Cloud Security & Cryptography I

Cloud Computing

SENY KAMARA
MICROSOFT RESEARCH
Computing as a Service

- Computing is a vital resource
  - Enterprises, governments, scientists, consumers, ...
- Computing is manageable at small scales...
  - e.g., PCs, laptops, smart phones
- ...but becomes hard to manage at large scales
  - build and manage infrastructure, schedule backups, hardware maintenance, software maintenance, security, trained workforce, ...
- Why not outsource it?
Computing Architecture

Applications
- Email, WWW, Social Net.,....

Platform
- Windows, Linux, MacOSX,...

Infrastructure
- memory, disk, network,
Cloud Services

- **Software as a service**
  - Gmail, Hotmail, Flickr, Facebook, Office365, Google Docs, …
  - Service: customer makes use of provider applications
  - Customer: consumers & enterprise

- **Platform as a service**
  - MS SQL Azure, Amazon SimpleDB, Google AppEngine
  - Service: customer makes use of provider’s software stack
  - Customer: developers

- **Infrastructure as a service**
  - Amazon EC2, Microsoft Azure, Google Compute Engine
  - Service: customer makes use of provider’s (virtualized) infrastructure
  - Customer: enterprise, developers
Cloud Deployment Models

Public

Private
Why the Hype?
Why Providers Care

- Spare capacity
  - most providers have underutilized data centers
  - might as well monetize it
- Potentially huge market
- Major *infrastructure shift*
  - Comparable to the Internet (?)
  - MS, Apple, Google, Amazon, Facebook
  - Can’t risk missing it
Why Clients Care

- Consumers
  - Convenience: backups, synchronization, sharing
- Startups/SME
  - Low CAPEX: low risk, less VC
  - Focus on product/service
  - Elasticity (can scale fast)
- Enterprise
  - Turn CAPEX into OPEX
  - Cheaper & more reliable services (email, payroll, …)
Why Researchers Care

- Papers!
- Grants!
- Interesting research
  - Distributed systems: fault-tolerance, cluster & parallel computing
  - Storage systems: GFS, HDFS, ...
  - Databases: Big Data, analytics, NoSQL, GraphDBs
  - Operating systems: virtualization
  - Algorithms: resource allocation, cluster algorithms, parallel algs
  - Economics: pricing, auctions
  - Security: forensics, VM isolation,
  - Networking: data center networks, architectures, protocols
  - Cryptography: new types of encryption, signatures, protocols, ...
Why Governments Care

- Cloud will impact cost of hardware and software
  - will impact the cost structure of many industries
  - will impact business creation
  - will impact economic performance of countries
- Cloud can provide cost savings for public sector
  - Hospitals, healthcare, education
  - Agencies that have periodic peaks (e.g., IRS)
  - Improved energy efficiency
    - Europe: 1.75% of carbon emissions due to IT usage
What are the Risks?
Cloud Policy

- What is the legal definition of a Cloud?
- Determines regulatory & policy frameworks
- What if
  - cloud's computation is wrong?
  - data stored is tampered with or lost?
  - customer goes out of business?
Cloud Policy [Jaeger-Lin-Grimes08]

- Should Telecom laws apply?
- Entities in telecom laws
  - ISP, telecomm providers, common carrier
- Telco laws assume purpose of technology is to ship bits
  - Do not offer legal compensation framework
  - If call or packets are dropped, just resend
- Cloud stores, computes and ships
  - What happens if data is lost?
Cloud Policy

- If Clouds are Telcos should net neutrality apply?
  - Net neutrality is good for Clouds
    - Cloud relies on stable and high quality Internet access
    - Prevents ISPs from extracting profits from providers
    - Prevents ISPs from gaining unfair advantage for own clouds
  - Net neutrality could be disastrous for Clouds
    - No differential pricing
    - No QoS
Cloud Policy

- Is a Cloud responsible for its tenants?
  - EC2 hosted Wikileaks and spammers
  - What if DoS attacks are launched from the Cloud?
  - What if hackers use cloud as stepping stone?
Cloud Insurance

Should customers be insured?

