A Light for Science

European Synchrotron Radiation Facility

The XRAY Grid

ESRFUP-WP11

Outline

- Our Project: ESRFUP-WP11
- Our Collaboration: XRAY VO
- Our Resources: Hardware and grid components
- Our Deployment : Grid services @ ESRF and partner sites
- Our Issues: Network
- Our Figures: Initial GridFTP transfer rates

ESRFUP-WP11

- ESRF is getting a Major Upgrade
  - Ambitious Program of 177M € over next 7 years
  - Budget decision had been taken only two weeks ago
- In preparation of this Upgrade Program
  - 13 Work packages already funded and ongoing
- Work Package 11: Grid Study and Test Bed
  - Funding for 50PMs over 18 months
  - Up to150k€ for hardware

The biggest incentive for the grid…
…is probably the data problem

- How to handle TeraBytes and possibly 10s of TBs per experiment?
- How will users access those data sets from home institutions distributed all over Europe?
- How to transfer, replicate and archive the data?
- How to guarantee security while allowing to share more widely between groups
- How best to manage the data analysis?

Current Grid Technology ...

- is able to leverage data analysis and curation of these large data sets
  - It provides highly scalable tools to store, archive, organize, as well as search, access and transfer data in a reliable way
- can make better use of existing compute cycles
  - offering potentially large amounts of resources for compute intensive applications
- comes with a proven security Infrastructure (GSI)
- provides the framework to foster collaboration within large virtual organizations by sharing the compute, storage and software resources and ultimately knowledge

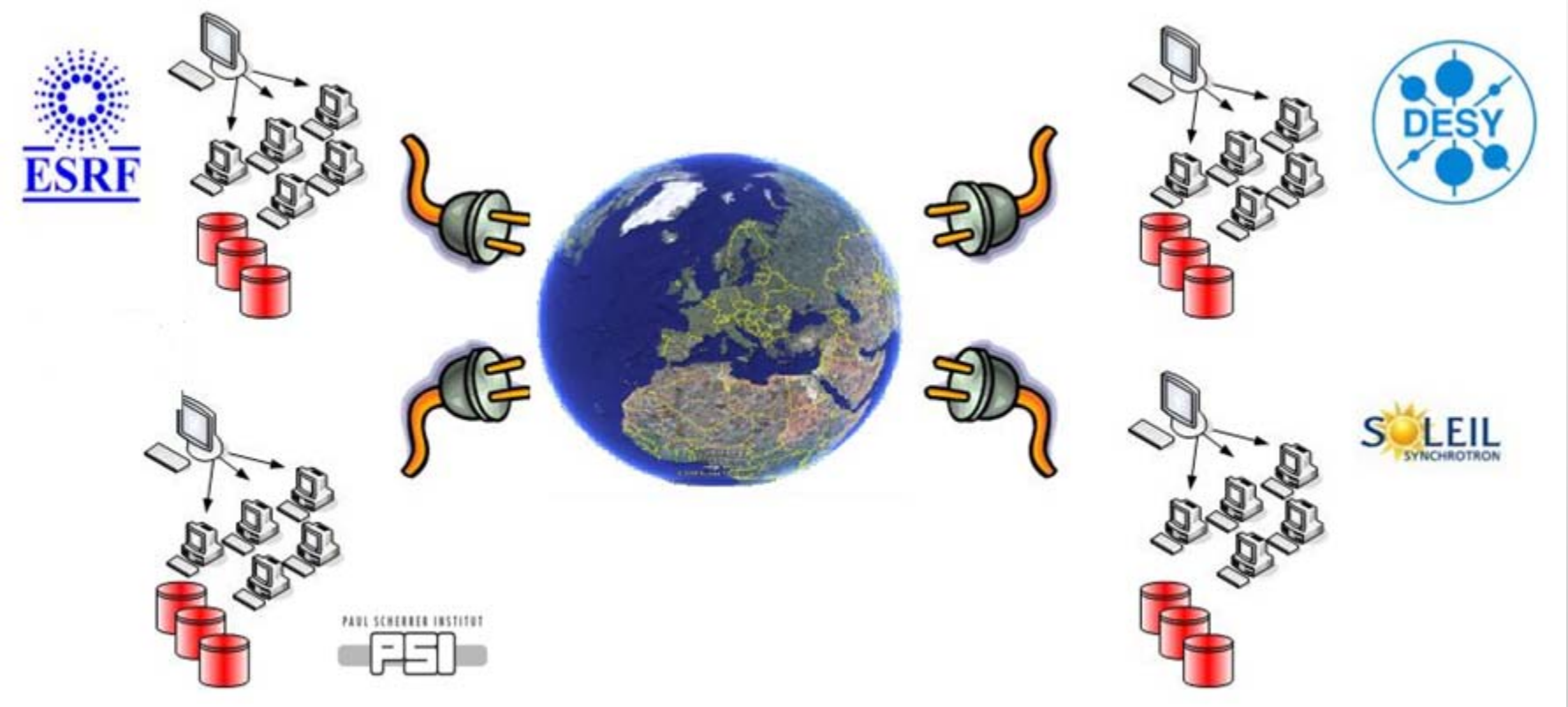
The deliverables of ESRFUP-WP11

- Deliverable: Collaboration Agreement between 3(-4) partner labs

for the creation of a Synchrotron Radiation Virtual Organization (Dec '08)

