CDF Grid (UK)

Dave Waters, University College London

- What is Grid ?
- Why do we want a Grid for CDF ?
- Hardware Resources available in the UK
- What do we want to do with the CDF Grid ?
- Requirements and Tools
- Possible Solutions; D0 and SAM
- Next steps

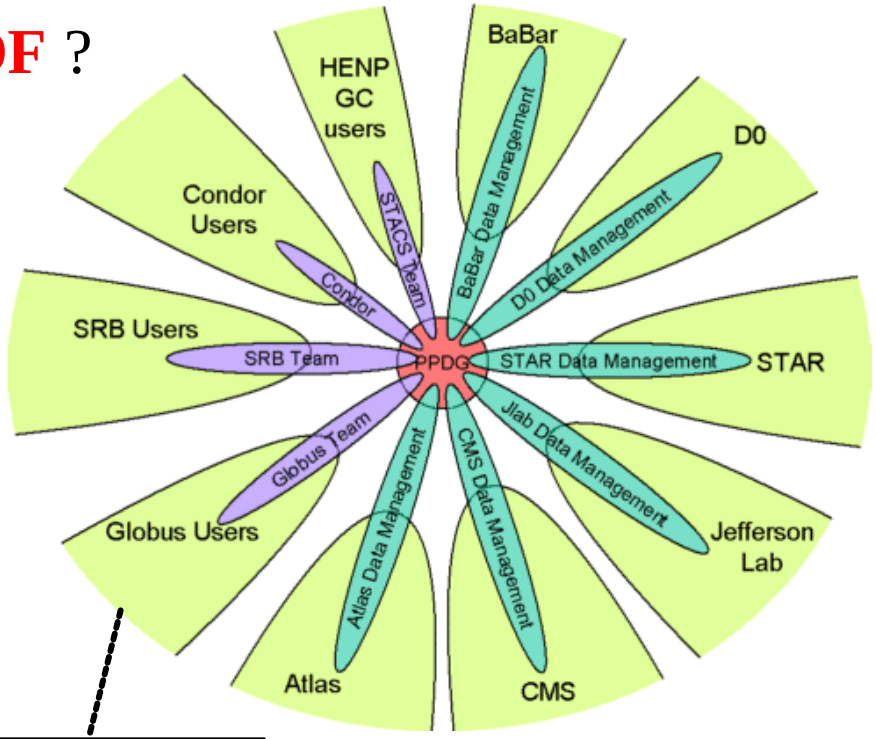
What is Grid ?

- Don’t really know.
- A language : distributed computing, virtual organisations, …
- An illuminating comparison is with the web :
  - **Web** : uniform access to HTML documents.
  - **Grid** : access to and sharing of all computing resources (storage, CPU, databases and catalogues, …)
- An emerging multi-layer architecture and set of protocols. There is an analogy with the Internet Protocol, through which a small number of protocols allow the development of a large range of applications that don’t care about the underlying fabric :
  - **IP** : allows you to write internet app’s without hardware knowledge.
  - **Grid** : allows you to construct distributed computing applications without worrying about details of individual resources.

What is Grid ?

- There is a Grid of organisations undertaking Grid work :

CDF ?



EU DataGrid

GridPP → Funds UK Particle Physics Grid Projects

Astrogrid

BioGRID

MeteoGRID

GriPhyN

NASA IPG

What is Grid ?

