Cloud Security and Privacy

From Prof Shan Lin
Stony Brook University

Outline

- Part I: Introduction
- Part II: Security and Privacy Issues in Cloud Computing
- Part III: Possible Solutions

Part I. Introduction

- Cloud Computing Background
- Cloud Models
- Why do you still hesitate to use cloud computing?
- Causes of Problems Associated with Cloud Computing
- Taxonomy of Fear
- Threat Model

Cloud Computing Background

- Features
  - Use of internet-based services to support business process
  - Rent IT-services on a utility-like basis
- Attributes
  - Rapid deployment
  - Low startup costs/ capital investments
  - Costs based on usage or subscription
  - Multi-tenant sharing of services/ resources
- Essential characteristics
  - On demand self-service
  - Ubiquitous network access
  - Location independent resource pooling
  - Rapid elasticity
  - Measured service
- "Cloud computing is a compilation of existing techniques and technologies, packaged within a new infrastructure paradigm that offers improved scalability, elasticity, business agility, faster startup time, reduced management costs, and just-in-time availability of resources"

From [1] NIST

A Massive Concentration of Resources

- Also a massive concentration of risk
  - expected loss from a single breach can be significantly larger
  - concentration of "users" represents a concentration of threats
- "Ultimately, you can outsource responsibility but you can't outsource accountability."

Cloud Computing: who should use it?

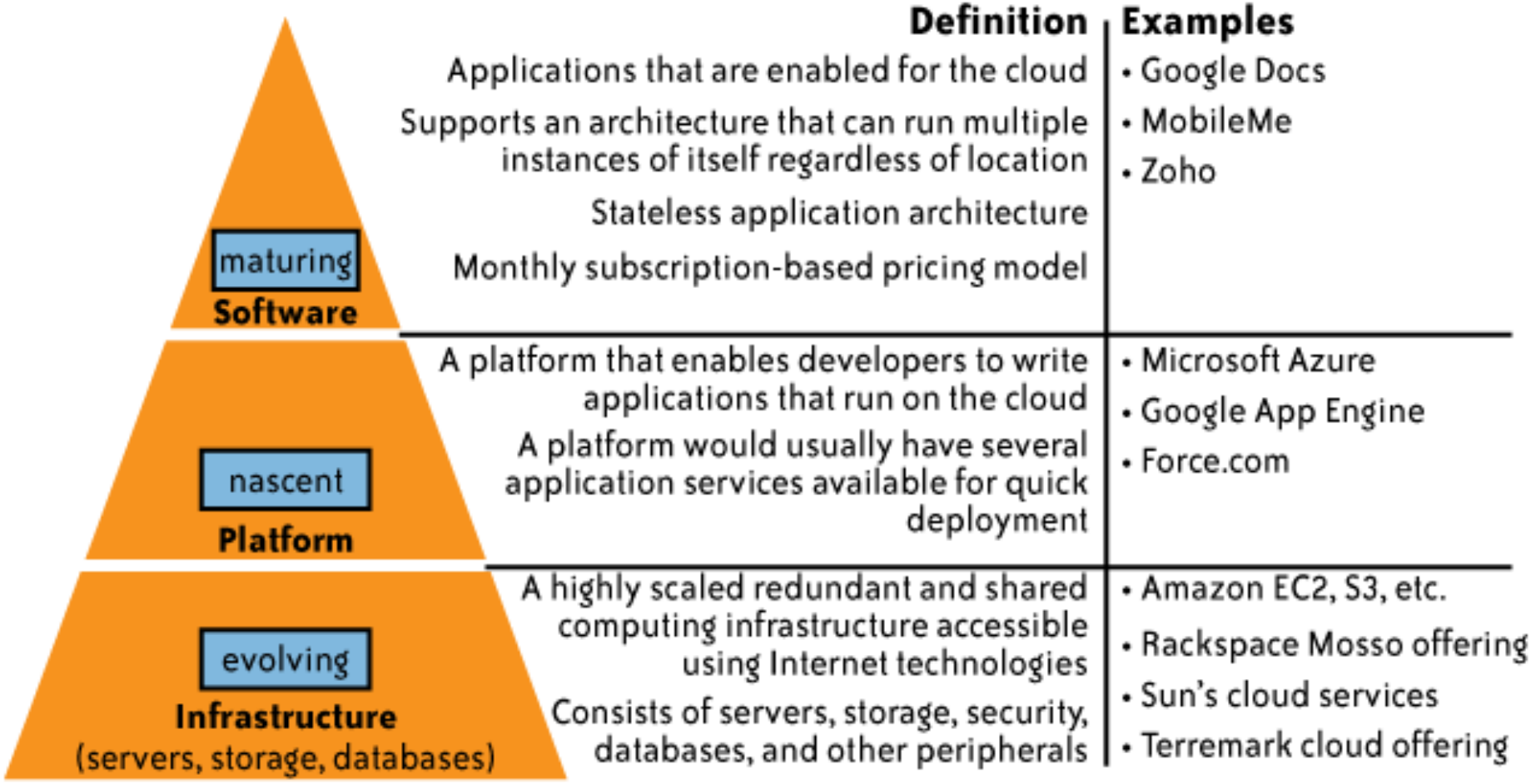
- Cloud computing definitely makes sense if your own security is weak, missing features, or below average.
- Ultimately, if
  - the cloud provider's security people are "better" than yours (and leveraged at least as efficiently),
  - the web-services interfaces don't introduce too many new vulnerabilities, and
  - the cloud provider aims at least as high as you do, at security goals,

  then cloud computing has better security.

Cloud Models

- Delivery Models
  - SaaS
  - PaaS
  - IaaS
- Deployment Models
  - Private cloud
  - Community cloud
  - Public cloud
  - Hybrid cloud
- Management Models (trust and tenancy issues)
  - Self-managed
  - 3rd party managed (e.g. public clouds and VPC)

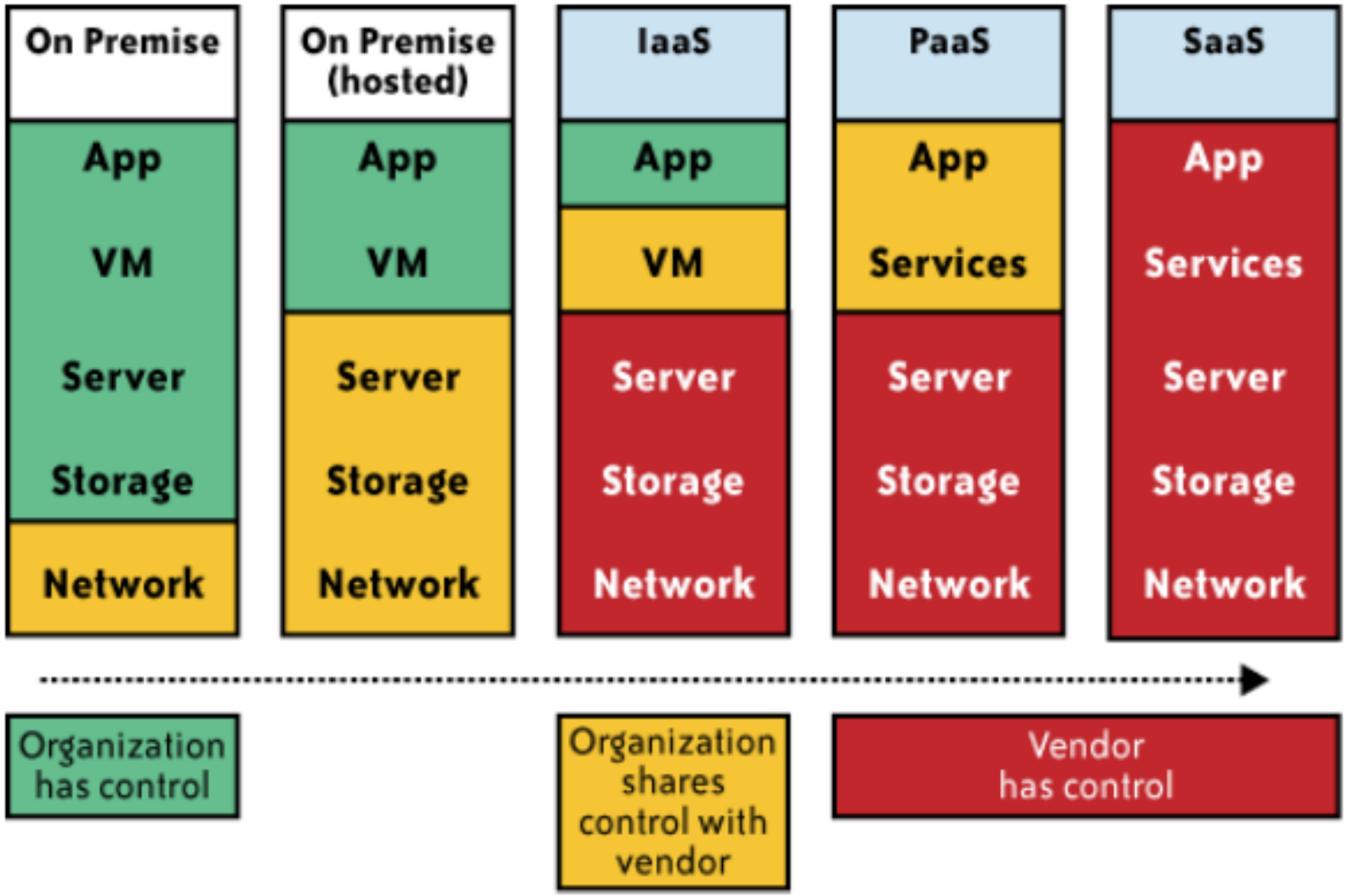
Delivery Models



While cloud-based software services are maturing,
Cloud platform and infrastructure offering are still in their early stages !

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Impact of cloud computing on the governance structure of IT organizations

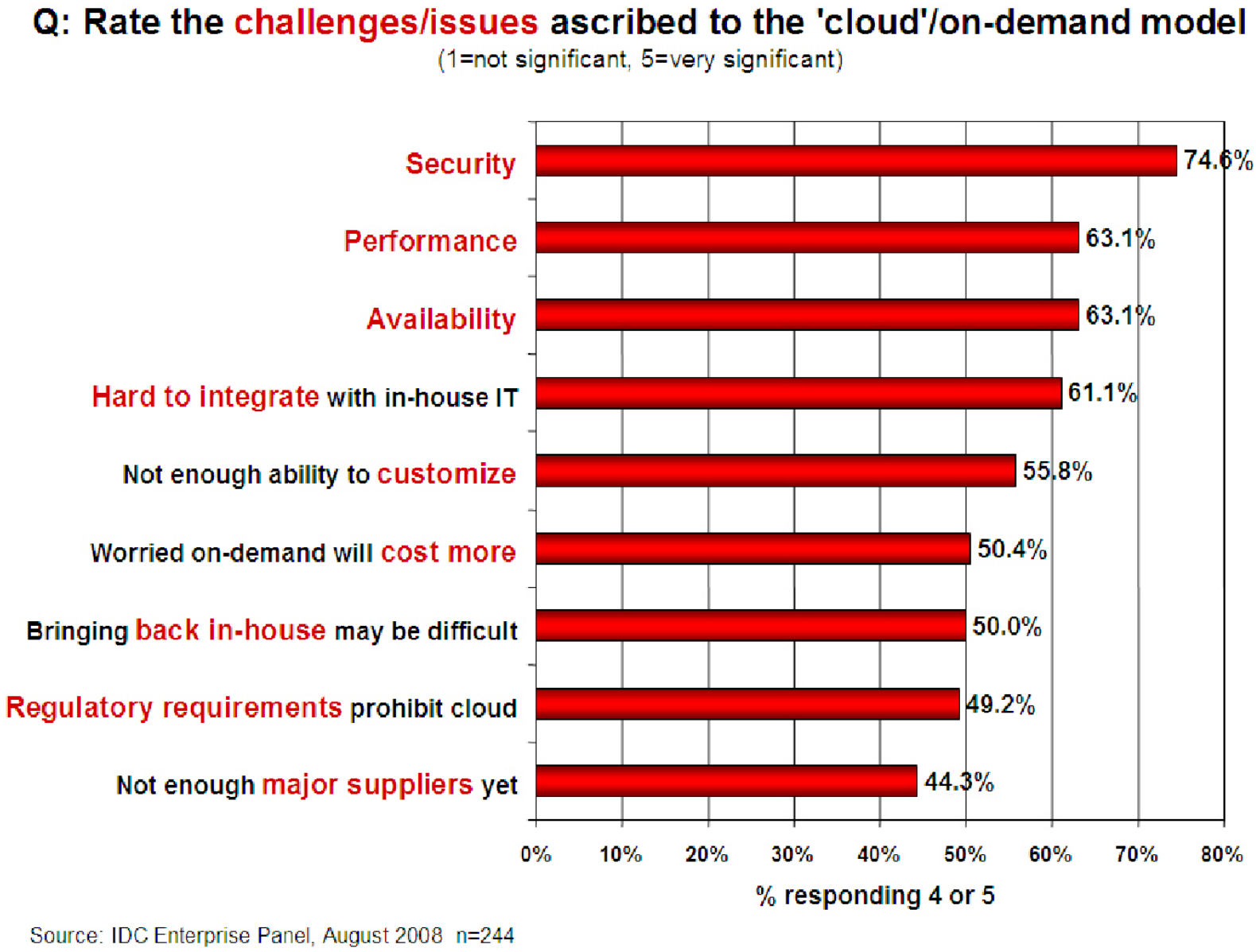


From [6] Cloud Security and Privacy by Mather and Kumaraswamy

If cloud computing is so great, why isn't everyone doing it?

