



LAVOISIER

A Data Source Composition Service

CC-IN2P3



CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE



S. Reynaud, G. Mathieu, P. Girard, F. Hernandez, IN2P3/CNRS Computing Centre, France
O. Aidel, CS-SI, France

How to deal with data sources heterogeneity ?

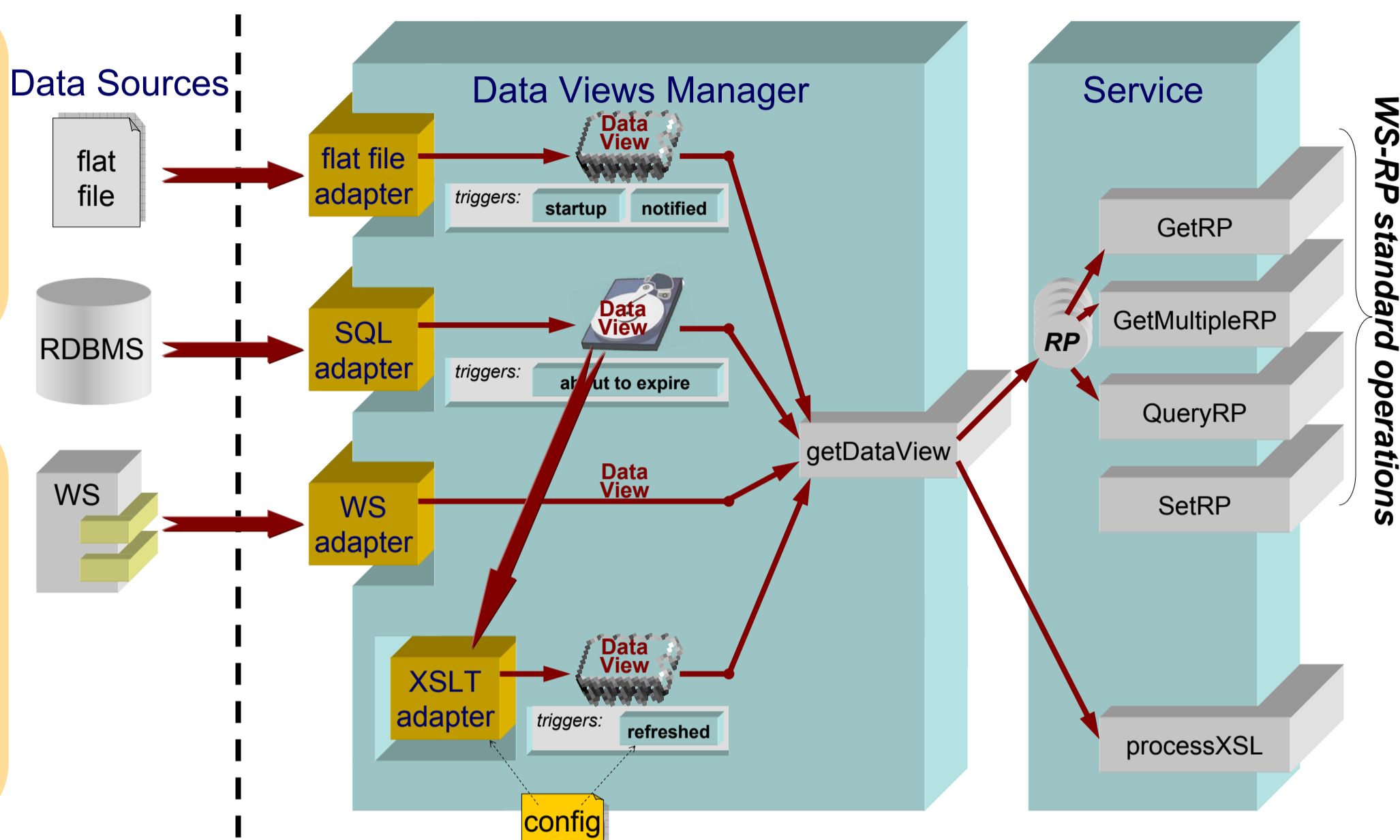
By providing an **unified view** of data collected from multiple heterogeneous data sources.

Lavoisier is an extensible **service** designed to provide such a view, by aggregating and potentially caching XML **data views** of the data collected from the legacy **data sources**.

