



Support to MPI, Schedulers and Complex Workflows

Compiled by Isabel Campos
Spanish NGI Director
Presented by Francisco Castejón



ES-NGI: Enol Fernandez (IFCA-CSIC, Santander) and
Ruben S. Montero (U. Complutense de Madrid)
GRID-IRELAND: John Walsh (TCD, Dublin)
PL-Grid: Marcin Plociennik (PSNC, Poznan)



EGI-InSPIRE proposal

- **MPI**
 - MPI Tools based on mpi-start
- **Schedulers**
 - GRIDWAY
- **Complex Workflows**
 - SOMA (Life Sciences env.)
 - TAVERNA (Life Sciences env.)
 - KEPLER-RAS (Fusion env.)

WE NEED TO PROGRESS NOW TOWARDS A
DESCRIPTION
OF THE WORK TO BE DONE TO SUPPORT THE HEAVY
USER
COMMUNITIES IN SA3

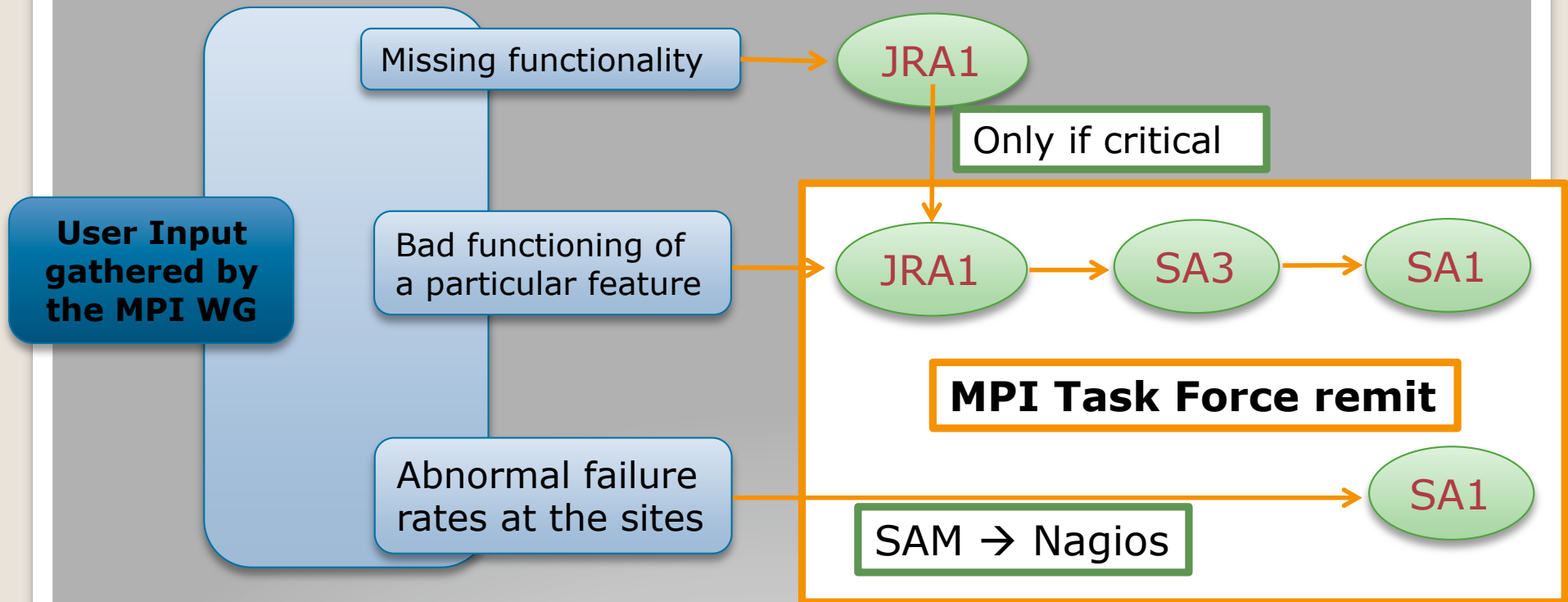


Support to MPI

Input from Enol Fernandez (IFCA-CSIC)
John Walsh (TCD) + MPI Working Group



Support to MPI: final steps in EGEE





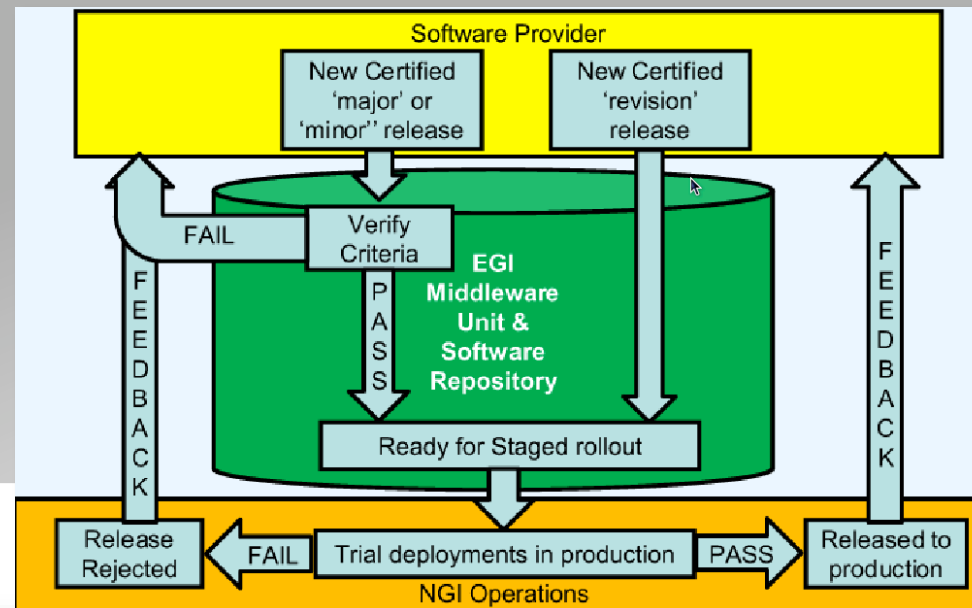
Support to MPI: final steps in EGEE

- CLOSING EGEE-III WITH A **STABLE NUMBER OF SITES** WITH PROPER MPI SUPPORT FROM WHICH TO GROW A WELL DEFINED MPI SUPPORTING INFRASTRUCTURE IN EGI
 - CURRENTLY **94 SITES SUPPORT MPI**, OF WHICH **84% ARE SUCCESSFULLY** PASSING THE SAM/NAGIOS TESTS
