



WS-PGRADE/gUSE

Supporting e-Science communities in Europe

Zoltan Farkas

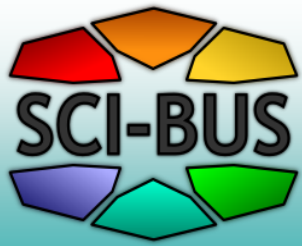
MTA SZTAKI LPDS, Hungary

zoltan.farkas@sztaki.mta.hu



Outline

- WS-PGRADE/gUSE in a nutshell
- Flexibility of:
 - Parallelism
 - Compute infrastructure access
 - Data storage access
 - Customization possibilities
- Customized gateway examples
- Interested in the details?



WS-PGRADE/gUSE

Generic-purpose gateway framework

- Based on Liferay
- General purpose
- **Workflow-oriented** gateway framework
- Supports the development and execution of workflow-based
- Supports the fast specific gateway technology
- Most important design aspect is **flexibility**





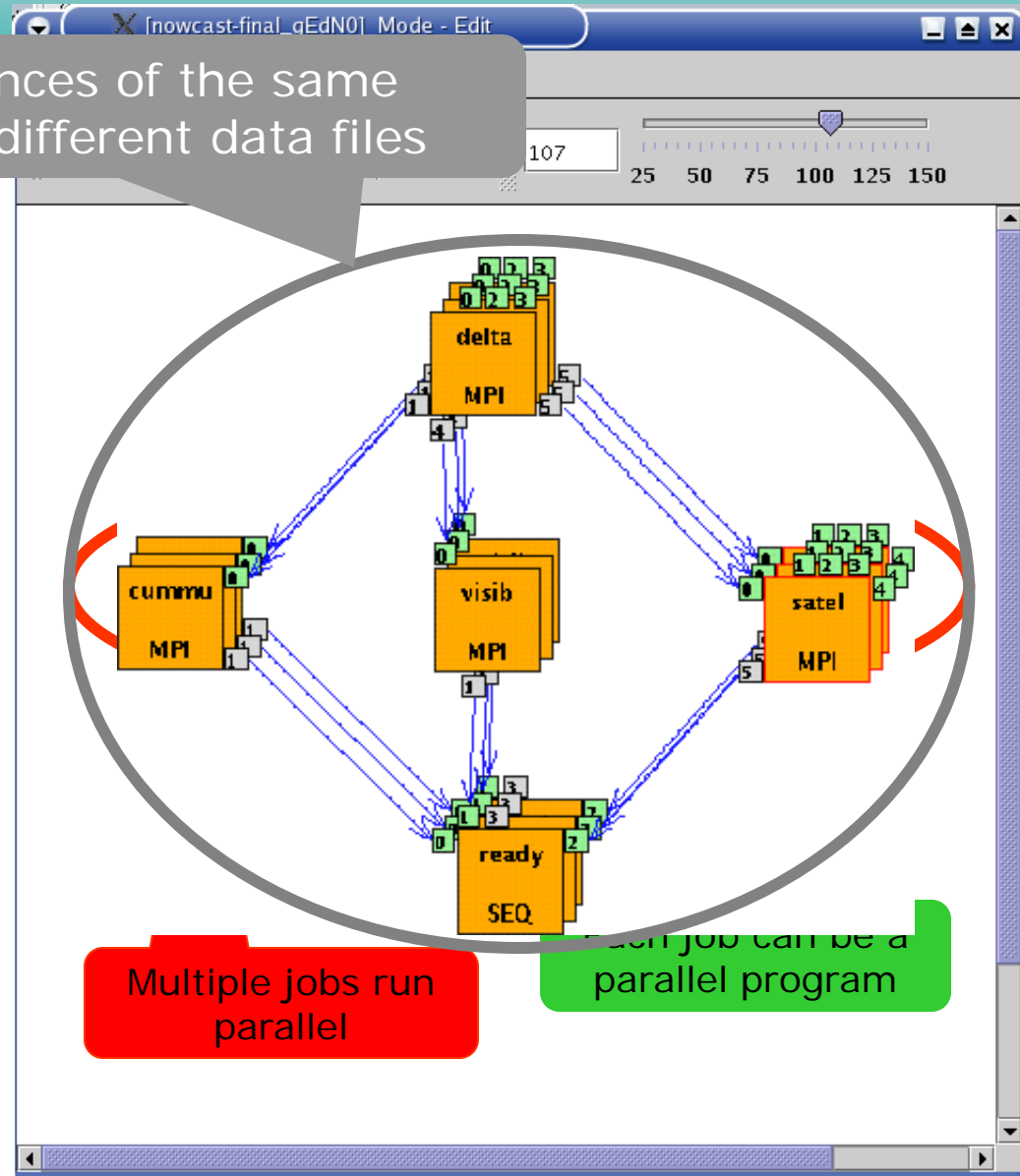
Flexibility in exploiting parallelism

Multiple instances of the same workflow with different data files

– Parallel execution inside a workflow node

– Parallel execution among workflow nodes

– Parameter study execution of the workflow





Flexibility of using compute infrastructures

- Flexible management of **Security**:
 - Individual users' certificate (X.509, SAML, pubkey, ...)
 - Robot certificates
- Flexible 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 (Direct Access, EGI FedCloud, CloudBroker Platform)
 - XSEDE



