Intro to Software as a Service (SaaS) and Cloud Computing

Armando Fox, UC Berkeley
Reliable Adaptive Distributed Systems Lab
© 2009-2012

Image: John Curley http://www.flickr.com/photos/jay_que/1834540/
Software as a Service: SaaS

• Traditional SW: binary code installed and runs wholly on client device
• SaaS delivers SW & data as service over Internet via thin program (e.g., browser) running on client device
  — Search, social networking, video
• Now also SaaS version of traditional SW
  — E.g., Microsoft Office 365, TurboTax Online
SaaS in 1996

One Person Can Immediately Affect Millions

Welcome to eBay's AuctionWeb.

Welcome to today's online marketplace...
...the market that brings buyers and sellers together in an honest and open environment...

Welcome to eBay's AuctionWeb.

From the founder:

February 26, 1996
Why SaaS?

1. No installation hassles
2. No worries about data loss
3. Easy for groups to interact with same data
4. Data is large or changes frequently
5. No compatibility hassles for developers
6. Easier upgrades
SaaS Needs Infrastructure

1. Communication: allow customers to interact with service
2. Scalability: fluctuations in demand during + new services to add users rapidly
3. Dependability: service and communication continuously available 24x7
Your PC vs. Datacenter Computer Smackdown

Sun E-10000 “supermini” c.1996

<table>
<thead>
<tr>
<th>Machine</th>
<th>Processor cores</th>
<th>RAM</th>
<th>Disk</th>
</tr>
</thead>
<tbody>
<tr>
<td>E10000, 1996</td>
<td>64 x 250MHz</td>
<td>64 GB</td>
<td>20 TB</td>
</tr>
<tr>
<td>PC, 1996</td>
<td>1 x 250 MHz</td>
<td>32 MB</td>
<td>4 GB</td>
</tr>
<tr>
<td><strong>Ratio</strong></td>
<td><strong>64:1</strong></td>
<td><strong>2000:1</strong></td>
<td><strong>5000:1</strong></td>
</tr>
<tr>
<td>Datacenter computer, 2010</td>
<td>8 x 1 GHz</td>
<td>16 GB</td>
<td>2 TB</td>
</tr>
<tr>
<td>PC, 2010</td>
<td>2 x 3 GHz</td>
<td>4 GB</td>
<td>0.5 TB</td>
</tr>
<tr>
<td><strong>Ratio</strong></td>
<td><strong>&lt; 2:1</strong></td>
<td><strong>4:1</strong></td>
<td><strong>4:1</strong></td>
</tr>
</tbody>
</table>

Modern datacenters use commodity computers.
“Workstation price-performance is improving at 80% per year, while that of supercomputers is improving at only 20-30% per year.”

Why?

“Instead of small computers for interactive use and larger computers for demanding applications, we propose using NOWs for all the needs of computer users.”

Whoa.
NOW-0

1994

Four HP-735’s
NOW-1

1995

32 Sun SPARC-stations
NOW-2

1997

60 Sun SPARC-2
• The first mobile Web browser with graphics was developed by:

(a) Apple Stanford
(b) Google Berkeley
(c) Motorola MIT
(d) None of the above
“Access Is the Killer App”
Project Daedalus, 1994-1999

• Faculty: Profs. Katz & Brewer
• Idea: Use the “cloud” for services!
  – First truly scalable search engine (Inktomi)
  – First mobile Web browser enabled by content transformation (TopGun)
  – Vision: Anywhere, anytime access to data & services, supported by the “cloud”
Datacenter is new “server”

- “Program” => Web search, email, map/GIS, …
- “Computer” => 1000’s computers, storage, network
- Warehouse-sized facilities and workloads

photos: Sun Microsystems, CNET, & datacenterknowledge.com
Public Cloud Computing Arrives
(Amazon Elastic Compute Cloud, 2007)

• What: Pay-as-you-go access to racked commodity servers
  – from 0.02/server-hour, no minimum
  – 100 servers x 1h costs same as 1 server x 100h

• Eliminates financial barrier to deploy SaaS
  – FarmVille: 4 days =1M players; 2 months = 10M; 9 months = 75M!
  – A cloud-based system is world’s 42nd fastest supercomputer, at $700/hr
  – IBM Watson would cost about $290/hr
• The first working prototypes of the Internet and of computing as a utility were demonstrated in:

(a) 1969
(b) 1978
(c) 1983
(d) 1990
Why Now?

• Communication

• Scalability

• Dependability
• Provisioning for peaks: wasteful, but necessary

"Statically provisioned" data center

"Virtual" data center in the cloud

Unused resources
Risk Transfer
(or: who remembers Friendster?)
Cost Associativity

• 1,000 CPUs for 1 hour same price as
  1 CPU for 1,000 hours

• Washington Post converted Hillary Clinton’s travel documents to post on WWW
  – Conversion time: <1 day after released
  – Cost: less than $200

• RAD Lab graduate students demonstrate improved MapReduce scheduling—on 1,000 servers
What cool things can we do with the cloud in academia?
CS 169: cloud supports SW development too!

- Develop your app
- Keep track of your code
- Test your app on different browsers
- Deploy it to the world

Total UCB computer resources: zero
2012: Cloud Computing and a MOOC*

- saas-class.org: first 5 weeks of CS 169
- >75,000 students learning SaaS programming!
- Cloud computing for autograding
  - What happens when 25,000 students submit a programming HW on the same day…?
- July 24, 2012: UC Berkeley joins edX

* massive open online class
Future of Software:

- Cloud *democratizes access* to large-scale computing, accelerates “SaaS-ification”
- Students, researchers, entrepreneurs can now have even greater impact
- New research & education opportunities
Going back to NOW...

- **2000**: using medium-sized clusters for Internet services => several PhD’s
- **2010**: CS169 students do it in 6-8 weeks and deploy on cloud computing, which is also used to grade their work.
- **2020**: ?
Thank you!

RAD Lab, 2011