Summary of the LHC Computing Review

http://lhc-computing-review-public.web.cern.ch

John Harvey
CERN/EP

May 10th, 2001
LHCb Collaboration Meeting
The Scale

- Data taking rate: 50,100, 200 Hz (ALICE, ATLAS-CMS, LHCb)
- Raw event size: 0.15 / 1 / 1-25 MB (LHCb/ATLAS-CMS / ALICE)
- Total raw data storage: 7.0 PB/yr
- Total simulated data storage: 3.2 PB/yr
- World-wide tape storage: 28.5 PB/yr 40 million CD-Rom’s
- World-wide disk storage: 10.4 PB/yr 100k disks @ 100 GB
- World-wide CPU capacity: 7350 kSI95 360k today’s PCs
- WAN bandwidth (Tier-0/-1): 5000 Mbps 4 experiments
Multi-Tier Hierarchical Model

Experiments

Tier-0 1

Tier-1 ~5 (national)
Region I
Region F
Region UK
Region D
CERN

Tier-2 ~15 (regional)
Institute
Institute
Institute MAP
Institute
Institute

Tier-3 ~50 Desktop server
Institute Server
Institute Server
Institute Server
Institute Server

..
Multi-Tier Model (MONARC)

- Tier 0 (CERN)
  - Production centre for real data, large storage capacity
  - Data distribution to Tier 1s (AOD, samples of RAW+ESD)
- Tier 1 (CERN)
  - Physics analysis
  - Production centre for simulation (shared with regional centres)
- Tier 1 regional centres
  - Production centre for simulation
  - Data storage and distribution (robotics and network)
  - Physics analysis
  - Collaboration wide resource (GRID) - access policy needed!
- Tier 2 - special facilities for restricted production work
  - Production analysis and simulation samples, physics analysis
  - Data distribution to Tier 1 (network)
- Distribution guideline - 1/3 each for Tier 0, Tier 1, Tier 2
## Rates and Installed Capacities

<table>
<thead>
<tr>
<th></th>
<th>ALICE</th>
<th>ATLAS</th>
<th>CMS</th>
<th>LHCb</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event size (MB)</td>
<td>25</td>
<td>2</td>
<td>1</td>
<td>0.125</td>
<td></td>
</tr>
<tr>
<td>Raw data/year (PB)</td>
<td>2.7</td>
<td>8.1</td>
<td>1.7</td>
<td>0.25</td>
<td>13.0</td>
</tr>
<tr>
<td>MC data/year (PB)</td>
<td>0.2</td>
<td>1.5</td>
<td>1.2</td>
<td>0.36</td>
<td>3.3</td>
</tr>
<tr>
<td>Tape at CERN (TB)</td>
<td>3200</td>
<td>8959</td>
<td>1540</td>
<td>912</td>
<td>14611</td>
</tr>
<tr>
<td>Disk at CERN (TB)</td>
<td>534</td>
<td>410</td>
<td>1143</td>
<td>330</td>
<td>2417</td>
</tr>
<tr>
<td>CPU at CERN (kSI95)</td>
<td>824</td>
<td>690</td>
<td>820</td>
<td>225</td>
<td>2559</td>
</tr>
<tr>
<td>Tape worldwide (TB)</td>
<td>4700</td>
<td>19800</td>
<td>10500</td>
<td>2800</td>
<td>37900</td>
</tr>
<tr>
<td>Disk worldwide (TB)</td>
<td>1600</td>
<td>2570</td>
<td>5900</td>
<td>1100</td>
<td>11070</td>
</tr>
<tr>
<td>CPU worldwide (kSI95)</td>
<td>1758</td>
<td>1944</td>
<td>2907</td>
<td>925</td>
<td>7535</td>
</tr>
<tr>
<td>WAN Tier0/Tier1 (Mb)</td>
<td>1500</td>
<td>1500</td>
<td>1500</td>
<td>310</td>
<td>4810</td>
</tr>
</tbody>
</table>

See spreadsheets for details of LHCb numbers
lhcb-comp.web.cern.ch/lhcb-comp/computingmodel/Requirements&Costs/requirements.htm
Hardware costs

