LHCb Computing Strategy

Nick Brook

University of Bristol

- Computing Model
  - 2008 needs
  - Physics software
- Harnessing the Grid
  - DIRAC
  - GANSA
- Experience & Readiness
**Dataflow**

RAW data is reconstructed:
- e.g.
  - Calo. Energy clusters
  - Particle ID
  - Tracks ...

At reconstr time only enough info is stored to allow physics pre-selection to run at a later stage - reduced DST (rDST) - stored separately from RAW

Reconstruction performed twice a year:
- quasi real time
- after LHC shutdown
**Dataflow - Stripping**

Stripping performed 4 times per year

rDST is analysed in production-mode → event streams for further analysis; 20-30 streams

Algorithm developed by physics working groups - use as i/p rDST & RAW

Event to be output will have additional reconstructed info added: (full) DST+ RAW data

Event Tag Collection - created to allow “quick” access to data; contain “metadata”
Dataflow - Analysis

User physics analysis will be primarily performed on the output of the stripping.

Output from stripping is self-contained i.e. no need to navigate between files.

Analysis generates quasi-private data e.g. Ntuple and/or personal DSTs.

Data publicly accessible - enable remote collaboration.
Use of computing centres

Main user analysis supported at CERN+6 “Tier-1” centres

Tier-2 centres essentially Monte Carlo production facilities

Plan to make use of LHCb online farm for re-processing
2008 Resource Summary

Estimated $4 \times 10^6$ secs physics (including machine efficiency)
Assume $8 \times 10^9$ events from event filter farm to CERN computer centre

<table>
<thead>
<tr>
<th></th>
<th>CPU (2.8GHz P4 years)</th>
<th>Disk (TB)</th>
<th>Tape (TB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CERN</td>
<td>360</td>
<td>350</td>
<td>631</td>
</tr>
<tr>
<td>Tier-1’s</td>
<td>1770</td>
<td>1025</td>
<td>860</td>
</tr>
<tr>
<td>Tier-2’s</td>
<td>4550</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
LHCb software

LHCb data processing applications and data flow
LHCb software framework

Object diagram of the Gaudi architecture
**LHCb software framework**

- Gaudi is architecture-centric, customisable framework
  - Adopted by ATLAS; used by GLAST & HARP
  - Same framework used both online & offline
- Algorithmic part of data processing as a set of OO objects
  - decoupling between the objects describing the data and the algorithms allows programmers to concentrate separately on both.
  - allows a longer stability for the data objects (the LHCb event model) as algorithms evolve much more rapidly
- An important design choice has been to distinguish between a transient and a persistent representation of the data objects
  - changed from persistency solutions without the algorithms being affected.
- Event Model classes only contain enough basic internal functionality for giving algorithms access to their content and derived information
  - Algorithms and tools perform the actual data transformations
DIRAC - A community Grid solution

• The DIRAC Workload & Data Management System (WMS & DMS) is made up of Central Services and Distributed Agents

• The main aims of DIRAC are:
  - To integrate all of the heterogeneous compute resources available to LHCb
  - Minimize the human intervention at sites
  - Use worldwide LHC Computing Grid services wherever possible

• DIRAC realizes these goals via:
  - Pilot Agent paradigm
  - Overlay Network paradigm
**DIRAC Overlay Network Paradigm**

- DIRAC Agents are deployed close to resources
- Forms an overlay network of Agents masking the underlying diversity of the available compute resource
- Services interact with Agents

**Computing Resources**

- PCs
- Grid
- Site Clusters
DIRAC Overlay Network Paradigm

- DIRAC Agents are deployed close to resources
- Forms an overlay network of Agents masking the underlying diversity of the available compute resource
- Services interact with Agents

### Computing Resources

- Grid
- Site Clusters
- PCs
DIRAC Overlay Network Paradigm

- DIRAC Agents are deployed close to resources
- Forms an overlay network of Agents masking the underlying diversity of the available compute resource
- Services interact with Agents

