

Information System

Valeria Ardizzone INFN EGEE NA4 Generic Applications Meeting Catania, 09-11 January 2006







www.eu-egee.org

INFSO-RI-508833



> Introduction to R-GMA and Grid Monitoring Architecture (GMA).

- R-GMA within Testbed
- **R-GMA** in depth:
 - Schema, Registry, Producer and Consumer
 - Query and Storage Types
 - R-GMA Browser
- > A use case of R-GMA for user application.



- Relational Grid Monitoring Architecture (R-GMA)
 - Developed as part of the EuropeanDataGrid Project (EDG)
 - Now as part of the EGEE project.

Enabling Grids for E-sciencE

- Based the Grid Monitoring Architecture (GMA) from the Global Grid Forum (GGF).
- Uses a relational data model.
 - Data are viewed as tables.
 - Data structure defined by the columns.
 - Each entry is a row (tuple).
 - Queried using Structured Query Language (SQL).

Grid Monitoring Architecture(GMA)

- The Producer stores its location (URL) in the Registry.
- The Consumer looks up producer URLs in the Registry.
- The Consumer contacts the Producer to get all the data or the Consumer can listen to the Producer for new data.



R-GMA within Testbed

Enabling Grids for E-sciencE



EGEE NA4 Meeting - Catania, 09-11 January 2006

eeee

CGCC R-GMA: Schema-Registry-Mediator

Enabling Grids for E-sciencE

R-GMA Server

VIRTUAL DATABASE



SCHEMA : it holds the names and definitions of all of the tables in the virtual database, and their authorization rules.

REGISTRY: It holds the details of all producers that are publishing to tables in the virtual database and it also holds the details of "continuous" consumers.

MEDIATOR: a set of rules for deciding which data providers to contact for any given query.

R-GMA: Producer-Consumer

Enabling Grids for E-sciencE

Producers: are the data providers for the virtual database. Writing data into the virtual database is known as publishing, and data is always published in complete rows, known as tuples. There are three types of producer: Primary, Secondary and On-demand.



Consumer: represents a single SQL SELECT query on the virtual database. The query is matched against the list of available producers in the Registry. The consumer service then selects the best set of producers to contact and sends the query directly to each of them, to obtain the answer tuples.

INFSO-RI-508833

eGee



Query and Storage Types

Enabling Grids for E-sciencE

- **Continuous:** as soon as new data becomes available it is broadcast to all interested parties.
- Latest: correspond to intuitive idea of current information.
- **History:** return time sequenced data.

Tuple-store can be in Memory or Database





LATEST RETENTION PERIOD (LRP) and HISTORY RETENTION PERIOD (RTP)

allow producers to periodically purge old tuples, and to give a precise meaning to the "current state".



INFSO-RI-508833



Continuous

Enabling Grids for E-sciencE





History or Latest

Enabling Grids for E-sciencE



https://rgmasrv.ct.infn.it:8443/R-GMA

Enabling Grids for E-science

• • • • • • • • • • • • • • • • • • •							
📄 mozilla.org 📄 Latest Builds 📄 CardLayoutDemo							
mozilla.org Latest R-GMA Browser Home Predefined: GhueServices GhueSite RGMALogs Table Sets	Builds CardLayoutDemo All tables GLUE Info Providers Network Monitoring Service Discovery CMS AppIMONIT GAMIAppStart GhueBatchJob GhueBatchJob GhueBatchQueue GhueBatchSystem GhueCE GhueCEAccessControlBaseRule GhueCESEBind GhueCESEBind GhueHostLocalFileSystem GhueHostNetworkAdapter GhueHostPoolAccount GhueHostRemoteFileSystem GhueHostRole Cite C	Click on a table name to get information about the table.					
Enabling Grids For E-sciencE	GlueSA GlueSAAccessControlBaseRule GlueSE GlueSEAccessProtocol GlueSEAccessProtocolSupportedSec GlueSL GlueService						

