Collaborative Computing at CERN: The distributed development of Grid infrastructure.

Dr. Will Venters and Dr Tony Cornford.

Information Systems and Innovation Group
Department of Management
The London School of Economics
w.venters@lse.ac.uk

This work was undertaken as part of Pegasus
EPSRC: Grant No: EP/D049954/1
www.pegasus.lse.ac.uk
The Pegasus Project

- 4yrs project.
- Looking at work practices
- Qualitative longitudinal research through studies of work practices – ethnographic focus.
Run Number: 152221, Event Number: 383185
Date: 2010-04-01 00:31:22 CEST

$\mathbf{W} \rightarrow \mu\nu$ candidate in 7 TeV collisions

$p_T(\mu^+) = 29$ GeV
$\eta(\mu^+) = 0.66$
$E_T^{\text{miss}} = 24$ GeV
$M_T = 53$ GeV
Why do particle physicists need the Grid?

Starting from this event

looking for this “signature”

Selectivity: 1 in $10^{13}$

Like looking for 1 person in a thousand world populations
Or for a needle in 20 million haystacks!

One year's data from LHC would fill a stack of CDs 20km high

- Concorde (15 Km)
- Mt. Blanc (4.8 Km)

- $\sim 100,000,000$ electronic channels
- $800,000,000$ proton-proton interactions per second
- $0.0002$ Higgs per second
LHC Computing Grid

- 12-15 PB per year (at full LHC).
- 40PB disk, 40PB tape. 100,000 processors.
- 170 sites, 34 countries, 8000 physicists.
- Our interest is in the working practices of those involved in the design and use of this Grid.
Background Context

• **Building the LHC Computing Grid (LCG):**
  - Highly distributed, complex and poorly defined systems development task.
  - Cutting edge hardware and software used.
  - New software standards being negotiated.
  - Middleware and support software being developed in a range of languages.
  - Grid is distributed and proceed at different paces because of funding.

• Particle physics has a long tradition of such large scale global collaborations (Traweek 1988).

**GridPP (UK Contribution to LCG):**

• To a significant degree agile...
• Collaboration of 230 people in 19 UK universities, RAL and CERN.
• Decisions are made democratically and consensually, and implemented by influence and persuasion.
• Network rather than hierarchy
• Virtual, federated, overlapping and inter-connected.
• Virtual meetings, wikis, blogs, mailinglists
Pegasus Framework

- Technical Grid
- Practices
- Knowledge Infrastructure
- Socio-technical Grid
Pegasus Framework

Technical Grid

Practices

Knowledge Infrastructure

Socio-technical Grid
Technical Grid

Computing Elements (processors)
Storage Elements (disks and tapes)
Networking (Fibre-optic, Gigabit).
Middleware – “the operating system”
Monitoring systems...
Pegasus Framework

Socio-technical Grid

Technical Grid

Knowledge Infrastructure

Practices
Practice of Scaled Agility

We questioned:
– Who they are?
– Their values and motivations?
– How they work?

The sense of agility – their practices are agile...

• “amethodological” (Truex et al 2000).
• Improvisational, bricolage, fluid.
• Scaled, distributed, loosely controlled.
• Pragmatic “make it work”
• Aesthetic of imperfection (Weick)
• Our view is of paradox –
## Practices: Improvisational Paradoxes

<table>
<thead>
<tr>
<th>Category</th>
<th>Theme</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belonging</td>
<td>Collective Individuality</td>
<td>Tension between individual freedom and group cohesion, emergent from &quot;handed out&quot; and membership of the collective.</td>
</tr>
<tr>
<td></td>
<td>Anxious Confidence</td>
<td>to live with ambiguity, complexity and challenges of working in an improvisational collective.</td>
</tr>
<tr>
<td>Learning</td>
<td>Learned Improvisation</td>
<td>Improvisation based on past experience and situated within environmental constraints</td>
</tr>
<tr>
<td></td>
<td>Reflective Spontaneity</td>
<td>Making sense via <em>ex post</em> interpretation and rationalization.</td>
</tr>
<tr>
<td>Organizing</td>
<td>Planned Agility</td>
<td>Tension between deliberate action of planning and the uncontrolled processes of drifting and unfolding.</td>
</tr>
<tr>
<td></td>
<td>Structured Chaos</td>
<td>Tension between the chaotic day-to-day practices and minimal structures which serve as a medium of the practices of exploration and trial-and-error.</td>
</tr>
</tbody>
</table>
Pegasus Framework

