ALICE Grid monitoring with MonALISA

Catalin Cirstoiu

7/06/2006
ALICE Offline Week
Contents

- Why we need monitoring
- Requirements
- What we monitor
- MonALISA overview
- Monitoring architecture in AliEn
- Examples
Why we need monitoring

- Worldwide distributed system
  - Various platforms
  - Various error conditions
  - Various software

- Key parameters
  - System status
  - Resource usage

- Debugging

- Correlating events
Requirements

- Global view of the distributed system
- Non-intrusiveness
- Accuracy
- Near real-time image of the status
- Long-term history of aggregated data
What we monitor

- **AliEn Components**
  - Central Services
    - Task Queue, Information Service, Optimizers, API etc.
  - Site Services
    - Cluster Monitor, Computing & Storage Elements
    - Job Agents
    - Jobs status & resource usage

- **Other Services**
  - CastorGrid staging & migration, Xrootd, MySQL

- **Nodes**
  - Central, site, worker nodes

- **Network traffic – inter & intra site**
  - Via Xrootd
  - Via FTD
MonALISA overview

- MonALISA is a Dynamic, Distributed Service Architecture capable to collect any type of information from different systems, to analyze it in near real time and to provide support for automated control decisions and global optimization of workflows in complex grid systems.
ML Discovery System & Services

- The framework is based on a hierarchical structure of loosely coupled agents acting as distributed services which are independent & autonomous entities able to discover themselves and to cooperate using a dynamic set of proxies or self-describing protocols.
ApMon – Application Monitoring

- Lightweight library of APIs (C, C++, Java, Perl, Python) that can be used to send any information to MonALISA Services
- High comm. performance
- Flexible
- Accounting
- Sys Mon

- Application Monitoring
  - Time:IP:procID
  - parameter1: value
  - parameter2: value
  - ...

- System Monitoring
  - load1: 0.24
  - processes: 97
  - pages_in: 83

- MonALISA hosts
  - Config Servlet dynamic reloading

- MonALISA Service
  - UDP/XDR Monitoring Data

- No Lost Packages
- MonALISA CPU Usage (%)
- Messages per second

ApMon configuration generated automatically by a servlet / CGI script
Monitoring architecture in AliEn

MonALISA @Site

Aggregated Data

MonALISA LCG Site

MonALISA Repository

http://pcalimonitor.cern.ch:8889/
Deployment and configuration

- MonALISA is packaged and prepared for installation by the AliEn Build System (BITS)
- From site administrators point of view, it is just like any other AliEn service
  - You start it with `alien StartMonaLisa`
  - You stop it with `alien StopMonaLisa`
  - Check status with `alien StatusMonaLisa`
- Configuration files for ML are generated automatically from AliEn LDAP
Job status monitoring

- Global summaries
  - For each/all statuses
  - For each/all sites
  - For each/all users
  - Running & cumulative

- Error statuses
- From job agents
- From central services
- Real-time map view
- Integrated pie charts
- History plots
Job status – real-time map

MonALISA Repository
ALICE

MonALISA Client
Logo on the button below to start the Monalisa Client.

ALICE Repository
Interactive Map
Job Information
Site views
Summary plots
Jobs status
Resource usage
User views
Summary plots
Jobs status
TaskQueue
SE Information
Services
Central Services
Site Services
VO Events
MONA
Network Traffic
AIIEn Tests
FTD Transfers
Original

Action Configuration
AliEn Configuration
Site Administration
Farm Colour Configuration

Job Status
Utilisation: Busy / Free Nodes

Running Jobs
Zombie Jobs
No Active Jobs
ML Service Down

Show links

7/06/2006
Catalin.Cirstoiu@cern.ch
Job status – integrated pie charts

Done jobs statistics

- Torino: 2.2%
- CERN: 31.3%
- Muenster: 23.0%
- Catania: 0.2%
- BNGL: 0.1%
- RAL: 3.3%
- PNNL: 0.1%
- OSC: 4.2%
- NIHAM: 3.4%
- LHCb: 1.5%
- INR: 0.3%
- ITEP: 1.8%
- CNAF: 17.6%
- Bari: 2.3%
- CERN-L: 1,191
- IHEP: 137
- ITEP: 110
- INR: 19
- Legnaro: 99
- Muenster: 1,591
- NIHAM: 229
- OSC: 281
- PNNL: 31
- RAL: 52
- RRC-KI: 38
- SARA: 408
- Torino: 147
Job status – history plots
Job resource usage monitoring

- **Cumulative parameters**
  - CPU Time
  - Wall time
  - Input & output traffic (xrootd)
  - Read & written files

- **Running parameters**
  - Resident memory
  - Virtual memory
  - Open files
  - Workdir size
  - Disk usage
  - CPU usage

- Aggregated per site
Job agents monitoring

- From Job Agent itself
  - Requesting job
  - Installing packages
  - Running job
  - Done
  - Error statuses

- From Computing Element
  - Available job slots
  - Queued Job Agents
  - Running Job Agents
Job network traffic monitoring

