Security

15-719/18-847b
Advanced Cloud Computing

Garth Gibson
Greg Ganger
Majd Sakr
Security readings


Public Infrastructure Cloud Services

- $130B?
- < 10 years
- Little VMM or stack diversity
- Data derived from Gartner 2013

Table II. Cloud Service Providers Surveyed

<table>
<thead>
<tr>
<th>Service Provider</th>
<th>Launch Year</th>
<th>Hypervisor(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verizon Cloud</td>
<td>2014</td>
<td>Xen/VMware</td>
</tr>
<tr>
<td>Google</td>
<td>2013</td>
<td>KVM</td>
</tr>
<tr>
<td>Savvis Direct</td>
<td>2012</td>
<td>Xen/VMware</td>
</tr>
<tr>
<td>HP Public Cloud</td>
<td>2012</td>
<td>KVM</td>
</tr>
<tr>
<td>Dimension Data</td>
<td>2011</td>
<td>VMware</td>
</tr>
<tr>
<td>Tier 3</td>
<td>2011</td>
<td>VMware</td>
</tr>
<tr>
<td>Microsoft Azure</td>
<td>2010</td>
<td>Custom(Hyper-V)</td>
</tr>
<tr>
<td>Fujitsu Trusted Public S5</td>
<td>2010</td>
<td>Xen</td>
</tr>
<tr>
<td>GoGrid Cloud Platform</td>
<td>2009</td>
<td>Xen</td>
</tr>
<tr>
<td>Joyent Compute/Manta Storage</td>
<td>2009</td>
<td>SmartOS</td>
</tr>
<tr>
<td>Amazon EC2/S3</td>
<td>2008</td>
<td>Xen</td>
</tr>
<tr>
<td>Rackspace Public Cloud</td>
<td>2008</td>
<td>Xen</td>
</tr>
<tr>
<td>SoftLayer</td>
<td>Unknown</td>
<td>Xen</td>
</tr>
</tbody>
</table>

[Huang15]
Survey state of art in IaaS security

• Consider IaaS cloud services, based on hypervisors & cloud environment software stacks (e.g. AWS, OpenStack)

• Consider attackers to be of two types:
  o Attacker runs on same machine, but not in cloud service provider (CSP)
  o Attacker has breached CSP on same or other machines in cloud services

• Security issues (Confidentiality, Integrity, Availability)
  o Customer/user data is protected, plus its existence & access patterns
  o Cloud service provider (CSP) provides correct and honest functions
    • Failure to meet Contract SLA (CS) is a security issue (some SLA goals are hard to verify, e.g. how reliable is my cloud data storage?)
## Survey of Public IaaS CSP Offerings

**Table III. Summary of Areas where the IaaS Industry has Established Best-Practices** [Huang15]

<table>
<thead>
<tr>
<th>CSP</th>
<th>Hypervisor</th>
<th>Firewall</th>
<th>Data-center Op.</th>
<th>Dedicated VM</th>
<th>Segregation</th>
<th>Crypto-Transit</th>
<th>Crypto-Rest</th>
<th>Auth</th>
<th>Delegation/ACL</th>
<th>CSP security</th>
<th>Location Constraints</th>
<th>SLA</th>
<th>Bill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Object</td>
<td>2 factor passwords SSO</td>
<td>Role-based</td>
<td>No spoofing</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Verizon</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Object</td>
<td>2 factor passwords, PIN</td>
<td>Role-based</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SavvisDirect</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Object</td>
<td>Passwords Keystone (SSO)</td>
<td>Access with tokens</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Rackspace</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Passwords Keystone (SSO)</td>
<td>Role-based</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Azure</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Passwords Keystone (SSO)</td>
<td>Rest-API</td>
<td>Some</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>DimensionData</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Object</td>
<td>Passwords Keystone (SSO)</td>
<td>Role-based</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Tier 3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Object</td>
<td>Passwords REST</td>
<td>Create users</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Joyent</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
<td>Password SSH Key</td>
<td>Sub users</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Fujitsu</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Certificate based</td>
<td>Sub users</td>
<td>Mentioned as risk</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>GoGrid</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Password/ API key</td>
<td>×</td>
<td>Prevent spam</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SoftLayer</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Password, API key</td>
<td>Sub users</td>
<td>Sub accounts</td>
<td>Not mentioned</td>
<td>✓</td>
</tr>
<tr>
<td>HP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Password Keystone (SSO)</td>
<td>Sub users</td>
<td>Sub accounts</td>
<td>Not mentioned</td>
<td>✓</td>
</tr>
<tr>
<td>Google</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Passwords, OAuth2</td>
<td>via OAuth2</td>
<td>Prevent spam</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
CSP Security Mechanisms (inherited from pre-cloud)