 - A **KNOWLEDGE-DATABASE** FOR SITE SUPPORT IS IN PLACE
[HTTP://WIKI.IFCA.ES/E-CIENCIA/INDEX.PHP/MPI_ERRORS](http://wiki.ifca.es/e-ciencia/index.php/mpi_errors)
- **DEFINE THE SET OF REQUIREMENTS** THAT USERS GROUPS FIND NECESSARY FOR **MORE ADVANCED MPI FEATURES** IN THE **EGI ERA**
 - A DOCUMENT IS BEING WORKED OUT INSIDE THE **MPI WORKING GROUP**



Providing MPI Support to EGI

- **mpi-start** will be maintained by **CSIC** inside the **EMI** project
- **Testing & Certification** of middleware components will be organized from **Ibergrid** (Spain + Portugal)
 - For MPI components **LIP** (Portugal), **CESGA** (Spain) will count on the support from **TCD** for the certification effort





Recent developments to improve user support in *mpi-start*

BASIC FEATURES OF MPI-START

- Supports OpenMPI and MPICH
- Supports file distribution in non-shared filesystems
- Hooks mechanism in place to ease I/O at pre- and post-run time

CURRENT VERSION 0.61 (ALREADY CERTIFIED)

- Weaknesses in error reporting identified and fixed
- Improved file distribution mechanism (allows using \$HOME and also other more generic i/o spaces)
- Automatic detection of 32bit or 64bit compiled libraries