- Hardware costs of initial setup of LHC distributed computer centres (Tiers 0, 1 and 2) is **240 MSFr**
  - LHCb cost estimate is **27 MSFr** i.e. ~11% of total
- CERN-based Tier 0/Tier 1 centre ~ 1/3 of total
- Significant uncertainties in performance of LHC, detectors, triggers, backgrounds, algorithms etc.
- Investment for initial system to be spent in 2005, 2006 and 2007 in ~equal portions (30,30,40)
- Maintenance & Operations (M&O) of LHC computing system
  - Rolling replacement within constant budget
  - Requires ~1/3 of initial investment per year (~80 MSFr)
  - Includes steady evolution of capacity
- Current cost estimates based on forecast evolution of price and performance of computer hardware
### Hardware costs of CERN Computing ’05-’07

#### Projected Evolution of Processor Performance, Price/Performance

- **Estimated Cost - CHF/195**
  - 1999: 10, 20, 30, 40, 50, 60, 70
  - 2000: 20, 40, 60, 80, 100, 120
  - 2001: 30, 60, 90, 120, 150
  - 2002: 40, 80, 120, 160, 200
  - 2003: 50, 100, 150, 200, 250
  - 2004: 60, 120, 180, 240, 300
  - 2005: 70, 140, 210, 280, 350
  - 2006: 80, 160, 240, 320, 400
  - 2007: 90, 180, 270, 360, 450
  - 2008: 100, 200, 300, 400, 500
  - 2009: 110, 220, 330, 440, 550
  - 2010: 120, 240, 360, 480, 600

- **Cost - CHF/GB**
  - Year: 2000: 0.01, 0.02, 0.03, 0.04, 0.05, 0.06
  - Year: 2001: 0.07, 0.08, 0.09, 0.10, 0.11, 0.12
  - Year: 2002: 0.13, 0.14, 0.15, 0.16, 0.17, 0.18
  - Year: 2003: 0.19, 0.20, 0.21, 0.22, 0.23, 0.24
  - Year: 2004: 0.25, 0.26, 0.27, 0.28, 0.29, 0.30
  - Year: 2005: 0.31, 0.32, 0.33, 0.34, 0.35, 0.36
  - Year: 2006: 0.37, 0.38, 0.39, 0.40, 0.41, 0.42
  - Year: 2007: 0.43, 0.44, 0.45, 0.46, 0.47, 0.48
  - Year: 2008: 0.49, 0.50, 0.51, 0.52, 0.53, 0.54
  - Year: 2009: 0.55, 0.56, 0.57, 0.58, 0.59, 0.60
  - Year: 2010: 0.61, 0.62, 0.63, 0.64, 0.65, 0.66

<table>
<thead>
<tr>
<th>Units</th>
<th>ALICE</th>
<th>ATLAS</th>
<th>CMS</th>
<th>LHCb</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>11069</td>
<td>10667</td>
<td>12667</td>
<td>3479</td>
</tr>
<tr>
<td>Disk</td>
<td>2188</td>
<td>1907</td>
<td>5314</td>
<td>1535</td>
</tr>
<tr>
<td>Robotic Tape</td>
<td>3200</td>
<td>9407</td>
<td>1617</td>
<td>958</td>
</tr>
<tr>
<td>Shelf Tape</td>
<td>0</td>
<td>0</td>
<td>1816</td>
<td>214</td>
</tr>
<tr>
<td>Total Cost</td>
<td>18073</td>
<td>23692</td>
<td>23135</td>
<td>7040</td>
</tr>
</tbody>
</table>

LHCb Tier-1/2’s 20152 kSFr (74%)
Common Prototype

- Intended to setup a common prototype as a joint project
  - Experiments, CERN, major regional centres all involved
  - Reaching ~50% (in complexity) of overall computing structure of 1 of the large LHC experiments by ~2003/4
- Use as testbed to test at realistic scales
  - Scalability tests of CPU and I/O performance
  - Evaluate new technologies - Copper gigabit; new tapes, IA-64
  - Software tests - fabric management, grid middleware
- To be used in LHCb data challenges
  - Stress test of data processing software - simulation, reconstruction and analysis
  - Stress test of production tools
  - Stress test of ‘chaotic’ access patterns to event database via analysis jobs
  - Perform data challenges of increasing size and complexity
    - July ’02, July ’03, ‘July ’04
Software