- Purchase compute and storage system, install and test gLite to form Grid Sites at the ESRF and 2(-3) other partner sites
  - Deliverable: ESRF Grid Site operational (Dec '08)
  - Deliverable: Partner Grid Sites operational (Jun '09)
- Organize a workshop with possible future partners
- Gridify one or two resource intensive applications on the test bed, write wrapper software, make added value analysis
  - Deliverable: Test case software operational (Oct '09)
- Final report on operational experiences with the international test bed installation including future orientations for photon science grid activities (Dec '09)

XRAY Grid Testbed of ESRFUP-WP11



Creation of a Synchrotron Radiation Virtual Organization

- The **XRAY VO** has been created
  - .. and registered with EGEE https://cic.gridops.org/index.php?section=vo&vo=xray.vo.eu-egee.org
- **V**irtual **O**rganization **M**anagement **S**ervice was set up
  - Enrolment URL:
    - https://grid-voms.esrf.eu:8443/voms/xray.vo.eu-egee.org
- Registered users are granted access to the resources of a VO according to their group membership and assigned role

Hardware specs

- Storage Components
  - Sun Fire X4500 Server (Thumper)
  - 2 dual core AMD Opteron, 16 GB
  - OpenSolaris based OS (SunOS 5.10)
  - 24 TB internal storage

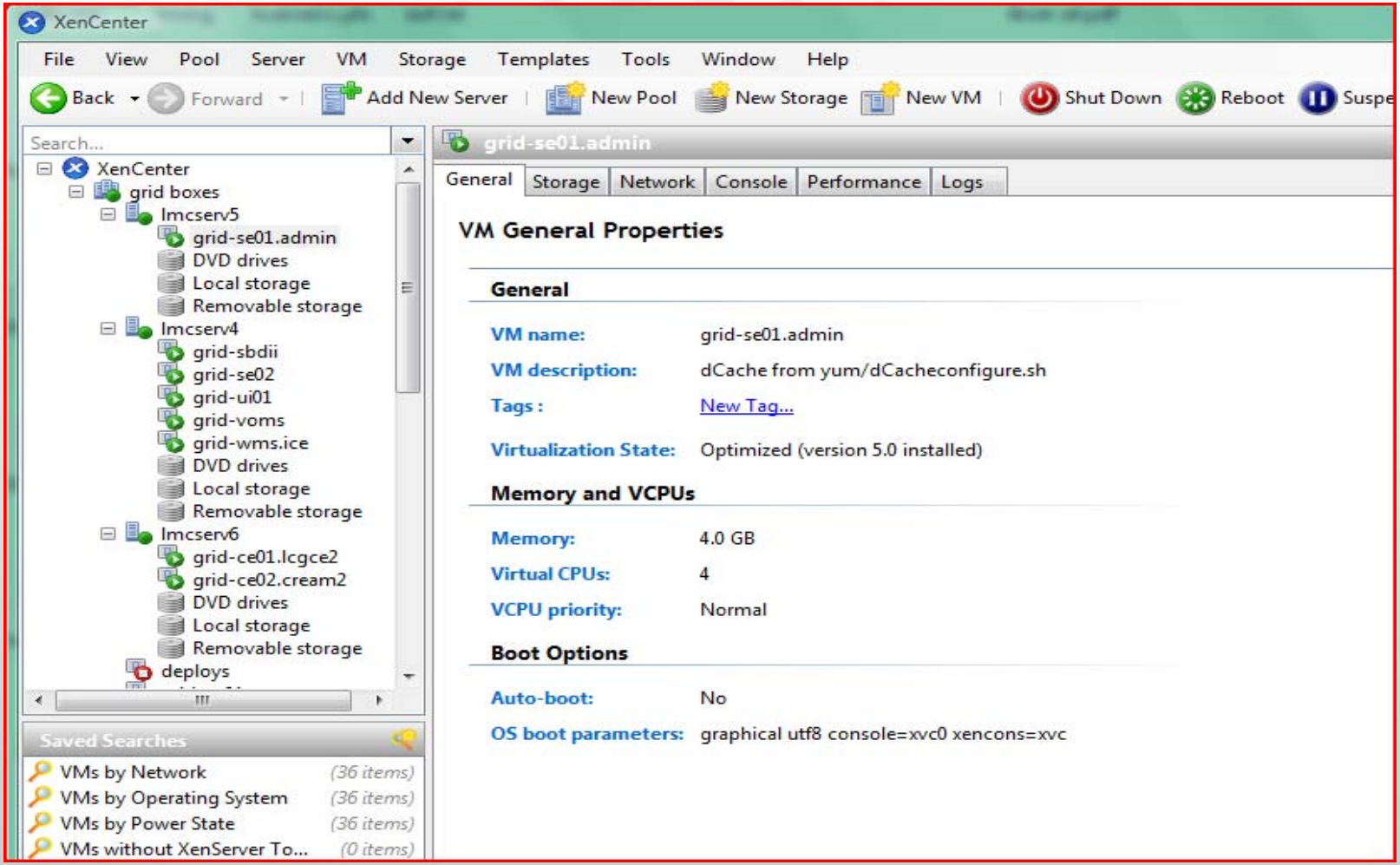
- Compute Nodes
  - Sun Fire X2200 M2 Server
  - 2 quad core AMD Opteron
  - 16 GB main memory, 250 GB HDD
  - Scientific Linux 4.7