- The cloud acts as a big black box, nothing inside the cloud is visible to the clients
- Clients have no idea or control over what happens inside a cloud
- Even if the cloud provider is honest, it can have malicious system admins who can tamper with the VMs and violate confidentiality and integrity
- Clouds are still subject to traditional data confidentiality, integrity, availability, and privacy issues, plus some additional attacks

Companies are still afraid to use clouds



Source: IDC Enterprise Panel, August 2008 n=244

Causes of Problems Associated with Cloud Computing

- Most security problems stem from:
  - Loss of control
  - Lack of trust (mechanisms)
  - Multi-tenancy

Loss of Control in the Cloud

- Consumer's loss of control
  - Data, applications, resources are located with provider
  - User identity management is handled by the cloud
  - User access control rules, security policies and enforcement are managed by the cloud provider
  - Consumer relies on provider to ensure
    - Data security and privacy
    - Resource availability
    - Monitoring and repairing of services/resources

Lack of Trust in the Cloud

- A brief deviation from the talk
  - (But still related)
  - Trusting a third party requires taking risks
- Defining trust and risk
  - Opposite sides of the same coin (J. Camp)
  - People only trust when it pays (Economist's view)
  - Need for trust arises only in risky situations
- Defunct third party management schemes
  - Hard to balance trust and risk
  - e.g. Key Escrow (Clipper chip)
  - Is the cloud headed toward the same path?

Multi-tenancy Issues in the Cloud

- Conflict between tenants' opposing goals
  - Tenants share a pool of resources and have opposing goals
- How does multi-tenancy deal with conflict of interest?
  - Can tenants get along together and 'play nicely' ?
  - If they can't, can we isolate them?
- How to provide separation between tenants?

- Cloud Computing brings new threats
  - Multiple independent users share the same physical infrastructure
  - Thus an attacker can legitimately be in the same physical machine as the target

Taxonomy of Fear

- Confidentiality
  - Fear of loss of control over data
    - Will the sensitive data stored on a cloud remain confidential?
    - Will cloud compromises leak confidential client data
  - Will the cloud provider itself be honest and won't peek into the data?
- Integrity
  - How do I know that the cloud provider is doing the computations correctly?
  - How do I ensure that the cloud provider really stored my data without tampering with it?

Taxonomy of Fear (cont.)

- Availability
  - Will critical systems go down at the client, if the provider is attacked in a Denial of Service attack?
  - What happens if cloud provider goes out of business?
  - Would cloud scale well-enough?
  - Often-voiced concern
    - Although cloud providers argue their downtime compares well with cloud user's own data centers

Taxonomy of Fear (cont.)

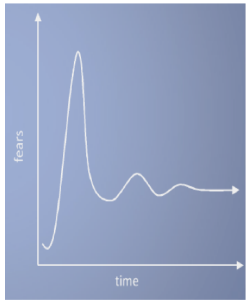
- Privacy issues raised via massive data mining
  - Cloud now stores data from a lot of clients, and can run data mining algorithms to get large amounts of information on clients
- Increased attack surface
  - Entity outside the organization now stores and computes data, and so
  - Attackers can now target the communication link between cloud provider and client
  - Cloud provider employees can be phished

Taxonomy of Fear (cont.)

- Auditability and forensics (out of control of data)
  - Difficult to audit data held outside organization in a cloud
  - Forensics also made difficult since now clients don't maintain data locally
- Legal quagmire and transitive trust issues
  - Who is responsible for complying with regulations?
    - e.g., SOX, HIPAA, GLBA ?
  - If cloud provider subcontracts to third party clouds, will the data still be secure?

Taxonomy of Fear (cont.)

Cloud Computing is **a security nightmare** and it can't be handled in traditional ways.
John Chambers
CISCO CEO



- Security is one of the most difficult task to implement in cloud computing.
  - Different forms of attacks in the application side and in the hardware components
- Attacks with catastrophic effects only needs one security flaw

(http://www.exforsys.com/tutorials/cloud-computing/cloud-computing-security.html)

Threat Model

- A threat model helps in analyzing a security problem, design mitigation strategies, and evaluate solutions
- Steps:
  - Identify attackers, assets, threats and other components
  - Rank the threats
  - Choose mitigation strategies
  - Build solutions based on the strategies

Threat Model

- Basic components
  - Attacker modeling
    - Choose what attacker to consider
      - insider vs. outsider?
      - single vs. collaborator?
    - Attacker motivation and capabilities
  - Attacker goals
  - Vulnerabilities / threats

What is the issue?

- The core issue here is the levels of trust
  - Many cloud computing providers trust their customers
  - Each customer is physically commingling its data with data from anybody else using the cloud while logically and virtually you have your own space
  - The way that the cloud provider implements security is typically focused on they fact that those outside of their cloud are evil, and those inside are good.
- But what if those inside are also evil?

Attacker Capability: Malicious Insiders

- At client
  - Learn passwords/authentication information
  - Gain control of the VMs
- At cloud provider
  - Log client communication
  - Can read unencrypted data
  - Can possibly peek into VMs, or make copies of VMs
  - Can monitor network communication, application patterns
  - Why?
    - Gain information about client data
    - Gain information on client behavior
    - Sell the information or use itself

Attacker Capability: Outside attacker

- What?
  - Listen to network traffic (passive)
  - Insert malicious traffic (active)
  - Probe cloud structure (active)
  - Launch DoS
- Goal?
  - Intrusion
  - Network analysis
  - Man in the middle
  - Cartography

Challenges for the attacker

- How to find out where the target is located?
- How to be co-located with the target in the same (physical) machine?
- How to gather information about the target?

Part II: Security and Privacy Issues in Cloud Computing - Big Picture

- Infrastructure Security
- Data Security and Storage
- Identity and Access Management (IAM)
- Privacy

- And more…

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Infrastructure Security

- Network Level
- Host Level
- Application Level

The Network Level

- Ensuring confidentiality and integrity of your organization's data-in-transit to and from your public cloud provider
- Ensuring proper access control (authentication, authorization, and auditing) to whatever resources you are using at your public cloud provider
- Ensuring availability of the Internet-facing resources in a public cloud that are being used by your organization, or have been assigned to your organization by your public cloud providers
- Replacing the established model of network zones and tiers with domains

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

The Network Level - Mitigation

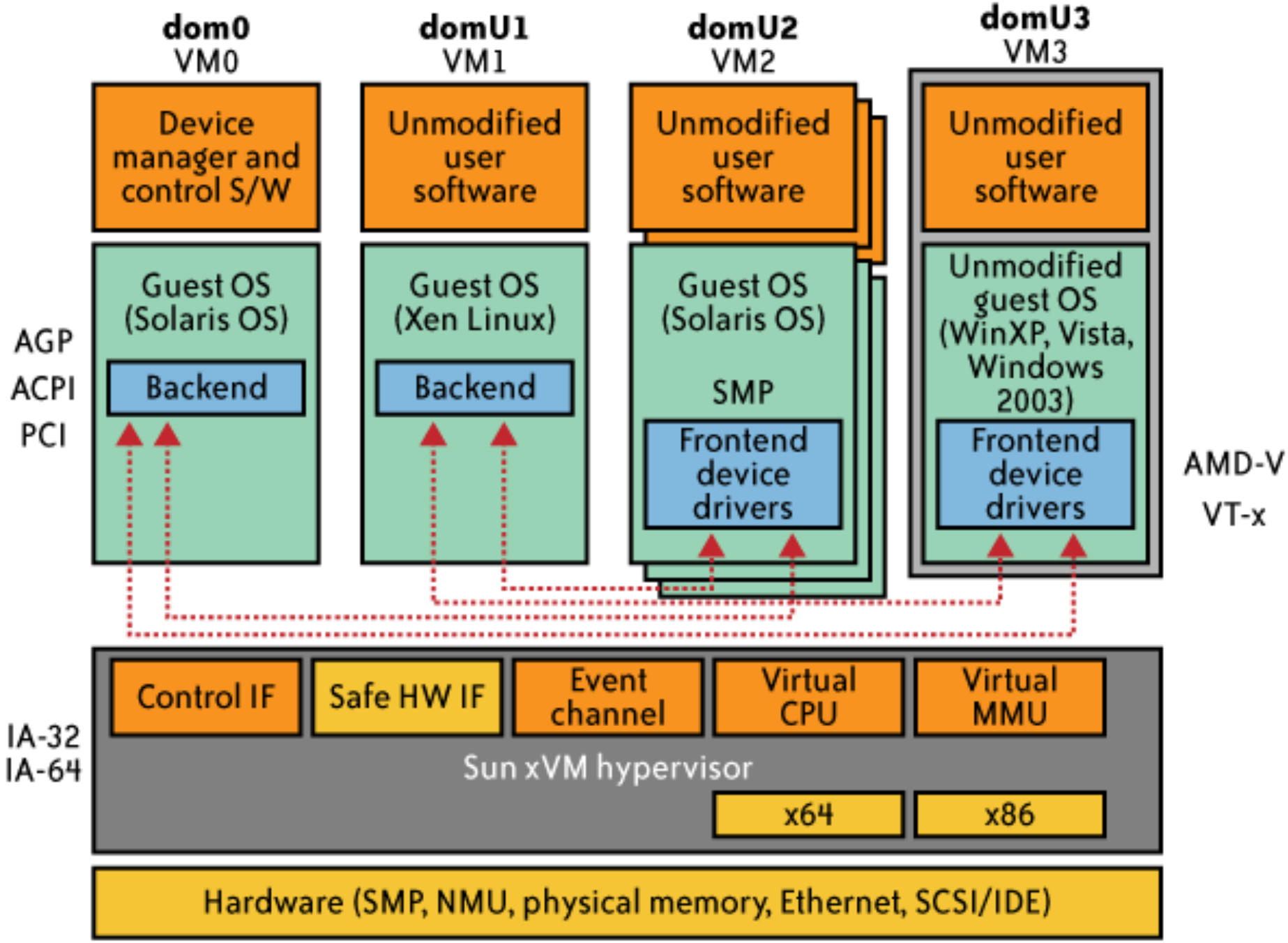
- Note that network-level risks exist regardless of what aspects of "cloud computing" services are being used
- The primary determination of risk level is therefore not which *aaS is being used,
- But rather whether your organization intends to use or is using a public, private, or hybrid cloud.

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

The Host Level

- SaaS/PaaS
  - Both the PaaS and SaaS platforms abstract and hide the host OS from end users
  - Host security responsibilities are transferred to the CSP (Cloud Service Provider)
    - You do not have to worry about protecting hosts
  - However, as a customer, you still own the risk of managing information hosted in the cloud services.

From [6] Cloud Security and Privacy by Mather and Kumaraswamy



From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Case study: Amazon's EC2 infrastructure

- "Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds"
  - Multiple VMs of different organizations with virtual boundaries separating each VM can run within one physical server
  - "virtual machines" still have internet protocol, or IP, addresses, visible to anyone within the cloud.
  - VMs located on the same physical server tend to have IP addresses that are close to each other and are assigned at the same time
  - An attacker can set up lots of his own virtual machines, look at their IP addresses, and figure out which one shares the same physical resources as an intended target
  - Once the malicious virtual machine is placed on the same server as its target, it is possible to carefully monitor how access to resources fluctuates and thereby potentially glean sensitive information about the victim

Local Host Security

- Are local host machines part of the cloud infrastructure?
  - Outside the security perimeter
  - While cloud consumers worry about the security on the cloud provider's site, they may easily forget to harden their own machines
- The lack of security of local devices can
  - Provide a way for malicious services on the cloud to attack local networks through these terminal devices
  - Compromise the cloud and its resources for other users

Local Host Security (Cont.)

- With mobile devices, the threat may be even stronger
  - Users misplace or have the device stolen from them
  - Security mechanisms on handheld gadgets are often times insufficient compared to say, a desktop computer
  - Provides a potential attacker an easy avenue into a cloud system.
  - If a user relies mainly on a mobile device to access cloud data, the threat to availability is also increased as mobile devices malfunction or are lost
- Devices that access the cloud should have
  - Strong authentication mechanisms
  - Tamper-resistant mechanisms
  - Strong isolation between applications
  - Methods to trust the OS
  - Cryptographic functionality when traffic confidentiality is required

The Application Level

- DoS
- EDoS(Economic Denial of Sustainability)
  - An attack against the billing model that underlies the cost of providing a service with the goal of bankrupting the service itself.
- End user security
- Who is responsible for Web application security in the cloud?
- SaaS/PaaS/IaaS application security
- Customer-deployed application security

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Data Security and Storage

- Several aspects of data security, including:
  - Data-in-transit
    - Confidentiality + integrity using secured protocol
    - Confidentiality with non-secured protocol and encryption
  - Data-at-rest
    - Generally, not encrypted , since data is commingled with other users' data
    - Encryption if it is not associated with applications?
      - But how about indexing and searching?
      - Then homomorphic encryption vs. predicate encryption?
  - Processing of data, including multitenancy
    - For any application to process data, not encrypted

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Data Security and Storage (cont.)

