September 2006

Total: External Grid Computing Project

Hannah Cumming

07/2006

Overview

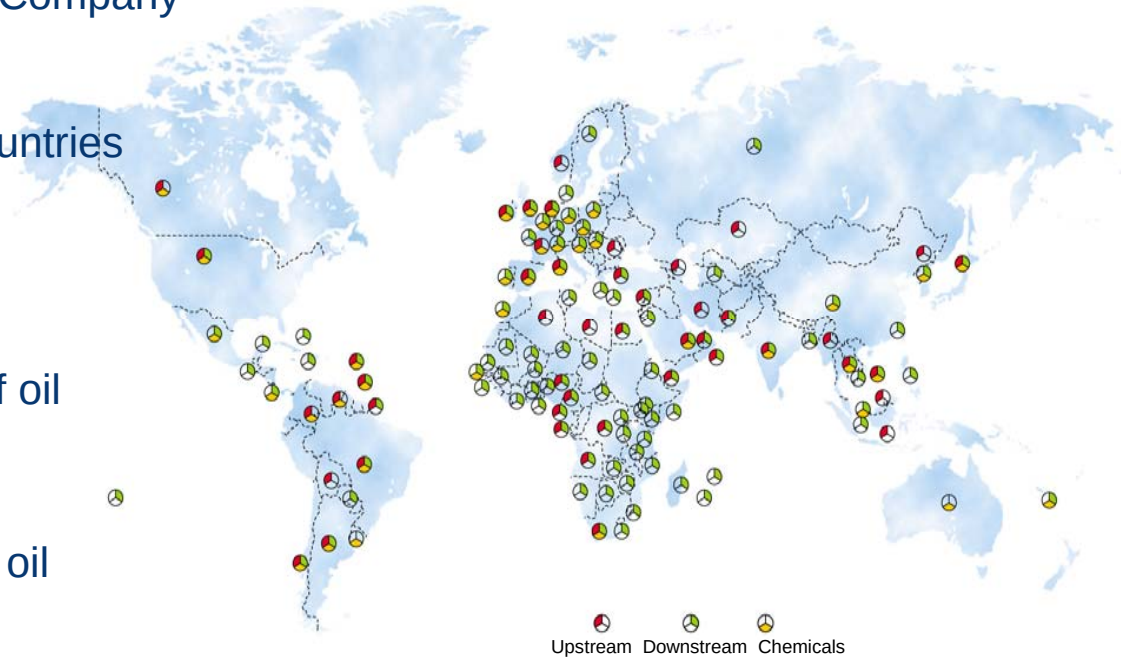
- **Introduction**
  - Total Overview
  - Totals use of High Performance Computing
  - Meeting Future High Performance Computing Needs
  - Why is Total interested in External Grids?
- **External Grid Project**
  - External Grid Computing Project
  - Why EGEE?
  - Current Set-up
  - Grid Application
  - Application Development
  - Current Status
  - Future Work
- **Conclusion**

TOTAL

Total Overview

Global Oil & Gas Company

- Worlds fourth largest Oil & Gas Company
- 95,000 employees world wide
- Operations in more than 130 countries

