GridICE Monitoring:
a VO-oriented perspective

Enrico Fattibene
INFN-CNAF (Italy)
enrico.fattibene <at> cnaf.infn.it

NA4 Generic Application Meeting
Catania 2006, 10 January

www.eu-egee.org

Information Society

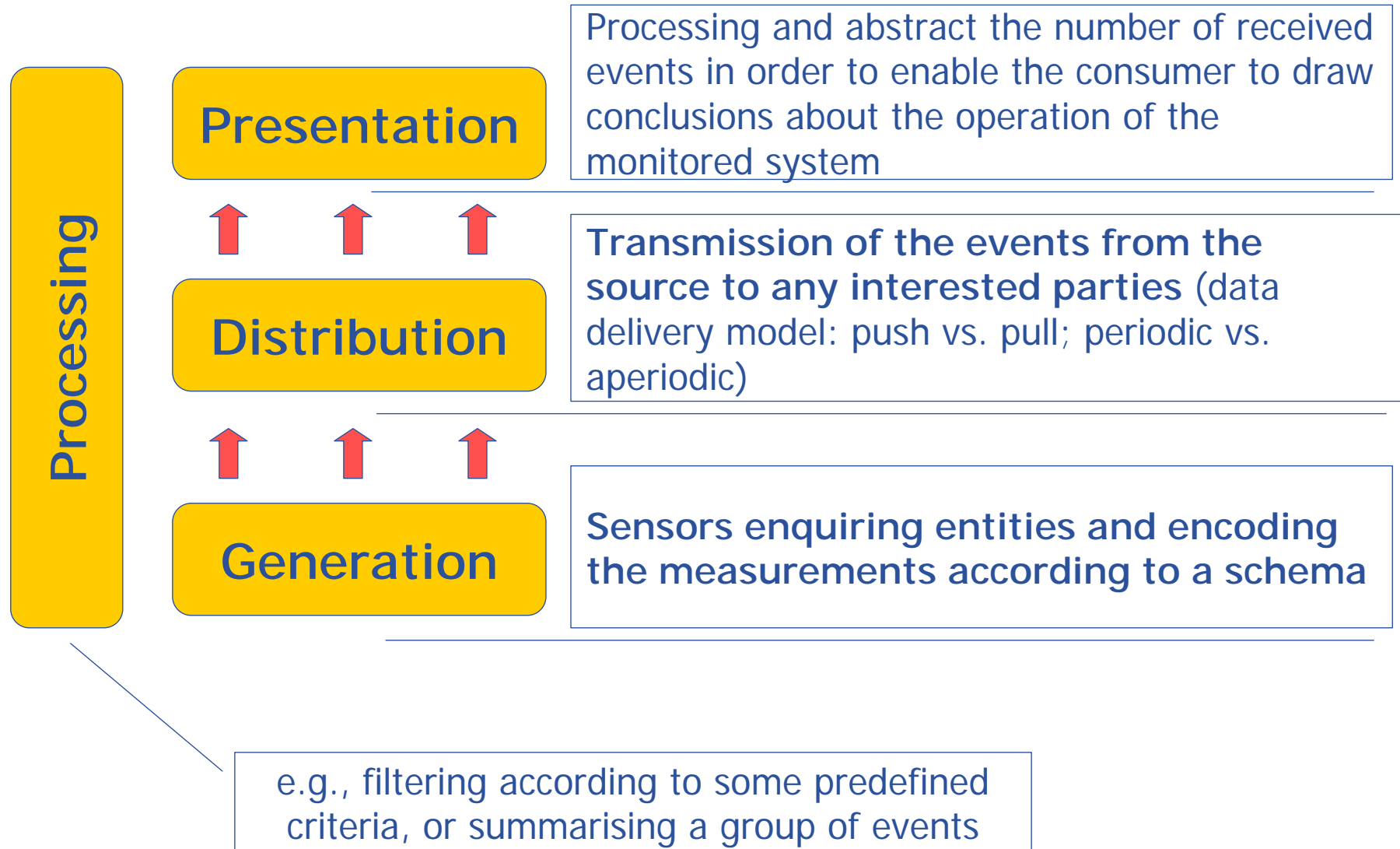Enabling Grids for E-sciencE

- **GridICE tool**
  - Brief introduction
  - Monitoring phases
  - GridICE architecture
  - VO manager viewpoint

- **Live demo**
  - Web interface organization
  - Common features
  - VO view

Enabling Grids for E-sciencE

- – **Grid resources availability is subject to failures.**
- – **Resources observability is necessary for the Grid utilization.**

egee

Enabling Grids for E-sciencE

GridICE:

- is a distributed monitoring tool for grid systems

- integrates with local monitoring systems

- offers a web interface for publishing monitoring data at the Grid level

- fully integrated in the LCG-2 Middleware
  - gridice-clients data collector installation and configuration for each site realized by the Yaim scripts.

