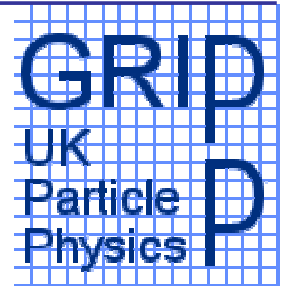


# QCDgrid status report

UKQCD data grid

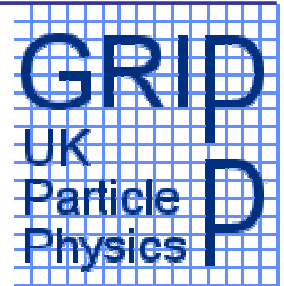
Chris Maynard

# QCDgrid outline



- QCDgrid is primarily a **data grid**
  - Store data online (disk)
  - Security through redundancy
    - Multiple copies of files in geographically distinct locations
  - Replica catalogue service (RC) links copies to logical file name (LFN)
- Data access through metadata catalogue service (MC)
  - Configurations marked up in QCDml1.1

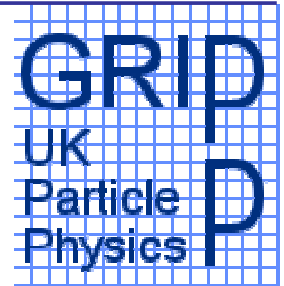
# Machines – LHC terminology



- Tier 1 system
  - 50TByte SAN system
    - QCDOC front end directly connected
  - Hold one copy of all UKQCD data
    - Machine exists, not yet part of grid
- Tier 2 systems
  - Edinburgh, Liverpool, Southampton, Swansea
    - Linux server + RAID array disk
    - Now: 5 TByte. Upgrade to 12.5 TByte
- 50 TByte capacity + 50 TByte replicated