Reusable adapters

Reusable adapters for data sources relying on the following technologies are provided:

- JNDI (LDAP, ...)
- SQL
- Flat files
- Local XML files
- Remote XML files (through HTTPS, ...)
- Processing XSL on data views
- Index of data views
- ...

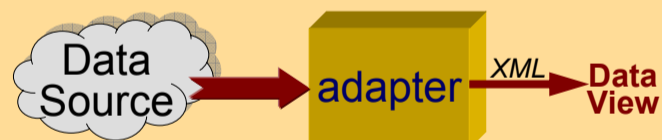


Roles with Lavoisier

The developer role:

Develop **new adapters** supporting part or all of the following capabilities (if needed):

- Generate XML view as a DOM tree
- Generate XML view as a stream
- Update data source with Lavoisier
- Notify Lavoisier when the data source is updated without it
- Provide index of argument values



The service administrator role:

Configure adapters, and **tune** cache mechanisms according to the characteristics and usage profile of both the data sources and the data views (access frequency, amount of data, latency...):

➢ Rules for triggering cache refresh:

- triggers:**
 - startup** ➢ When the data view is initialized
 - time** ➢ When a time-based event occurs ("cron-like")
 - updated** ➢ When the data source is updated through Lavoisier
 - notified** ➢ When the data source is updated by an external actor
 - about to expire** ➢ When the data view is about to expire (proactive)
 - accessed & expired** ➢ When the data view is both accessed and expired
 - refreshed** ➢ When the cache of one of the specified data views has been refreshed

➢ Cache type:



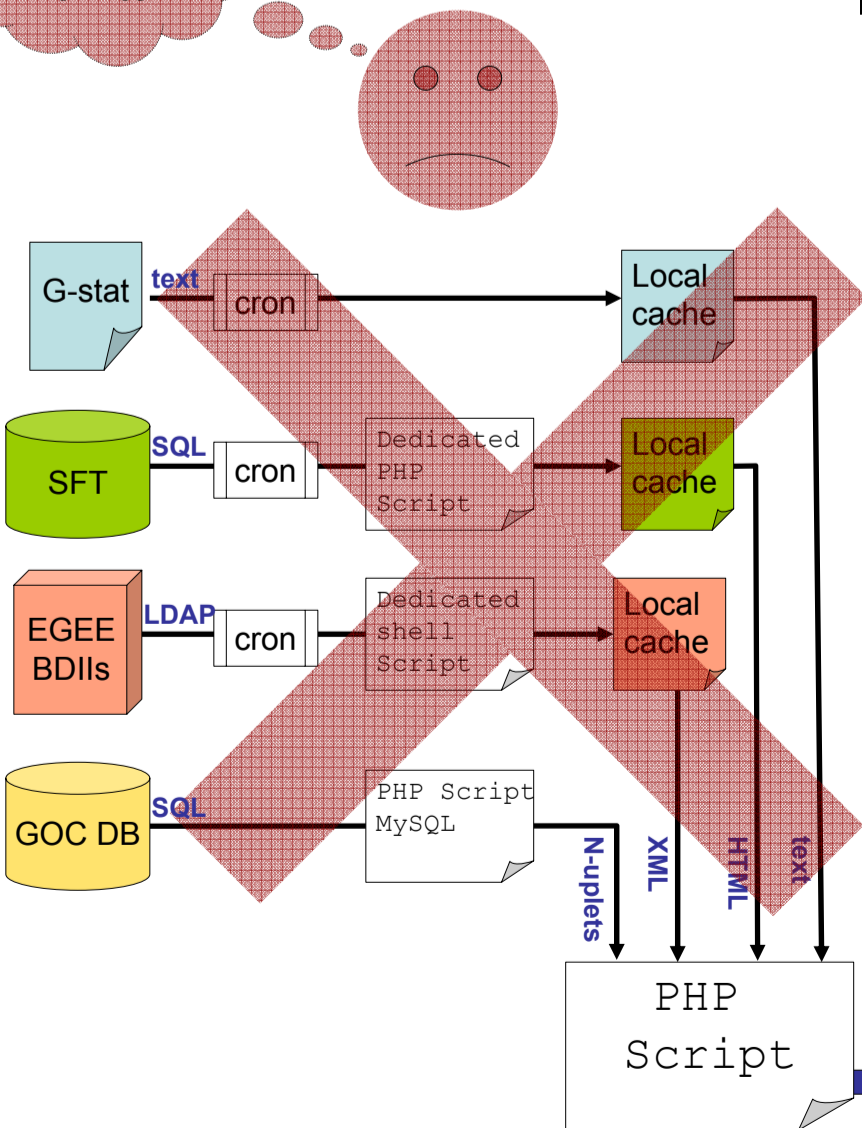
The user role:

