



Enabling Grids for E-science

Practical using R-GMA

www.eu-egee.org



Please download this talk – in this practical we are following PPT slides

- **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.**

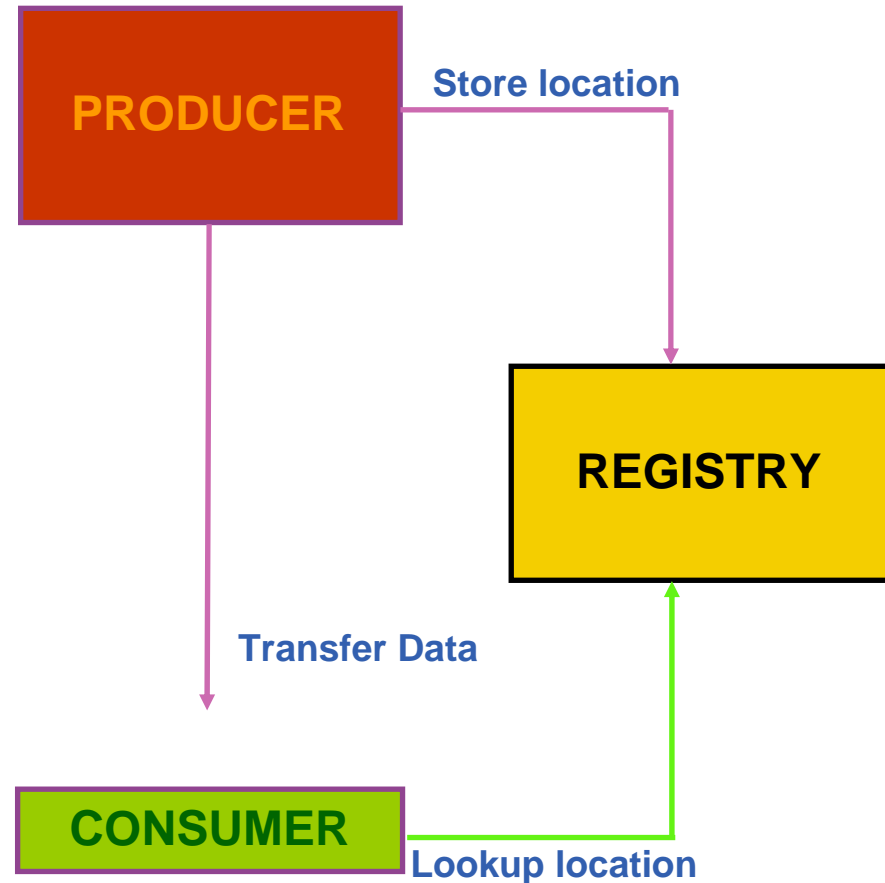
- **Relational Grid Monitoring Architecture (R-GMA)**
 - Developed as part of the EuropeanDataGrid Project (EDG)
 - Now as part of the EGEE project.
 - Based the Grid Monitoring Architecture (GMA)

- **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).

