

Overview about the SCI-BUS Project and WS-PGRADE/gUSE

http://www.sci-bus.eu

Peter Kacsuk, Zoltán Farkas MTA SZTAKI

Start date: 2011-10-01

Duration: 36 months





How to build a science gateway?



1. Build from scratch

2. Adapt and customize an existing gateway framework

- Significantly reduces development time
- Requires limited manpower and development cost
- Produces a robust and usable service
- The open source community is driving force for further development and extensions

SCI-BUS provides the required core gateway and customization technology

Who are the members of an e-science community regarding Option 2?





Science Gateway (SG) Framework Developers (5-10)

- Develop generic SG framework
- SCI-BUS project



SG Instance Developers (50-100)

- Develop application domain specific SG instance
- SCI-BUS project





WF (Application) Developers (500-1.000)

- Develop WF applications
- Publish the completed WF applications for end-users
- SHIWA project



 Execute the published WF applications with custom input parameters by creating application instances using the published WF applications as templates

SCI-BUS EU FP7 project

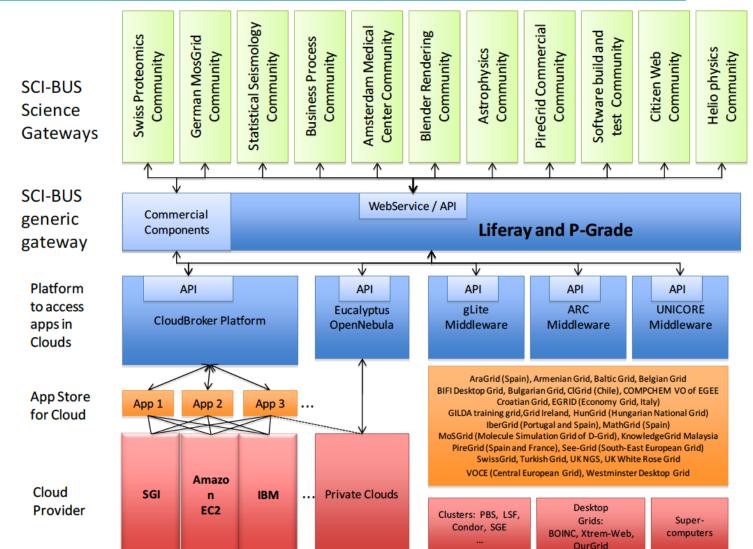


- SCI-BUS (SCIence gateway Based User Support) provides gateway framework and science gateway building technology
- 3-year project: 1 Oct 2011 30 Sep 2014
- Objectives of SCI-BUS
 - Support both WF developers and end-user scientists
 - Create a generic-purpose science gateway framework
 - Elaborate a science gateway instance development technology
 - Establish production SG instance services both for national grids (horizontal user communities) and various science communities (vertical user communities)
 - Develop business models to guarantee sustainability and commercial exploitation

4

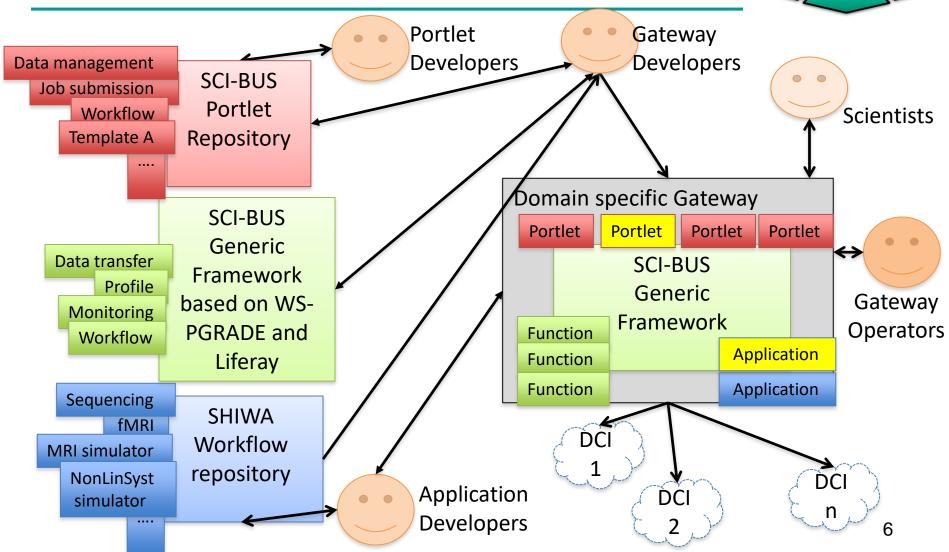
SCI-BUS Architecture





Community tools, activities and user roles in building gateways





WS-PGRADE/gUSE Generic-purpose gateway framework

- Based on Liferay
- General purpose
- Workflow-oriented gateway framework
- Supports the development and execution of workflow-based applications
- Enables the multi-cloud and multi-DCI execution of any WF
- Supports the fast development of SG instances by a customization technology
- Provides access to
 - internal repository
 - external SHIWA Repository

Scalable architecture based on collaborating services



Graphical User Interface: WS-PGRADE

Presentation tier



Workflow Storage File Storage Workflow Interpreter

Application Repository

gUSE Information System Middle tier

gUSE services



NEW

Job Submission Service: DCI-BRIDGE

gLite resource, ARC resources, Unicore resources, Globus resources (v2 and v4), LSF and PBS clusters, BOINC, GBAC, web services, Gemlca services, Local resource, Google App Engine Architectural tier

Available middlewares



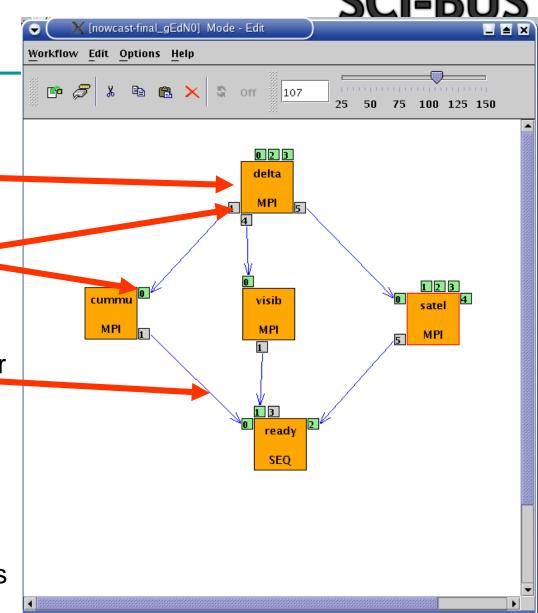
What is a WS-PGRADE workflow

A directed acyclic graph where

- Nodes represent jobs,
 services, embedded
 workflows
- Ports represent input/output files the jobs/services expect/produce
- Arcs represent file transfer operations and job dependencies

semantics of the workflow:

 A node can fire, i.e. its job can be executed if all of its input files are available

