Announcements

- **Teaching Evaluations**
  - Monday, May 26 at 12:01 am and closing Sunday, June 8 at 11:59 pm.

- **Final Business Papers Due 6/5**

- **Final Exam**
  - Wednesday June 11, 8am - 11am
Cloud Computing

Cloud Computing: refers to both

- applications delivered as services over internet
  - aka Software as a Service (SaaS)
- hardware / software in data centers providing those services -- a cloud
Cloud Computing

- 2 flavors:
  - Public Cloud, available to public
    - provides utility computing
  - Private Cloud
    - internal to company
Advantages

- **SaaS**
  - Control of Versioning
  - Users access anywhere
  - Ease of data sharing
  - Pay as you go

- **Additional +’s of Cloud Computing**
  - Deploy new services without building and provisioning data centers
    - E.g. Zynga Farmville
  - Scale up/down resources as needed
**Context**

- **2000's**
  - Large investments by web giants (e.g. Google, Amazon) in infrastructure
    - Giant data centers
    - Software Infrastructure for such data centers
      - MapReduce -- allows computations to be distributed to multiple machines “map”, and then results collected for further processing “reduce.”
      - Hadoop - open source version of above
  - Above pieces prerequisites to become a cloud provider
Reasons to be a cloud provider

- A big player enjoys economy of scale advantage
- Leverage existing investments for new revenue stream (e.g. Amazon)
- Defend existing markets (e.g. MS enterprise apps with Azure)
- Become a platform (facebook)
- Leverage relationships (IBM)
Why is the Cloud becoming big only now?

- shift from large commitment models to contactless short term model
- Mobile interactive applications that need huge data sets
- Parallel batch processing - software like Hadoop makes it easier to do this
- Analytics - less growth in plain transaction processing, more growth in analyzing trends / predictions from large data sets
Types of Utility Computing

- Amazon EC2 - to programmer, each instance looks like physical hardware
  - Can control whole layer stack
  - Other managed services provided (e.g. SimpleDB)
- Application Domain specific platforms
  - Google AppEngine (software dev. platform for web applications)
  - Force.com (Salesforce.com) - platform for business apps that use salesforce.com DB
- MS Azure -
  - Provides developers a general purpose software framework .NET
  - Compiled to a managed environment (rather than to specific hardware)
Economics

- “pay as you go” model
- add and remove resources at a fine time scale
  - proprietary data centers have to provision for peak
  - hard to predict demand of new services
  - poor service quality can alienate customer
  - large data centers have significant eco. of scale advantage
Challenges

- **Availability**
  - Can actually be better than in-house data centers
  - More robust to DDOS (Distributed denial of service) attacks by being so large

- **Lock-in**
  - Data lock-in - online storage services have gone bust
  - Application programming interfaces not common
Challenges

- **Confidentiality and Audits**
  - Sarbanes Oxley, HIPPA
  - Can use encryption
  - Audibility can be added as layer

- **Data Transfer bottlenecks**
  - Slow transfer can offset faster processing
  - Ship hard drives
  - Upload once, use multiple times
Amazon Web Services
Genesis

- Associates Group
  - Business that advertised Amazon products on their websites
- Amazon releases product data to associates group
  - API
  - Associates determine how to present it
  - Successful
Amazon Technology

- Proprietary tech to run at their scale
- Services low in the stack
  - Computing, storage, messaging, database
- Experience in running large data centers
- Leverage in acquiring software/hardware at low prices
AWS  Start

Feedback from
- Internal teams
- Company leaders
- External developers

Feedback:
- Requirements: Scalable, reliable, low-latency, simple to use
AWS First services

- Focus on Infrastructure
  - Simple Storage Service (S3)
  - Elastic Compute Cloud (EC2)
  - Simple Database (SimpleDB)
  - Simple Queue Service (SQS)
S3

- Launched March 2006
- Redundant copies
- $.12 - $.15 per GB per month
- $.10 per GB uploaded, $.17 per GB downloaded

Usage
- Small companies (SmuMug photo sharing)
- Microsoft (Vista distribution)
EC2

- Launched August 2006
- $0.10 - 0.80 per instance hour
- 10 per GB uploaded, .17 per GB downloaded

Users:
- Powerset natural language search
- AideRSS news story filtering
SQS

- 2004
- Messaging service
- .01 per 10000 requests
- .10 GB uploaded, .10 to .17 GB downloaded
Other services

- Flexible Payment Service
  - August 2007
  - Usable by developers
  - More flexibility in when transactions take place than Paypal

- Premium Support
  - Silver: MAX[ $100 per month, .10Xservive bill]
  - Gold: MAX[ $300 per month, .10-.20 X servive bill]
Market

- Digital Infrastructure Market
    - Deutsche Bank
- AWS has potential for $200 million incremental revenue
  - Deutsche Bank
- AWS ~revenue: $800 million a year today
  - (vs. 61 billion for Amazon as a whole)
Competition