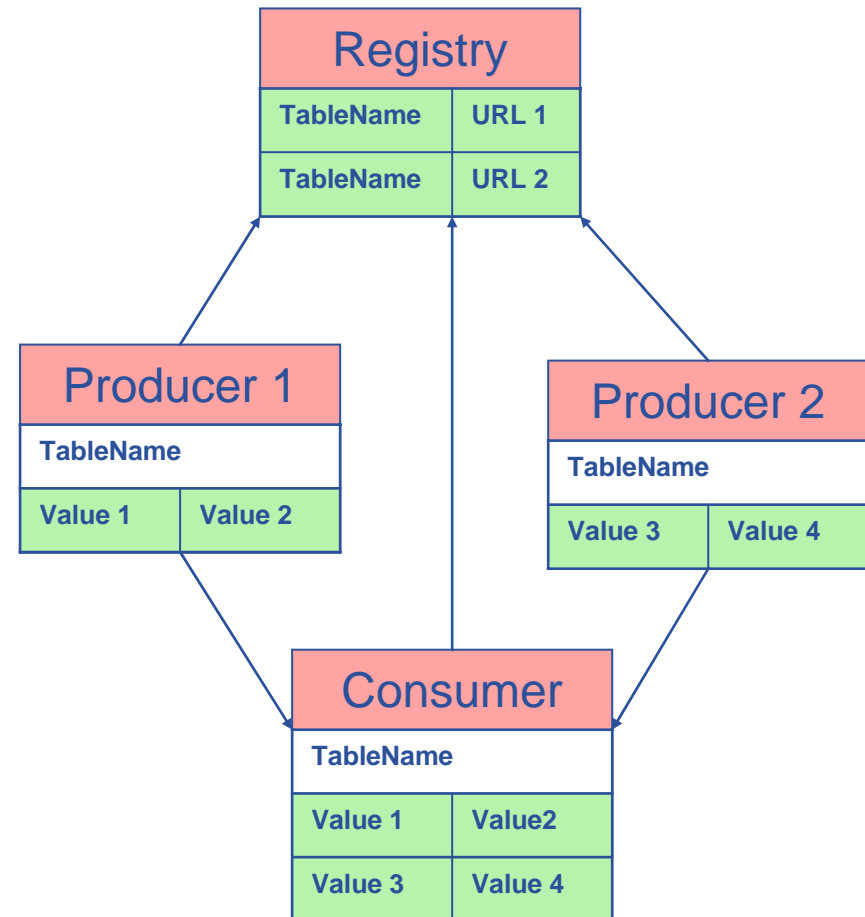
name	ID	birth	Group
Tom	4	1977-08-20	HR

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

- 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.

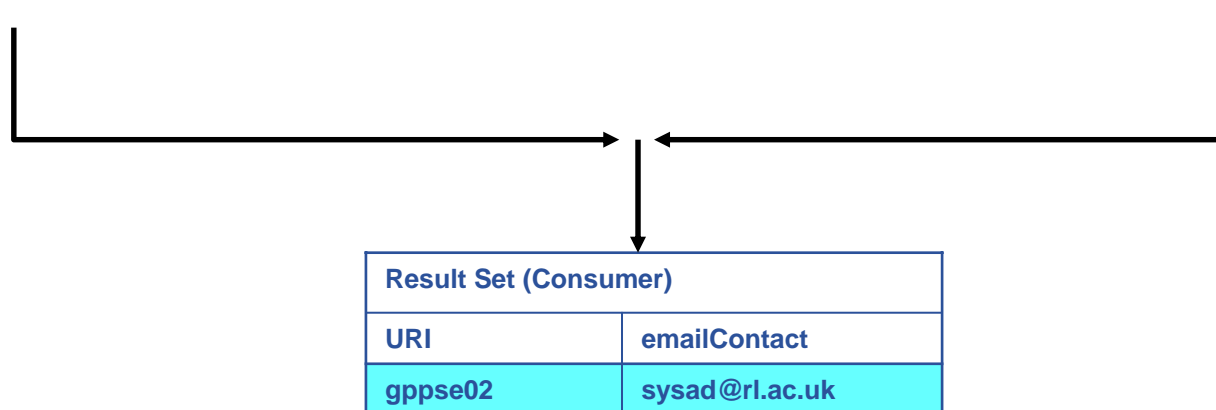


- The Consumer interrogates the Registry to identify all Producers that could satisfy the query.
- Consumer connects to the Producers.
- Producers send the tuples to the Consumer.
- The Consumer will merge these tuples to form one result set.

