



**Data**

# *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 Carpena
- 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



## EMI Factsheet

Budget : about 24 Million Euros

Funding : about 50% by EU-FP7, rest by partners

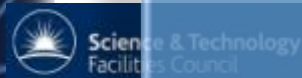
Covers : JRA, SA and NA

Partners : 22

Middlewares: Arc, gLite, UNICORE and dCache



UPPSALA  
UNIVERSITET



NIKHEF



TECHNISCHE  
UNIVERSITÄT  
DRESDEN

CINECA



CESCA

SWITCH

INFN



16/09/2010

EMI Overview - EGI TF, Amsterdam

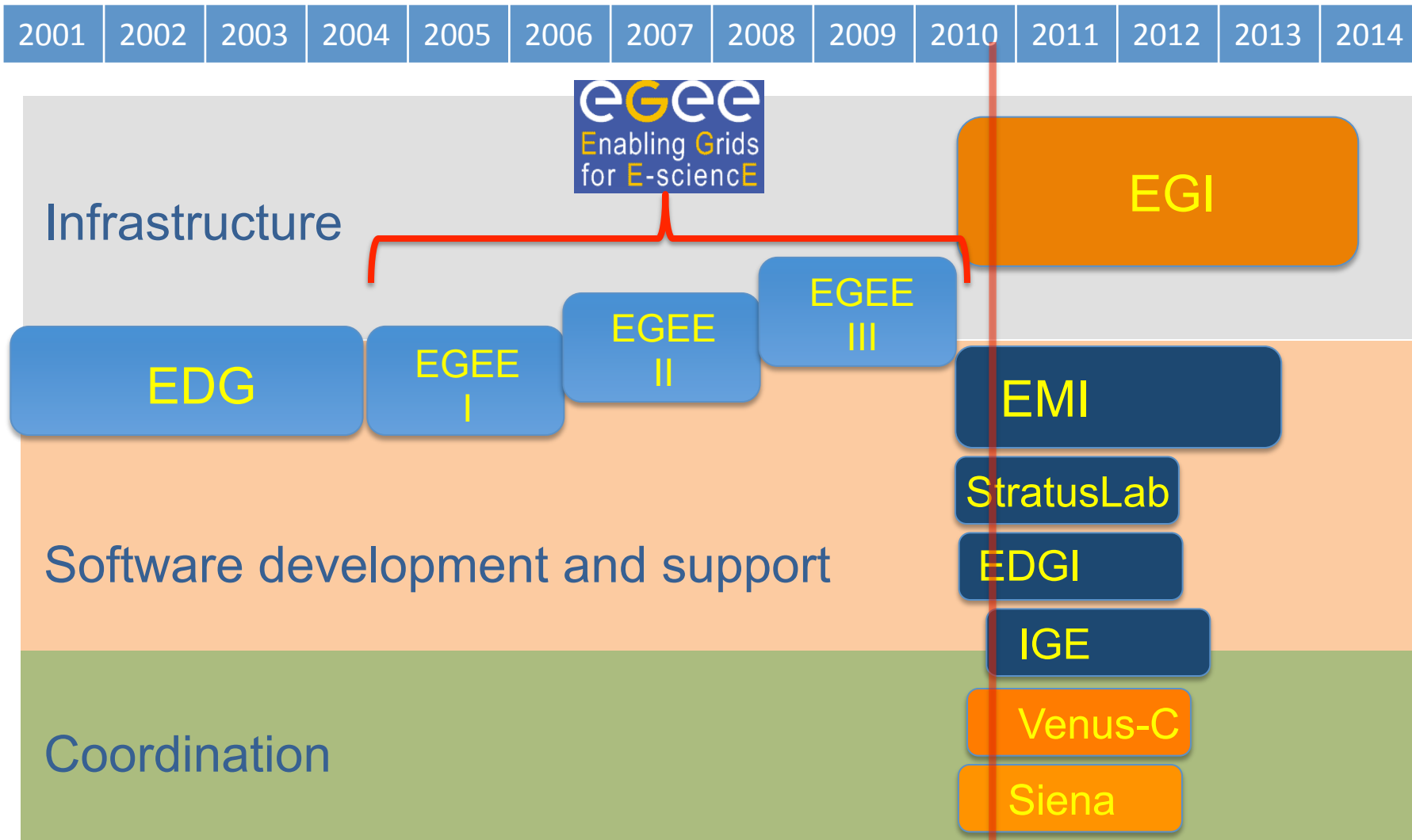


May 25, 2011

EMI Data, IEEE and MSST, Denver



# The last Decade in Europe (HTC)



# Project details

## StratusLab

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).

[Statuslab.eu](http://Statuslab.eu)

## VENUS-C

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

[Venus-c.eu](http://Venus-c.eu)

## SIENA

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.

[sienainitiative.eu](http://sienainitiative.eu)

## EDGI

**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.

[Edgi-project.eu](http://Edgi-project.eu)

## IGE

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.

