



Enabling Grids for E-scienceE

Introduction to R-GMA: Relational Grid Monitoring Architecture

www.eu-egee.org



Slides are taken/derived from

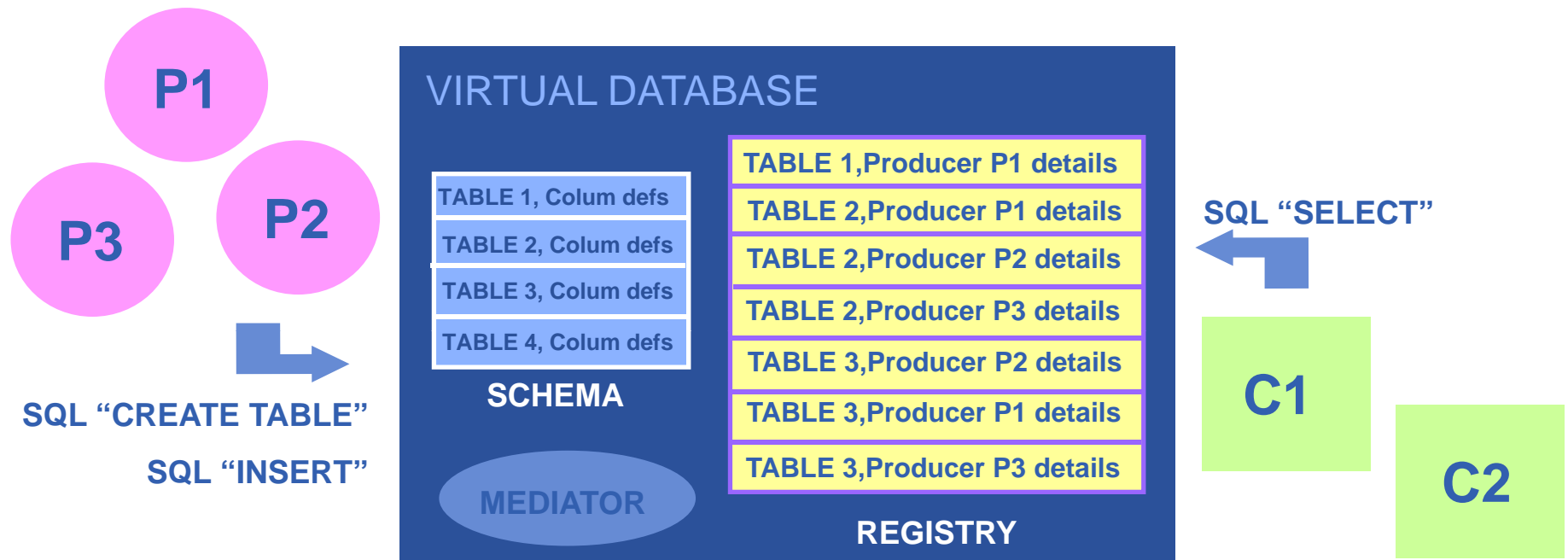
- **the GILDA team**
- **Steve Fisher (RAL, UK) and the R-GMA team**

- **Uniform method to access and publish both information and monitoring data.**
- **From a user's perspective, an R-GMA installation currently appears similar to a single relational database.**
- **GMA (Grid Monitoring Architecture) was developed by the GGF**
- **R-GMA (Relational GMA) was created:**
 - To simplify use of GMA (servers “know” about registries, not the client software)
 - To give a relational view

- **Relational Grid Monitoring Architecture (R-GMA)**
 - Developed as part of the EuropeanDataGrid Project (EDG)
 - Now as part of the EGEE project.
 - Evolution from the Grid Monitoring Architecture (GMA)
- **Uses a relational data model.**
 - Data are viewed as a table.
 - Data structure defined by the columns.
 - Each entry is a row (tuple).
 - Queried using Structured Query Language (SQL).

name	ID	birth	Group
Tom	4	1977-08-20	HR

```
SELECT * FROM people WHERE group='HR'
```



There is no central repository!!! There is only a “*Virtual Database*”.

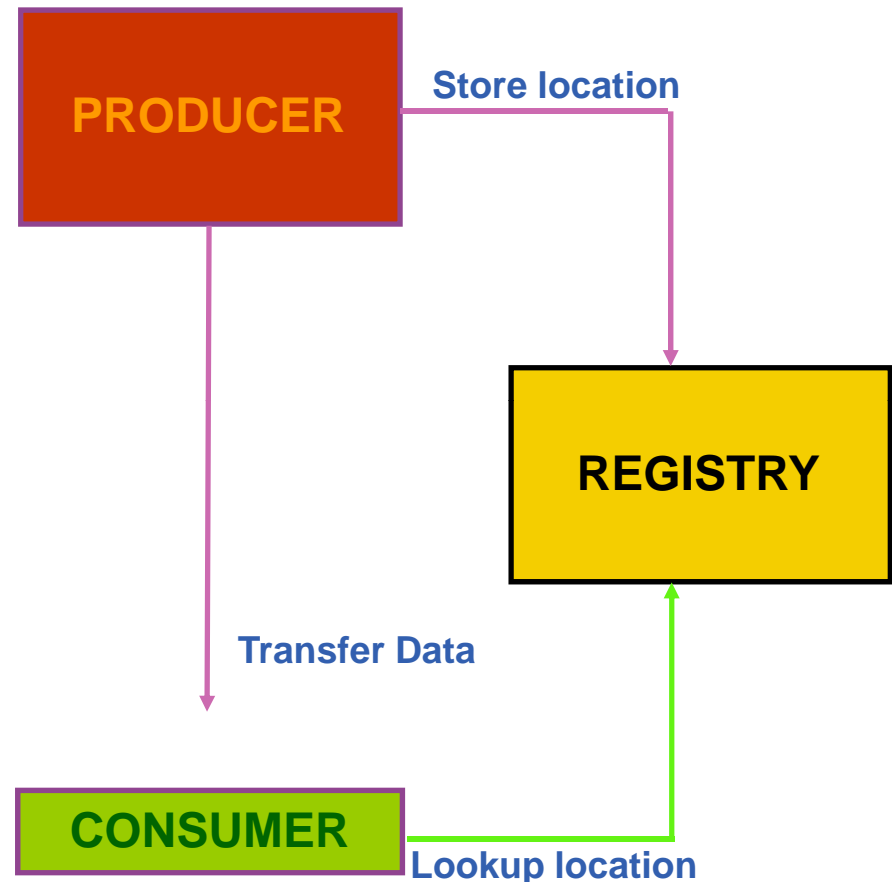
Schema is a list of table definitions: additional tables/schema can be defined by applications

Registry is a list of data producers with all its details.

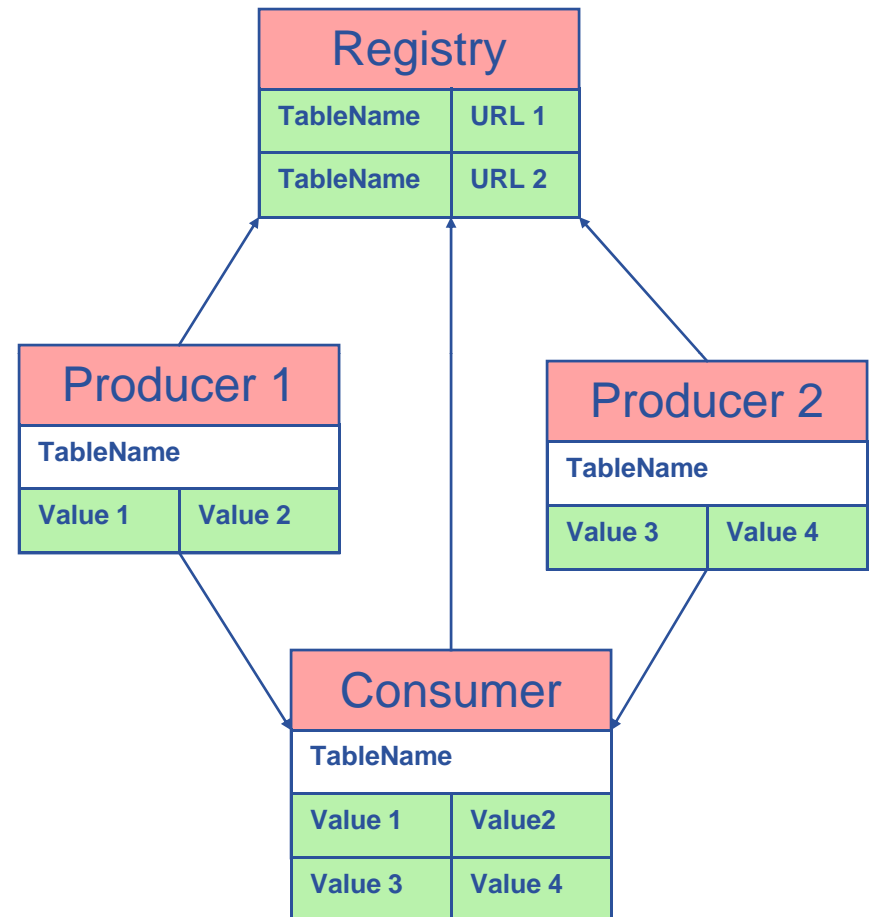
Producers publish data.

Consumers read data published.

- 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.

