



Improvements of the grid infrastructure and services within SEE-GRID

Anastas Misev
MARNET/MARGI/UKIM
Macedonia



- SEE-GRID
 - Project series
 - SEE-GRID – establish infrastructure
 - SEE-GRID-2 – extend infrastructure, grid applications
 - Regional infrastructure
- SEE-GRID-SCI leverages the SEE e-Infrastructure to enable new scientific collaborations among user communities
 - Stimulates widespread e-Infrastructure usage by new user groups
 - Aims to enlarge the regional e-Infrastructure
 - Helps mature and stabilize the National Grid Initiatives in the region

Joint Research Activity 1



SEE-GRID-SCI
SEE-GRID infrastructure for regional eScience

- Aims to improve the usability of the infrastructure and Grid services for the end-users from target communities and manageability of the infrastructure by focusing in
 - Services - Addressing some issues common to several user communities and their applications that are not addressed by the existing middleware and infrastructure, thus contributing to enhancement of services provided to end-users.
 - Operational tools - Researching in the application-focused features of operational tools and development of new or extending of existing operational tools.

JRA1 Tools and services



SEE-GRID-SCI
SEE-GRID infrastructure for regional eScience

- 1 Operational Tools
 - BBmSAMeX (BBmSAM Extensions)
 - AMS (Alert Messaging Service)
 - GSSVA (Security Assessment Tool)
 - JTS (Job Track Service)
 - Logwatch-G (Logwatch Plug-ins for Grid Services)
 - NMTT (No Mercy Ticketing System)
 - USGIME (User/application specific grid infrastructure monitoring extension of P-GRADE Portal)
- 2 Application Services
 - CWRE (Common Workflow Repository Extension of P-GRADE Portal)
 - DM-Web (Data Management Web Portal)
 - ESIP Platform (Environment oriented Satellite Data Processing Platform)
 - Event Logger
 - FM-J-API (SEE-GRID File Management Java API)
 - MEWS (Mathematical Expressions Web Service)
 - RAS (Rendering Application Service)
 - SDSAS (SDS (Seismic Data Server) Applications Service)
 - ULMON (User Level Monitoring Tool)
 - Work Binder

BBmSAM Extensions



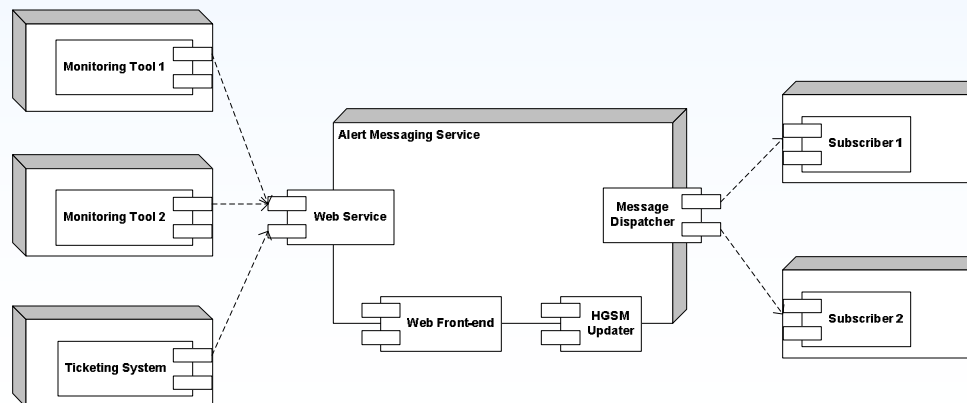
SEE-GRID-SCI
SEE-GRID infrastructure for regional eScience

- BBmSAM periodically synchronizes with HGSM server the database containing information about sites, nodes, services and other GRID data
- SAM Client component performs scheduled tests of defined sites/nodes/services in cooperation with SAM Web Services
- BBmSAM SLA component periodically calculates SLA according to SEE-GRID-SCI SLA specifications
- XML Extension collects real-time data from BBmSAM and SLA database and reports it to external entities

Alert Messaging Service (AMS)