Flexibility in data storage access

- **Use 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
- **Data Avenue Liferay portlet** to access the data transfer services of Data Avenue Blacktop
- See details: <https://data-avenue.eu/>
- Currently supported protocols:
 - HTTP(s), SFTP, GSIFTP, SRM, iRODS, S3



Data Avenue services



Two panel view Edit favorites History

Protocol: gsiftp Go Protocol: s3 Go
URL: dpm.hpcc.sztaki.hu:2811/tmp/gridtest/ Favorites URL: s3.lpsd.sztaki.hu/gridtestbucket/ Favorites

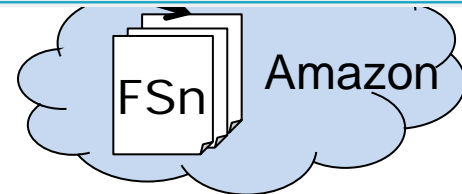
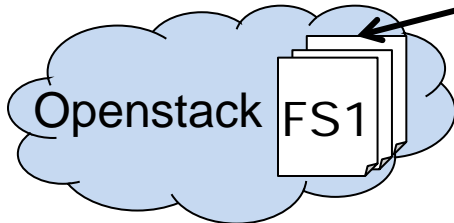
Name	Size	Last modified	Name	Size	Last modified
..			..		
job1_inputs		26.02.2014 10:12:18	job1_output		
job2_inputs		26.02.2014 10:12:32	job2_output		
input.dat	399 B	20.02.2014 10:43:30	output.dat	37.1 MB	26.02.2014 09:14:13
temp_results1.dat	361.3 MB	26.02.2014 10:11:09			
temp_results1.zip	16.9 KB	26.02.2014 10:15:50			

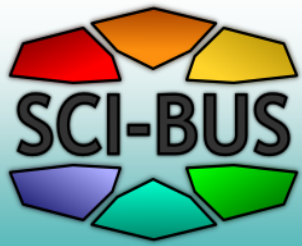
Refresh Mkdir Rename Delete Copy Move Download Upload

Copy from gsiftp://dpm.hpcc.sztaki.hu:2811/tmp/gridtest/temp_results1.dat to s3://s3.lpsd.sztaki.hu/gridtestbucket/ Transferring Details Cancel
Started: 26.02.2014 10:43:55 175270232 bytes/378800128 bytes