Access/query data views content through Web Service operations

- Access data views through WS-Resource Properties standard operations (GetRP, GetMultipleRP, QueryRP, SetRP)
- Process XSL stylesheets on multiple data views

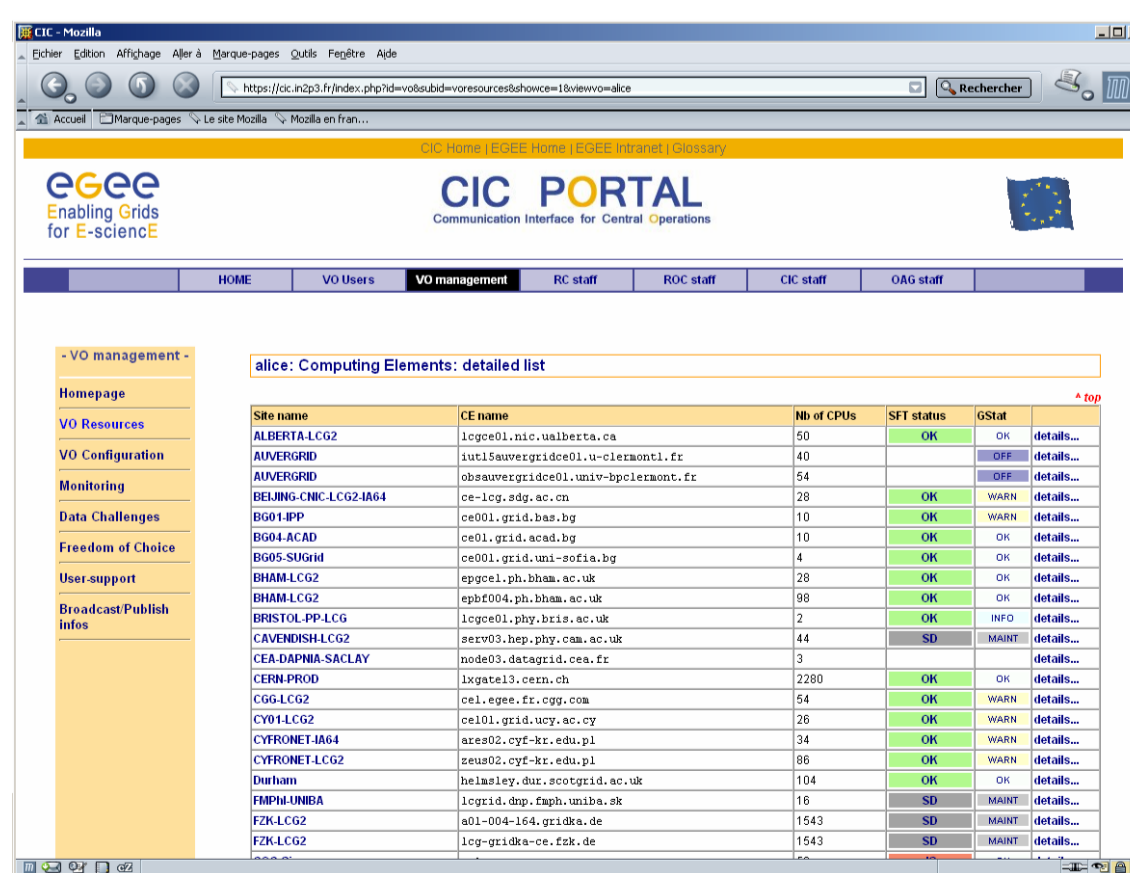
Hard to maintain

without Lavoisier



Usage example

Information summary on the EGEE CIC Portal



with Lavoisier

Cool!