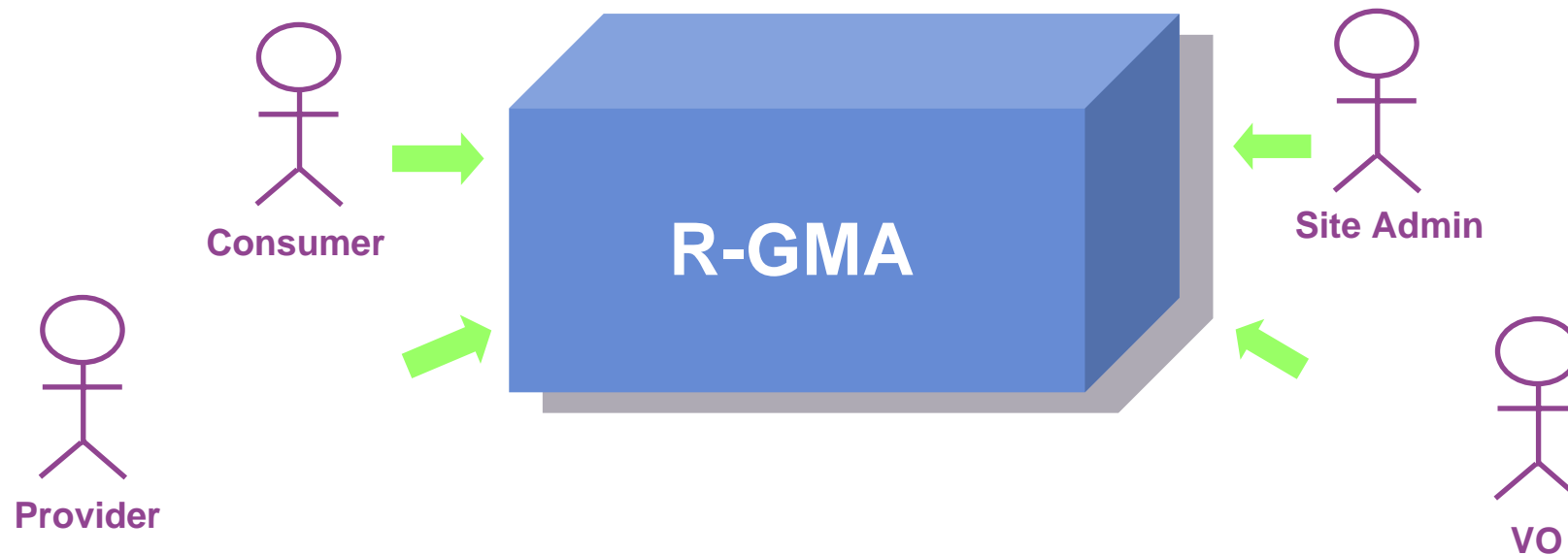


Service				
URI	VO	type	emailContact	site
gppse01	alice	SE	sysad@rl.ac.uk	RAL
gppse01	atlas	SE	sysad@rl.ac.uk	RAL
gppse02	cms	SE	sysad@rl.ac.uk	RAL
lxshare0404	alice	SE	sysad@cern.ch	CERN
lxshare0404	atlas	SE	sysad@cern.ch	CERN

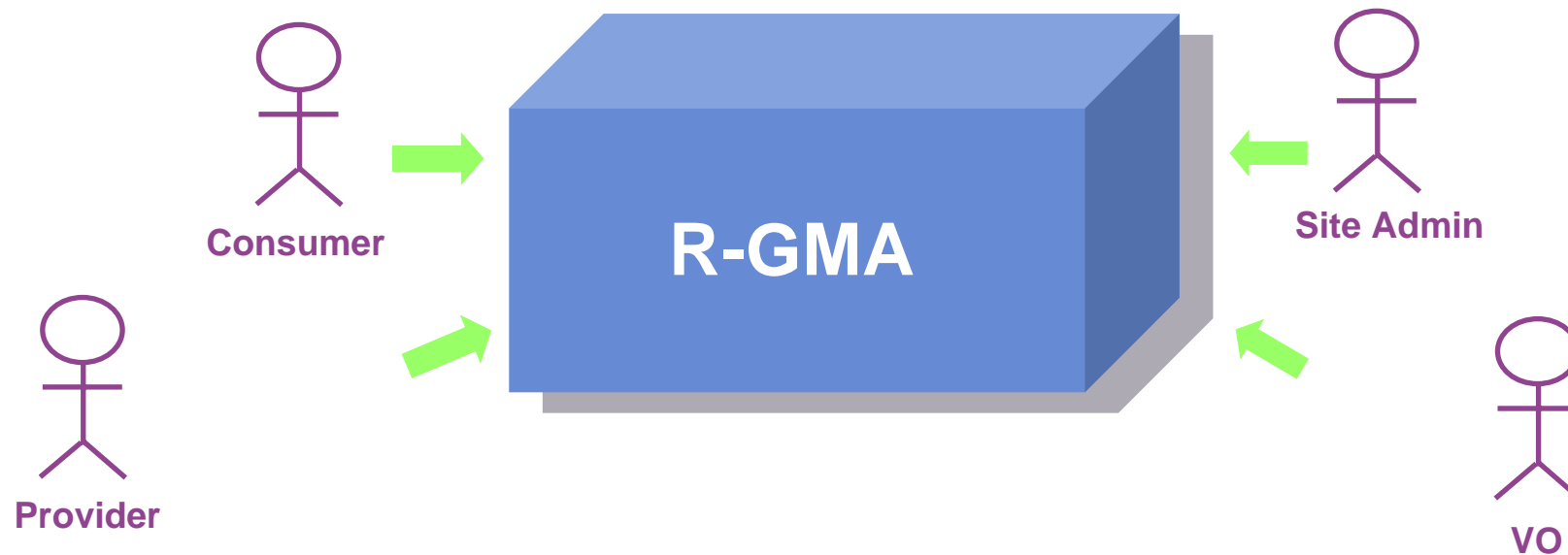
ServiceStatus				
URI	VO	type	up	status
gppse01	alice	SE	y	SE is running
gppse01	atlas	SE	y	SE is running
gppse02	cms	SE	n	SE ERROR 101
lxshare0404	alice	SE	y	SE is running
lxshare0404	atlas	SE	y	SE is running