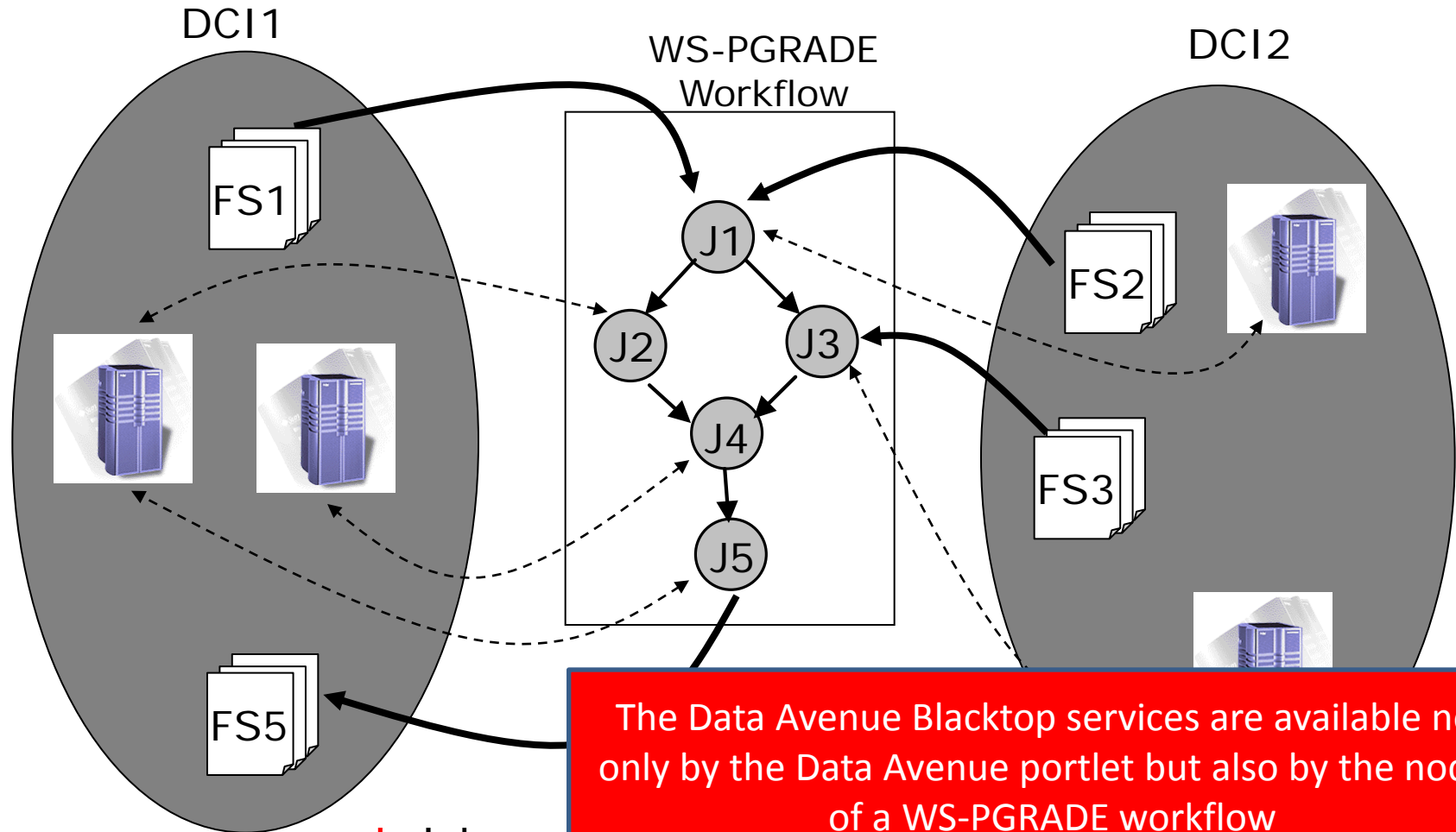
Copy from gsiftp://dpm.hpcc.sztaki.hu:2811/tmp/gridtest/temp_results1.zip to s3://s3.lpsd.sztaki.hu/gridtestbucket/ Done Details Delete
Started: 26.02.2014 10:43:15 17352 bytes/17352 bytes

Copy from gsiftp://dpm.hpcc.sztaki.hu:2811/tmp/gridtest/temp_results1.dat to s3://s3.lpsd.sztaki.hu/gridtestbucket/ Canceled Details Delete
Started: 26.02.2014 10:42:41 268431552 bytes/378800128 bytes





Generic data transfer among WS-PGRADE workflow nodes



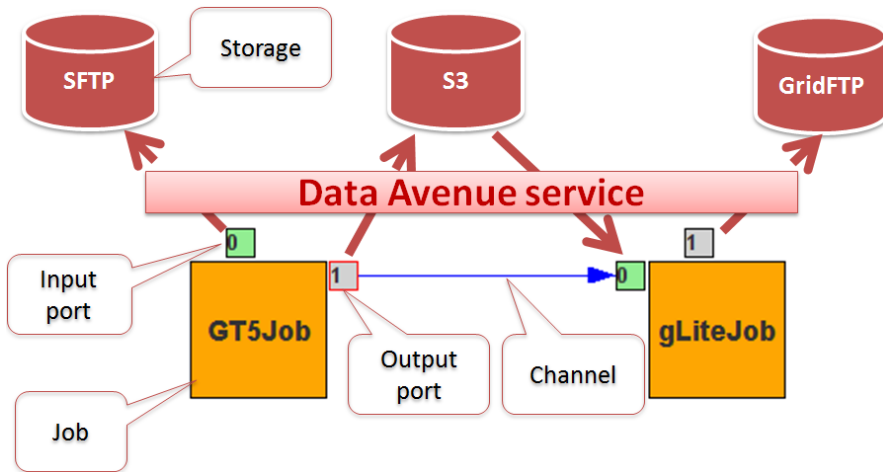
J: Job

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



Data Avenue in WS-PGRADE/gUSE

- Data sources and destinations of jobs can be selected
- gUSE automatically manages data transfers using Blacktop
- Actual transfer delegated up to the worker node wherever possible, *bypassing* the Blacktop service if the middleware is capable of handling the protocol



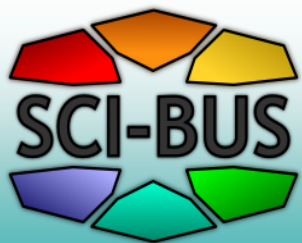
The screenshot shows the configuration interface for a job. The top window is titled "Configure" and shows fields for "Job's name" (Job0) and "Optional note" (Description of Job). Below these are icons for "Job Executable", "Job KV", "JDL(RSL)", and "History". The main area shows "Port Number0" with "Port Name: PORT0" and "Description of Port". It includes a section for "Input Port's Internal File Name" (PORT0) and "Port dependent condition allowing the run of the job:". Below this is a "Source of input directed to this port:" section with a "URL" field (http://*****.*****.port) and "Authentication id: 1000-1334". A red arrow points from the "DATA AVENUE" icon in this section to the bottom window. The bottom window is also titled "Configure" and shows a "1. First Step" tab. It contains a table with columns "Name", "Size", and "Last modified".