[Edgi-project.eu](http://Edgi-project.eu)

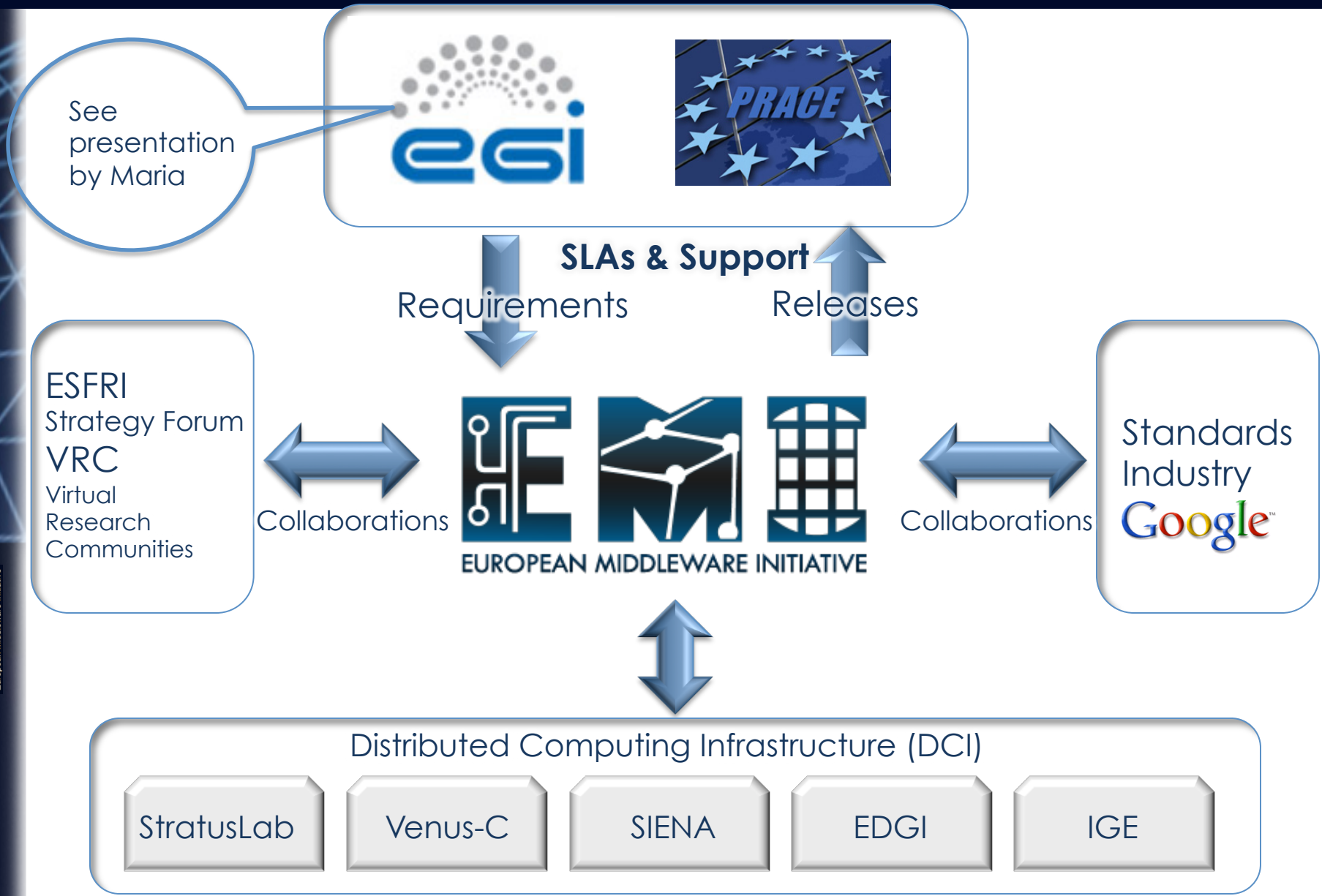
# Mantra

*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



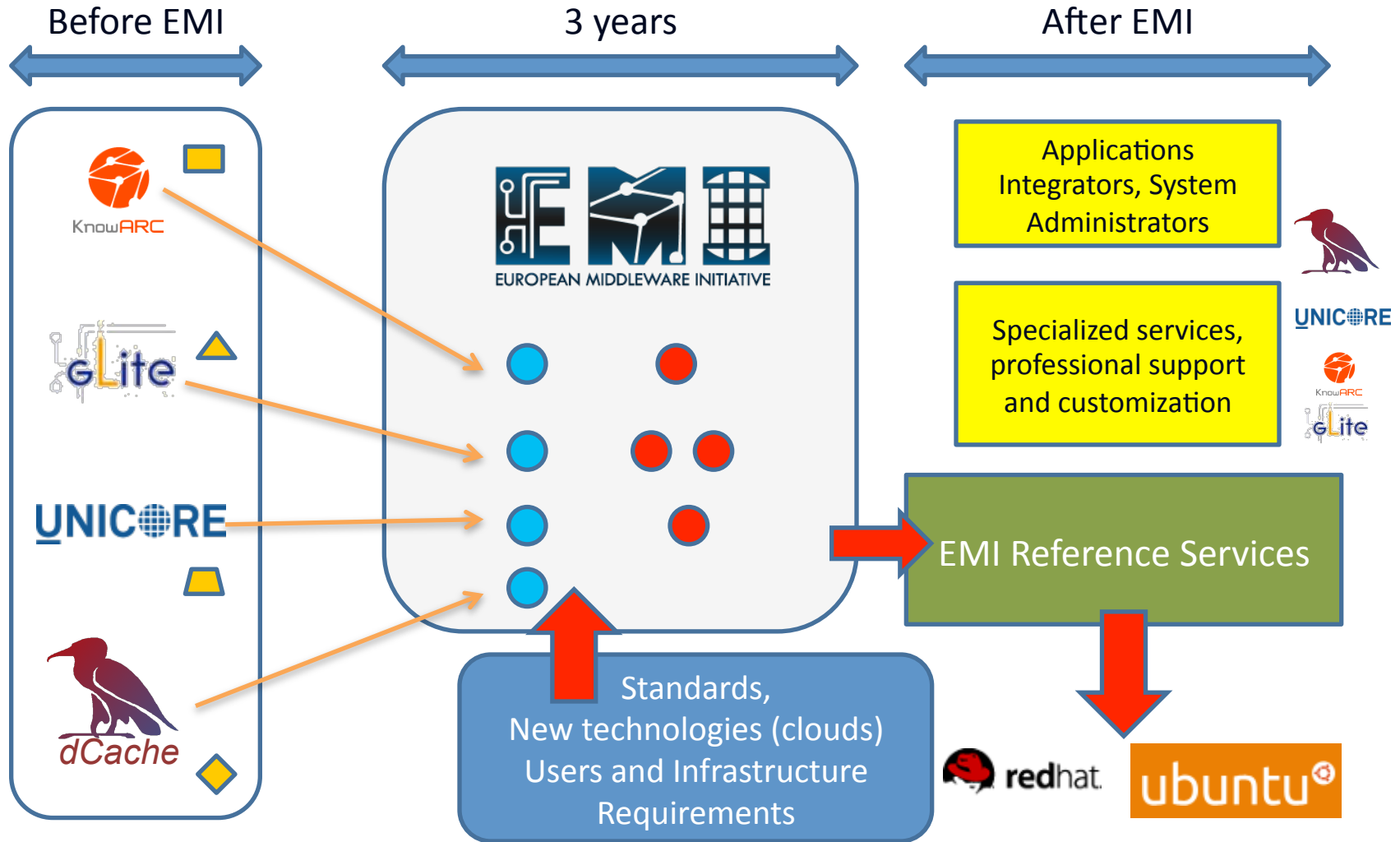
# EMI in context



# EMI, the product

# EMI Middleware Evolution

Stolen from Alberto Di Meglio



# The EMI Middlewares

Computing

