dCache.ORG

Storage at DESY and more

Patrick Fuhrmann et al.

and some slides stolen from
Tigran Mkrtchyan

additional funding, support or contributions by

dCache.ORG

Did you notice yesterday ?

dCache.ORG

Yesterday, 10:28 a proton bunch for the first time passed the entire 27 Km LHC tunnel.

This was the start of the largest experiment people ever made.

Hopefully not the last one.

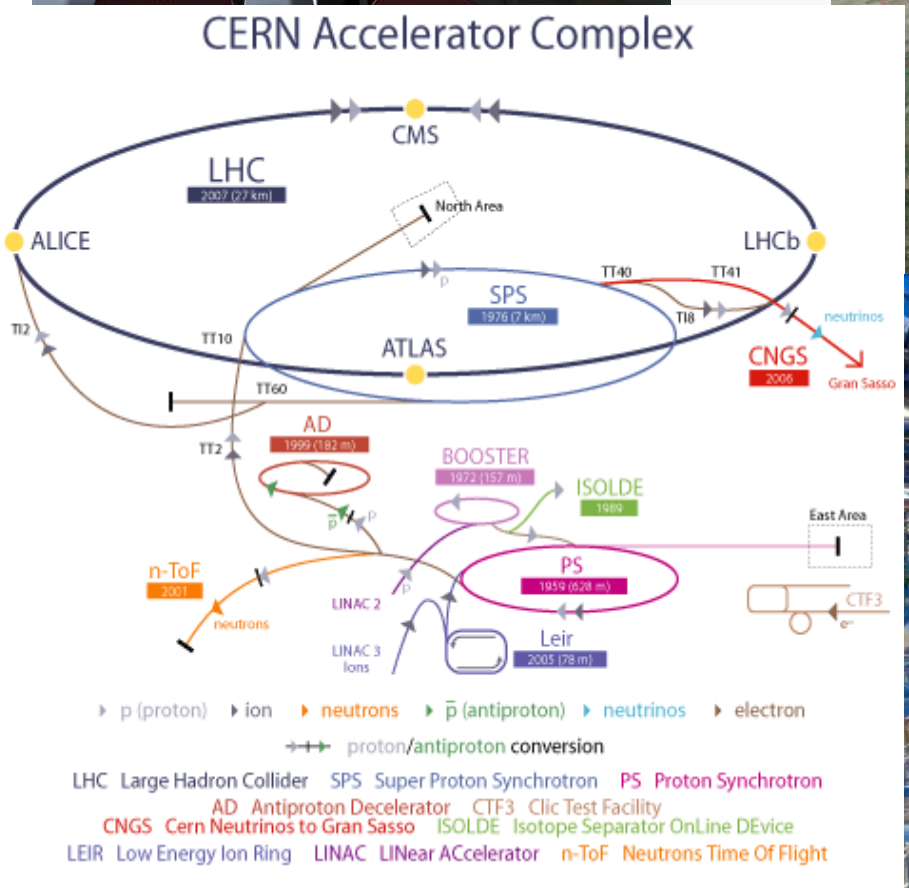
(I'm getting a lot of trouble saying this)

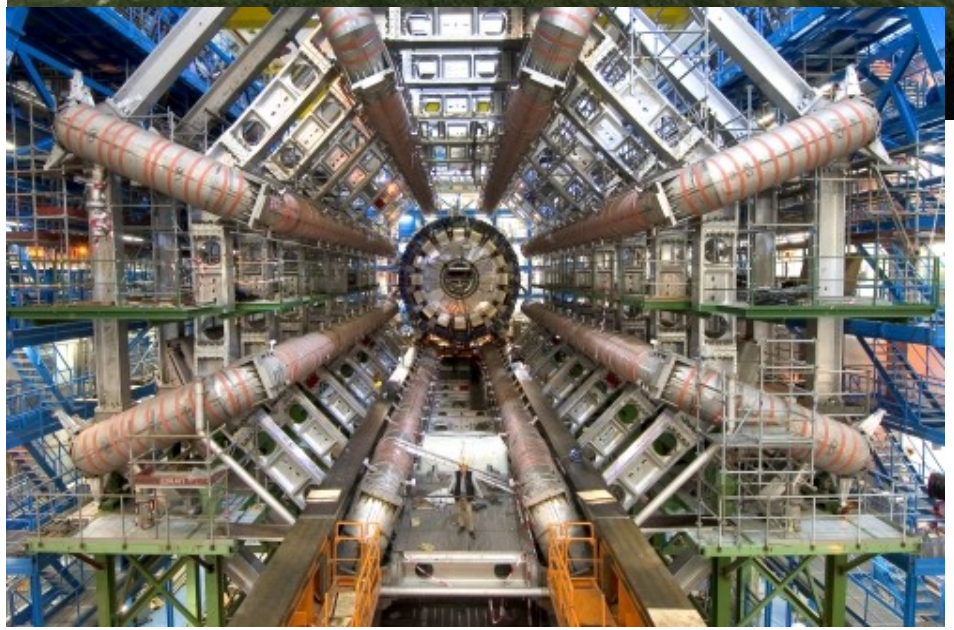
And you may ask : What does this has to do with this meeting ?