Name	Size	Last modified
extras		
apache-tomcat-6.0.39-deployer.tar.gz	987.0 KB	28.01.2014 00:01:33
apache-tomcat-6.0.39-deployer.zip	989.9 KB	28.01.2014 00:01:33
apache-tomcat-6.0.39-halldocs.tar.gz	3.5 MB	28.01.2014 00:01:33
apache-tomcat-6.0.39-windows-#64.zip	8.3 MB	28.01.2014 00:01:33
apache-tomcat-6.0.39-windows-x64.zip	7.8 MB	28.01.2014 00:01:33
apache-tomcat-6.0.39-windows-x86.zip	7.7 MB	28.01.2014 00:01:33
apache-tomcat-6.0.39.exe	7.7 MB	28.01.2014 00:01:33
apache-tomcat-6.0.39.tar.gz	6.7 MB	28.01.2014 00:01:33
apache-tomcat-6.0.39.zip	7.1 MB	28.01.2014 00:01:33



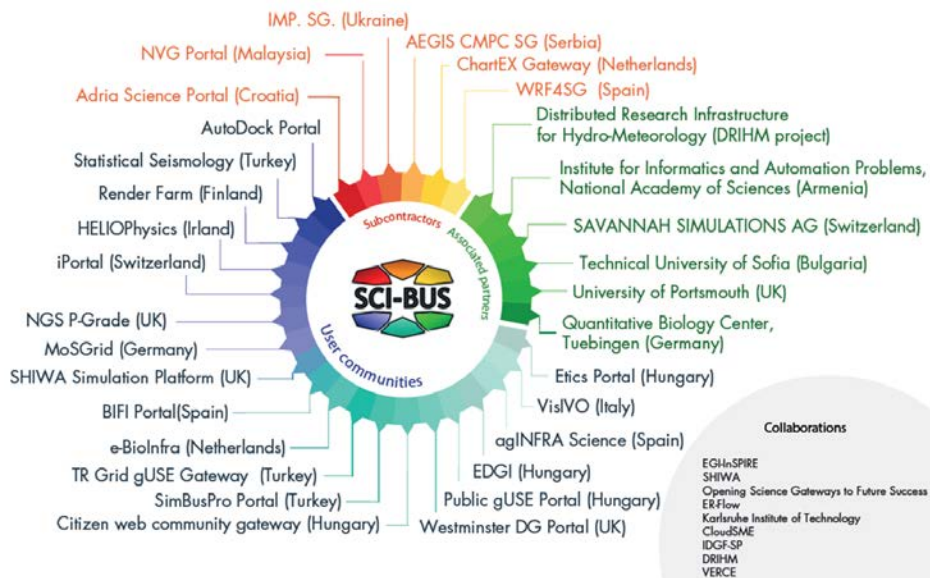
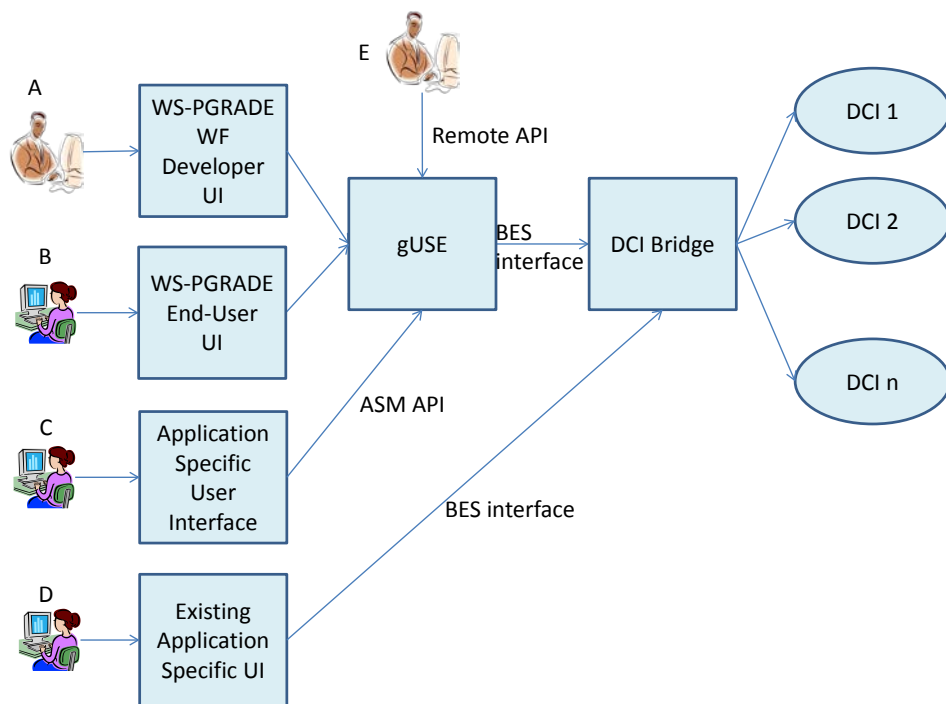
Flexibility of gateway types and user views

