

# Italian Grid Infrastructure Portal (for CTA)

[portal.italiangrid.it](http://portal.italiangrid.it)

Marco Bencivenni

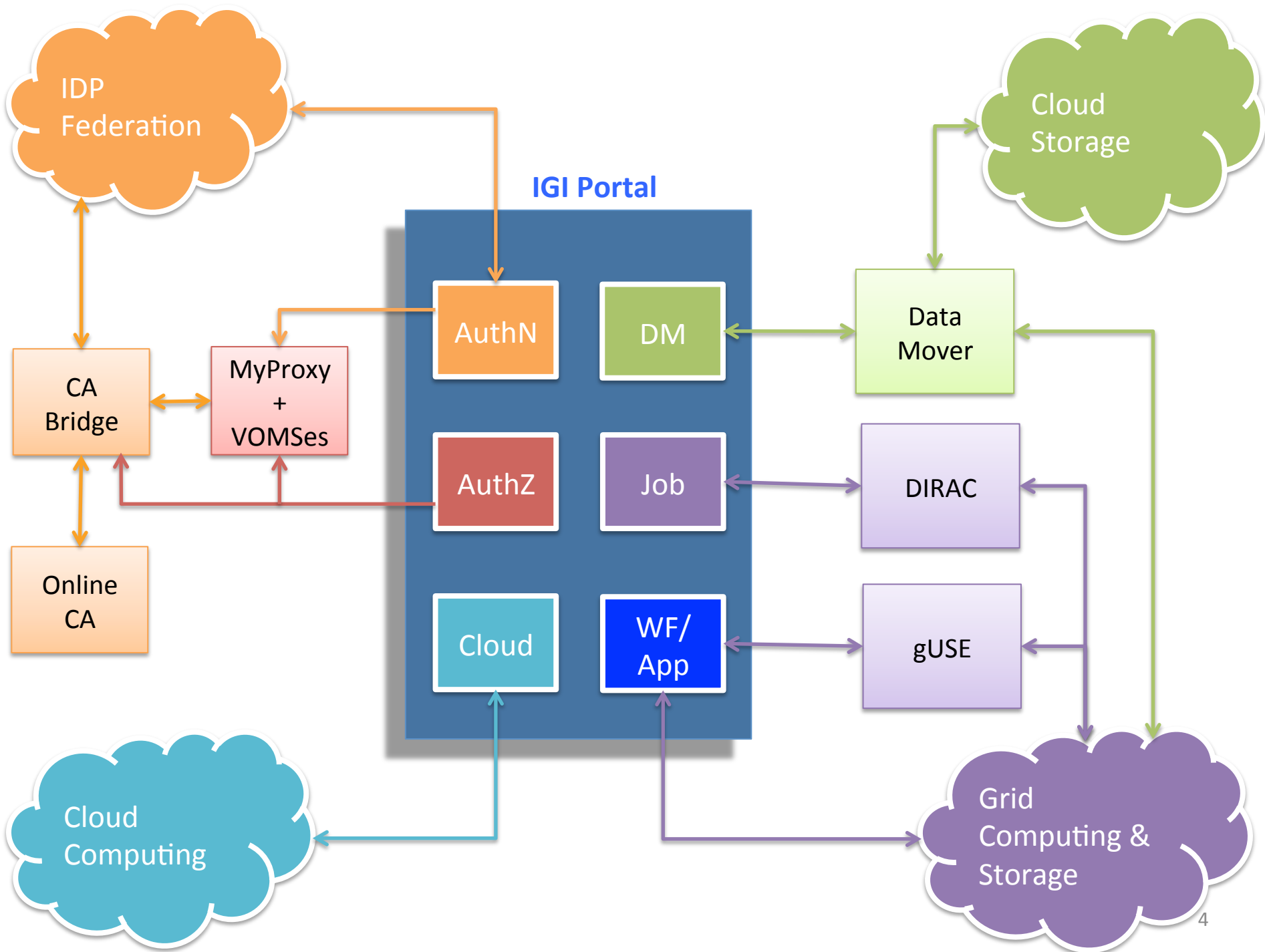
21/06/2013

# Introduction

- **Main goals**
  - Hiding the Grid Complexity
  - Facilitating the credential management
  - Being a unique access point for several services
- **Main Features**
  - Federated authentication
  - Online-CA integration (Certificate on demand)
  - Simple job, workflow, application submission (different middleware supported)
  - Virtual machine instantiation (Wnodes, OpenNebula, OpenStack)
  - Grid and Cloud data management
  - Modular structure

