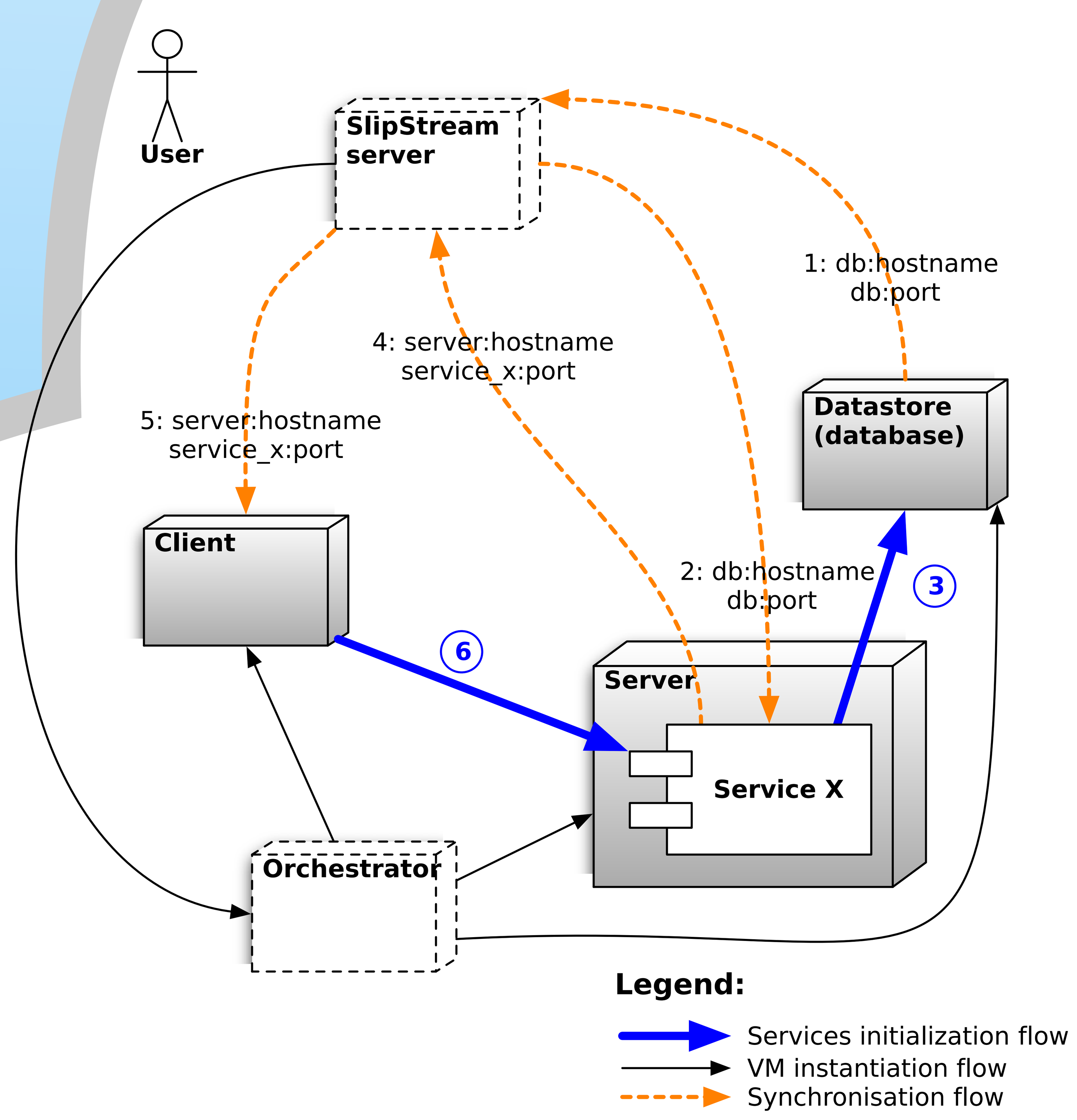
How does it work?

SlipStream is accessed via a secure Web 2.0 application. Users can organize their work in project hierarchies. Each SlipStream artifact is version controlled, providing complete history of changes. Two main workflows are available in SlipStream:

1. virtual machine (and disk) creation;
2. deployment and service synchronization.

The first flow allows a user to specify the content of each virtual machine. The second flow includes the synchronization between inter-dependent software components.