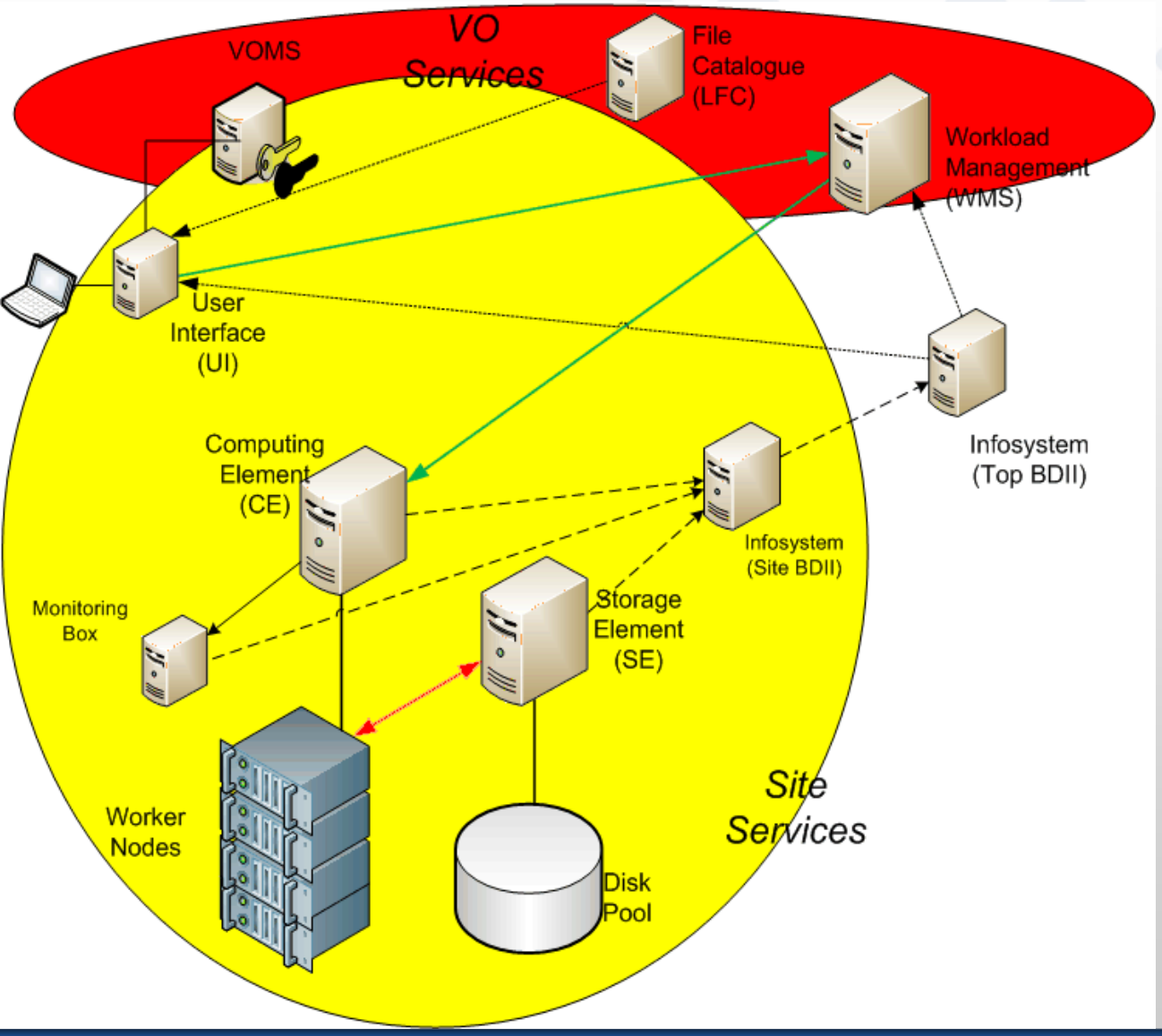
Hardware Specs

- Middleware servers
  - Sun Fire X4150
  - 2 x quad core Intel Xeon E5440
  - 16GB main memory, 2x73 GB HDD
  - Citrix XenServer 5.0 Enterprise Edition
  - Virtual Machines on Scientific Linux 4.7 (i386 or x86_64)