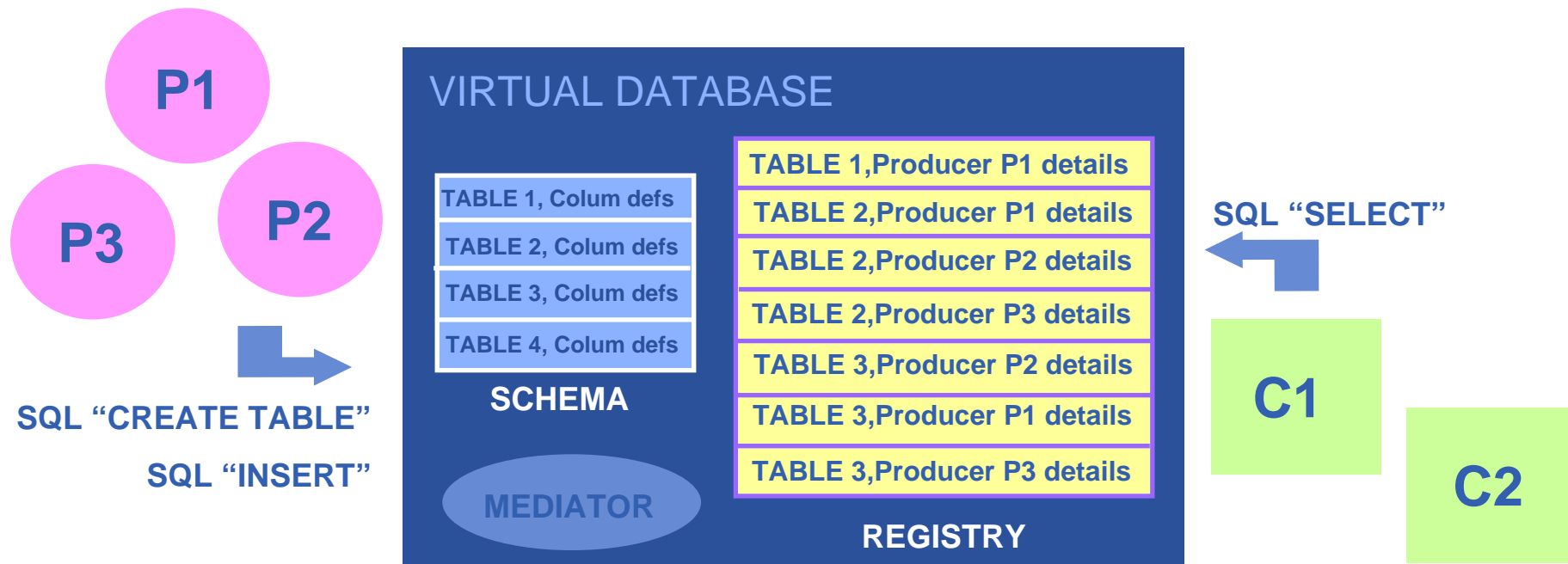
```
SELECT Service.URI Service.emailContact FROM Service S, ServiceStatus SS
WHERE (S.URI= SS.URI and SS.up='n')
```



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



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

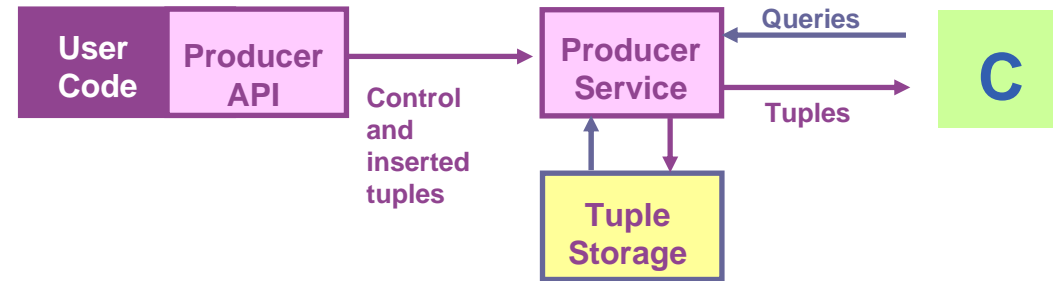
Registry is a list of data producers with all its details.

Producers publish data.

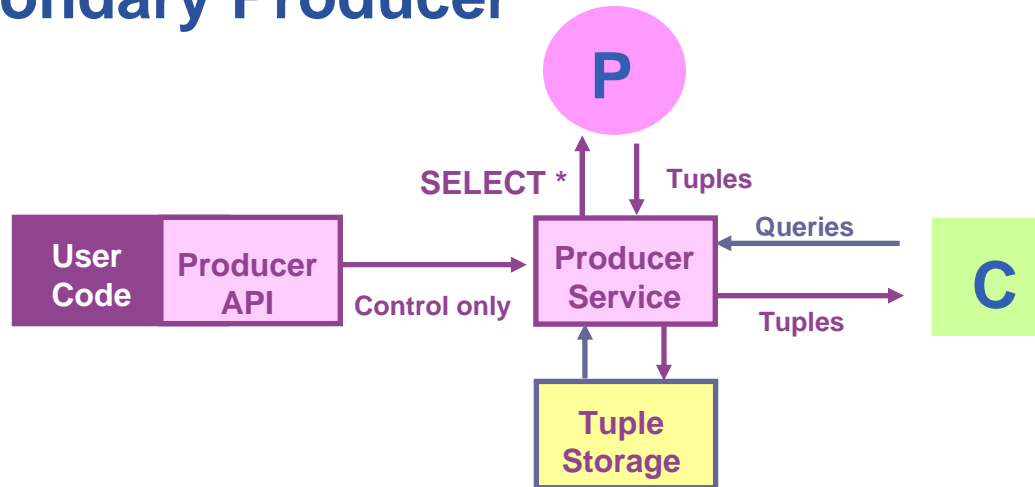
Consumer read data published.