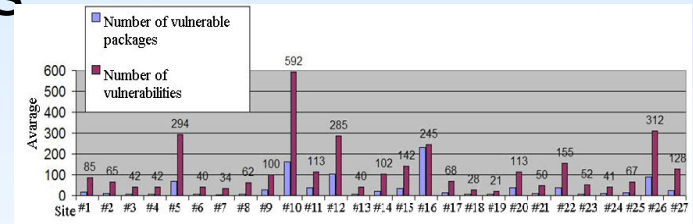
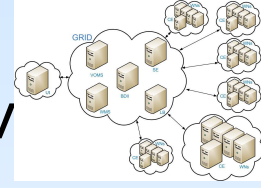


- Area(s) of usage
 - Alert & notification API
 - Monitoring tools
- Description
 - AMS is an operational tool designed to unify and simplify the process of sending and receiving different kinds of alerts in the grid (or any other similar) infrastructure.
 - Java EE application running on a JEE Application Server.
 - A single Web Service Interface is provided for all the monitoring tools, ticketing systems, etc. with the purpose of receiving alerts.
 - Web Front-end for administrators and subscribers is provided.
 - Alerts are filtered and grouped based on subscriber's preferences and messages are then dispatched via the chosen transport mechanism (E-Mail, IM, SMS).
 - Updater mechanism for automatic configuration updates.
 - Authentication and authorization is based on X.509 certificates.



GSSVA - Grid Site Software Vulnerability Analyzer

- Automation of security/vulnerability analysis of service grid systems
- Monitors the grid sites from a security point of view
- Based on PAKITI with modifications
- Basic features:



- centralized status monitoring Server,
- using grid infrastructure (grid's protocol),
- scalable,
- user friendly GUI,
- DN based authentication
- restricted vulnerability status information access on user level

Pakiti: "security" view - hosts for TESZT (31 October 2008 22:20)

Order by: Choose view:

Tag: test vers

Scientific Linux 4

vulnerable RPMs	CVEs	hostname	current kernel	last report	version
12	15		2.6.18.8-xen-3.2.1-2	7 September 2008 18:04	g-ena_vers
4	4		2.6.10.0-xen-3.2.1-2	30 September 2008 17:22	g-ena_vers
7	11		2.6.18.8-xen-3.2.1-2	24 October 2008 22:51	G-E_vers

Formatted in 0.01 seconds

■ Handy tool to make grids more secure

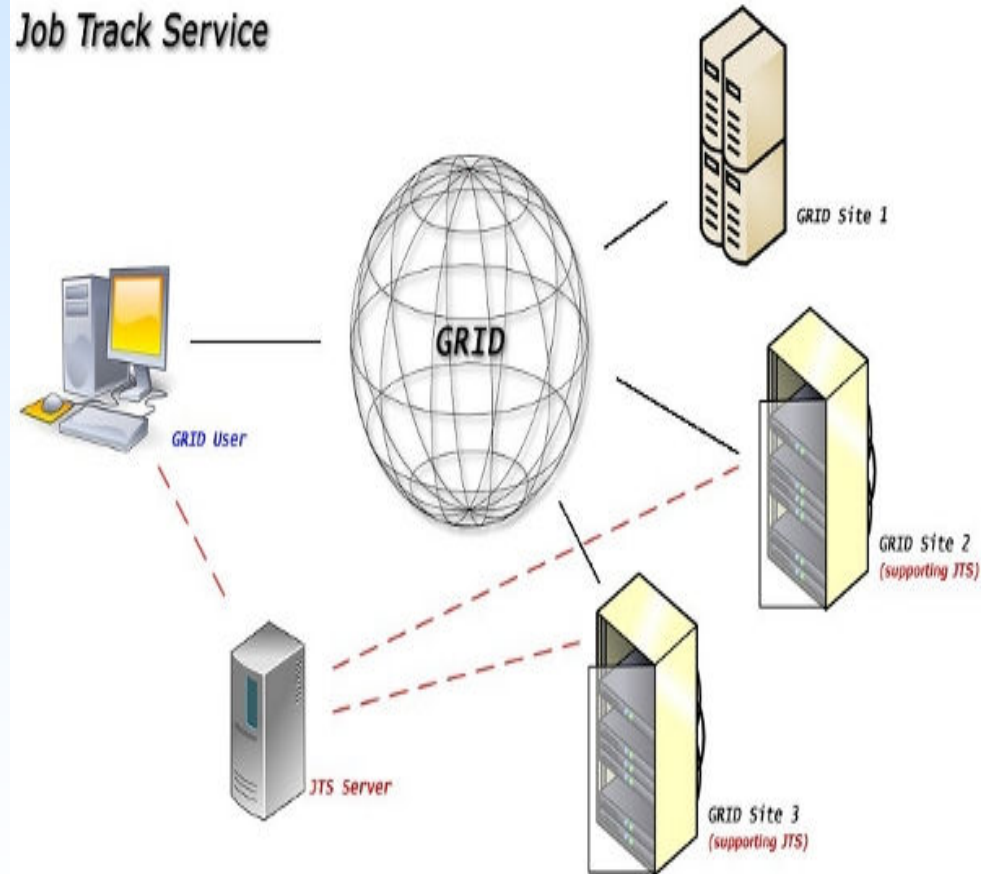
JTS (Job Track Service)



