Cloud Computing

Prof. Ravi Sandhu
Executive Director and Endowed Chair

April 12, 2013

ravi.sandhu@utsa.edu
www.profsandhu.com
The Cloud is the Computer
- IEEE Spectrum, 2008

Datacenter as a Computer
- Barroso and Hölzle, 2009

The Network is the Computer
- Sun Microsystems, early 1990s
Cloudwashing

I NEED YOU TO CLOUDWASH OUR SOFTWARE.

MOVING SOME OF ITS FUNCTIONS ONTO THE INTERNET, BUT CALL THE INTERNET A CLOUD.

NO ONE WILL TAKE US SERIOUSLY UNLESS WE'RE DOING SOMETHING IN THE CLOUD.

WILL PEOPLE TAKE US SERIOUSLY IF WE MAKE TECHNOLOGY DECISIONS BASED ON JARGON?

WE DON'T CARE WHAT SMART PEOPLE THINK. THERE AREN'T THAT MANY OF THEM.

WE ONLY NEED TO CONVINCE OUR DUMB CUSTOMERS.

DO YOU BELIEVE I MOVED OUR SOFTWARE TO THE CLOUD YESTERDAY?

YOU DID?

I'M GOING TO SAY YES.
The Cloud: Perspectives and Forces

Science

Business

Engineering

World-Leading Research with Real-World Impact!
NIST Cloud Computing 3-4-5 Definition

5 Essential Characteristics

2009-2011
16 versions

3 Service Models

4 Deployment Models
NIST Cloud Computing 3-4-5 Definition

5 Essential Characteristics

2009-2011
16 versions

3 Service Models
Software as a Service (SaaS)
Platform as a Service (PaaS)
Infrastructure as a Service (IaaS)

4 Deployment Models
Public
Private
Community
Hybrid
NIST Cloud Computing 3-4-5 Definition

5 Essential Characteristics

- On-demand self service
- Broad network access
- Resource pooling (multi-tenant)
- Rapid elasticity
- Measured service

3 Service Models

- Software as a Service (SaaS)
- Platform as a Service (PaaS)
- Infrastructure as a Service (IaaS)

4 Deployment Models

- Public
- Private
- Community
- Hybrid
NIST Cloud Computing 3-4-5 Definition

5 Essential Characteristics

On-demand self service
Broad network access
Resource pooling (multi-tenant)
Rapid elasticity
Measured service

2009-2011
16 versions

Other Common Characteristics

Geographic distribution
Homogeneity
Resilience
Massive scale
Virtualization
Security

3 Service Models

Software as a Service (SaaS)
Platform as a Service (PaaS)
Infrastructure as a Service (IaaS)

4 Deployment Models

Public
Private
Community
Hybrid
IaaS: Amazon Web Services (AWS)
- Compute (EC2)
- Storage (S3)
- Database (multiple)
- Plus more

IaaS: other players
- Eucalyptus: open source with EC2 APIs
- OpenStack: industry backed open source consortium (Rackspace genesis)
- Joyent: SmartOS based

SaaS: SalesForce, CRM (Customer Relationship Management)
- Any web application can be pitched as SaaS
- But don’t forget “On-demand self service” and “Rapid elasticity”

PaaS: least well defined service
- Salesforce: force.com plus more
- Microsoft Azure
- Google AppEngine
“We argue that Cloud Computing not only overlaps with Grid Computing, it is indeed evolved out of Grid Computing and relies on Grid Computing as its backbone and infrastructure support.”

I don’t think so
Cloud and Grid: Foster et al 2008

Scale

Distributed Systems

Supercomputers

Grids

Clusters

Clouds

Web 2.0

Application Oriented

Services Oriented

World-Leading Research with Real-World Impact!

