EMI Data, the unified European Data Management Middleware

Patrick Fuhrmann (DESY)

EMI Data Area lead

(on behalf of many people and slides stolen from all over the place)
Credits

- Alejandro Alvarez
- Alex Sim
- Claudio Cacciari
- Christian Bernardt
- Christian Loeschen
- Elisabetta Ronchieri
- Fabrizio Furano
- Giuseppe Fiameni
- Giacinto Donvito
- Giuseppe Lo Presti
- Jon Kerr Nilsen
- Jan Schaefer
- Jean-Philippe Baud
- Dmitry Ozerov
- Yves Kemp
- 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

• EMI, the facts
• EMI, the product
• EMI-Data, the components
• EMI-Data, the mission
• EMI-Data, selected Topics
  – Interoperability and reduction of components
    • Client library consolidation
  – Standardization
    • WebDAV
    • NFSv4.1/pNFS
EMI, the project

EMI, the facts
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
The last Decade in Europe (HTC)

Infrastructure
- EDG
- EGEE I
- EGEE II
- EGEE III

Software development and support
- EGI
- EMI
- StratusLab
- EDGI
- IGE
- Venus-C
- Siena

Coordination

May 25, 2011
EMI Data, IEEE and MSST, Denver
StratusLab is developing and deploying cloud technologies with the aim of simplifying and optimizing the use and operation of distributed computing infrastructures such as the European Grid Infrastructure (EGI).

VENUS-C is focused on a reliable, industry-quality, sustainable platform: letting scientists be scientists and supporting small & medium enterprises.

SIENA will support Europe’s Distributed Computing Infrastructure (DCI) initiatives and the European Commission in working towards the delivery of a future e-Infrastructures roadmap that will be aligned with the needs of European and national initiatives.

Desktop Grids: EDGI will develop DG-Cloud bridge middleware with the goal to get instantly available additional resources for DG systems if the application has some QoS requirements that could not be satisfied by the available resources of the DG system.

IGE wants to knit a tight European network between the European Globus developers and users, thus ensuring a fast response time to European user requests and the provision of up-to-date information to the European developers of the European user requirements.
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

See presentation by Maria

ESFRI Strategy Forum VRC Virtual Research Communities

SLAs & Support
Requirements
Releases

Collaborations

EUROPEAN MIDDLEWARE INITIATIVE

Distributed Computing Infrastructure (DCI)

StratusLab Venus-C SIENA EDGI IGE

Standards Industry Google

May 25, 2011
EMI Data, IEEE and MSST, Denver
EMI, the product
EMI Middleware Evolution

Before EMI

3 years

After EMI

Applications Integrators, System Administrators

Specialized services, professional support and customization

EMI Reference Services

Standards, New technologies (clouds) Users and Infrastructure Requirements

Stolen from Alberto Di Meglio
The EMI Middlewares

Computing  Security  Storage  Clients

ARC

GLite

UNICORE

dCache

May 25, 2011  EMI Data, IEEE and MSST, Denver
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

Stolen from Alberto Di Meglio

May 25, 2011  EMI Data, IEEE and MSST, Denver
EMI, the components
The EMI Pie

63 components and about 350 packages

UAS-Compute
A-REX
MPI
WMS
ARGUS
VOMS
UNICORE-Gate
gridSite
UNICORE-SMS
StoRM
dCache
Arc-libs
FTS, LFC, DPM, GFAL
Information system
accounting
bookkeeping
What does EMI-Data provide?

The EMI-Data shopping basket

CART
The EMI shopping cart

Reliable File
Transport Service

Professional Storage Solutions

Fits all size

dCache

DPM

File Location and meta data Service (LFC)
EMI, the storage elements

!!! Something for everyone

- SRM
- Disk Storage Layer
- dCache
- StoRM
- DPM
- GPFS
- TSM

- 100 PBytes world-wide
- Most WLCG Tier I’s
- Holds 50% of WLCG data
- Featuring
  - File replication on hot-spot detection
  - Draining of pools
  - Resilient dataset management
  - Replication on arrival

- In use at the Italien Tier I plus about 40 Tier II’s
- Makes use of features of underlying storage system (GPFS, Lustre …)