- **Systems**
  - Hypervisor – Xen, KVM, VMWare, Hyper-V – maturing through use and abuse
  - Firewall – separate from VM to emulate separate hardware
  - Data center operations – best practice human/physical practices from hosting market
  - Dedicated VMs – virtual but not shared, sometimes offered (vs private cloud)
  - Corporate Segregation – usually protected from CSP’s other business/staff

- **Cryptographic**
  - In transit – SSL/TLS, riding Web 2.0 technology (e.g. certificates)
  - At rest – does encryption at rest by CSP offer real protection? CSPs disagree

- **Access control**
  - Authentication – single sign on (SSO) service diversity (Oauth, Keystone, SAK, IAM)
  - User creation and ACLs – “file system like” plus time, IP address controls; disagreement on how constrained/delegatable/revocable?
Security Mechanisms con’t

• CSP Security – how transparent is CSP security of themselves?
  o Eg. IP address management; banned/blacklist of CSP IP protected

• CSP contract
  o Location of VM – country restrictions, geo-replication for avail/latency
  o SLAs – all vary price based on SLA availability, but few other properties
  o Billing – bytes/hours verify easily, compute quality not so much
    • Project 2 – multiple AZ when you didn’t ask for costs more
Survey of Public IaaS CSP Offerings

Table III. Summary of Areas where the IaaS Industry has Established Best-Practices

<table>
<thead>
<tr>
<th>CSP</th>
<th>Hypervisor</th>
<th>Firewall</th>
<th>Data-center Op.</th>
<th>Dedicated VM</th>
<th>Segregation</th>
<th>Crypto-Transit</th>
<th>Object</th>
<th>Delegation/ACL</th>
<th>CSP security</th>
<th>Location Constraints</th>
<th>SLA</th>
<th>Bill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Object</td>
<td>2 factor passwords SSO</td>
<td>Role-based</td>
<td>No spoofing</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Verizon</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Object</td>
<td>2 factor passwords, PIN</td>
<td>Role-based</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SavvisDirect</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Object</td>
<td>Passwords Keystone (SSO)</td>
<td>Access with tokens</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Rackspace</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Passwords Keystone (SSO)</td>
<td>Role-based</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Azure</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Passwords Keystone (SSO)</td>
<td>REST-API</td>
<td>Some</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>DimensionData</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Object</td>
<td>Passwords Keystone (SSO)</td>
<td>Role-based</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Tier 3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Object</td>
<td>Passwords Keystone (SSO)</td>
<td>Create users</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Joyent</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
<td>Password SSH Key</td>
<td>Sub users</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Fujitsu</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Certificate based</td>
<td>Sub users</td>
<td>Mentioned as risk</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>GoGrid</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Password/ API key</td>
<td>Prevent spam</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>SoftLayer</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Password, API key</td>
<td>Sub users</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>HP</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Password Keystone (SSO)</td>
<td>Sub users</td>
<td>Not mentioned</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Google</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Block</td>
<td>Passwords, OAuth2</td>
<td>via OAuth2</td>
<td>Prevent spam</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