# Introduction

- Liferay based
  - Based on Portlet JSR286 standard for easier reuse and integration of components
- Mostly integration of existing services
  - Focus on reuse vs. new development for sustainability: DIRAC, gUSE/WS-Pgrade, WNoDeS, StoRM, EJBCA etc.
- General Purpose → Multi VOs/Communities portal
  - Interfaces ad hoc for specific communities or applications
- Some development needed to link components together
- User driven Functionalities development



# AuthN/AuthZ

- Authentication based on identity federation (IDEM – eduGAIN extendible with other federations or single IDP )
  - No new username and password to remember
  - Sensible information not in the portal
  - IDP standard attributes used also for other services
  - Each IDP provides its authentication method (username and password, certificate etc)
- To use the portal features an X.509 certificate and a VO membership are needed
- The portal provides an IDP catch-all for the users with certificate but not members of IDEM – eduGAIN
  - The data retrieved from the certificate are used to register the user in the IDP
- Interfacing to an online-CA to provide certificate on-demand in a transparent way for the members of IDEM – eduGAIN (**work in progress**)
  - The data retrieved from the IDP are used for the certificate generation

# Computing - Simple Job

- Dirac based --> Pilot job
- Simple JDL building (normal job, MPI, parametric)
- Upload of: input data, executable and parameters (thanks to the use of personal certificate)
- Output and log retrieve
- Job status monitoring

The screenshot displays the 'My Jobs' web interface. On the left, a yellow-bordered form titled 'JDL' contains the following fields:

- Job Name:** Portal\_Job\_Test
- Executable:** /Users/dmichelotto/Docum Sfoggia... or Specify your executable
- Arguments:** 200 | tee log.txt
- Output Sandbox (For multiple files separate with ';'):** log.txt
- Input SandBox:** /Users/dmichelotto/Docum Sfoggia... (with a red 'X' icon)

At the bottom of the form are 'Submit' and 'Back' buttons. On the right, a green-bordered sidebar titled 'Add or Remove fields' lists the following options:

- + Standard Output
- + Standard Error
- Output Sandbox
- + Input Sandbox

Below this list, under the 'MPI' section, are the following options:

- + CPU Number
- + Host Number
- + Whole Nodes
- + SMP Granularity

A vertical dashed line is present in the background of the interface.

# Computing - Workflow

- gUSE/WS-Pgrade based
- Graphical editor to build the workflow structure
- Multi middleware
- Simple JDL building



Delete old instances  
 Do not delete old instances



## Import Sample Job

gridit:

Simple ▾

Import Job

Workflow name:	ANSYS
Note:	2012-9-12
Workflow Graph:	ANSYS_2012-09-03-092854 <input type="button" value="Edit"/> <input type="button" value="Fit copied workflow to a new graph"/>
Workflow Template:	--



# Computing: Specific Applications

- gUSE/WS-Pgrade based – Application Specific Module API (ASM)
- Ad hoc interfaces
- Runtime output and log retrieve
- Licenses manager
- Data management integration
- Checkpoint