- Rack
  - Sun Rack 900-38
  - Network Switch, Extreme x450e-48p

Virtualization using Citrix XenServer 5.0 EE





Site and VO Services

The different choices

Computing Elements	
LCG-CE	**Cream CE**
Most widely deployed in EGEE	The next generation
Based on Globus gatekeeper	Based on web services

Storage Resource Manager	
Disk Pool Manager (DPM)	**dCache**
Most popular in EGEE	Popular with the largest Sites
Medium to high scalability	High scalability and versatility
Disk only storage	Disk storage + backend
Lightweight	High system complexity

XRAY setup @ ESRF

- **14 Worker Nodes** with altogether **80 Cores**
- **12 TB of disk** space (RAID-Z2)
- **2 Computing Elements** in Test Bed
  - **Lcg-CE** (3.1.21, 32bit)
  - **Cream CE** (1.8.4, 32bit)
- **2 Storage Elements**
  - **dCache** (1.8.0-15p12, 64bit, chimera)
  - **DPM** (1.6.11, 64bit, mysql)
- **1 Site BDII**
- **1 User interface** 32bit
- **1 VOMS** (v1.8.8, 32bit, mysql)
- *Soon 1 WMS*
- *Pending transition to new network setup*

Hardware to be set up @ partner sites

XRAY setup @ DESY

- DESY enabled XRAY on their existing grid site
- Shipped one thumper (**~20TB**) for their dCache pool

XRAY setup @ PSI

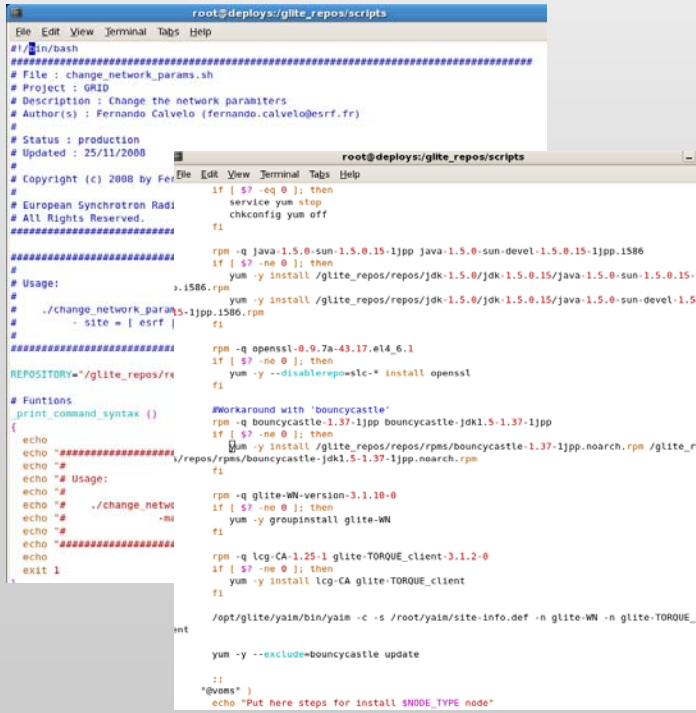
- 6 Worker Nodes with **48 Cores**
- Up to **20TB** of disk space
- 1 Computing Element: **lcg-CE**
- 1 Storage Element: **dCache**
- 1 User interface, Site-BDII and Monbox
- Cluster is currently on its way to PSI