- Data lineage
  - Knowing when [illegible] cloud is important for [illegible]
  - e.g., Amazon AWS
    - Store <d1, t1, ex1.s3.amazonaws.com>
    - Process <d2, t2, ec2.compute2.amazonaws.com>
    - Restore <d3, t3, ex2.s3.amazonaws.com>

Where is (or was) that system located?
What was the state of that physical system?
How would a customer or auditor verify that info?

- Data provenance
  - Computational accuracy (as well as data integrity)
  - E.g., financial calculation: sum ((((2*3)*4)/6) -2) = $2.00 ?
    - Correct : assuming US dollar
    - How about dollars of different countries?
    - Correct exchange rate?

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Data Security and Storage

- Data remanence
  - Inadvertent disclosure of sensitive information is possible
- Data security mitigation?
  - Do not place any sensitive data in a public cloud
  - Encrypted data is placed into the cloud?
- Provider data and its security: storage
  - To the extent that quantities of data from many companies are centralized, this collection can become an attractive target for criminals
  - Moreover, the physical security of the data center and the trustworthiness of system administrators take on new importance.

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Why?

- Organization's trust boundary will become dynamic and will move beyond the control and will extend into the service provider domain.
- Managing access for diverse user populations (employees, contractors, partners, etc.)
- Increased demand for authentication
  - personal, financial, medical data will now be hosted in the cloud
  - S/W applications hosted in the cloud requires access control
- Need for higher-assurance authentication
  - authentication in the cloud may mean authentication outside F/W
  - Limits of password authentication
- Need for authentication from mobile devices

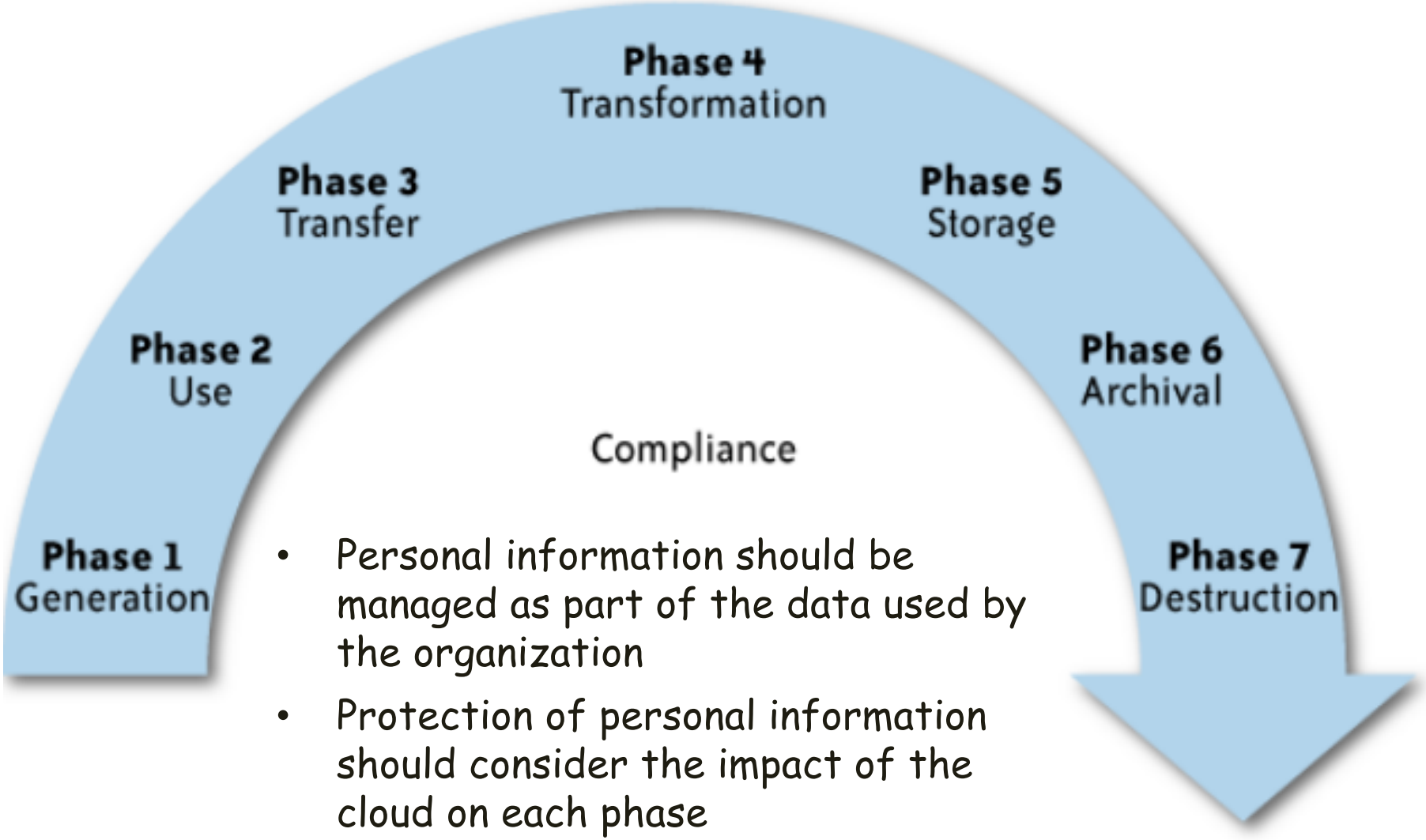
From [6] Cloud Security and Privacy by Mather and Kumaraswamy

What is Privacy?

- The concept of privacy varies widely among (and sometimes within) countries, cultures, and jurisdictions.
- It is shaped by public expectations and legal interpretations; as such, a concise definition is elusive if not impossible.
- Privacy rights or obligations are related to the collection, use, disclosure, storage, and destruction of personal data (or Personally Identifiable Information—PII).
- At the end of the day, privacy is about the accountability of organizations to data subjects, as well as the transparency to an organization's practice around personal information.

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

What is the data life cycle?



- Personal information should be managed as part of the data used by the organization
- Protection of personal information should consider the impact of the cloud on each phase

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

What Are the Key Privacy Concerns?

- Typically mix security and privacy
- Some considerations to be aware of:
  - Storage
  - Retention
  - Destruction
  - Auditing, monitoring and risk management
  - Privacy breaches
  - Who is responsible for protecting privacy?

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Storage

- Is it commingled with information from other organizations that use the same CSP?
- The aggregation of data raises new privacy issues
  - Some governments may decide to search through data without necessarily notifying the data owner, depending on where the data resides
- Whether the cloud provider itself has any right to see and access customer data?
- Some services today track user behaviour for a range of purposes, from sending targeted advertising to improving services

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Retention

- How long is personal information (that is transferred to the cloud) retained?
- Which retention policy governs the data?
- Does the organization own the data, or the CSP?
- Who enforces the retention policy in the cloud, and how are exceptions to this policy (such as litigation holds) managed?

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Destruction

- How does the cloud provider destroy PII at the end of the retention period?
- How do organizations ensure that their PII is destroyed by the CSP at the right point and is not available to other cloud users?
- Cloud storage providers usually replicate the data across multiple systems and sites—increased availability is one of the benefits they provide.
  - How do you know that the CSP didn't retain additional copies?
  - Did the CSP really destroy the data, or just make it inaccessible to the organization?
  - Is the CSP keeping the information longer than necessary so that it can mine the data for its own use?

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Auditing, monitoring and risk management

- How can organizations monitor their CSP and provide assurance to relevant stakeholders that privacy requirements are met when their PII is in the cloud?
- Are they regularly audited?
- What happens in the event of an incident?
- If business-critical processes are migrated to a cloud computing model, internal security processes need to evolve to allow multiple cloud providers to participate in those processes, as needed.
  - These include processes such as security monitoring, auditing, forensics, incident response, and business continuity

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Privacy breaches

- How do you know that a breach has occurred?
- How do you ensure that the CSP notifies you when a breach occurs?
- Who is responsible for managing the breach notification process (and costs associated with the process)?
- If contracts include liability for breaches resulting from negligence of the CSP?
  - How is the contract enforced?
  - How is it determined who is at fault?

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Who is responsible for protecting privacy?

- Do
-
-
stew
- Organizations can trans[illegible]ility, but not accountability
- Risk assessment and mitigation throughout the data life cycle is cr[illegible]cal.
- Many new risks and unknowns
  - The overall complexity of privacy protection in the cloud represents a bigger challenge.

e.g., Suppose a hacker breaks into Cloud Provider A and steals data from Company X. Assume that the compromised server also contained data from Companies Y and Z.

- **Who investigates this crime?**
- **Is it the Cloud Provider, even though Company X may fear that the provider will try to absolve itself from responsibility?**
- **Is it Company X and, if so, does it have the right to see other data on that server, including logs that may show access to the data of Companies Y and Z?**

From [6] Cloud Security and Privacy by Mather and Kumaraswamy

Part III. Possible Solutions

- Minimize Lack of Trust
  - Policy Language
  - Certification
- Minimize Loss of Control
  - Monitoring
  - Utilizing different clouds
  - Access control management
  - Identity Management (IDM)
- Minimize Multi-tenancy

Security Issues in the Cloud

- In theory, minimizing any of the issues would help:
  - Third Party Cloud Computing
  - Loss of Control
    - Take back control
      - Data and apps may still need to be on the cloud
      - But can they be managed in some way by the consumer?
  - Lack of trust
    - Increase trust (mechanisms)
      - Technology
      - Policy, regulation
      - Contracts (incentives): topic of a future talk
  - Multi-tenancy
    - Private cloud
      - Takes away the reasons to use a cloud in the first place
    - VPC: its still not a separate system
    - Strong separation

Third Party Cloud Computing

Like Amazon's EC2, Microsoft's Azure

- Allow users to instantiate Virtual Machines
- Allow users to purchase required quantity when required
- Allow service providers to maximize the utilization of sunk capital costs
- Confidentiality is very important

Known issues: Already exist

- Confidentiality issues
- Malicious behavior by cloud provider
- Known risks exist in any industry practicing outsourcing
- Provider and its infrastructure needs to be trusted

New Vulnerabilities & Attacks

- Threats arise from other consumers
- Due to the subtleties of how physical resources can be transparently shared between VMs
- Such attacks are based on placement and extraction
- A customer VM and its adversary can be assigned to the same physical server
- Adversary can penetrate the VM and violate customer confidentiality

More on attacks…

- Collaborative attacks
- Mapping of internal cloud infrastructure
- Identifying likely residence of a target VM
- Instantiating new VMs until one gets co-resident with the target
- Cross-VM side-channel attacks
- Extract information from target VM on the same machine

More on attacks…

- Can one determine where in the cloud infrastructure an instance is located?
- Can one easily determine if two instances are co-resident on the same physical machine?
- Can an adversary launch instances that will be co-resident with other user instances?
- Can an adversary exploit cross-VM information leakage once co-resident?

Answer: Yes to all

Minimize Lack of Trust

- **POLICY LANGUAGE**
- **CERTIFICATION**

Minimize Lack of Trust: Policy Language

- Consumers have specific security needs but don't have a say-so in how they are handled
  - What the heck is the provider doing for me?
  - Currently consumers cannot dictate their requirements to the provider (one-sided)
- Standard language to convey one's policies and expectations
  - Agreed upon and upheld by both parties
  - Standard language for representing SLAs
  - Can be used in a intra-cloud environment to realize overarching security posture

Minimize Lack of Trust: Policy Language (Cont.)

- Create policy language with the following characteristics:
  - Machine-understandable (or at least processable),
  - Easy to combine/merge and compare
  - Examples of policy statements are, "requires isolation between VMs", "requires geographical isolation between VMs", "requires physical separation between other communities/tenants that are in the same industry," etc.
  - Need a validation tool to check that the policy created in the standard language correctly reflects the policy creator's intentions (i.e. that the policy language is semantically equivalent to the user's intentions).

Minimize Lack of Trust: Certification

- Certification
  - Some form of reputable, independent, comparable assessment and description of security features and assurance
  - Sarbanes-Oxley, DIACAP, DISTCAP, etc (are they sufficient for a cloud environment?)
- Risk assessment
  - Performed by certified third parties
  - Provides consumers with additional assurance

Minimize Loss of Control

- MONITORING
- UTILIZING DIFFERENT CLOUDS
- ACCESS CONTROL MANAGEMENT
- IDENTITY MANAGEMENT (IDM)

Minimize Loss of Control: Monitoring

- Cloud consumer needs situational awareness for critical applications
  - When underlying components fail, what is the effect of the failure to the mission logic
  - What recovery measures can be taken (by provider and consumer)
- Requires an application-specific run-time monitoring and management tool for the consumer
  - The cloud consumer and cloud provider have different views of the system
  - Enable both the provider and tenants to monitor the components in the cloud that are under their control

Minimize Loss of Control: Monitoring (Cont.)