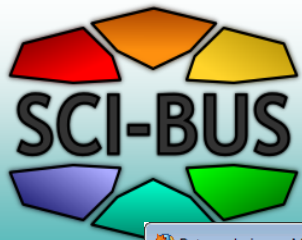
1. **Generic purpose gateways (workflow view)**
 - Core WS-PGRADE/gUSE (e.g. Greek NGI)
2. **Generic purpose gateway for specific technologies (workflow view)**
 - SHIWA gateway for workflow sharing and interoperation
3. **Domain-specific science gateway instance**
 - Autodock gateway (end-user view)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



WS-PGRADE/gUSE customization possibilities

- Application Specific Module: ASM API
- Remote API
- End User View
- Job submission





Some examples of SCI-BUS domain-specific gateways

Data analysis - - Mozilla Firefox

File Edit View History Bookmarks

https://

File Edit View F

amC

Welcome Data analysis

projects data processing

processing

date	project	description
2013-06-03 18:08:13.0	fMRI obs	
2013-06-03 18:04:21.0	fMRI obs	
2013-06-03 17:17:44.0	fMRI obs	
2013-06-03 17:12:49.0	fMRI obs	
2013-06-03 15:08:23.0	fMRI obs	
2013-06-03 11:33:24.0	fMRI obs	
2013-05-30 16:58:20.0	fMRI obs	
2013-05-30 16:38:30.0	fMRI obs	
2013-05-30 15:46:55.0	fMRI obs	
2013-05-30 15:14:56.0	fMRI obs	
2013-05-30 14:50:38.0	fMRI obs	
2013-05-30 14:43:53.0	fMRI obs	
2013-05-30 14:40:21.0	fMRI obs	
2013-05-30 14:08:09.0	fMRI obs	
2013-05-29 16:51:56.0	fMRI obs	
2013-05-29 16:32:39.0	fMRI obs	
2013-05-29 16:03:00.0	fMRI obs	
2013-05-29 16:00:36.0	fMRI obs	

Simulation control panel:

- Start
- Start Time
- End Time
- Manual
- Pause
- Earth
- Stop

STATISTICAL SEISMOLOGY SCIENCE GATEWAY

Welcome Simple Use Storage Security Discussion Statistics Help

SSS-Gateway > Simple Use > SSF4

SSF4 - Complex predictive modeling of earthquake phenomena in time-space domains

Catalog and Sources

Catalog: CombinedCatalog.txt Change

Source: Sources02.txt Change

Site Region

Latitudes: Top: 40.8 Bottom: 38.9

Longitudes: Left: 30.8 Right: 33.7

Select From Map

Models

PSHA Model: Poisson (circled in red)