INFSO-RI-508833



USE CASE TIMELINE: To submit the JDL file from the GENIUS portal and monitoring its status. In the meantime, from RGMA Browser, monitoring the table and if there is any producers that are publishing tuples. If there is one, to send a query with a predicate to obtain the answer tuples.

eGee

Create a table in Schema

Enabling Grids for E-sciencE

	<u>All tables</u>	\Rightarrow	A
R-GMA	GLUE Info Providers		
Browser	Network Monitoring	≣	Q
	Service Discovery		N
Home	CMS		II
Desdecad	ApplMONIT		Jo
Chus Services	GAMIAppStart		P
GlueServices	GlueBatchJob		H
RGMAL ors	GlueBatchQueue		0
T-11- C-4-	GlueBatchSystem	E	N
Table Sets	<u>GlueCE</u>		N
	<u>GlueCEAccessControlBaseRule</u>		1.
	<u>GlueCESEBind</u>		
	<u>GlueCluster</u>		2
	<u>GlueHost</u>		
	<u>GlueHostLocalFileSystem</u>		
	<u>GlueHostNetworkAdapter</u>		
	<u>GlueHostPoolAccount</u>		
	<u>GlueHostProcess</u>		
	<u>GlueHostRemoteFileSystem</u>		
	GlueHostRole		
	<u>GhueSA</u>		
ACAA	GlueSAAccessControlBaseRule		
COCC	GheSE		
Enabling Grids	GlueSEAccessProtocol		
FOR E-SCIENCE	GlueSEAccessProtocolSupportedSec		
	CiheSL		

ApplMONIT

Query this table

Name	ID	Туре
ID	530	VARCHAR(20)
JobDone	531	VARCHAR(10)
Param	532	INT
HostCE	533	VARCHAR(80)
Owner	534	VARCHAR(20)
MeasurementDate	535	DATE
MeasurementTime	536	TIME

Query this table

INFSO-RI-508833

ege	R-G Enabling Grids for E-sciencE	MA Browser as Consumer
R-GMA Browser	All tables GLUE Info Providers Network Monitoring Service Discovery CMS ApplMONIT GAMIAppStart GhueBatchJob GhueBatchJob GhueBatchSystem GhueCE GhueCEsEBind GhueHostLocalFileSystem GhueHostNetworkAdapter GhueHostNetworkAdapter GhueHostRemoteFileSystem GhueHostRole GhueHostRole GhueHostRole	SELECT D JobDone Param HostCE Owner FROM AppIMONIT WHERE Description of table Type of query: • History • Latest • Continuous • Continuous & old Queries wait for 5 seconds • Use Mediator • Select producers you want to query: History Latest
CGCC Enabling Grids For E-sciencE	GlueSE GlueSEAccessProtocol GlueSEAccessProtocolSupportedSec GlueSL GlueService	Continuous Query

INFSO-RI-508833



API available for Java, C, C++ and Python

Users may by-pass API if they wish, but API is the easiest way to use R-GMA services



Producer Application (in Java)(1)

Enabling Grids for E-sciencE

Producer Properties

egee

rype. Frinary
Storage type: Database
Termination Interval: 300 (seconds)
Predicate: Where
Query type: HISTORY
Latest Retention Period: 300 (seconds)
History Retention Period: 300 (seconds)

```
ProducerProperties producerProps = null;
if (producerType.equals("CONTINUOUS"))
{ producerProps = new ProducerProperties(Storage.MEMORY, 0); }
else if (producerType.equals("LATEST"))
{ producerProps = new ProducerProperties(Storage.DATABASE, ProducerProperties.LATEST); }
else if (producerType.equals("HISTORY"))
{ producerProps = new ProducerProperties(Storage.DATABASE, ProducerProperties.HISTORY); }
else
{ System.err.println("Invalid producer type (" + producerType + ").");
 System.exit(1);
}
```