- Easy to install
- Very little maintenance
- Majority of WLCG sites
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
- EGI : Attracting resp. enabling new communities.
- **Becoming competitive and attractive by :**
  - Standards
  - Professional Support
  - Strict quality monitoring
Some more examples

- Defining (with OGF) and implementing an Storage Accounting Record
- Migrating the security of the Storage Resource Manager protocol from GSI (httpg) to standard SSL/X509.
- Fixing the catalogue synchronization problem.
- Migrating to next version of the information provider schema GLUE2.0
- Improving the File Transfer Service by integrating the load of the network and the storage element backend.
- For the entire list, have a look at :
  - https://twiki.cern.ch/twiki/bin/view/EMI/EmiJra1T3Data
EMI, some selected topics
Component consolidation

ARC Data Lib

gLite Data Lib

UNICORE Data Interface

Storage Interface (SRM, LFC)

dCache

DPM

File Catalogue

May 25, 2011
EMI Data, IEEE and MSST, Denver
Component consolidation

**Diagram:**

- **ARC**
- **dCache**
- **Data Interface**
- **EMI Data Lib**
- **EMI jData Lib**
- **GLite**
- **Storage Interface (SRM, LFC)**
- **StoRM**
- **DPM**
- **File Catalogue**

**Notes:**
- May 25, 2011
- EMI Data, IEEE and MSST, Denver
My preferred topic

Standards
Standardization

Applying industry standards

Standards are the key for sustainability of Open Source Projects

EMI in terms of data access and control:

- WebDAV
- Posix file system: NFSv4.1 / pNFS
- (SRM security, https instead of httpg)
Standardization: WebDAV

WebDAV

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

With EMI-2 (mostly already with EMI-1) we provide WebDAV support from our SE’s
Another standard

NFS v4.1 / pNFS
What’s NFSv4.1/pNFS?

CITI, at the University of Michigan, is funded by major storage providers to coordinate the pNFS effort and provide reference implementations.

Industry Support - Implementations

- **Clients**
  - Linux
  - Sun (Solaris)

- **Servers**
  - Desy
  - EMC
  - IBM
  - Linux
  - NetApp
  - Panasas
  - Sun (Solaris)

Several other implementations have been tested at Bake-a-thons and Connectathons

Group meets three times a year to check interoperability.
How does it work?

Stolen from: http://www.pnfs.com/
Why would one need it?

Benefits of Parallel I/O

✓ Delivers Very High Application Performance
✓ Allows for Massive Scalability without diminished performance

Benefits of NFS (or most any standard)

• Ensures Interoperability among vendor solutions
• Allows Choice of best-of-breed products
• Eliminates Risks of deploying proprietary technology

Stolen from:
http://www.pnfs.com/
Why would we need it?

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

May 25, 2011
Why so interesting for EMI-Data

The NFS 4.1/pNFS design is a one to one fit with dCache and DPM design.
NFS v4.1 / pNFS availability

✓ EMI server
  o dCache : production version with EMI 1
  o DPM : prototype, ready for EMI-2

✓ Linux Kernel
  o Completed in 2.6.39
  o Back-port of pNFS into RH 6.2

✓ Industry
  o NetApp OnTab 8.1
  o Other vendors : code ready but not officially available
NFS 4.1 / pNFS is a great opportunity for Open Source Projects (EMI) to compete with industry and of course the other way around.
Results: NFS 4.1 / pNFS

The DESY Grid Lab

Operated by
Yves Kemp
Dmitri Ozerov

DESY Grid Lab available for more than 9 months to evaluate protocols and systems. Publications at CHEP and HEPIX

Characteristics
• 32 nodes = 265 cores = Small Tier II
• 1 GB resp. 10 GB network
• 80 TBytes in 5 pools
• Real Compute Element with pNFS dCache storage element
• Realistic conditions, various tests (applications).

Results
• Published at CHEP and Hepix
• Extremely stable
• Performance identical to protocols currently in use in HEP
• Key to performance: client side caching.
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, which is especially important for EGI.
- EMI-Data will become THE competitor in Storage Management in Europe 😊.
- Everybody can join or may provide suggestions through EGI.eu. (next talk)
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