SEE-GRID-SCI
SEE-GRID infrastructure for regional eScience

- JTS implements messaging system to allow QoS for jobs, gathering and analysis of performance data and provision of application specific info from central point. Emphasis on MPI and multi-CPU jobs relevant to Environmental VO.
- It provides enhancements to the current monitoring and accounting system, incorporating performance information and allows provision of QoS by management of batch system reservations.

Job Track Service





- Area(s) of usage
 - Local monitoring of grid services.
- Description
 - Develop modules for logwatch and similar tools to look for and report anomalies in log files generated by grid middleware.
- Additional details
 - Logwatch is a customizable log analysis system. Logwatch parses through your system's logs for a given period of time and creates a report analyzing areas that you specify, in as much detail as you require. Logwatch is easy to use and will work right out of the package on most systems.
 - <http://www.logwatch.org/>

NMTT (“No Mercy” Trouble Ticketing System)



SEE-GRID-SCI
SEE-GRID infrastructure for regional eScience

- NMTT is an enhancement of Oneorzero application, that has been used since SEE-GRID1, and SEE-GRID2 projects, and it was also used in the EGEE1, EGEE2 projects as a regional help desk application which was interfaced with the GGUS system providing for the South-Eastern Europe ROC support system
- NMTT is integrated with GGUS, the EGEE support system, by using GGUS's web services to escalate local tickets to GGUS, and by using email notifications to import GGUS updates.
- NMTT is integrated with Nagios and operators can automatically create a ticket for a corresponding Nagios alarm directly from the Nagios web interface, reducing the effort required to track all the errors reported by the monitoring systems.
- Trac repository (source code available):
<http://nomercy.grid.ici.ro/>

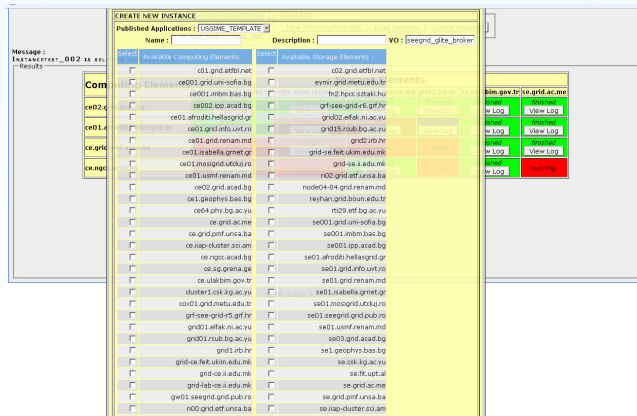
USGIME - User/application Specific Grid Infrastructure Monitoring Extension