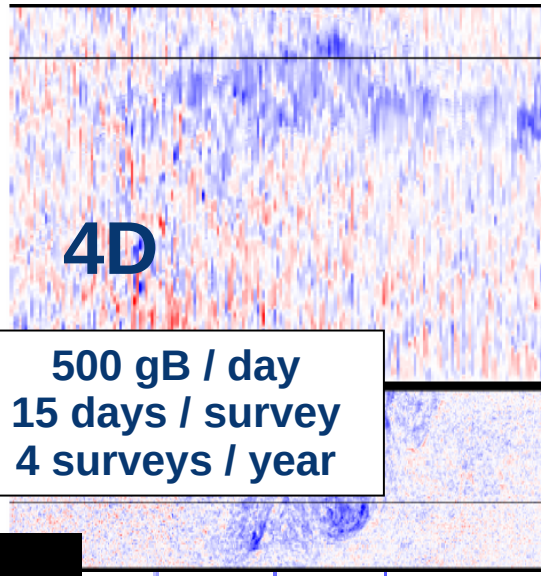


Exploration & Production

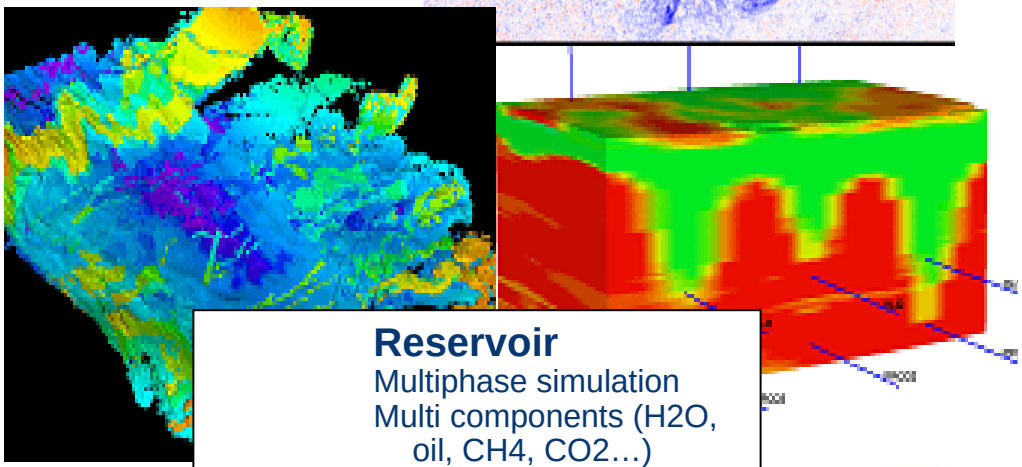
- Production: 2.5 million barrels of oil equivalent per day
- Reserves: 11.1 billion barrels of oil
- Producing in 29 countries
- Interests in five of the worlds largest liquefaction plants, accounting for 40% of the worlds LNG production

Totals use of High Performance Computing (HPC)

- **Seismic Processing**
  - Heavy data processing (in TB / projects)
  - Heavy consumer of computing resources (**80%** of TOTAL computing capacity)
  - Heavy R&D efforts to improve seismic tools (ex imaging)
- **Interpretation of subsurface structures**
  - Instant computing power required (surface/volumetric interpretation)
  - Heavy parametric studies
  - Visualisation of large models
- **Reservoir / Field modelling**
  - Reservoir simulation
  - Probabilistic production forecasts
  - Sensitivity analysis



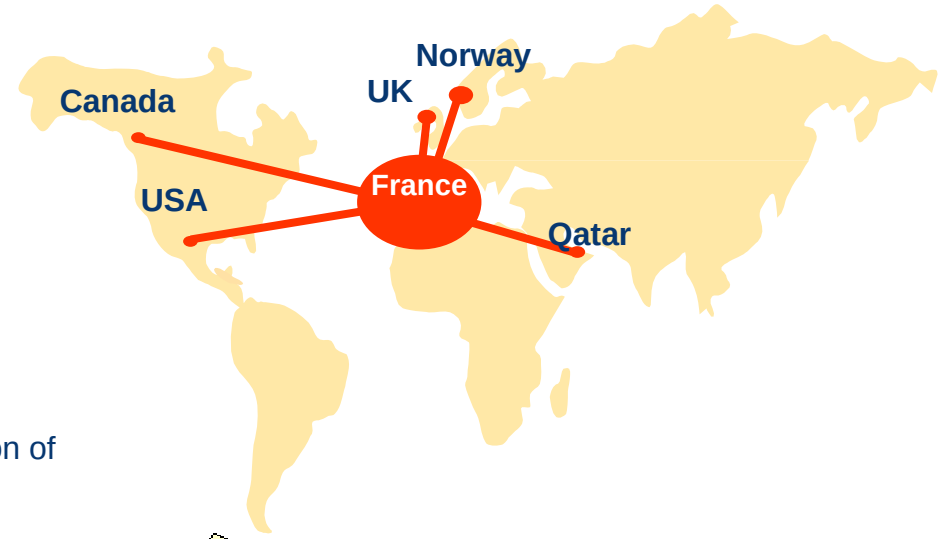
500 gB / day
15 days / survey
4 surveys / year