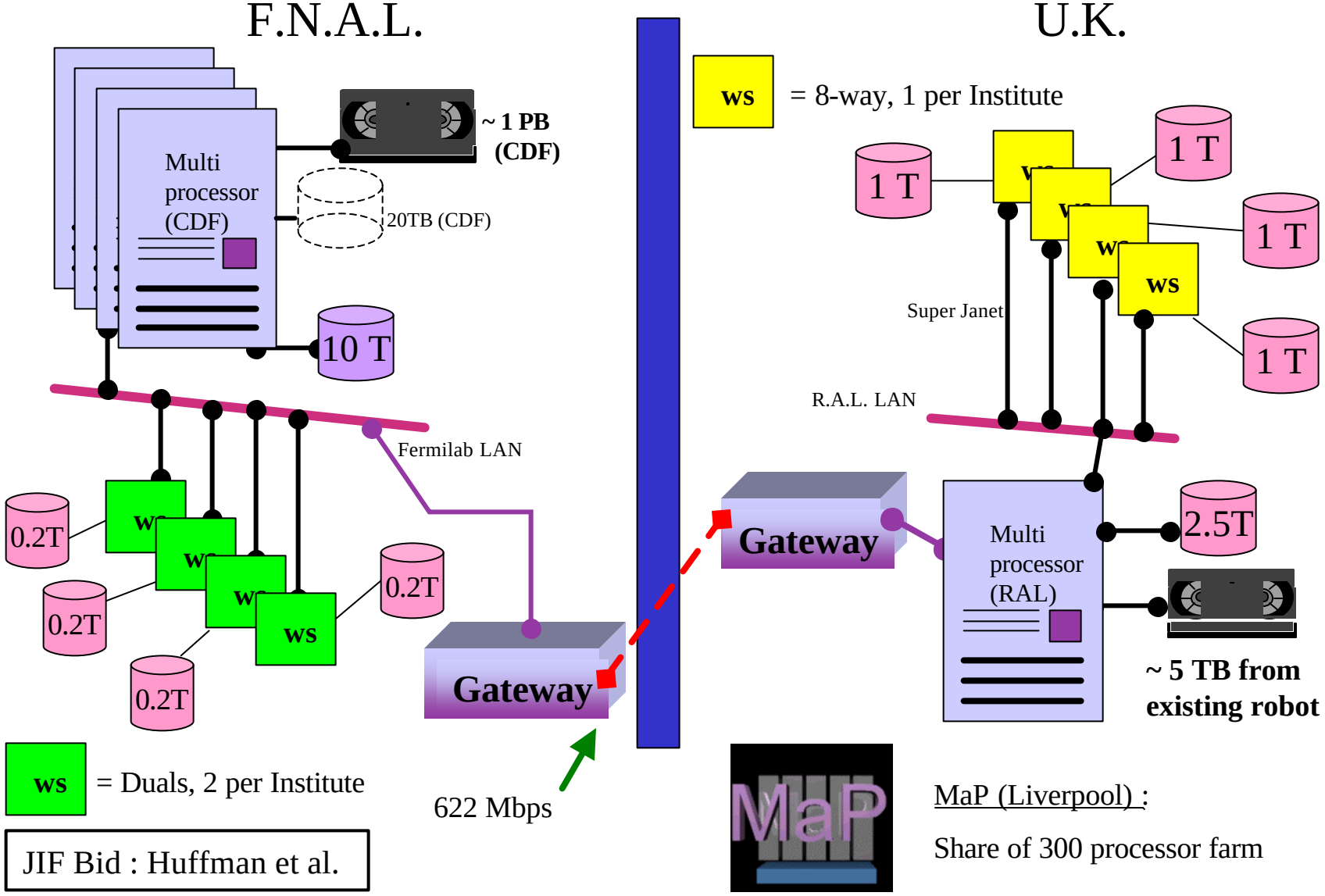
- Well funded :
  - $DOE, $NSF.
  - EU DataGrid : ~10-20M Euro
  - eSciences in the UK : £100M +
  - GridPP : ~£17M
  - Many others ...
- Well populated :
  - 100-200 people working on EU DataGrid alone.
- The effort in HEP is naturally focussed on LHC era experiments. However substantial funds are available for “prototype Grids”, including D0 and CDF :
  - Experience with a an experiment taking real data is worth many Mock Data Challenges.

Why do we want a Grid for CDF ?

- We have extensive distributed computing resources available (for example in the UK) that we want to utilise as efficiently and transparently as possible.
- Pressure on central facilities means that remote computing power should be harnessed (for LHC experiments this is actually a requirement - there will not be enough computing power at CERN to do everything centrally).
- Experience on previous experiments has been mixed. Remote computing power often not used because :
  - Latest code, calibrations, reprocessed data is always on-site.
  - Too much effort to set-up environment, pull data to remote sites (N.B. this is already much better in CDF in my experience).

⇒ Can Grid can help us overcome these problems ?

Hardware Resources available in the UK



Hardware Resources available in the UK

IBM e-server X series 370

- 8 times 700 MHz Xeon
- 4GB RAM
- 1 TB Fibre Channel Disk

Tape Store

Fermilab :

- 8 dual 800 MHz PIII machines
- 10 TB disk (CDF)

What do we want to do with the CDF Grid ?

- Almost everything you can do on the central systems :
  - Populate UK disks with secondary datasets.
  - Reprocess secondary datasets.
  - Skim to create tertiary datasets.
  - Create standard and user ntuples from datasets.
  - Large volume Monte Carlo simulations.
- More importantly, we want these operations to be transparent to the user and the results to be available to everyone on CDF :
  - Metadata describing UK resident datasets visible everywhere (for example, by being logged in the Data File Catalogue).
  - Reprocessed and Monte Carlo data to be transferred back to Fermilab if required.

What do we want to do with the CDF Grid ?