[Huang15]
Broader academic issues

• Attack targets
  o Cache channel – measure cache access time to deduce what data a co-
    resident victim code is touching (Ristenpart09)
  o Storage channel – even encrypted, shared storage access pattern can be
    inferred, unless real data “moves around”
  o Covert channels – conspiring VMs can hide communication using above
  o Image sharing – shared VM root images very useful, so dangerous
  o Leak prevention – hybrid apps split over different CSP/private clouds
Broader academic issues con’t

• Integrity, Availability Issues
  o Proof of Possession/Retrieval – challenge/response statistical sampling
  o Storage Integrity – digital signature, P2P tests of mutation ordering
  o Compromised VMM – TPM HW hash of running code,
  o Hardened VMMs – reduce complexity/size of VMM code/API
  o Coping with VMMs – verifiable computation or homomorphic encrypt

• Contractual Issues – trusted auditors, proof of location by latency
Discussion of academic vs industry defenses

<table>
<thead>
<tr>
<th>Problem</th>
<th>Current Solutions and Results</th>
<th>Summary</th>
</tr>
</thead>
</table>
| Malicious or compromised CSPs               | **Academia:** POR/PDP, hypervisor integrity, auditing  
**Industry:** Marketing                        | CSPs assume they are trusted and disclaim liability for damages due to security compromises. Academia serves an important role for identifying weaknesses in CSP security that would be hard for CSPs themselves to talk about publicly. |
| Cross-VM leakage, cache timing channels, covert channels | **Academia:** Evaluation of feasibility of attack, memory placement, deterministic execution  
**Industry:** Dedicated instances           | Industry proposes dedicated instances. Main question is the value of this solution since it costs more for a dedicated instance. Academia could help by definitively establishing the feasibility of such attacks and the associated costs with mounting the attack. |
| Malicious VM images and leakage through VM images | **Academia:** Demonstrated the size and scope of the problem through measurement.  
**Industry:** Provide documentation and warnings to users.                                | A simple tool to detect leakage and malicious VMs has been constructed in academia. There is potential for a more comprehensive solution to solve this problem.                                               |
| Hypervisor integrity and compromised hypervisors | **Academia:** Hypervisor hardening, using untrusted hypervisors, TPM-based remote attestation by customers.  
**Industry:** Hypervisor customizations and keeping hypervisor patched.                       | Direct TPM-based remote attestation by customers is of limited use due to use of industry use of customized hypervisors. Protecting against an untrusted hypervisor is the only viable solution against a malicious CSP. |
| Billing Integrity                            | **Academia:** Demonstrate attacks against billing integrity, auditing techniques to ensure that difficult to verify security properties of cloud service are upheld.  
**Industry:** Bill based on easily verifiable quantities.                                      | Unfortunately, industry billing practice multiplies easily verifiable quantities with rates based on difficult to verify properties. There is potential for research into better auditing and metering techniques. |
Attestation

- If you control original code in secure container and secure container signs a computation in secure container, then you are trusting manufacturer

Figure 53: The chain of trust in software attestation. The root of trust is a manufacturer key, which produces an endorsement certificate for the secure processor’s attestation key. The processor uses the attestation key to produce the attestation signature, which contains a cryptographic hash of the container and a message produced by the software inside the container.

[Costan16, IACR Cryptology]
Next day plan

• Monday: Greg will talk about learning to love spot pricing
• Wednesday: Guest lecture on formal verification of cloud services
  • Bryan Parno, CMU
• Monday after: TBD
• Wednesday after: second midterm
  o Closed book, no electronics, one 8.5x11” cheat sheet allowed
  o In-class, sit with an empty sit on both sides of you
  o Topics covered include project 3 (scheduling) and all lectures after exam 1
Cloud Security

Marcus Peinado
Microsoft Research
Can you trust the cloud?

• Huge trusted computing base
  • Privileged software
    Hypervisor, firmware, ...
  • Management stack
  • Staff
    Sysadmins, cleaners, security, ...
  • Law enforcement

• Hierarchical security model
  • Observe or modify any data
  • Even if encrypted on disk / net

Application
Operating system
Hypervisor
Firmware/bootloader
Management tools
People

Trust
Homomorphic Encryption: Toy Example

- RSA encryption: \( E(x) = x^e \mod m \)
- \( E(x) \cdot E(y) = (x^e \mod m) \cdot (y^e \mod m) = (xy)^e \mod m = E(xy) \)
- We can multiply on encrypted data!
Fully Homomorphic Encryption (FHE)

• Can we do more than just multiply?