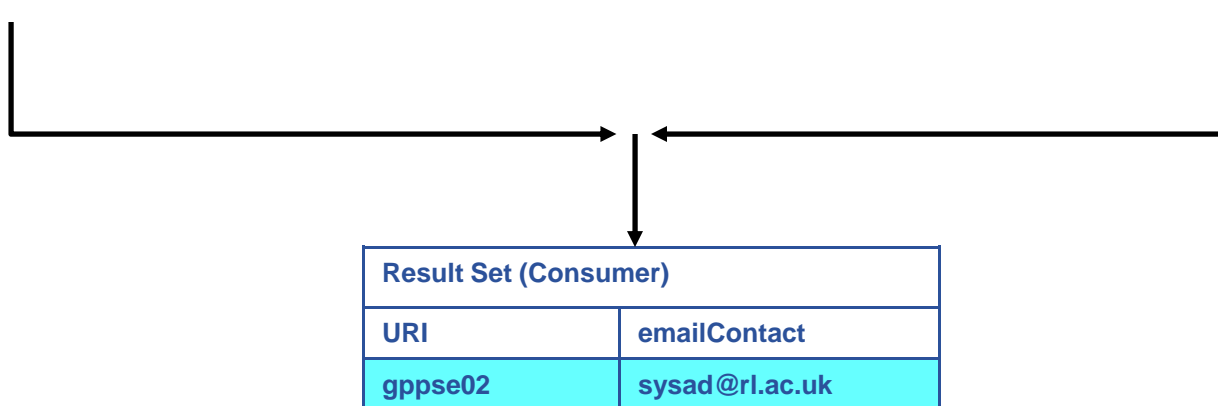


- The Consumer interrogates the Registry to identify all Producers that could satisfy the query.
- Consumer connects to the Producers.
- Producers send the tuples to the Consumer.
- The Consumer will merge these tuples to form one result set.

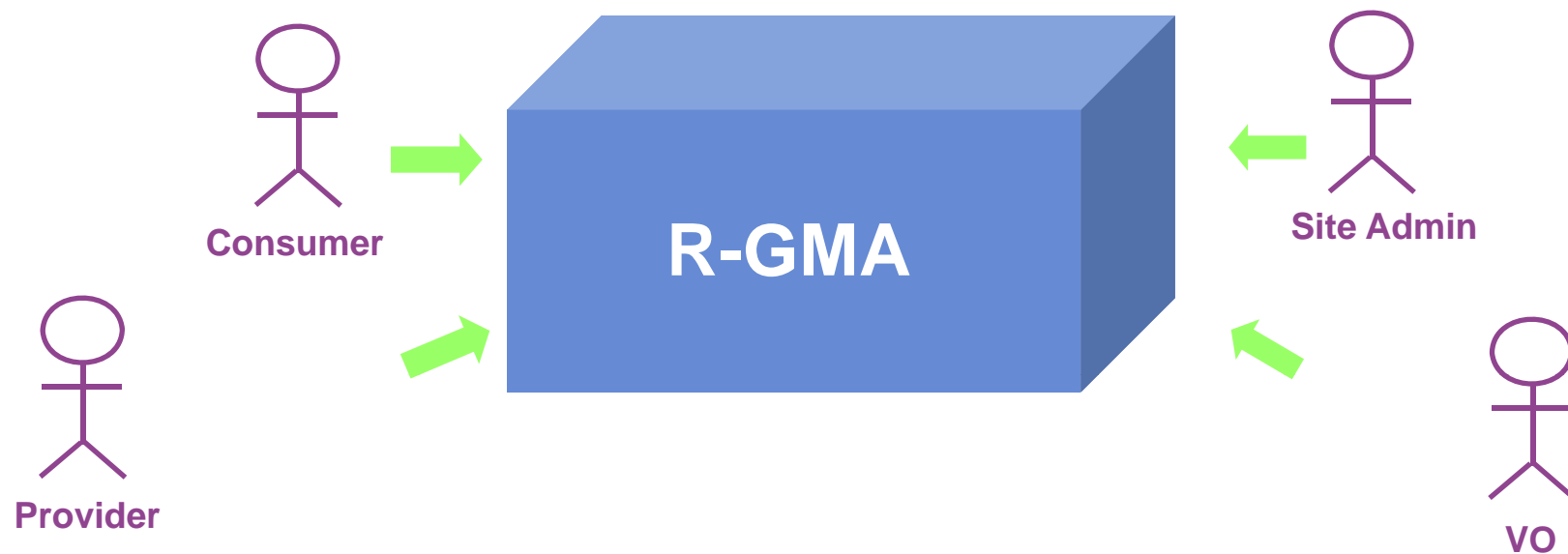


Service				
URI	VO	type	emailContact	site
gppse01	alice	SE	sysad@rl.ac.uk	RAL
gppse01	atlas	SE	sysad@rl.ac.uk	RAL
gppse02	cms	SE	sysad@rl.ac.uk	RAL
lxshare0404	alice	SE	sysad@cern.ch	CERN
lxshare0404	atlas	SE	sysad@cern.ch	CERN

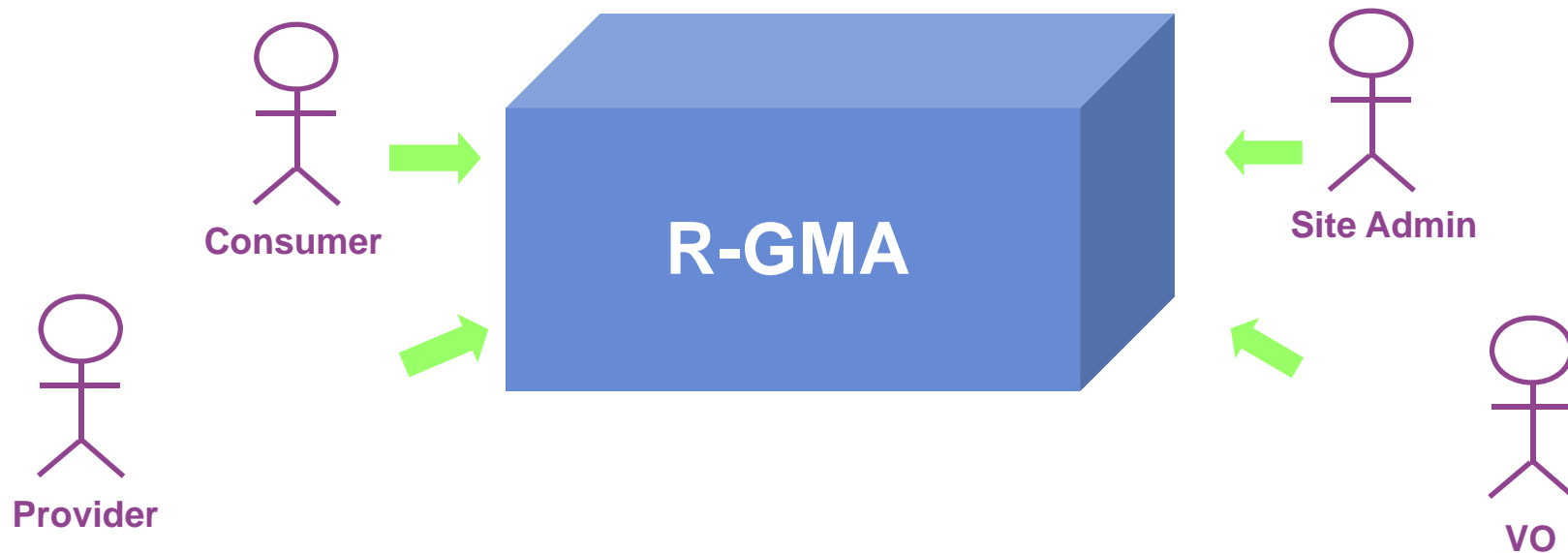
ServiceStatus				
URI	VO	type	up	status
gppse01	alice	SE	y	SE is running
gppse01	atlas	SE	y	SE is running
gppse02	cms	SE	n	SE ERROR 101
lxshare0404	alice	SE	y	SE is running
lxshare0404	atlas	SE	y	SE is running



```
SELECT Service.URI Service.emailContact FROM Service S, ServiceStatus SS
WHERE (S.URI= SS.URI and SS.up='n')
```

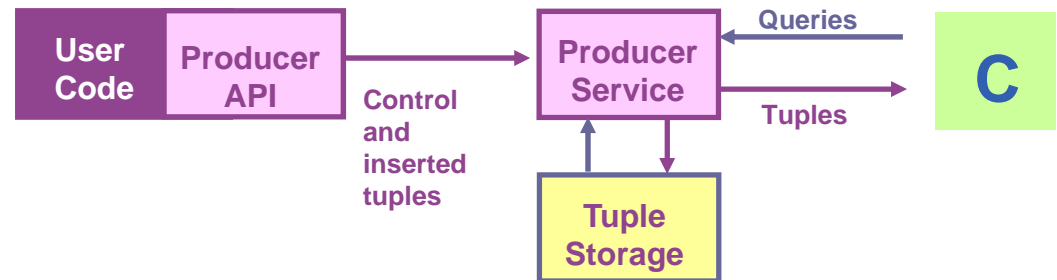
- **Consumer users:** who request information.
- **Producer users:** who provide information.
- **Site administrators:** who run R-GMA services.
- **Virtual Organizations:** who “own” the schema and registry.