Magnitude-Frequency Relation: Renewal

Attenuation Relation: Renewal Hybrid

Options

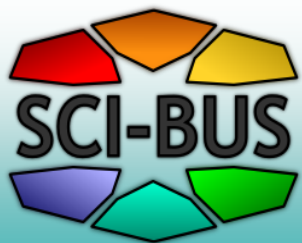
Execution

Run Status: FINISHED Refresh

Results

Download Output Show Probability Curve Show Probability Map

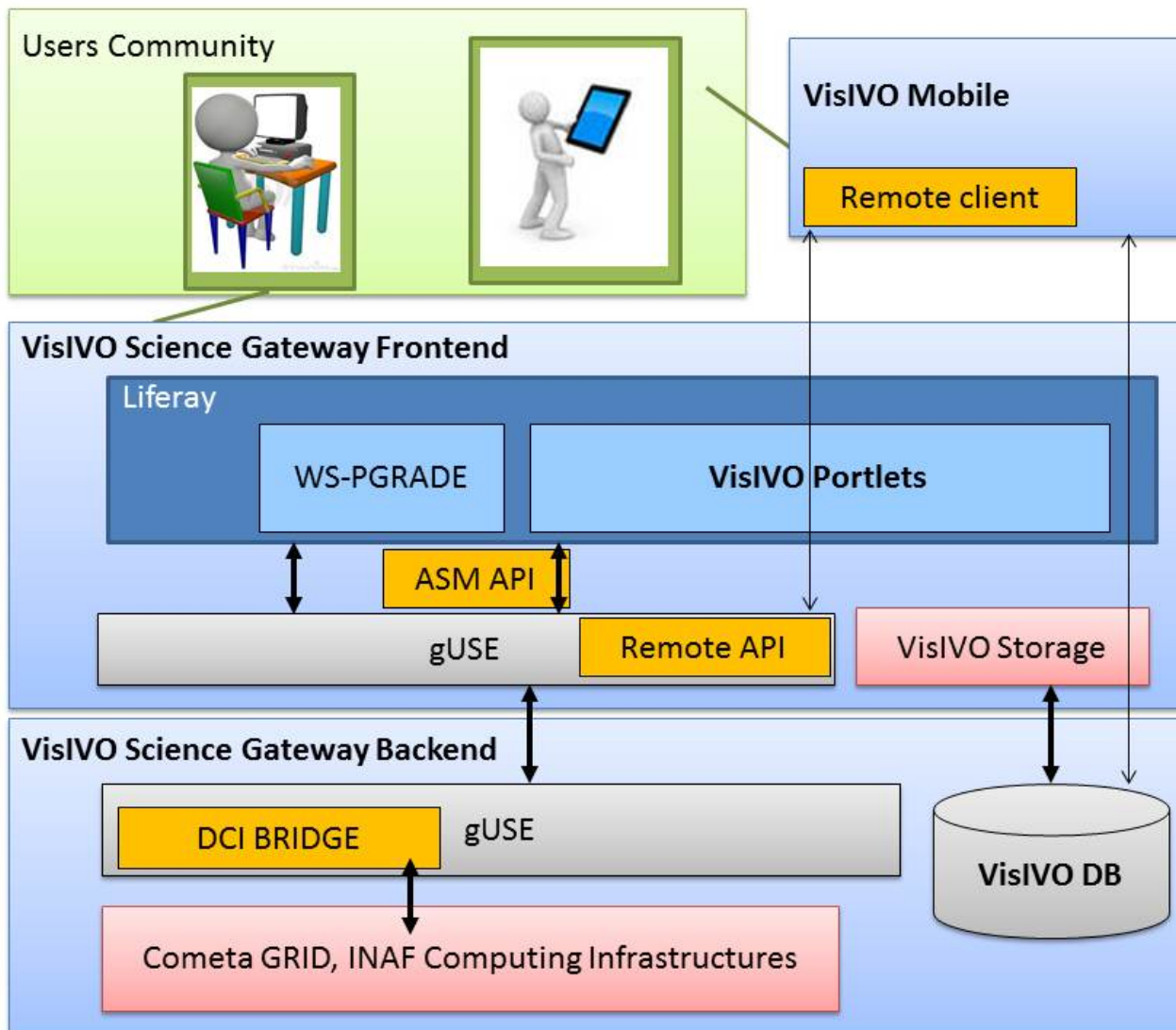
Probability of exceeding PGA level of 0.2 g in 10 years