FUTURE SUPPORT FOR ADVANCED SELECTION OF CORES/ NODE

- Important for a proper MPI process allocation
- Important for OpenMP support (multithreaded codes)



Summary of actions

- We expect to get **feedback and requirements from the EGI user communities**
- **EGI Requirements** will be transmitted to the **Software Providers**
 - EMI will provide: **mpi-start** (CSIC) and **MPI-utils** (TCD)
- **Testing & Certification** will take place organized by **Ibergrid (CESGA, LIP) and TCD**
- **Rollout of RPMs** will take place with the general mechanism foreseen in EGI
- **User Support and Site support** will be organized in the **EGI Helpdesk**



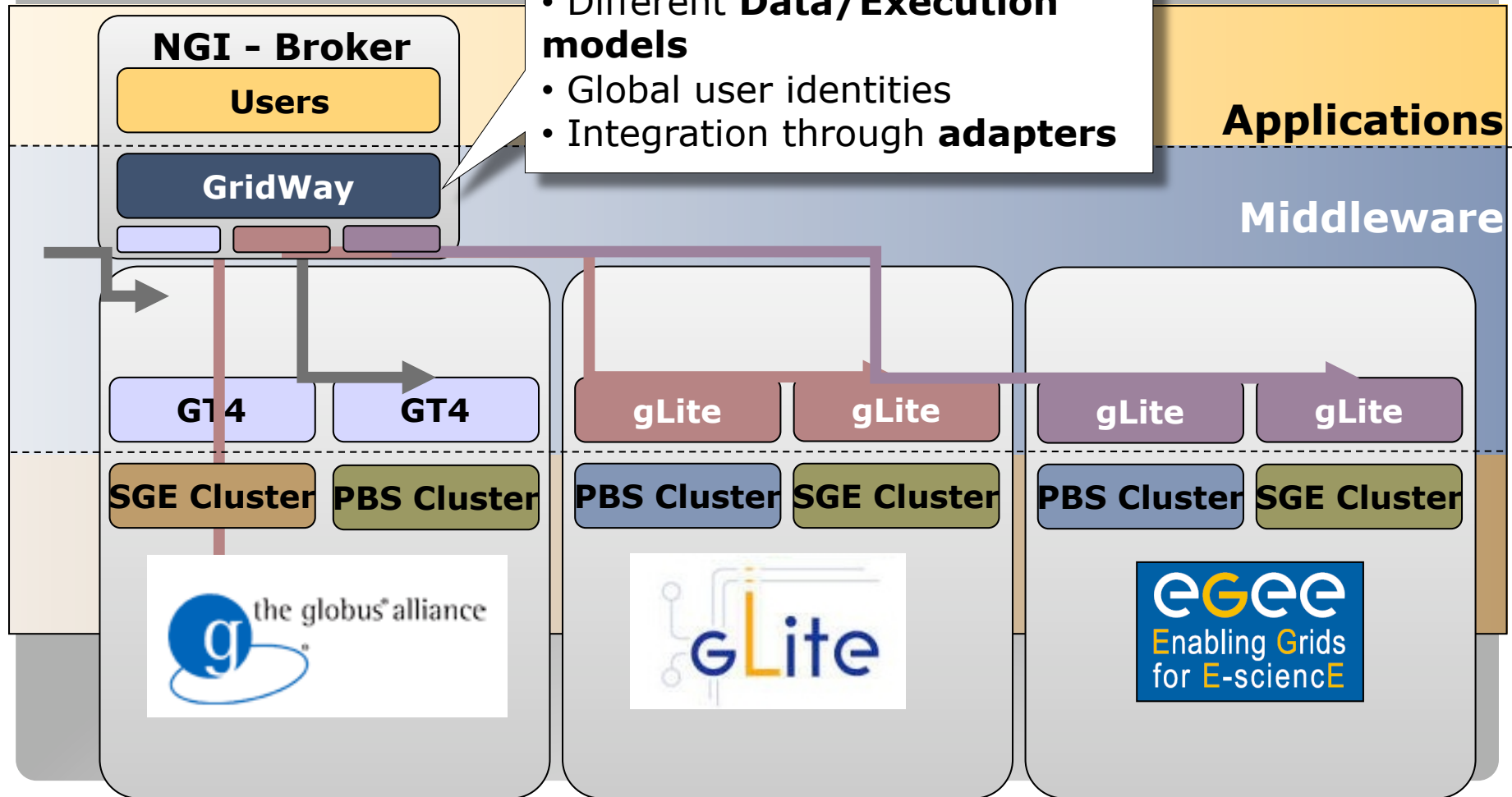
Schedulers: GridWay

Input from Ruben Santiago
Universidad Complutense de Madrid



A Metascheduler to enable interoperation with GT4 and clouds

- Different **Middleware stacks**
- Different **Data/Execution models**
- Global user identities
- Integration through **adapters**





Integration of glite + GT4 in ES-NGI

Integrate NGI-GT4 Resources

- GridWay Broker instance deployed (RedIRIS)
- Backup/Testing available at UCM

Integrate with NGI global services

- User access to GT4 resources

Actual output of the GridWay Broker for the NGI resources (GT4 + glite)

HID	OS	ARCH	MHZ	MEM (F/T)	DISK (F/T)	N (U/F/T)	LRMS	HOSTNAME
0	Linux2.6.24	x86_6	1995	546/2014	5812/16362	0/2/4	PBS	hydrus.dacya.ucm.es
2	Linux2.6.9-	x86_6	3200	15/2007	162867/216680	0/8/8	SGE	aristoteles.inf-cr.uclm.
3	Linux2.6.9-	x86_	3000	146/2007	17977/21643	0/10/146	PBS	ce01.macc.unican.es
4	Linux2.6.20	x86	1595	14/765	8262/10822	0/335/335	SGE	test01.egee.cesga.es
5	ScientificS	i686	4400	16000/16000	0/0	0/50/1344	jobmanager-lcgpbs	ce07.pic.es
6	ScientificS	i686	4400	16000/16000	0/0	0/17/1272	jobmanager-lcgpbs	ce05.pic.es