Security

Storage

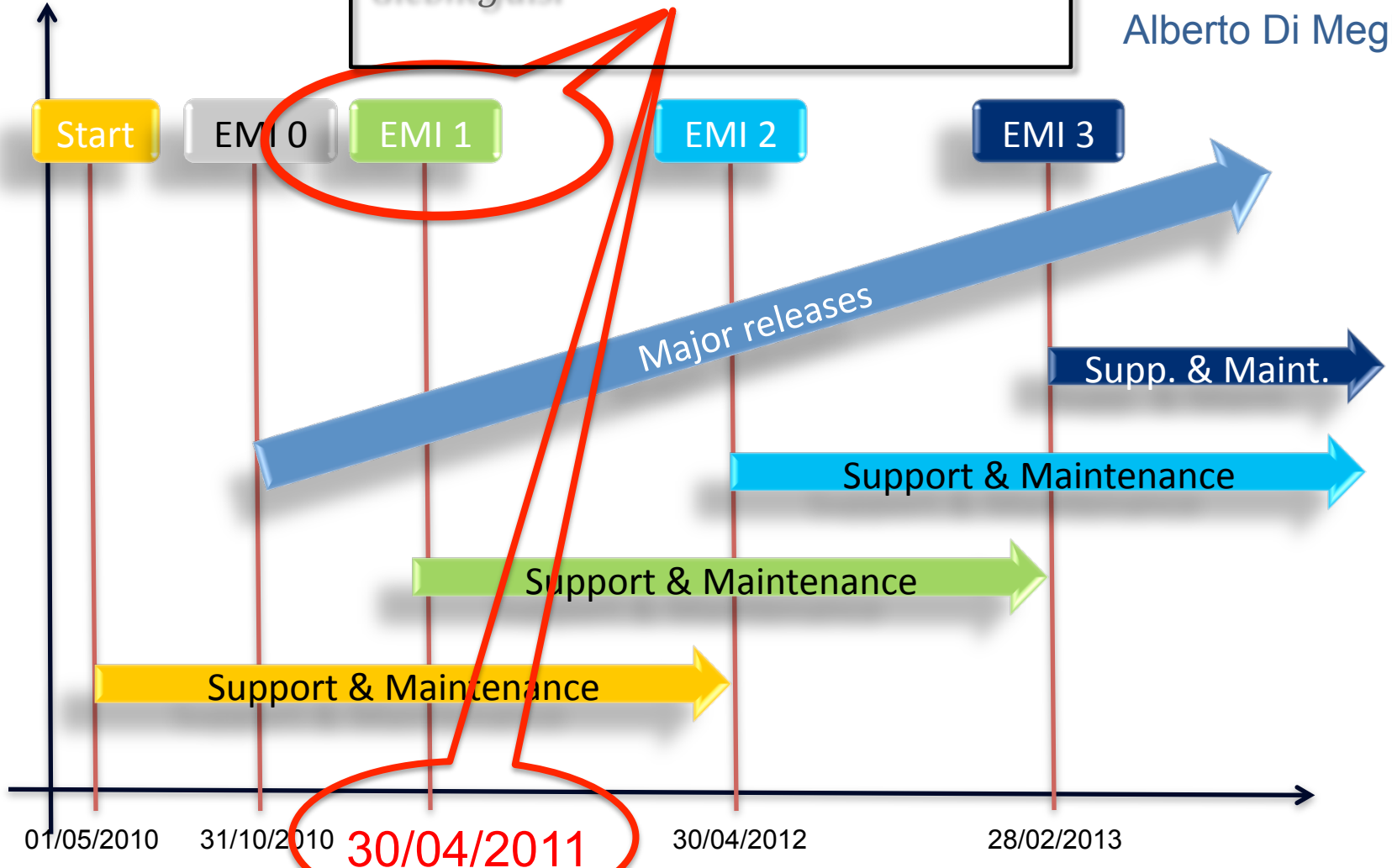
Clients



# Release Plan

Kebnekaise, Lappland, Sw, 2100m  
*Giebnegáisi*

Stolen  
from  
Alberto Di Meglio

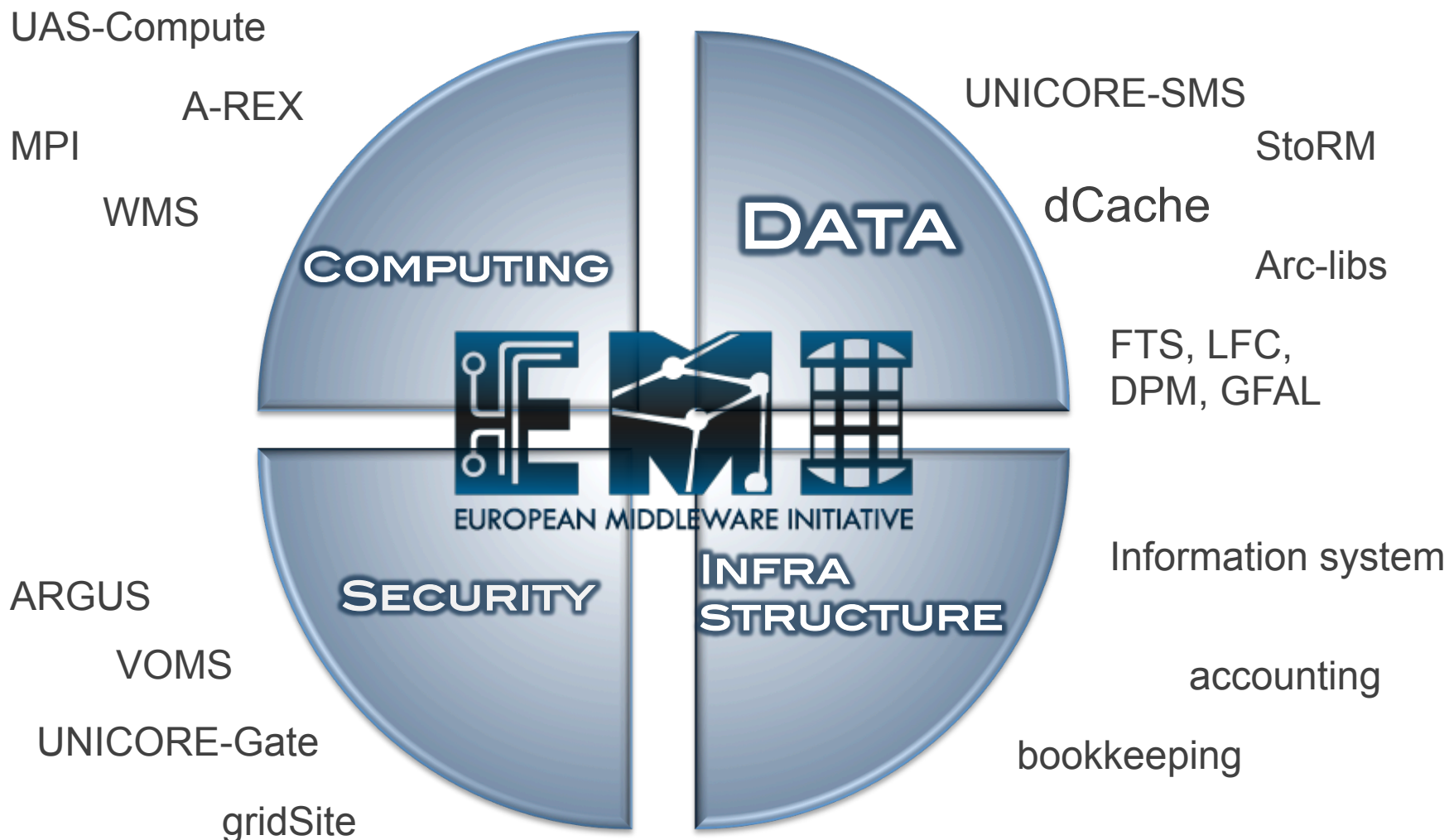


# Components

## EMI, the components

# The EMI Pie

## 63 components and about 350 packages



# What does EMI-Data provide ?