XRAY setup @ Soleil

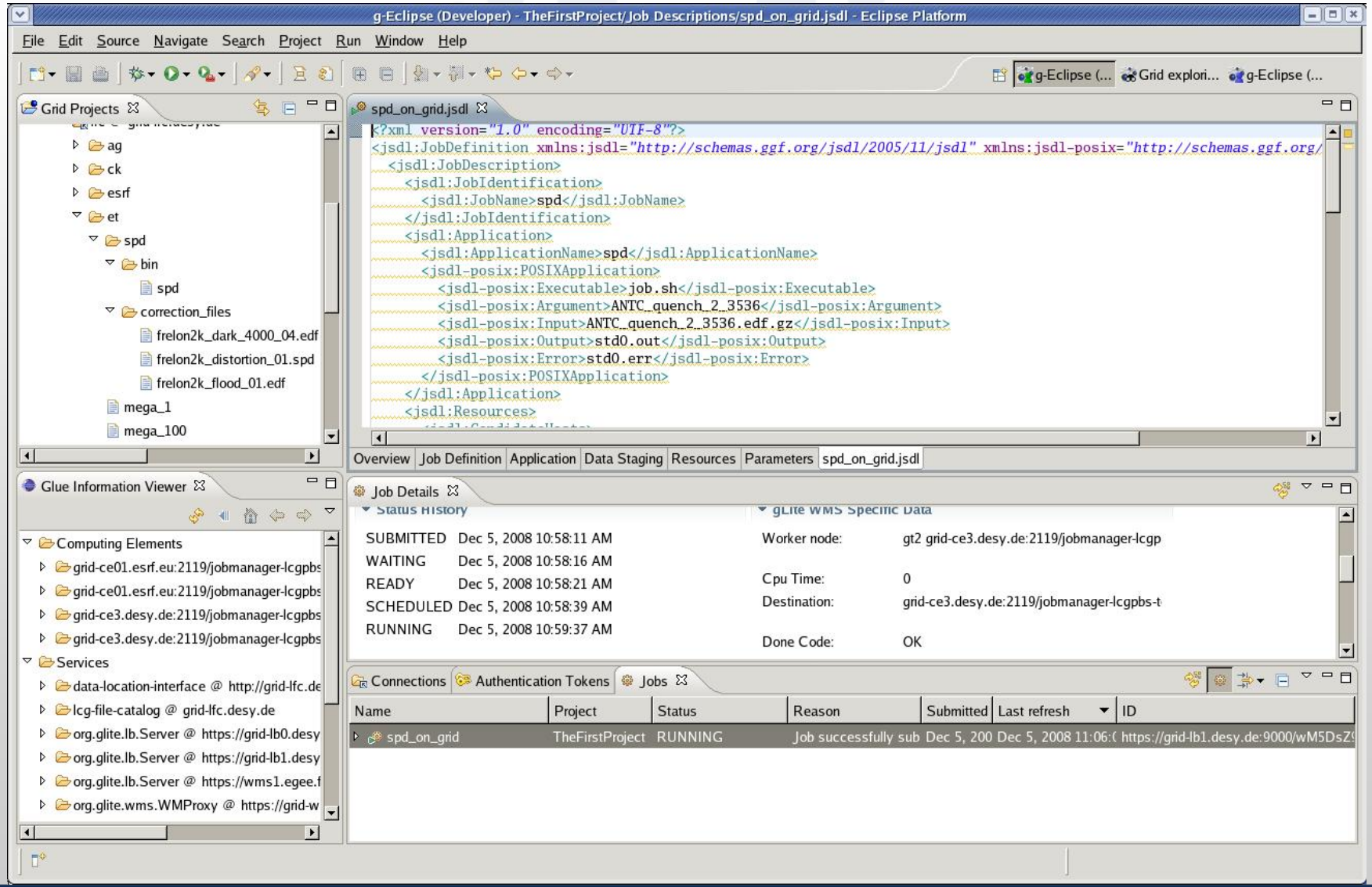
- 6 Worker Nodes with **48 Cores**
- Up to **20TB** of disk space
- 1 Computing Element: lcg-CE
- 1 Storage Element: **DPM**
- 1 User interface, Site-BDII, and Monbox
- Cluster not yet installed
- Pending delivery of Thumper