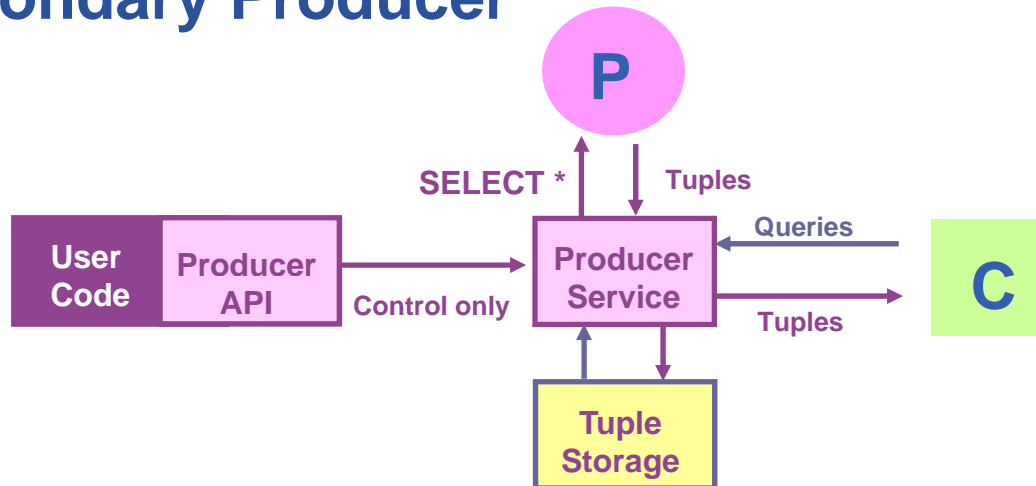
- **Mutual Authentication:** guaranteeing who is at each end of an exchange of messages.
- **Encryption:** using an encrypted transport protocol (HTTPS).
- **Authorization:** implicit or explicit.

- **Producer and Consumer Services are typically on a one per site basis**
- **Centralized Registry and Schema.**
- **The Registry and Schema may be replicated, to avoid a single point of failure**
 - ... when you use RGMA CLI you will see which are being used

- Primary Producer

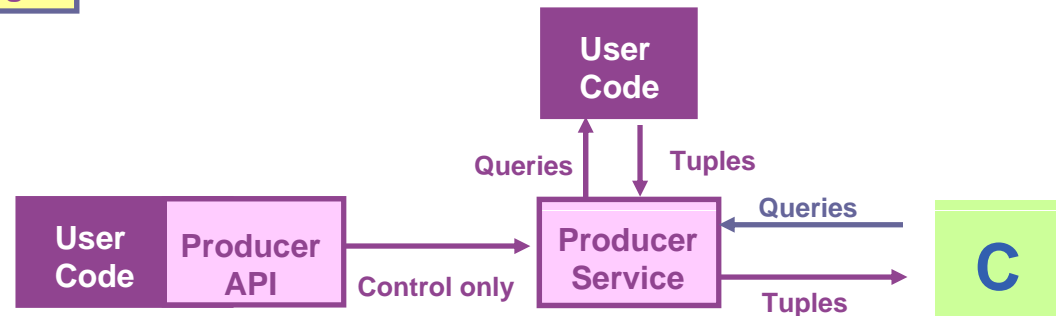


- Secondary Producer



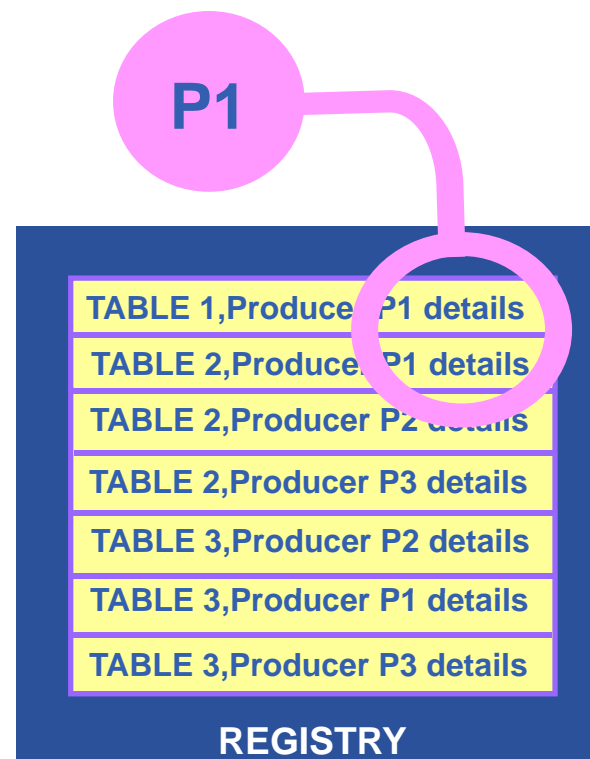
- On-Demand Producer

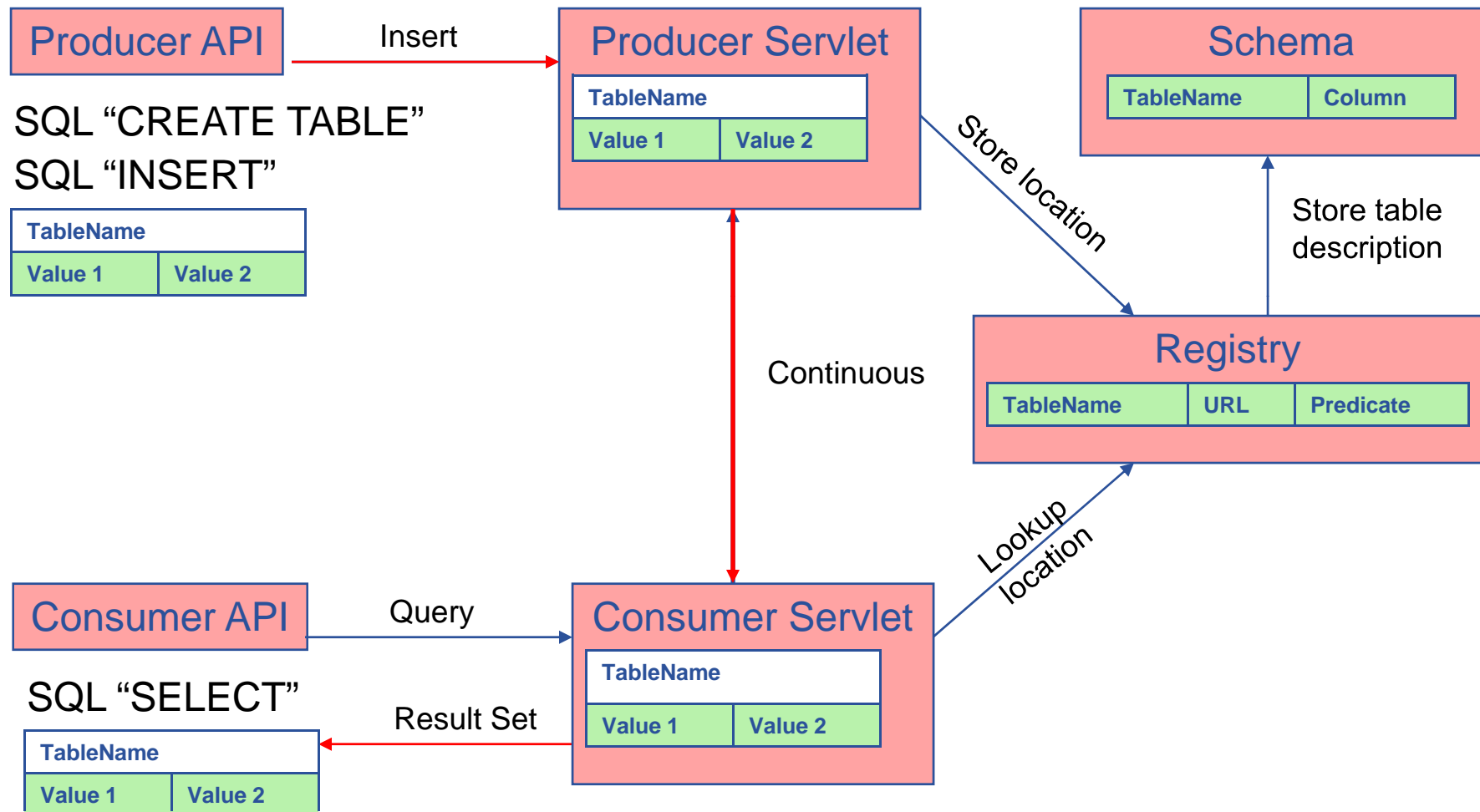
- No internal storage
- Queries passed to user code