Computing Resources

- PCs
- Site Clusters
- Grid
- Agents

HCP, Elba - May'07
DIRAC Overlay Network Paradigm

- DIRAC Agents are deployed close to resources
- Forms an overlay network of Agents masking the underlying diversity of the available compute resource
- Services interact with Agents

![Diagram of DIRAC Overlay Network Paradigm]

**Computing Resources**

- **Grid**
- **Site Clusters**
- **PCs**
- **Agents**

**Services**

- Service 1
- Service 2
- Service 3
DIRAC Pilot Agent Paradigm

- DIRAC is a **PULL** scheduling system
  - Agents first occupy a resource and then request jobs from a central task queue
  - This 'late binding' allows execution environment to be checked in advance
- Pilot Agents are sent to the gLite Resource Broker as normal jobs
  - Facilitate PULL approach on PUSH system
- LCG jobs are Pilot jobs in the context of the DIRAC WMS
  - Actual workload management performed by DIRAC
DIRAC Workload Management System

• Heterogeneous groupings of resources such as clusters / Grids become homogeneous via DIRAC
• DIRAC can therefore be viewed as a (very) large batch system
  - Accounting
  - Priority Mechanism
  - Fair share
**GANGA** - user interface to the Grid

- **Goal**
  - Simplify the management of analysis for end-user physicists by developing a tool for accessing Grid services with built-in knowledge of how Gaudi works

- **Required user functionality**
  - Job preparation and configuration
  - Job submission, monitoring and control
  - Bookkeeping browsing, etc.

- **Done in collaboration with ATLAS**

- **Use Grid middleware services**
  - Interface to the Grid via Dirac and create synergy between the two projects
**Ganga jobs**

- A job in Ganga is constructed from a set of building blocks, not all required for every job.
Ganga: how the pieces fit together

- Ganga has built-in support for ATLAS and LHCb
- Component architecture allows customisation for other user groups

Applications

- ATLAS applications
- LHCb applications
- Other applications

User interface for job definition and management

Metadata catalogues
- File catalogues
- Tools for data management

Data storage and retrieval

Ganga job archives

Ganga monitoring loop

Remote repository
Local repository

Experiment-specific workload-management systems

Local batch systems
Distributed (Grid) systems

Processing systems (backends)

Tools for data management
LHCb Simulation Production

- Typical MC Production job lasts 24hrs
- Recently achieved 10K concurrent production jobs
  - Throughput only limited by available capacity of LCG
- ~80 distinct sites accessed via Grid or directly

Sustained resource usage over extended periods of time
- System is stable for simulation
Breakdown of Production
LHCb Reconstruction Results

- April 2007, reconstruction jobs successfully running at all LHCb Tier-1 sites
  - CERN (Switzerland)
  - IN2P3 (France)
  - GridKa (Germany)
  - CNAF (Italy)
  - NIKHEF (Netherlands)
  - PIC (Spain)
  - RAL (U.K.)

- Reconstruction challenge is ongoing
  - Current issues include: site service instability; tape failures...
LHCb Analysis

596 unique GANGA Users
  • 99 users from LHCb
  • ~41k GANGA sessions since start of year
  • ~10k LHCb sessions
393k jobs passed through DIRAC analysis system since start of year

- Users happy with efficiency
- Access to large amount of resources
Summary

Computing model being finalised currently under stress test...
... particularly access to data

Software framework robust & mature
final version of reconstruction s/w available
3 different persistency solutions without major upheaval

LHCb DIRAC system
allows efficient use of Grid resources
Monte Carlo production now routine
Reconstruction under stress test

User analysis on the Grid
seeing increase in use
GANGA interface between s/w framework & DIRAC
Summary

Computing model being finalised
  currently under stress test...
  ... particularly access to data
Software framework robust & mature
  final version of reconstruction s/w available
  3 different persistency solutions without major upheaval
LHCb DIRAC system
  allows efficient use of Grid resources
  Monte Carlo production now routine
  Reconstruction under stress test
User analysis on the Grid
  seeing increase in use
  GANGA interface between s/w framework & DIRAC

Confident LHCb computing will be ready for data taking