EMI Data, the unified European Data Management Middleware

Patrick Fuhrmann (DESY)
EMI Data Area lead
Credits

- Alejandro Alvarez
- Alex Sim
- Claudio Cacciari
- Christion Bernardt
- Christian Loeschen
- Dirk Duellmann
- Elisabetta Ronchieri
- Fabrizio Furano
- Giuseppe Fiameni
- Giacinto Donvito
- Giuseppe Lo Presti
- Jon Kerr Nilsen
- Jan Schaefer
- Jean-Philippe Baud
- Karsten Schwank
- Michele Carpene
- Michele Dibenedetto
- Michail Salichos
- Mischa Salle
- Oscar Koeroo
- Oliver Keeble
- Paul Millar
- Ralph Mueller-Pfefferkorn
- Ricardo Rocha
- Riccardo Zappi
- Tigran Mkrtchyan
- Zsolt Molnar
- Zsombor Nagy

Our wiki: https://twiki.cern.ch/twiki/bin/view/EMI/EmiJra1T3Data
Outline

• The EMI Factsheet
• From 22 Partners to EMI-Data (quick and boring)
• The EMI-Data shopping cart.
• The EMI(-Data) release plan.
• The EMI-Data Mission
• Conclusions
EMI Factsheet

- **Budget**: about 24 Million Euros
- **Funding**: about 50% by EU-FP7, rest by partners
- ** Covers**: JRA, SA and NA
- **Partners**: 22
- **Middlewares**: Arc, gLite, UNICORE and dCache
According to our Project Director, Alberto Di Meglio:

The European Middleware Initiative (EMI) project represents a close collaboration of the major European middleware providers - ARC, gLite, UNICORE and dCache - to establish a sustainable model to support, harmonise and evolve distributed computing middleware for deployment in EGI, PRACE and other distributed e-Infrastructures.
EMI in context

EGI, PRACE, WLCG, OSG

Requirements

SLAs & Support

Releases

EMI

Collaborations

ESFRI, VRCs

Stolen from Alberto Di Meglio

Standards Industry

DCI collaborations

StratusLab

VENUS-C

SIENA

EDGI

IGE

Mar 25, 2011

EMI Data, ..., ISGC 2011, ASGC, Taipei, TW
EMI Data in context

dCache, StoRM, DPM, FTS, LFC, GFAL, arc-libs, UNICORE-SMS, etc
A-REX, UAS-Compute, WMS, CREAM, MPI, etc
ARGUS, VOMS, UNICORE-Gate, gridSite, etc
Information system, accounting, bookkeeping

Mar 25, 2011
EMI Data, .... ISGC 2011, ASGC, Taipei, TW
What does EMI-Data provide?

The EMI-Data shopping basket
The EMI shopping cart

Reliable File Transport Service

File Location and meta data Service (LFC)

Professional Storage Solutions
*Fits all sizes (IKEA approach)*

Data Access and control library(s)

Generic Meta Data Service (AMGA)

- DOG
- Happy
- Dancing

Mar 25, 2011
EMI Data, .... ISGC 2011, ASGC, Taipei, TW
**Release Plan**

**Kebnekaise, Lappland, Sw, 2100m**

Giebnegáisi

**Stolen from Alberto Di Meglio**

01/05/2010
31/10/2010
30/04/2011
30/04/2012
28/02/2013

Support & Maintenance

Support & Maintenance

Support & Maintenance

Major releases

Supp. & Maint.
The Mission
The Mission

• Fixing of issues based on the experience of operating the infrastructures for some years.
• Improving or creating interoperability between components and middle-ware.
• Reducing components by merging functionality or removing duplication.
• Applying standards where available
• Standardizing EMI-Data mechanisms with “standardization bodies” e.g. OGF
• Attracting new communities.
• Becoming competitive and attractive
  – Standards
  – Professional Support
  – Strict quality monitoring
Applying industry standards

- Posix file system: NFS 4.1 / pNFS
- WebDAV
- SSL security for SRM
WebDAV

- Very useful for new (non-LHC) communities.
- IETF Standard
- Allows “File system like” access with
  - Mac OS
  - Linux
  - Windows
Standardization: NFS 4.1 / pNFS

Simplicity

✓ Regular mount-point and real POSIX I/O
✓ Can be used by unmodified applications (e.g. Mathematica..)
✓ Data client provided by the OS vendor
✓ Smart caching (block caching) development done by OS vendors

Performance

✓ pNFS: parallel NFS (first version of NFS which support multiple data servers)
✓ Clever protocols, e.g. Compound Requests

Why pNFS

EMI-1  EMI-2  EMI-3

dCache  ✓ ✓ ✓
StoRM  ✓ ✓ ✓
DPM  ✓ ✓ ✓
Availability for production use

<table>
<thead>
<tr>
<th>Industry</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>NetApp</td>
<td>Q4</td>
<td>Q2</td>
</tr>
<tr>
<td>Blue Arc</td>
<td></td>
<td>Q3</td>
</tr>
<tr>
<td>DPM</td>
<td></td>
<td>Q4</td>
</tr>
<tr>
<td>StoRM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pNFS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Industry**
  - NetApp
  - Blue Arc