CALCENTION OF CONTROL OF CONTRON

```
. . . . . . .
PrimaryProducer pp = null;
ResourceEndpoint endpoint = null;
Try
  ProducerFactory pf = new ProducerFactoryStub();
{
   TimeInterval ti = new TimeInterval(terminationInterval, Units.SECONDS);
   pp = pf.createPrimaryProducer(ti, producerProps, null);
   endpoint = pp.getResourceEndpoint();
                                                                         Schema
   String predicate = "WHERE ID = '" + Id + "'";
                                                                         Service
   pp.declareTable(tableName,
                   predicate,
                                                                        Registry
                   new TimeInterval(historyRP, Units.SECONDS),
                                                                         Service
                   new TimeInterval(latestRP, Units.SECONDS));
                                                                           register
                                                                           producer
```





```
pp.insert(insert);
  . . . . . . . . .
                                                           Schema
pp.close();
                                                           Service
                                                          Registry
                                                           Service
                                                             register
                                                             producer
                                          declare
                                                      Producer
                          Producer
                                     API
                                                       Service
                         application
                                          insert
```

GGCC JDL with User Producer Application

```
Type = "Job";
JobType = "Normal";
Executable="startPP.sh";
Arguments = "100 HISTORY Valeria_Ardizzone";
StdOutput="stdout.log";
StdError="stderr.log";
InputSandbox={"startPP.sh","pp.class"};
OutputSandbox={"stdout.log","stderr.log"};
. . . . . . .
```

```
]
```

eGee

JDL Submission from GENIUS

Enabling Grids for E-sciencE



Single Job

🖿 up

- Job Submission
- Job Queue
- Job Data
- Clean Job Queues
- Close Interactive Job Session

Single Job

🖿 up

- Job Submission
- Job Queue
- Job Data
- Clean Job Queues
- Close Interactive Job Session

powered by EnginFrame 3.2 compliant with LCG-2 GRID.IT gLite-1





Enabling Grids for E-sciencE

Grid Enabled web eNvironment for site Independent User job Submission

RB: gilda-glite	VO: gilda	Catalog: GILDA	Your Data				
Now you may also choose a	specific Computing Element	for your job.					
JDL File Selected /h	nome/vardizzo/UseCasesRGI	MA/RGMAPP_prop.jdl 💟					
Specify the CE Resource	et the GILDA-GLITE Resource	Broker choose 🔽					
Submit Job							
RB: gilda-glite	VO: gilda	Catalog: GILDA	Your Data I				
Selected Virtual Organisation name (from proxy certificate extension): gilda Connecting to host glite-rb.ct.infn.it, port 7772 Logging to host glite-rb.ct.infn.it, port 9002							
The job has been succ Use glite-job-status	essfully submitted to command to check job c	ubmit Success =========== the Network Server. urrent status. Your job id	entifier is:				
- https://glite-rb.ct	.infn.it:9000/ChED7MwR	8kbK9FzhKUeQ x A					
The job identifier ha /home/vardizzo/.geniu	s been saved in the fo s/.tmp_submittedjob_va	llowing file: rdizzo					