- **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

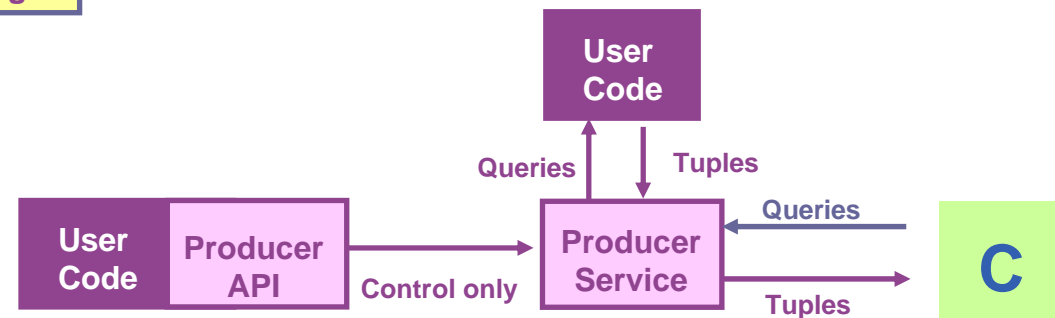


- Secondary Producer



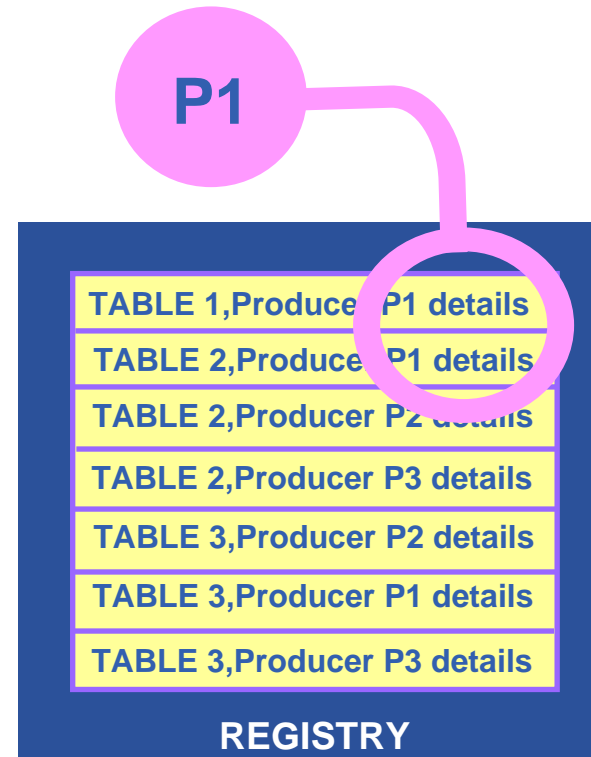
- On-Demand Producer

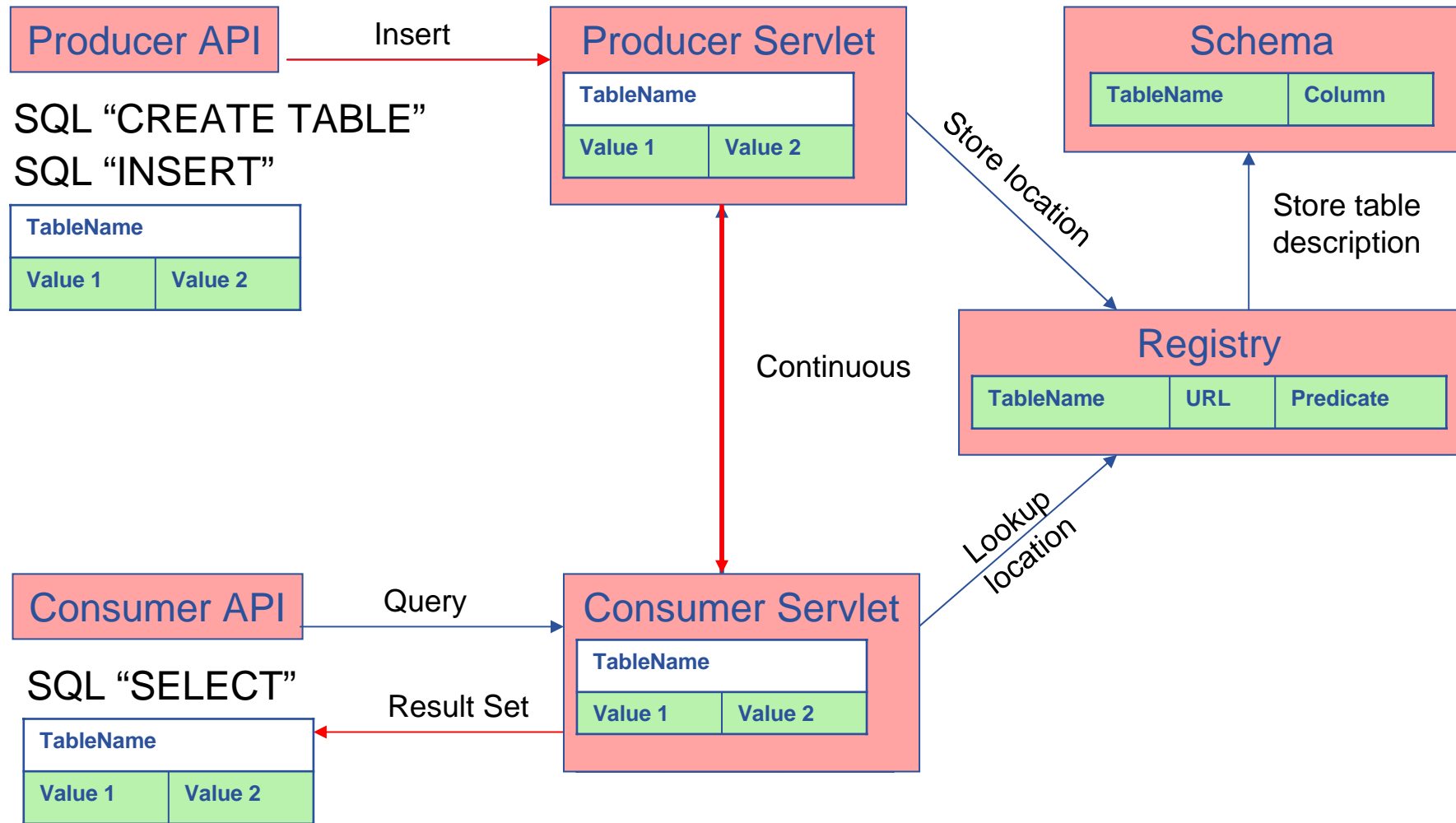
- No internal storage
- Queries passed to user code



