

Implementing Open Source Grid in Oil and Gas Industry

EGEE London Business Day - Distributed Computing Workshop

Hannah Cumming



TOTAL

What is our interest in Grids?

Many of the projects at GRC require simulations that are computationally expensive and produce huge datasets.

History Matching

- Trying to match production data from by using our reservoir model. Multi-realisation can requires large amounts of simulations per run.

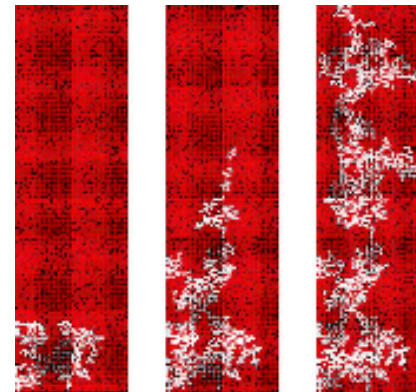
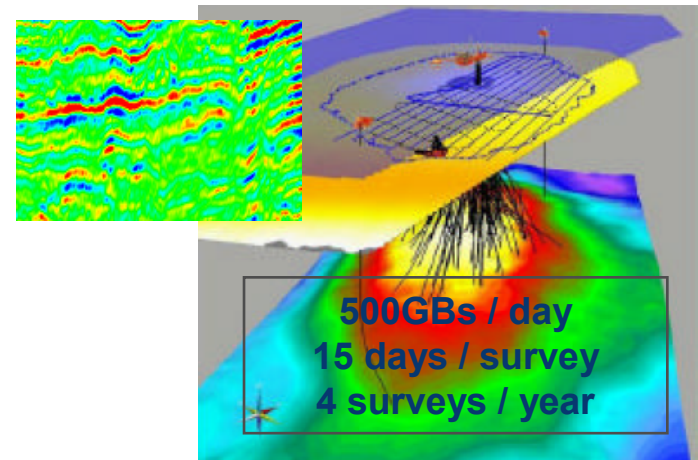
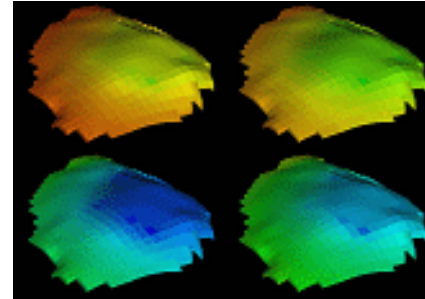
Seismic Processing

- Huge datasets that require many teraflops and number of machines in order to process the data in acceptable timescale (i.e. days not years!)

Pore Network Modelling

- Currently we simulate the interaction of oil, gas and water in pore space for core samples sizes but would like to simulate on a much bigger scale.

Need to look at new Computing methods to deal with future computing demands!



Grids in commercial and research worlds

Commercial

- Mainly Internal or Company wide (Enterprise)
- Small amount of Vendors selling to lots of users (On-Demand)

- Commercial Software – LSF, United Devices ...

- Less information available

vs

vs

vs

Research

- Sharing of resources and large amount of organisations working together (External)

- OpenSource software – gLite, Globus..

- Very open lots of information available

Types of Grids:

- Internal/Enterprise Grids
- External Grids
- On-demand Computing
- Data Grid
- cycle-scavenging networks
- Grid Enabling Software

Although it is not always this case as a general rule the two communities of research and enterprise grids are quite separate

External Grid Computing Project

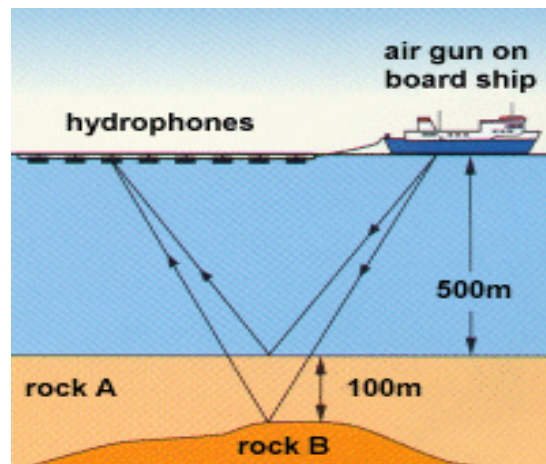


External Grid: A shared distributed computing resource in terms of computing power and data storage and also distributed in administrative domains, that behaves as if it was one resource.

