Emerging Privacy Issues

Privacy Considerations in Cloud Computing

Lewis Oleinick, CIPP/G
Chief Privacy and FOIA Officer
Defense Logistics Agency
Disclaimer

The views presented herein are my own and do not represent the views of DoD or the Defense Logistics Agency.
Agenda

• What is “cloud computing?”
• Are there different types of “clouds?”
• What are the economic benefits of cloud computing?
• What are the privacy issues of cloud computing?
• What are some examples?
What is “cloud computing?”

- Wikipedia defines “cloud computing” as:
  - a paradigm of computing in which dynamically **scalable** and often **virtualized** resources are provided **as a service** over the **Internet**.[1][2] Users need not have knowledge of, expertise in, or control over the technology infrastructure in the "cloud" that supports them.[3]
“Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three delivery models, and four deployment models.”
A Picture of “Cloud Computing”

SIMPLY EXPLAINED - PART 17:
CLOUD COMPUTING
3 Types of Cloud Computing

Tim O’Reilly’s definitions:

1. **Utility computing (Infrastructure as a Service - IaaS)**
   - provides virtual machine instances, storage, and computation at pay-as-you-go utility pricing, e.g., Amazon’s AWS.
   - developers, not end-users, are the target of this kind of cloud computing.

2. **Platform as a Service - PaaS**
   - one step up from pure utility computing are platforms like Google AppEngine and Salesforce's force.com, which hide machine instances behind higher-level APIs.

3. **Cloud-based end-user applications (Software as a Service – SaaS)**
   - applications that were formerly delivered locally on a PC, like spreadsheets, word processing, databases, and even email that are now delivered over the web, e.g., gmail or Google docs and spreadsheets are "cloud computing applications."
   - Google search or Google maps, while on the same servers, are not.
4 Cloud Deployment Models*

- **Private cloud**
  - *enterprise owned or leased*
  - *may, or may not, exist on premise*

- **Community cloud**
  - *shared infrastructure for specific community*

- **Public cloud**
  - *Sold to the public, mega-scale infrastructure*

- **Hybrid cloud**
  - *composition of two or more clouds*

Risks for Each Deployment Model

**Cloud Deployment Models**

- **PUBLIC CLOUD**: Highest risk due to lack of security control, multi-tenancy, data management, limited SLA and lack of common regulatory controls.

- **PRIVATE CLOUD**: Least risk due to single ownership and strong shared mission goals and legal/regulatory requirements.

- **COMMUNITY CLOUD**: Moderate risk due to multi-tenancy, however less risk than public cloud due to shared legal/regulatory compliance issues.

- **HYBRID CLOUD**: Risk dependent upon combined models. Combination of private/community is lowest risk, while combination of public is greatest risk.
Cost/Benefit of Cloud Computing Models

Clouds - What’s the Business Case?

- **Data-as-a-Service (DaaS) Cloud(s)** (e.g., Google Analytics, DNI)
- **Software-as-a-Service (SaaS) Cloud(s)** (e.g., Salesforce.com, DKO, Intelink, A-Space)
- **Infrastructure-as-a-Service (IaaS) Cloud(s)** (e.g., Amazon WS, RACE)

Relative Per Problem Cost

Problem Agility
Privacy Policy Issues in the Cloud

- Data Security
- Privacy Act
- E-Gov’t Act (PIA)
- Breach Reporting
- International Privacy Law
- Federal Records Act
Some Technology Solutions for Security in the Cloud*

- Encrypted Swap / No Swap
- Encrypted File Systems
- Encrypted Data Transit (In/Out)
- Secured, Fit for Purpose Machine Image

“The prospect of ‘my’ data disappearing or being unavailable is far more alarming than, for example, the disappearance of a service that merely hosts an aggregated view of data that is available elsewhere say Yahoo!

search or Microsoft live maps.”
Possible Structures of a Public “Cloud” or, “Dude! Where is my Data?”

What you think you are getting.

What you may actually be getting.

To India, Malaysia or China?