INFSO-RI-508833

CGCC User Producer in R-GMA Browser

Enabling Grids for E-sciencE

	All tables	
R-GMA	GLUE Info Providers	SELECT ID
Browser	Network Monitoring	Param
	Service Discovery	HostCE
	CMS	Owner V
Home		FROM ApplMONIT
Predefined:	GAMIAppStart	WHERE
GlueServices	ChaPatablab	
<u>GlueSite</u>	ChieBatchOvene	
<u>RGMALogs</u>	ChueDatchQueue	
Table Sets	ChueDatchSystem	Query
	ChueCE	Description of table
	GlueCEAccessControlBaseRule	·
	GneCESEBind	Type of query:
	GheChister	○ History ③ Latest ○ Continuous ○ Continuous & old
	GhueHost	Oueries wait for 5 seconds
	GlueHostLocalFileSystem	
	GlueHostNetworkAdapter	
	GlueHostPoolAccount	• Use Mediator
	<u>GlueHostProcess</u>	Select producers you want to guard
	<u>GlueHostRemoteFileSystem</u>	Select producers you want to query.
	<u>GlueHostRole</u>	History
	GlueSA	https://rgmasrv.ct.infn.it:8443/R-GMA/DBProducerServlet 1452780494
	<u>GlueSAAccessControlBaseRule</u>	
	GhueSE	
ICCCCC	GlueSEAccessProtocol	Latest
Enabling Grids	GlueSEAccessProtocolSupportedSec	
For E-sciencE	GlueSL	Continuous
	<u>GlueService</u>	https://ramastry.ct.infn.it.8443/R_GMA/DBDrochucarSatzlat 1452780404
		intps/rgmasiv.cc.iiiiii.it.o++5/iC-OiviA/DBF10duCciScivict 1452760494

INFSO-RI-508833

C G	BG	Query from R-GMA Brows	ser
R-GMA Browser	All tables GLUE Info Provide Network Monitorin Service Discovery Service Discovery AppIMONIT GAMIAppStart GhueBatchJob GhueBatchQueue GhueBatchQueue GhueBatchQueue GhueBatchSystem GhueCE GhueCEAccessCont GhueCESEBind GhueCESEBind GhueChster GhueHostLocalFileS GhueHostNetworkA GhueHostPoolAccou GhueHostProcess GhueHostProcess GhueHostRole GhueSA GhueSAAccessCont	SELECT Dobone Param HostCE Owner FROM AppIMONIT WHERE Param > 30 and Param < 40 Query Description of table Type of query: • History ● Latest ● Continuous ⑧ Continuous & old Queries wait for 5 seconds • Use Mediator • Select producers you want to query: History Intps://rgmasrv.ct.infn.it.8443/R-GMA/DBProducerServlet 1452780494	
Enabling Grids For E-sciencE	GheSEAccessProte GheSEAccessProte GheSL GheService	Latest Continuous https://rgmasrv.ct.infn.it:8443/R-GMA/DBProducerServlet 1452780494	

INFSO-RI-508833



Query Results

Ouery: SELECT ID, JobDone, Param, HostCE, Owner, MeasurementDate, MeasurementTime FROM All tables AppIMONIT WHERE Param > 30 and Param < 40 **R-GMA** GLUE Info Provide Browser Network Monitorin ID JobDone Param HostCE Owner MeasurementDate MeasurementTime Service Discovery PP 14659 31% 31 Valeria Ardizzone 2006-01-06 12:17:25 grid036.ct.infn.it $\mathbf{\Sigma}$ Home PP 14659 32% 32 grid036.ct.infn.it Valeria Ardizzone 2006-01-06 12:17:25 ApplMONIT Predefined: PP 14659 33% Valeria Ardizzone 12:17:25 33 grid036.ct.infn.it 2006-01-06 GAMIAppStart **GlueServices** GlueBatchJob PP 14659 34% 34 grid036.ct.infn.it Valeria Ardizzone 2006-01-06 12:17:25 GlueSite GlueBatchOueue PP 14659 35% grid036.ct.infn.it 35 Valeria Ardizzone 2006-01-06 12:17:25 **RGMALogs** GlueBatchSystem PP 14659 36% 36 grid036.ct.infn.it Valeria Ardizzone 12:17:25 Table Sets 2006-01-06 GlueCE 37 GlueCEAccessCont PP 14659 37% Valeria Ardizzone 2006-01-06 12:17:25 grid036.ct.infn.it GlueCESEBind PP 14659 38% grid036.ct.infn.it Valeria Ardizzone 2006-01-06 12:17:25 38 GlueCluster PP 14659 39% 39 2006-01-06 12:17:25 grid036.ct.infn.it Valeria Ardizzone GlueHost GlueHostLocalFileS