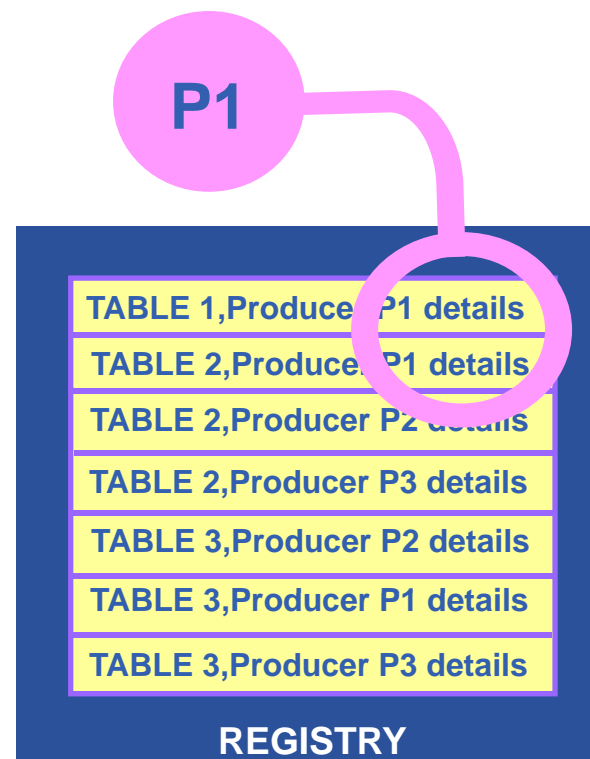
Continuous

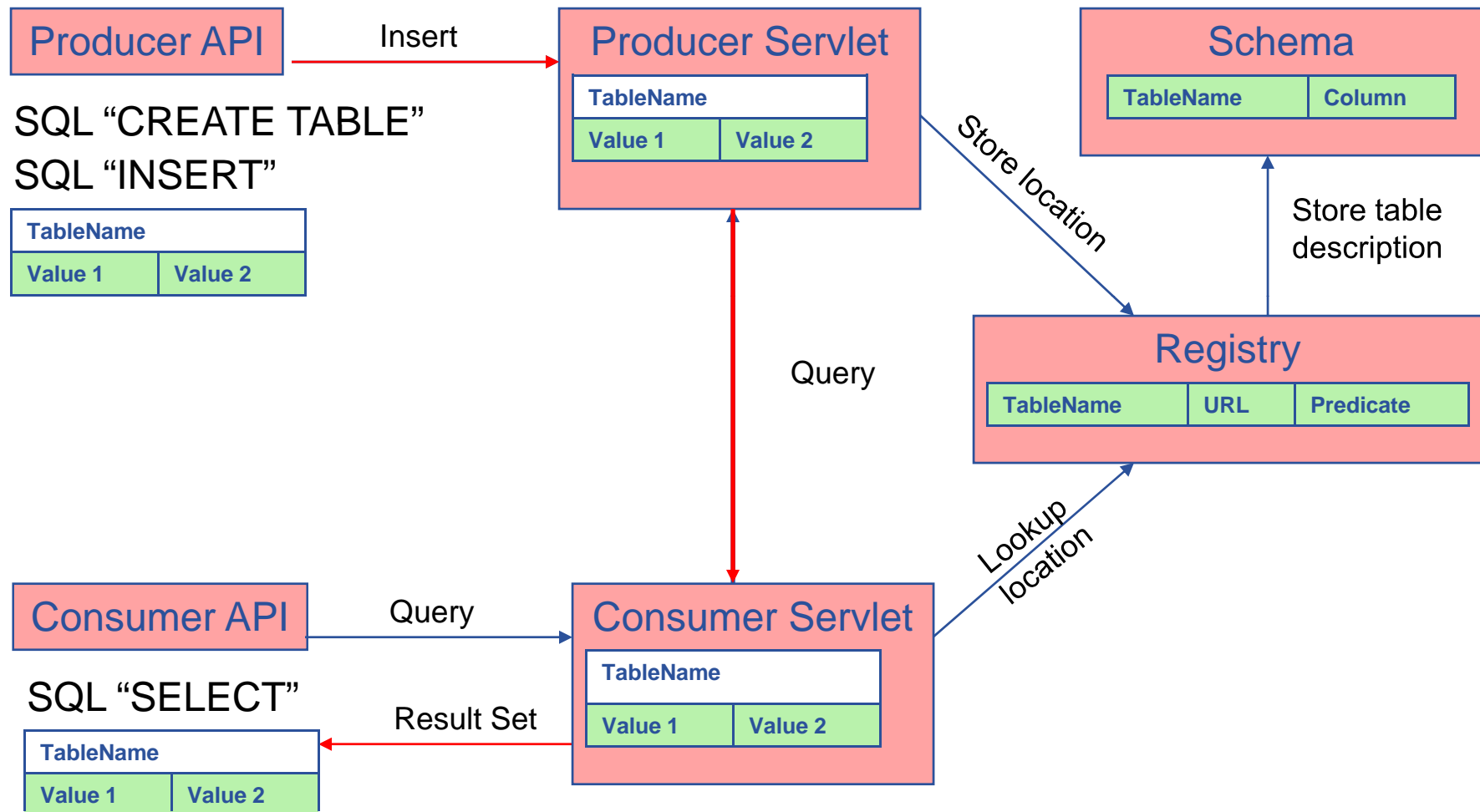
- Latest
- History
- Static



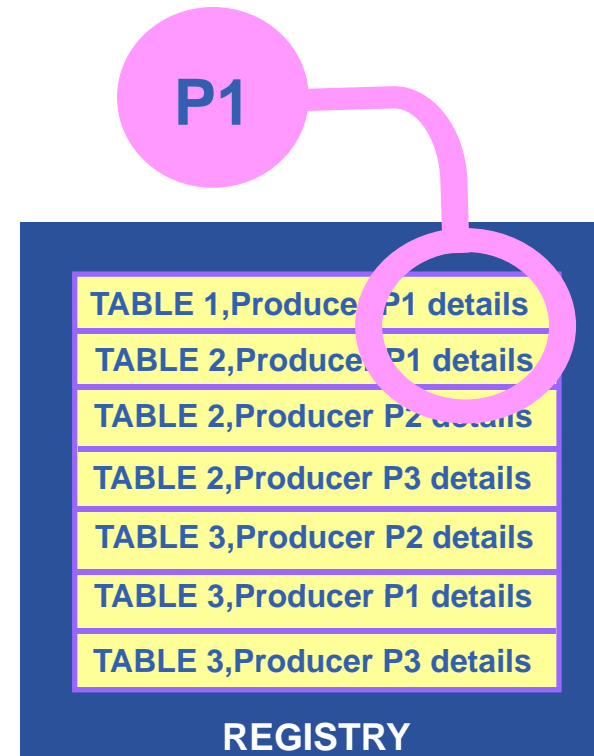
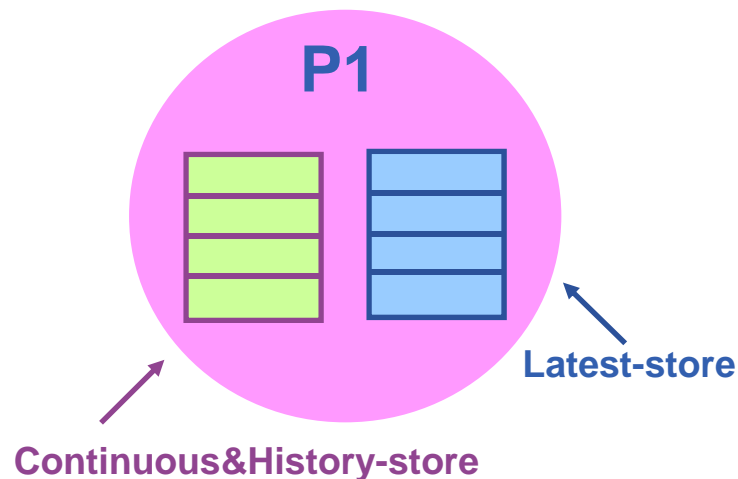


- Continuous
- Latest
- History
- Static





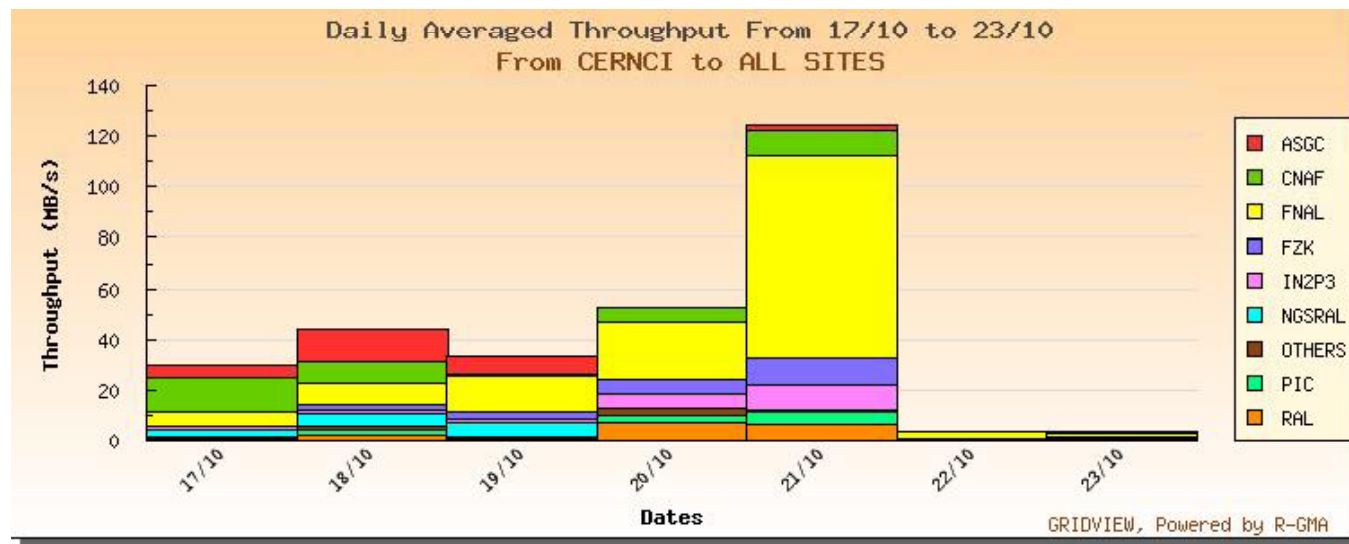
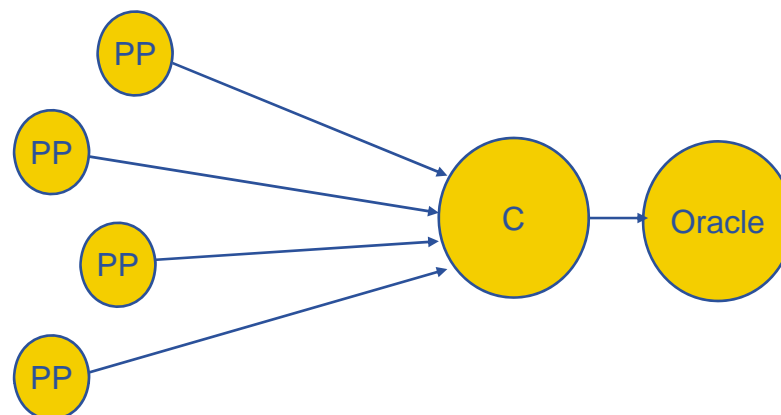
- Continuous
- Latest
- History
- Static