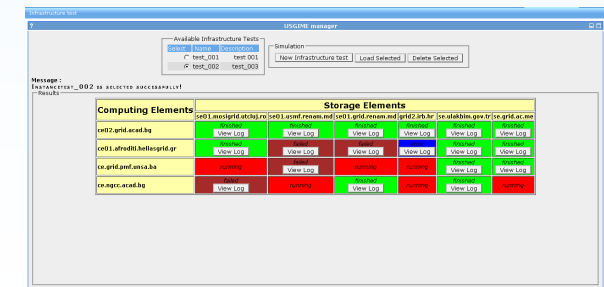
SEE-GRID-SCI
SEE-GRID infrastructure for regional eScience

- User/application specific grid infrastructure monitoring tool
- Allows not only site administrators but normal grid users to explore and monitor the status of available grid infrastructure and grid application services.
- Integrated with P-GRADE Portal

Create infrastructure test



Getting status results (with color codes and error logs)





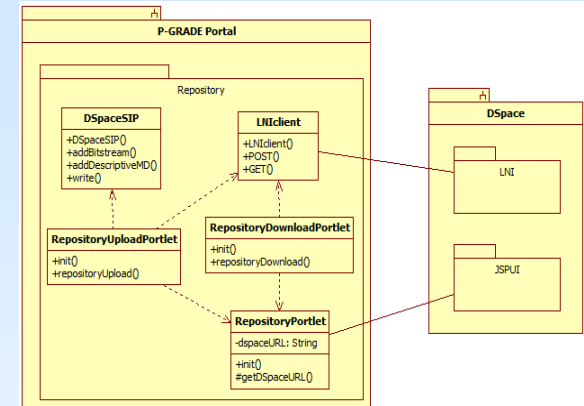
- Area(s) of usage
 - Workflow specification, verification and execution
- Description
 - The service enables execution of complex petri-net like Grid Workflows specified in GWorkflowDL language
 - Current workflows consist of simple DAG
 - Complex workflows enable definition of workflows containing:
 - Loops
 - Choices (Non-determinism)
 - Workflow execution enables gridification of wider range of applications
 - Data can be processed in parallel independently
 - Parametric studies
 - Jobs that involve different programs to run in a series

CWRE – Common Workflow Repository Extension



SEE-GRID-SCI
SEE-GRID infrastructure for regional eScience

- Workflow repository
- Allows users to share their application workflows
- Supports collaborative development among P-GRADE Portal users
- Based in DSPACE
- Integrated with P-GRADE Portal



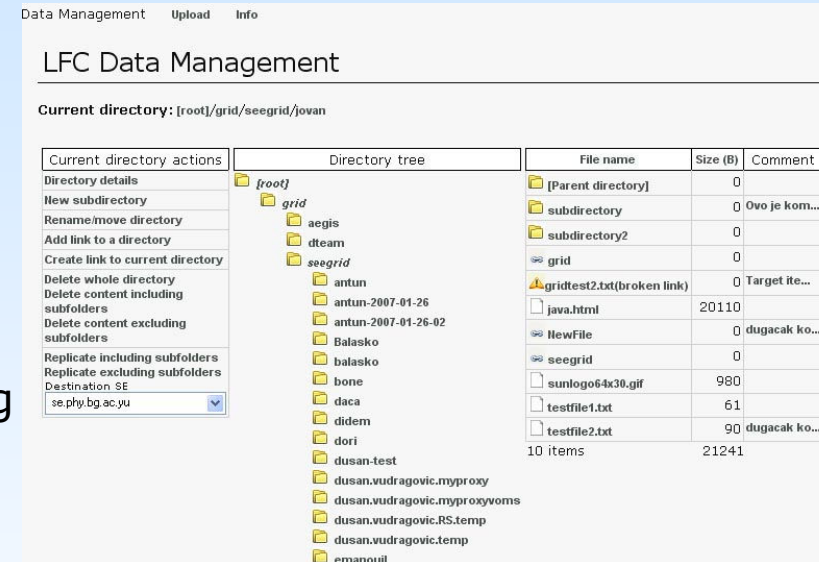
Data Management Web Portal (DM-Web)

■ Area(s) of usage

- Data Management
- Logical Files
- Development Environments and Portals

■ Description

- DM-Web is a portal that enables easy interaction with grid file catalogue using simple Web-based user interface.
- Based on SEE-GRID File Management Java API, offers many file and directory management operations.
- Can be integrated with other Web applications.
- Portal appearance and the set of allowed operations can be customized by using application-specific configuration files, the way it was done for WRF-ARW and REFS applications.