VisIVO Astrophysics Science Gateway Architecture



INAF



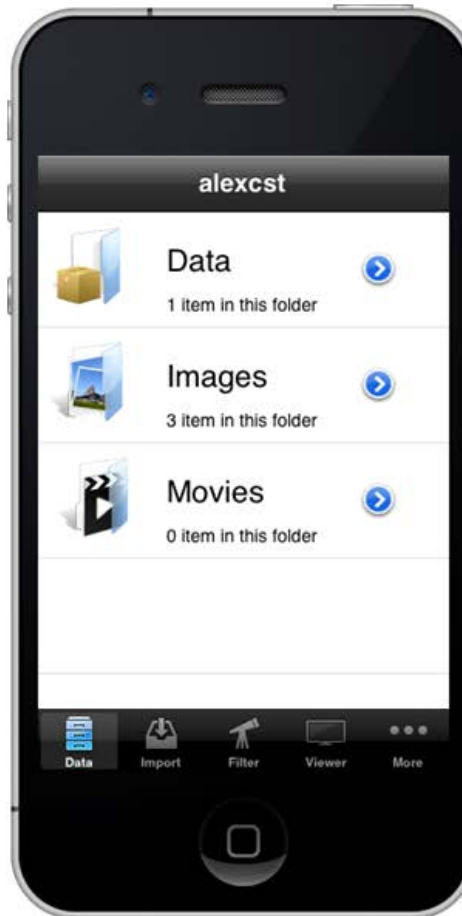


VisIVO Mobile



INAF

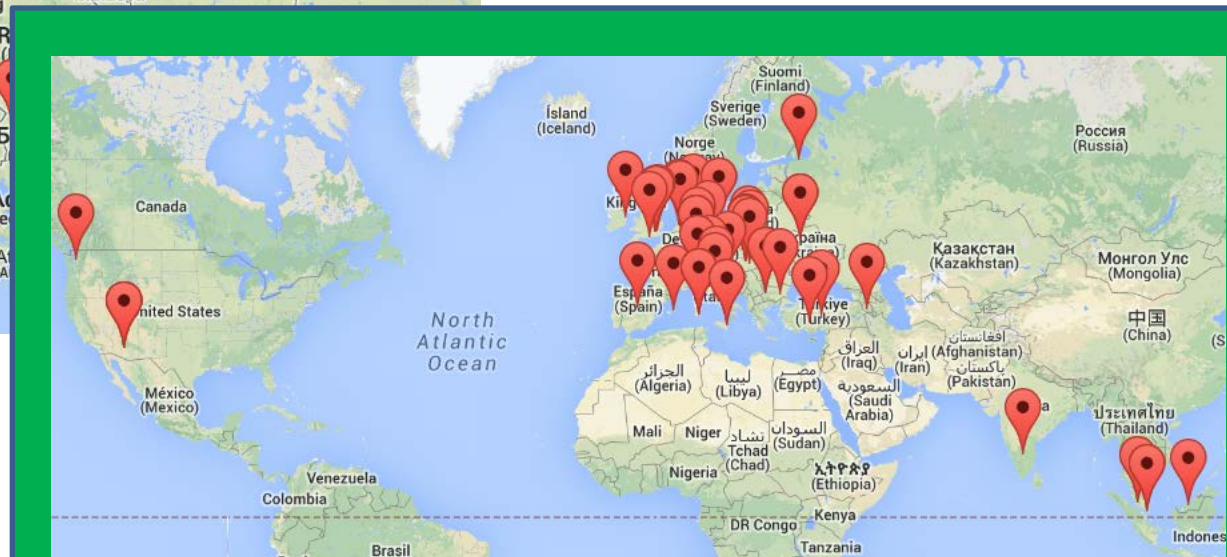
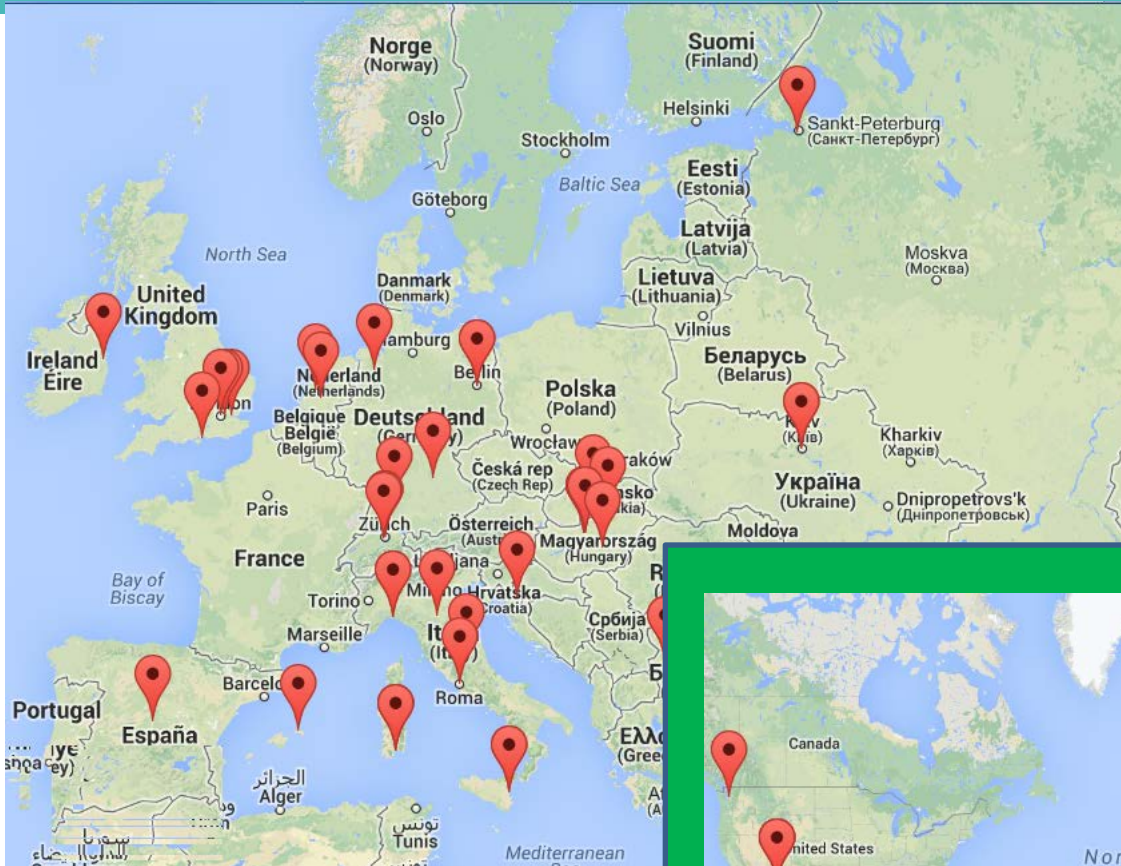
INSTITUTO NAZIONALE DI ASTRONOMIA
NATIONAL INSTITUTE FOR ASTRONOMY