7	ScientificS	i686	4400	16000/16000	0/0	0/50/1344	jobmanager-lcgpbs	ce06.pic.es
8	ScientificS	i686	3000	1024/1024	0/0	0/335/335	jobmanager-lcgsge	test03.egee.cesga.es
9	ScientificC	i686	3200	2048/2048	0/0	0/777/1645	jobmanager-lcgpbs	gridce01.ifca.es
10	ScientificS	i686	3200	513/513	0/0	0/0/100	jobmanager-lcgpbs	ce-eela.ciemat.es
11	ScientificS	i686	3200	2048/2048	0/0	0/4/62	jobmanager-pbs	ce01.ific.uv.es
12	ScientificC	i686	3194	513/513	0/0	0/24/24	jobmanager-lcgpbs	ce-ieg.bifi.unizar.es
13	ScientificS	i686	1200	1024/1024	0/0	0/3/8	jobmanager-pbs	lcg2ce.ific.uv.es
14		i686	0	0/0	0/0	0/0/0		ngiesce.itaca.upv.es



Complex Workflows: Kepler/RAS

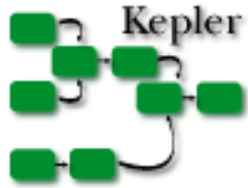
Input from Marcin Plociennik
PSNC, Poznan (PL-GRID)



Handling Scientific workflows



A screenshot of the Kepler workflow editor interface. The main workspace shows a complex workflow diagram with several components: a 'Web Service Actor' connected to 'Web Service' components, which then feed into 'MISHKA' and 'MISHKA2' components. These components are connected to 'Itm Plotter2d', 'Itm Plotter2d2', and 'Itmplot' components. There are also 'String Constant' and 'Constant' components. The workflow is connected to 'Elements To Array' and 'Valeurs propres' components. The interface includes a menu bar (File, Edit, View, Workflow, Tools, Window, Help), a toolbar with various icons, and a 'Components' panel on the left with a search bar and search results. The search results show various plotter and statistical analysis components. The bottom status bar shows the current workflow name 'Unnamed' and several open files including 'root4.png' and 'Valeurs propres'. The top right corner shows the system clock 'Mon Dec 18, 15:30'.