- Based on Jobs’ xrootd transfers
- Aggregated statistics for
  - Sites (incoming, outgoing, site to site, internal)
  - Storage Elements (incoming, outgoing)
- On
  - Read and written files
  - Transferred MB/s

![Written files to SEs](image)
AliEn & LCG Services monitoring

- AliEn services
  - Periodically checked
  - PID check + SOAP call

- LCG environment and tools
  - Proxy, gsiscp, LCG CE/SE, Job submission
  - BDII, Local catalog, software area etc.
  - Error messages in case of failure
CastorGrid scripts monitoring

- Migration
  - Amount, speed, errors
- Staging
  - Amount, speed, errors
- Nodes
  - Host parameters
- Xrootd resource usage
- File cache status
  - Used space, no. of files
API Services monitoring

- API Service sessions
  - Established, active
- API Service users
  - Active, total
- Statistics
  - Executed commands
## VOBox monitoring

- **Machine parameters, real-time & history**
  - Load, Mem usage, processes, sockets
- **Pledged resources**

### VO Box machine status

<table>
<thead>
<tr>
<th>Site name</th>
<th>Last seen online</th>
<th>ML name</th>
<th>Load5</th>
<th>Usage</th>
<th>User</th>
<th>System</th>
<th>Cnt</th>
<th>MHz</th>
<th>System</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. CERN (CAF)</td>
<td>2006-06-25 11h</td>
<td>CERNPS</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2. CERN (CAF)</td>
<td>2006-06-01 04h</td>
<td>CERNL</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3. CERN + FZK</td>
<td>2006-03-30 02h</td>
<td>FZK</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4. INFN (CAF)</td>
<td>2006-06-06 00h</td>
<td>CNAP</td>
<td>0.083</td>
<td>4.236</td>
<td>3.92</td>
<td>1.313</td>
<td>2</td>
<td>1.367</td>
<td>99</td>
</tr>
<tr>
<td>5. INFN (DB)</td>
<td>2006-06-06 00h</td>
<td>INFNDB</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>6. ML Trento</td>
<td>2006-06-06 00h</td>
<td>MLTRENT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7. UK Trento</td>
<td>2006-06-02 10h</td>
<td>UKTRENT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13. FZU AS Prague</td>
<td>2006-06-06 00h</td>
<td>FZUAS</td>
<td>0.025</td>
<td>1.427</td>
<td>0.552</td>
<td>0.575</td>
<td>2</td>
<td>3.666</td>
<td>86</td>
</tr>
<tr>
<td>14. GSI</td>
<td>2006-06-06 00h</td>
<td>GSI</td>
<td>0.033</td>
<td>0.869</td>
<td>0.58</td>
<td>0.286</td>
<td>4</td>
<td>2.967</td>
<td>87</td>
</tr>
<tr>
<td>15. INFN Tier2 Federation</td>
<td>2006-06-06 00h</td>
<td>INFNT2</td>
<td>0.844</td>
<td>2.031</td>
<td>1.406</td>
<td>0.746</td>
<td>2</td>
<td>1.860</td>
<td>97</td>
</tr>
<tr>
<td>16. Madrid (Spain)</td>
<td>2006-06-06 00h</td>
<td>MADRID</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>17. Munich</td>
<td>2006-06-06 00h</td>
<td>MUNCHEN</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>18. ROXO</td>
<td>2006-06-06 00h</td>
<td>ROXO</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>19. RMK (Hungary)</td>
<td>2006-06-06 00h</td>
<td>RMK</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>20. Romanian Tier-2 Federation</td>
<td>2006-06-06 00h</td>
<td>ROMN2</td>
<td>0.117</td>
<td>0.859</td>
<td>0.505</td>
<td>0.364</td>
<td>4</td>
<td>2.503</td>
<td>86</td>
</tr>
<tr>
<td>21. SARA</td>
<td>2006-06-06 00h</td>
<td>SARA</td>
<td>0.036</td>
<td>2.115</td>
<td>1.537</td>
<td>0.578</td>
<td>2</td>
<td>3.657</td>
<td>51</td>
</tr>
<tr>
<td>22. Sejong (Korea)</td>
<td>2006-06-06 00h</td>
<td>SEO</td>
<td>0.259</td>
<td>4.004</td>
<td>2.559</td>
<td>2.465</td>
<td>1</td>
<td>3.192</td>
<td>59</td>
</tr>
</tbody>
</table>

### VOBox parameters

- **CPU Usage**
- **Memory Usage**
- **Swap Usage**
- **Disk Usage**
- **Network**
- **Processes**
- **Sockets**

### Date
7/06/2006

Catalin.Cirstoiu@cern.ch
Summary

- Monitoring is vital in a large distributed system
  - Distributed systems should have built-in monitoring capabilities
- We have to deal with a lot of data
  - We need decentralization
  - And aggregation
- MonALISA is very flexible and powerful
  - Monitoring is just the beginning
  - Gathered data can be used for intelligent control and decisions taking
Questions?

Thank you!

http://alien.cern.ch

http://monalisa.caltech.edu