- Insufficient support for simulation packages and analysis tools (e.g. FLUKA and ROOT)
- Core software teams in experiments severely understaffed
- Planned reduction of CERN-IT staff incompatible with CERN-based LHC computing system and software support
### Manpower needs (FTEs) for CORE Software

<table>
<thead>
<tr>
<th></th>
<th>2000 Have (miss)</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALICE</td>
<td>12(5)</td>
<td>17.5</td>
<td>16.5</td>
<td>17</td>
<td>17.5</td>
<td>16.5</td>
</tr>
<tr>
<td>ATLAS</td>
<td>23(8)</td>
<td>36</td>
<td>35</td>
<td>30</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>CMS</td>
<td>15(10)</td>
<td>27</td>
<td>31</td>
<td>33</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>LHCb</td>
<td>14(5)</td>
<td>25</td>
<td>24</td>
<td>23</td>
<td>22</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>64(28)</td>
<td>105.5</td>
<td>106.5</td>
<td>103</td>
<td>100.5</td>
<td>99.5</td>
</tr>
</tbody>
</table>

**Only computing professionals counted**

- CERN/IT - current staff complement 187
- minimum required to run centre 157
- predicted complement in 2006 137
### Manpower LHCb Core software and Computing

<table>
<thead>
<tr>
<th>Task</th>
<th>Profile</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Software Framework</strong></td>
<td>Engineer</td>
<td>8(3)</td>
<td>9(4)</td>
<td>8(3)</td>
<td>6(2)</td>
<td>5(1)</td>
<td>5(1)</td>
</tr>
<tr>
<td>basic software components</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application Frameworks</td>
<td>Physicist</td>
<td>6(2)</td>
<td>9(4)</td>
<td>9(4)</td>
<td>8(4)</td>
<td>8(4)</td>
<td>8(4)</td>
</tr>
<tr>
<td>simulation,reconstruction,analysis, event display</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software Support</td>
<td>Engineer</td>
<td>2(0)</td>
<td>4(1)</td>
<td>4(1)</td>
<td>4(1)</td>
<td>4(1)</td>
<td>4(1)</td>
</tr>
<tr>
<td>code mgt &amp; distribution, testing, quality control, documentation, production tools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computing Facilities</td>
<td>Engineer</td>
<td>3(1)</td>
<td>3(2)</td>
<td>3(2)</td>
<td>5(3)</td>
<td>5(3)</td>
<td>5(3)</td>
</tr>
<tr>
<td>Event Filter Farm, LAN, CDR, GRID, OS management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total CORE Computing</td>
<td></td>
<td>19(6)</td>
<td>25(11)</td>
<td>24(10)</td>
<td>23(10)</td>
<td>22(9)</td>
<td>22(9)</td>
</tr>
</tbody>
</table>
## Manpower DAQ and ECS