- Provide mechanisms that enable the provider to act on attacks he can handle.
  - infrastructure remapping (create new or move existing fault domains)
  - shutting down offending components or targets (and assisting tenants with porting if necessary
  - Repairs
- Provide mechanisms that enable the consumer to act on attacks that he can handle (application-level monitoring).
  - RAdAC (Risk-adaptable Access Control)
  - VM porting with remote attestation of target physical host
  - Provide ability to move the user's application to another cloud

Minimize Loss of Control: Utilize Different Clouds

- The concept of 'Don't put all your eggs in one basket'
  - Consumer may use services from different clouds through an intra-cloud or multi-cloud architecture
  - Propose a multi-cloud or intra-cloud architecture in which consumers
    - Spread the risk
    - Increase redundancy (per-task or per-application)
    - Increase chance of mission completion for critical applications
  - Possible issues to consider:
    - Policy incompatibility (combined, what is the overarching policy?)
    - Data dependency between clouds
    - Differing data semantics across clouds
    - Knowing when to utilize the redundancy feature (monitoring technology)
    - Is it worth it to spread your sensitive data across multiple clouds?
      - Redundancy could increase risk of exposure

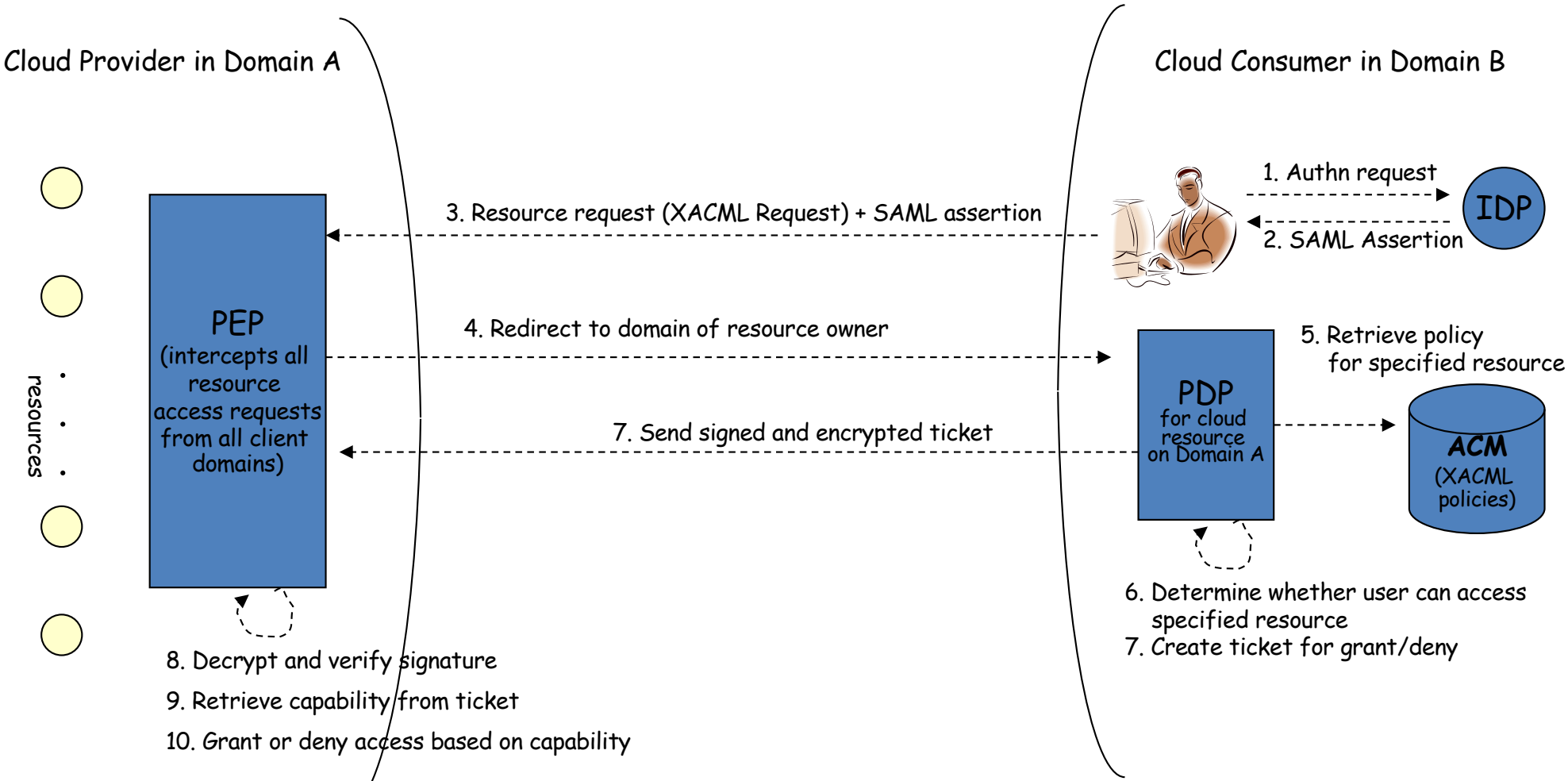
Minimize Loss of Control: Access Control

- Many possible layers of access control
  - E.g. access to the cloud, access to servers, access to services, access to databases (direct and queries via web services), access to VMs, and access to objects within a VM
  - Depending on the deployment model used, some of these will be controlled by the provider and others by the consumer
- Regardless of deployment model, provider needs to manage the user authentication and access control procedures (to the cloud)
  - Federated Identity Management: access control management burden still lies with the provider
  - Requires user to place a large amount of trust on the provider in terms of security, management, and maintenance of access control policies. This can be burdensome when numerous users from different organizations with different access control policies, are involved

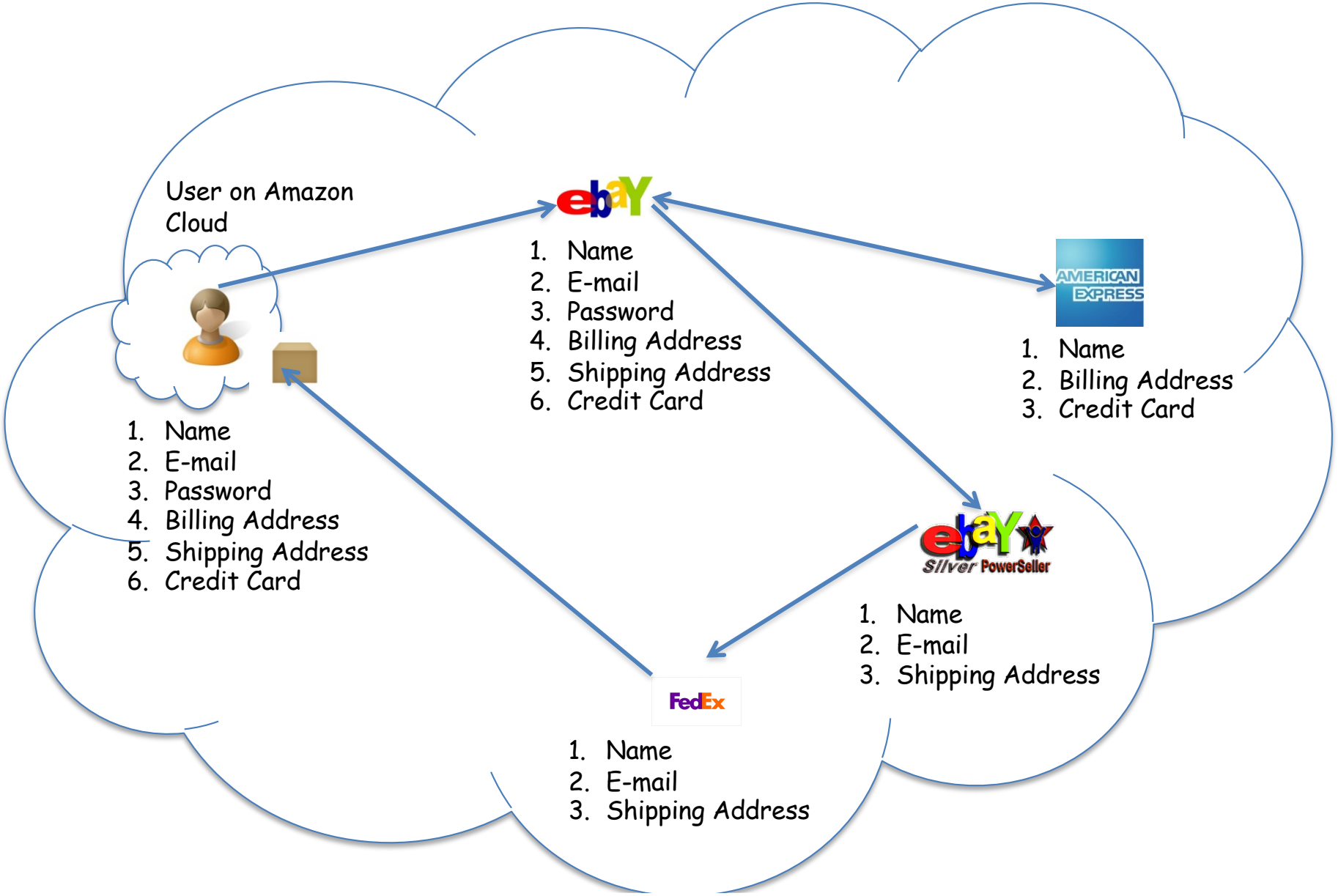
Minimize Loss of Control: Access Control (Cont.)

- Consumer-managed access control
  - Consumer retains decision-making process to retain some control, requiring less trust of the provider (i.e. PDP is in consumer's domain)
  - Requires the client and provider to have a pre-existing trust relationship, as well as a pre-negotiated standard way of describing resources, users, and access decisions between the cloud provider and consumer. It also needs to be able to guarantee that the provider will uphold the consumer-side's access decisions.
  - Should be at least as secure as the traditional access control model.
  - Facebook and Google Apps do this to some degree, but not enough control
  - Applicability to privacy of patient health records

Minimize Loss of Control: Access Control



Minimize Loss of Control: IDM Motivation



Minimize Loss of Control: IDM Identity in the Cloud

Minimize Loss of Control: IDM
Present IDMs

- IDM in traditional application-centric IDM model
  - Each application keeps track of identifying information of its users.
- Existing IDM Systems
  - Microsoft Windows CardSpace [W. A. Alrodhan]
  - OpenID [http://openid.net]
  - PRIME [S. F. Hubner]

These systems require a **trusted third party** and
do not work on an **untrusted host**.

If Trusted Third Party is compromised, all the identifying information of the users is also compromised

[Latest: **AT&T iPad leak**]

Minimize Loss of Control: IDM Issues in Cloud Computing

- Cloud introduces several issues to IDM
  - Users have **multiple accounts** associated with **multiple service providers.**
  - Lack of trust
    - Use of Trusted Third Party is not an option
    - Cloud hosts are untrusted
  - Loss of control
    - Collusion between Cloud Services
      - Sharing sensitive identity information between services can lead to undesirable **mapping of the identities to the user.**

IDM in Cloud needs to be user-centric

Minimize Loss of Control: IDM
Goals of Proposed User-Centric IDM for the Cloud

1. Authenticate without disclosing identifying information
2. Ability to securely use a service while on an untrusted host (VM on the cloud)
3. Minimal disclosure and minimized risk of disclosure during communication between user and service provider (Man in the Middle, Side Channel and Correlation Attacks)
4. Independence of Trusted Third Party

Minimize Loss of Control: IDM Approach - 1

- **IDM Wallet:**
  - Use of AB scheme to protect PII from untrusted hosts.
- **Anonymous Identification:**
  - Use of Zero-knowledge proofing for authentication of an entity without disclosing its identifier.

Minimize Loss of Control: IDM Components of Active Bundle (Approach – 1)

- **Identity data:** Data used during authentication, getting service, using service (i.e. SSN, Date of Birth).
- **Disclosure policy:** A set of rules for choosing Identity data from a set of identities in IDM Wallet.
- **Disclosure history:** Used for logging and auditing purposes.
- **Negotiation policy:** This is Anonymous Identification, based on the Zero Knowledge Proofing.
- **Virtual Machine:** Code for protecting data on untrusted hosts. It enforces the disclosure policies.

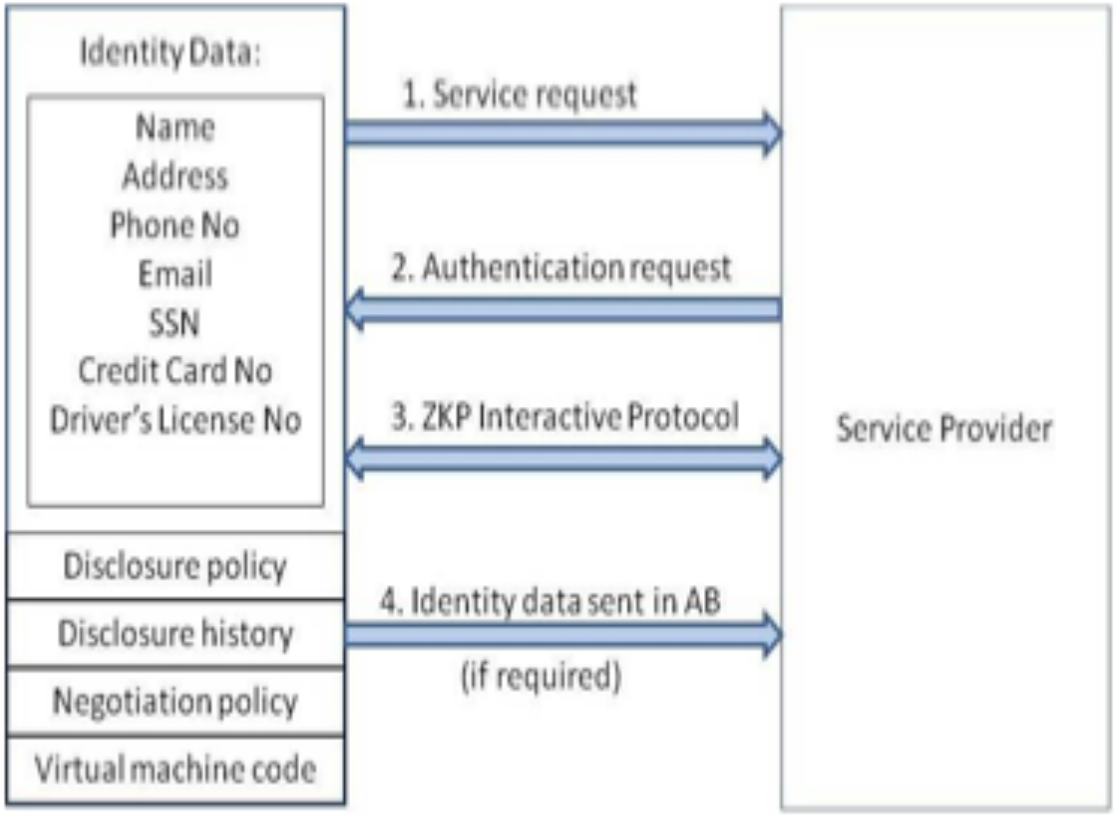
Minimize Loss of Control: IDM
Anonymous Identification (Approach – 1)

Anonymous Identification

(Shamir's approach for Credit Cards)

- IdP provides Encrypted Identity Information to the user and SP.
- SP and User interact
- Both run IdP's public function on the certain bits of the Encrypted data.
- Both exchange results and agree if it matches.