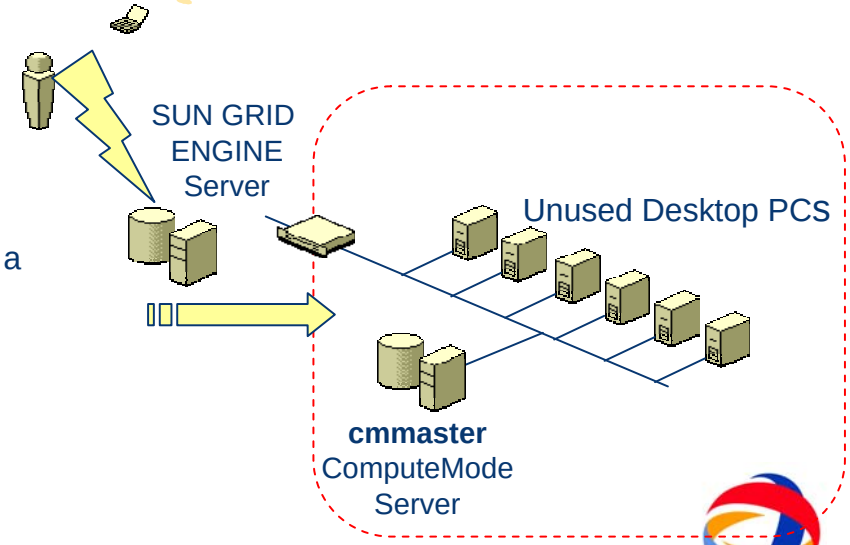


Reservoir
Multiphase simulation
Multi components (H2O, oil, CH4, CO2…)
Sweeping (EOR),
Chemical reactions
Geomechanics (seal efficiency…)

Meeting Future HPC Demands

- **Internal Resources**
  - Clusters, Workstation, Servers
  - Workstation Farms
    - Utilising Internal Unused Workstations
  - Internal Total Grid
    - Compute resource sharing within one administrative domain (maximizing utilisation of heterogeneous clusters)
- **External Resources**
  - On-Demand Computing
    - Processing jobs or treating large data sets remotely on a hosted environment temporarily to a customer
  - External Grid Computing
    - Distributed in terms of administrative domains as well as having physically distributed resources





Why External Grids - Benefits

Future Benefits

- Instant access to external computing resources
- Use for peak demand type situations
- Access to potentially very large computing resources
- A different way of working use computing resources at the location where the data is generated no need to transfer large data sets.
- Optimising cost of HPC Ownership – shared resources
- Contribute spare computing resources to the Grid

Now

- Middle ware and technologies being developed in the external grid community that can be used now on Internal Grids

TOTAL

Why External Grids - Obstacles

Network and Bandwidth

- Traditionally companies do not have high speed links to external networks

Licensing issues

- Need ISP to provide viable Grid Licensing options (per user, per CPU, per use, per VO, per site?)

SECURITY!!!

- Programs and data would have to be stored on external computing resources .
- Need to prove/convince IT departments internally that it is safe

