Grid Mashups
Gluing grids together with Condor and BOINC

Mark Silberstein, Artyom Sharov, Assaf Schuster, Dan Geiger
Technion – Israel Institute of Technology
Problem ....
Problem ....
Problem ....
Parallelization

From tens to millions of subtasks
Parallelization

From tens to millions of subtasks

Where to find the computers?????
Let's build a cluster!!!
Let's build a cluster!!!
Let's use several clusters
Let's use several clusters
Let's also ask for help from
Let's also ask for help from @
Let's also ask for help from @HOME community!
Can we glue these grids together?
Can we glue these grids together?

Yes
We also want to hide the complexity..
By building virtual supercomputer for domain researchers
Virtual supercomputer for geneticists
Superlink-online: http://bioinfo.cs.technion.ac.il/superlink-online

Superlink-online is powered by Condor, a recently developed genetic linkage program. Tasks are automatically divided into small pieces and executed simultaneously using many computers. The running environment is provided by Condor. High Throughput Computing systems which allow the usage of idle cycles of hundreds of desktops. Jobs are submitted and computed all at the same time by Superlink on many computers. Longer jobs may be run in parallel and be divided among the computers in front of other users. Results are sent by e-mail upon job completion. Details of analysis are kept confidentially and responsibly by our servers.
Main components

- BOINC Server
  - HTTP frontend
  - Scheduler
  - Database
    - jobs, monitoring, system statistics
Main components

- Database: jobs, monitoring, system statistics
- Scheduler
- HTTP frontend

BOINC Server
Main components

- **BOINC Server**
  - HTTP frontend
  - Scheduler
  - Database (jobs, monitoring, system statistics)

- **BOINC clients**
  - Submitter for Condor in Madison
  - Submitter for EGEE
  - Submitter to Technion Condor

- **Database**
  - jobs, monitoring, system statistics

- **Submitters**
  - EGEE
  - Condor

- **World Map**
  - Distributed computing nodes worldwide

- **Technical University of Eindhoven**
  - Computer Science Department

- **EGEE Logo**
  - Enabling Grids for E-sciences
Main components

- **BOINC Server**
  - **HTTP frontend**
  - **Scheduler**
  - **Database**
    - jobs, monitoring, system statistics

- **Web Portal**
  - Task state

- **BOINC clients**
  - Submitter for Condor in Madison
  - Submitter for EGEE
  - Submitter to Technion Condor
Main components

- **BOINC Server**
  - **HTTP frontend**
  - **Database** jobs, monitoring, system statistics
  - **Scheduler**

- **Task execution and monitoring workflow**

- **Web Portal**
  - Task state

- **BOINC clients**
  - Submitter for Condor in Madison
  - Submitter to Technion Condor

- **BOINC clients**
  - Submitter for EGEE
Main components

**BOINC Server**
- **HTTP frontend**
- **Scheduler**
- **Database**
  - jobs, monitoring, system statistics

**Virtual cluster maintainer**
- BOINC clients submitter for Condor in Madison
- BOINC clients submitter for EGEE
- Submitter to Technion Condor

**Task execution and monitoring workflow**

**Web Portal**
- Task state

**Condor Week 2008**

Silberstein
Main components

**BOINC Server**
- **HTTP frontend**
- **Scheduler**
  - Database
    - jobs, monitoring, system statistics

**Task execution and monitoring workflow**

**Virtual cluster maintainer**
- BOINC clients submittor for Condor in Madison
- BOINC clients submittor for EGEE
- Submittor to Technion Condor

**Dedicated cluster fallback**

**Web Portal**
- Task state

Condor Week 2008

Mark Silberstein
BOINC Server

- Database
  - jobs, monitoring, system statistics

Scheduler

- HTTP frontend

Web Portal

- Task state

Task execution and monitoring

Virtual cluster maintainer

Dedicated cluster fallback

Submittor to Technion Condor

BOINC clients submittor for Condor in Madison

BOINC clients submittor for EGEE

Submitter to Technion Condor

Condor INSIDE

Condor INSIDE

Condor DAGman reliability

Condor priority management

Condor Week 2008

Mark Silberstein
Condor STARTD as a stand-alone resource manager (Condor over BOINC)

- BOINC Server
  - HTTP frontend
  - Database
    - jobs, monitoring
    - system statistics
  - Scheduler
- Virtual cluster maintainer
- Dedicated cluster fallback
- BOINC clients submitter for Condor in Madison
- BOINC clients submitter for EGEE
- Submitter to Technion Condor
- Condor INSIDE

- Web Portal
  - Task state
  - Task execution and monitoring
- Condor INSIDE
- Condor DAGman reliability
- Condor priority management
Welcome to Superlink@Technion!