CART

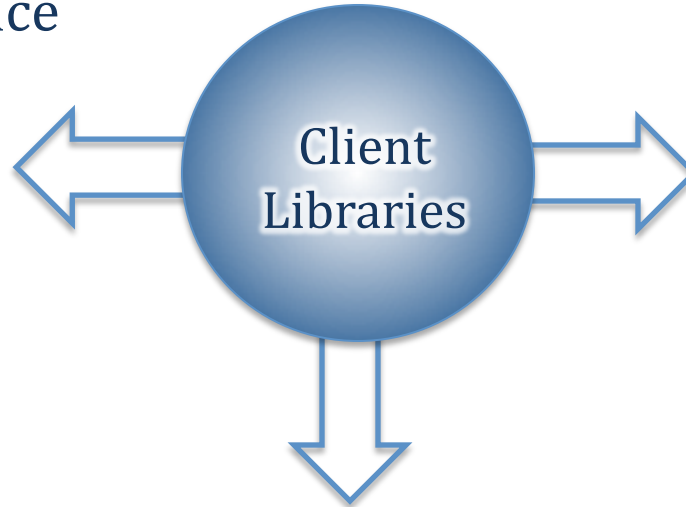
The EMI-Data shopping ~~basket~~





# The EMI shopping cart

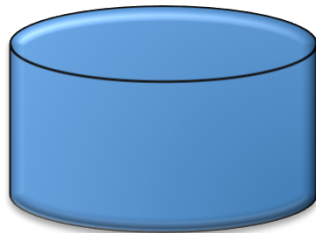
Reliable **F**ile  
Transport **S**ervice



File Location  
and meta  
data Service  
(LFC)

Professional Storage Solutions