• Yes. FHE systems can perform arbitrary operations on encrypted data [Gentry 2009].

• But they are very slow (at present)
  • Orders of magnitude slowdown

• Research directions
  • Can we reduce the overhead? (Crypto research)
  • Can we trade off functionality and security for speed?
Example: CryptDB [Popa 2011]

• SQL database that runs on encrypted data.
• DB systems perform only a limited set of operations on the data.
• Weaken encryption for some DB columns to enable interesting queries.
Encrypting Databases

Select * from PatientRecords
where **SSN = dwom58anb**

Need **deterministic encryption** for columns used in simple filters.

This may leak some information.

<table>
<thead>
<tr>
<th>SSN</th>
<th>Patient Name</th>
<th>Age</th>
<th>diagnosis</th>
<th>doctor’s comments</th>
</tr>
</thead>
</table>
| xz6f73bdfk| 4jruaiu B4ur8w | kjewh | dsasdfs   | kljaenfja al ksdfisegi i ooi aoei wrg io |}
| 45geio809d | jdjncs 7832ibhe | ;kds   | gfhgf      | df, adsmkladfkl slkjdn lkjdf ndkfu isdfdiufuisads sd |
| kf9sh23kdfb | kwjek 9e89cjf | run8f  | dsfadf     | df, adsmkladfklslg slkjdn lkjdf ndkfu isdfdiufuisads sd |
| 8dnklklsjfb3 | 9r09f 90sd9f | 98udc  | hghj       | tghtrstghfg rweakosaen ser wert24rt gwyew5y |
| dwom58anb  | 0j98fi0f ikdfm90er | ;kds   | fgdfg      | w545425 456 567 56g wrt bgh rt er t bw5b th wey55by |
| h698sab4klb | pwoe0.9fuwe | 890uefw | ythhgfpg | eoir vwer wrey6 7iwt w454 34 74t t drjksd fg re ge |

Select * from PatientRecords
where **Age > rur8f**

Need **order-preserving encryption (OPE)** for columns used in range queries.

This leaks more information.

Select * from PatientRecords

If no query depends on a column then we can encrypt it at full strength.
Trusted Hardware

• Assume special hardware features.
• Then we can run regular programs at native speed.
• And the CSP cannot access our data.
Trusted Hardware

Cloud Storage

1. Isolation

Virtualization Stack

User VM

2. Attestation

Hypervisor

Other people’s VMs

Good cloud user: Alice
Intel SGX

- Hardware isolation for an enclave
  - New instructions to establish, protect
  - Call gate to enter
- Attestation
- Available in the most recent Intel CPUs (Skylake).
  - Only client processors
SGX: Isolation (conceptually)

- **Administrators** control software
- Hypervisor
- Main operating system
- SGX

**Administrators cannot alter SGX**
Initialization: Building an Enclave

- Enclaves are built by untrusted code.
  - Enclave code
  - Enclave data
- CPU/SGX records enclave hash.

- Untrusted code can tamper with enclave creation
- But any tampering will be recorded in the enclave hash
Enclave Software

• Can run arbitrary code in enclaves.

• Challenges:
  • External dependencies: Cannot trust the system beyond the enclave boundaries.
  • Only Ring 3 (user mode): Hard to run a VM/OS this way.

• First results:
  • Haven [Baumann 2014]: Arbitrary Windows applications.
  • VCCC [Schuster 2015]: Hadoop
Trusted Hardware: Problems

• No SGX hardware available
• SGX does not support VMs
• Limitations
• Still have to trust
  • An entire CPU
  • Intel Corporation
  • Manufacturing plant
• Compromised attestation keys
• Hardware attacks