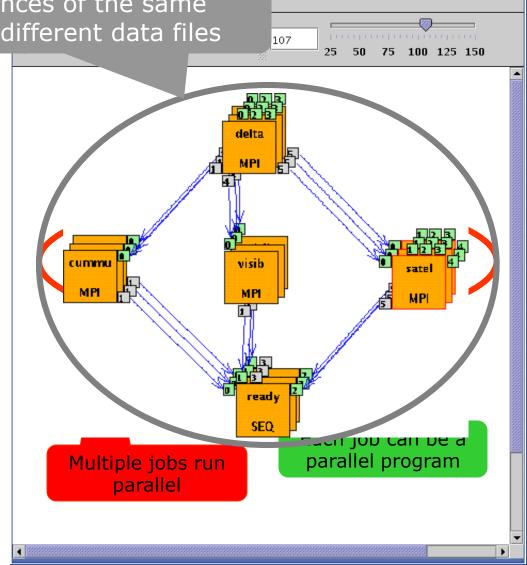


Introducing three levels of parallelism

Multiple instances of the same workflow with different data files

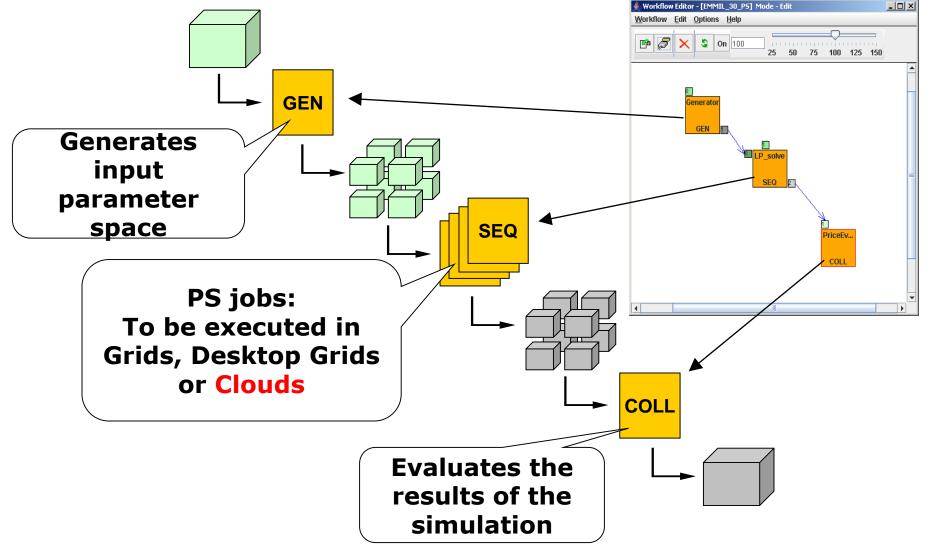
X [nowcast-final_gEdN0] Mode - Edit

- Parallel execution inside a workflow node
- Parallel execution among workflow nodes
- Parameter study execution of the workflow



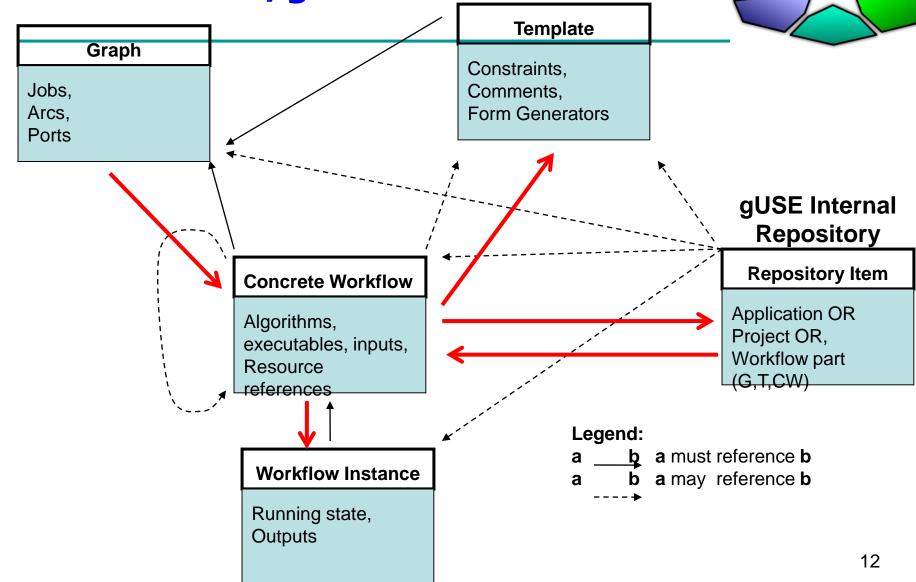
Executing Parameter Sweep (PS) application in various DCIs without changing the application workflow





Workflow related concepts

in WS-PGRADE/gUSE



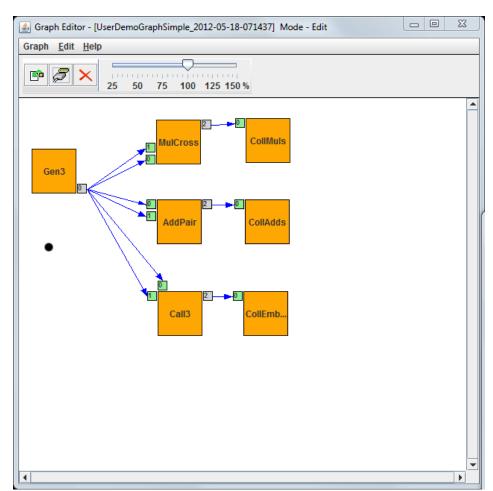
Support for workflow developers



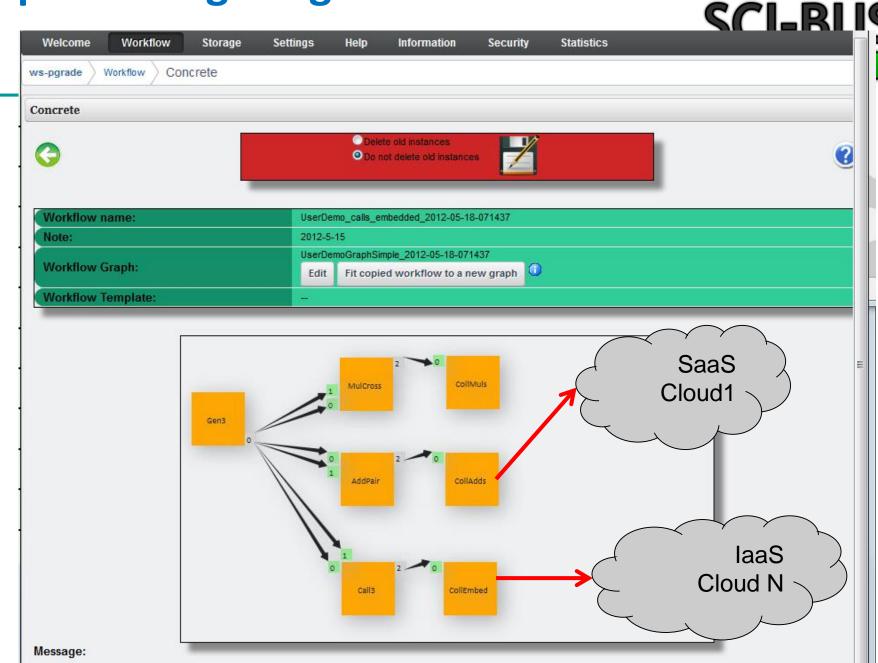
Creating and running WS-PGRADE workflows



Step 1: Edit workflow

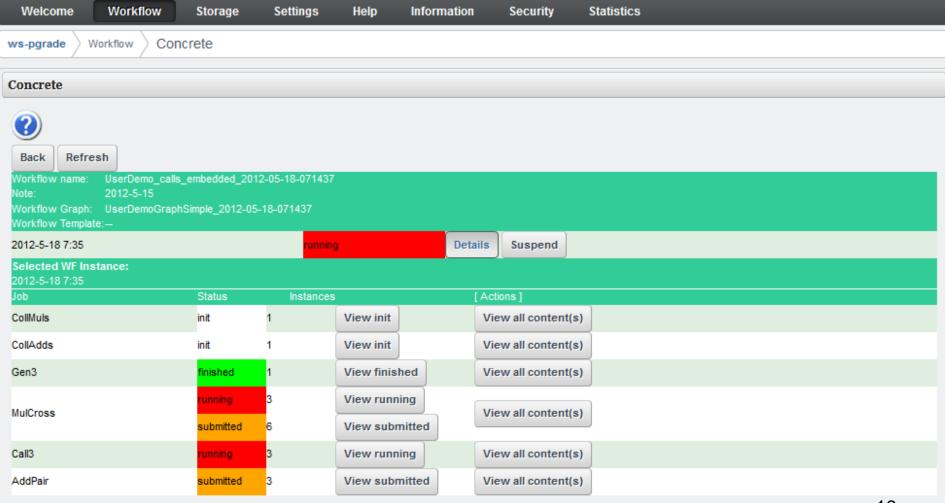


Step 2: Configuring the workflow



Step 3: Running workflow instance





Flexibility of using various DCIs by WS-PGRADE/gUSE



- Security: Flexible management of
 - Individual users' certificate
 - Robot certificates

Flexibility of using various DCIs by WS-PGRADE/gUSE



- Seamless access to various types of DCIs:
 - Clusters (PBS, LSF, MOAB, SGE)
 - Cluster grids (ARC, gLite, GT2, GT4, GT5, UNICORE)
 - Supercomputers (e.g. via UNICORE)
 - Desktop grids (BOINC)
 - Clouds (Via CloudBroker Platform)