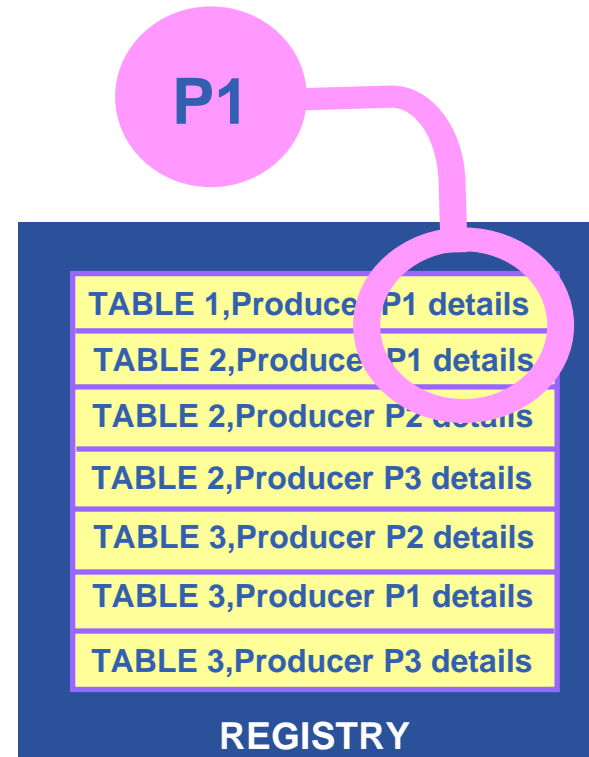
Continuous

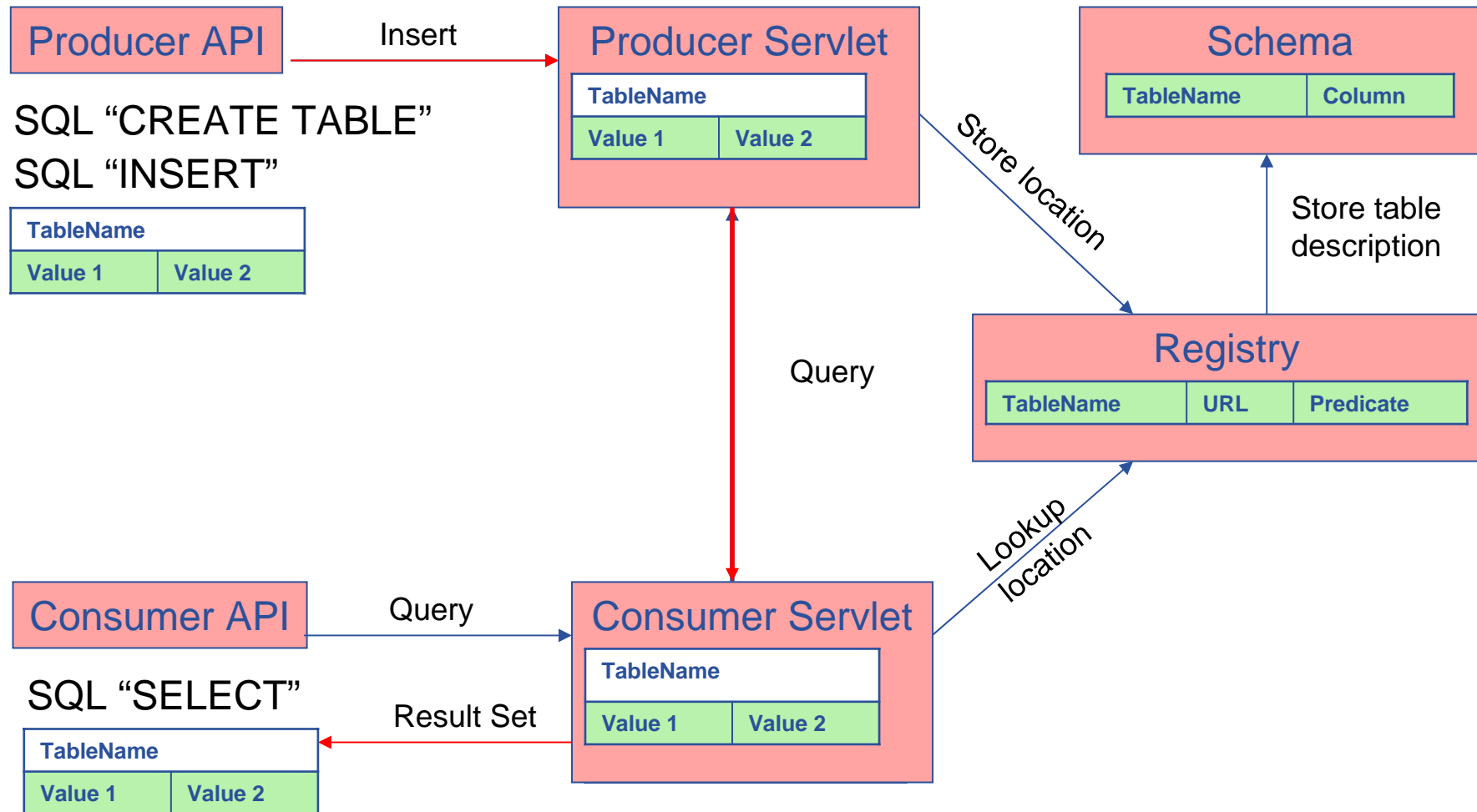
- Latest