Trans-border data flow of personal information?
Cloud Computing Examples

• In Government
  – OMB/GSA Cloud “one-stop” apps.gov site
  – Dept of Energy Lab’s use of Cloud
  – Dept. of Interior National Business Center's Cloud strategy
  – Army and Census use of SalesForce.com
  – DoD DISA Cloud infrastructure.

• In the Private Sector
  – GE use of Google Apps

• By the Bad Guys
• Announced 09/15/2009
• CIO Vivek Kundra
  – satisfying security and privacy concerns would be the biggest barrier to adoption of cloud solutions
  – agencies that deal with less sensitive data should shift to the cloud first
  – government would continue to own and operate systems that manage classified or sensitive data
• U.S. Department of Energy
  – Argonne National Lab Cloud Computing project
• New computation intensive simulations needed fast for critical presentation.
• All DOE resources committed.
• Argonne Lab Team used Amazon’s IaaS on a public cloud to do in minutes what would have required months.
<table>
<thead>
<tr>
<th>NBC’s Cloud Offerings include:</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBC’s IaaS offering. Will allow end-user provisioning of a variety of types of servers and operating systems through a single website.</td>
<td>NBCGrid will provide technology-agnostic server hosting, with a variety of pricing models, including metered and pre-paid, based on the customer’s usage of RAM or CPU per hour.</td>
</tr>
<tr>
<td>NBC’s Cloud storage offering.</td>
<td>NBCFiles’ Cloud storage offering will provide large capacity on a metered, pay-per-gigabyte price model. Multiple security tiers drive pricing.</td>
</tr>
<tr>
<td>NBC’s Cloud-based application marketplace.</td>
<td>NBCApps’ Cloud-based application marketplace.</td>
</tr>
<tr>
<td>Messaging, collaboration and Web 2.0 tools like wikis and blogs.</td>
<td>Gov’t-wide PA SORN and PIA?</td>
</tr>
<tr>
<td>AQD LoB SaaS: “on-demand” version of ESE.</td>
<td>Messaging, collaboration and Web 2.0 tools like wikis and blogs.</td>
</tr>
<tr>
<td>HR LoB SaaS: “on-demand” version of HR’s Onboarding, LMS, Performance &amp; Competency Management and Time and Attendance packages.</td>
<td>Gov’t-wide PA SORN and PIA?</td>
</tr>
<tr>
<td>Agency / Application</td>
<td>Use</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----</td>
</tr>
</tbody>
</table>
| US Census Bureau     | • Replaces 3000 user legacy CRM system  
| Partnership mgmt.    | • Tracks partner commitments  
|                     | • Provides near real-time reporting for congressional inquiries | 3 months |
| US Army Recruiting   | • Helps recruiters identify persons likely to join Army  
|                     | ➢ Coincided with launch of Army Experience Center  
|                     | ➢ Fosters communication  
|                     | ➢ Allows detailed statistical/ marketing analyses  
|                     | • Future capabilities  
|                     | ➢ Google Maps integration  
|                     | ➢ Link to Facebook | 3 months |
DoD DISA Cloud infrastructure

• Rapid Access Computing Environment
  - Agile and responsive computing
  - Authorized customers order and gain access to a Server in less than 24 hours
  - DoD Certification and Accreditation
  - Provides flexible development platform for Web, application or database
  - Windows, Red Hat, SUSE Servers in less than 30 minutes
  - MIPR or government credit card

User Self-service
A Senior Engineer at the Washington Post received 17,481 pages of data as non-searchable PDF files of a former White House official’s public schedule.

Using Optical Character Recognition (OCR) tools to convert and reformat the PDF pages into machine-readable text would take about 30 minutes per page using a standard PC.

Using Amazon EC2, he launched 200 server instances to process the images, at a speed of approximately 60 seconds per page, the project was completed within nine (9) hours.

Project used 1,407 hours of virtual machine time and cost $144.62
GE moved 400,000 desktops from Microsoft Office to Google Apps.