Processing

Presentation

Distribution

Generation

Processing and abstract the number of received events in order to enable the consumer to draw conclusions about the operation of the monitored system

Transmission of the events from the source to any interested parties (data delivery model: push vs. pull; periodic vs. aperiodic)

Sensors enquiring entities and encoding the measurements according to a schema

e.g., filtering according to some predefined criteria, or summarising a group of events

- **Generation of events:**
  - Sensors: typically perl scripts or c programs.
  - Schema:
    - GLUE Schema v.1.1 + GridICE extension.
      - *System related (e.g., CPU load, CPU Type, Memory size).*
      - *Grid service related (e.g., CE ID, queued jobs).*
      - *Network related (e.g., Packet loss).*
      - *Job usage (e.g., CPU Time, Wall Time).*
  - All sensors are executed in a periodic fashion.

- **Distribution of events:**
  - Hierarchical model.
    - **Intra-site**: by means of the local monitoring service
      - *default choice, LEMON (http://www.cern.ch/lemon).*
    - **Inter-site**: by offering data through the Grid Information Service.
    - **Final Consumer:** depending on the client application.
  - Mixed data delivery model.
    - **Intra-site**: depending on the local monitoring service (push for lemon).
    - **Inter-site**: depending on the GIS (current choice, MDS 2.x, pull).
    - **Final consumer**:
      - *pull (browser/application)*
      - *push (publish/subscribe notification)*

Enabling Grids for E-sciencE

- **Data stored in a RDBMS used to build aggregated statistics.**

- **Data retrieved from the RDBMS are encoded in XML files.**

- **XSL to XHTML transformations to publish aggregated data in a Web context.**

Enabling Grids for E-sciencE

- **Location bar.**
- **Context-aware tabbed views.**
- **Possibility of sorting rows by any column attribute.**
- **Common background colour for rows showing the same value for the sorting attribute.**
- **Monitoring data available as XML document.**
- **Help pages.**

Enabling Grids for E-sciencE

Need for analyzing the usage, behavior and performance of a Grid depending on different users:

- **VO manager**

  actual set of resources accessible to its members

- **Grid operations manager**

  all resources for what GOC is responsible

- **Site administrator**

  own site resources

eG**ee**

Enabling Grids for E-sciencE

- **Visualization of the actual set of resources accessible to its members.**

- **Evaluation of how the VO members requests are being distributed over the available resources.**

- **Evaluation of the Service Level Agreement (SLA) for the global Grid service offers.**

- **Run retrospective analysis**

Mostly interested in:

– **Resources available to the VO**

▪ Computing elements where VO users can submit jobs.

▪ Storage elements where VO users can store/retrieve data.

– **Job monitoring**

▪ How many jobs are running or queued?

- *For the whole VO? In each site? Submitted by a certain RB?*

▪ How many jobs have been executed?

- *For the whole VO? In each site? Submitted by a certain RB?*

- **GridICE server installed and configured to highlight the monitoring of the GILDA VO and sites in terms of:**

  1. *Grid resources.*
  2. *Fault detection.*
  3. *Job activity.*

References

- **Dissemination Web site:** http://grid.infn.it/gridice
- **GridICE Server for GILDA testbed:**

   http://alifarm7.ct.infn.it:50080/gridice

- **GridICE Server for LCG-Grid:**

   http://gridice2.cnaf.infn.it:50080/gridice

GridICE publications:

- S. Andreozzi, N. De Bortoli, S. Fantinel, A. Ghiselli, G. L. Rubini, G. Tortone, M. C. Vistoli GridICE: a monitoring service for Grid systems, Future Generation Computer System 21 (2005) 559–571
- S. Andreozzi, A. Ciuffoletti, A. Ghiselli, C. Vistoli. Monitoring the connectivity of a Grid. Proceedings of the 2nd International Workshop on Middleware for Grid Computing (MGC 2004) in conjunction with the 5th ACM/IFIP/USENIX International Middleware Conference, Toronto, Canada, October 2004.
- S. Andreozzi, N. De Bortoli, S. Fantinel, G.L. Rubini, G. Tortone. *Design and Implementation of a Notification Model for Grid Monitoring Events.* CHEP04, Interlaken (CH), Sep 2004.