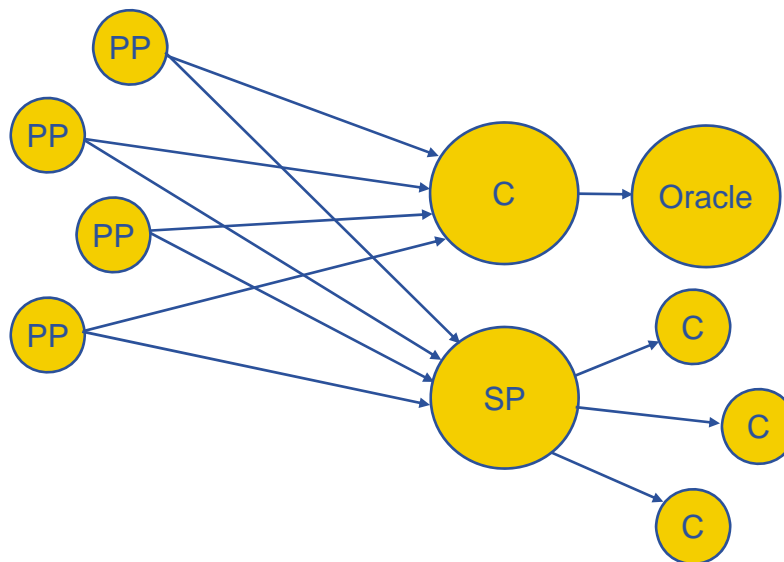
Latest Retention Period
History Retention Period

GridFTP Monitoring (gridView)

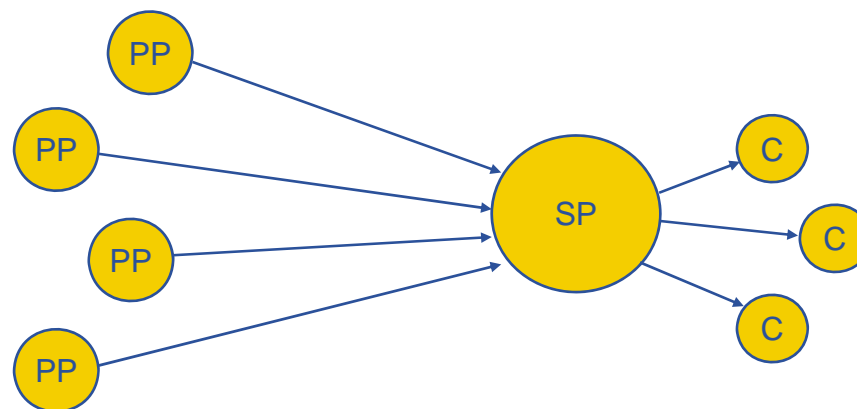
- SA1 have written script to “tail” FTP logs and publish via PP on gridFTP server nodes
- Continuous query pulls all the data to a central location and writes to an Oracle database for analysis
- Used for Service Challenge 3
- <http://gridview.cern.ch/GRIDVIEW/>



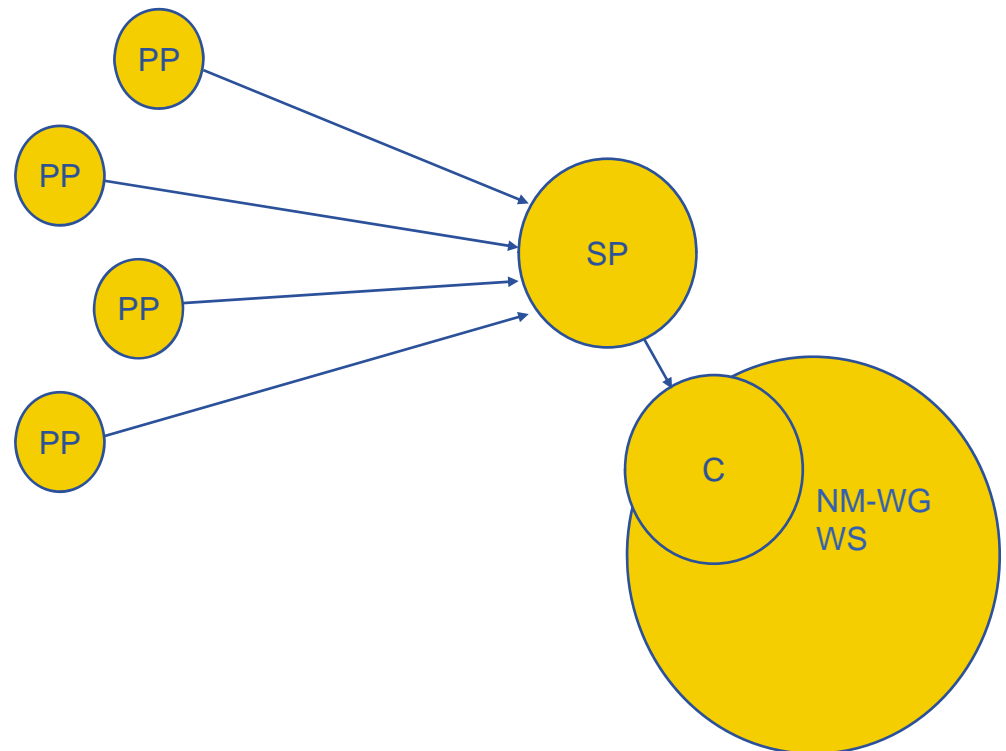
- Reads L&B logs on the resource broker nodes.
- Publishes data on state of jobs
- A database secondary producer is used to aggregate the data as well as a gridView consumer.
- CMS dashboard
 - <http://lxarda09.cern.ch/dashboard/request.py/jobsummary?>



- On the WNs, the Job Wrapper (if enabled by JDL) periodically publishes information about the state of the process running the job and its environment.
- A database secondary producer is used to aggregate the data.
 - <https://rgma13.pp.rl.ac.uk:8443/R-GMABrowser/Browser.do/queryTable?selectQueryType=latest&duration=20&tableName=JobMonitor>



- **Network performance data important:**
 - to detect and resolve network problems.
 - to intelligently schedule jobs based on network load and reliability.
- **active measurements between end-sites, using tools such as**
 - iperf,
 - udpmon
 - ping.



• <https://egee.epcc.ed.ac.uk:28443/npm-dt/query.jsp>

NPM Diagnostic Tool

NPM Diagnostic Tool - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

Query

Time Range Set

Start: 2005-10-22 14:30:00 End: 2005-10-24 14:30:00 Focus: 2005-10-24 14:30:00

End: 2005-10-24 14:30:00 Period: 2 Days Tolerance [s]: minus 172800 plus 0

Max Results

Maximum number of results unbounded

Test Path

Source: Destination:

doe2-pos-dc
doe2-pos-dc.es.net
e2emonit.mrs.grid.cn
e2emonit.nesc.ed.ac.uk

e2emonit.mrs.grid.cn

e2emonit.nesc.ed.ac.uk => e2emonit.mrs.grid.cn

Choose New Source Add Path Delete Path Clear All Paths

Find Data For This Query

Metric Set

Metric: No Data Packet Size: byte Packets: Packet Gap: byte

Statistic Set

Statistic: No Statistics Available Interval: Minutes

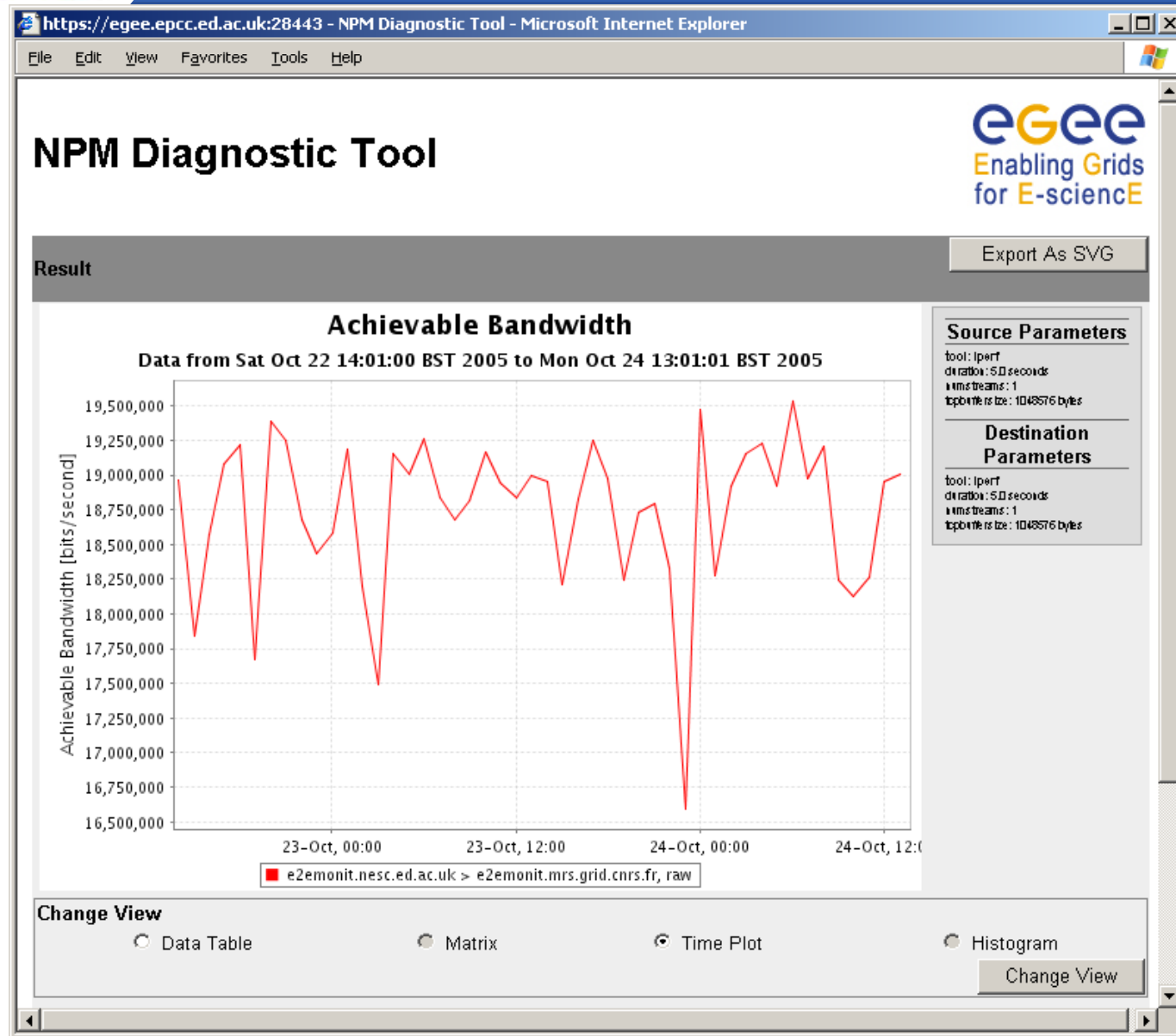
View As

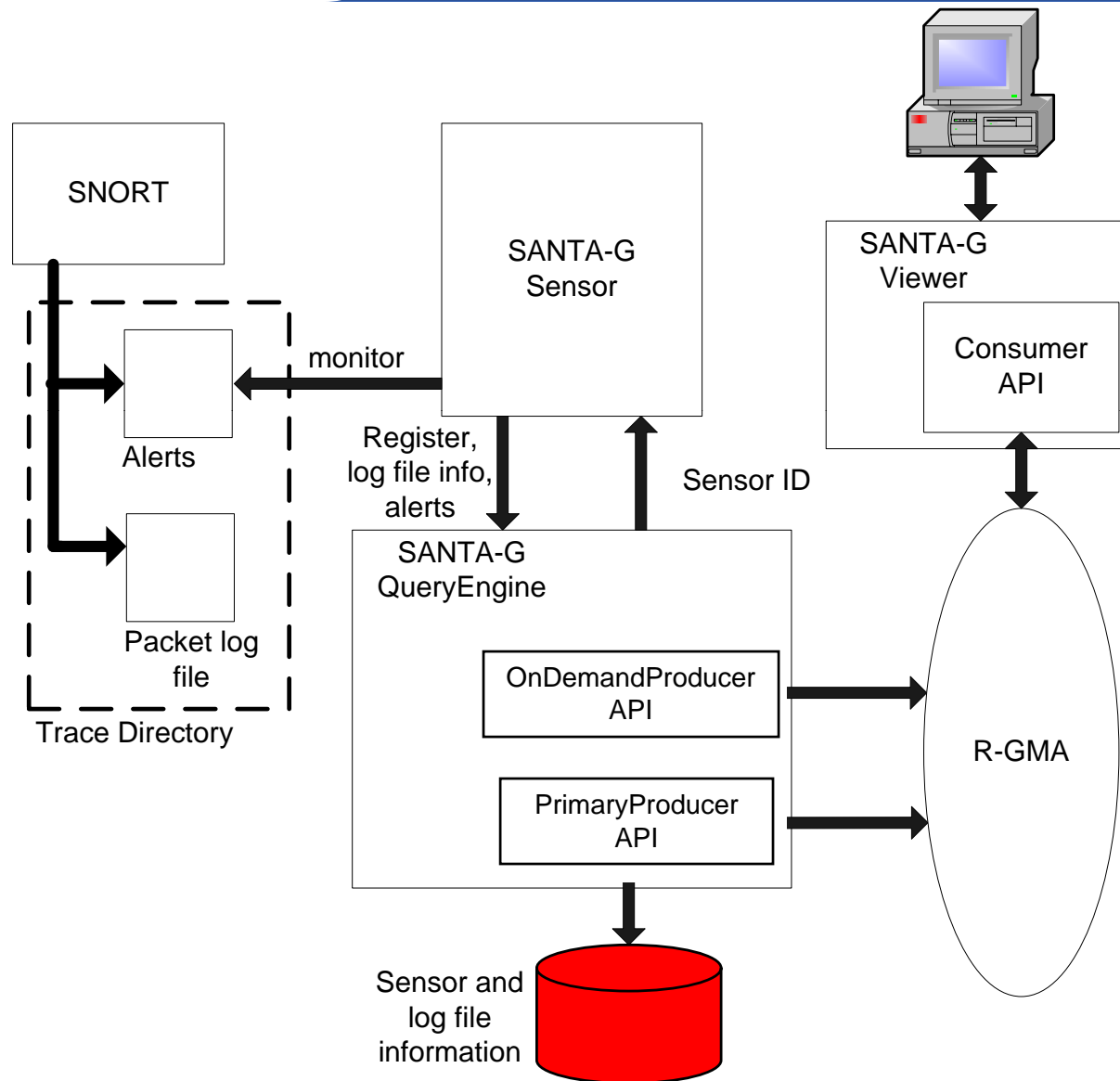
☐ Data Table ☐ Matrix ☒ Time Plot ☐ Histogram

Submit Query

[Create a new query.](#)

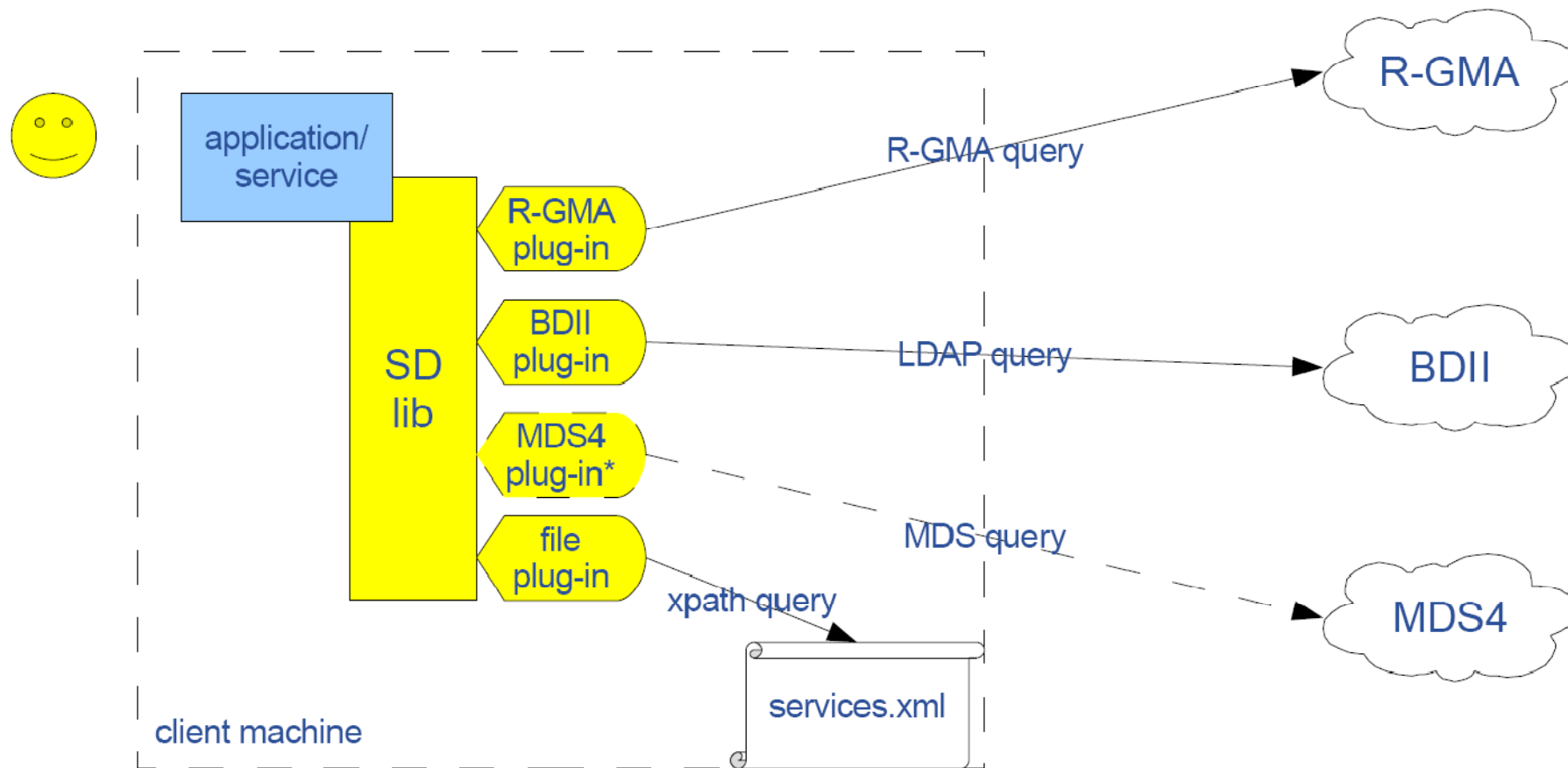
You are logged in as: CN=alstair.phipps, L=NESC, OU=Edinburgh, O=e Science, C=UK
NPM Diagnostic Tool (1.1) © Members of the EGEE Collaboration 2005
[Email the DT administrator](#) | [Download the DT User Guide](#)