Data Management Upload Info

LFC Data Management

Current directory: [root]/grid/seegrid/jovan

Current directory actions	Directory tree	File name	Size (B)	Comment
Directory details	[root]	[Parent directory]	0	
New subdirectory	grid	subdirectory	0	Ovo je kom...
Rename/move directory	aegis	subdirectory2	0	
Add link to a directory	dteam	grid	0	
Create link to current directory	seegrid	gridtest2.txt(broken link)	0	Target ite...
Delete whole directory	antun	java.html	20110	
Delete content including subfolders	antun-2007-01-26	NewFile	0	dugacak ko...
Delete content excluding subfolders	antun-2007-01-26-02	seegrid	0	
Replicate including subfolders	Balasko	sunlogo64x30.gif	980	
Replicate excluding subfolders	balasko	testfile1.txt	61	
Destination SE	bone	testfile2.txt	90	dugacak ko...
	daca			
	didem			
	dori			
	dusan-test			
	dusan.vudragovic.myproxy			
	dusan.vudragovic.myproxyvoms			
	dusan.vudragovic.RS.temp			
	dusan.vudragovic.temp			
	emanouil			

10 items 21241



- ESIP is based on the gProcess platform developed by the MedioGrid national research project providing the user with the possibility to explore the optimal solutions for Grid processing and information searching in the multispectral bands of the satellite images.
- The gProcess platform is an interactive toolset supporting the flexible description, instantiation, scheduling and execution of the Grid processing.
- ESIP layers on gProcess a set of Web and Grid services, and image oriented basic operators
 - EditorWS – used to retrieve information on available resources (e.g. lists of operations, subgraphs, satellite image types, and data types).
 - ManagerWS – used to retrieve information on workflows (i.e. PDG, and iPDG), and fetches and uploads related resources (i.e. workflows, operators, services, subgraphs, data).
 - ExecutorWS - executes the instantiated workflows (iPDG) over the Grid, and monitors this execution.
 - ViewerWS - gets and formats the input and output data (e.g. initial and processed satellite images).

Event Logger Application



- Area(s) of usage
 - Application-level event logging
 - Performance benchmarking
- Description
 - Generic Java EE application for creation, analysis, and aggregation of various application-generated events.
 - Events are generated using an API that accesses the service using WS or EJB interface.
 - HTML-based web application for querying of the archive and display of reports.
 - Implemented generation and display of aggregations.

Working Set New Query About

Event Logger - Domain: TestApp

Target type: applicationTest

Event type	finish.test	
Event time interval (optional)	From	2007/12/04 11:08:47
	To (inclusive)	2008/03/22 11:09:00

Get Addresses

Addresses	Source	Group
	Destination	Any
Value (optional)	From	
	To (inclusive)	
Value qualifier (optional)	From	
	To (inclusive)	
Event table size	Display all	
Event count limit	Default	
Aggregation Level	All	