Still research based

- No payment structure in use

TOTAL

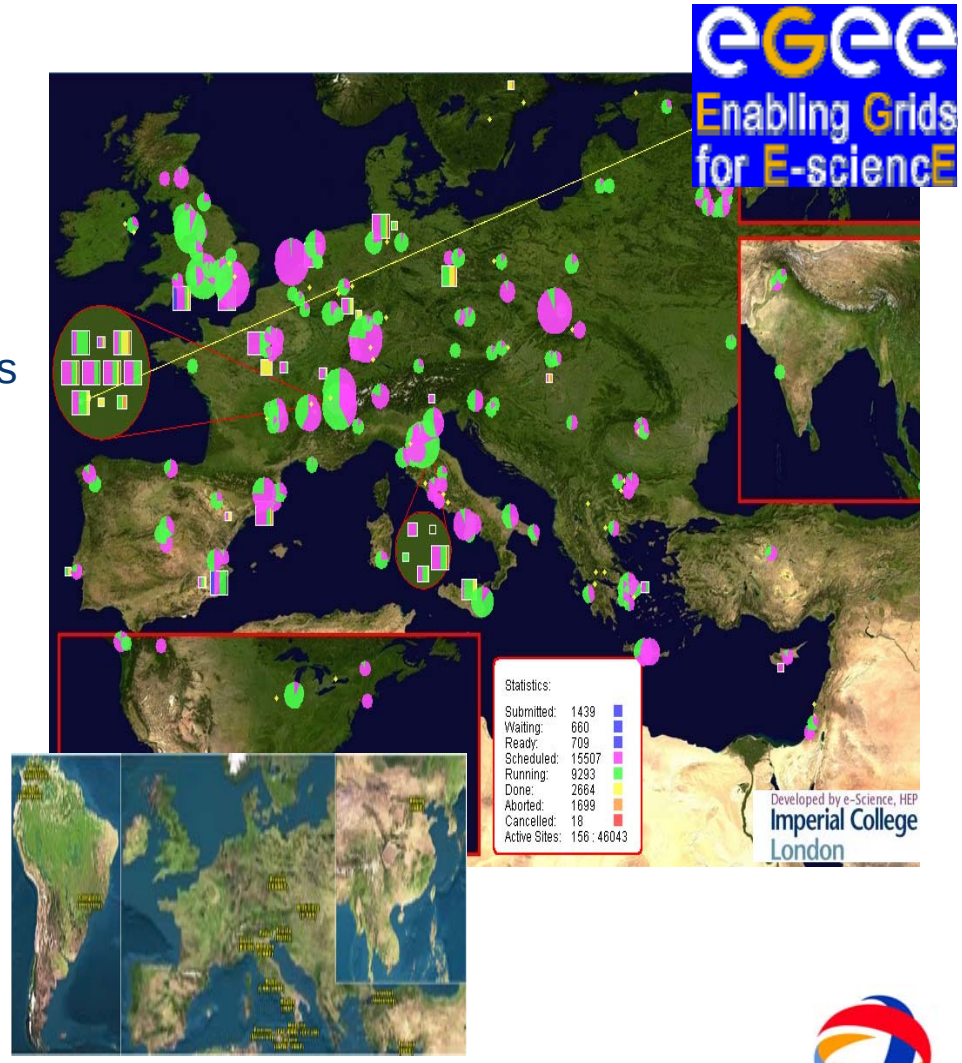
External Grid Computing Project

Aims of External Grid Computing Project Initiative:

A research based project to assess the usefulness to Total of an ‘External’ GRID. By adapting one of our own applications (Twist++) for an external GRID.

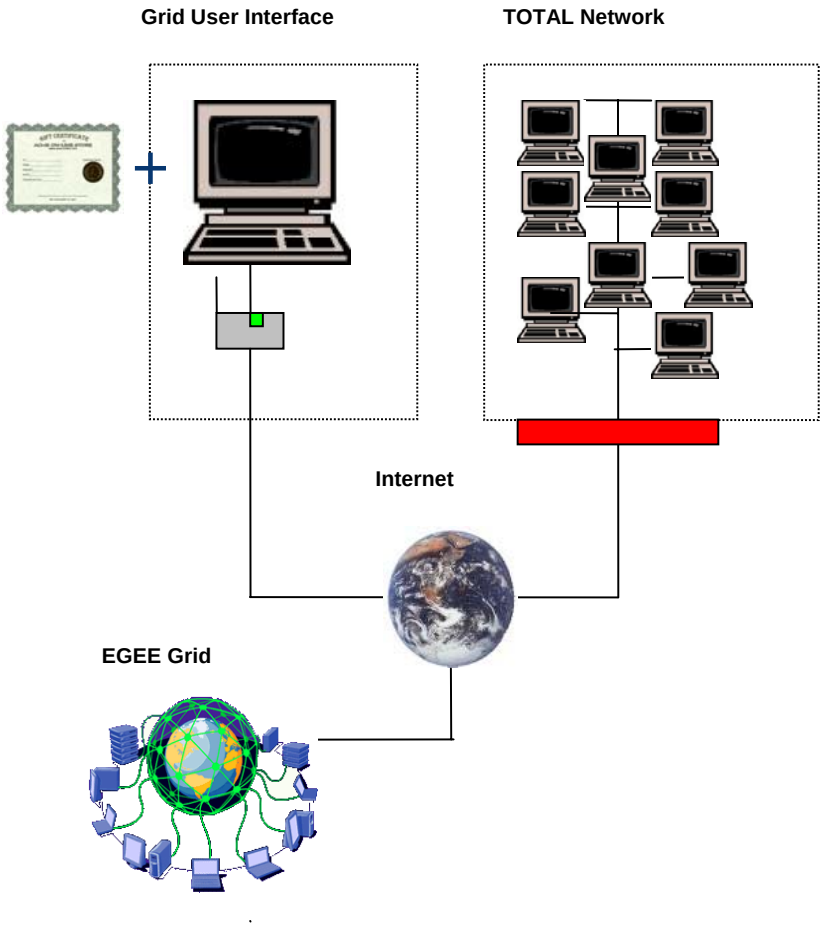
Why EGEE?