Minimize Loss of Control: IDM Usage Scenario (Approach – 1)



Minimize Loss of Control: IDM Approach - 2

- Active Bundle scheme to protect PII from untrusted hosts
- Predicates over encrypted data to authenticate without disclosing unencrypted identity data.
- Multi-party computing to be independent of a trusted third party

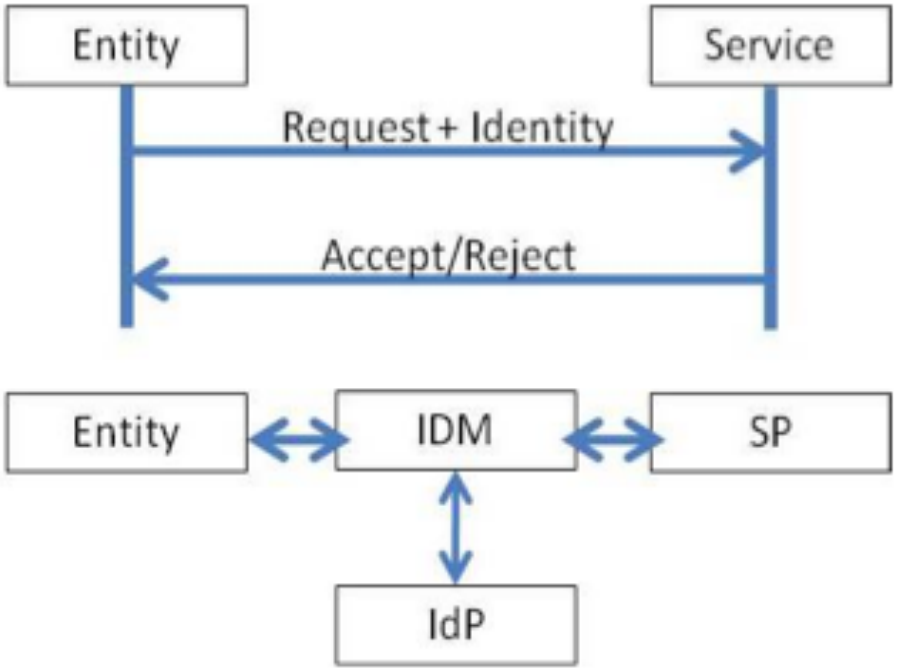
Minimize Loss of Control: IDM Usage Scenario (Approach – 2)

- Owner O encrypts Identity Data(PII) using algorithm Encrypt and O's public key PK. Encrypt outputs CT—the encrypted PII.
- SP transforms his request for PII to a predicate represented by function p.
- SP sends shares of p to the n parties who hold the shares of MSK.
- n parties execute together KeyGen using PK, MSK, and p, and return TKp to SP.
- SP calls the algorithm Query that takes as input PK, CT, TKp and produces p(PII) which is the evaluation of the predicate.
- The owner O is allowed to use the service only when the predicate evaluates to "true".

Minimize Loss of Control: IDM
Representation of identity information for negotiation

- Token/Pseudonym
- Identity Information in clear plain text
- **Active Bundle**

Minimize Loss of Control: IDM Motivation-Authentication Process using PII



Problem: Which information to disclose and how to disclose it.

Proposed IDM: Mechanisms

- [16] *Protection of Identity Information in Cloud Computing without Trusted Third Party* - R. Ranchal, B. Bhargava, L.B. Othmane, L. Lilien, A. Kim, M. Kang, Third International Workshop on Dependable Network Computing and Mobile Systems (DNCMS) in conjunction with 29th IEEE Symposium on Reliable Distributed System (SRDS) 2010
- [17] *A User-Centric Approach for Privacy and Identity Management in Cloud Computing* - P. Angin, B. Bhargava, R. Ranchal, N. Singh, L. Lilien, L.B. Othmane 29th IEEE Symposium on Reliable Distributed System (SRDS) 2010
- *Privacy in Cloud Computing Through Identity Management* - B. Bhargava, N. Singh, A. Sinclair, International Conference on Advances in Computing and Communication ICACC-11, April, 2011, India.

- Active Bundle
- Anonymous Identification
- Computing Predicates with encrypted data
- Multi-Party Computing
- Selective Disclosure

Proposed IDM: Active Bundle

- Active bundle (AB)
  - An encapsulating mechanism protecting data carried within it
  - Includes data
  - Includes metadata used for managing confidentiality
    - Both privacy of data and privacy of the whole AB
  - Includes Virtual Machine (VM)
    - performing a set of operations
    - protecting its confidentiality

Proposed IDM: Active Bundle (Cont.)

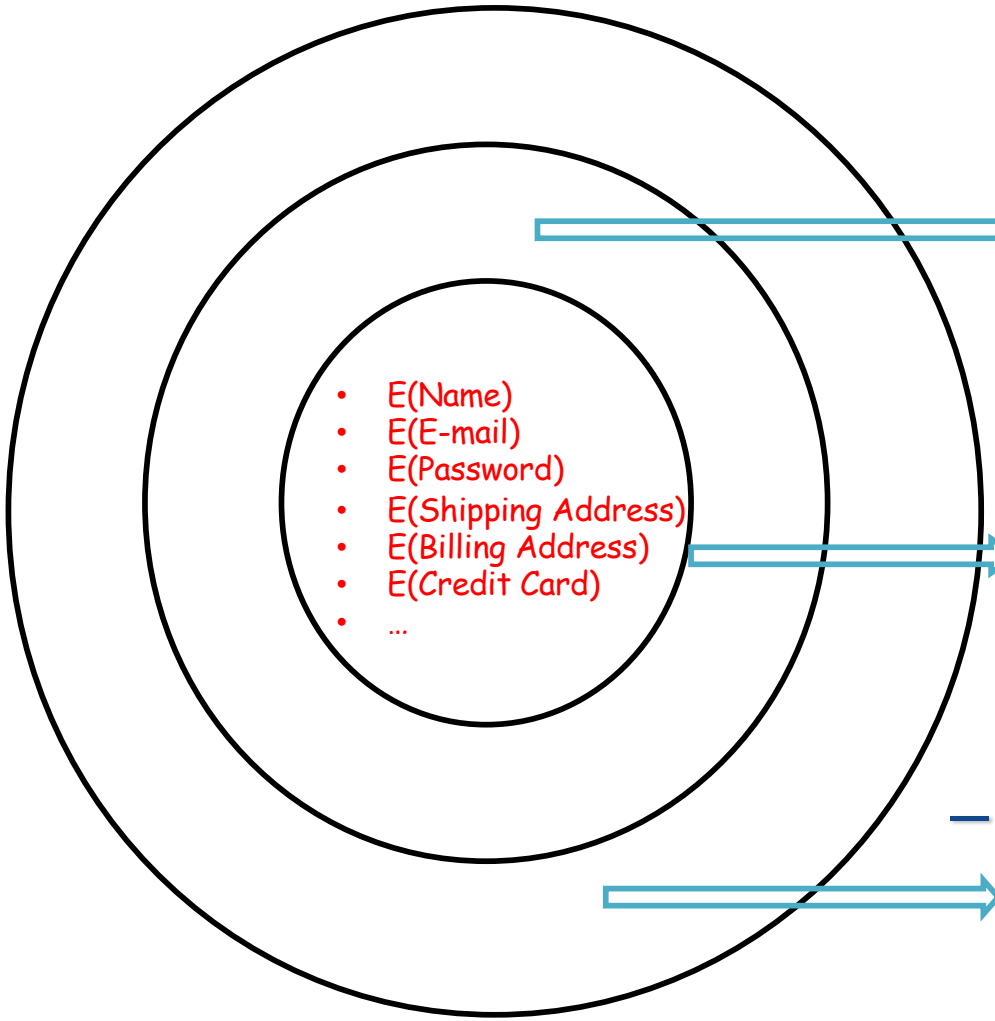
- Active Bundles—Operations
  - Self-Integrity check

    E.g., Uses a hash function
  - Evaporation/ Filtering

    Self-destroys (a part of) AB's sensitive data when threatened with a disclosure
  - Apoptosis

    Self-destructs AB's completely

Proposed IDM: Active Bundle Scheme

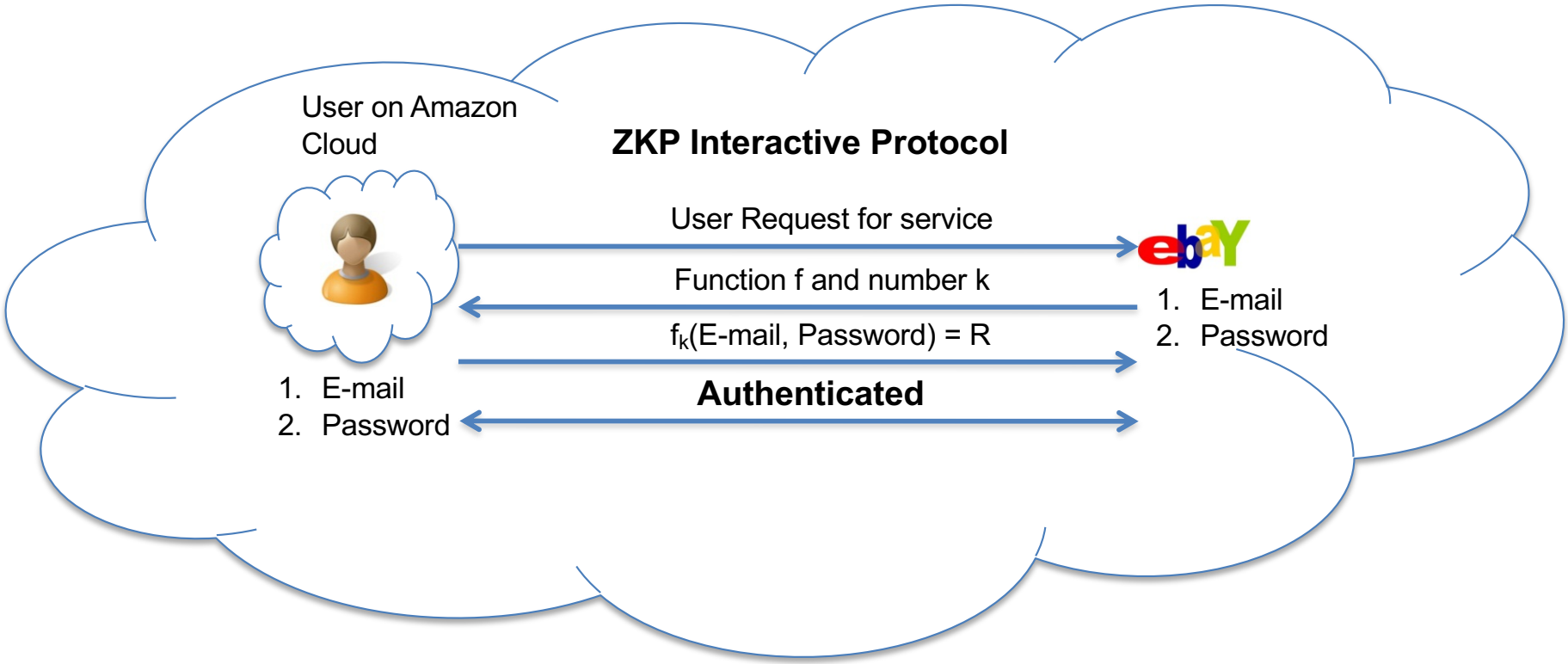


- **Metadata:**
  - Access control policies
  - Data integrity checks
  - Dissemination policies
  - Life duration
  - ID of a trust server
  - ID of a security server
  - App-dependent information
  - …
- **Sensitive Data:**
  - Identity Information
  - …
- **Virtual Machine (algorithm):**
  - Interprets metadata
  - Checks active bundle integrity
  - Enforces access and dissemination control policies
  - …