Get Events

Event Logger - Domain: TestApp

Results

Group: Source 147.91.4.104

Event time	Source	Source detail	Destination	Destination detail	Value (ms)	Value qualifier (nodes)	Count	Aggr. interval (s)	Max (ms)	Min (ms)	Mean (ms)	Dev (ms)	Sum (ms)
2008/03/09 10:01:01	147.91.4.104	choppa-pc	grid02.rcub.bg.ac.yu:8080	grid02:8080	12.30000	0.000	1	-	-	-	-	-	-
2008/03/09 10:02:01	147.91.4.104	choppa-pc	grid02.rcub.bg.ac.yu:8080	grid02:8080	24.60000	0.000	1	-	-	-	-	-	-
2008/03/09 10:03:01	147.91.4.104	choppa-pc	grid02.rcub.bg.ac.yu:8080	grid02:8080	36.90000	0.000	1	-	-	-	-	-	-
2008/03/09 10:04:01	147.91.4.104	choppa-pc	grid02.rcub.bg.ac.yu:8080	grid02:8080	49.20000	0.000	1	-	-	-	-	-	-
Summary	Count	Mean (ms)	Dev (ms)	Max (ms)	Min (ms)	Sum (ms)							
	4	30.75000	11.58560	49.20000	12.30000	123.00000							

1 groups

Query

Domain	TestApp
Target type	applicationTest
Event type	finish.test
Event time	From 2007/12/04 11:08:47 To 2008/03/22 11:09:00
Address	Source Group Destination Any
Value	From To
Value qualifier	From To
Event table size	Display all
Event count limit	Default
Aggregation level	All

Repeat Query Edit Query

SEE-GRID File Management Java API (FM-J-API)



■ Area(s) of usage

- Data Management, Logical Files.
- Programmatic grid file management in Java.

■ Description

- FM-J-API is a data management Java API that supports grid file management.
- Supports most of data and logical file management operations available in gLite environment through LFC and lcg-utils C APIs.
- Capable of managing sets of files and replicas within LFC Catalog.
- Provides the grid infrastructure an effective way to implement file manipulation on the grid for any application using the simple API functions within their code.



■ Area(s) of usage

- Running Jobs
- Logical Files
- MPI
- Interactive & Pilot Jobs
- Java networking

■ Description

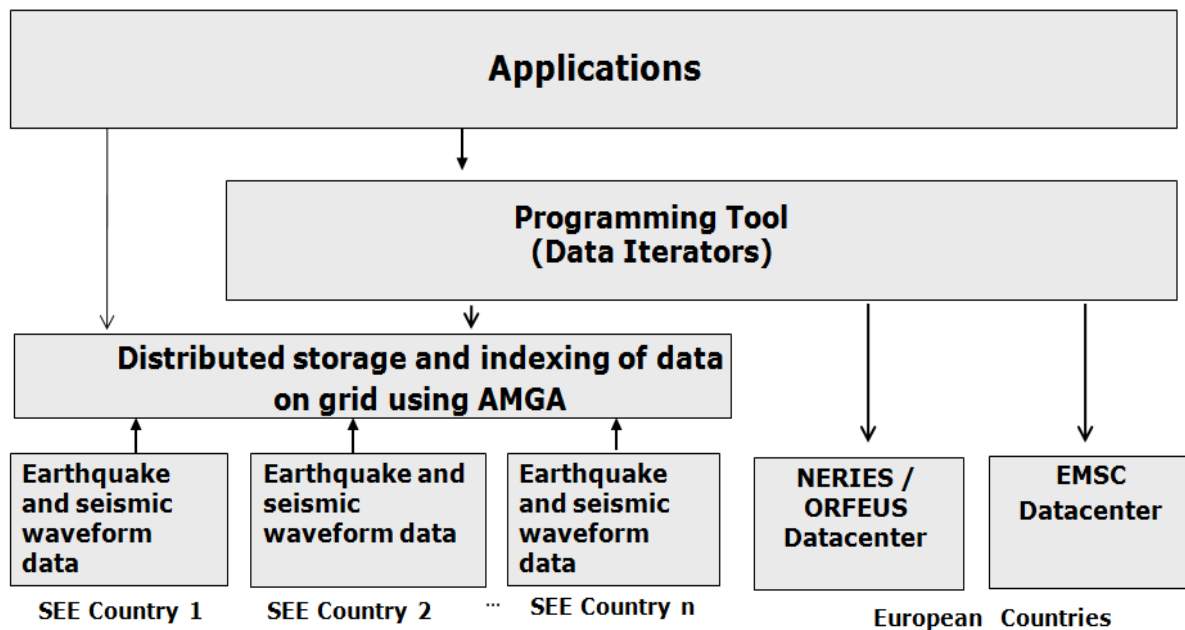
- This Web service could be used for calculation of mathematical expressions. Users instead of using PETSc, or some similar, library for parallel calculation could create XML task and to submit it to the Web server. This Web service would in background use grid and libraries for parallel expression calculation to gain result.
- Instead of using/writing parallel applications for calculating mathematical expressions users could use this Web service that in background does the parallelization using some existing libraries without direct contact with grid.