*Fits all size*

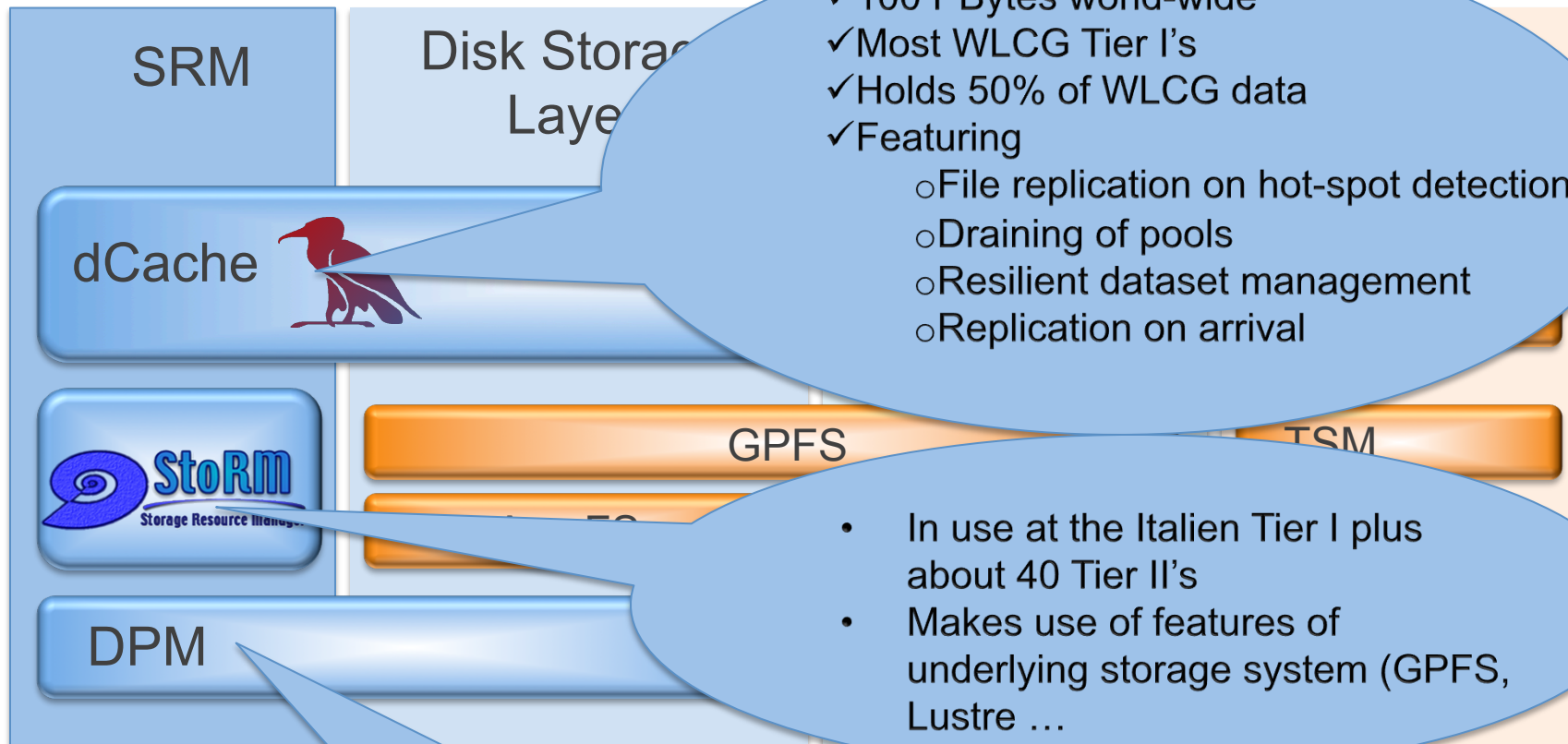


DPM



# EMI, the storage elements

!!! Something for everyone



- ✓ 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 Italian 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-wares.
- 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