- History
- Static



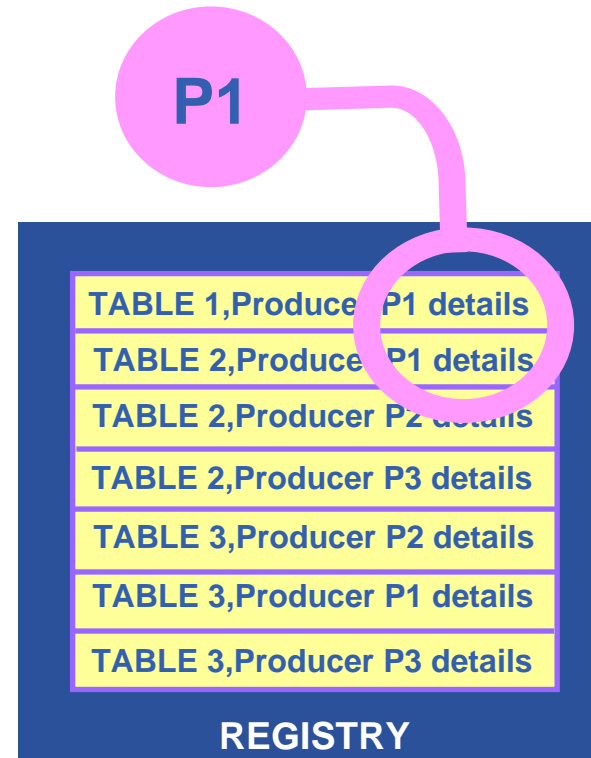
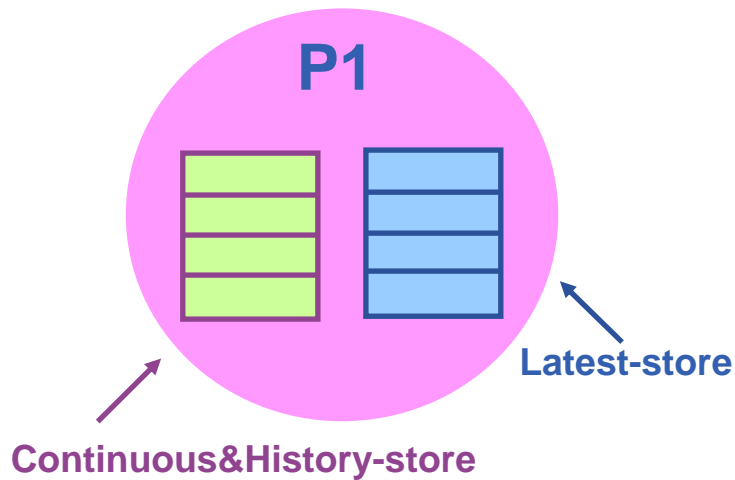