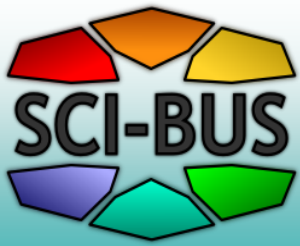




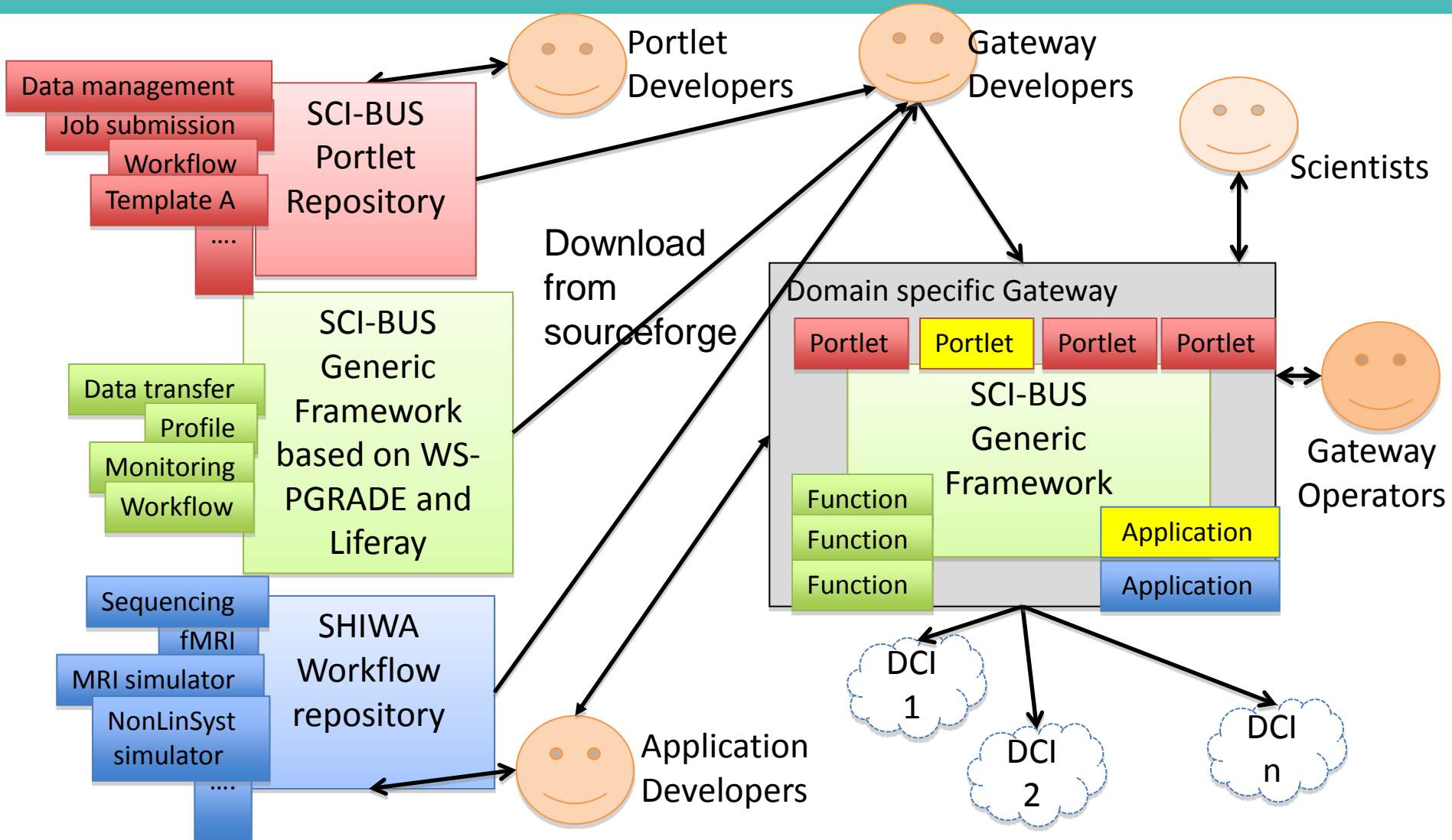
gUSE-based gateways

- Close to 130 deployments world-wide
- More than 17.000 downloads from SourceForge





Summary: SCI-BUS solution to build science gateways by communities





Additional info

- SCI-BUS: <http://sci-bus.eu/>
- WS-PGRADE/gUSE: <http://guse.hu>,
<http://guse.sf.net/>
- Data Avenue: <https://data-avenue.eu/>
- More in-depth talk:
 - When: 2:30-4 p.m. Friday, Oct. 3
 - Where: Purdue University, LWSN, Room 1106
- Thank you for your attention!