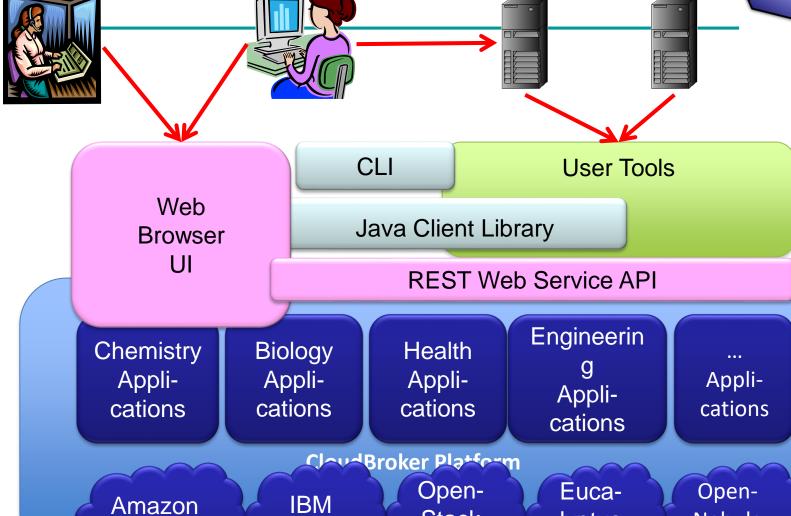
CloudBroker Platform



- Web-based application store for the deployment and execution of scientific and technical software in the cloud
- On demand, pay per use, browser / programmatic / command-line access, crossdomain
- Uses infrastructure as a service (laaS) from resource providers and offers these laaS resources for users
- Offers software as a service (SaaS) for end users

CloudBroker Platform Architecture





Cloud

Cloud

Stack

Cloud

lyptus

Cloud

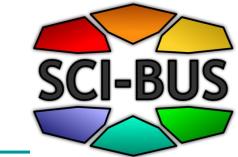
Nebula

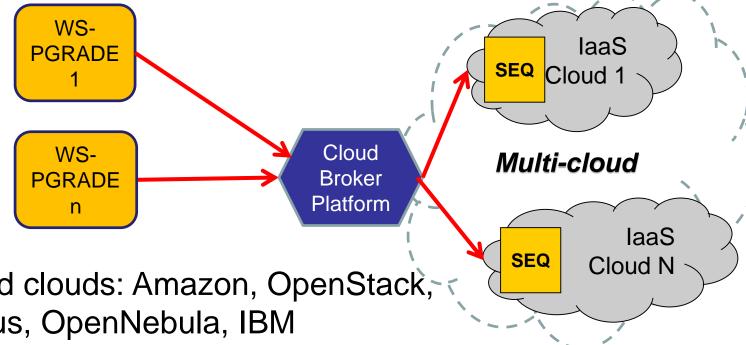
Cloud

20

CloudBroker Integration

Integrated WS-PGRADE/CloudBroker Platform to access multi-clouds





- Supported clouds: Amazon, OpenStack, Eucalyptus, OpenNebula, IBM
- SaaS solution:
 - Preregistered services/jobs can run from WS-PGRADE Supported from gUSE 3.5.0
- laaS solution:
 - any services/jobs (e.g. BoT jobs) can run from WS-PGRADE Supported from gUSE 3.5.1

WS-PGRADE UI to configure nodes for cloud services





Integration features



- Support for commercial clouds with costs (prices configured in CloudBroker Platform):
 - Estimated job cost before submission
 - Actual job and workflow cost after execution



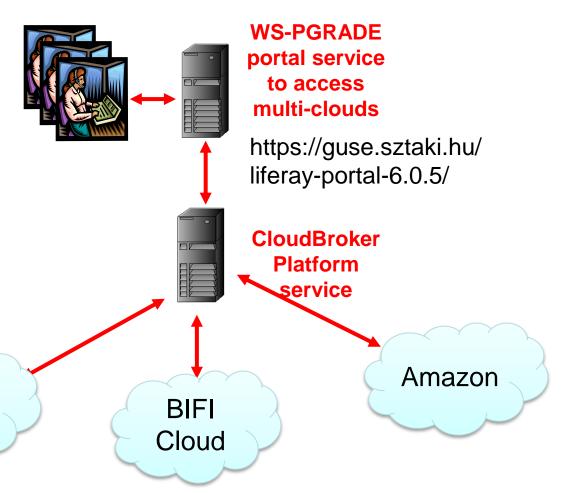
SCI-BUS multi-cloud access service

SZTAKI

Cloud



- Anyone can use the system (after registration) both for WS-PGRADE and CBP:
 - free for the academic clouds
 - To be paid for the commercial clouds



Demo (15 min)



- Workflow creation
- Workflow configuration for cloud access
- CloudBroker authentication management
- Workflow execution in cloud
- Workflow execution monitoring
- Robot certificate usage
- Upload workflow to gUSE Internal Repository
- Reuse robot-enabled workflow from gUSE Internal Repository

The flexibility of using WS-PGRADE/gUSE



- Flexible data storage access via the Data Avenue Blacktop service
 - to access data storages in different DCIs
 - To transfer files among the storages of different DCIs
 - To upload/download files to/from the storages of different DCIs
- Currently supported protocols:
 - http, https, ftp, gsiftp, srm
- Soon coming protocols:
 - iRODS, LFC, cloud storage protocols

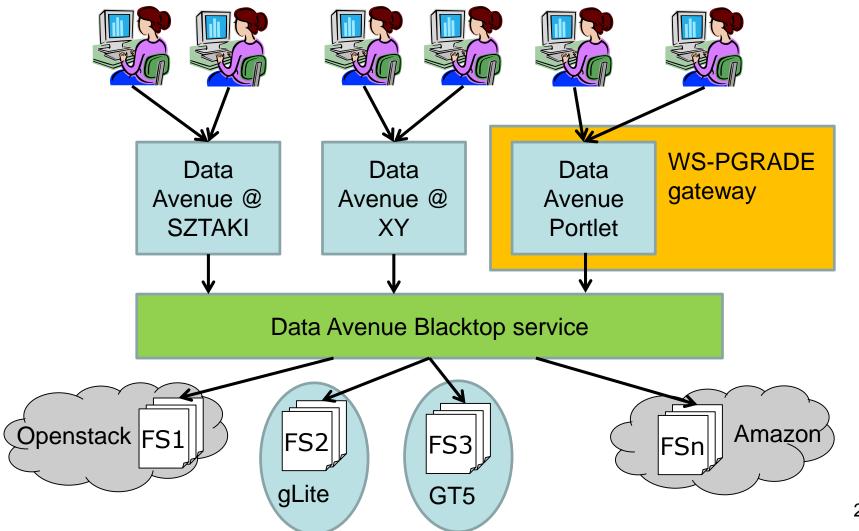
Access to Data Avenue Blacktop service



- Data Avenue Blacktop is a service that provides the basic types of data transfer operations (move, copy, delete, etc.) for different types of grid and cloud storages like: SFTP, GridFTP, HTTP, HTTPS, SRM, LFC
- See details: http://data-avenue.eu/home
- The Data Avenue Liferay portlet has been developed and will be added to WS-PGRADE in order to access the data transfer services of Data Avenue Blacktop
- As a result the WS-PGRADE users will be able to transfer data in a very flexible way between grid and cloud storages