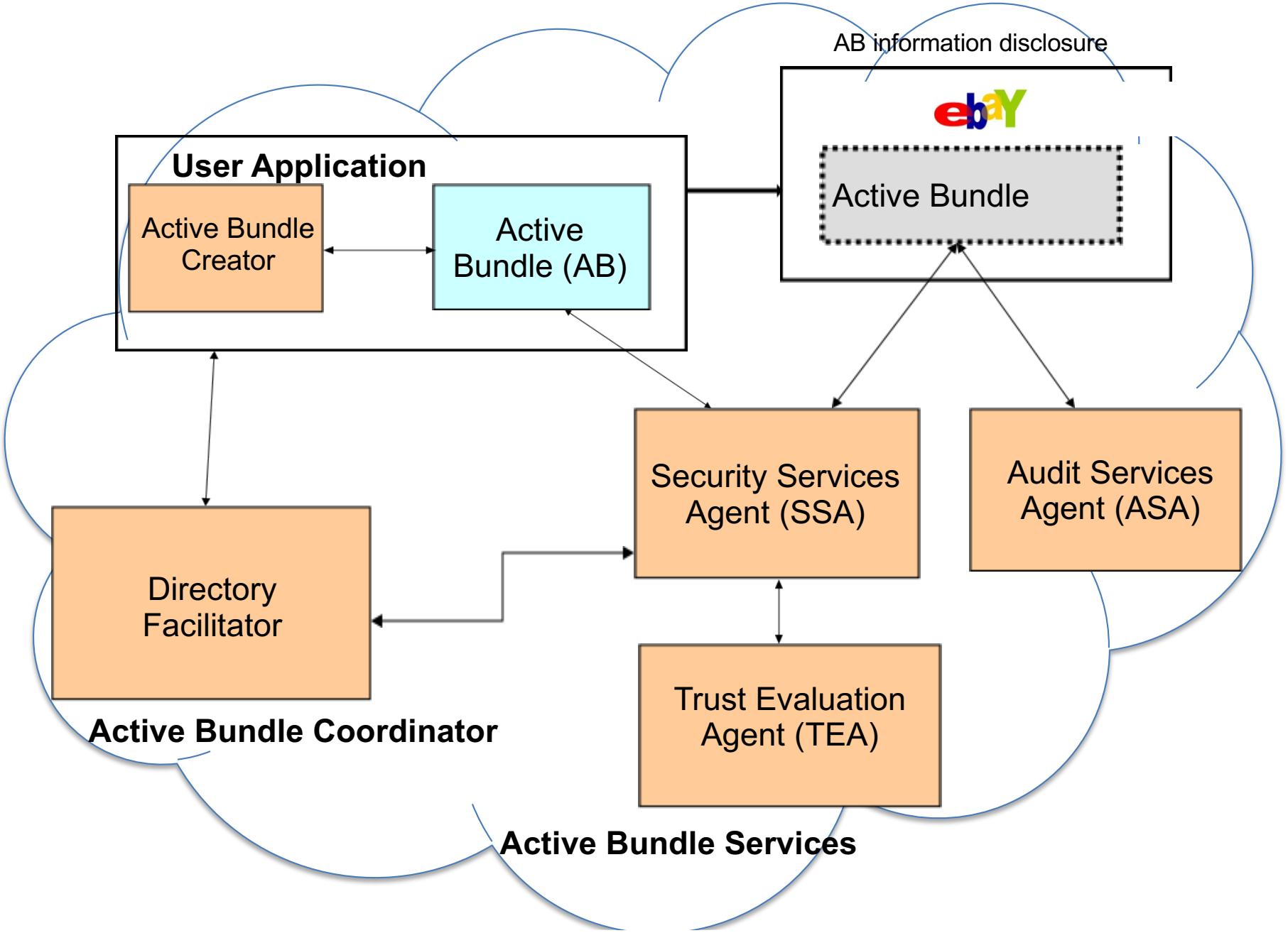
* E( ) - Encrypted Information

Proposed IDM: Anonymous Identification

- Use of Zero-knowledge proofing for user authentication without disclosing its identifier.

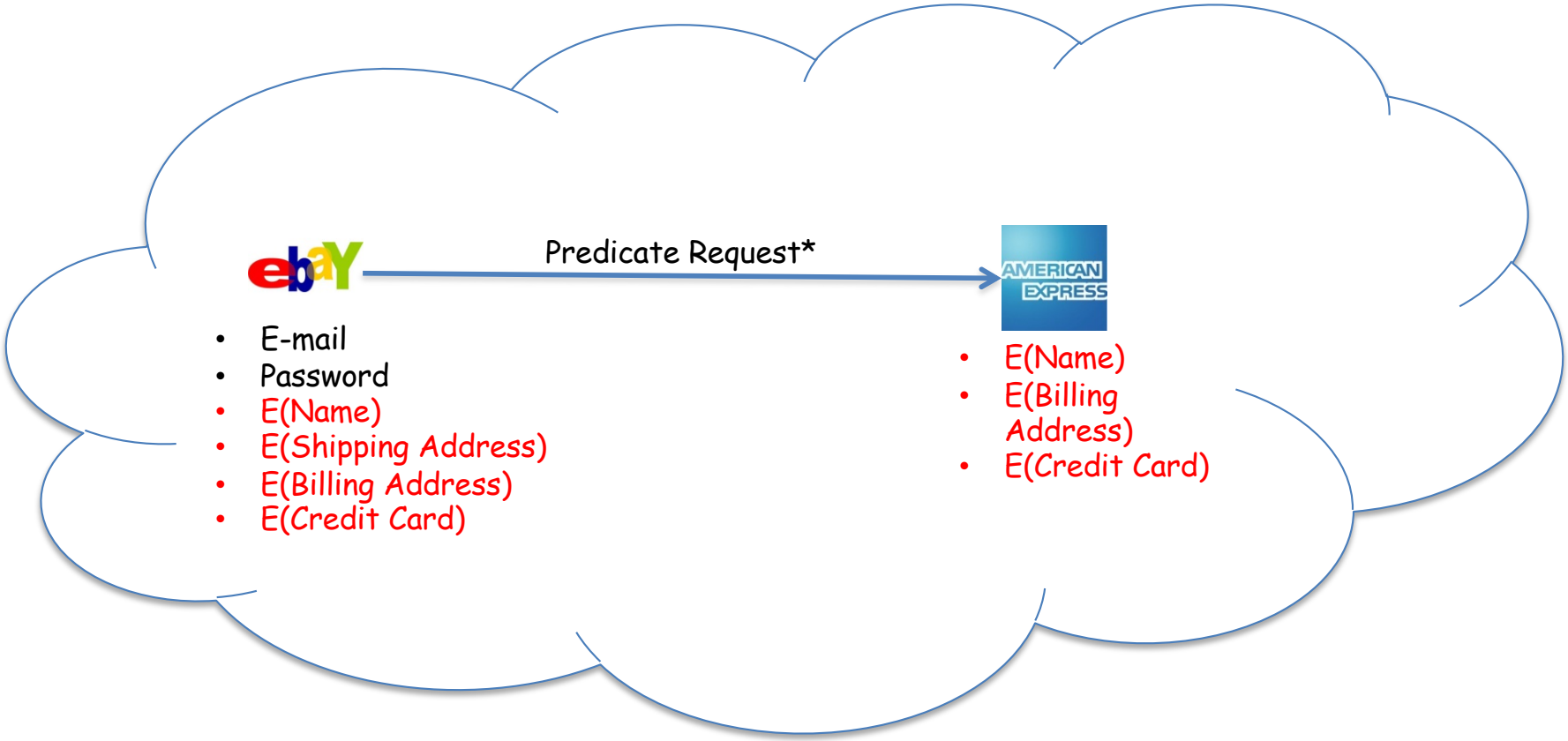


Proposed IDM: Interaction using Active Bundle



Proposed IDM: Predicate over Encrypted Data

- Verification without disclosing unencrypted identity data.