SlipStream™
Release with confidence



Deployment overview

Hybrid Cloud Solutions

SlipStream is ready for multi-cloud usage, which is an important use-case for the second year of StratusLab. This means that users will be able to leverage their virtual machine creation recipes, as well as their deployment models, across different clouds. This feature also gives site administrators the possibility to re-generate virtual machines for different cloud services, on-the-fly.

"SlipStream is to software what the robotic assembly line is to manufacturing"

Continuous integration and deployment on clouds case studies

SlipStream is currently been evaluated in the context of the critical software systems at the European Space Agency. This work includes streamlining their development and release processes, such that their time-to-market is dramatically reduced.

Using SlipStream users can automate the creation of virtual applicances and deployments, such that full sites, can be deployed within minutes. For example, deployment of European Grid Infrastructure (EGI) sites via StratusLab open-source private IaaS cloud distribution.

"Declare war on pre-release stress. Release often and under your own terms. Improve software quality. Reduce time-to-market. These are a few reasons why we hope you'll enjoy using SlipStream!"

Website: http://sixsq.com
Support: support@sixsq.com
Contact: contact@sixsq.com

* http://stratuslab.eu StratusLab is developing a complete, open-source IaaS cloud distribution, and is particularly focused on enhancing distributed computing infrastructures such as the European Grid Infrastructure (EGI).

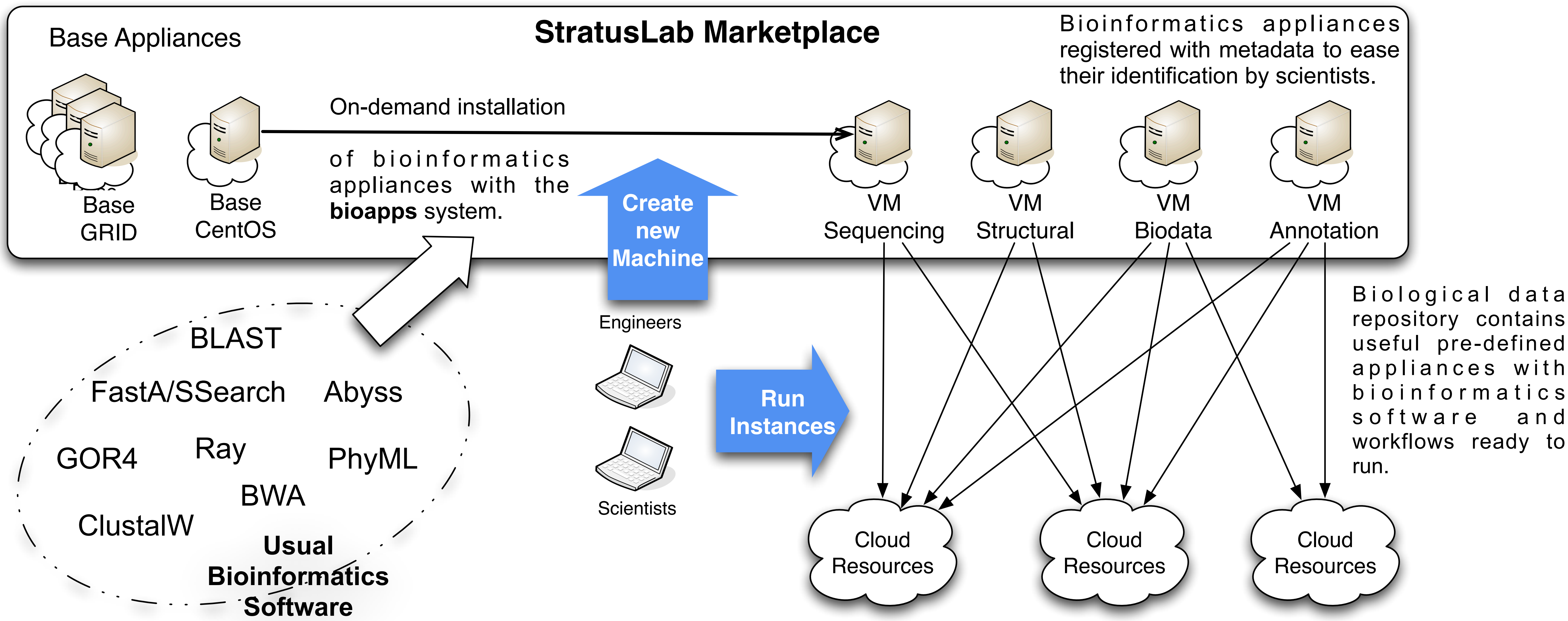
SixSq Sàrl
Switzerland

Virtualization of Bioinformatics Applications on Cloud Infrastructures

Christophe Blanchet[1] and Charles Loomis[2]