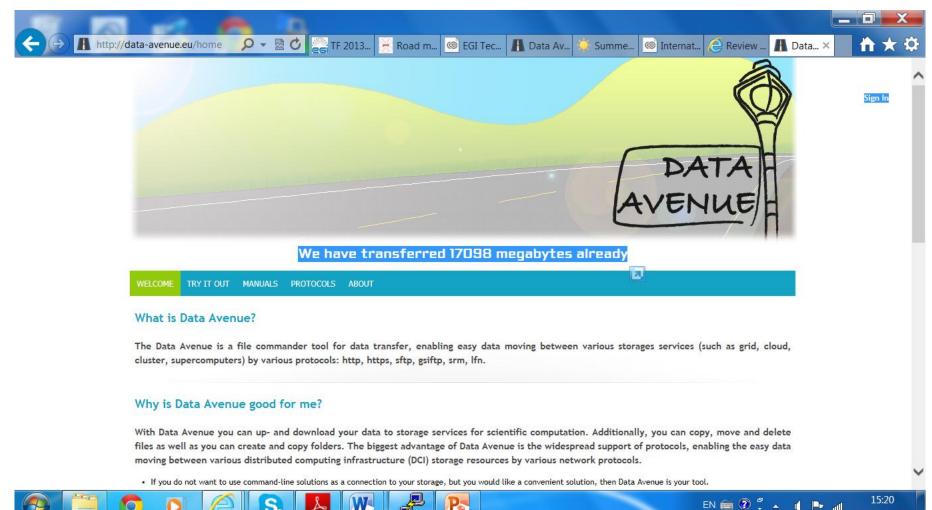
Data Avenue services





Data Avenue @ SZTAKI

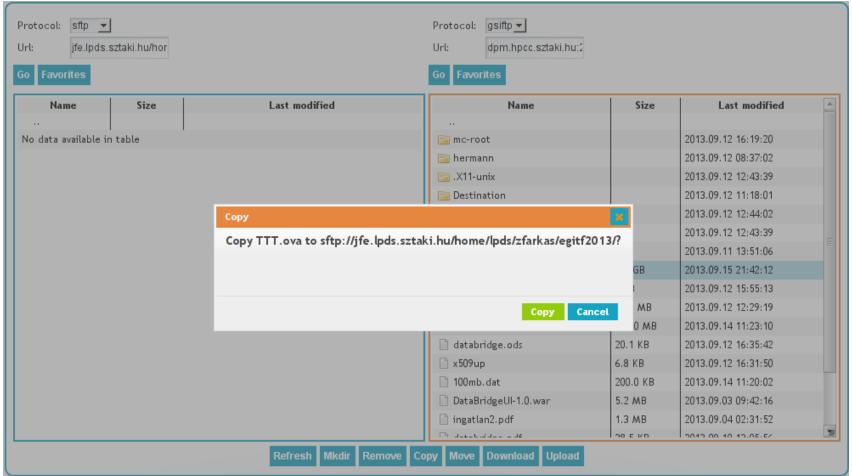




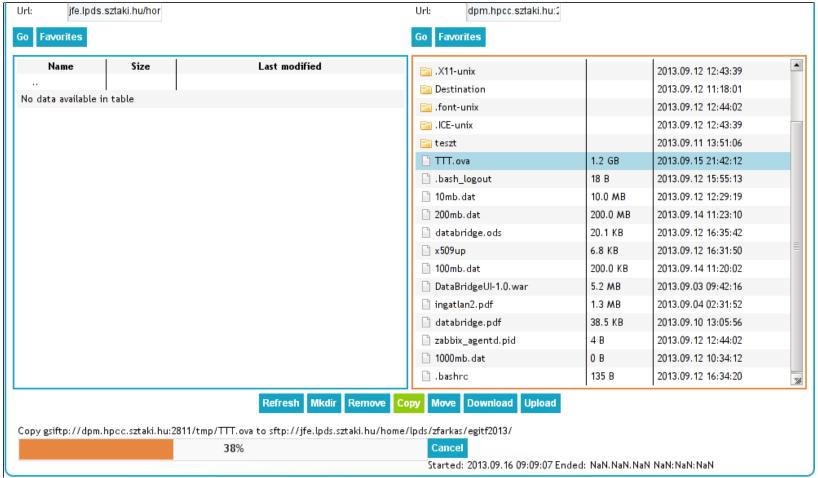


Protocol: sftp 🔻		Protocol: gsiftp ▼				
Urt: jfe.lpds.sztaki.hu/hor	Authentication usin	ng				
Go Favorites	 MyProxy serve 	er account				
i bin	Server:		: 7512			
ib ib ib ii i	Username:					
inks ar71xx	Password:					
Cloud	Delegation Life	e Time: 12	н			
.bash_profile		'				
.config.backfire	Existing VOMS	extended proxy certific	ate			
.lesshst	► Existing proxy	▶ Existing proxy certificate without VOMS extension (GlobusRFC820)				
ADSL	▶ Existing proxy	▶ Existing proxy certificate without VOMS extension (GlobusLegacy)				
input-ligands.zip	Existing proxy certificate without VOMS extension (Globus)					
id_rsa_jrm	_					
.Xauthority iliferay_before.sql	Add to favorites		Cancel Ok			
TTT. ova	1.2 GB 2013.0	3.20 16:54:46				
as_key.pm		7.05 13:08:25				
liferay_after.sql	521.3 KB 2013.0	9.04 14:27:58				
install-crossover-11.3.1.bin	41.9 MB 2012.1	0.31 08:34:36				
🗋 .viminfo	6.0 KB 2012.1	1.08 15:25:37				







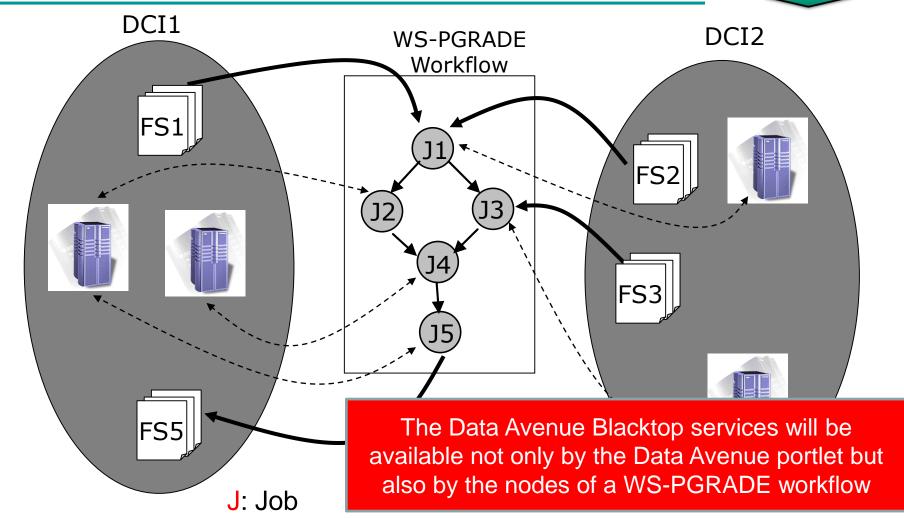




Protocol: Url: Go Favori	jfe.lpds.sztaki.h	nu/hor		Protocol: Url: Go Favor	gsiftp _ dpm.hpcc.sztaki.hu:2					
N	Name	Size	Last modified	📴 .X11-u	nix		2013.09.12 12:43:39			
 D		4.2.CD	2042 00 47 00 40 57	☐ Destination☐ .font-unix			2013.09.12 11:18:01			
TTT. ov	/a	1.2 GB	2013.09.16 09:10:57				2013.09.12 12:44:02			
				📴 . ICE-ur	nix		2013.09.12 12:43:39			
				📴 teszt			2013.09.11 13:51:06			
				☐ TTT. ov	/a	1.2 GB	2013.09.15 21:42:12			
				🖺 .bash_	logout	18 B	2013.09.12 15:55:13			
				🗎 10mb.	dat	10.0 MB	2013.09.12 12:29:19			
				🗎 200mb	.dat	200.0 MB	2013.09.14 11:23:10			
				🗎 databı	idge.ods	20.1 KB	2013.09.12 16:35:42			
				🗎 x509up		6.8 KB	2013.09.12 16:31:50			
				🗎 100mb	.dat	200.0 KB	2013.09.14 11:20:02			
				🗎 DataBı	idgeUl-1.0.war	5.2 MB	2013.09.03 09:42:16			
				🗎 ingatla	n2.pdf	1.3 MB	2013.09.04 02:31:52			
				🗎 databı	idge.pdf	38.5 KB	2013.09.10 13:05:56			
				🗎 zabbix	_agentd.pid	4 B	2013.09.12 12:44:02			
				🗎 1000m	b. dat	0 B	2013.09.12 10:34:12			
				🗎 .bashr	С	135 B	2013.09.12 16:34:20			
Refresh Mkdir Remove Copy Move Download Upload										
Copy gsiftp://dpm.hpcc.sztaki.hu:2811/tmp/TTT.ova to sftp://jfe.lpds.sztaki.hu/home/lpds/zfarkas/egitf2013/										
			Completed	Delete						
Started: 2013.09.16 09:09:07 Ended: 2013.09.16 09:10:57										

Generic data transfer among WS-PGRADE workflow nodes





FS: File storage system, e.g. gsiftp, iRODS, SRM

Demo of Data Avenue services (5 min)