Superlink@Technion helps geneticists all over the world find disease-provoking genes causing some types of diabetes, hypertension (high blood pressure), cancer, schizophrenia and many others. Press here to learn more.

To contribute the idle cycles of your computer you need to download and install a small client, called BOINC. Please note that this project is still in beta, meaning that this site can be shut down for maintenance of work units may fail. We would appreciate your feedback via our forums.

Join now!

User of the day

Superlink@clusters

This project is a part of Superlink-online genetic linkage analysis research, conducted jointly by Distributed Systems Lab and Computational Biology Lab. Superlink@clusters serves as a submission platform to clusters at EGEE, UW Madison and Technion, as well as for the use of idle cycles of the computers at the CS department at the Technion.

If you are using one of the BOINC-enabled computers at the CS department and encounter any problem, please contact Mark Silberstein via e-mail. Any feedback will be much appreciated.

News

Jan 16, 2009 Merging hosts by name
Similar hosts are being merged on a daily basis.

Oct 22, 2007 Start of the project
Superlink@clusters project started

News is available as an RSS feed here.
Preliminary results

110 CPU Years consumed and a few millions jobs completed in 4 months ~ 350 CPUs working for us around the clock

- 49 (clusters: EGEE, UW Madison, Technion CS > 20,000 CPUs)
- 61 (Contributors of Superlink@Technion > 8000 CPUs)
## Top participants

### Superlink@Technion

<table>
<thead>
<tr>
<th>Rank</th>
<th>Name</th>
<th>Recent average credit</th>
<th>Total credit</th>
<th>CPU time used (hh:mm:ss)</th>
<th>Country</th>
<th>Participant since</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>zombie87</td>
<td>3,454.76</td>
<td>280,636.87</td>
<td>1907:54:20</td>
<td>United States</td>
<td>25 Jun 2007 4:07:20 UTC</td>
</tr>
<tr>
<td>2</td>
<td>The Swordfish</td>
<td>2,781.76</td>
<td>223,736.97</td>
<td>13465:21:36</td>
<td>United States</td>
<td>12 Nov 2007 5:20:42 UTC</td>
</tr>
<tr>
<td>4</td>
<td>Steve of Boinc UK</td>
<td>2,339.91</td>
<td>155,414.76</td>
<td>6239:46:02</td>
<td>United Kingdom</td>
<td>4 Nov 2007 5:16:28 UTC</td>
</tr>
<tr>
<td>6</td>
<td>smoked_trout</td>
<td>2,098.63</td>
<td>38,418.90</td>
<td>1853:35:58</td>
<td>United States</td>
<td>17 Apr 2008 16:00:41 UTC</td>
</tr>
<tr>
<td>7</td>
<td>thelmores</td>
<td>1,699.26</td>
<td>35,921.30</td>
<td>979:53:26</td>
<td>United States</td>
<td>17 Apr 2008 1:44:15 UTC</td>
</tr>
<tr>
<td>8</td>
<td>biancaw</td>
<td>1,559.84</td>
<td>47,191.36</td>
<td>2516:25:01</td>
<td>Germany</td>
<td>31 Mar 2008 19:10:43 UTC</td>
</tr>
<tr>
<td>9</td>
<td>Kloedhil</td>
<td>1,464.10</td>
<td>34,373.61</td>
<td>1443:11:38</td>
<td>Czech Republic</td>
<td>1 Apr 2008 18:14:49 UTC</td>
</tr>
<tr>
<td>10</td>
<td>tomba</td>
<td>1,380.41</td>
<td>89,284.65</td>
<td>5567:17:28</td>
<td>United States</td>
<td>13 Jan 2008 17:10:30 UTC</td>
</tr>
<tr>
<td>11</td>
<td>UDT-Timby</td>
<td>1,032.09</td>
<td>14,264.51</td>
<td>761:34:54</td>
<td>United Kingdom</td>
<td>30 Jun 2007 19:08:56 UTC</td>
</tr>
<tr>
<td>12</td>
<td>ANCHULA-MARK 🍀</td>
<td>976.76</td>
<td>22,500.27</td>
<td>2031:35:01</td>
<td>United Kingdom</td>
<td>2 Mar 2006 14:37:28 UTC</td>
</tr>
<tr>
<td>13</td>
<td>false 🍀</td>
<td>861.52</td>
<td>34,270.90</td>
<td>1939:59:33</td>
<td>United Kingdom</td>
<td>15 Mar 2006 22:35:16 UTC</td>
</tr>
<tr>
<td>14</td>
<td>BobCat13</td>
<td>810.72</td>
<td>51,659.19</td>
<td>3051:42:48</td>
<td>United States</td>
<td>20 Jun 2007 0:36:20 UTC</td>
</tr>
</tbody>
</table>