© Ravi Sandhu
1. Coordinates resources that are not subject to centralized control
   - Virtual Organization (VO)
2. Uses standard, open, general-purpose protocols and interfaces
   - Globus toolkit
3. Delivers non-trivial qualities of service
1. Coordinates resources that are not subject to centralized control
   - Virtual Organization (VO)
2. Uses standard, open, general-purpose protocols and interfaces
   - Globus toolkit
3. Delivers non-trivial qualities of service
   - Yes

On-demand self service  Geographic distribution
Broad network access   Homogeneity
Resource pooling (multi-tenant) Resilience
Rapid elasticity        Massive scale
Measured service        Virtualization
Security

No but VOs may be enabled on demand
No but standard open source software and APIs may emerge (OpenStack is the current contender)
## Grid versus Cloud Drivers

<table>
<thead>
<tr>
<th>Cloud</th>
<th>Grid</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Commercially developed</td>
<td>- DoD funded, no commercial traction</td>
</tr>
<tr>
<td>- Little or no academic input</td>
<td>- Mainly academic driven</td>
</tr>
<tr>
<td>- Pay-per-use</td>
<td>- Pay-per-seat (one-time payment</td>
</tr>
<tr>
<td>- Payment driven</td>
<td>- Project oriented, proposal driven</td>
</tr>
<tr>
<td>- Centrally owned hardware</td>
<td>- Multiply owned hardware</td>
</tr>
<tr>
<td>- Centrally scheduled</td>
<td>- Distributed scheduling</td>
</tr>
<tr>
<td>- Single point of trust</td>
<td>- Multiple trust points</td>
</tr>
<tr>
<td>- Simple security</td>
<td>- Complex PKI based security</td>
</tr>
<tr>
<td>- Interactive</td>
<td>- Batch</td>
</tr>
<tr>
<td>- Commodity computing</td>
<td>- High performance computing</td>
</tr>
<tr>
<td>- Small and medium businesses</td>
<td>- High end organizations</td>
</tr>
<tr>
<td>- Virtualization essential</td>
<td>- Virtualization often not used</td>
</tr>
<tr>
<td>- Not so predictable performance</td>
<td>- Predictable performance</td>
</tr>
<tr>
<td>- Cost associativity</td>
<td>- Not to same degree</td>
</tr>
</tbody>
</table>
Cloud and Grid: Foster et al 2008

Data

Cloud Computing

Client Computing

Consume
Produce

Consume
Produce

Communication
Figure 1. Users and providers of cloud computing. We focus on cloud computing’s effects on cloud providers and SaaS providers/cloud users. The top level can be recursive, in that SaaS providers can also be a SaaS users via mashups.
Not IaaS or PaaS but classes of utility computing
# Berkeley View of Cloud: 2010

## Table 1. Comparing public clouds and private data centers.

<table>
<thead>
<tr>
<th>Advantage</th>
<th>Public Cloud</th>
<th>Conventional Data Center</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appearance of infinite computing resources on demand</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Elimination of an up-front commitment by Cloud users</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Ability to pay for use of computing resources on a short-term basis as needed</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Economies of scale due to very large data centers</td>
<td>Yes</td>
<td>Usually not</td>
</tr>
<tr>
<td>Higher utilization by multiplexing of workloads from different organizations</td>
<td>Yes</td>
<td>Depends on company size</td>
</tr>
<tr>
<td>Simplify operation and increase utilization via resource virtualization</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
Table 2. Top 10 obstacles to and opportunities for growth of cloud computing.

<table>
<thead>
<tr>
<th>Obstacle</th>
<th>Opportunity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Availability/Business Continuity</td>
<td>Use Multiple Cloud Providers</td>
</tr>
<tr>
<td>2 Data Lock-In</td>
<td>Standardize APIs; Compatible SW to enable Surge or Hybrid Cloud Computing</td>
</tr>
<tr>
<td>3 Data Confidentiality and Auditability</td>
<td>Deploy Encryption, VLANs, Firewalls</td>
</tr>
<tr>
<td>4 Data Transfer Bottlenecks</td>
<td>FedExing Disks; Higher BW Switches</td>
</tr>
<tr>
<td>5 Performance Unpredictability</td>
<td>Improved VM Support; Flash Memory; Gang Schedule VMs</td>
</tr>
<tr>
<td>6 Scalable Storage</td>
<td>Invent Scalable Store</td>
</tr>
<tr>
<td>7 Bugs in Large Distributed Systems</td>
<td>Invent Debugger that relies on Distributed VMs</td>
</tr>
<tr>
<td>8 Scaling Quickly</td>
<td>Invent Auto-Scaler that relies on ML; Snapshots for Conservation</td>
</tr>
<tr>
<td>9 Reputation Fate Sharing</td>
<td>Offer reputation-guarding services like those for email</td>
</tr>
<tr>
<td>10 Software Licensing</td>
<td>Pay-for-use licenses</td>
</tr>
</tbody>
</table>