- Traditional IT infrastructure
  - IBM, Sun, HP
  - Sun Grid - Compute power at $1 per hour

- Storage
  - Network Appliance, EMC, IBM, HP, Sun
  - Storage Service:
    - A few small players
Competition

- **Google Apps for your domain**
  - Offload email systems
  - Office applications

- **Google App Engine**
  - Allow developers to build, host web apps on Google infrastructure
    - Web apps only
    - Python required
Competition

- Salesforce.com
  - CRM app service
  - Force.com
    - Platform for creating business apps
Competition

- **Microsoft**
  - **Windows Live**
    - Consumer software Services: email, news headlines, blog, audio feed
  - **Microsoft Azure (Beta Release October 08)**
    - SharePoint Services (doc management)
    - .net services (app framework)
    - SQL services (data storage)
    - Live Services (data synchronization)
    - DynamicsCRM (CRM)
## Comparison

<table>
<thead>
<tr>
<th>Feature</th>
<th>Microsoft</th>
<th>Amazon</th>
<th>Google</th>
<th>VMware</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability</td>
<td>Early private CTP</td>
<td>Yes, commercially available</td>
<td>In public beta</td>
<td>Announced</td>
</tr>
<tr>
<td>Computing</td>
<td>You provide .NET code for front-end and back-end servers which Microsoft then runs on Windows 2008 virtual machines according to your environment specifications (how many machines of each kind you need, and so on.)</td>
<td>Elastic Compute Cloud (EC2) allows you to upload your XEN virtual machine images to the infrastructure and gives you APIs to instantiate and manage them.</td>
<td>You write your web application in Python or Django with a specific set of limitations set by Google and submit the application code to them.</td>
<td>Lets you easily move your virtual machines between environments and run them on premise or at any partner datacenter.</td>
</tr>
<tr>
<td>Load balancing</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Not announced</td>
</tr>
<tr>
<td>Storage</td>
<td>Yes: application storage and SQL services</td>
<td>Yes: Simple Storage Service (S3) and SimpleDB</td>
<td>Yes: database Datastore APIs</td>
<td>Not announced</td>
</tr>
</tbody>
</table>
## Comparison

<table>
<thead>
<tr>
<th>Feature</th>
<th>Microsoft</th>
<th>Amazon</th>
<th>Google</th>
<th>VMware</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Storage</strong></td>
<td>Yes: application storage and SQL services</td>
<td>Yes: Simple Storage Service (S3) and SimpleDB</td>
<td>Yes: database Datastore APIs</td>
<td>Not announced</td>
</tr>
<tr>
<td><strong>Message queuing for machine communications</strong></td>
<td>Yes: queues in Windows Azure storage</td>
<td>Yes: Simple Queue Service (SQS)</td>
<td>No</td>
<td>Not announced</td>
</tr>
<tr>
<td><strong>Integration with other services</strong></td>
<td>So called .NET services (aka BizTalk in the cloud): Access control services, workflow service, service bus, Live Mesh, Various Live services (contacts, mail, maps and so on.) At the moment, all these components do not seem to be integrated with the solution but rather bundled.</td>
<td>No</td>
<td>Yes, with existing Google services: authentication, mail, base, calendar, contacts, documents, pictures, spreadsheets, YouTube.</td>
<td>No</td>
</tr>
<tr>
<td><strong>Tied to the vendor datacenter</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No, the VMs can be hosted by any of the partners or used on-premise.</td>
</tr>
<tr>
<td><strong>Development tools</strong></td>
<td>Yes, integration into Visual Studio, support for any .NET languages.</td>
<td>Not applicable. Amazon simply runs your virtual machines and does not care which development platform you are using on top of the base OS.</td>
<td>Yes, have basic editing, local simulation, and deployment tools. Language selection limited to Python and Django. Application-level tools such as Google Web Toolkit (GWT) do not seem to have any integration with Google App Engine.</td>
<td>Not applicable. VMware simply runs your virtual machines and does not care which development platform you are using on top of the base OS.</td>
</tr>
</tbody>
</table>

AWS Pricing

- Low
- No upfront fee
**AWS reaction**

- 400K developers in 08
- Wall street reaction
  - “I have yet to see how these investments are producing any profit. They’re probably more of a distraction than anything else.” Safa Rashtchy, Piper Jaffray
  
  “Amazon is years ahead of anyone else when it comes to ‘cloud computing.’ Tomorrow’s computing environment is being dictated by Amazon.” Trip Chowdhry, Global Equities Research
Questions

- Imagine you are launching a new service and smartphone app to find green businesses in a user’s local area.
  - Would you host your new service on EC2 or build a traditional data center? Why?

- Imagine you are a large medical insurance company
  - Would you host your enterprise application for processing medical claims on the cloud or in a private data center?