[1] IDB IBCP FR3302 CNRS, 7 passage du Vercors, 69007, Lyon
[2] LAL, UMR8607 CNRS, Bât. 200 BP 34, 91898, Orsay

Several experimental technologies have been improved to such a degree that obtaining data is easy, causing a deluge of data for the Bioinformatics community. The challenge is to be able to analyze these data with the relevant applications, for example, sequencing a whole genome obtained from Next Generation Sequencing (NGS) instrument. Many projects are working on the genome sequence of different organisms, continuously providing new sequences for analysis. Some bioinformatics algorithms like BLAST, FastA or ClustalW are used for that analysis and are usually classified as data-intensive, processing gigabytes of data stored in flat-file databases like UNIPROT, EMBL or PDBseq via a shared filesystem. Others like Abyss, BWA or Ray take the output sequences of sequencing machines and assemble them to get the complete sequence of the studied genome.
In the context of the StratusLab project (EU-FP7, www.stratuslab.org), we have built two bioinformatics virtual appliances: a "Biological databases repository" and a "Bioinformatics compute node".



Cloud Advantages for Bioinformatics

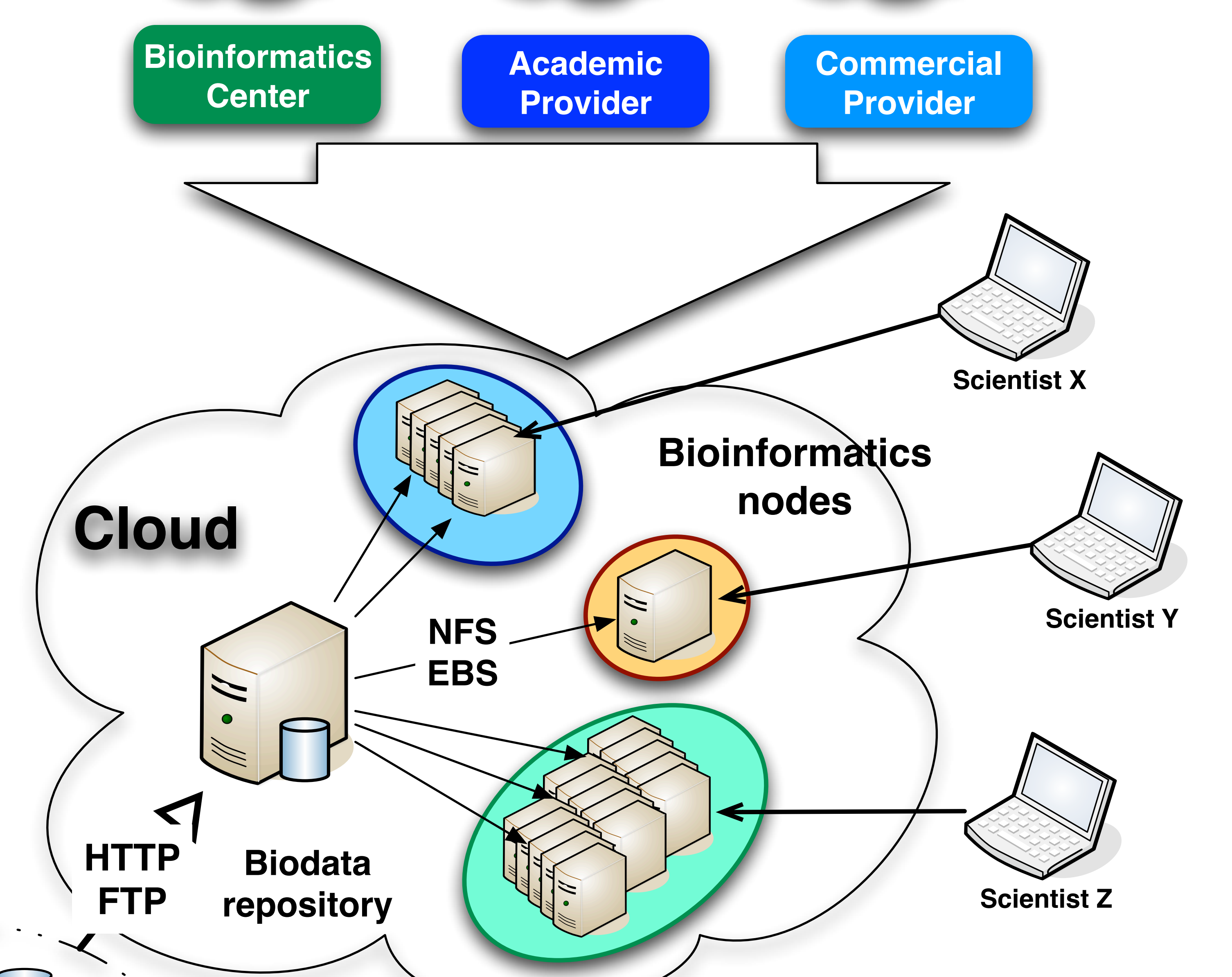
Cloud technologies provide scientists with the flexibility to deploy bioinformatics applications on different virtual machines. But clouds have to be connected with public bioinformatics infrastructures, especially the network of international biological databases, to be useful.