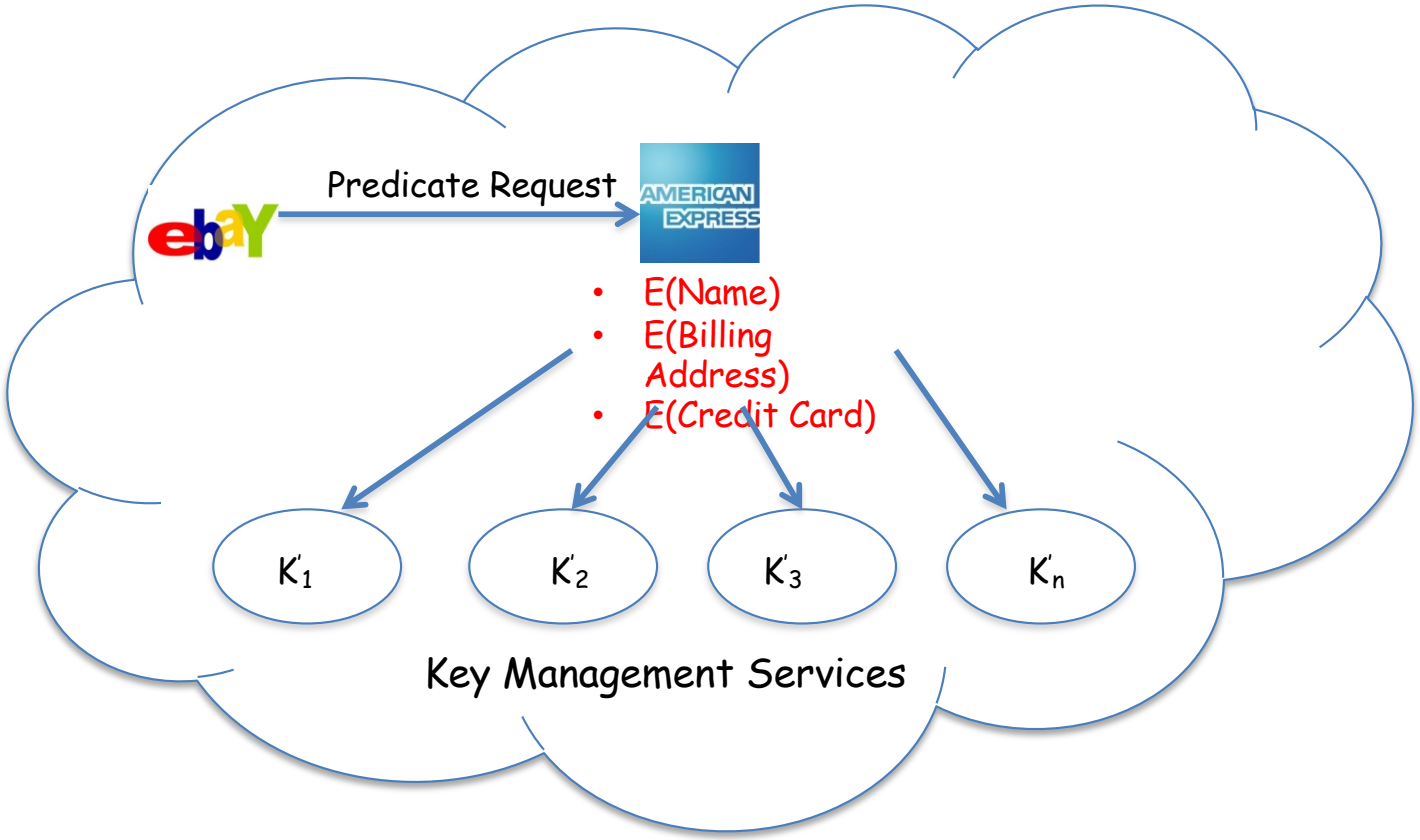


*Age Verification Request

*Credit Card Verification Request

Proposed IDM: Multi-Party Computing

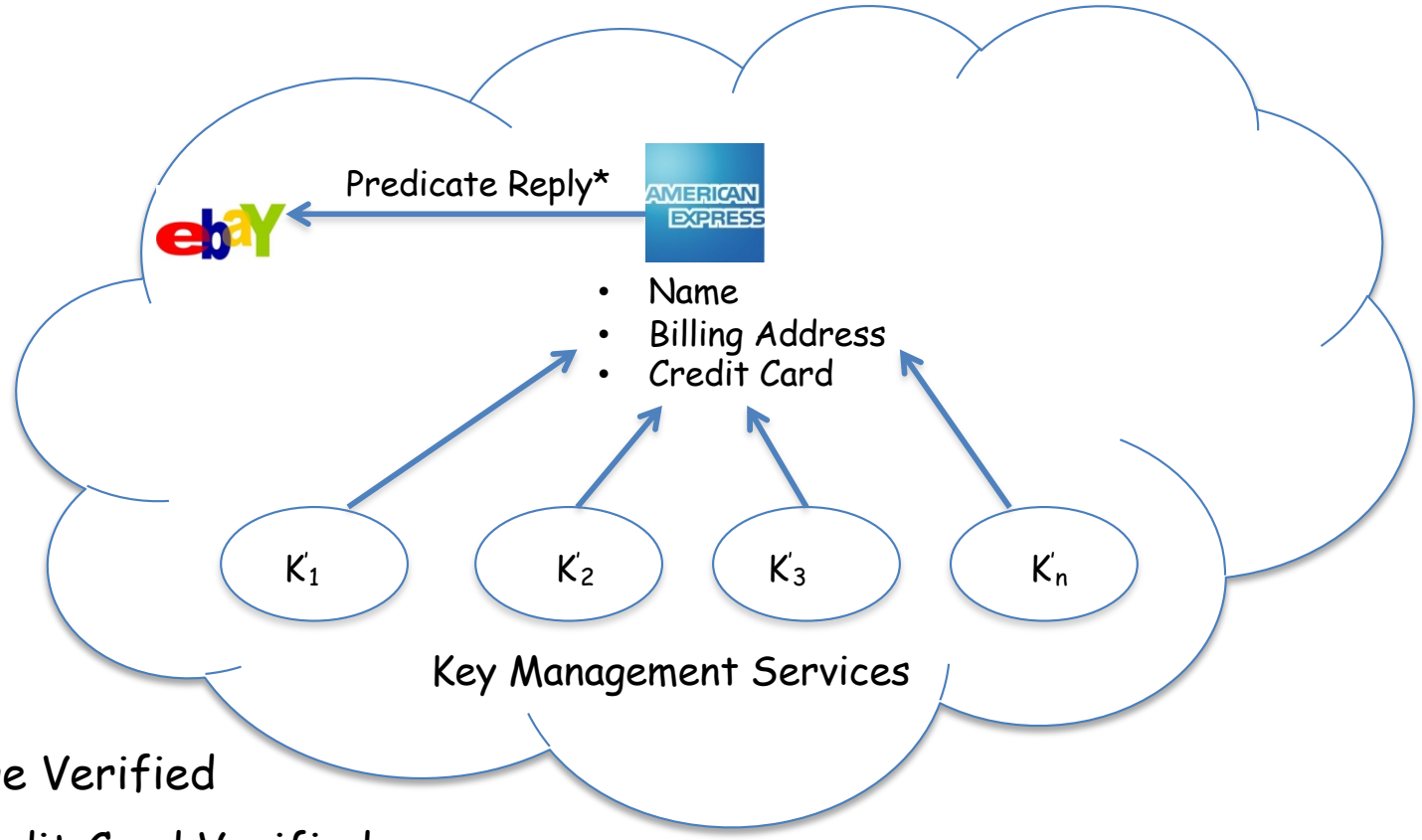
- To become independent of a trusted third party
  - Multiple Services hold shares of the secret key
  - Minimize the risk



* Decryption of information is handled by the Key Management services

Proposed IDM: Multi-Party Computing

- To become independent of a trusted third party
  - Multiple Services hold shares of the secret key
  - Minimize the risk

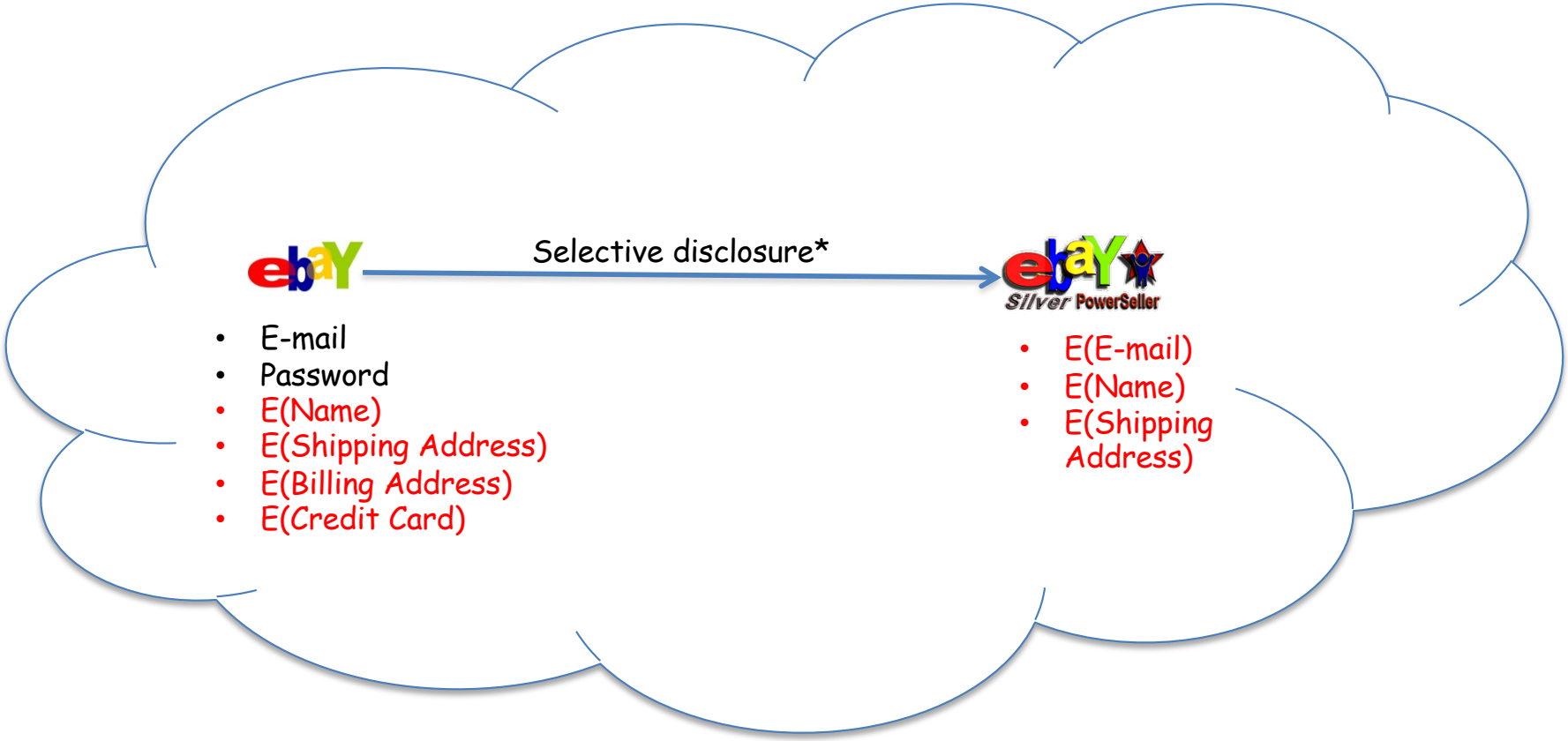


*Age Verified

*Credit Card Verified

Proposed IDM: Selective Disclosure

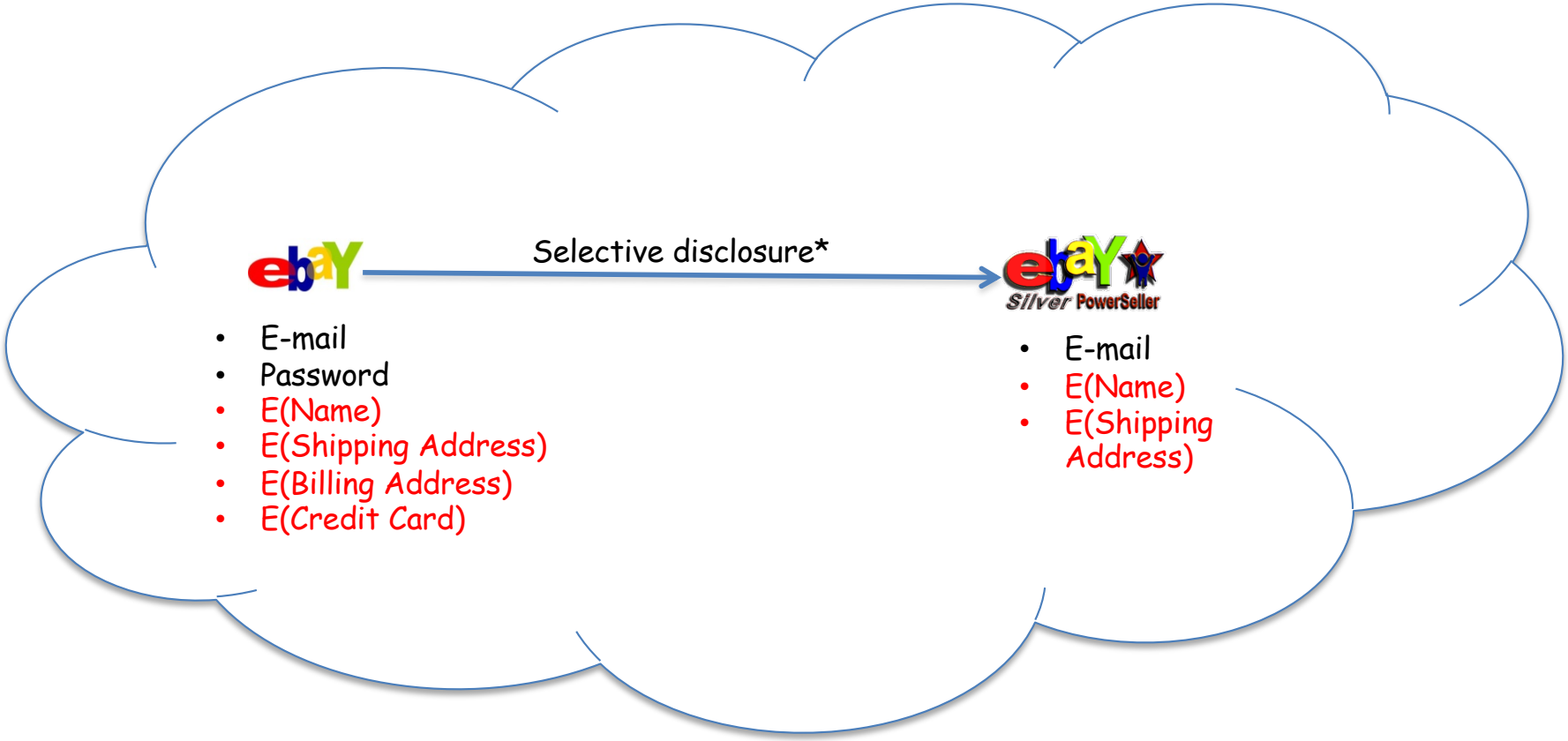
- User Policies in the Active Bundle dictate dissemination