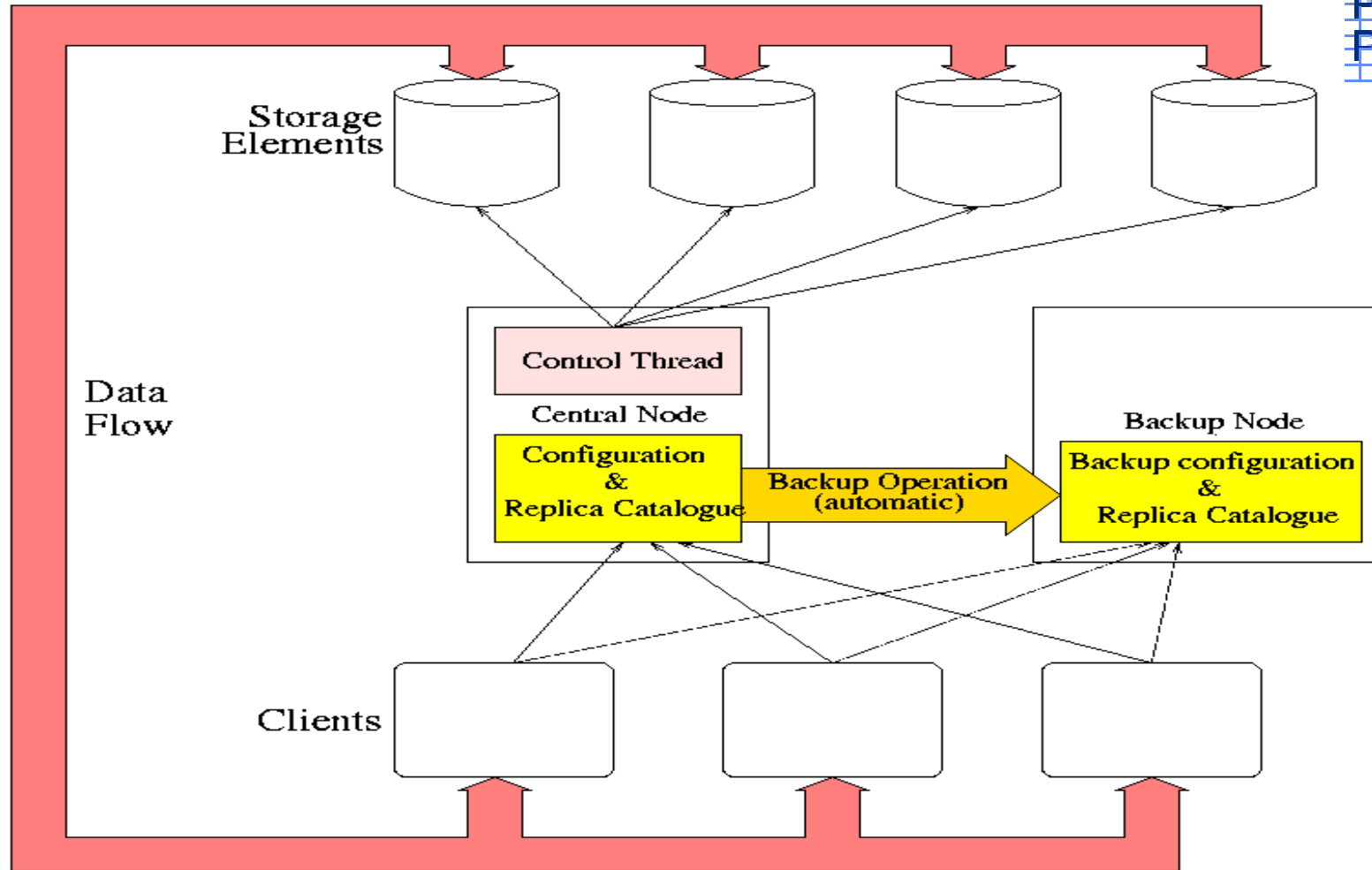
# Software

---



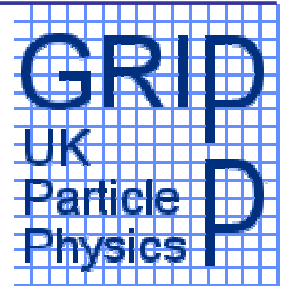
- Middleware: Globus 2.4
- RC: Globus RC
- MC: Native XML database eXist
- Bespoke Control Thread (CT)
- Command line client tools
- GUI client tools
  - Metadata browser
    - Based on OGSA-DIA browser
  - Job submission tool

# Control Thread



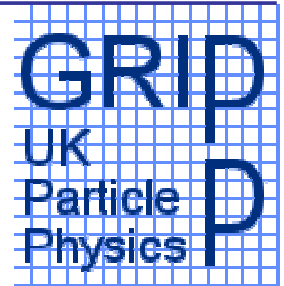
# Command line tools

---



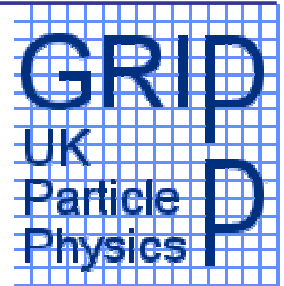
- For grid administrator
  - Add/remove machine
  - Disable/enable machine
  - Add/remove users (see later)
  - Remove Data
- For user
  - Add, retrieve, store locally, data
  - Job submission post-production on grid data
    - Remote machine not required to be QCDgrid

# QCDgrid GUI



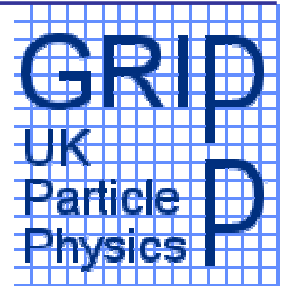
- Same functionality as command line
- Metadata browser
  - Search and read metadata
- Based on OGSA-DIA browser
  - Freely available from [qcdgrid.forge.nesc.ac.uk](http://qcdgrid.forge.nesc.ac.uk)
- Undergoing modification to take account of split schema in QCDml
- Demonstration
  - Can read MC without authentication

# MC Browser and QCDml



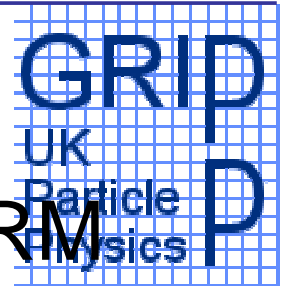
- QCDml1.1
  - Two schemata
    - Configuration XML
    - Ensemble XML
- MD catalogue and browser being modified
- QCDml split is logical split
  - Nearly always search on **Ensembles** not
  - Native XML database has few ensemble files
    - Relational DB good, Hierarchical DB Bad
    - Quick searching!

# Authenticated access



- QCDgrid data access
  - X509 certificate from **trusted** CA
    - Currently only UK e-Science CA
    - We can trust your CA!
  - QCDgrid is not closed
    - Not anonymous, but authenticated access.
  - ILDG data sharing
    - Technically could move to anonymous access
    - Require data sharing policy from ILDG board

# QCDgrid and ILDG



- ILDG RC services are SRM
  - QCDgrid RC catalogue published as SRM
    - This is not yet implemented
    - QCDgrid will appear as single ILDG node
- ILDG MD services
  - UKQCD configurations marked up in QCDml1.1
  - MC catalogue public anonymous access
  - Aggregate to obtain ILDG MC