- **pNFS**
  - pNFS in official 2.6.38 kernel
  - pNFS in official 2.6.38 kernel for SL5

- **Beta read-only**
  - Next Golden Release (1.9.12)

- **GPFS native**
  - Production (OnTap 8.1)
  - GPFS pNFS

- **pNFS Enabled Kernel**
  - Linux distributions (RH6.2…)
  - DESY Linux pNFS kernel for SL5

---

Mar 21, 2011  Standardization in EMI Data, Taipei OGF 31  17
Performance and Stability

The DESY Grid Lab
Operated by
Yves Kemp
Dmitri Ozerov

DESY Grid Lab available for more than 9 months to evaluate protocols and systems.

Cern NFS 4.1 evaluation setup
Contact
Andrea Sciaba

Has been enlarged and will start testing end of March’11
Interoperability between EMI components and

- Integration of ARGUS (Blacklisting)
- GLUE 2.0 migration
- UNICORE client integration
Interoperability

ARGUS integration

✓ ARGUS: Authorization system
✓ Allows local and centralized configuration (and both)
✓ SE’s, LFC and FTS starting with ARGUS blacklisting

Migration

✓ Already agreed before EMI
✓ Starting with GLUE 1.3 information published with GLUE 2.0
✓ Followed by clients (between EMI-1 and EMI-2)
✓ Running both systems in parallel for awhile.
UNICORE integration

UNICORE access to EMI-Data components

SRM client interface
LFC Client interface
AMGA client

EMI-1  EMI-2  EMI-3
Unicore  ✔  ✔

Mar 25, 2011  EMI Data, ... ISGC 2011, ASGC, Taipei, TW
Standards : SRM over SSL

Replacing SRM over httpg (GSI) by https

Stolen from activity leader : Paul Millar

- GSI (httpg) is not a standard (SSLv3 is)
- Coupling delegation with transport negotiation is inflexible.
- It’s not widely used outside of Grid.
- Libraries are coming from a single vendor : GLOBUS
- E.g. no hardware acceleration.
- Delegation has to be solved. (see “spin off” standards)
- Prototype (server and client) w/o delegation available in EMI-1 for dCache.
EMI involvement in OSG (wg)

- Storage accounting record WG
- Delegation
Storage Accounting Record (StAR)

Stolen from activity leader: Jon Kerr Nielsen

- If there is no standard for storage accounting yet, do anyone actually need it?
- Developing a storage infrastructure
  - We need to know how much storage space is used, by which group/user, on which storage media
  - To know where to put the money when increasing the storage space
  - To know who to ask for the money to increase the storage space
- Basis for billing
  - Storage is expensive
  - Some non-academic resource owners may not like to give it away for free
- Jon prepared a draft definition of a StAR
- -> OGF existing working group (Jon becoming Co-Chair)
Delegation using SSLv3

Delegation

- SRM needs delegation to perform operations on behalf of the initiator. (e.g. 3rd party copy, bring online, ..)
- GSI (httpg) allows delegation w/o additional services.
- For SSLv3 delegation is not defined.
- There are several options
- -> OSG working group (Forming a research group: Paul)
Fixing an intended design flaw

The (in)famous catalogue synchronization

• Catalogues and SE name-spaces get out of sync over the time. (They interact by non atomic operations)
• Mechanisms to get them synchronized exists but involve SE resp. catalogue dump and are painful and certainly don’t scale.
• The plan is to use message passing to synchronizes them in ‘real time’
• First step (EMI-1) allow to ‘register’ lost files manually.
• DPM and LFC interaction already done. Now working on experiment catalogues (ATLAS DDN).
• StoRM and dCache a bit behind.
Fixing an intended design flaw

Stolen from activity leader: Fabrizio Furano

Uses a virtual destination, e.g.:

Consumer.LFC1.SEMsg_upstream

the broker queues messages for this endpoint if it disconnects momentarily.

LFC

Fix info!

Downstream topic

“SEMsg_downstream”

Chmod(sfn)

Adapter

Msg brokers

Fix info!

Upstream topic

“SEMsg_upstream”

NotAvailable(sfn)

FileCreated(sfn)

Uses a virtual destination, e.g.:

Consumer.DPNS1.SEMsg_downstream

the broker queues messages for this endpoint if it disconnects momentarily.

Adapter

SE or other Catalogue

A file can be N/A if:
- it was requested to a DB that does not know it
- OR if it was requested to a GridFTP that does not find it (trickier)
Conclusions

- **EMI Data** is a good opportunity to get our storage management middleware into a maintainable shape.
- Standardization is the way to get broader acceptance by other communities.
- EMI-Data will become THE competitor in Storage Management in Europe.
- Everybody can join or may provide suggestions through WLCG or EGI.eu.
Further reading

https://twiki.cern.ch/twiki/bin/view/EMI/EmiJra1T3Data

EMI is partially funded by the European Commission under Grant Agreement INFSO-RI-261611