Good question.

The LHC Tier model and the Grid.

dCache.ORG





dCache.ORG

dCache.ORG

The DESY involvement in this historic event is two-folded.

DESY stores and processes LHC data for 3 of the 4 LHC experiments.

DESY is the leading force of an international collaboration, providing storage software, which holds the largest share of LHC data in 7 Tier I labs and more than 70 Tier II's.

dCache

dCache.ORG

Let us begin with the worldwide
LHC Computing Grid = (W)LCG

Or, to be slightly more specific :
The LCG data model.

Die obligatorischen Zitate

10 Sep 2008
(Yesterday)

Jamie Shiers back in 2005

When the LHC starts operating in 2007, it will be the most data-intensive physics instrument on the planet, producing more than ***1500*** megabytes of data every second for over a decade.

Bemerkung von Patrick am 11 Sep 2008

Wenn Sie das hier lesen, hat der LHC kein Schwarzes Loch erzeugt. (Oder nur ein ganz kleines)

dCache.ORG

Content

dCache.ORG

The LCG Tier model.

The Storage Element, workhorse of the LCG data model.

dCache

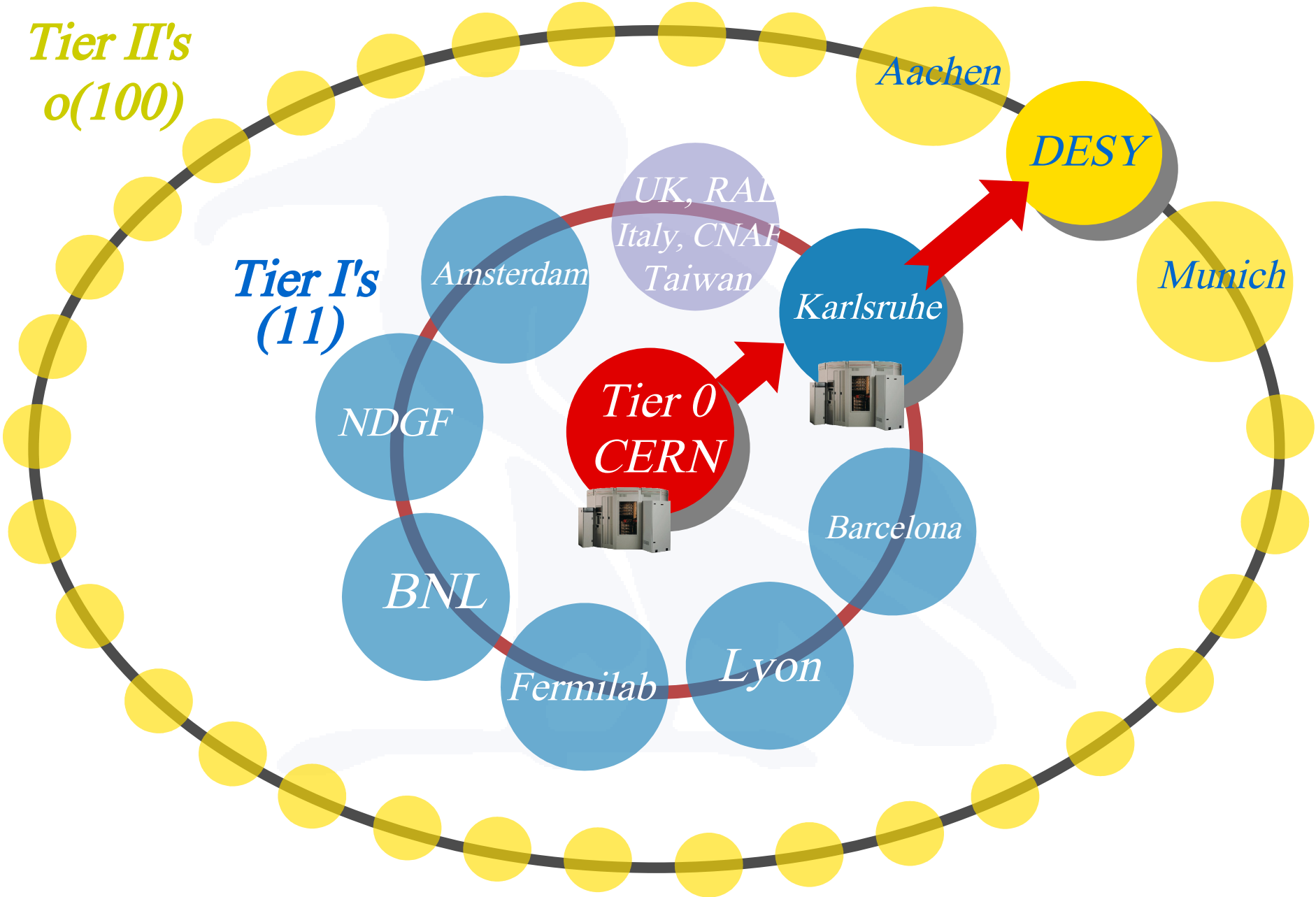
- The LCG Storage Element
- Project topology
- ... in a nutshell