- International Grid with strong links to the UK. Important as a UK based research centre.
- The EGEE Grid is one of the worlds largest external Grids. Also has large influence over future of Grid developments
- gLite Middleware had all the features we were interested in testing- Resource Broker
- The EGEE grid has a wide user community physics, medicine, chemistry....and also geoscience community
- GILDA test grid was useful starting point, allowing Total to test the Grid with little or no risk. Support available through GILDA.



TOTAL

Current Set-up at the Geoscience Research Centre

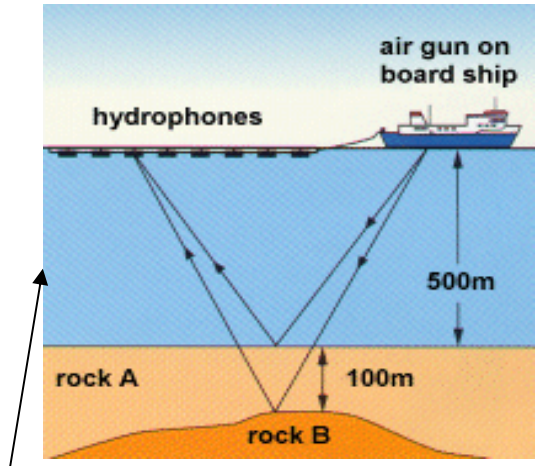


Current Set-up:

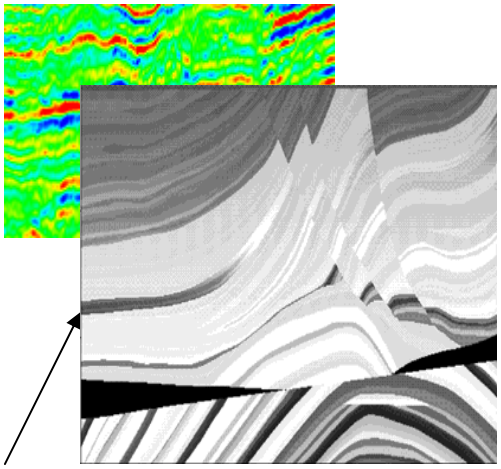
- Dedicated User Interface machine to the EGEE GILDA Grid on separate internet connection using plug and play user interface
- Grid Certificate and membership to the gilda Virtual Organisation
- No resources contributed at this time to the EGEE Grid
- Use of the GILDA Grid for development and small scale tests