The flexibility of using WS-PGRADE/gUSE

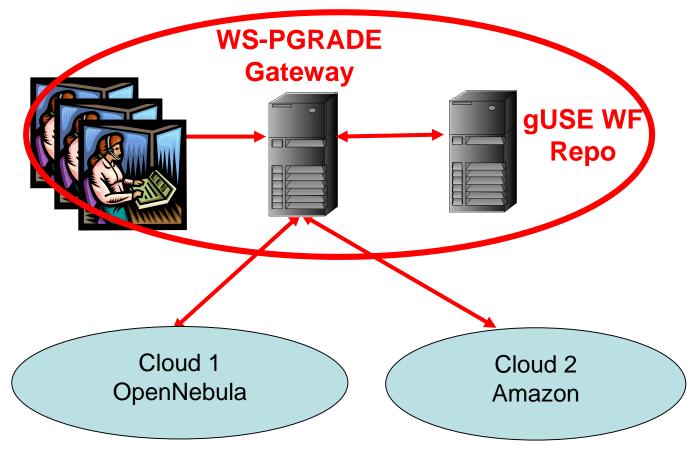


- Support for collaboration inside and among WS-PGRADE user communities
 - Internal Workflow Repository supports collaboration inside a WS-PGRADE user community
 - External SHIWA Workflow Repository supports collaboration among WS-PGRADE user communities



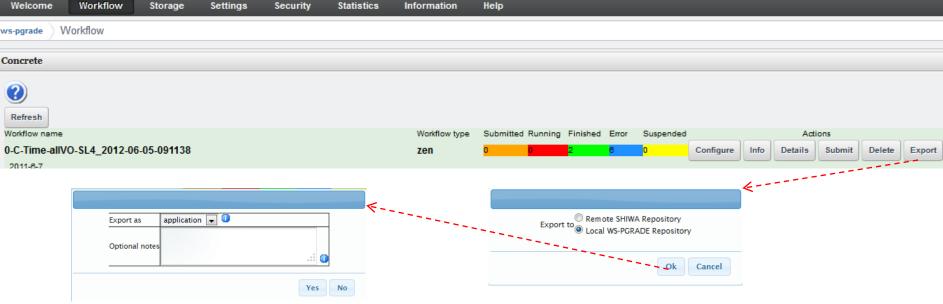
Collaboration among community members based on WS-PGRADE gateway





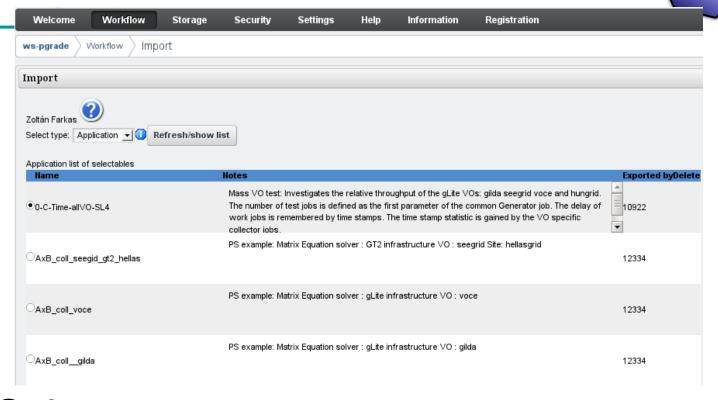
Share your workflow! - Export





- Define destination of export (local or global)
- Define type in "Export as" (application is OK)
- Add some description of your WF for others

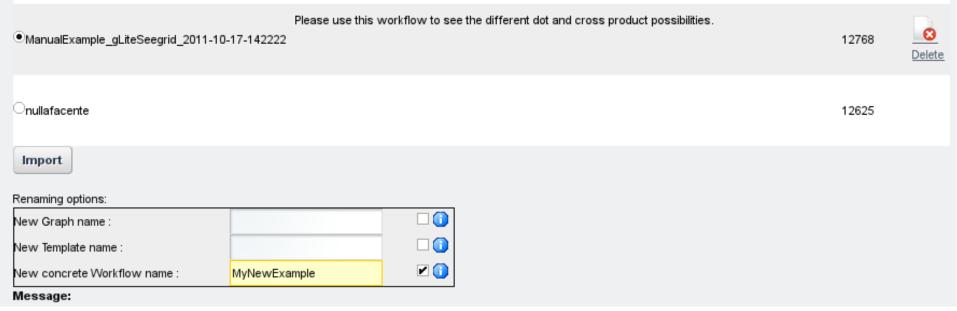
Use others' workflows! - Import SCI-B



- Select type
- Click refresh
- Select application from the list

Use others' workflows! - Import SCI-B

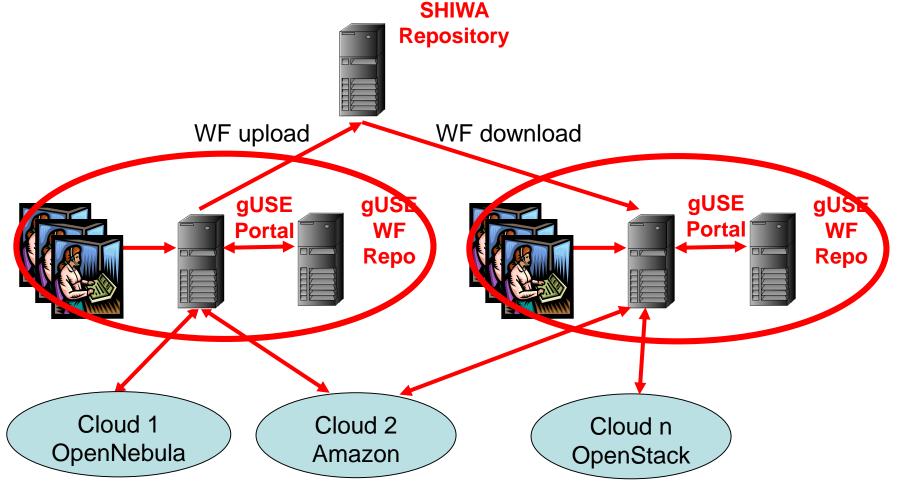




- Override properties you like
- Click "Import"

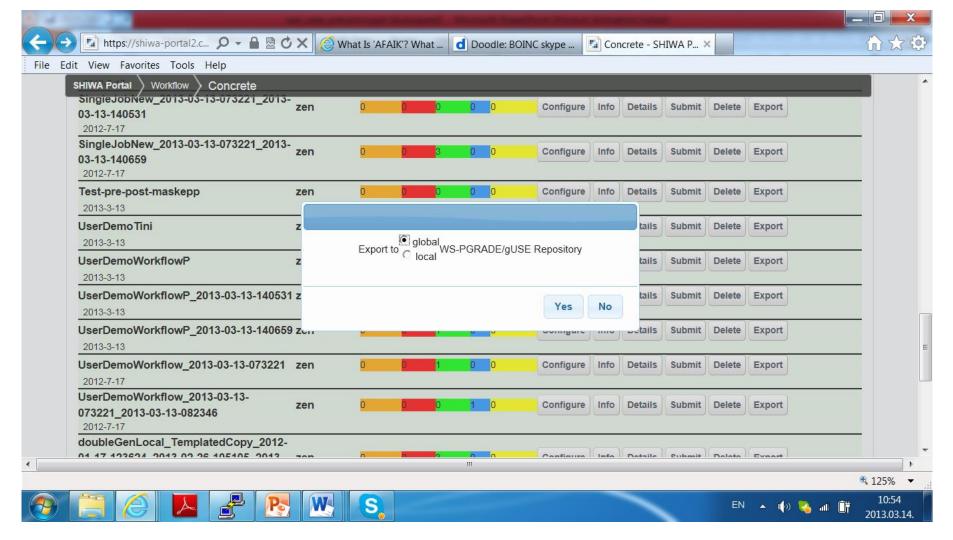
Collaboration among different community members based on gUSE gateway and SHIWA Repository





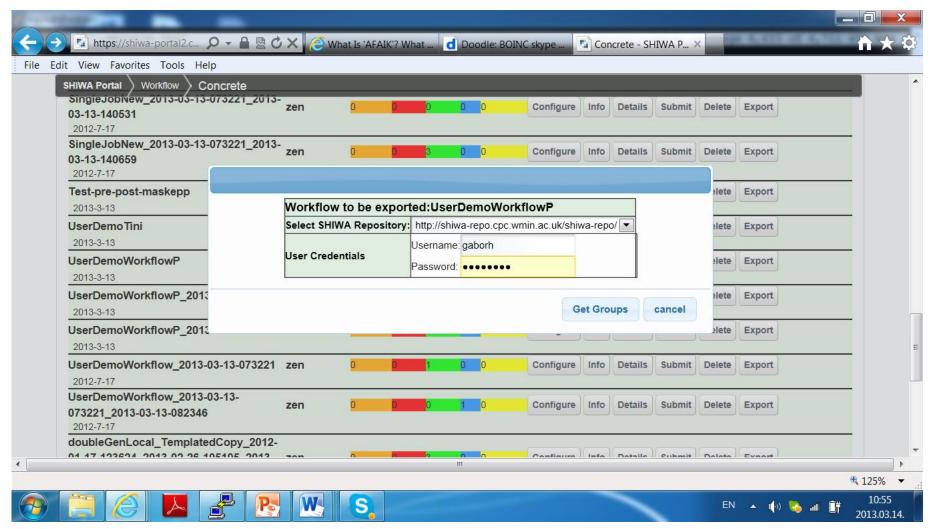
Upload WS-PGRADE WF to SHIWA Repository





Upload WS-PGRADE WF to SHIWA Repository





Demo (5 min)