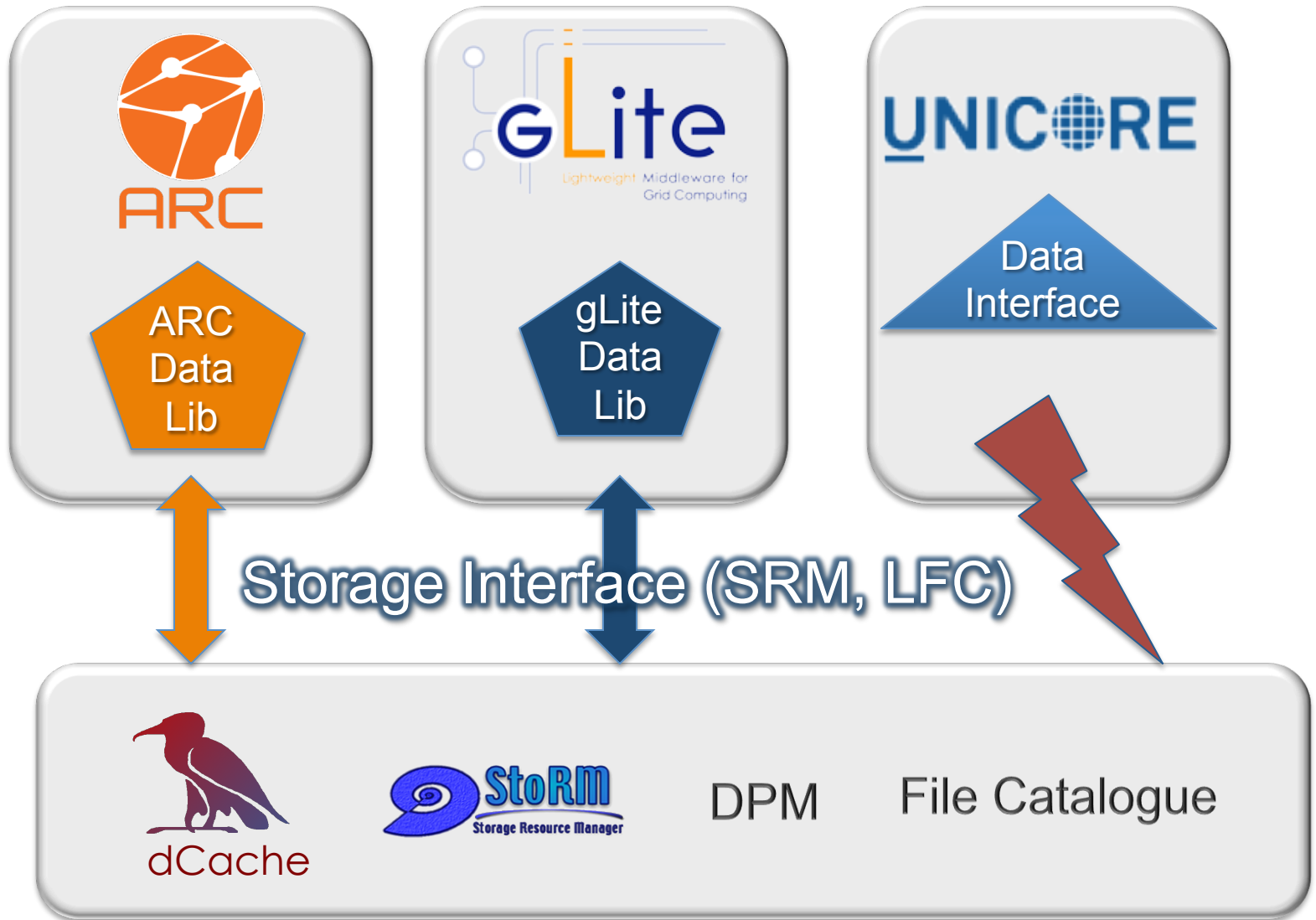
# Some more examples

- 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>

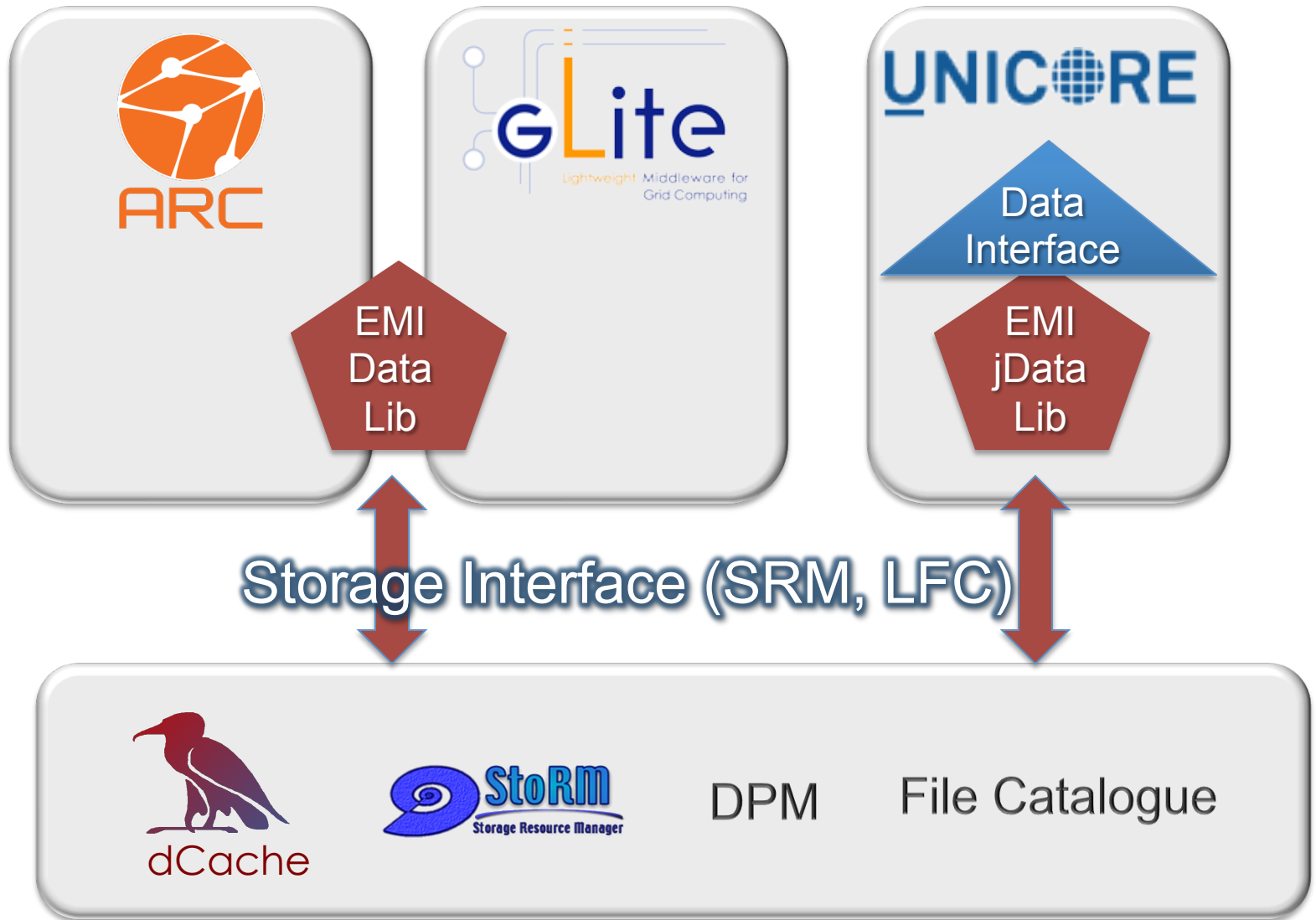
# Selected Topics

EMI, some selected topics

# Component consolidation



# Component consolidation





# Selected Topics

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



**webdav.dcache.org**  
Connected as: WebDAV

Disconnect