The Nordic Data Grid Facility Approach

Once again : The LCG data model.

LCG Tier Model (data view)

dCache.ORG



The LHC Storage Element is defined by the protocols it supports

dCache.ORG

Remote Storage Resource Management

SRM (which is an Open Grid Forum Standard)

- Allocation of space
- Assigning space attributes (Tape, Disk, Offline ...)
- Transfer Protocol Negotiation
- Limited name-space operations

The LHC Storage Element is defined by the protocols it supports (cont.)

dCache.ORG

Wide Area Transport Protocol (IETF standards)

- gsiFtp
- http(s)

Local Access Protocol (Kein standard)

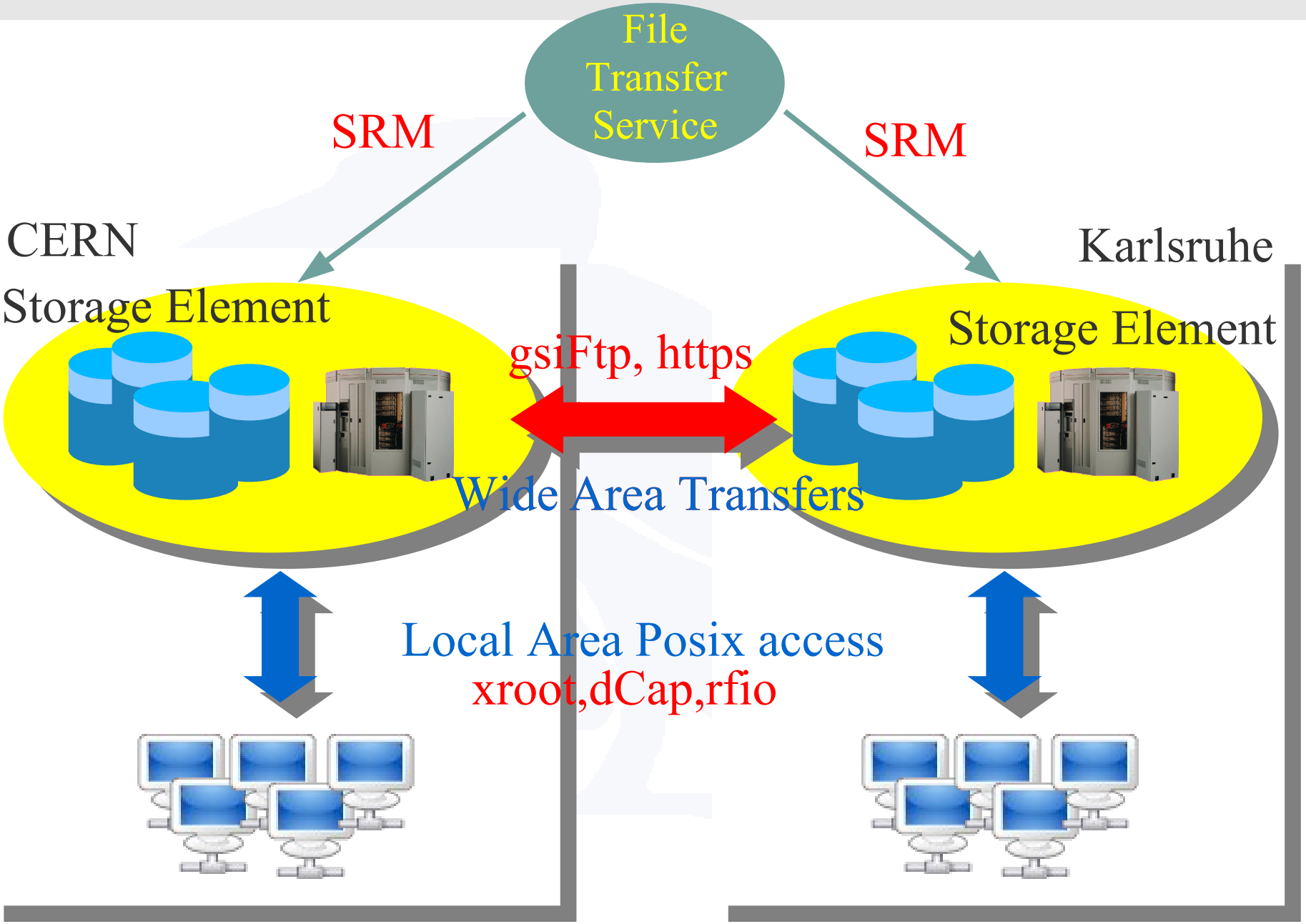
- (gsi)dCap
- rfio
- xRoot

Information Publishing

- Content : GLUE (OGF)
- Transport : LDAP

Storage Element interactions.

dCache.ORG



dCache.ORG