The Grid intrusion detection work is now within the Interactive European Grid (<http://www.interactive-grid.eu>) project, as part of the JRA workpackage, and is known as Active Security (<http://www.grid.ie/i2g>)

- **Questions to answer:**
 - “I am at CERN, in 'dteam' VO. Where is a MyProxy server?”
 - `glite-sd-query -t myproxy -s CERN-PROD`
- **Service Discovery offers:**
 - client API (library) to hide the differences
 - plug-in architecture to simplify dependencies
 - uses the subset of Glue schema as data model
 - simple API, no complex queries
 - CLI for other tools and testing
- **Plug-ins for:**
 - BDII
 - R-GMA
 - MDS4 (not yet)
 - File (only for testing)



TCD: Trinity College Dublin

- **gridFS: a grid filesystem**
- **InfoGrid: a grid using an information model**
- **Keith Rochford's work on grid service monitoring**
- **Adaptive eLearning: R-GMA is the first course**
- **Shared memory for grids (SMG)**

- **APIs exist in Java, C, C++, Python.**
 - For clients (servlets contacted behind the scenes)
- **They include methods for...**
 - Creating consumers
 - Creating primary and secondary producers
 - Setting type of queries, type of produces, retention periods, time outs...
 - Retrieving tuples, inserting data
 - ...
- **You can create your own Producer or Consumer.**

- We will use a client that gives command-line interfaces to both consumers and producers
- We will explore the tables on the R-GMA service provided on GILDA
- Use a table that is set up for training purposes to produce and consume data

Now please follow the “more information” link

R-GMA practical html page

- The command line tool can be used in batch mode in three ways:
 - **rgma -c <command>**
Executes <command> and exits.
The -c option may be specified more than once.
 - **rgma -f <file>**
Executes commands in <file> sequentially then exits.
Each line should contain one command.
 - Embedded in a shell script

R-GMA Browser

requires certificate in browser

Table description

R-GMA Browser Home Page - Mozilla Firefox

File Modifica Visualizza Vai Segnalibri Strumenti ?

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

R-GMA Browser

[Home](#)

Predefined:

- [GlueServices](#)
- [GlueSite](#)
- [RGMALogs](#)

Table Sets

[All tables](#)

- [GLUE Info Providers](#)
- [Network Monitoring](#)
- [Service Discovery](#)
- [CMS](#)
- [JobMonitor](#)
- [JobStatusRaw](#)
- [NetworkCE](#)
- [NetworkFileTransferThroughput](#)
- [NetworkICMPPacketLoss](#)
- [NetworkOneWayIPDV](#)
- [NetworkRTT](#)
- [NetworkSE](#)
- [NetworkTCPThroughput](#)
- [NetworkUDPPacketLoss](#)
- [NetworkUDPThroughput](#)
- [RGMALogs](#)
- [RLSService](#)
- [Service](#)
- [ServiceStatus](#)
- [Site](#)
- [SiteInfo](#)
- [T1](#)
- [UseCasePP](#)
- [bossJobExOutMessage](#)
- [bossJobExOutStandardInfo](#)
- [grmHeader](#)
- [grmTrace](#)
- [userCasePP](#)
- [userTable](#)

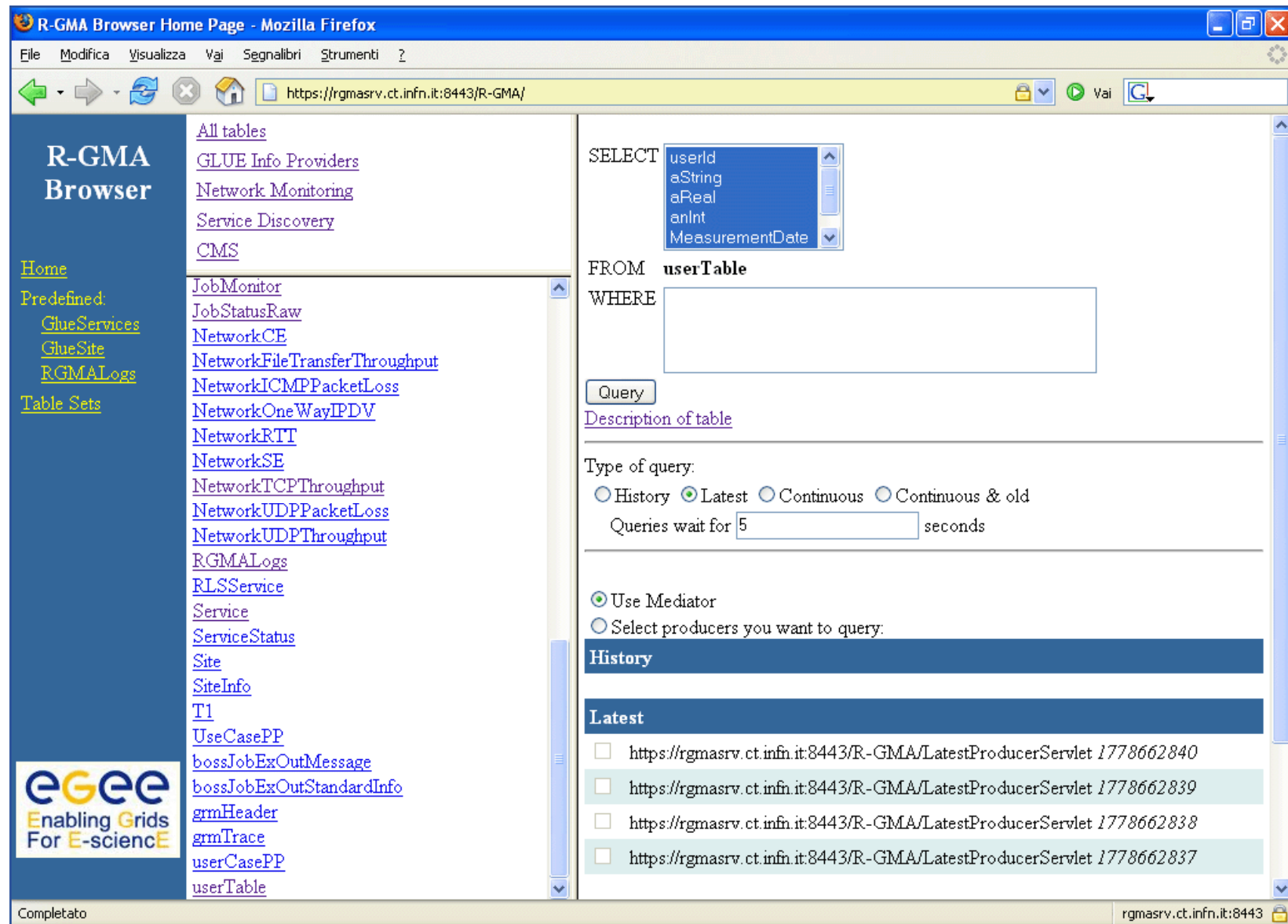
userTable

[Query this table](#)

Name	ID	Type
userId	374	VARCHAR(255)
aString	375	VARCHAR(255)
aReal	376	REAL
anInt	377	INTEGER
MeasurementDate	378	DATE
MeasurementTime	379	TIME

[Query this table](#)