NEW JOB Refresh

Executables: life.bash View Exe

Arguments: -pnode -vww -- life-mpi -c 1000 -r 1000 -s 1000 -W 0

THEOPHYS: 05-04-2013 05:21:42

Input: life.tgz Submit

Requirements: emi-ce.pr.infn.it

Output: TestDiego

Host Number: 1

RUNNING

Show Details

STDERR

STDOUT

Abort

Executables: life.bash View Exe

Arguments: -pnode -vww -- life-mpi -c 1000 -r 1000 -s 1000 -W 0

THEOPHYS: 03-06-2013 10:49:48

Input: life.tgz Submit

Requirements: emi-ce.pr.infn.it

Output: TestDiego2

Host Number: 1

FINISHED

Show Details

STDERR

STDOUT

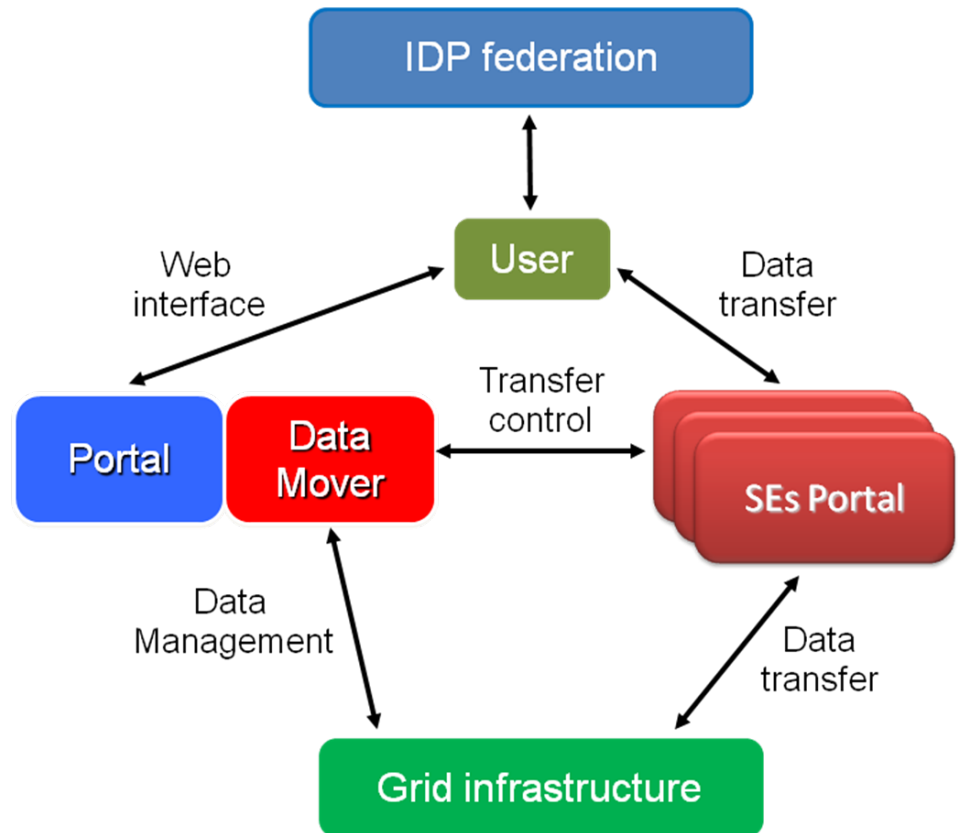
Delete



Institute	Application	Description	State
INFN-Legnaro	<ul style="list-style-type: none"> <li>• Ansys</li> <li>• Fluka</li> </ul>	<ul style="list-style-type: none"> <li>• Simulation Software</li> </ul>	<ul style="list-style-type: none"> <li>• <b>OK</b></li> <li>• <b>Developing</b></li> </ul>
INGV	<ul style="list-style-type: none"> <li>• NEMO</li> </ul>	<ul style="list-style-type: none"> <li>• Oceanographic modeling</li> </ul>	<ul style="list-style-type: none"> <li>• <b>OK</b></li> </ul>
INFN-Parma/Pisa	<ul style="list-style-type: none"> <li>• Executables</li> </ul>	<ul style="list-style-type: none"> <li>• Theoretical Physics</li> </ul>	<ul style="list-style-type: none"> <li>• <b>OK</b></li> </ul>
CNR ISAC	<ul style="list-style-type: none"> <li>• Globo</li> </ul>	<ul style="list-style-type: none"> <li>• Weather Forecasting and Climate study</li> </ul>	<ul style="list-style-type: none"> <li>• <b>To Do</b></li> </ul>
CNR ISOF	<ul style="list-style-type: none"> <li>• Namd</li> <li>• Gaussian</li> <li>• Amber</li> </ul>	<ul style="list-style-type: none"> <li>• Dynamic Molecular Simulation</li> <li>• Electronic Structure Modeling</li> <li>• Dynamic Molecular Simulation</li> </ul>	<ul style="list-style-type: none"> <li>• <b>OK</b></li> <li>• <b>Developing</b></li> <li>• <b>To Do</b></li> </ul>
Compchem VO	<ul style="list-style-type: none"> <li>• Namd</li> <li>• Gaussian</li> <li>• DI_poly</li> <li>• abc</li> </ul>	<ul style="list-style-type: none"> <li>• Dynamic Molecular Simulation</li> <li>• Electronic Structure Modeling</li> <li>• Dynamic Molecular Simulation</li> <li>• Analytical chemistry</li> </ul>	<ul style="list-style-type: none"> <li>• <b>OK</b></li> <li>• <b>Developing</b></li> <li>• <b>To Do</b></li> <li>• <b>To Do</b></li> </ul>
Bio Computing Group - Bologna University	<ul style="list-style-type: none"> <li>• Blast</li> </ul>	<ul style="list-style-type: none"> <li>• DNA sequences alignment</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Developing</b></li> </ul>