- Upload workflow to gUSE Internal Repository
- Download workflow from gUSE Internal Repository
- Usage of gUSE Internal Repository

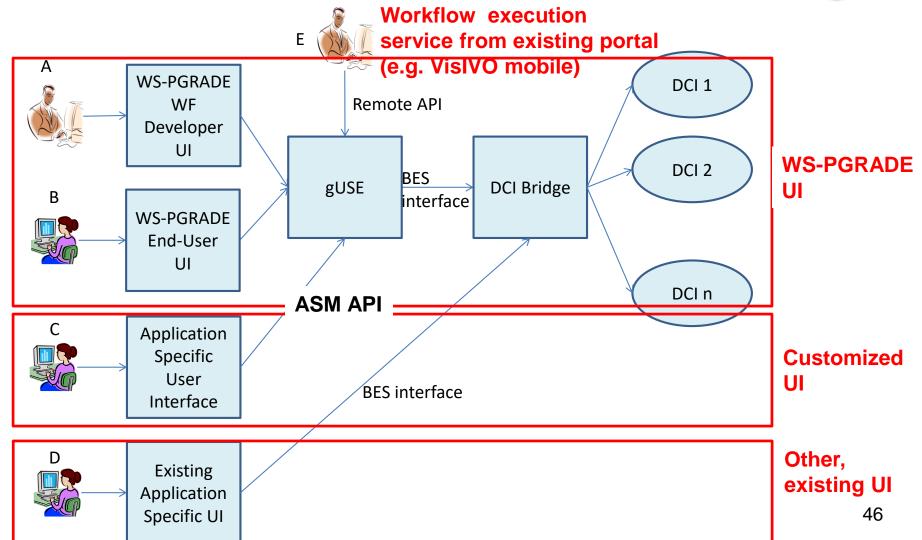
Flexible usage scenarios/business models by WS-PGRADE/gUSE



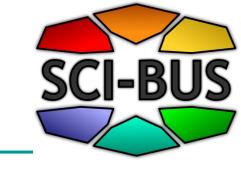
- Workflow developer view (full gateway framework view)
- End-user view (limited portlets, hidden WFs)
- Customized user interface to support the creation of domain specific gateways (ASM API)
- Provide workflow execution service on top of many different DCIs (Remote API)

Typical usage scenarios of WS-PGRADE/gUSE





End-user view based gateways



What is required from the end-user?

- Import workflow from repository
- Customise, execute and monitor application using simple web forms



What needs to be done by the gateway/application provider (system administrator + workflow developer)?

- Deploy gateway out of box
- Develop and configure workflows
- Create templates and applications
- Export application to repository



Demo of End-User View via Auto-Dock Gateway (5 min)



Types of gateways to be developed from the core WS-PGRADE/gUSE framework

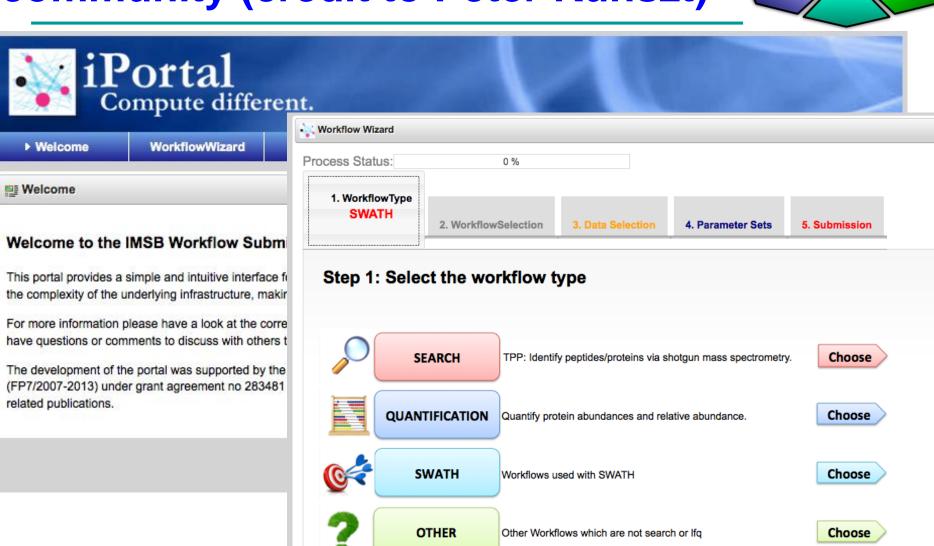


- Generic purpose gateways for clouds (workflow view)
 - Core WS-PGRADE/gUSE
- 2. Generic purpose gateway for specific technologies (workflow view)
 - SHIWA gateway for workflow sharing and interoperation
- 3. Application-oriented science gateway instance
 - Autodock gateway (end-user view)
 - Swiss proteomics portal (customized GUI using ASM API)
 - VisIVO Mobile (use of Remote API)

https://wiki-bsse.ethz.ch/display/iP

iPortal for proteomics user community (credit to Peter Kunszt)

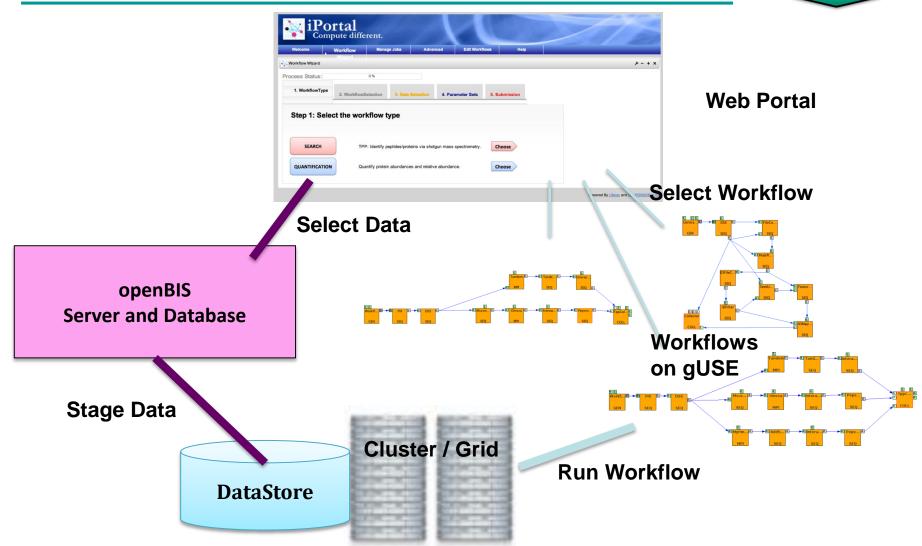




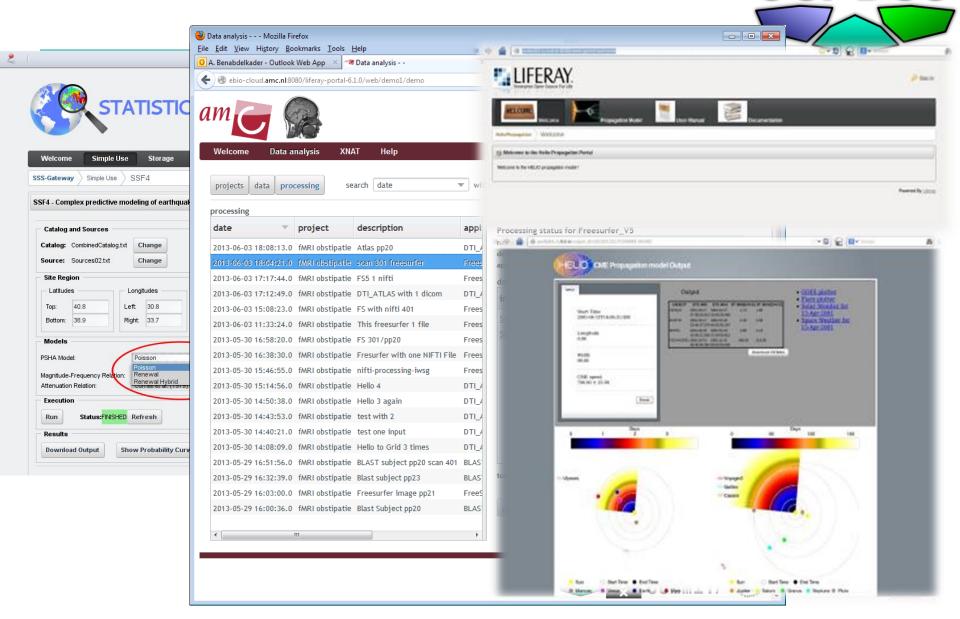
https://wiki-bsse.ethz.ch/display/iP

iPortal for proteomics user community (credit to Peter Kunszt)



