Digital Forensics for IaaS Cloud Computing

Josiah Dykstra

June 26, 2012
The views expressed in this presentation are mine alone. Reference to any specific products, process, or service do not necessarily constitute or imply endorsement, recommendation, or favoring by the United States Government or the Department of Defense.
Outline

- Today: What’s the problem?
- Trust: Can you believe the data?
- Tests: Experiments in forensic acquisition
- Trouble: Results and alternatives
- Tomorrow: Future work
Bio

- Ph.D. candidate, “Digital Forensics for Infrastructure-as-a-Service Cloud Computing”
- 8 years in network security, malware analysis, intrusion detection, forensics
First Cloud Crime?

Amazon's EC2 Cloud Service Fueled PlayStation Network Attack

By David Murphy  |  May 14, 2011 03:23pm EST  |  7 Comments  |  Email  |  Print

If you're looking for the source of the network attacks that brought down Sony's PlayStation Network—yes, it's still down—look no further than Amazon. The online retail giant didn't bring down the PlayStation Network per se, but an undisclosed source speaking to Bloomberg News has indicated that hackers used Amazon's cloud services to fuel the break-in.

According to the source, the hackers posed as a normal business and signed up for a legitimate server rental through Amazon's EC2 service—otherwise known as Amazon Elastic Compute Cloud. It's unclear how the hackers specifically used EC2 to push the attack out, which is almost as unknown a figure as the exact treasure trove of data the attackers were able to access within Sony's network.
An Investigator’s View
\[ x = \frac{1}{0} \]
Truth or Fiction?

“Incident response and computer forensics in a cloud environment require fundamentally different tools, techniques, and training...”

“Cloud computing is a model for enabling ubiquitous, 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 is composed of five essential characteristics, three service models, and four deployment models.”

http://www.nist.gov/customcf/get_pdf.cfm?pub_id=909616
September 2011
That which we call a rose...

- On-demand self-service
- Elastic
- Utility consumption
- Location independence
- Resource abstraction and pooling
You are here
## Conflicting Goals

<table>
<thead>
<tr>
<th>Cloud</th>
<th>Forensics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location independence</td>
<td>Discovery of computational structure</td>
</tr>
<tr>
<td></td>
<td>Legal jurisdiction</td>
</tr>
<tr>
<td>Rapid elasticity</td>
<td>Evidence preservation</td>
</tr>
<tr>
<td></td>
<td>Data integrity</td>
</tr>
<tr>
<td>Data reliability (replication)</td>
<td>Chain of custody</td>
</tr>
<tr>
<td></td>
<td>Evidence integrity</td>
</tr>
<tr>
<td>Multi-tenancy</td>
<td>Attribution of data</td>
</tr>
<tr>
<td></td>
<td>Chain of custody</td>
</tr>
<tr>
<td>General, abstract data structures</td>
<td>Best evidence</td>
</tr>
<tr>
<td></td>
<td>Presentation/Visualization of evidence</td>
</tr>
</tbody>
</table>
Q: Where is my data stored?

Amazon S3 offers storage in the US Standard, US West (Oregon), US West (Northern California), EU (Ireland), Asia Pacific (Singapore), Asia Pacific (Tokyo), South America (Sao Paulo), and AWS GovCloud (US) Regions. You specify a Region when you create your Amazon S3 bucket. Within that Region, your objects are redundantly stored on multiple devices across multiple facilities.

http://aws.amazon.com/s3/faqs/#Where_is_my_data_stored
Location of Customer Data

Microsoft may transfer Customer Data within a major geographic region (e.g., within Europe) for data redundancy or other purposes. For example, Windows Azure Storage geo-replication feature will replicate Windows Azure Blob and Table data, at no additional cost, between two sub-regions within the same major region for enhanced data durability in case of a major data center disaster. However, customers can choose to disable this feature.

No Tested Tools

REPORTS

- Test Results for Digital Data Acquisition Tool: Tableau Imager (TIM) Version 1.11 (March 2011)
- Test Results for Digital Data Acquisition Tool: SubRosaSoft MacForensics Lab 2.5.5 (September 2010)
- Logicube Forensic Talon Software Version 2.43 (January 2010)
- BlackBag MacQuisition 2.2 (September 2009)
- EnCase 6.5 (September 2009)
- EnCase LinEn 6.01 (October 2008)
- EnCase 5.05f (June 2008)
- FTK Imager 2.5.3.14 (June 2008)
- DCCIdd (Version 2.0, June 1 2007) (January 2008)
- EnCase 4.22a (January 2008)
- Encase Linen 5.05f (January 2008)
- dd FreeBSD (January 2004)
- Encase 3.20 (June 2003)
- Safeback 2.18 (June 2003)
- Safeback (Sydex) 2.0 (April 2003)
- dd GNU fileutils 4.0.36, Provided with Red Hat Linux 7.1 (August 2002)
No Case Law
“Forensics” Today

*Expanded Subscriber Content* (sometimes referred to as Neoprint) will be delivered in PDF format and may include:

- Profile Contact Information
- Mini-Feed
- Status Update History
- Shares
- Notes
- Wall Postings
- Friend Listing, with Friends Facebook ID’s
- Groups Listing, with Facebook Group ID’s
- Future and Past Events
- Video Listing, with filename

*User Photos* (sometimes referred to as User Photoprint) is delivered in PDF format and may include photos uploaded by the user and photos uploaded by other users that have the requested user tagged in them.
Resellers

Dropsync

Dropbox

Amazon S3
Polly traffics in child pornography.
He stores contraband images in the cloud.
He uses a pre-paid credit card.
His cloud-hosted website shares the images.