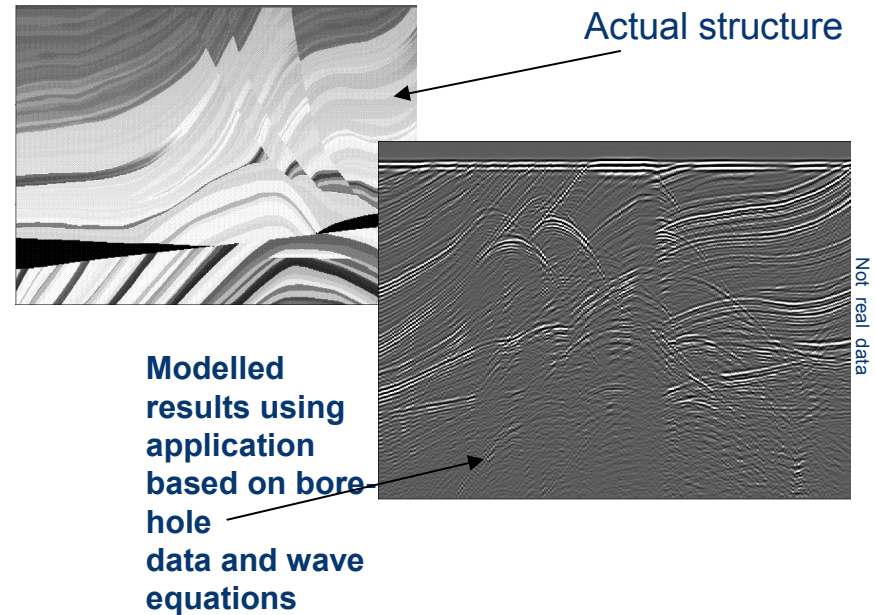
- ▶ The external grid project is one of a number of projects within Total looking into computing technologies.
- ▶ A research project to assess the usefulness to Total of an 'External' GRID.
- ▶ Using hands on approach by adapting one of our own applications for an external Grid.
- ▶ Although specifically aimed at External Grids also looking at transferable parts to other areas such as internal Grids

Twist++ Application Details

Twist++ is a fast 2D finite difference modelling code suitable for 2D marine experiments.



Marine Experiment



Software chosen for External Grid Project because:

- The code was developed at the Geoscience Research Centre
- It is already designed to use distributed computing
- It requires a large amount of computational time week to run

Why EGEE?

EGEE is an European Project has created and integrated a Computational GRID infrastructure that is the largest of its kind in the world.

- 120 partners
- 48 Countries
- 250 Sites
- 60,000 CPUs
- 20PB
- 8000 Users every day
- Open Source Middleware gLite
- Many different disciplines including a few Geosciences
- Mix of Companies and Research Organisation Involved
- Industry Forum and Task Force
- EGEE now into 3rd Phase

Our Grid Requirements

- Multi-national
- Longevity
- Multiple Organisations
- Production service
- Strong links to UK



How to get access to an External Grid

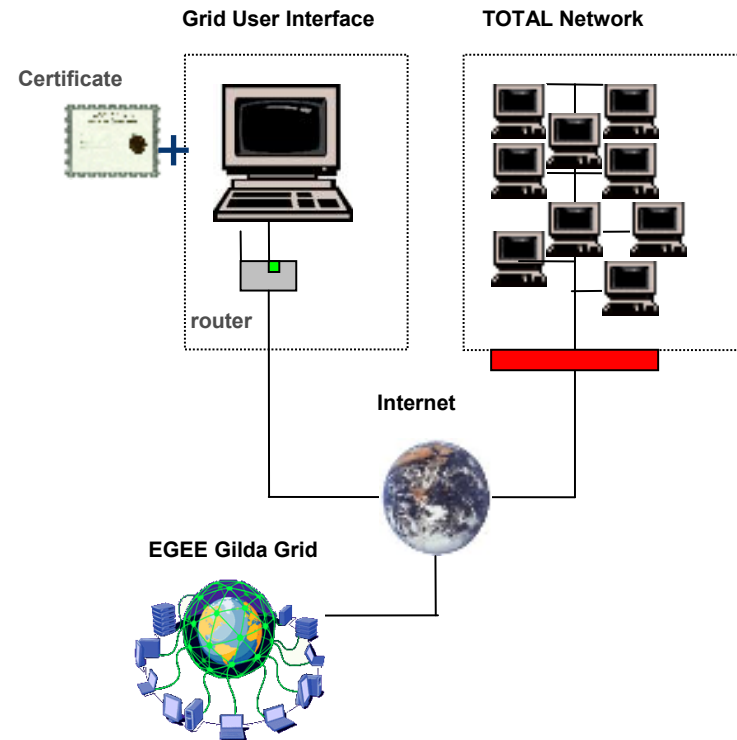
EGEE training GRID (GILDA)

- Set up access over the internet/ email.
- Allowed trial of grid
- Wiki Documentation
- Experience of Grid and gLite Middleware
- Development of our Grid Application
- Small scale testing