- 100% reliability is impossible
- Downtime can be costly (startups can go out of business)

AWS outages

- **December 12th, 2010**: EC2 down for 30 mins (Europe)
- **April 21, 2011**: storage down for 10-12 hours (N. Virginia)
  - Foursquare, Reddit, Quora, BigDoor and Hootsuite affected
- **August 6th, 2011**: storage down for 24 hours (Ireland)
- **August 8th, 2011**: network connectivity down for 25 mins (N. Virginia)
  - Reddit, Quora, Netflix and FourSquare affected
- **July 7th, 2012**: storage down for few hours (Virginia)
  - Instagram, Netflix, Pinterest affected

What is the right model for Cloud insurance?
Data-Related Issues

- Where is the data?
  - In which legal jurisdiction?
  - Does that government have access?
  - Which regulations apply?
- Compliance
  - If I store data of type X, am I compliant with regulation Y?
- Licensing
  - If I store licensed data and/or code, am I violating terms?
Data-Related Issues [Reed10]

- Who owns the data?
  - No notion of property rights for information
  - Property rights only for physical object that stores information
  - “owner” can control information through mix of IP, privacy rights and contracts

- Typical Cloud scenario
  - Customer entrusts own data + data of clients to cloud
  - Cloud stores and processes data
  - Client uses cloud services to create new data
  - Cloud generates metadata and new data
Data-Related Issues

What can the Cloud do with Data?
- Can Cloud mine tenant data to improve its cloud services?
- Can Cloud mine tenant data to improve its other products
  - Can MS mine cloud data to improve Bing, Office, ... ?
Data-Related Issues

- Google Drive
  - Released April 24th, 2012
  - Similar to Dropbox, Skydrive, etc...
  - Media firestorm with respect to license
  - User retains intellectual property rights
  - Google retains rights to
    - reproduce, use, and create derivative works
    - Extract content to customize advertising and other services
    - *perpetually*...even after removal of content!
Data-Related Issues

- Entropy reduction [Ohm09]
  - anonymized data sets can be de-anonymized using auxiliary information
  - Cloud providers hold a large amount of auxiliary information!
  - Therefore can have large effect on privacy
  - Should they be regulated?
Government Surveillance

- Gordon Frazer
  - managing director of Microsoft UK
  - Office 365 Launch (July, 2011):
    - “cloud data is not protected against US Patriot Act...”
    - “…no matter where it is stored, …”
    - “and we might give data without telling you”
  - Huge controversy!
Government Surveillance

- Ivo Opstelten [Dutch minister of safety & justice]
  - US providers could be excluded from bidding on Dutch contracts
- Sophie in ‘t Veld [Dutch member of European Parliament]
  - asked European Commission to clarify jurisdictional issues urgently!
  - But banning transfer of European (citizen) data to U.S. could violate WTO agreements…
Government Surveillance

- France
  - invested 150/225M euros in SFR & Orange
  - so CloudWatt & Numergy have *local* data centers?
The Patriot Act

- 1968: Omnibus Crime Control and Safe Streets Act
  - Prohibits interstate gun sales, set 21 as minimum age to buy guns, ...
  - Also set rules for obtaining wiretap orders in the United States

- 1986: Electronic Communications Privacy Act
  - Amendment to OCCSSA
  - Prevents unauthorized government access to private electronic communications

- 2001: “Patriot Act”
  - Series of amendments to previous acts including ECPA
  - Increased law enforcement's ability to recover data and communications
The Patriot Act

- EU allows private data to be exported to
  - Argentina, Israel, most of Canada, ...
  - ...but not to US or most of Asia

- Safe Harbor
  - US companies promise to enact certain security & privacy measures
  - Most US companies agree
  - SH has exception for national security...
  - But SH was enacted before 911 and PA
  - EU would have never agreed to SH if it knew PA was coming
Patriot Act

Effects of controversy

- EU enterprises and govs nervous about US clouds
- Great for EU cloud providers!
- US cloud providers asked Obama administration to clarify scope of PA
Cloud Adversarial Models
The Cloud Abstraction
The Cloud Abstraction
The Cloud Abstraction
The Cloud Abstraction
The Cloud Abstraction
The Cloud Abstraction
The Cloud Abstraction
The Cloud Abstraction
The Cloud Abstraction
Cloud Adversarial Models

- Clouds must protect against traditional adversaries
  - Hackers, malware, botnets, spammers, ...