<table>
<thead>
<tr>
<th>Task</th>
<th>Profile</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Readout Unit and detector links</td>
<td>Engineer</td>
<td>2(0)</td>
<td>2(1)</td>
<td>2(1)</td>
<td>2(1)</td>
<td>2(1)</td>
<td>2(1)</td>
</tr>
<tr>
<td>Event Building</td>
<td>Engineer</td>
<td>2(0)</td>
<td>2(0)</td>
<td>3(1)</td>
<td>3(1)</td>
<td>3(1)</td>
<td>2(1)</td>
</tr>
<tr>
<td>Timing and Fast Control (TFC)</td>
<td>Engineer</td>
<td>1(0)</td>
<td>2(0)</td>
<td>2(0)</td>
<td>2(0)</td>
<td>1(0)</td>
<td>1(0)</td>
</tr>
<tr>
<td>ECS interface to electronics</td>
<td>Engineer</td>
<td>0(0)</td>
<td>1(1)</td>
<td>1(1)</td>
<td>1(1)</td>
<td>0(0)</td>
<td>0(0)</td>
</tr>
<tr>
<td>Hardware support and installation</td>
<td>Technician</td>
<td>0(0)</td>
<td>0(0)</td>
<td>0(0)</td>
<td>1(1)</td>
<td>1(1)</td>
<td>1(1)</td>
</tr>
<tr>
<td>Data monitoring framework and DAQ applications</td>
<td>Engineer</td>
<td>0(0)</td>
<td>0(0)</td>
<td>0(0)</td>
<td>2(2)</td>
<td>2(2)</td>
<td>2(2)</td>
</tr>
<tr>
<td>Controls framework, database and applications</td>
<td>Engineer</td>
<td>1(0)</td>
<td>3(1)</td>
<td>3(1)</td>
<td>4(2)</td>
<td>4(2)</td>
<td>4(2)</td>
</tr>
<tr>
<td>Operations</td>
<td>Physicist</td>
<td>0(0)</td>
<td>0(0)</td>
<td>0(0)</td>
<td>0(0)</td>
<td>1(1)</td>
<td>2(2)</td>
</tr>
<tr>
<td>Total DAQ/ECS</td>
<td></td>
<td>7(0)</td>
<td>10(3)</td>
<td>11(4)</td>
<td>15(8)</td>
<td>14(8)</td>
<td>14(9)</td>
</tr>
</tbody>
</table>
General recommendations

- Setup committee (SC2) to oversee LHC Computing Project composed of highest level software and computing management in experiments, CERN-IT and regional centres to oversee the deployment of the entire LHC computing infrastructure
  - Response from CERN management in preparation
- Each collaboration must prepare an MoU for LHC computing describing funding and responsibilities for hardware and software including human resources.
- Interim MoUs or software agreements should be setup by the end of 2001 to ensure appropriate development of the software
  - CMS have in mind an IMoU
  - ATLAS have pursued the idea of formal software agreements for some time
Software Projects

Software Framework (GAUDI)
- Event model - development and optimisation
- Detector description - development and optimisation of geometry
- Scripting component to allow interactive analysis based on PYTHON
- Grid services
- Data management (event data, conditions data, bookkeeping)

Software support
- software test, quality and performance; data quality monitoring
- Documentation support: workbooks, templates, web

Computing Facilities
- Development of analysis model
- Control and management of event filter farm
- Technical support at pit - farm, LAN, installation, commissioning etc

Physics application frameworks
- Simulation program - project leader
- High Level Trigger - project leader, HLT framework
- Analysis Program - project leader
- Event Display - project leader
DAQ/ECS Projects

- Readout unit, links - engineer
- Event builder prototyping and testing
- ECS interface to electronics (CC-PC) - software engineer
- Slow controls software framework and utilities
- Configuration databases and utilities
- Hardware support and installation; from '03
- Data monitoring framework and utilities; from '03
- DAQ applications - run control, error handling; from '04
- Operations - LHC interface, utilities; from '04
Observations and conclusions

- Waiting for response from CERN management
  - guidelines on construction and cost sharing of prototype
  - timescale for Computing TDR and MoU
  - allocation of additional new effort to IT and experiments
  - role and composition of SC2 and timescale for launch
    - Data management project already in preparation

- Communication with funding agencies
  - Discussions at LHCC, RRBs - preparation of IMoU
  - Responsibilities for core software (sharing policy)
  - Advance notice of long term computing plan (cost sharing)
  - Policy of access to centres outside CERN

- Preparation of distributed computing infrastructure
  - Development of analysis model - physics use-cases
  - Development of grid services - integration in GAUDI
  - Preparation of data challenges
Missing Manpower for CORE Computing

- Provision of core software and computing infrastructure is a collaboration wide responsibility
- Entering intensive development phase now - commitments needed soon
- Agreement on how to share responsibilities will greatly facilitate process of filling missing roles

**Formal**

1. IMoU approach (CMS) - sharing algorithm based on number of scientists
2. Software Agreements (ATLAS) - formal contracts whereby an institute agrees to take responsibility for a project
3. Level of effort - a country (or institute) agrees to maintain a contribution of n FTEs for core computing activities