DRIHM.eu Project	<ul style="list-style-type: none"> <li>• WRF</li> </ul>	<ul style="list-style-type: none"> <li>• Weather Forecasting</li> </ul>	<ul style="list-style-type: none"> <li>• <b>To Do</b></li> </ul>
Virgo Experiment	<ul style="list-style-type: none"> <li>• Executables</li> </ul>	<ul style="list-style-type: none"> <li>• Detection of gravitational wave</li> </ul>	<ul style="list-style-type: none"> <li>• <b>Developing</b></li> </ul>

# Data Management

- To avoid the portal being a bottleneck for the data moving, files transfer is external to the portal through a dedicated storage space
- The **Portal SEs** acts as a cache memory until the file has been transferred to/from Grid (StoRM based)
- The **Data Mover** monitors and manages every transfer steps and communicates directly with the Grid Infrastructure for some operations: replicas, deletion, etc.



# Data Management

- LFC catalog browsing
- File replica
- Renaming
- File sharing
- File Deleting
- Folder making
- Details retrieving
- It acts at Logical and Physical level.

The screenshot displays the Grid Catalog web interface. At the top, there is a navigation bar with tabs: Welcome, Workflow, Storage (selected), My Data, Statistics, Help, CLOUD, and MPI. Below this is a toolbar with icons for Upload, Download, Rename, New Dir, Delete, Move, Share, Replicate, More Info, Set as Home, and Back. The interface shows the current location as Grid Catalog / VO: gridit / /grid/gridit/. On the left, a 'Folders' sidebar lists various directories including 1dmichelotto, ARCEM, Auger, LHCf, Lecce, MATLAB, MNEGRI, Parma, alex, and ansys. A 'Details' panel at the bottom left shows information for the selected file '20070130.1', including its name, size (42 b), and last modification date (2007/01/30 00:00). The main area contains a table of files and folders.

Filename	Shared with	Size	Modified
Page 1 / 8 >>			
1dmichelotto	All VO	-	2013/03/07 09:10
ARCEM	All VO	-	2011/12/23 00:00
Auger	All VO	-	2009/07/11 00:00
LHCf	All VO	-	2011/10/22 00:00
Lecce	All VO	-	2007/10/26 00:00
MATLAB	All VO	-	2012/01/04 00:00
MNEGRI	All VO	-	2012/05/04 00:00
1dmichelottoSilverlight.dmg	All VO	14.24 Mb	2013/09/27 12:26
20070130	All VO	42 b	2007/01/30 00:00
20070130.1	All VO	42 b	2007/01/30 00:00
20070130.2	All VO	42 b	2007/01/30 00:00
20070226	All VO	42 b	2007/02/26 00:00
20070227	All VO	42 b	2007/02/26 00:00
20070612	All VO	42 b	2007/06/12 00:00
2007061202	All VO	42 b	2007/06/12 00:00
2007061203	All VO	42 b	2007/06/12 00:00
2007061204	All VO	42 b	2007/06/12 00:00