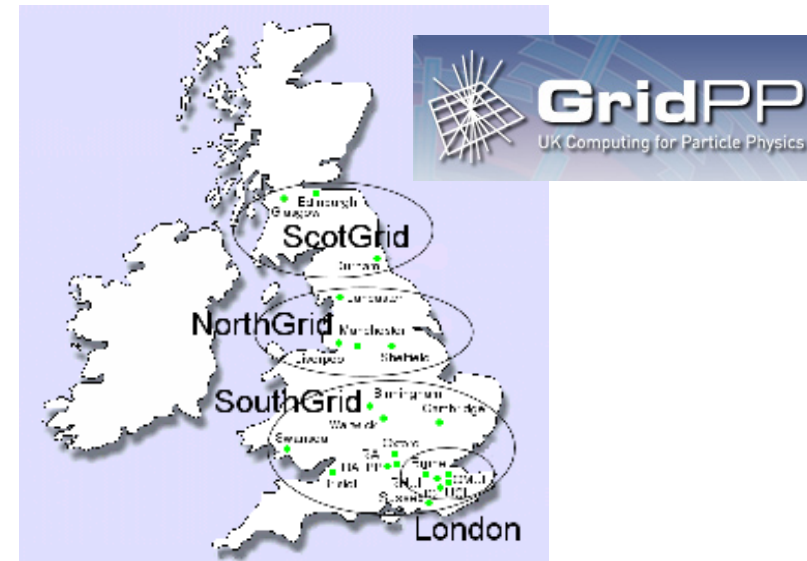
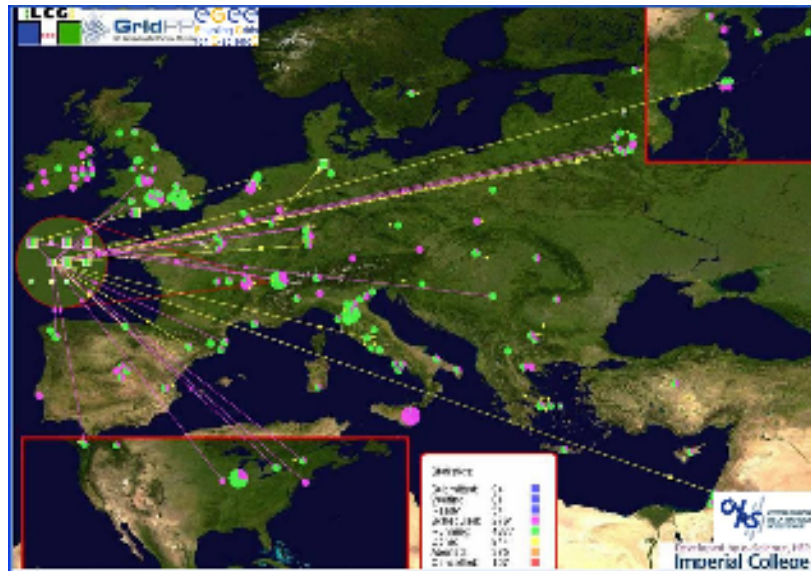
Allowed us to gain access with out making huge commitments

Allowed us to access risk and other obstacles

Gain Experience and knowledge through EGEE Community



Main EGEE Grid - VO and Certificates



- ▶ Access to the main Grid through GridPP
- ▶ Created a Virtual Organisation totalep with the help of GridPP
- ▶ To enable us to run larger runs
- ▶ These results will allow us to make decisions as to what to do next.

Pros and Cons of External Grids

Pros

- ▶ It works able to submit jobs to external Grid and get results
- ▶ The External Grid community are at the forefront driving standards and developing new ideas in the Grid Community
- ▶ A lot of these are transferable and could be useable in other domains
- ▶ External Grids are already proven in research domain and research collaborations with commercial companies works well.

Cons

- ▶ A lot of applications we use are commercial simulators so can't use in External Grid Environment
- ▶ In commercial environment security is big issue
- ▶ Also have to think carefully about the type of application and how/who is using it. As it takes effort to make a good Grid application
- ▶ Still focus strongly on research domain less clear what role commercial companies can play.
- ▶ External grid use is restricted by the capacity of the underlying network to carry data volume (Good research network but could be less true in commercial world)

Conclusions

- ▶ **GILDA was very useful way to test the grid and gain experience without any commitments**
- ▶ **It works and the External Grid is being used by the research community on production bases**
- ▶ **The External Grid community are at the forefront driving standards and developing new ideas within Grid Community**
- ▶ **Many of technologies and standards being developed can be used internally as well**
- ▶ **Research collaborations between commercial and External Grids work well now**
- ▶ **Already commercial vendors are grouping together to provide grid services to very specific fields**

Questions?