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

# Selected Topics

Another standard

NFS v4.1 / pNFS


# What's NFSv4.1/pNFS ?



center for  
information  
technology  
integration

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

## Industry Support - Implementations

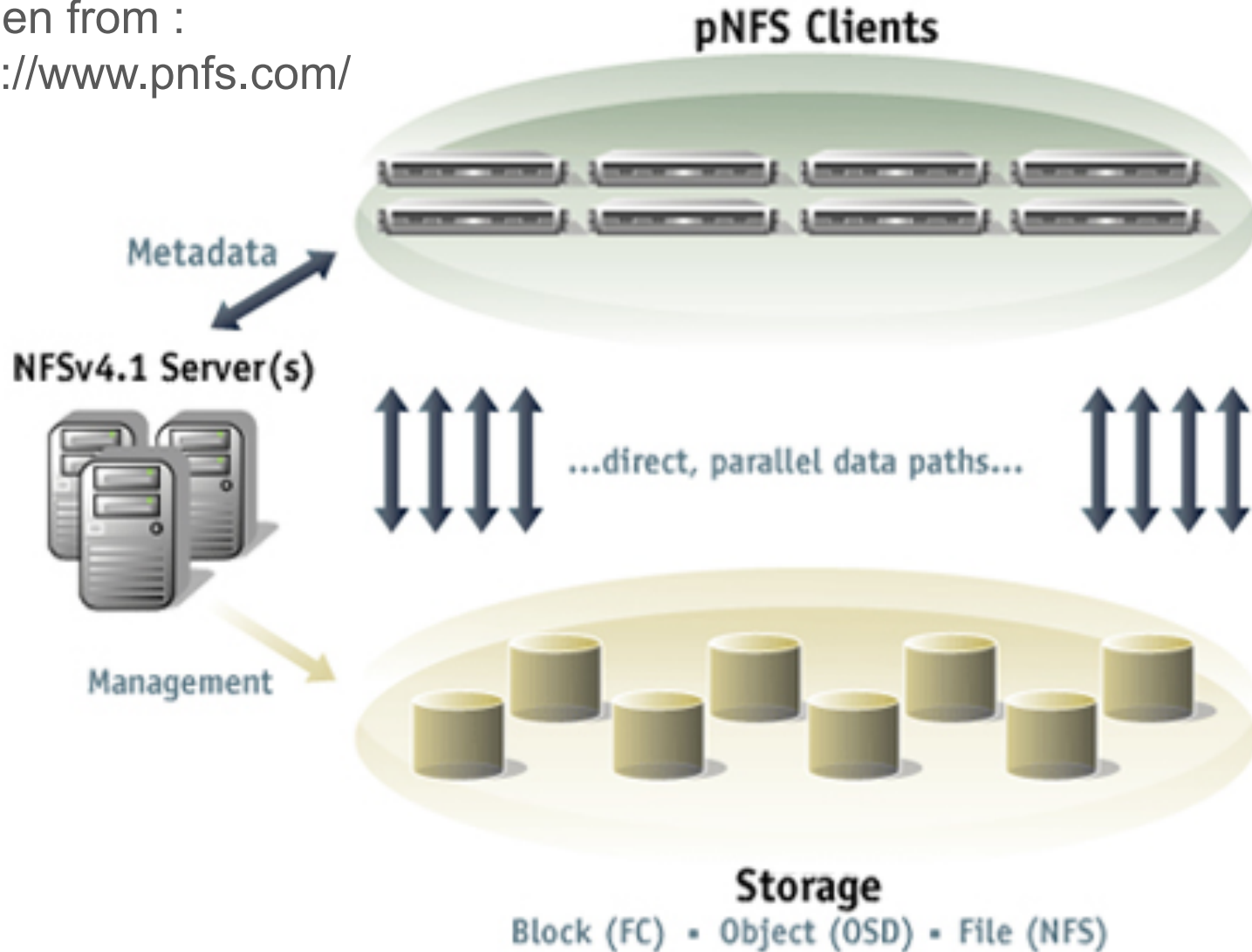
- Clients
  - Linux
  - Sun (Solaris)
- Servers
  - Desy 
  - EMC
  - IBM
  - Linux
  - NetApp
  - Panasas
  - Sun (Solaris)