# Data Management - Upload

- SE Upload and Catalog registration
- SE destination choosing
- Automatic retry in case of failure
- Final state information and error
- Resumable upload

The screenshot displays a web interface for uploading files. At the top, there is a green button labeled '+ Add files...', a red dropdown menu currently set to 'Random destination', and a blue button labeled 'Start upload'. To the right of these controls is a progress bar and status text: '83.81 Mbit/s | 00:00:00 | 64.10 % | 10.60 MB / 16.54 MB'. Below this is a table of files being uploaded:

File Name	Size	Progress	Status
jQuery_Tutorial.zip	1.27 GB	100%	Transfer OK
demo-nemo2.screenflow	0.20 KB	0%	Transfer failed view details
Single_Signon_Kerberos_LDAP.pdf	2.53 MB	100%	Transfer OK
Poster_Duck_R2B_completo1.pptx	6.41 MB	~25%	Start Cancel
jira-cli-2.4.0-distribution.zip	3.78 MB	~50%	Start Cancel
Manual and Developer Tutorial for Application Specific Module of gUSE.doc	6.35 MB	~10%	Start Cancel

# Data Management - Download

- Download from Grid to user's PC
- Download from Grid to remote server (FTP,SFTP webDAV)
- Mail notification when the transfer is completed.

Download

Where do you prefer to save the files selected?

On my PC

On a remote Server

Please indicate an FTP or SFTP server where to copy them.

You will be notified by mail as soon as the transfer ends.

Protocol: SFTP

Server url:  (Es: server.example.it)

Path:  (Es: /home/user/folder/)

Username:

Password:

- *Demonstration video:* <https://portal.italiangrid.it/web/guest/data-management>

# Cloud

- The portal uses a Cloud CLI to allow the users to interact with several platforms Cloud (OpenStack, OpenNebula and WNoDeS ) in a transparent way using, to instantiate virtual machines choosing the image from a central repository (EGI Market Place for FedCloud task force)
  - Users can set the virtual machine size: RAM, hard disk, core numbers
  - Uploading or auto-generation of SSH key pair
  - Login via web console or via SSH client from user's PC
  - OCCl standard
- The mechanism can be easily extended for other use cases using repositories and images ad hoc

 EGI-Demo-INFN

**OS:** Debian : 6.0.5 : x86\_64

**Endorser:** elisabetta.ronchieri@cnafe.infn.it

**Description:** Debian appliance for the EGI demo available at INFN

Size

- ✓ small
- large
- medium
- extralarge

How many Instances

1

Selected VO:

fedcloud.egi.eu

Create Instance

# Compliance status key

- [**READY**] – the Science Gateway is able to fulfill this requirement but some application developments could be needed for complete fulfillment
- [**FUTURE**] – the requirement is not fulfilled now, work on fulfilling it was not started
- [**EXTERNAL DEPENDENCY**] – fulfillment of the requirement depends completely on external (to the Gateway) factors – e.g. existence of a certain application or policy
- [**CLARIFICATION NEEDED**] – the requirement needs further clarification
- [**DIFFERENT IMPLEMENTATION**] – From an architectural point of view the functionality is present but implemented in a different way