Hier kommen wir denn endlich zu dCache

dCache.ORG

Global Overview

What is a dCache

dCache.ORG

dCache is a software package.

dCache is an international collaboration.

dCache is developed at DESY, FNAL and NDGF.

dCache is packaged and managed at DESY.

dCache manages storage on ~1000 of disk nodes.

dCache drives HSM storage back-end(s).(eg Tsm,Hpss,...)

dCache is a (certified) LHC Storage Element.

dCache is in use at 7 Tier I's and > 70 Tier II's.

dCache will store the largest share of LHC data outside CERN.

dCache.ORG

Some technical details (in a nutshell)

dCache in a nutshell

dCache.ORG

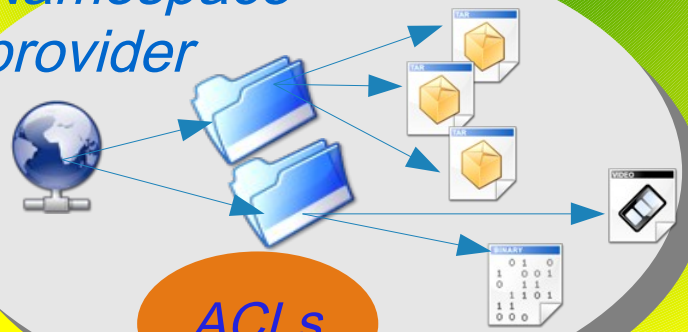
dCache.ORG

Tape Storage

OSM, Enstore
Tsm, Hpss, DMF

dCache Core

Namespace provider



ACLs

Protocol Engines

Information Protocol(s)