Law enforcement discovers the website and wants to terminate the service and prosecute the criminal.
Cloud Crime Scene

Cloud Service Provider

Internet

Polly

Law Enforcement

Forensic Investigator

Provider Technician

web-exposed crime scene
Key Issues in Cloud Forensics

1. Acquisition of data is more difficult.
2. Cooperation from cloud providers is paramount.
3. Cloud data may lack key forensic attributes.
4. Current forensic tools are unprepared to process cloud data.
5. Chain of custody is more complex.
<table>
<thead>
<tr>
<th>Layer</th>
<th>Cloud Layer</th>
<th>Acquisition Method</th>
<th>Cloud Trust required</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Guest Application</td>
<td>Depends on data</td>
<td>OS, HV, Host, Hardware, Network</td>
</tr>
<tr>
<td>5</td>
<td>Guest OS</td>
<td>Remote forensic software</td>
<td>OS, HV, Host, Hardware, Network</td>
</tr>
<tr>
<td>4</td>
<td>Virtualization</td>
<td>Introspection</td>
<td>HV, Host, Hardware, Network</td>
</tr>
<tr>
<td>3</td>
<td>Host OS</td>
<td>Access virtual disk</td>
<td>Host, Hardware, Network</td>
</tr>
<tr>
<td>2</td>
<td>Physical Hardware</td>
<td>Access physical disks</td>
<td>Hardware, Network</td>
</tr>
<tr>
<td>1</td>
<td>Network</td>
<td>Packet capture</td>
<td>Network</td>
</tr>
</tbody>
</table>
Experiments

Experiment 1 (Guest OS)
- Launch and “hack” a virtual machine in EC2
- Use EnCase and FTK agents to acquire disk images remotely
- Use Fastdump, FTK Imager, Memoryze to acquire memory images remotely
- Analyze data to determine success

Experiment 2 (Virtualization)
- Launch and “hack” a virtual machine on a local cloud
- Use introspection to inject an EnCase agent to acquire disk image
- Create virtual machine snapshot and analyze “live” offline
- Analyze data to determine success

Experiment 3 (Host OS)
- Launch and “hack” a virtual machine in EC2
- Use AWS Export to obtain a disk image
- Analyze data to determine success
Experiments
## Results

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Tool</th>
<th>Evidence Collected</th>
<th>Time (Hrs)</th>
<th>Trust Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>EnCase</td>
<td>Success</td>
<td>12</td>
<td>OS, HV, Host, Hardware, Network</td>
</tr>
<tr>
<td>1</td>
<td>FTK</td>
<td>Success</td>
<td>12</td>
<td>OS, HV, Host, Hardware, Network</td>
</tr>
<tr>
<td>1</td>
<td>FTK Imager (disk)</td>
<td>Success</td>
<td>12</td>
<td>OS, HV, Host, Hardware, Network</td>
</tr>
<tr>
<td>1</td>
<td>Fastdump</td>
<td>Success</td>
<td>2</td>
<td>OS, HV, Host, Hardware, Network</td>
</tr>
<tr>
<td>1</td>
<td>Memoryze</td>
<td>Success</td>
<td>2</td>
<td>OS, HV, Host, Hardware, Network</td>
</tr>
<tr>
<td>1</td>
<td>FTK Imager (memory)</td>
<td>Success</td>
<td>2</td>
<td>OS, HV, Host, Hardware, Network</td>
</tr>
<tr>
<td>1</td>
<td>Volume block copy (dd)</td>
<td>Success</td>
<td>14</td>
<td>OS*, HV, Host, Hardware, Network</td>
</tr>
<tr>
<td>2</td>
<td>Agent Injection</td>
<td>Success</td>
<td>1</td>
<td>HV, Host, Hardware, Network</td>
</tr>
<tr>
<td>3</td>
<td>AWS Export</td>
<td>Success</td>
<td>120</td>
<td>AWS Technician, Technician’s Host, Hardware &amp; Software, AWS Hardware, AWS Network</td>
</tr>
</tbody>
</table>
Trouble

- Forensic workstation online
- Security of remote agent
- Cost – time and $$$
- Changes to cloud environment
- Legal questions
Alternatives

- Root Trust in the Host/VM with TPMs
- Collection from Management Plane
- Forensics Support as a Service
- Contract and Legal Solutions
Management Plane

Welcome

The AWS Management Console provides a graphical interface to Amazon Web Services. Learn more about how to use our services to meet your needs, or get started by selecting a service.
Potential Forensic Data

- Billing records
- Netflow, Packet Capture
- API/Management access logs
- Security logs (firewall, IDS, etc.)
- Physical drives
- Virtual drives
- Guest OS data
- Cloud data storage
Legal Considerations

- Expectation of privacy
- Possession, custody, control
- Data preservation
- Jurisdiction
- Seizing Data
- Defenses
  - Complexity, production, *Daubert*, ripeness
Tomorrow: Future work

- Corroborate from multiple layers
- “Live” Forensics with Snapshots
- Parallel analysis of PaaS and SaaS
- Consumer-driven forensic capabilities
  - Log retrieval, metadata (volume checksums), volume download
- Legal analysis
  - How do you legally obtain cloud documents?
  - Who legally owns the data/IP address/etc?
  - Who’s law applies to the data? To the forensics?
Summary

- Cloud challenges forensic acquisition
- Need to trust the data acquired
- Some tools work. Some are needed.