eGee

lstituto Nazionale di Fisica Nucleare

INFN

Job Output in GENIUS

Enabling Grids for E-sciencE







Grid Enabled web eNvironment for site Independent User job Submission

Single Job									_	_		
⊒ up	la-glite	VO: gilda	(Catalog: G	ILDA		Your D	ata	Logout			
Job Submission	Job Queue											
Job Ououo	ID	J	DL Name		Last		Des	tination		Status	Exit	Action
Job Queue					Updat	e					Code	
Job Data												~ .
Clean Job Queues	-VOE-bVU-O-	the survey for an alterna of the second	DOM DO	MADD	Jan 6) 00 -114		110/51-5	in Carita	Dene		Get
Close Interactive Job	<u>5K9FZnKUeQX</u> A	/nome/vardizzo/UseCa	asesKGMA/KG	MAPP pro	2006 2006	J8 gute-ce	ct.mm.nt2	119/blan-pbs-	-infinite	Done	0	Output
Session					CET							
nowered by					021							
EnginFrame 3.2												
compliant with												
LCG-2 GRID.IT												
<u>gLite-1</u>												
Single Teb												
Single Job	*******	######### STAR	T ######	******	*****							
- up	COMM_EXE	= pp.class										
Job Submission	/usr/java	/j2sdk1.4.2_08//	bin/java -	classpat	h :/opt/	/glite/s	share/ja	va/glite-	rgma-api	-java.	jar:/o	opt/gl
	INSERT IN	TO ApplMONIT (ID	, JobDone,	Param,	HostCE,	Owner)	VALUES	('PP_14659	9','1%',	'1','g:	rid03(6.ct.i
Job Queue	INSERT IN	TO ApplMONIT (ID	, JobDone,	Param,	HostCE,	Owner)	VALUES	('PP_14659	9','2%',	'2','g:	rid03(6.ct.i
Job Data	INSERT IN	TO ApplMONIT (ID	, JobDone,	Param,	HostCE,	Owner)	VALUES	('PP_14659	9','3%',	'3','g:	rid03(6.ct.i
Clean Job Onenes	INSERT IN	TO ApplMONIT (ID	, JobDone,	Param,	HostCE,	Owner)	VALUES	('PP_14659	9','4%',	'4','g:	rid03(6.ct.i
Clean 500 Queues	INSERT IN	TO ApplMONIT (ID	, JobDone,	Param,	HostCE,	Owner)	VALUES	('PP_14659	9','5%',	'5','g:	rid03(6.ct.i
Close Interactive Job	INSERT IN	TO ApplMONIT (ID	, JobDone,	Param,	HostCE,	Owner)	VALUES	('PP_14659	9','6%',	'6','g:	rid03(6.ct.i
Session	INSERT IN	TO ApplMONIT (ID	, JobDone,	Param,	HostCE,	Owner)	VALUES	('PP_14659	9','7%',	'7','g:	rid03	6.ct.i
	INSERT IN	TO ApplMONIT (ID	, JobDone,	Param,	HostCE,	Owner)	VALUES	('PP_14659	9','8%',	'8','g:	rid03	6.ct.i



- Consumer users: who requests information.
- Producer users: who provides information.
- Site administrators: who runs R-GMA services.
- Virtual Organizations: who "owns" the schema and registry.



- Mutual Autentication: guaranteeing who is at each end of an exchange of messages.
- Encryption: using an encrypted transport protocol (HTTPS).
- Authorization: implicit or explicit.



More information

- R-GMA overview page.
 - http://www.r-gma.org/
- **R-GMA documentation in EGEE**
 - <u>http://hepunx.rl.ac.uk/egee/jra1-uk/</u>
- R-GMA in GILDA
 - http://hepunx.rl.ac.uk/egee/jra1-uk/