Handling the site administration

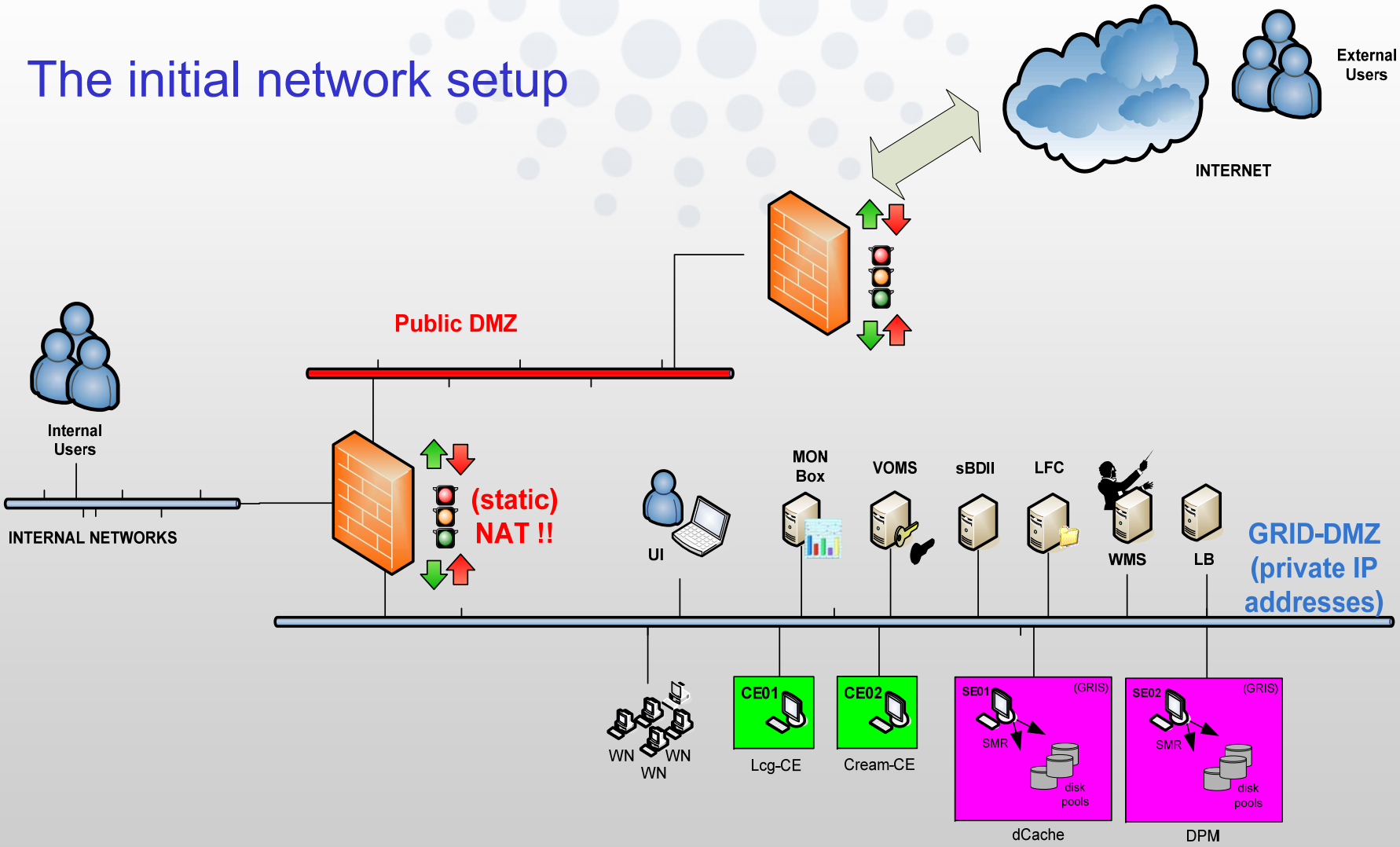
- Created a couple of deployment and configuration scripts
  - Similar to the current practice at the ESRF
- Facilitates the deployment and update of services
- Enables us to keep the configuration homogeneous