Platform as a Service - PaaS

- Provide scientists with Bioinformatics appliances on academic or commercial data centers, e.g. RENABI, or on their own computer or private cloud
- The mission of bioinformatics centers switches from providing services to providing virtual appliances

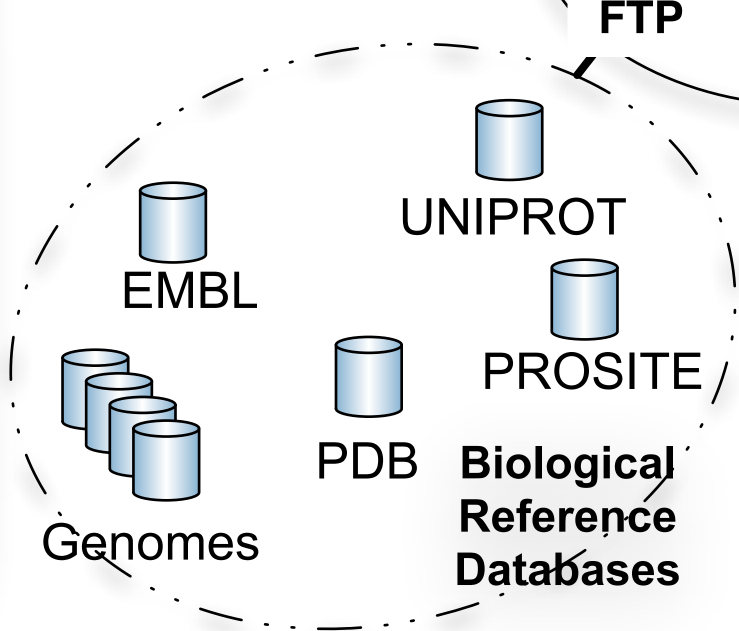
Infrastructure as a Service - IaaS

- Bioinformaticians can pre-define and deploy clusters or web service infrastructures
- Biologists can deploy different virtual servers to run a complete analysis pipeline



Access to Reference Databases

Access from any compute node is a strong requirement. Such databases contain, for example, protein or gene sequences and associated data, protein structures or complete genomes. We have built a virtual appliance that acts as a proxy between the internet where all the reference databases are published and the cloud where the bioinformatics analyses will be computed. To import and maintain these databases, we use the BioMaJ system (biomaj.genouest.org). This virtual machine stores the data on a local disk (in the future on a cloud-persistent disk), and then exports it with NFS to all the bioinformatics computing machines of the cloud.



Ready-to-run Bioinformatics Appliances

Appliances containing common bioinformatics software is also an important requirement. We are building such appliances, keeping them up-to-date, and making them available from the StratusLab Marketplace. Scientists with a cloud login can then launch as many instances of these appliances as required by their analysis pipelines. Because bioinformatics applications require access to reference data to process their analyses, the compute node mounts the exported volumes containing the biological databases.

Bioinformatics Cloud Usage

Usage must to be connected with public bioinformatics infrastructures like the French Bioinformatics Network RENABI (www.renabi.fr) and especially its grid infrastructure GRISBI (www.grisbio.fr). The adoption of clouds for bioinformatics applications will be strongly correlated to the capability of cloud infrastructures to provide ease-of-use and access to reference biological databases and common bioinformatics software. In that sense, StratusLab is collaborating with RENABI to help fulfill the requirements of the Bioinformatics community.

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