- For example, a use-case for creation of a tertiary dataset :
  1. User prepares a “skimming” executable on own machine.
  2. User runs a web based browser to specify metadata parameters and hence obtain a list of datasets (e.g. “All Stream B data from runs containing > 1M events taken during January 2002”).
  3. User runs an application to specify and describe a new dataset (e.g. using the Dataset Registry GUI)
  4. User submits executable and dataset specifications to an application which directs the job to where the data resides.
  5. Data are written to disk as the skim proceeds. The location of the new dataset is updated concurrently in the Data File Catalogue.
  6. Monitoring tools enable the user to track the progress of the job.
  7. Upon completion the status of the job is returned to the user.

Requirements

- Use-cases such as these lead to many requirements on the "middleware" used to perform these operations :
  - Mechanisms must exist for optimising file transfer between sites (e.g. presenting enough information to decide whether tape or network transfer is most efficient).
  - Remote sites must make available in a standard fashion information about their resources (CPU, storage, network connectivity, etc.). Resource availability should be translated into costs (e.g. duration of specified tasks at different sites).
  - Databases must be capable of tracking data held at remote sites as well as at Fermilab, making it available in a uniform fashion.
  - Mechanisms to allow jobs from participating remote sites to be run locally, with appropriate priorities.
  - Many monitoring and security requirements.

CDF Tools

- Many of the required ingredients potentially already exist, at least in some form :
  - The **CDF code distribution** does a very good job of creating a uniform software environment at remote sites. **L3**, **runMC**.
  - The **Data File Catalogue** is a metadata store, in principal usable at many sites and capable of storing secondary dataset information.
  - The **Database GUI** allows the metadata to be queried in a very flexible fashion.
  - A **Dataset Registry GUI** allows creation of new dataset descriptions.
  - The **Disk Inventory Manager** allows control of local disk caches, in principal at many sites.
- However very little of this infrastructure has yet been used in a distributed fashion. There are many missing pieces.

Grid Tools

- There are many Grid tools available that can help us :
  - **Directory Services** : mechanisms by which participating sites can advertise available resources (CPU, data etc.)
  - **Data Replication Services** : distributed storage and caching for optimal performance of the applications requiring the data.
  - **Monitoring and Diagnostics Services** : convenient means of keeping track of distributed workloads.
- For example those provided by the Globus Toolkit :
  - **GridFTP** (parallel FTP for optimal file transfer)
  - **G**rid **R**esource **I**nformation **S**ervice (English : finding out what and where computing resources are available)
  - **G**rid **R**esource **A**ccess and **M**anagement (English : allocation of specific resources and monitoring the usage of these resources)
  - **Condor-G** (scheduling and remote job submission).

Possible Solutions

- In principle there are many ways to implement a Grid.
- [Mosix is a promising fabric component (D.Kant, QMW)]
- A detailed proposal was put forward (McArthur, Huffman, Reichold, Watts, Fisher, Sansum) that :
  - Would use local **DFC**s and **DIM**s at participating sites.
  - Would use **LDAP** (**L**ightweight **D**irectory Access Protocol) as a means of querying remote sites for available resources.
  - Could use other Grid tools such as GRIS, GridFTP, etc.
- D0 have a very sophisticated and advanced product in the shape of **SAM** which is already being used to harness distributed computing resources Þ Vicky White's talk.

Next Steps

- There are many avenues we would like to explore given sufficient manpower :
  - Getting the CDF Data Handling system working in a distributed, Grid like way.
  - Can CDF use SAM ? It is a very well developed product and does not assume anything about the nature of the underlying data. It can accommodate different metadata catalogues.
  - Using a framework for prototype Grids that has been developed by EU DataGrid ?
  - Combining different elements of the above ? (SAM is already being enhanced by the addition of Globus tools I believe).

Next Steps

- Right now we're crawling :
  - Looking on web pages to find which UK machines hold which files.
  - FTP'ing files around by hand.
  - Calling people on the phone to ask if we can have accounts on their machines.
- We're learning to walk :
  - Installation of the Globus Toolkit on UK machines.
  - Parallel FTP'ing of datasets across the Atlantic and within the UK.
- And we might even learn how to run :
  - There is a possibility of limited UK funding to contribute towards the development of a Grid for CDF.

Conclusions

- With significant exceptions, previous experiments have not utilised distributed computing resources in the most effective way.
- Future experiments have to do better.
- We are motivated by the presence of very significant computing resources both in Fermilab and in the UK to attempt to construct a CDF-UK Grid.
- D0 are a long way ahead with SAM.
- We have several ideas for how to proceed. As usual the constraints are temporal and financial.

Possible Solutions