## Top participants

### Superlink@Clusters

<table>
<thead>
<tr>
<th>Rank</th>
<th>Name</th>
<th>Recent average credit</th>
<th>Total credit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>egaee</td>
<td>19,669.21</td>
<td>2,607,640.88</td>
</tr>
<tr>
<td>2</td>
<td>technioncondor</td>
<td>3,560.51</td>
<td>412,529.58</td>
</tr>
<tr>
<td>3</td>
<td>farm</td>
<td>2,277.24</td>
<td>272,766.05</td>
</tr>
<tr>
<td>4</td>
<td>madisoncondor</td>
<td>2,259.92</td>
<td>274,493.93</td>
</tr>
<tr>
<td>5</td>
<td>lccn</td>
<td>805.61</td>
<td>201,605.33</td>
</tr>
<tr>
<td>6</td>
<td>dsi</td>
<td>803.94</td>
<td>48,187.52</td>
</tr>
<tr>
<td>7</td>
<td>assistant</td>
<td>500.92</td>
<td>51,887.40</td>
</tr>
</tbody>
</table>
### Superlink@Technion

<table>
<thead>
<tr>
<th>Rank</th>
<th>Name</th>
<th>~CPU time</th>
<th>Total credit</th>
<th>CPU time used (hh:mm:ss)</th>
<th>Country</th>
<th>Participant since</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>zombie67</td>
<td>3,454.76</td>
<td>280,696.87</td>
<td>19079:54:20</td>
<td>United States</td>
<td>25 Jun 2007 4:07:20 UTC</td>
</tr>
<tr>
<td>2</td>
<td>The Swordfish</td>
<td>2,781.76</td>
<td>223,736.97</td>
<td>13465:21:36</td>
<td>United States</td>
<td>12 Nov 2007 5:20:42 UTC</td>
</tr>
<tr>
<td>5</td>
<td>VO/HvH</td>
<td>2,279.92</td>
<td>94,535.72</td>
<td>4621:39:10</td>
<td>United States</td>
<td>17 Apr 2008 16:00:41 UTC</td>
</tr>
<tr>
<td>6</td>
<td>smoked_trout</td>
<td>2,098.63</td>
<td>38,418.90</td>
<td>1583:35:58</td>
<td>United States</td>
<td>17 Apr 2008 1:44:15 UTC</td>
</tr>
<tr>
<td>7</td>
<td>thlmores</td>
<td>1,669.26</td>
<td>35,921.30</td>
<td>979:53:26</td>
<td>United States</td>
<td>31 Mar 2008 19:10:43 UTC</td>
</tr>
<tr>
<td>8</td>
<td>blancau</td>
<td>1,559.84</td>
<td>47,191.36</td>
<td>2516:25:01</td>
<td>Germany</td>
<td>1 Apr 2008 18:14:49 UTC</td>
</tr>
<tr>
<td>9</td>
<td>Kioehilf</td>
<td>1,464.10</td>
<td>34,378.61</td>
<td>1443:11:38</td>
<td>United States</td>
<td>13 Jan 2007 17:10:30 UTC</td>
</tr>
<tr>
<td>10</td>
<td>tomba</td>
<td>1,380.41</td>
<td>89,284.65</td>
<td>5567:17:28</td>
<td>United States</td>
<td>30 Jun 2007 19:08:56 UTC</td>
</tr>
<tr>
<td>11</td>
<td>UCT-Timby</td>
<td>1,032.09</td>
<td>14,264.51</td>
<td>761:34:54</td>
<td>United Kingdom</td>
<td>2 Mar 2006 14:37:28 UTC</td>
</tr>
<tr>
<td>12</td>
<td>ANCHULA-MARK 🌐</td>
<td>976.76</td>
<td>22,500.27</td>
<td>2031:35:01</td>
<td>United Kingdom</td>
<td>15 Mar 2006 22:35:15 UTC</td>
</tr>
<tr>
<td>13</td>
<td>falsie 🌐</td>
<td>861.52</td>
<td>34,276.93</td>
<td>1939:59:33</td>
<td>United States</td>
<td>20 Jun 2007 0:36:20 UTC</td>
</tr>
</tbody>
</table>