Technical Grid

Practices

Knowledge Infrastructure

Socio-technical Grid
Knowledge Infrastructure

- Stuff/Matter
- Social Structures
- Knowledge
- History & Culture

Knowledge Infrastructure
Knowledge

• A clearly understood Purpose; Higgs boson, and Nobel Prizes
• ‘Humming’ ‘with itself, about itself’ (Knorr-cetina 1999)
• Key ‘knowledgeable’ individuals.
• Language and acronyms of the project
• Informal discussions “in the pub”.

www.pegasus.lse.ac.uk – SSIT Workshop June 2010
Stuff/Matter
GridPP2 Goal: To develop and deploy a large scale production quality grid in the UK for the use of the Particle Physics community.
# UK Grid Status at 26 Aug 2008 09:02:07


**Resource Broker Summary**

<table>
<thead>
<tr>
<th>Institute</th>
<th>CPU Tot</th>
<th>CPU Free</th>
<th>Jobs Cur</th>
<th>Jobs Wait</th>
<th>Disk Tot</th>
<th>Disk Free</th>
<th>GOC Status</th>
<th>SAM Tests</th>
<th>LHCb Tests</th>
<th>SE</th>
<th>FCR</th>
<th>ATLAS Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trinity Dublin</td>
<td>692</td>
<td>368</td>
<td>326</td>
<td>159</td>
<td>35.5</td>
<td>33.1</td>
<td>P P P</td>
<td>W P P</td>
<td>100% 99%</td>
<td>P 100% 100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brunel</td>
<td>409</td>
<td>8</td>
<td>190</td>
<td>1152</td>
<td>10.7</td>
<td>2.2</td>
<td>P P P</td>
<td>P P P</td>
<td>100% 100%</td>
<td>P 100% 94%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imperial HEP</td>
<td>245</td>
<td>308</td>
<td>64</td>
<td>391</td>
<td>47.0</td>
<td>14.0</td>
<td>W P P</td>
<td>W P P</td>
<td>100% 78%</td>
<td>P 100% 95%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imperial LeSC</td>
<td>200</td>
<td>303</td>
<td>40</td>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
<td>P P P</td>
<td>P P P</td>
<td>100% 93%</td>
<td>P 0% 0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QMUL</td>
<td>1480</td>
<td>1709</td>
<td>18</td>
<td>0</td>
<td>33236.1</td>
<td>29935.0</td>
<td>P P P</td>
<td>P P P</td>
<td>50% 79%</td>
<td>P 100% 100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RHUL</td>
<td>400</td>
<td>1</td>
<td>375</td>
<td>305</td>
<td>149.0</td>
<td>72.2</td>
<td>P P P</td>
<td>P P P</td>
<td>100% 99%</td>
<td>P 100% 96%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCL CCC</td>
<td>2560</td>
<td>1026</td>
<td>938</td>
<td>0</td>
<td>309.1</td>
<td>304.0</td>
<td>M P P</td>
<td>M P P</td>
<td>0% 0%</td>
<td>0% 0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCL HEP</td>
<td>115</td>
<td>110</td>
<td>5</td>
<td>0</td>
<td>1.0</td>
<td>0.3</td>
<td>M P P</td>
<td>M P P</td>
<td>0% 12%</td>
<td>0% 0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lancaster</td>
<td>380</td>
<td>333</td>
<td>47</td>
<td>1</td>
<td>85.4</td>
<td>70.4</td>
<td>P P P</td>
<td>P P P</td>
<td>23% 78%</td>
<td>P 100% 100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liverpool</td>
<td>479</td>
<td>407</td>
<td>73</td>
<td>0</td>
<td>34.6</td>
<td>30.1</td>
<td>F P P</td>
<td>P P P</td>
<td>100% 100%</td>
<td>P 100% 94%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manchester</td>
<td>1740</td>
<td>795</td>
<td>797</td>
<td>2</td>
<td>1038.1</td>
<td>41.4</td>
<td>P P P</td>
<td>P P P</td>
<td>100% 100%</td>
<td>P 100% 94%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheffield</td>
<td>200</td>
<td>188</td>
<td>13</td>
<td>0</td>
<td>11.8</td>
<td>8.6</td>
<td>P P P</td>
<td>P P P</td>
<td>100% 98%</td>
<td>P 100% 94%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Click below on an institute name for a summary for that institute.

