BOINC and Cloud Computing

Derrick Kondo
INRIA, France
Cloud Background

• Vision
  • Hide complexity of hardware and software management from a user.

• Benefits
  • Pay as you go
  • Scale up or down dynamically
  • Root-access configurability (no lock-in)
Platform Performance vs. Costs

What is the relationship?

BOINC + Cloud?
Outline

• Performance tradeoffs
• Monetary tradeoffs
• Client hosting
• Server hosting
Method

- Use real costs
- Large BOINC project (SETI@home)
- Small BOINC project (XtremLab)
- Amazon Elastic Computing Cloud (EC2)
- Use real measurements
  - Server logs
  - BOINC statistical data
Stages of Project & Application

Platform Construction ➔ Application Deployment ➔ Application Execution ➔ Application Completion
How long before I get X TeraFLOPS?

Strategy:
- Add to BOINC project list
- Respond to users (leverage volunteers)
- Google Ad Sense
- Forum Announcements

Can get over 20 TeraFLOPS within 6 months
How long to deploy my batch of tasks?

Strategy: Specify lower request times [Heien et al.]
How many volunteer nodes are equivalent to 1 cloud node?

Strategy: Use statistical prediction of availability (see my ISGW talk)

2.8 active volunteer hosts per 1 cloud node
How long should I wait for task completion?

Strategy: Use statistical prediction of availability (see my ISGW talk)

Stretch > 5.
Median project deadline: 9 days.
Execution time on 3GHz host: 3.67 hours.
Good success rates: 96.1% of WCG met.
Monetary Tradeoffs

- Client hosting on cloud
  - Not worth it and never will
- Server hosting on the cloud
  - Possible solution
## Monthly Project Costs

<table>
<thead>
<tr>
<th>Component</th>
<th>Project</th>
<th>SETI@home</th>
<th>XtremLab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salaries</td>
<td>10K for sys admins</td>
<td></td>
<td>5K $\rightarrow$ 1K $\rightarrow$ Students</td>
</tr>
<tr>
<td>Electricity</td>
<td>90 for 6 servers</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>Network</td>
<td>2K for 100 Mbit</td>
<td></td>
<td>covered by university</td>
</tr>
<tr>
<td>Hardware</td>
<td>18K for servers, 25K for air conditioner</td>
<td></td>
<td>4K $\rightarrow$ 1K</td>
</tr>
<tr>
<td>Total startup</td>
<td>43K</td>
<td></td>
<td>4K $\rightarrow$ 1K</td>
</tr>
<tr>
<td>Total monthly</td>
<td>12K</td>
<td></td>
<td>5k $\rightarrow$ 1K</td>
</tr>
</tbody>
</table>
EC2 Pricing

**Table 1. Pricing for EC2 Instances**

<table>
<thead>
<tr>
<th>Instance Type</th>
<th>Cost/hour (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Small</td>
<td>0.10</td>
</tr>
<tr>
<td>Standard Large</td>
<td>0.40</td>
</tr>
<tr>
<td>High-CPU</td>
<td>0.20</td>
</tr>
</tbody>
</table>

**Table 2. Pricing for EC2 Data Transfer**

<table>
<thead>
<tr>
<th>Transfer Type</th>
<th>Cost/GB-Month (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inbound transfer</td>
<td>0.10</td>
</tr>
<tr>
<td>first 10 TB</td>
<td>0.17</td>
</tr>
<tr>
<td>next 10-50TB</td>
<td>0.13</td>
</tr>
<tr>
<td>next 50-150TB</td>
<td>0.11</td>
</tr>
<tr>
<td>over 150 TB</td>
<td>0.10</td>
</tr>
</tbody>
</table>

**Table 3. Pricing for EBS**

<table>
<thead>
<tr>
<th>Resource</th>
<th>Rate (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storage</td>
<td>0.10 / GB-Month</td>
</tr>
<tr>
<td>IO request</td>
<td>0.10 / million</td>
</tr>
</tbody>
</table>

...
What about Client Hosting on the Cloud?
When is BOINC more cost effective than Clouds?

After ~13 days
Cost of Clouds versus BOINC

What are total costs over time?

Within <3 days, BOINC is cheaper
Cost of Clouds versus BOINC

How many months of BOINC can X months of Cloud buy me?

12 months of Cloud can pay for > 125 BOINC years. BOINC way more sustainable.
Equivalent Clouds given BOINC Costs

Given BOINC costs, what size Cloud can I buy?

Orders of magnitude lower than BOINC
What about Server Hosting on the Cloud?
Number of Hosts over Time

Load variation exits.
Clouds take care of server management.
Potential to exploit clouds.
Server Costs on a Cloud

How much to host BOINC server on cloud?

Cheaper to host on Cloud. Bandwidth is expensive.
Cloud Resources with Given Budget

How big of a server can I support with given budget?

Most project servers are sustainable on the cloud.
Summary

- Performance tradeoffs
  - 20 BOINC TeraFLOPS within 6 months

- Monetary tradeoffs
  - Client hosting
    - After 13 days, BOINC more cost effective
  - Server hosting
    - BOINC server on cloud is cost-effective
    - Savings of at least 40%