- Continuous
- Latest
- History
- Static





- Continuous
- Latest
- History
- Static



Latest Retention Period
History Retention Period

- **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.**

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

R-GMA practical

- **CHECK YOU HAVE A VOMS PROXY CERTIFICATE**
- To Start the R-GMA command line tool run the following command:

```
>rgma
```

- On startup you should receive the following message:

```
Welcome to the R-GMA virtual database for Virtual Organisations.  
You are connected to the R-GMA registry service at
```

```
http://<registry-host>:8080/R-GMA/RegistryServlet
```

```
Type "help" for a list of commands.
```

```
rgma>
```

- Commands are entered by typing at the **rgma>** prompt and hitting 'enter' to execute the command.
- A **history** of the commands executed can be accessed using the Up and Down arrow keys.
- To **search a command from history** use CTRL-R and type the first few letters of the command to recall.
- **Command autocompletion** is supported (use Tab when you have partly entered a command).

General Commands

- **exit or quit**

Exit from R-GMA command line interface.

- **help**

Display general help information.

- **help <command>**

Display help for a specific command.

- **Show tables**

Display the name of all tables existing in the Schema

- **Describe <tablename>**

Show all information about the structure of a table

- Querying data uses the standard SQL SELECT statement, e.g.:

```
rgma> SELECT * FROM GlueService
```

The behaviour of SELECT varies according to the type of query being executed. In R-GMA there are three basic types of query:

- **LATEST Queries** only the most recent tuple for each primary key
- **HISTORY Queries** all historical tuples for each primary key
- **CONTINUOUS Queries** returns tuples continuously as they are inserted.

- The type of query can be changed using the SET QUERY command as follow:

```
rgma> SET QUERY LATEST
```

or

```
rgma> SET QUERY CONTINUOUS
```

- The current query type can be displayed using

```
rgma> SHOW QUERY
```

1. Display all the table of the Schema

```
rgma>show tables
```

2. Display information about GlueSite table

```
rgma>describe GlueSite
```

3. Basic select query on the table named GlueSite

```
rgma>set query latest
```

```
rgma>show query
```

```
rgma> select Name,Latitude,Longitude from  
GlueSite
```

- The maximum age of tuples to return can also be controlled. To limit the age of latest or historical tuples use the **SET MAXAGE** command. The following are equivalent:

```
rgma> SET MAXAGE 2 minutes
```

```
rgma> SET MAXAGE 120
```

- The current maximum tuple age can be displayed using **rgma> SHOW MAXAGE**
- To disable the maximum age, set it to none:

```
rgma> SET MAXAGE none
```

- **The final property affecting queries is timeout.**
 - For a latest or history query the timeout exists to prevent a problem (e.g. network failure) from stopping the query from completing.
 - For a continuous query, timeout indicates how long the query will continue to return new tuples. Default timeout is 1 minute and it can be changed using

rgma>SET TIMEOUT 3 minutes or SET TIMEOUT 180

- **The current timeout can be displayed using**
rgma>SHOW TIMEOUT

- The SQL INSERT statement may be used to add data to the system:

```
rgma> INSERT INTO userTable VALUES ('a', 'b', 'c', 'd')
```

- In R-GMA, data is inserted into the system using a **Producer** component which handles the INSERT statement.
- Using the command line tool you may work with one producer at a time.
- The current producer type can be displayed using:

```
rgma>show producer
```
- The producer type can be set using:

```
rgma>set producer latest
```

Choose a role for the exercise as consumer or as producer (alternate if you wish)

PRODUCERS

```
rgma> set producer continuous
```

```
rgma> set maxage 3 minutes
```

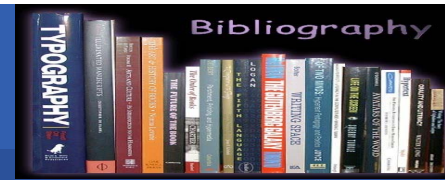
```
rgma> insert into userTable values('edinburghxx','any string',1.4,66)
```

CONSUMERS

```
rgma> set query continuous OR set query history
```

```
rgma> set timeout 5 seconds
```

```
rgma> select * from userTable
```



LCG-2 User Guide Manual Series

<https://edms.cern.ch/file/454439/LCG-2-UserGuide.html>