Storage Management Protocol(s)
SRM 1.1 2.2

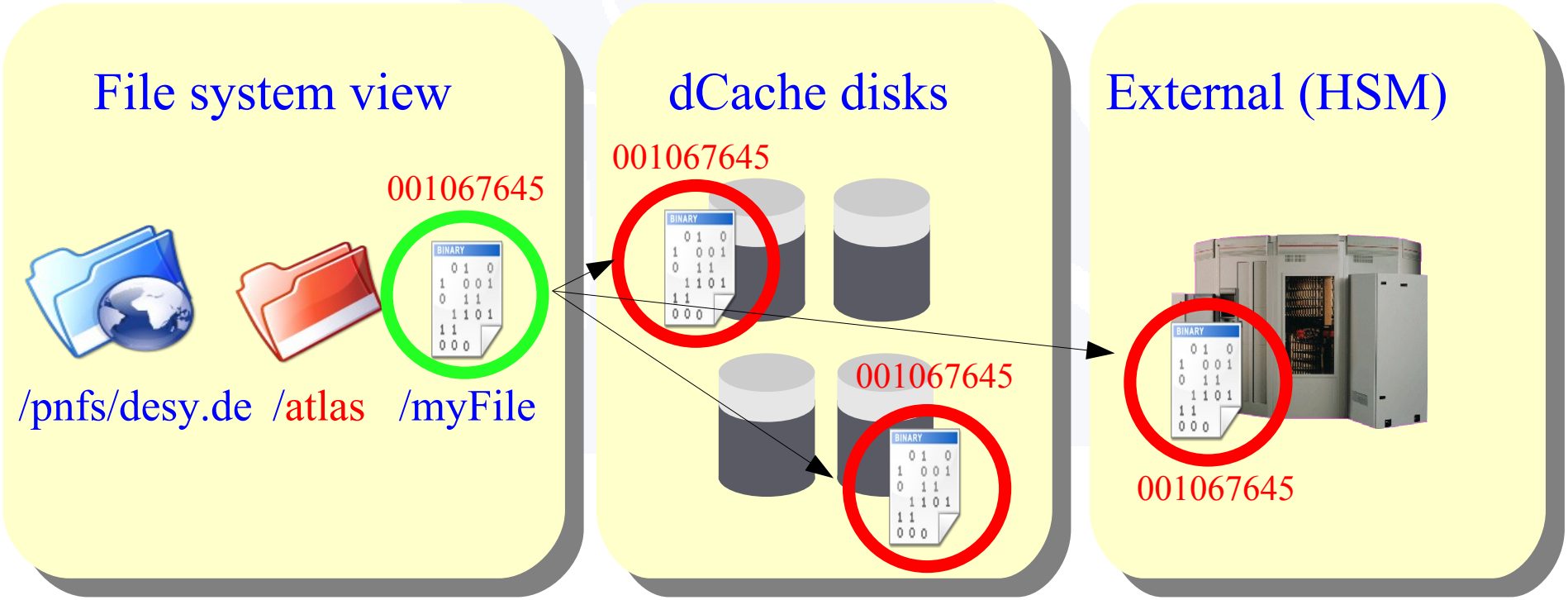
Data & Namespace Protocols
(NFS 4.1)
ftp (V2)
dCap
gsiFtp
xRoot
(http)

Namespace ONLY
NFS 2 / 3

dCache in a nutshell

dCache.ORG

- Strict name space and data storage separation, allowing
  - consistent name space operations (mv, rm, mkdir e.t.c)
  - consistent access control per directory resp. file
  - managing multiple internal and external copies of the same file
  - convenient name space management by nfs (or http)



dCache in a nutshell

dCache.ORG

- **Overload and meltdown protection**
  - Request Scheduler.
  - Primary Storage pool selection by protocol, IP, directory, IO direction
  - Secondary selection by system load and available space considerations.
  - Separate I/O queues per protocol (load balancing)
- **Supported protocols :**
  - (gsi)ftp
  - (gsi)dCap
  - xRoot
  - SRM
  - nfs2/3 (name space only)
  - nfs 4.1 very soon

dCache in a nutshell

dCache.ORG

- **File hopping on**
  - automated hot spot detection
  - configuration (read only, write only, stage only pools)
  - on arrival (configurable)
  - outside / inside firewalls
- **dCache partitioning for very large installations**
  - Different tuning parameter for different parts of dCache
- **Resilient Management**
  - at least n but never more than m copies of a file

dCache.ORG

dCache project topology

dCache.ORG

Development

Code management
System verification
Deployment

dCache.ORG

Open Science Grid

VDT Virtual Data Toolkit

US

gLite

Europe
Asia
South America

The Team

dCache.ORG

dCache.ORG

Head of dCache.ORG

Patrick Fuhrmann

Core Team (Desy, Fermi, NDGF)

Andrew Baranovski
Gerd Behrmann
Bjoern Boettscher
Ted Hesselroth
Alex Kulyavtsev
Iryna Koslova
Dmitri Litvintsev
David Melkumyan
Dirk Pleiter
Martin Radicke
Owen Synge
Neha Sharma
Vladimir Podstavkov

Head of Development FNAL :

Timur Perelmutov

Head of Development DESY :

Tigran Mkrtchyan

Head of Development NDGF:

Gerd Behrmann

External Development

Abhishek Singh Rana, SDSC
Jonathan Schaeffer, IN2P3

Support and Help

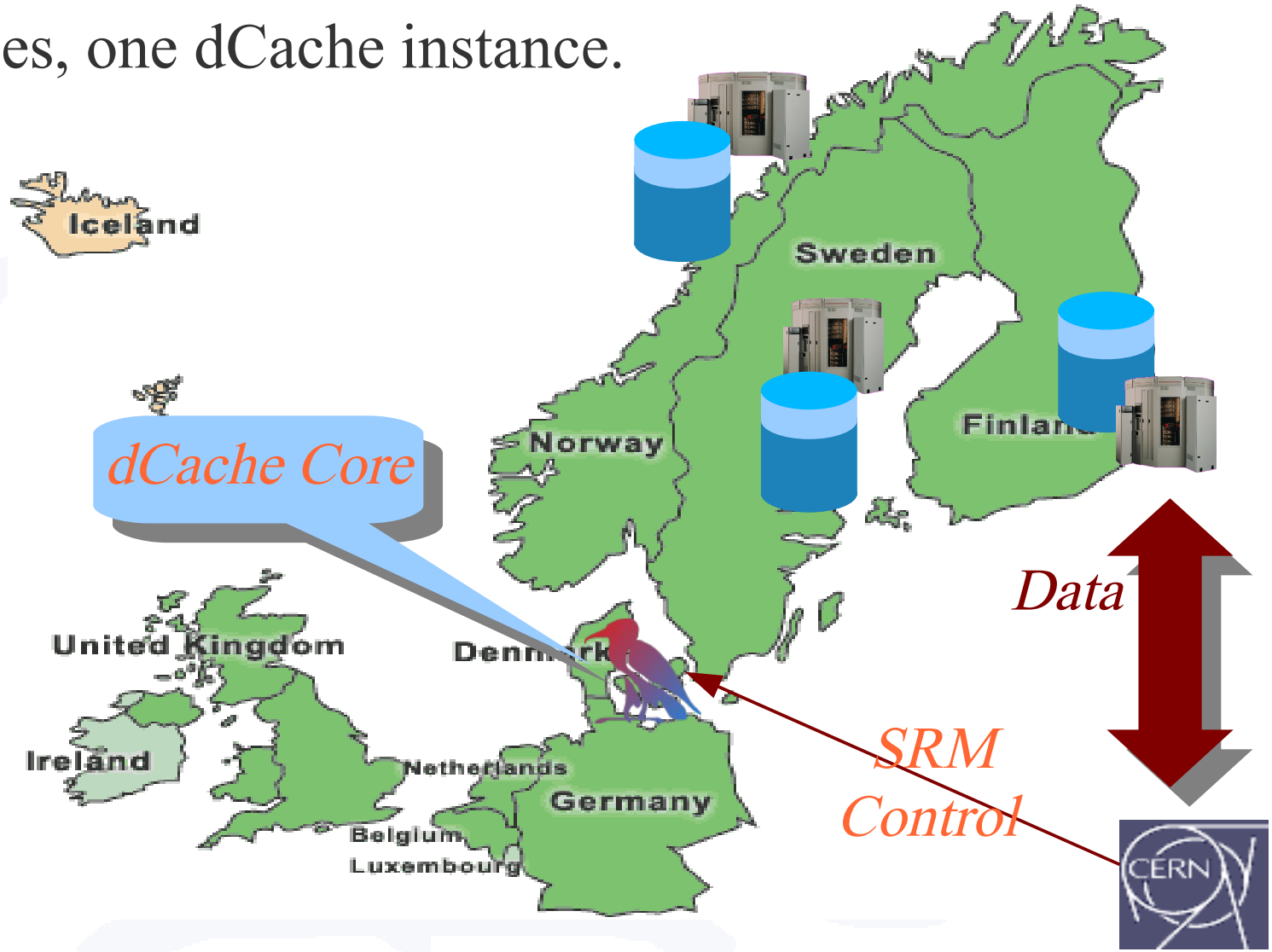
Greig Cowan, gridPP
Stijn De Weirdt (Quattor)
Maarten Lithmaath, CERN
Flavia Donno, CERN