Due to privacy concerns, migrated from Google Apps to Zoho.com.

N.B. -- Zoho.com Privacy Notice includes the following:

- Contents of your Account. We store and maintain files, documents, to-do lists, emails and other data stored in your Account at our facilities in the United States or any other country. Use of Zoho Services signifies your consent to such transfer of your data outside of your country. In order to prevent loss of data due to errors or system failures, we also keep backup copies of data including the contents of your Account. Hence your files and data may remain on our servers even after deletion or termination of your Account. We assure you that the contents of your Account will not be disclosed to anyone and will not be accessible to employees of Zoho in this capacity. We also do not process the contents of your Account for serving targeted advertisements.
Password Cracking in the Cloud.

• Researchers combined COTS Password Recovery software with Amazon's cloud computing services to show:

<table>
<thead>
<tr>
<th># Characters in Password</th>
<th>Cost to Crack Password in 1 Hour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eight-character password without special characters.</td>
<td>$45</td>
</tr>
<tr>
<td>Eight-character password with 1 special character.</td>
<td>$100,000</td>
</tr>
<tr>
<td>10-character complex password with special characters</td>
<td>$10 million</td>
</tr>
</tbody>
</table>

• The researchers rented multiple dual-core virtual machines for $0.30 an hour each. The total $45 cost is equivalent to running 150 machines for one hour.
• The report includes step-by-step instructions on how to configure Amazon's cloud services COTS Distributed Password Recovery software.
Where is this all going?

Web OS 2009
A Self-Organizing, Organic Network Platform Nears the Tipping Point

Cloud Computing
Web 2.0
Logistics Info
Data Collection

Where is the Privacy?
Take-aways

1. “Cloud Computing” will have privacy implications in the Federal sector – both intended and unintended.

2. Community “Clouds,” using IaaS, such as DISA’s RACE and Ag’s NBCgrid provide good opportunities to test the waters.

3. Use of commercial “clouds” present potential international privacy issues.

4. CIO and Privacy Officer’s will need to work together to address privacy issues.
Questions

Lew Oleinick, CIPP/G
Chief Privacy and FOIA Officer
Defense Logistics Agency
lewis.oleinick@dla.mil
BACKGROUND SLIDES
<table>
<thead>
<tr>
<th>Cloud User Tools</th>
<th>Reporting &amp; Analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td>API’s</td>
<td>Analytic Tools</td>
</tr>
<tr>
<td>Workflow Engine</td>
<td>Data Mgmt</td>
</tr>
<tr>
<td>EAI</td>
<td>Reporting</td>
</tr>
<tr>
<td>Mobile Device Integration</td>
<td>Knowledge Mgmt</td>
</tr>
<tr>
<td>Data Migration Tools</td>
<td></td>
</tr>
<tr>
<td>ETL</td>
<td>Product Catalog</td>
</tr>
</tbody>
</table>

**Software as a Service (SaaS) / Applications**
- Citizen Engagement: Wikis / Blogs, Social Networking, Agency Website Hosting
- Gov Productivity: Email / IM, Virtual Desktop, Office Automation
- Gov Enterprise Apps: Business SaaS Apps, Core Mission Apps, Legacy Apps (Mainframes)

**Platform as a Service (PaaS)**
- Database, Testing Tools, Developer Tools
- DBMS, Directory Services

**Infrastructure as a Service (IaaS)**
- Storage, Virtual Machines
- CDN, Web Servers, Server Hosting

**Service Mgmt & Provisioning**
- Service Provisioning, SLA Mgmt, Performance Monitoring, DR / Backup, Operation Mgmt

**Security & Data Privacy**
- Data/Network Security, Data Privacy, Certification & Compliance, Authentication & Authorization, Auditing & Accounting

**Data Center Facilities**
- Routers / Firewalls, LAN/WAN, Internet Access, Hosting Centers

**Government Cloud Computing Framework**

"Cloud Computing" as Seen by OMB/GSA for the Cloud Line of Business