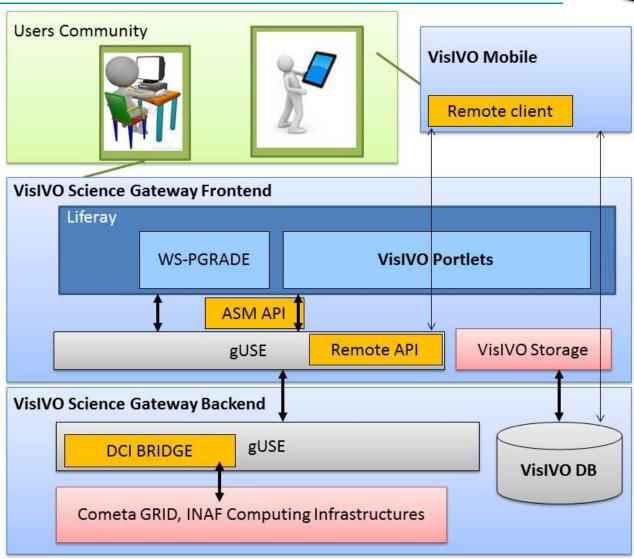


Further examples of SCI-BUS domainspecific gateways



VisIVO Gateway and VisIVO Mobile using Remote API





VisIVO Mobile



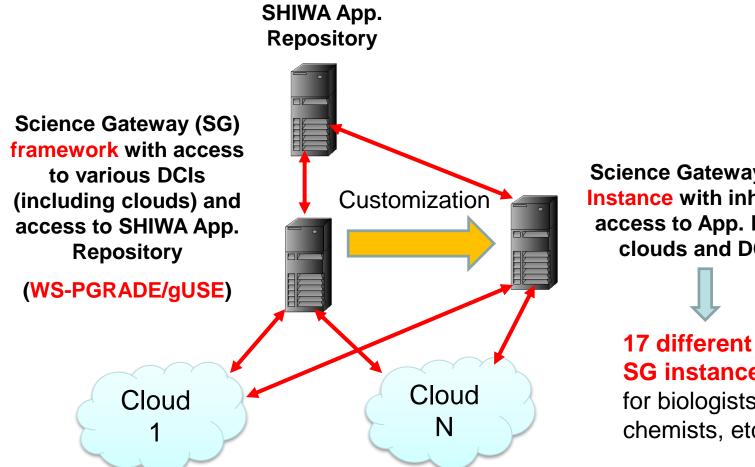






Support for science gateway instance developers





Science Gateway (SG) Instance with inherited access to App. Repo, clouds and DCIs



SG instances for biologists, chemists, etc.

Types of gateways to be developed from the core WS-PGRADE/gUSE framework



							_		
Adria Science portal	RBI	CRO-NGI	Generic- purpose	Geo-science, meteorology	3.5.2	1	1	Dec 2012	
AEGIS CMPC Portal	<u>IPB</u>	AEGIS, cmpc.aegis.rs	Application Specific	Condensed Matter Physics Community	3.5.2	1	3	Nov 2012	10
agINFRA Science Gateway	agINFRA project	aqINFRA VO	Generic- purpose	Agricultural Community	3.5.1	1	2	Dec 2012	10
AutoDock portal	MTA SZTAKI LPDS, developed by <u>UoW</u>	EDGeS@home		bioinformatics (molecular docking)	3.4.3	2	3	May 2012	~110
<u>BIFI-Unizar</u> portal	BIFI-Unizar	AraGRID, PireGRID			3.4.2				23
Citizen Web Community Gateway	E-Group	CloudBroker	Application specific	Public administration/government	3.5.1	1	2	Dec 2012	10
AMC e-BioInfra Gateway	AMC e-BioScience Group	BiGGrid via gLite, VL-e Med (vlemed) VO, local clusters	both	bioinformaticians, neuroscientists, biochemists	3.5.2	1 HW 1 Virtual	3	Nov 2012 Mar 2013	34
Etics Portal	4D Soft	-			-				
<u>iportal</u> (internal access only)	ETH Zürich	Local cluster	Application Specific	Life Science / Proteomics	3.4.4	1	3	Jun 2012 3.4.4 Jan 2012 3.4 Nov 2011 3.3.3.1	~50

Communities developing WS-PGRADE/gUSE based gateways



- 11 partner communities
- 6 subcontractors
- 7 associated members
- EU projects
 - agINFRA, DRIHM,
 - VERCE, VIALACTEA,
 - EDGI, IDGF-SP,
 - SHIWA, ER-Flow,
 - CloudSME



G.V. Kurdyumov Institute for Metal Physics

(Ukraine)



Institute of Physics Belgrade (IPB)

(Serbia)



Leiden Institute for Advanced Computer Science

(Netherlands)



NVG Scientific Sdn Bhd

(Malaysia)



Ruđer Bošković Institute

(Croatia)

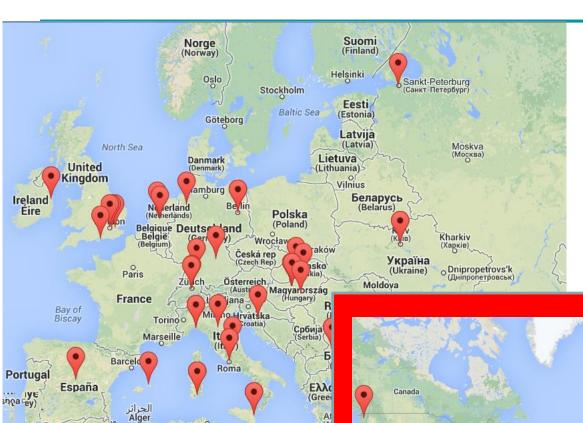


The Meteorology Group of Universidad de Cantabria

(Spain)

