XRAY Grid

TO BE OR NOT TO BE?
I was not always a Grid sceptic!

- I started off as a grid enthusiast e.g. by insisting that Grid be part of the ESRF Upgrade Program outlined in the Purple Book.

- In this talk I will try to explain why nowadays I oscillate between being a Grid sceptic and Devils' advocate.
Which Grid are we talking about?

- **EGEE / gLite grid**:
Synchrotron Requirements

• Lots of small jobs – 100000's running for minutes

• Jobs often read and write images – 100's images / job

• Experiments generate lots of data – up to Tera Bytes / day

• Large number of small users groups – 100's / year

• Many diverse experiments – 1000's / year
XRAY Science is DATA INTENSIVE!

- ESRF has currently 400 TB, end of 2009 800 TB, in the future Petabytes ...

- Data per Experiment and per Beamline since January 2009:

  ![Graphs showing TB per Experiment and TB per Beamline](image-url)
Test program - spd

- **Spd is a program to correct 2D images for:**
  - Spatial distortion: 2d spline curve
  - Flood field: image division
  - Background field: image subtraction

- One image takes about 17 seconds
- Additional images take a fraction of a second

- Simple but typical of many programs used on 2d images:
  - Low on CPU, High in I/O

- Typical data set is 180 images x 8 i.e. 1.44 GB
- Typical experiment measures HUNDREDS of data sets
grid_spd

- grid_spd is a script written by Emmanuel Taurel to test running spd on the XRAY grid setup by Clemens Koerdt and Fernando Calvelo

- Test scenarios:
  - Data on User Interface, submit Jobs to WMS
  - Data on LFC, submit Jobs to WMS
  - Data on LFC, submit jobs directly to CREAM-CE
  - Data on LFC, submit Parametric jobs
  - Data on NFS, submit jobs to CREAM-CE
  - Data on NFS, run job from command line
grid_spd results
XRAY Grid

TO BE ?

We could assume:

- Grid is OK for synchrotron science
- Programs will run slower but still run
- External users will be able to use the grid resources
- Data can be exported and accessed from UI
- Middleware (gLite) will get better (and faster) with time
- National Grid Initiatives will provide free resources
XRAY Grid

OR NOT TO BE ?

• We could assume:
  - Grid is NOT OK for synchrotron science
  - Data intensive applications are not adapted to Grid
  - LFC and WMS are too slow
  - SE does not provide fast, direct access to data
  - Exporting data is not adapted to occasional intense usage
  - Middleware (gLite) will not change drastically
  - Future of gLite is uncertain, funding stops in 2010
XRAY issues with EGEE/Glite

- LFC:
  - Very slow for many files

- WMS:
  - Very slow, optimised for long jobs

- SE:
  - dcache + DPM are slow, not directly mountable

- WN:
  - Users cannot run with their UID/GID
  - Not adapted for MPI intensive applications

- GridFTP:
  - Needs tuning by guru's, not reliable, no/poor Windows support

- Grid certificates:
  - Not clear that they scale to 1000's of XRAY users
Why is (EGEE) Grid not adapted?

- Grid assumes each node is independent
- File systems are not mounted locally on WN's
- Grid file I/O is much slower than locally mounted files
- Grid does not provide any improvement over clusters
- Why would you want to use the Grid in this case?
XRAY grid next steps

- Gridify and benchmark more applications:
  - Penelope – a radiation dose monte-carlo simulation
  - Peaksearch – a 2D and 3D peaksearching code
  - Laminography – a tomography technique
  - FDMNES – an xray spectra calculation MPI code
  - Fit2dCake – a radial integration program
  - XDS – a protein structure refinement program

- Try to use GridFTP for transferring real data, test iRods

- Install a faster scheduler e.g. Condor, and connect to ESRF local batch system
XRAY – A Hybrid Grid

• Setup local optimised clusters with high performance local file systems and batch submission systems

• Provide Grid access via direct login and grid certificates

• Export data via gridFTP and iRods:

• Use gLite for long jobs or remote batch submission

• Preinstall common xray software
HPC-grid versus HEP Grid

- High Performance Computing refers to optimised clusters for running cpu + data intensive and mpi applications

- Users have an account on the worker nodes

- File systems and network is optimised

- Computation is done on:
  - Multi-core CPU's
  - GPU's

- Hardware and software is installed and managed locally
XRAY Science and Grid?
Conclusions*

- EGEE/Glite based Grid has a steep learning curve, is not user friendly however it does seem to work for certain applications

- First impression of EGEE/gLite based Grid is that it is not suited to XRAY science

- A big thanks to DESY for help so far but we need more if we want to run data intensive applications on the grid

- A better return on investment for data intensive applications seems to be to invest in HPC e.g. GPU's, multi-core CPU's

* we reserve the right to change these conclusions in the future 8-)