- And against
  - Physical attackers
  - Rogue employees: can access part of infrastructure
    - Steal hard drives, see PII
  - Tenants: are like traditional adversaries but inside the cloud
    - DoS, cross-VM attacks
  - Providers: control entire infrastructure
    - hardware, OS, HV, network, data center
  - Governments: can issue subpoenas, get warrants, ...
    - Get keys, hard drives, servers, monitor communications
Cloud Attacks
Overview of EC2

- Infrastructure cloud (IaaS)
- 1st generation compute instances
  - **M1 Small**: 1.7GB, 1 v-core & 1 ECU, 160GB storage (6 c/hr)
  - **M1 Medium**: 33.75GB, 1 v-core & 2 ECU, 410GB storage
  - **M1 Large**: 7.5 GB, 2 v-cores & 2 ECU each, 850GB storage, 64-bit
  - **M1 XLarge**: 15GB, 4 v-cores w/ 2 ECU each, 1690GB storage, 64-bit (1 $/hr)
- 2nd generation compute instances
  - **M3 XLarge**
  - **M3 Super XLarge**
Overview of EC2

- More instances
  - High memory instances
  - High CPU instances
  - Cluster compute instances
  - Cluster GPU instances
Overview of EC2

- **Storage**
  - Instance local storage (volatile)
    - Size depends on instance type
  - Elastic Block Store (≈ virtual hard drive)
    - Up to 1TB per volume

- **Pricing options**
  - On-demand instances (pay per use)
  - Reserved instances (pay up front) & marketplace
  - Spot instances (bid and use while < spot price)
Overview of EC2

- **Regions**
  - US East (Northern Virginia)
  - US West (Oregon)
  - US West (Northern California)
  - EU (Ireland)
  - Asia Pacific (Singapore)
  - Asia Pacific (Tokyo)
  - South America (Sao Paulo)
  - GovCloud (US)

- **Availability zones**
  - Insulated from each other
  - Zone 1 cannot affect Zone 2 & 3
Attacking EC2 Tenants

- [Ristenpart-Tromer-Shacham-Savage09]
- Cloud cartography
  - Map internal IP to instance parameters
- Co-location
  - Place an attack VM on same server as target
- Co-residency checks
  - Check if attack VM is co-located with target VM
- Cross-VM attacks
  - Steal keys using a cache-based side-channel attack
Cloud Cartography

- Map from internal IP to instance parameters
  - Launched 20 instances for every zone/type (3x5) in US
  - EC2 IP space partitioned by zone/type
- Using cartography
  - Get target’s external IP
  - Query internal DNS service for internal IP
  - Use map to guess instance type and zone
Co-Location

- Co-location strategy #1
  - Just launch as many VMs as possible in same zone+region
- EC2 co-location is biased towards
  - Sequentially launched VMs
  - Parrallely launched VMs from different accounts
- Co-location strategy #2
  - Launch attack VMs as soon as target VM is launched
  - Or overload target and wait for Autoscaling
Co-Residence Testing

- EC2 observations
  - First hop from any VM is Dom0
  - Numerically-close IPs typically assigned to same server
- Co-residency testing
  - If IPs are numerically close
  - Traceroute to target should include only 1 hop (Dom0)
Cross-VM Attacks

- Coarse-grained [Ristenpart-Tromer-Shacham-Savage09]
  - Recovers traffic rates, keystroke activity, ...
  - Single-core virtualized server (running Xen)
  - Cache attack (L1 data cache)
  - Requires co-locating 1 VM
Cross-VM Attacks

- **Fine-grained** [Zhang-Juels-Reiter-Ristenpart12]
  - Recovers El Gamal secret key (457-bit exp & 4096-bit p)
  - Multi-core (4) virtualized server (running Xen)
  - Cache attack (L1 instruction cache)
  - Requires co-locating 1 VM with 2 VCPUs
    - VCPU1 probes (measures victim through cache)
    - VCPU2 issues interrupts to force Xen to run VCPU1
  - Uses machine learning (SVMs) + HMMs extract signal
  - Training SVMs requires
    - machine with same architecture & victim code
More Cloud Attacks

- Amazon machine images [Bugiel et al.11]
  - Analyzed 1225 AMIs
  - Found source code, private keys, administrator pwds
- Topology inference [Raiciu-Ionescu-Niculescu12]
  - Mapped the EC2 US-EAST AvZ D data center network
  - Cloud-specific covert DoS attacks
Securing the Cloud

- How do we secure public infrastructure clouds?
  - Systems security: virtualization, isolation, access control, ...
  - Network security: firewalls, network intrusion detection, ...
- How do we protect against all adversaries?
  - New systems security mechanisms
  - New cryptographic techniques!