gUSE based gateways





Mediterranean

تونس Tunis



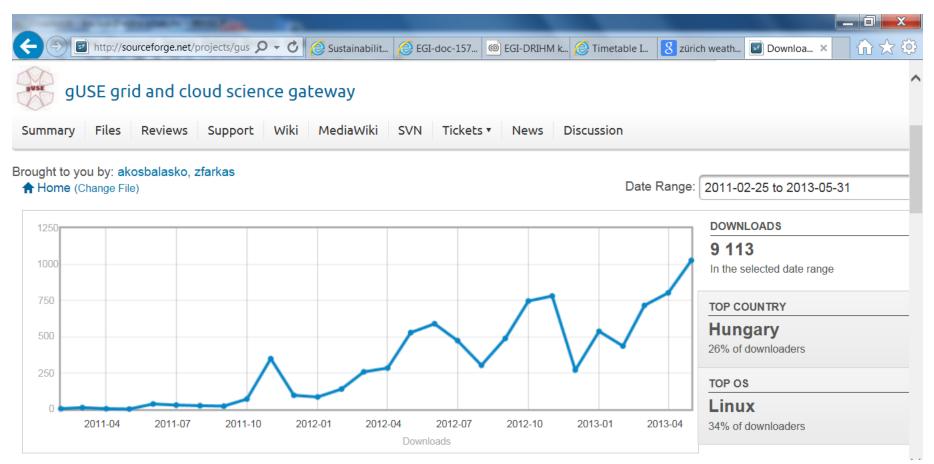
gUSE Roadmap based on community effort



→ http://www.sci-bus.eu/guse	ustainabi 餐 EG	I-doc-1 🍩 EGI-DRIH 🙋 Timetable 💈 züric	hwe 🖬 [Oownloa	e all	SE r ×	_ □ x
	17. CloudBroker Commercial component extension	portlets that provide cost display, billing and invoicing related to the CloudBroker Platform. The new portlets will enable users to check out their CloudBroker account balance, spendings, and can get an overview on the prices of using Resources, Softwares and Instances offered by the CloudBroker Platform.			3.5.8+	Under	^
	18. Integrated graph editor	The current Graph Editor can be run as a separate Java Web Start application, which can be run only on desktop computers (and in the basic setup, doesn't work with OpenJDK). The INAF team is working on a web-based version of the graph editor that integrates seamlessly into the set of WS-PGRADE portlets, so editing graphs will become an easier task, and mobile device users will be able to use this feature as well.	INAF			In progress	
	19. SGE DCI Bridge plugin	The development will extend the DCI support capabilities of DCI Bridge with SGE-based clusters, thus users of WS-PGRADE/gUSE will be able to run jobs on SGE-based local resource management systems directly. The SGE plugin is similar to the PBS and LSF DCI Bridge plugins, that is an SGE job submission node (from where SGE jobs can be submitted and managed) is accessed	Davor Davidović (RBI)			In progress	V
	P	Commission and forest the DCI Bridge mechine		EN 🖀 .	- (~ [j .d	23:24 2013.06.03.

gUSE download statistics at sourceforge





Conclusions

Join SCI-BUS as associated member



Why to select WS-PGRADE/gUSE and join the SCI-BUS community?

1.Robustness

Already large number of gateways used in production

2. Sustainability

 The SCI-BUS project and its sustainability and commercialization plan guarantees it

3. Functionalities

- Rich functionalities that are growing according to the SCI-BUS and sourceforge community needs
- 4. How easy to adapt for the needs of the new user community?
 - Already large number of gateways customized from gUSE/WS-PGRADE
- 5. You can influence the progress of WS-PGRADE/gUSE

Where to find further information?



- SCI-BUS web page:
 - http://www.sci-bus.eu/
- gUSE/WS-PGRADE:
 - http://www.guse.hu/
- gUSE on sourceforge
 - http://sourceforge.net/projects/guse/
 - http://sourceforge.net/projects/guse/forums/forum/
 - http://sourceforge.net/projects/guse/develop

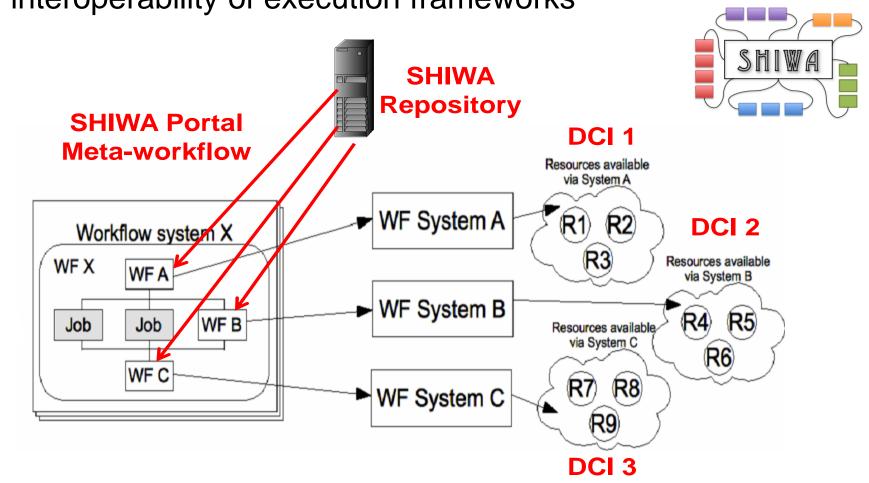
The flexibility of using WS-PGRADE/gUSE



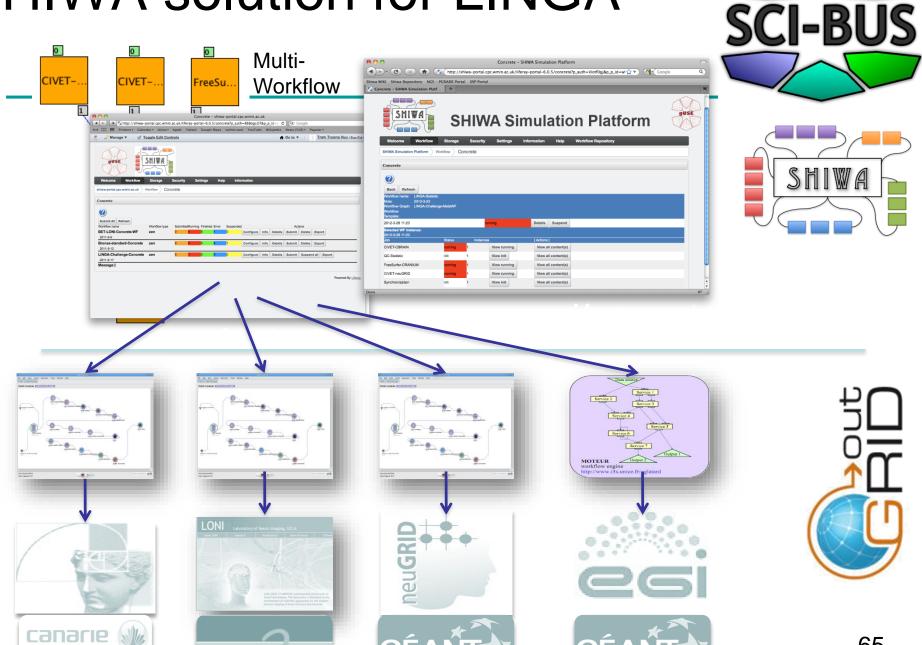
- Support for workflow interoperability based on SHIWA technology
 - Coarse Grained Interoperability approach for integrating many different types of workflows as metaworkflows
 - Using other workflows available in the intermediate IWIR format in the SHIWA repository

Coarse-grained interoperability **SCI-BUS**

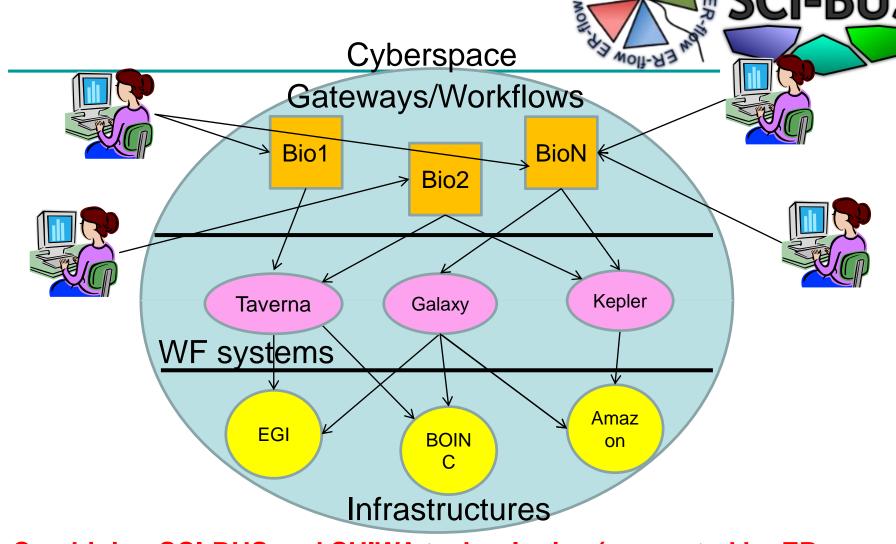
 CGI = Nesting of different workflow systems to achieve interoperability of execution frameworks



SHIWA solution for LINGA



SCI-BUS and ER-Flow



Sign Ex-flow

Combining SCI-BUS and SHIWA technologies (supported by ER- $_{66}$ Flow) users can access and use many WFs and many infrastructures in an interoperable way no matter which is their home WF system

Set of technologies to achieve the required Cyberspace



- DCI Bridge to access a large set of computing resources:
 - Grid, cluster and supercomputer
- CloudBroker Platform to access a large set of cloud computing resources:
 - OpenStack, OpenNebula, Eucalyptus, Amazon, IBM
- Data Avenue Blacktop to access a large set of data resources (service under development):
 - Cloud, grid, cluster, supercomputer
- SHIWA gateway service (based on WS-PGRADE/gUSE) to create and manage meta-workflows accessing all the infrastructures above