See also: [LondonGrid](#) [NorthGrid](#) [ScotGrid](#) [SouthGrid](#)
Social Structures

• “Clusters of Competence” (Kyriakidou 2009)
• Virtual, federated, overlapping and inter-connected.
• Drawing on long tradition of collaboration (Traweek 1988).
History and Culture

• “Down-to-earth and creative approaches embedded in particle physics tradition” (Traweek, 1988).
• Stories of past experiments – including deaths & mistakes.
• Academic publication as retrospective narrative.
• Long history of success in computing; Cray X-MP; WWW; Open source; linux; Gigabit networks…
• Culture of Science – even within the computing area.
Lessons we can learn

• An example of an unusual but effective form of “Scaled Agility”
• Competition… with redundancy.
• Clusters of competence.
• Aesthetic of imperfection.
• Anxious confidence.
• Culture across space and time!
References

Consequential Damage

Multi kA electrical arc

Magnets Displaced

Beam Screen (BS): The red color is characteristic of a clean copper surface. BS with some contamination by super-isolation (MLI multi layer insulation). BS with soot contamination. The grey color varies depending on the thickness of the soot, from grey to dark.

Vacuum Chamber Contamination
Surface activities
Ongoing Work

Add pressure release valves

Magnet Stand Improvements

Use of a foam plug wet with alcohol
Up to 50 passages in each direction

Clean Contaminated Vacuum Chamber
GridPP Middleware includes

Grid Data
- File Metadata
  - Logical File Name
  - GUID
  - System Metadata (owner, permissions, checksum, ...)
- User Metadata
  - User Defined Metadata
- File Replica
  - Storage File Name
  - Storage Host
- Symlinks
  - Link Name

Network Monitoring
- UK national R&D network infrastructure
  - International Connectivity
  - National Interconnection

Workload Management
- Integrated over all YOs and RBs:
  - Successes/Day: 15225
  - Success %: 69%
  - Improving from 42% to 70-80% during 2005
  - Problems Identified:
    - half WMS (Grid)
    - half JDL (User)

Information Services
- R-GMA Web Services
  - API available for Java, C, C++ and Python
  - Users may by-pass API if they wish, but API is the easiest way to use R-GMA services

Security
- Security in LCG/EGEE
  - JRA3
  - JRA1
  - NA4
  - Middle Security Group
  - defines/maintains policy and procedures
  - For LCG GGD and EGEE SA1

Storage Interfaces
- SRM
  - A single SRM server to service incoming file requests (this is implemented as a web service)
  - Multiple file servers with 3 NFS file systems on which data resides.
  - Data transfer is done to/from the file servers, thus inbound IP connectivity is essential to make the SRM SE available to the wider grid.

www.pegasus.lse.ac.uk – SSIT Workshop June 2010
WorldWide Resources

- **Worldwide**
  - 55 Countries
  - 283 Sites
  - 180,000 CPUs

- **UK**
  - 23 Sites
  - 20,000 CPUs

*Latest SAM results, Site Status, for OOPS VO, 10 Feb 2009 14:19 GMT. Size of site rectangles is number of CPUs from BDII. Certified Production sites, grouped by regions.*
Top 10 Facts

1. When the 27km long circular tunnel at CERN was excavated, between lake Geneva and the Jura mountain range, the two ends met up with just one centimetre of error.

2. When protons arrive in the LHC they are travelling at 0.999997828 times the speed of light. Each proton goes around the 27km ring over 11 000 times a second.

3. A nominal proton beam in the LHC will have an energy equivalent to a person in a car driving at 1700 kph.

4. The combined strands of the superconducting cable being produced for the LHC would go around the equator 6.8 times.

5. Part of the LHC will be the world's largest fridge. It could hold 150 000 fridges at a temperature colder than deep outer space.

6. The vacuum in the LHC is comparable to outer space, if it were a car tyre with a leak, there are so few gas molecules that it would take 10 000 years to go flat.

7. The ATLAS cavern could hold the nave of Notre Dame cathedral.

8. Understanding all of the different forces of nature within one framework is one of the ultimate goals of physics. Decades of experiments at CERN are working towards this.

9. On 1st October 2003 CERN and the California Institute of Technology set a new Internet Land Speed Record by transferring 1.1 terabytes of data in less than 30 minutes across 7000km of network. The equivalent of transferring a full length DVD movie in 7 seconds.

10. The Grid is a service for sharing computer power and data storage capability over the Internet. Its ultimate aim is to turn the global network of computers into one vast computational resource.