Kepler/RAS - overview

Kepler – workflow orchestration

- A framework for design, execution and deployment of scientific workflows
- Support for concurrent modelling, design and execution
- Precisely defined models of computation and component interaction
- An intuitive GUI that lets rapid workflow composition
- A modular, reusable and extendable object-oriented environment
- An XML based workflow definition – MoML
- Developed in US (UC Davis, UC Santa Barbara, and UC San Diego)
- In terms of Euforia project extended with GRID/HPC execution actors (with usage of RAS services)
- "MINIPROJECT": EGEE-EUFORIA-DEISA
- Chosen and used by fusion community (EFDA ITM)

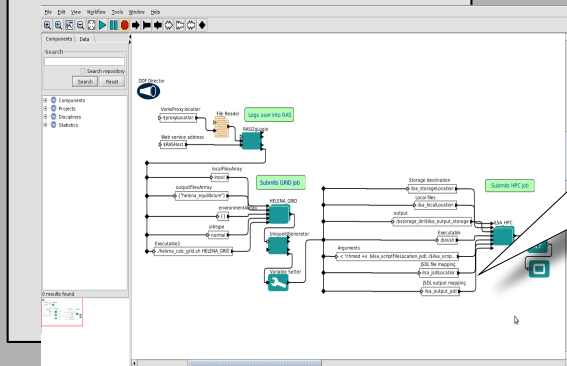
RAS – Roaming Access Server (part of Migrating Desktop)

- Support for different middleware stacks (gLite/UNICORE)
- Developed in terms of int.eu.grid/BalticGrid II/Euforia
- Integrated with VineToolkit/gLogin
- Providing interactive services

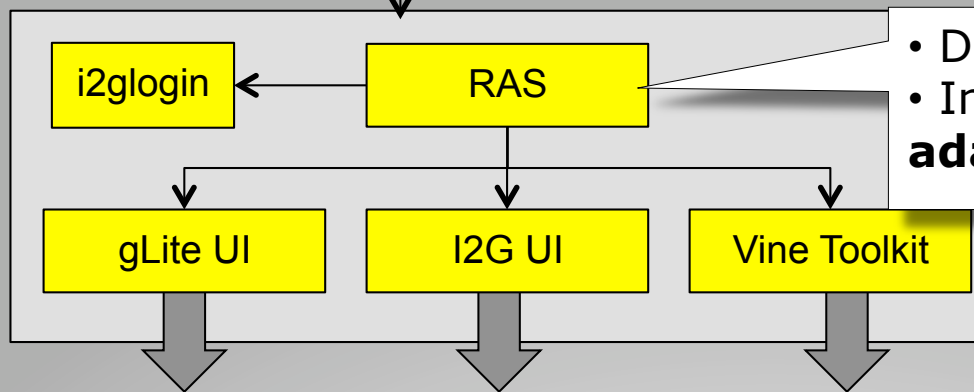


Support for Scientific workflows

Kepler Engine



- Support for different complex use cases
- Different level of integration with applications
- Mixed Grid & HPC workflows
- Support for workflows requiring visualisation and interactive access



- Different **middleware stacks**
- Integration through **plugins/adapters**



i2g



GRID and HPC infrastructures

Activities planned

- **To maintain the integration of Kepler/RAS with the different underlying middleware stacks**
- **To maintain Kepler/RAS services**

But also to

- **Support next application use cases**
- Supporting different workflow scenarios
- Customisation according to specific user's requirements
- Initial target – Fusion community, however since the framework provides generic services, open to support wider user communities (like coming from ES, A&A, LS or other)

FUSION needs

- **MPI**
 - MPI Tools based on mpi-start
- **Schedulers**
 - GRIDWAY
- **Complex Workflows**
 - KEPLER-RAS (Fusion env.)
- **Other tools:**
 - TAPAS
 - DRMAA
 - GIF PORTAL (Russian grid)
 - DATA MANAGEMENT (?)