- Area(s) of usage
 - Rendering animations and visualisation in science
- Description
 - Main aim is to help users to animate and visualize its application output results
 - Fast animation rendering at grid infrastructure based on POV-Ray (Persistence of Vision Ray-Tracer) with output in a specified format
 - The service aims to reduce the necessary Grid knowledge on users' part to minimum, by reducing the number of commands needed to operate to just two.
- Additional details
 - PovRay: <http://www.povray.org/>

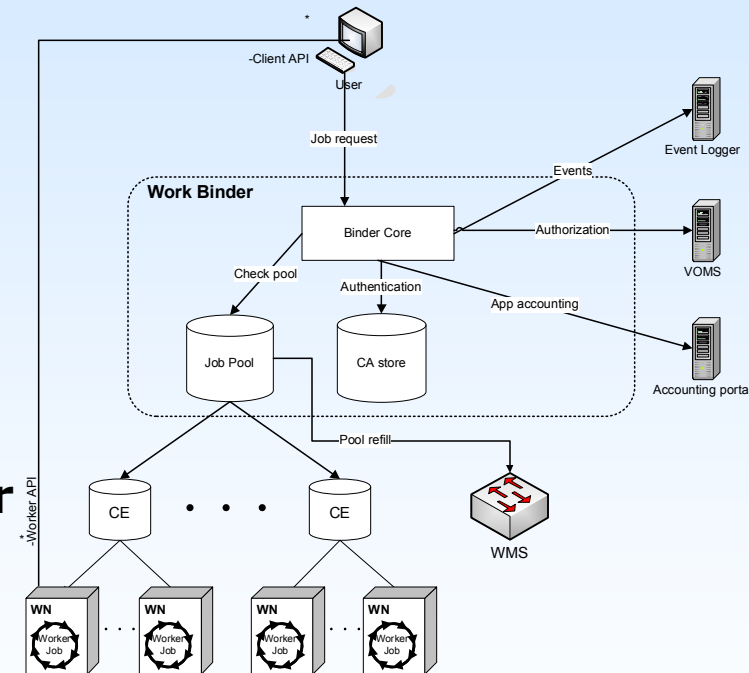
Seismic Data Server Application Service (SDSAS)

- SDSAS provides the following:
 - Distributed storage and serving of seismic data from different partner countries
 - Logical organization and indexing of distributed seismic data
 - C++ iterators that provide easy access to seismic data
 - High level client interface to NERIES web services for accessing ORFEUS waveform data and EMSC quakeml files for earthquake events.



Work Binder Application Service

- Area(s) of usage
 - Running Jobs
 - Interactive & Pilot Jobs
- Description
 - An infrastructural service and Java API for submission, management, and access to computing resources, suitable for interactive grid applications and on-demand computing with frequent usage of short jobs.
 - Allows clients to easily **allocate worker jobs without significant delay**.
 - Supports any number of applications.
 - Suitable for interactive applications and dynamic workflows.
 - **Client-worker communication** (via binder or direct).



Conclusion



SEE-GRID-SCI
SEE-GRID infrastructure for regional eScience

- The overall impact of JRA1 activity is
 - to ease the involvement of new users, applications and resource providers,
 - by automating and simplifying common and burdensome tasks.
- Its results are available to other regional and global initiatives, providing its most successful tools and components to other projects and contributing to optimal development and use of e-Infrastructures and establishing a feature rich grid environment.
- More info on
 - http://wiki.egee-see.org/index.php/JRA1_Commonalities



SEE-GRID-SCI
SEE-GRID eInfrastructure for regional eScience



Thank you for your attention