Distributed Systems Meet Economics: Pricing in the Cloud

Presenter: Rishan Chen
Peking University and Microsoft Research, Asia
June 2010, Boston, MA
Cloud is a distributed system

• System metrics
  – Throughput
  – Latency / response time
  – Failure rate
  – Power consumption, etc.

• As a pay-as-you-go service
  – Two parties connected by the pricing scheme
  – It’s all about the money!
Pricing in the Cloud

- It significantly changes the landscape of system design: **Cost** as an explicit and measurable system metric
  - How both parties optimize their logic
  - Is the pricing fair
  - How does the pricing interplay with the evolving system dynamics
  - How to measure the cost of failures, etc.
Methodology overview

• Approximate a typical workload in current cloud computing
  • Postmark (I/O-intensive)
  • PARSEC: Dedup, BlackScholes (CPU-intensive)
  • Hadoop (large-scale data processing)

• Complementary approaches for evaluations
  • A black-box approach with Amazon EC2
  • Built a cloud-computing test bed, Spring, to perform fully controlled experiments
Preliminary results

• Pricing may give different indices for users and providers for system optimizations (e.g., consolidation)
• System performance variations may lead to pricing fairness issues
• System evolution (e.g., adoption of new hardware like SSD) may affect pricing scheme
• Failures need to be better dealt with regarding to the cost
Highlights of our study

• Pricing (profit) versus throughput

<table>
<thead>
<tr>
<th>Number of concurrent VMs</th>
<th>One VM</th>
<th>Two VMs</th>
<th>Four VMs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average cost per task ($)</td>
<td>0.004</td>
<td>0.004</td>
<td>0.012</td>
</tr>
<tr>
<td>Profit ($)</td>
<td>-0.009</td>
<td>0.002</td>
<td>0.028</td>
</tr>
<tr>
<td>Throughput (tasks/h)</td>
<td>28.3</td>
<td>56.4</td>
<td>33.9</td>
</tr>
</tbody>
</table>

Run Postmark continuously and report the number for four tasks; we compare the consolidation of x VMs on a single physical machine.
Highlights of our study /2

- Optimizing for cost versus optimizing traditional system metrics
Highlights of our study /3

• Pricing fairness: performance variation

<table>
<thead>
<tr>
<th></th>
<th>Postmark</th>
<th>Dedup</th>
<th>BlackScholes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$cv$</td>
<td>9.1%</td>
<td>11.0%</td>
<td>3.9%</td>
</tr>
<tr>
<td>$maxDiff$</td>
<td>40.1%</td>
<td>38.8%</td>
<td>12.6%</td>
</tr>
</tbody>
</table>

Table 8: Variation of different runs on EC2

Figure 2: Variations among three instances (Postmark)
Open questions

• What are good properties for a pricing scheme?
• How do users and providers adapt the system design to evolving and even hybrid pricing schemes?
• How is the pricing scheme adapted to the evolving system dynamics and (new) technologies?
• How to deal with failures’ cost regarding to the pricing?
Related work

• Other pricing schemes
  – Bilateral
  – **Amazon EC2 Spot Instances**: Enable you to bid for unused Amazon EC2 capacity
    • Navraj Chohan, et al., *See Spot Run: Using Spot Instances for MapReduce Workflows*, June 2010
  – **Microsoft SQL Azure**: Make pricing more scalable and more predictable

• **Distributed computing w/ Economics**
  – Ang Li, et al., *CloudCmp: Shopping for a Cloud Made Easy*, June 2010
Summary

• Pricing is an important bridge between users and providers
• It significantly changes the dynamics in system design
• The interplay between economics and system design can be a fruitful research direction