Predicting The Future
...or...
What We’ve Done In The Past

Keith Norman
Tessella Support Services plc
Overview

- Tessella
- Some distributed projects
  - Climate Prediction
  - JAC @ JET
  - Tessella GTI
- Observations
Tessella’s Background

- Software services company serving a wide customer base
- Formed in 1980, independent
- 140+ Technical Staff
  - Programmers, Consultants & Project Managers
  - All trained in software engineering
- ~£10M Turnover (2005/06)
- ISO 9001 & TickIT since 1992
- “Investor in People” Status
- Long-term client partnerships
climateprediction.net

- Met Office Hadley Model
  - Widely used climatology code
  - UNIX-based
  - Various starting values and “tweaks” possible
  - Years of model time = many hours of supercomputer time
Distributed Approach

- Many years of model time = many weeks of PC time…
- …but many more PCs than supercomputers!
- Many PC CPU cycles unused (3GHz machines running Word + Outlook)
Collaboration Team

- Oxford University Atmospheric, Oceanic and Planetary Physics
- Oxford University ComLab
- Reading University, RAL
- Tessella, Met Office
- NAG, Open University, et al
- DTI funded
climateprediction.net launch
climateprediction.net launch
Application to climateprediction.net
Current status
Join the climateprediction.net experiment!

What is climateprediction.net?
Climateprediction.net is the largest experiment to try and produce a forecast of the climate in the 21st century. To do this, we need people around the world to give us time on their computers - time when they have their computers switched on, but are not using them to their full capacity.

Why?
Climate change, and our response to it, are issues of global importance, affecting food production, water resources, ecosystems, energy demand, insurance costs and much else. There is a broad scientific consensus that the Earth will probably warm over the coming century. climateprediction.net aims, for the first time, to tell us what is most likely to happen.

What do we want you to do?
You can download a climate model from this website. It will run automatically as a background process on your computer whenever you switch your computer on. It should not affect any other tasks you use your computer for. At the model runs, you can watch the weather patterns on your, unique, version of the world evolve. The results are sent back to us via the internet, and you will be able to see a summary of your results on this web site. Climateprediction.net uses the same underlying software, BOINC, as many other distributed computing projects. If you like, you can participate in more than one project at a time.

If you are in the climate research community and are interested in participating in the experiment in a research capacity, the research pages provide some basic background material.
Other Distributed Computing

- seti@home
- folding@home
- Many other internet based
- Beowulf clusters – different set of issues, but in many ways easier to control
Types of Implementation

Server room cluster (dedicated or non-dedicated)

Management control

Office-distributed cluster (non-dedicated)

Leverage existing resources

Geographically-distributed cluster (non-dedicated)

World’s your oyster
JET

- World’s largest fusion research facility
- Modelling & “number crunching”
- Simulating plasma inside torus in 3d
JET Analysis Cluster

- 99 Linux nodes (114 processors)
- Cheap PCs, no screen, no keyboard – few £100
- Centrally managed
- 85 Gflops
- Mainly MPI applications
- 200 concurrent users - requires load balancing
JAC

Early Days …

JAC in April 2002
JAC

- Rack mounted in server room
- Fast ethernet/Gigabit connections between nodes
- Dedicated Tessella systems administrator
- Tessella involved in many application codes
Tessella GTI

- “Grid Technology Infrastructure”
- Leverage existing desktop and server room resources
- LAN and WAN based
- Based on Condor
- Added web interface, security, web services, …
Tessella GTI

- Runs processes low priority on clients
  - Windows, Linux, Solaris, …
- Suitable for exes, dlls, standalone Excel spreadsheets, …
- Programmatic interface via web services
- Secure logon
Submit With Input Files
Review Progress

---

### GTI Job Status - Job Number: 29

<table>
<thead>
<tr>
<th>Process</th>
<th>Holdtime</th>
<th>Location</th>
<th>Run Start</th>
<th>Full details</th>
<th>SLCT?</th>
</tr>
</thead>
<tbody>
<tr>
<td>RandomWalk 1 Dm</td>
<td></td>
<td></td>
<td>2004-10-21 14:08:29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Powered by "Tessella" 14:29.18 12-Oct-04

---

[Image of job status screen with details of processes and job status]
Check Pool Status

List of all machines in currently running in the GTI pool

<table>
<thead>
<tr>
<th>Hostname</th>
<th>Location</th>
<th>OpSys</th>
<th>Architecture</th>
<th>Status</th>
<th>Memory</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>pika</td>
<td>Cambridge</td>
<td>INTEL</td>
<td>UNCLAIMED</td>
<td>129</td>
<td>Full details</td>
<td></td>
</tr>
<tr>
<td>puck</td>
<td>Warrington</td>
<td>INTEL</td>
<td>CLAIMED</td>
<td>304</td>
<td>Full details</td>
<td></td>
</tr>
<tr>
<td>paper0</td>
<td>Abingdon</td>
<td>INTEL</td>
<td>CLAIMED</td>
<td>511</td>
<td>Full details</td>
<td></td>
</tr>
<tr>
<td>vm1@phosphor</td>
<td>Abingdon</td>
<td>INTEL</td>
<td>CLAIMED</td>
<td>255</td>
<td>Full details</td>
<td></td>
</tr>
<tr>
<td>vm2@phosphor</td>
<td>Abingdon</td>
<td>INTEL</td>
<td>CLAIMED</td>
<td>255</td>
<td>Full details</td>
<td></td>
</tr>
<tr>
<td>pmb</td>
<td>Dyston (UQ)</td>
<td>INTEL</td>
<td>CLAIMED</td>
<td>511</td>
<td>Full details</td>
<td></td>
</tr>
<tr>
<td>vm2@glacier</td>
<td>Abingdon</td>
<td>INTEL</td>
<td>UNCLAIMED</td>
<td>511</td>
<td>Full details</td>
<td></td>
</tr>
<tr>
<td>penguin</td>
<td>Abingdon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

List of all machines missing from the GTI pool

<table>
<thead>
<tr>
<th>Hostname</th>
<th>Location</th>
<th>Code</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>penguin</td>
<td>Abingdon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>paper0</td>
<td>Abingdon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>paper1</td>
<td>Cambridge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>paper2</td>
<td>Abingdon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>paper3</td>
<td>Abingdon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>paper4</td>
<td>Abingdon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>paper5</td>
<td>Abingdon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pmb</td>
<td>Dyston (UQ)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vm2@glacier</td>
<td>Abingdon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>penguin</td>
<td>Abingdon</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Download Results
Applications

- Flood prediction
- Protein matches
- Particle physics
- Graphics rendering
- Biotechnology searches
- Business analytics
- Clash & crash analysis
Types of Solution

- Globally distributed
  - climateprediction.net, etc
- LAN Distributed
  - GTI, etc
- Controlled Cluster
  - Beowulf
- Lots of success with all