



## 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/](http://www.flickr.com/photos/jay_que/1834540/)



## Software as a Service: SaaS

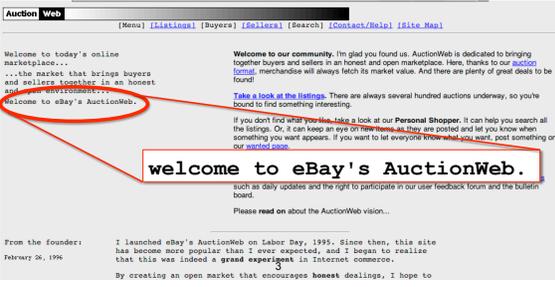
- Traditional SW: binary code installed and runs wholly on client device
- SaaS delivers SW & data as service over Internet via thin program (e.g., browser) running on client device
  - Search, social networking, video
- Now also SaaS version of traditional SW
  - E.g., Microsoft Office 365, TurboTax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

2



## SaaS in 1996

### One Person Can Immediately Affect Millions



welcome to eBay's AuctionWeb.



## Why SaaS?

1. No install worries about HW capability, OS
2. No worries about data loss (at remote site)
3. Easy for groups to interact with same data
4. If data is large or changed frequently, simpler to keep 1 copy at central site
5. 1 copy of SW, controlled HW environment => no compatibility hassles for developers
6. 1 copy => simplifies upgrades for developers *and* no user upgrade requests

4



## 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

5



## Your PC vs. Datacenter Computer Smackdown

Sun E-10000 "supermini" c.1996

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



**Modern datacenters use commodity computers.**

6

**“The Case for NOW (Networks of Workstations)”**

- “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.*

7

**“The Case for NOW (Networks of Workstations)”**

- The first Web search engine whose design made possible the huge scale of today’s search engines was invented at:

- (a) Stanford
- (b) Berkeley
- (c) Yahoo!
- (d) Google
- (e) IBM

8

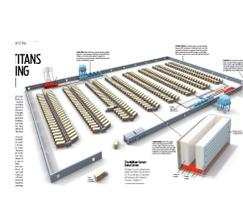
**Why Cloud Computing? (Amazon Elastic Compute Cloud, 2008)**

- 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 42<sup>nd</sup> fastest supercomputer, at \$700/hr
  - IBM Watson would cost about \$290/hr

9

**Datacenter is new “server”**

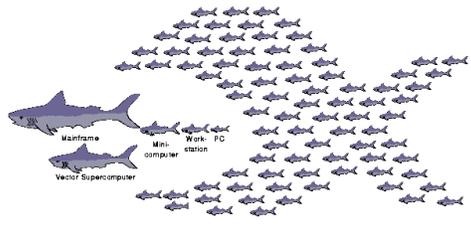
- “Program” => Web search, email, map/GIS, ...
- “Computer” => 1000’s computers, storage, network
- Warehouse-sized facilities and workloads

photos: Sun Microsystems, CNET, & datacenterknowledge.com

10

**UC Berkeley Networks Of Workstations (1994-1999)**



11

**NOW-0**

1994

Four HP-735's



 RAD Lab

NOW-1

1995

32 Sun SPARC-stations



 RAD Lab

NOW-2

1997

60 Sun SPARC-2



 RAD Lab  
UC Berkeley

Challenge: how do you program a NOW? How do you keep it running as individual machines fail?

15

 RAD Lab

Trivia Fact

- The first full Web browser running on a mobile device was developed by:

(a) Apple  
(b) Stanford  
(c) Berkeley  
(d) Nokia  
(e) Motorola

16

 RAD Lab

**“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”*



 RAD Lab

**2007: Public Cloud Computing Arrives**

- Amazon Elastic Compute Cloud (EC2)
- “Compute unit” rental: \$0.02-0.68/hr.
  - 1 CU  $\approx$  ~1 GHz x86 core
  - Virtual machine technology used to “slice up”
- No up-front cost, no contract, no minimum
- Billing rounded to nearest hour
  - pay-as-you-go storage also available
- “Computing as utility” – MULTICS, c.1969
- See [abovetheclouds.cs.berkeley.edu](http://abovetheclouds.cs.berkeley.edu)

18

**Why Now (not then)?**

- The Web “Space Race”: Build-out of extremely large datacenters (10,000’s of **commodity** PCs)
- Driven by growth in demand (more users)
  - Discovered **economy of scale**: 5-7x cheaper than provisioning a medium-sized (100’s machines) facility
  - Infrastructure software: e.g., Google File System
  - Operational expertise
- More pervasive broadband Internet
- Dominance of Intel x86 architecture in servers
- Free & open source software availability
- What’s new: **risk transfer** & **cost associativity**

**Cloud Economics 101**

- Provisioning for peaks: wasteful, but necessary

“Statically provisioned” data center      “Virtual” data center in the cloud

Unused resources

**Risk Transfer (or: who remembers Friendster?)**

Resources      Capacity      Demand      Time (days)

Resources      Capacity      Demand      Time (days)

Resources      Capacity      Demand      Time (days)

Lost revenue

Lost users

**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?**

**2005-2010: RAD Lab and Karl’s Long Weekend**

Presidents’ Day Weekend, Feb 21-13  
Final demo on Feb 24

**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**

25

**2012: Cloud Computing and a MOOC\***

- saas-class.org: first 5 weeks of CS 169
- 25,000 students learning SaaS programming!
- Cloud computing for autograding
  - What happens when 25,000 students submit a programming HW on the same day...?

\* massive open online class

26

**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  
– Everything delivered as SaaS now...
- **2020:** ?



27

**2011: Future=Mobile+Cloud**



28

**Summary**

- Cloud computing *democratizes access* to large-scale computing resources
  - Pay-as-you-go => low risk, low entry cost
- *Accelerates* “SaaS-ification”
  - Economic benefits of delivering software as a service now available to anyone
- Allows students, academia to have even greater impact on industry
- Open up research/innovation opportunities

29

**Thank you!**



RAD Lab Team

30