Interfacing Big Data: Death by a 25 Million Pinpricks

Stephen Mock
Director of Advanced Computing Interfaces
Texas Advanced Computing Center
Advanced Computing Interfaces (ACI)

Who are we?

Web & Mobile Applications
Matthew Hanlon*
Maytal Dahan
Carrie Arnold
David Choi
David Montoya
Walter Scarborough

Web & Cloud Services
Rion Dooley*
Warren Smith
Akhil Seth
Charlie Dey
Steve Terry
Clint Blythe
James McCurdy

*Manager
Advanced Computing Interfaces (ACI)

What do we do?

**Web & Mobile Applications**
- Portals & Gateways
- Web Applications
- User Interfaces
- “Front end”

**Web & Cloud Services**
- APIs
- Web Services
- Data Management
- “Back end”
Interface Expectations

• Set by the state of the web
HPC Jobs Data

• XSEDE’s Central DB has all job data since around 2004
  – 25M jobs
  – Going back to the TeraGrid days
• TACC has ~14M job records
• Big data
  – Many small data points, many pinpricks
Job Data Creation

- Rate increasing with larger, more capable systems
  - July 2011: ~19M job records
  - As of last week: 25,480,013 job records

- Stampede!
  - Increased capacity generating more jobs at faster rate
XSEDE Central DB Jobs

• Postgres DB

• Querying XSEDE central DB directly:
  – lacks the responsiveness required by an interactive web portal
  – individual queries 5+ seconds
    • dozens of queries to provide analytics across projects, resources, users, date ranges

• A different approach was needed
NoSQL: CouchDB

- Export all XSEDE jobs data
- Import into CouchDB instance
- Periodically import new jobs records to CouchDB
What is CouchDB?

- Document-oriented database
  - semi-structured documents
  - flat file
  - IDs and revisions, indexed in b-tree

- Queried using MapReduce
  - key-value “views” of the data

- RESTful JSON API
  - documents, queries, responses, configuration settings

- Written in Erlang
  - concurrency, distributed applications

- ACID compliant database
Comparing Postgres & CouchDB: jobs queries

<table>
<thead>
<tr>
<th>System</th>
<th>Queries</th>
<th>Total (m)</th>
<th>Average (s)</th>
<th>Min (s)</th>
<th>Max (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>postgres tgcdb.teragrid.org</td>
<td>4566</td>
<td>426.0573656</td>
<td>5.598651322</td>
<td>0.512231112</td>
<td>740.5359879</td>
</tr>
<tr>
<td>postgres autry.corral.tacc.utexas.edu</td>
<td>4566</td>
<td>114.9091518</td>
<td>1.509975714</td>
<td>0.554908037</td>
<td>75.0084548</td>
</tr>
<tr>
<td><strong>Speedup (tgcdb/new hardware)</strong></td>
<td></td>
<td><strong>3.707775741</strong></td>
<td><strong>3.707775741</strong></td>
<td><strong>0.923091895</strong></td>
<td><strong>9.872700215</strong></td>
</tr>
<tr>
<td>couchdb autry.corral.tacc.utexas.edu</td>
<td>4566</td>
<td>13.93821839</td>
<td>0.183156615</td>
<td>0.12539196</td>
<td>0.765941143</td>
</tr>
<tr>
<td><strong>Speedup (couchdb/tgcdb)</strong></td>
<td>30.56756277</td>
<td>30.56756277</td>
<td>4.085039511</td>
<td>966.8314525</td>
<td></td>
</tr>
<tr>
<td><strong>Speedup (couchdb/new hardware)</strong></td>
<td>8.244177887</td>
<td>8.244177887</td>
<td>4.425387693</td>
<td>97.92978936</td>
<td></td>
</tr>
</tbody>
</table>
XSEDE User Portal: MyJobs

My Jobs lets you view current and historical data for jobs run on XSEDE resources. Click on a resource below to view job data for that resource.

Historical Jobs

You can search for jobs you have run on longhorn.tacc.teragrid.org using the form below. **Note:** if you have a significant job history on this resource limit the search period for best performance.

**START DATE** 03/01/2012  **END DATE** 11/14/2012

<table>
<thead>
<tr>
<th>PROJECT NUMBER</th>
<th>JOB ID</th>
<th>NAME</th>
<th>START TIME (GMT)</th>
<th>END TIME (GMT)</th>
<th>QUEUE</th>
<th>PROCESSORS</th>
<th>WALL TIME (MINUTES)</th>
<th>CHARGE (SU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TG-STA060001N</td>
<td>169601</td>
<td></td>
<td>2012-08-23 13:52:51</td>
<td>2012-08-23 13:58:01</td>
<td>normal</td>
<td>8</td>
<td>5.1667</td>
<td>1.33128</td>
</tr>
<tr>
<td>TG-STA060001N</td>
<td>157569</td>
<td></td>
<td>2012-06-12 14:36:36</td>
<td>2012-06-12 14:45:44</td>
<td>normal</td>
<td>8</td>
<td>9.1333</td>
<td>2.354895</td>
</tr>
</tbody>
</table>

Showing 1 to 3 of 3 entries

- There can be some delay between actual job submission and when the job will appear in My Jobs.
- Some sites do not report job owner information. Although you may have jobs scheduled at these sites, those jobs will not appear in My Jobs.
Jobs Data: Messaging-Based

- Immediate job state updates
  - New job, running, finished, errors…
- Without polling the queue commands
- Push rather than pull
- Gatherers publish and don’t care who consumes
- Consumers subscribe for what they care about, and don’t care who publishes
Messaging Brokers

• AMQP (Advanced Message Queuing Protocol)
  – RabbitMQ
  – Apache QPID
Messaging Architecture

- Messaging Service
  - Information Gatherer
  - Database
  - Consumers

Common Representation Language

(messages) query/result (messages)
Job Messages

- Producers watch the queue log files for changes
- Send message upon job state change
- Periodically send entire queue listing
- JSON formatted
Result

• What’s that get us?
  – Immediate & up-to-date job state info
  – Very small impact on head nodes
    • Happy sysadmins!
  – Simple producers and consumers

…but mostly…
Mobile Notifications

- Enables mobile app job notifications
Stephen Mock
mock@tacc.utexas.edu
512-475-9411

For more information:
www.tacc.utexas.edu