gEclipse

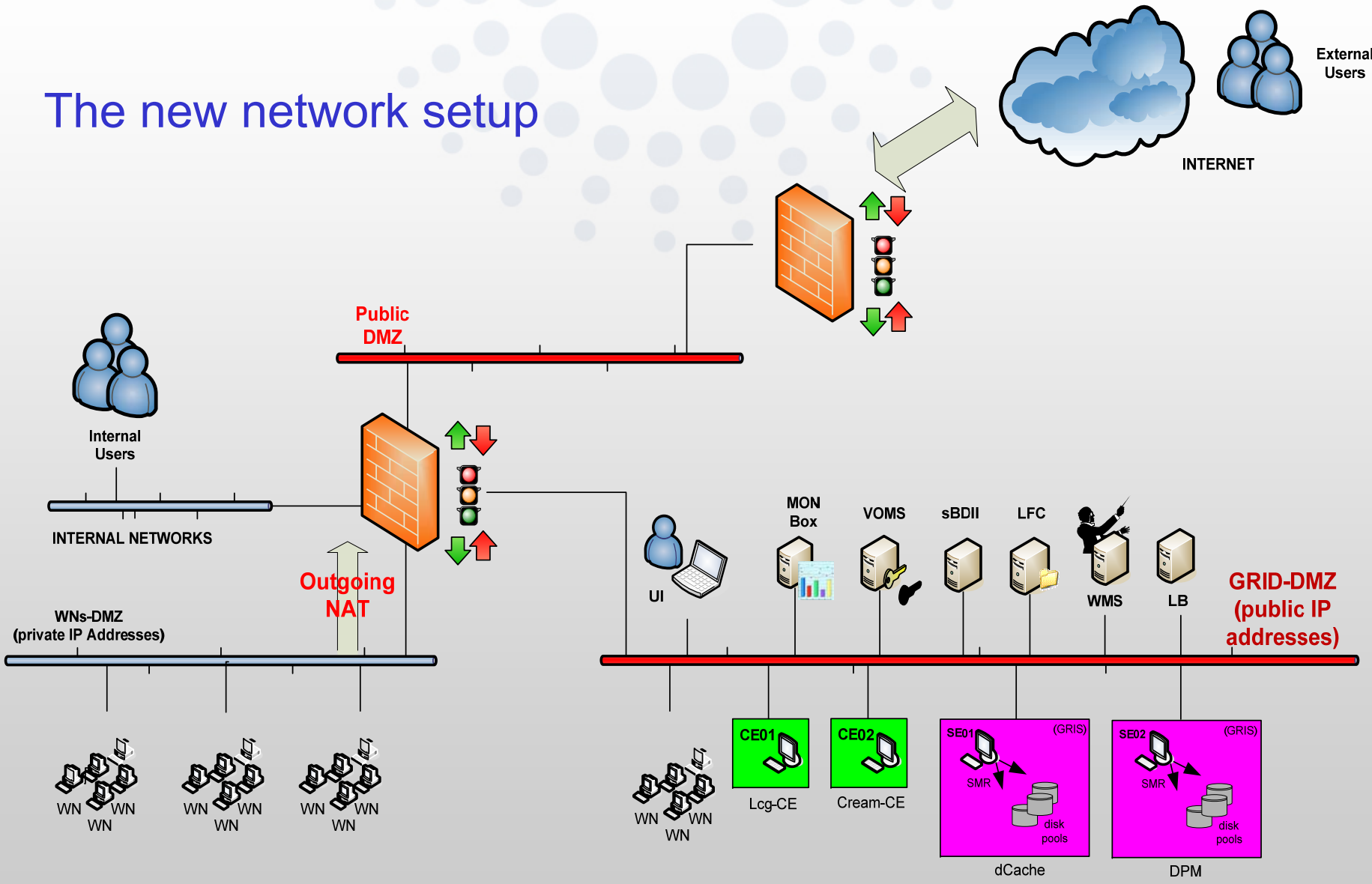

g-Eclipse (Developer) - TheFirstProject/Job Descriptions/spd_on_grid.jsdl - Eclipse Platform
File Edit Source Navigate Search Project Run Window Help
Grid Projects
ag
ck
esrf
et
spd
bin
spd
correction_files
frelon2k_dark_4000_04.edf
frelon2k_distortion_01.spd
frelon2k_flood_01.edf
mega_1
mega_100
spd_on_grid.jsdl
<?xml version="1.0" encoding="UTF-8"?>
<jsdl:JobDefinition xmlns:jsdl="http://schemas.ggf.org/jsdl/2005/11/jsdl" xmlns:jsdl-posix="http://schemas.ggf.org/
<jsdl:JobDescription>
<jsdl:JobIdentification>
<jsdl:JobName>spd</jsdl:JobName>
</jsdl:JobIdentification>
<jsdl:Application>
<jsdl:ApplicationName>spd</jsdl:ApplicationName>
<jsdl-posix:POSIXApplication>
<jsdl-posix:Executable>job.sh</jsdl-posix:Executable>
<jsdl-posix:Argument>ANTC_quench_2_3536</jsdl-posix:Argument>
<jsdl-posix:Input>ANTC_quench_2_3536.edf.gz</jsdl-posix:Input>
<jsdl-posix:Output>std0.out</jsdl-posix:Output>
<jsdl-posix:Error>std0.err</jsdl-posix:Error>
</jsdl-posix:POSIXApplication>
</jsdl:Application>
<jsdl:Resources>
Overview Job Definition Application Data Staging Resources Parameters spd_on_grid.jsdl
Glue Information Viewer
Computing Elements
grid-ce01.esrf.eu:2119/jobmanager-lcgpbs
grid-ce01.esrf.eu:2119/jobmanager-lcgpbs
grid-ce3.desy.de:2119/jobmanager-lcgpbs
grid-ce3.desy.de:2119/jobmanager-lcgpbs
Services
data-location-interface @ http://grid-lfc.de
lcg-file-catalog @ grid-lfc.desy.de
org.glite.lb.Server @ https://grid-lb0.desy
org.glite.lb.Server @ https://grid-lb1.desy
org.glite.lb.Server @ https://wms1.egee.f
org.glite.wms.WMProxy @ https://grid-w
Job Details
Status History
SUBMITTED Dec 5, 2008 10:58:11 AM
WAITING Dec 5, 2008 10:58:16 AM
READY Dec 5, 2008 10:58:21 AM
SCHEDULED Dec 5, 2008 10:58:39 AM
RUNNING Dec 5, 2008 10:59:37 AM
gLite WMS Specific Data
Worker node: gt2 grid-ce3.desy.de:2119/jobmanager-lcgp
Cpu Time: 0
Destination: grid-ce3.desy.de:2119/jobmanager-lcgpbs-t
Done Code: OK
Connections Authentication Tokens Jobs
Name Project Status Reason Submitted Last refresh ID
spd_on_grid TheFirstProject RUNNING Job successfully sub Dec 5, 200 Dec 5, 2008 11:06:(https://grid-lb1.desy.de:9000/wM5DsZ

The initial network setup



XRAY Grid

The new network setup



Some GridFTP transfer numbers

- GridFTP File Transfers ESRF → DESY SE
  - Single stream: ~ **1.5MB/s** (50 MB, 100 MB, or 500 MB files)
  - 8 parallel streams: **8 MB/s** for a 100 MB files
    - Translates in 35 h for a 1 TB dataset
- GridFTP File Transfer DESY SE → ESRF
  - ~ **1.5MB/s** for various files sizes
  - parallel streams blocked