Twist++

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

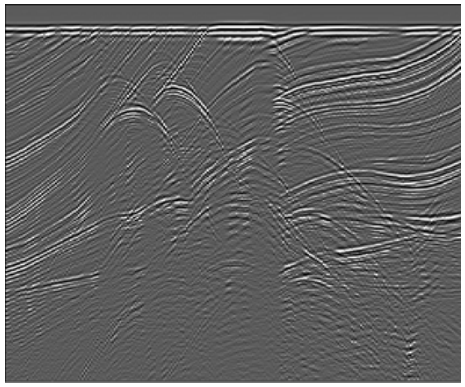


Marine Experiment



Results of marine experiments

Modelled results based on bore-hole data and wave equations



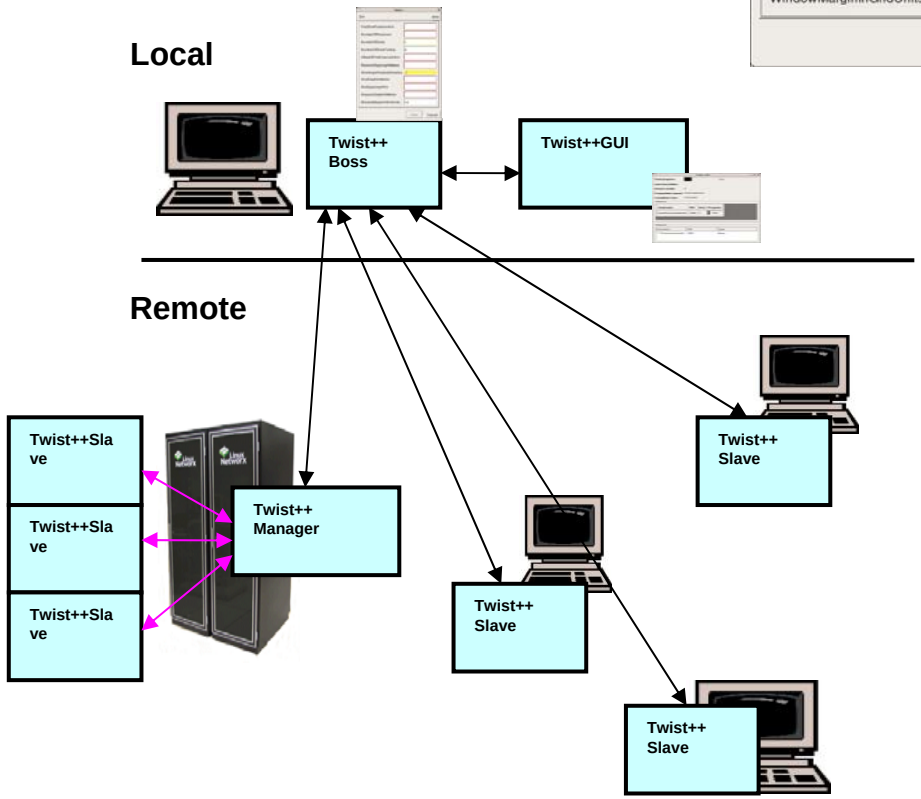
Not real data

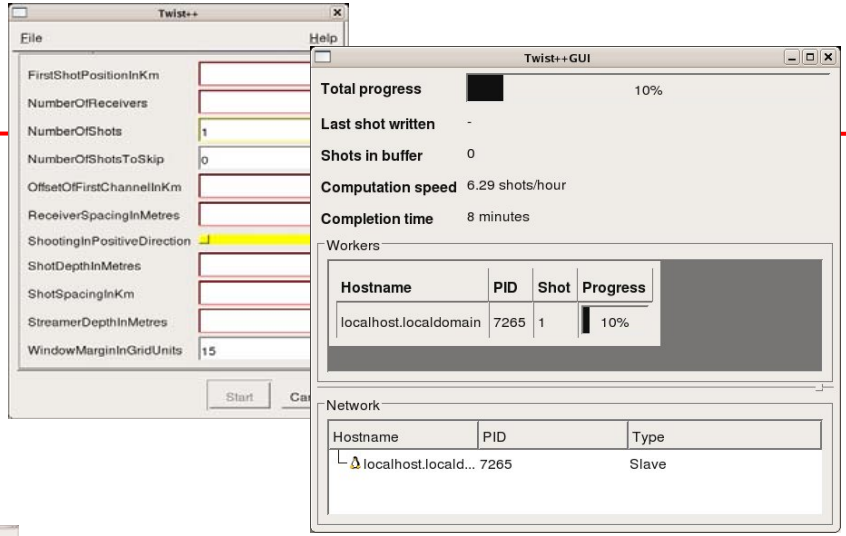
An example use it to use modelled data to validate results from marine experiments

TOTAL

Twist++

Original Design





Software chosen for Grid Project because:

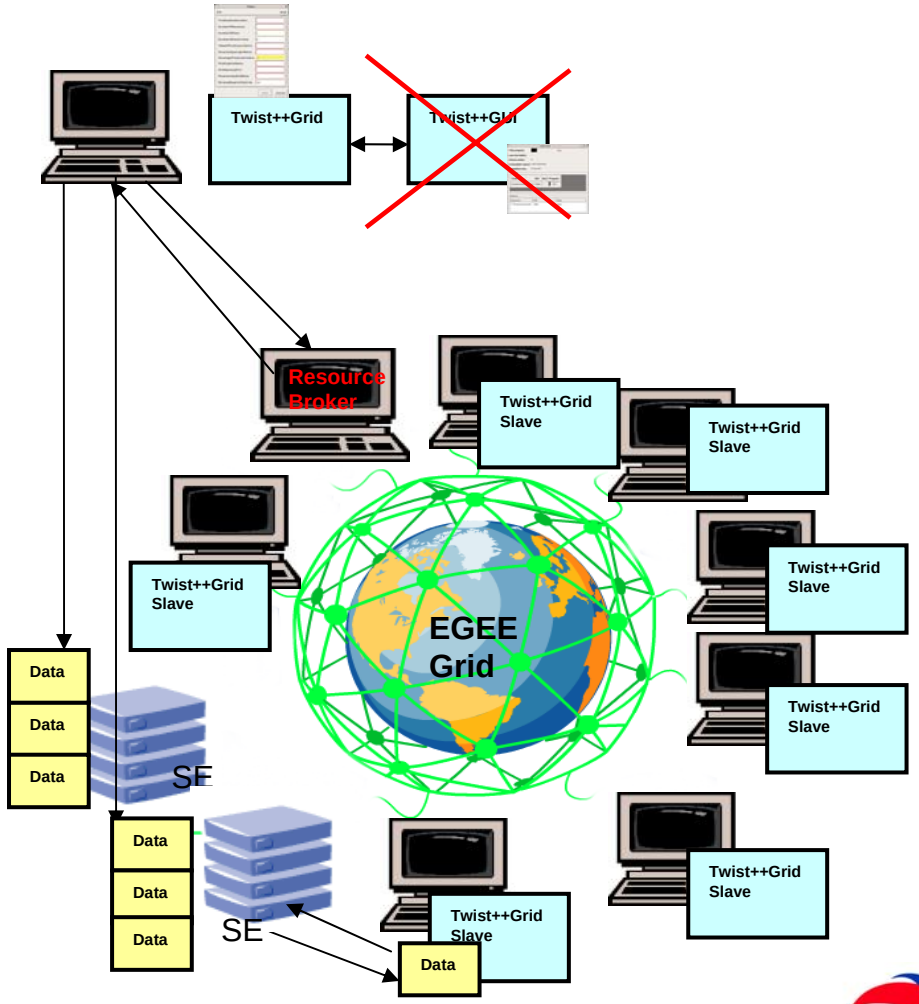
- It was developed at the Geoscience Research Centre
- It is already uses distributed computing
- It requires a large amount of computational time week to run

TOTAL

Twist++Grid Development

Grid Application Design

- Twist++Grid runs program from UI machine – distribute data and remote jobs to the Grid
- Twist++Grid Slave is smaller executable that is used to process sections of data on remote Grid nodes and produce a result file
- Use gLite Application Programming Interface to connect with the grid



Current Status

- Twist++ code running on the GILDA grid using links with gLite C++ APIs.
- Small scale tests conducted to test reliability of the application

Findings from application development:

- This particular application was relatively easy to adapt to grid architecture
- gLite APIs are easy to use so far
- APIs and features of Middleware still changing (glite IO -> gfal)
- Documentation on gLite is improving (Updated User Manual, new GILDA WIKI site, ticketed support system for GILDA) still can be difficult to know which features are currently working and what are the known bugs.

TOTAL

Future Work

- Further work on application add security measures, GUI
- Up-scale Application to Main EGEE Grid
- Conduct large scale tests with Twist++ Application (1,10,100… nodes)
- Project finishes End of Year
- Step by Step approach for future projects

TOTAL

Conclusions

- **Continued Business need for HPC resources**
- **External Grids have many potential benefits**
  - instant access, peak demand
- **Still many obstacles**
  - Not proven in commercial situations, licensing
- **Development of Grid Application relatively straight forward for *this* application**
- **Further work to be carried out on Main Grid before final conclusions**

TOTAL