### Superlink@Clusters

<table>
<thead>
<tr>
<th>Rank</th>
<th>Name</th>
<th>~CPU time</th>
<th>Total credit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>egae</td>
<td>19,669.21</td>
<td>2,607,640.88</td>
</tr>
<tr>
<td>2</td>
<td>technioncondor</td>
<td>3,560.51</td>
<td>412,529.58</td>
</tr>
<tr>
<td>3</td>
<td>farm</td>
<td>2,277.24</td>
<td>272,766.05</td>
</tr>
<tr>
<td>4</td>
<td>madisoncondor</td>
<td>2,259.92</td>
<td>274,493.93</td>
</tr>
<tr>
<td>5</td>
<td>lccn</td>
<td>895.61</td>
<td>201,605.33</td>
</tr>
<tr>
<td>6</td>
<td>dsi</td>
<td>803.94</td>
<td>48,187.52</td>
</tr>
<tr>
<td>7</td>
<td>assistant</td>
<td>500.92</td>
<td>51,887.40</td>
</tr>
</tbody>
</table>
Single framework for grid performance analysis

- Superlink@Technion environment is more predictable!

Since clusters are much more volatile because of policy

- Jobs left in queue

- Eviction rate

Condor Week 2008
Mark Silberstein
Single framework for grid performance analysis

- EGEE vs. Madison pool
  - EGEE behaves more as @Technion than Madison pool. Why?

High load, complex policies and centralized control result in rapid (statistically unpredictable) changes in resource allocation.

EGEE is much less centralized – evictions are less correlated, hence have steady state.
Lesson 1: Track the system performance from the application perspective

- Simple criterion - compare with the time it would take on a dedicated cluster

Using BOINC (glide-ins) helps separate application performance from batch system performance
Lesson 2: Learn the system parameters

- Example: Average running time of a BOINC client until preempted

Wrong job granularity

The same run with proper job granularity

- Using BOINC allows to tailor job size with no additional overhead: few minutes long jobs run as fast as few hours long
Lesson 3: Distinguish between high throughput and high performance runs

- Many-jobs: high throughput runs – number of jobs much more than number of running machines
  - Less sensitive to failures - overlapped with the execution of other jobs
- Few-jobs: high performance runs – number of jobs is about the number of running machines
  - Performance very sensitive to failures

Dedicated cluster comes handy!!!
Thinking out loud

- Will grids become less useful to opportunistic users when they become easy to use (higher utilized)?
- Should Condor policy be hostile to opportunistic users, or some guarantees are better to be provided?
- And if above is true, and you are going to have a lot of jobs – maybe it's better to buy your own cluster...
Future work.... Superlink@Mars.. Venus.. Moon

But before that ... If your grid can handle backfill – let us in!
If you want to contribute your PC(s) to the search for disease-provoking genetic mutations –

Join Superlink@Technion!