- GridFTP File Transfer: CERN → ESRF SE ESRF → CERN
  - (45MB file) **~3-5 MB/s** **~3-5 MB/s**
  - 8 parallel streams: **13 MB/s** (1TB/d)

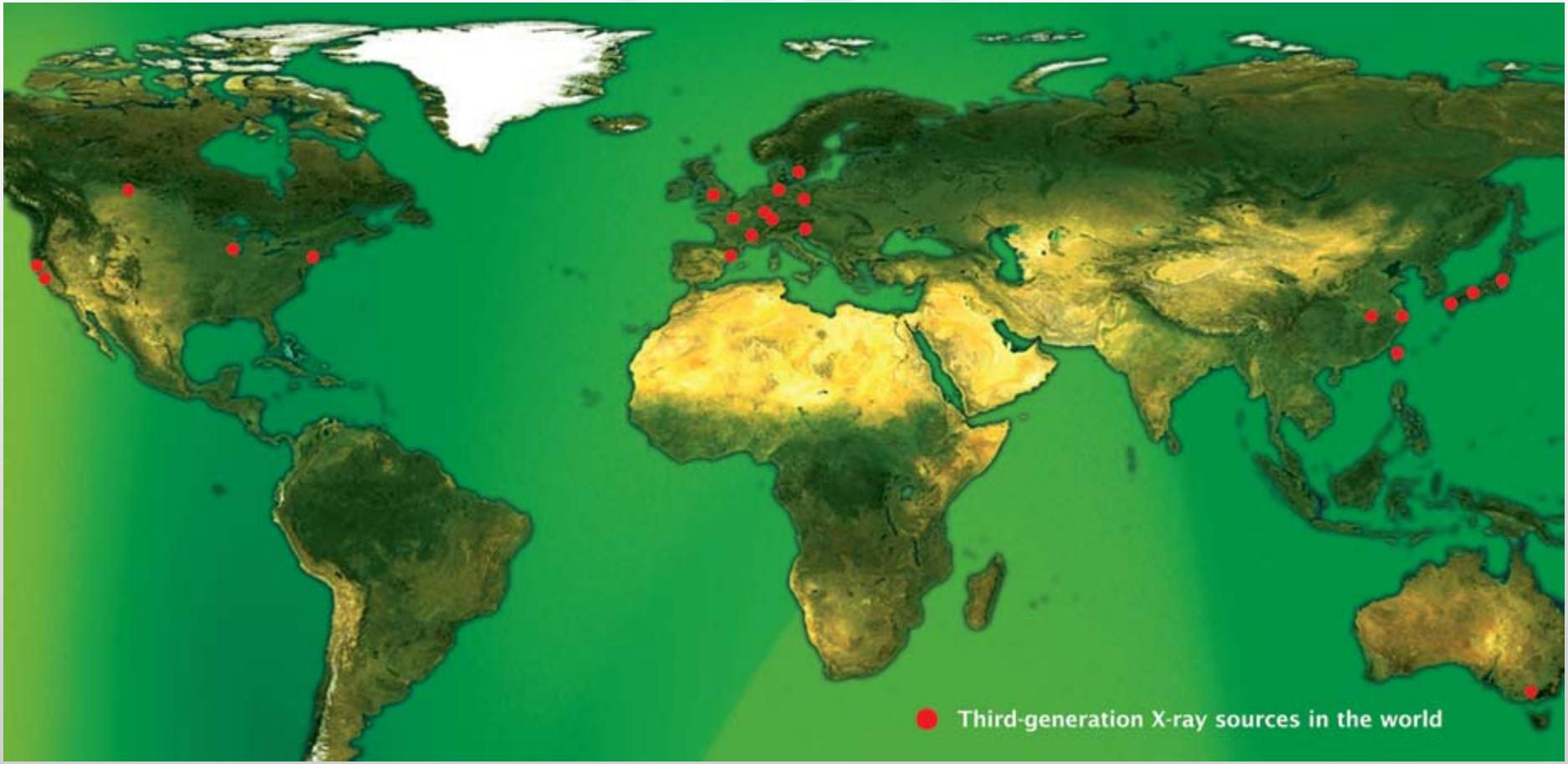
- Uplink to Renater is **1Gb/s (125MB/s),** but shared between ESRF, ILL, EMBL
  - Do we need a dedicated >1Gb/s uplink for the grid ?!

Test the grid yourself

- Want to test the grid yourself?

  → start here: http://www.esrf.fr/Infrastructure/Computing/Grid/how-to-become-an-xray-user

- Want to discuss grid?

  → contact grid-admin@esrf.fr

Tomorrow's XRAY Grid!?!