Group meets three times a year to check interoperability.

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

# How does it work ?

Stolen from :  
<http://www.pnfs.com/>



# Why would one need it ?

Stolen from :

<http://www.pnfs.com/>

## 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

# 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

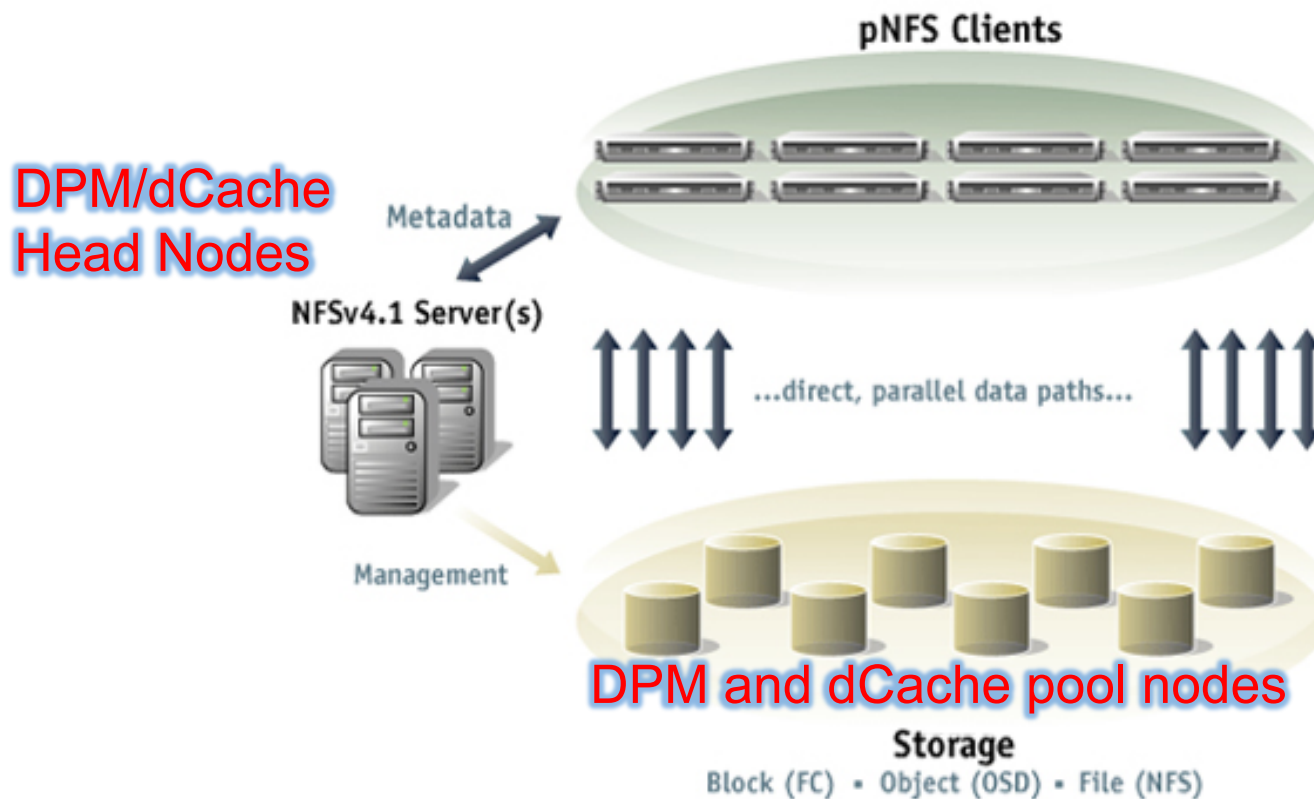
## Performance

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



# 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
  - dCache : production version with EMI 1
  - DPM : prototype, ready for EMI-2
- ✓ Linux Kernel
  - Completed in 2.6.39
  - Back-port of pNFS into RH 6.2
- ✓ Industry
  - NetApp OnTab 8.1
  - Other vendors : code ready but not officially available

# Competition

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