dCache.ORG

The dCache Nordic Data Grid Facility Approach

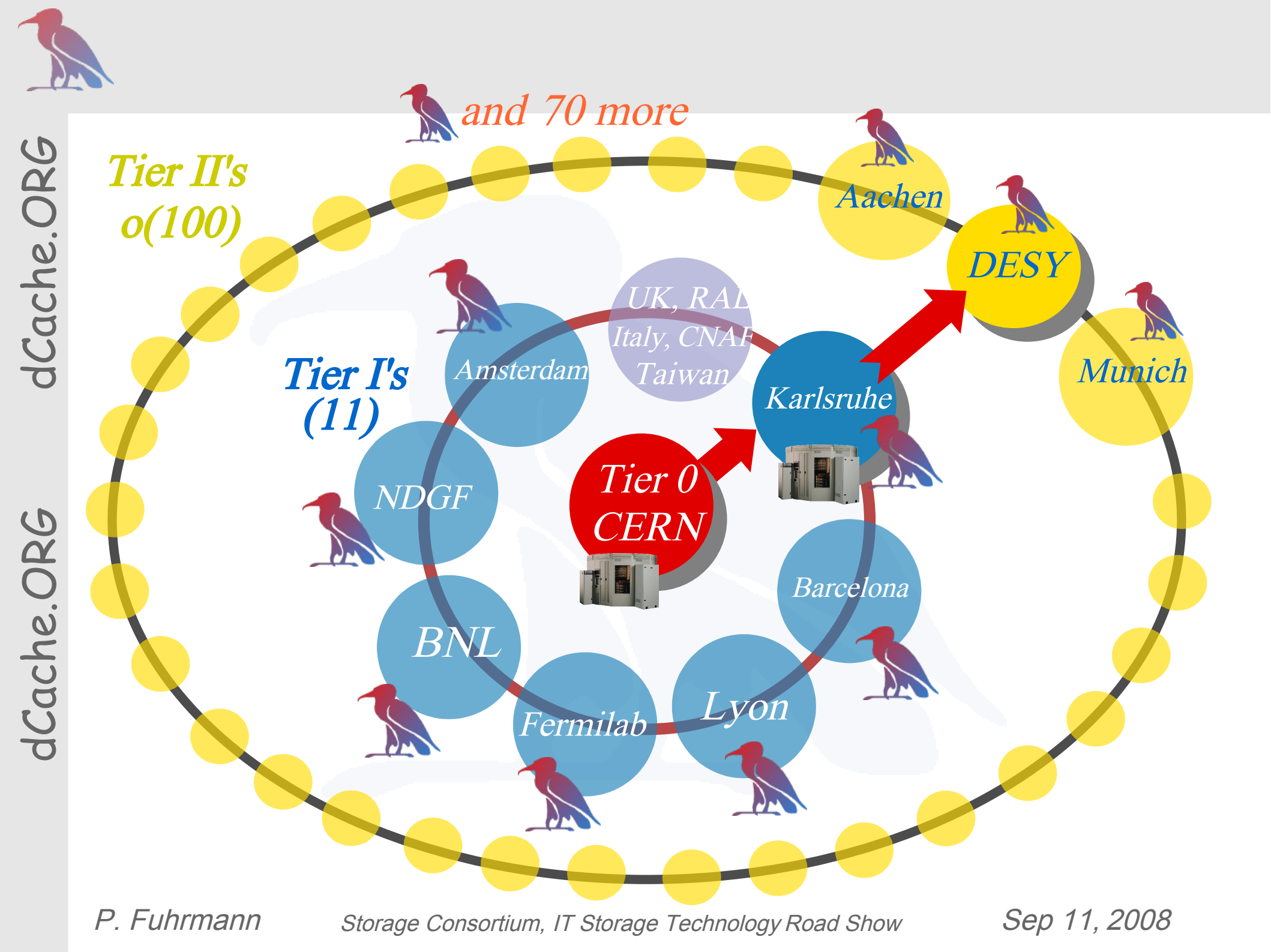
dCache.ORG

4 Countries, one dCache instance.



Idea :

At any time a country may 'go down' though raw data storage proceeds.


and 70 more
Tier II's
o(100)
Aachen
DESY
Munich
Tier I's
(11)
Amsterdam
UK, RAL
Italy, CNAF
Taiwan
Karlsruhe
NDGF
Tier 0
CERN
Barcelona
BNL
Fermilab
Lyon
dCache.ORG
dCache.ORG

dCache.ORG

Further reading

www.dCache.ORG

and of course enjoy your stay at desy and in hamburg.