# Capability requirements

- [READY] UR-SG-0010: Capabilities
- [READY] UR-SG-0020: Data access
- [EXTERNAL DEPENDENCY] UR-SG-0030: Applications access
- [EXTERNAL DEPENDENCY] UR-SG-0040: Data Management Monitoring Applications
- [READY] UR-SG-0045: Privileged Applications access
- [READY] UR-SG-0050: Documentation access
- [READY] UR-SG-0055: Science Gateway documentation – „About” page and FAQ available
- [EXTERNAL DEPENDENCY] UR-SG-0060: User support
- [EXTERNAL DEPENDENCY] UR-SG-0065: Help-desk
- [EXTERNAL DEPENDENCY] UR-SG-0070: Community feedback – functionality available in Liferay, presently no content available
- [READY] UR-SG-0080: Resource discovery service – the applications run only on the resources to which the user has access rights (decided on VO membership basis)
- [READY] UR-SG-0090: Job execution service – done for jobs executed on Grid infrastructure
- [READY] UR-SG-0095: Data for local computing
- [READY] UR-SG-0100: Input/output transfer between applications
- [READY] UR-SG-0110: Application workflow management – execution of predefined workflows is possible
- [READY] UR-SG-0150: Simultaneous users performance



# Constraint requirements

- **[READY]** UR-SG-0200: Internet Service Providers support – hosted in CNAF with very good Internet connection
- **[FUTURE]** UR-SG-0300: User devices support – mobile support is future work
- **[READY]** UR-SG-0400: Web browser support
- **[READY]** UR-SG-0410: User device Operating system dependency
- **[READY]** UR-SG-0500: Web portal
- **[READY]** UR-SG-0510: Science Gateway language
- **[READY]** UR-SG-0600: Maintainability over many years – the Gateway’s architecture is modular
- **[READY]** UR-SG-0610: Existing framework use – uses Liferay, DIRAC, WS-Pgrade etc., components reuse is possible
- **[CLARIFICATION NEEDED]** UR-SG-0620: Existing standards compliance – what kind of standards should be considered?
- **[CLARIFICATION NEEDED]** UR-SG-0630: EGI VO Portal policy?
- **[READY]** UR-SG-0700: Availability>98%
- **[EXTERNAL DEPENDENCY]** UR-SG-0800: Portability over 30-40 years – the Gateway is built from standard technologies which are expected to be maintained in foreseeable future, applications?
- **[CLARIFICATION NEEDED]** UR-SG-0910: Supported simultaneous user profiles – roles and privileges supported
- **[EXTERNAL DEPENDENCY]** UR-SG-0915: Restricted network access
- **[READY]** UR-SG-0916: Specific access rights per application – see UR-SG-0910
- **[EXTERNAL DEPENDENCY]** UR-SG-0920: CTA data collection access rights
- **[EXTERNAL DEPENDENCY]** UR-SG-0930: CTA access policy
- **[READY]** UR-SG-0940: Public access

# Single Sign On specific requirements

- **[DIFFERENT IMPLEMENTATION]** UR-SSO-0010: Local Account – Username and password are not local but external to the portal
- **[READY]** UR-SSO-0020: X509 certificate or institute login/password support
- **[DIFFERENT IMPLEMENTATION]** UR-SSO-0030: – CTA Federation OR IDP
- **[READY]** UR-SSO-0040: Grid certificate support - Personal certificate used
- **[READY]** UR-SSO-0050: User attributes – IDP attributes
- **[READY]** UR-SG-0060: User support
- **[EXTERNAL DEPENDENCY]** UR-SSO-0100: Performance - IDP dependent
- **[READY]** UR-SSO-0900: Password encrypted on network– Shibboleth protocol
- **[READY]** UR-SSO-0910: Identity check for each connection
- **[READY]** UR-SSO-0920: Connection time-out
- **[CLARIFICATION NEEDED]** UR-SSO-0930: Ability to keep user session opened even if browser is closed – Browser settings dependent
- **[CLARIFICATION NEEDED]** UR-SSO-0940: Session recovery
- **[READY]** UR-SSO-1000: Existing standards compliance

# Comments to the VT

- Some of the requirements need further clarification
- External authentication
  - CTA Identity federation
  - A new user should be saved in a CTA IDP (username and password)
- Summary table for Single Sign On requirements (pag. 14) → many errors

# Contacts

[portal.italiangrid.it/web/guest/contact-us](http://portal.italiangrid.it/web/guest/contact-us)

portal@lists.italiangrid.it