Completato rgmasrv.ct.infn.it:8443

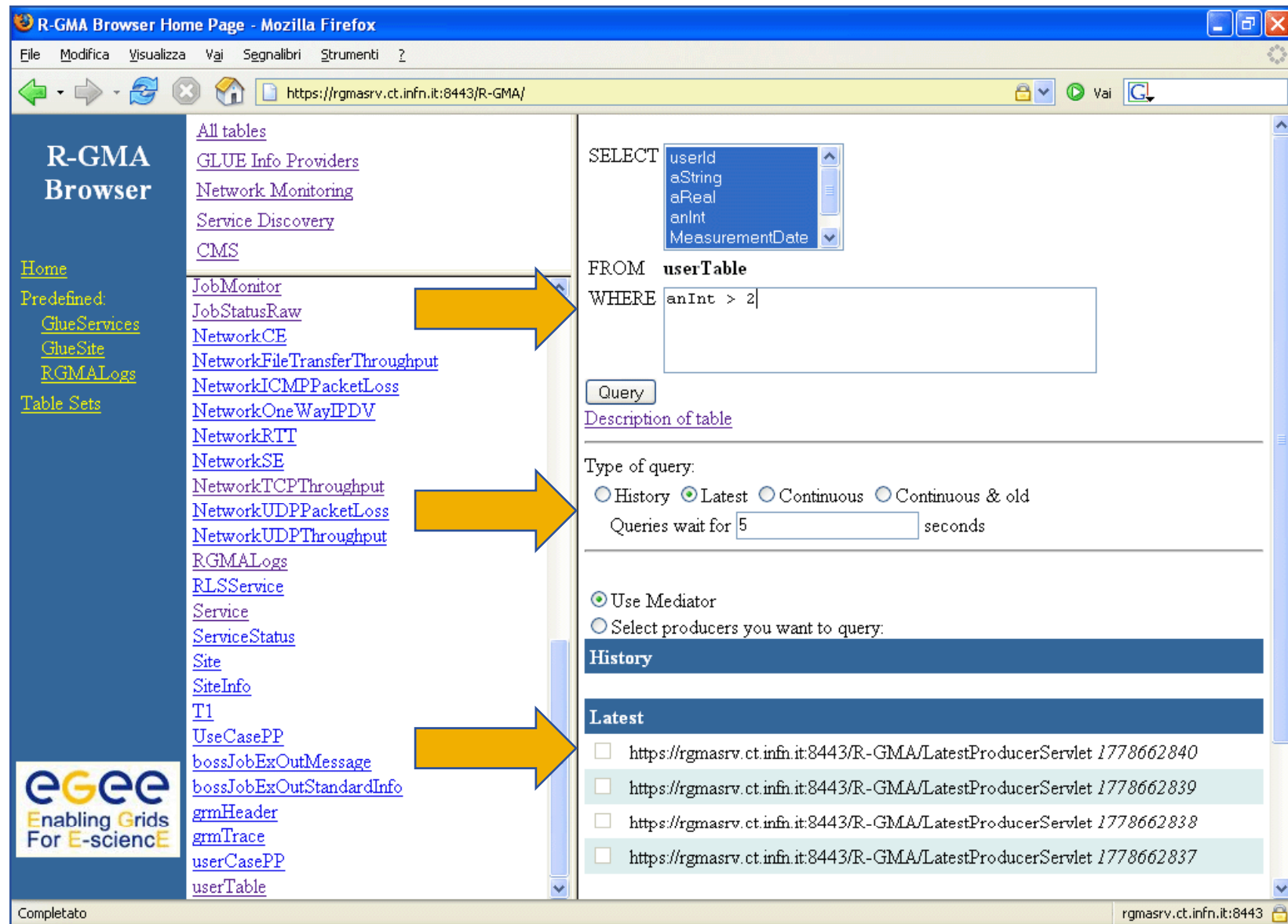


The screenshot shows the R-GMA Browser Home Page in Mozilla Firefox. The browser window title is "R-GMA Browser Home Page - Mozilla Firefox". The address bar shows the URL "https://rgmasrv.ct.infn.it:8443/R-GMA/".

The page layout includes a left sidebar with the "R-GMA Browser" logo and a list of links under "Predefined:" and "Table Sets". The main content area is divided into two columns. The left column lists various tables and services, including "All tables", "GLUE Info Providers", "Network Monitoring", "Service Discovery", "CMS", "JobMonitor", "JobStatusRaw", "NetworkCE", "NetworkFileTransferThroughput", "NetworkICMPPacketLoss", "NetworkOneWayIPDV", "NetworkRTT", "NetworkSE", "NetworkTCPThroughput", "NetworkUDPPacketLoss", "NetworkUDPThroughput", "RGMALogs", "RLSService", "Service", "ServiceStatus", "Site", "SiteInfo", "T1", "UseCasePP", "bossJobExOutMessage", "bossJobExOutStandardInfo", "grmHeader", "grmTrace", "userCasePP", and "userTable". The right column contains a query editor with a "SELECT" dropdown menu showing "userId", "aString", "aReal", "anInt", and "MeasurementDate". The "FROM" clause is set to "userTable". The "WHERE" clause is empty. Below the query editor is a "Query" button and a "Description of table" link. The "Type of query" section has radio buttons for "History", "Latest" (selected), "Continuous", and "Continuous & old". The "Queries wait for" field is set to 5 seconds. The "Use Mediator" section has radio buttons for "Use Mediator" (selected) and "Select producers you want to query:". The "History" section is empty. The "Latest" section shows a list of URLs and their corresponding IDs.

At the bottom of the browser window, the status bar shows "Completato" and the URL "rgmasrv.ct.infn.it:8443".

Query from R-GMA Browser



R-GMA Browser Home Page - Mozilla Firefox

File Modifica Visualizza Vai Segnalibri Strumenti ?

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

R-GMA Browser

Home

Predefined:

- GlueServices
- GlueSite
- RGMALogs

Table Sets

All tables

- GLUE Info Providers
- Network Monitoring
- Service Discovery
- CMS
- JobMonitor
- JobStatusRaw
- NetworkCE
- NetworkFileTransferThroughput
- NetworkICMPPacketLoss
- NetworkOneWayIPDV
- NetworkRTT
- NetworkSE
- NetworkTCPThroughput
- NetworkUDPPacketLoss
- NetworkUDPThroughput
- RGMALogs
- RLSService
- Service
- ServiceStatus
- Site
- SiteInfo
- T1
- UseCasePP
- bossJobExOutMessage
- bossJobExOutStandardInfo
- grmHeader
- grmTrace
- userCasePP
- userTable

SELECT

userId
aString
aReal
anInt
MeasurementDate

FROM userTable

WHERE anInt > 2

Query

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

- ☐ https://rgmasrv.ct.infn.it:8443/R-GMA/LatestProducerServlet 1778662840
- ☐ https://rgmasrv.ct.infn.it:8443/R-GMA/LatestProducerServlet 1778662839
- ☐ https://rgmasrv.ct.infn.it:8443/R-GMA/LatestProducerServlet 1778662838
- ☐ https://rgmasrv.ct.infn.it:8443/R-GMA/LatestProducerServlet 1778662837

Completato

rgmasrv.ct.infn.it:8443

R-GMA Browser Home Page - Mozilla Firefox

File Modifica Visualizza Vai Segnalibri Strumenti ?

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

R-GMA Browser

Home

Predefined:

[GlueServices](#)

[GlueSite](#)

[RGMALogs](#)

Table Sets

[All tables](#)

[GLUE Info Providers](#)

[Network Monitoring](#)

[Service Discovery](#)

[CMS](#)

[AppIMONIT](#)

[GAMIAAppStart](#)

[GlueBatchJob](#)

[GlueBatchQueue](#)

[GlueBatchSystem](#)

[GlueCE](#)

[GlueCEAccessControlBase](#)

[GlueCESEBind](#)

[GlueCluster](#)

[GlueHost](#)

[GlueHostLocalFileSystem](#)

[GlueHostNetworkAdapter](#)

[GlueHostPoolAccount](#)

[GlueHostProcess](#)

[GlueHostRemoteFileSystem](#)

[GlueHostRole](#)

[GlueSA](#)

[GlueSAAccessControlBase](#)

[GlueSE](#)

[GlueSEAccessProtocol](#)

[GlueSEAccessProtocolSupp](#)

[GlueSL](#)

[GlueService](#)

[GlueServiceAssociation](#)

Query: **SELECT UniqueID, TotalCPUs, Status, MeasurementDate, MeasurementTime FROM GlueCE WHERE TotalCPUs > 2**

UniqueID	TotalCPUs	Status	MeasurementDate	MeasurementTime
glite-ce.ct.infn.it:2119/blah-pbs-short	16	Production	2006-01-25	09:27:22
eggee008.cnaf.infn.it:2119/blah-pbs-long	4	Production	2006-01-25	10:01:23
eggee008.cnaf.infn.it:2119/blah-pbs-infinite	4	Production	2006-01-25	10:01:23
eggee008.cnaf.infn.it:2119/blah-pbs-short	4	Production	2006-01-25	10:01:23
glite-ce.ct.infn.it:2119/blah-pbs-infinite	16	Production	2006-01-25	09:27:22
lxcde01.pd.infn.it:2119/blah-pbs-long	6	Production	2006-01-25	09:36:15
lxcde01.pd.infn.it:2119/blah-pbs-short	6	Production	2006-01-25	09:36:15
lxcde01.pd.infn.it:2119/blah-pbs-infinite	6	Production	2006-01-25	09:36:15
glite-ce.ct.infn.it:2119/blah-pbs-long	16	Production	2006-01-25	09:27:22

Number of rows: 9

Wait for seconds

Completato

rgmasrv.ct.infn.it:8443

- **R-GMA overview page.**
 - <http://www.r-gma.org/>
- **R-GMA in EGEE**
 - <http://hepunix.rl.ac.uk/egee/jra1-uk/>
- **R-GMA command line tool**
 - <http://hepunix.rl.ac.uk/egee/jra1-uk/glite-r1/command-line.pdf>
- **R-GMA Browser Home Page**
 - <https://rgmasrv.ct.infn.it:8443/R-GMA/>