*e-bay shares the encrypted information based on the user policy

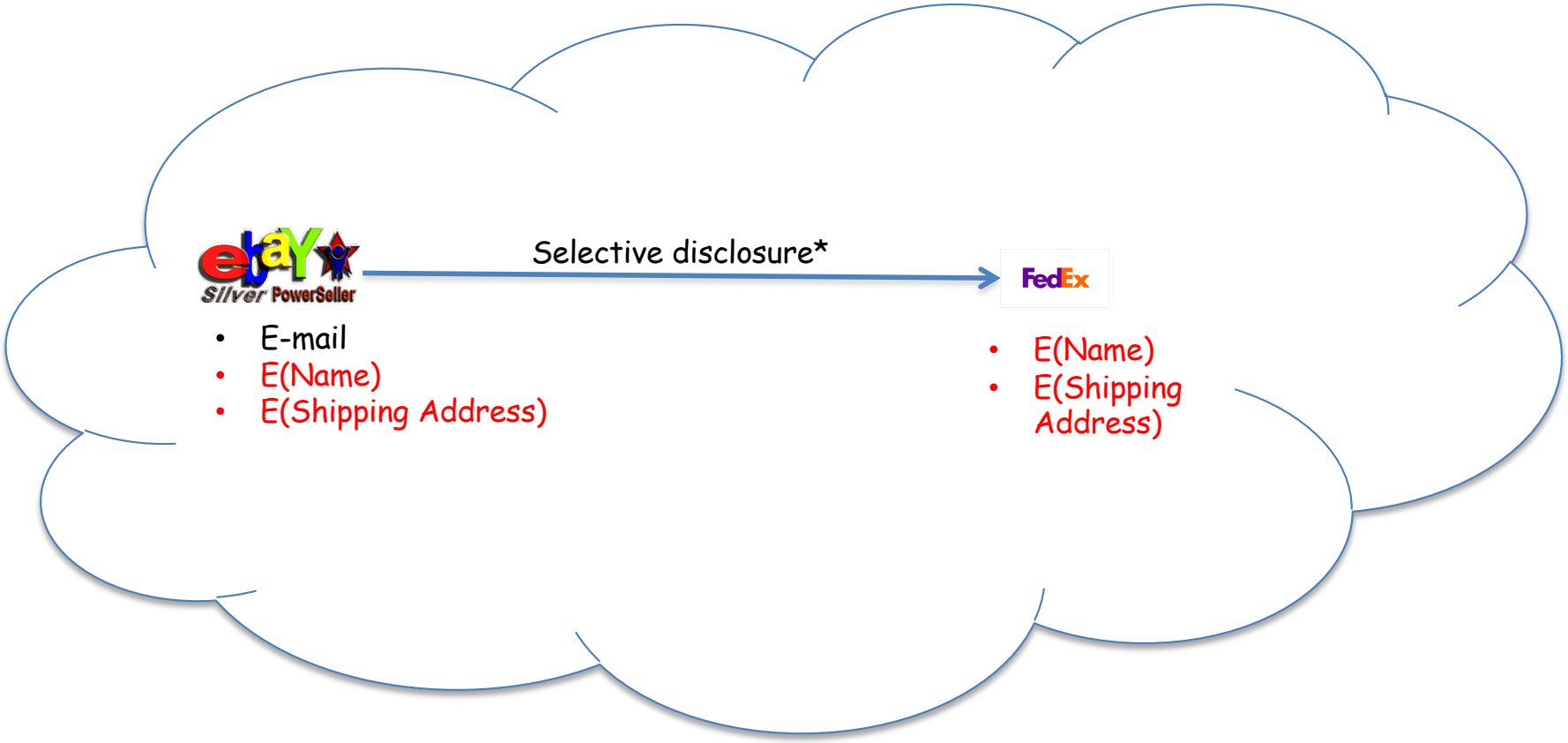
Proposed IDM: Selective Disclosure

- User Policies in the Active Bundle dictate dissemination



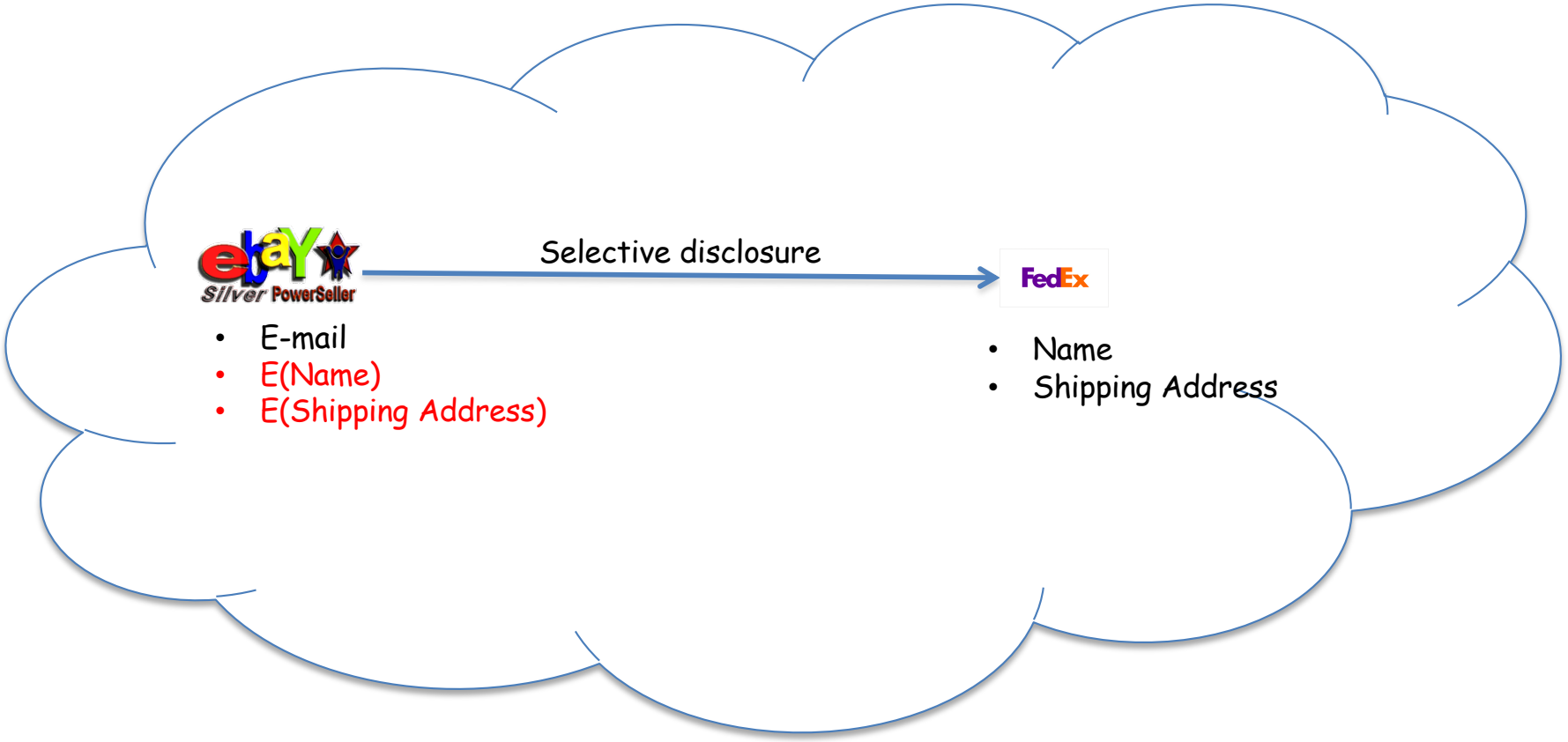
Decryption handled by Multi-Party Computing as in the previous slides

Proposed IDM: Selective Disclosure



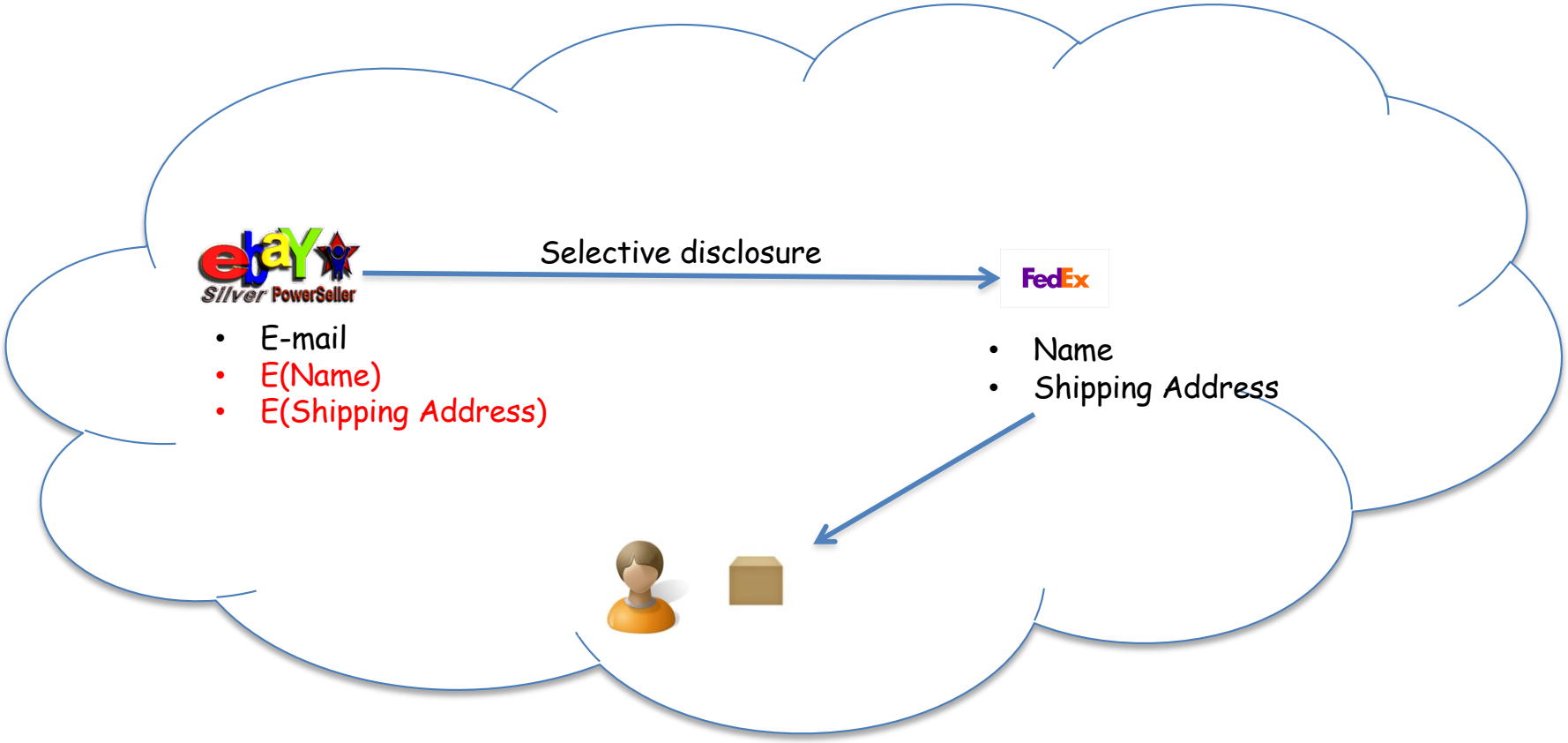
*e-bay seller shares the encrypted information based on the user policy

Proposed IDM: Selective Disclosure



- Decryption handled by Multi-Party Computing as in the previous slides

Proposed IDM: Selective Disclosure



- Fed-Ex can now send the package to the user

Proposed IDM: Identity in the Cloud

Proposed IDM: Characteristics and Advantages

- Ability to use Identity data on untrusted hosts
  - Self Integrity Check
  - Integrity compromised- apoptosis or evaporation
  - Data should not be on this host
- Independent of Third Party
  - Prevents correlation attacks
- Establishes the trust of users in IDM
  - Through putting the user in control of who has his data
  - Identity is being used in the process of authentication, negotiation, and data exchange.
- Minimal disclosure to the SP
  - SP receives only necessary information.

Proposed IDM: Conclusion & Future Work

- Problems with IDM in Cloud Computing
  - Collusion of Identity Information
  - Prohibited Untrusted Hosts
  - Usage of Trusted Third Party
- Proposed Approaches
  - IDM based on Anonymous Identification
  - IDM based on Predicate over Encrypted data
- Future work
  - Develop the prototype, conduct experiments and evaluate the approach

Minimize Multi-tenancy

Minimize Multi-tenancy

- Can't really force the provider to accept less tenants
  - Can try to increase isolation between tenants
    - Strong isolation techniques (VPC to some degree)
      - C.f. VM Side channel attacks (T. Ristenpart et al.)
    - QoS requirements need to be met
    - Policy specification
  - Can try to increase trust in the tenants
    - Who's the insider, where's the security boundary? Who can I trust?
    - Use SLAs to enforce trusted behavior

Conclusion

- Cloud computing is sometimes viewed as a reincarnation of the classic mainframe client-server model
  - However, resources are ubiquitous, scalable, highly virtualized
  - Contains all the traditional threats, as well as new ones
- In developing solutions to cloud computing security issues it may be helpful to identify the problems and approaches in terms of
  - Loss of control
  - Lack of trust
  - Multi-tenancy problems

CLOUD COMPUTING FOR MOBILE USERS: CAN OFFLOADING COMPUTATION SAVE ENERGY?

What cloud gives us, generally

Take Amazon cloud for example.

- store personal data

(Simple Storage Service (S3))

- perform computations on stored data

(Elastic Compute Cloud (EC2).)

What cloud gives us, generally

If you want to set up a business.

- low initial capital investment
- shorter start-up time for new services
- lower maintenance and operation costs
- higher utilization through virtualization
- easier disaster recovery

What about cloud computing for mobile users? Specifically

Two main concerns:

- mobile computing are limited energy
- wireless bandwidth

The importance of battery lifetime of mobile phones

Various studies have identified longer battery lifetime as the most desired feature of such systems.

- longer battery life to be more important than all other features, including cameras or storage.
- short battery life to be the most disliked characteristic of Apple's iPhone 3GS
- battery life was the top concern of music phone users.

Four basic approaches to saving energy and extending battery lifetime in mobile devices:

- Adopt a new generation of semiconductor technology.
- Avoid wasting energy. (when it is idle, sleep mode)
- Execute programs slowly. (When a processor's clock speed doubles, the power consumption nearly octuples).
- Eliminate computation all together. (offloading these applications to the cloud).

Can offloading these applications to the cloud save energy and extend battery lifetimes for mobile users?

How to implement a quantitative study. The amount of energy saved is

$$P_c \times \frac{C}{M} - P_i \times \frac{C}{S} - P_{tr} \times \frac{D}{B}.$$

S : the speed of cloud to compute C instructions

M : the speed of mobile to compute C instructions

D : the data need to transmit

B : the bandwidth of the wireless Internet

Can offloading these applications to the cloud save energy and extend battery lifetimes for mobile users?

p the energy cost per second when the mobile phone is doing computing

p_i the energy cost per second when the mobile phone is idle.

p_{tr} the energy cost per second when the mobile is transmission the data.

Can offloading these applications to the cloud save energy and extend battery lifetimes for mobile users?

Suppose the server is F times faster—that is, S = F × M. We can rewrite the formula as

$$\frac{C}{M} \times \left(P_c - \frac{P_i}{F} \right) \vdash P_{tr} \times \frac{D}{B}.$$

Energy is saved when this formula produces a positive number. The formula is positive if D/B is sufficiently small compared with C/M and F is sufficiently large.

sample applications benefiting from offloading

chess game.

A chessboard has 8 × 8 = 64 positions. Each player controls 16 pieces at the beginning of the game. Each piece may be in one of the 64 possible locations and needs 6 bits to represent the location. To represent a chess game's current state, it is sufficient to state that 6 bits × 32 pieces = 192 bits = 24 bytes; this is smaller than the size of a typical wireless packet.

sample applications benefiting from offloading

The amount of computation for chess is very large; Claude Shannon and Victor Allis estimated the complexity of chess to exceed the number of atoms in the universe. Since the amount of computation C is extremely large, and D is very small, chess provides an example where offloading is beneficial for most wireless networks.

sample applications not benefiting from offloading

- regions like national parks
- the basement of a building
- interior of a tunnel,
- subway.

In these cases,

where the value of B in Equation can become very small or even zero, cloud computing does not save energy.

Making computation offloading more attractive

There is a fundamental assumption under-lying this analysis with the client-server model: Because the server does not already contain the data, all the data must be sent to the service provider.

However, cloud computing changes that assumption: The cloud stores data and performs computation on it. For example, services like Amazon S3 can store data, and Amazon EC2 can be used to perform computation on the data stored using S3.

When considering Privacy and security

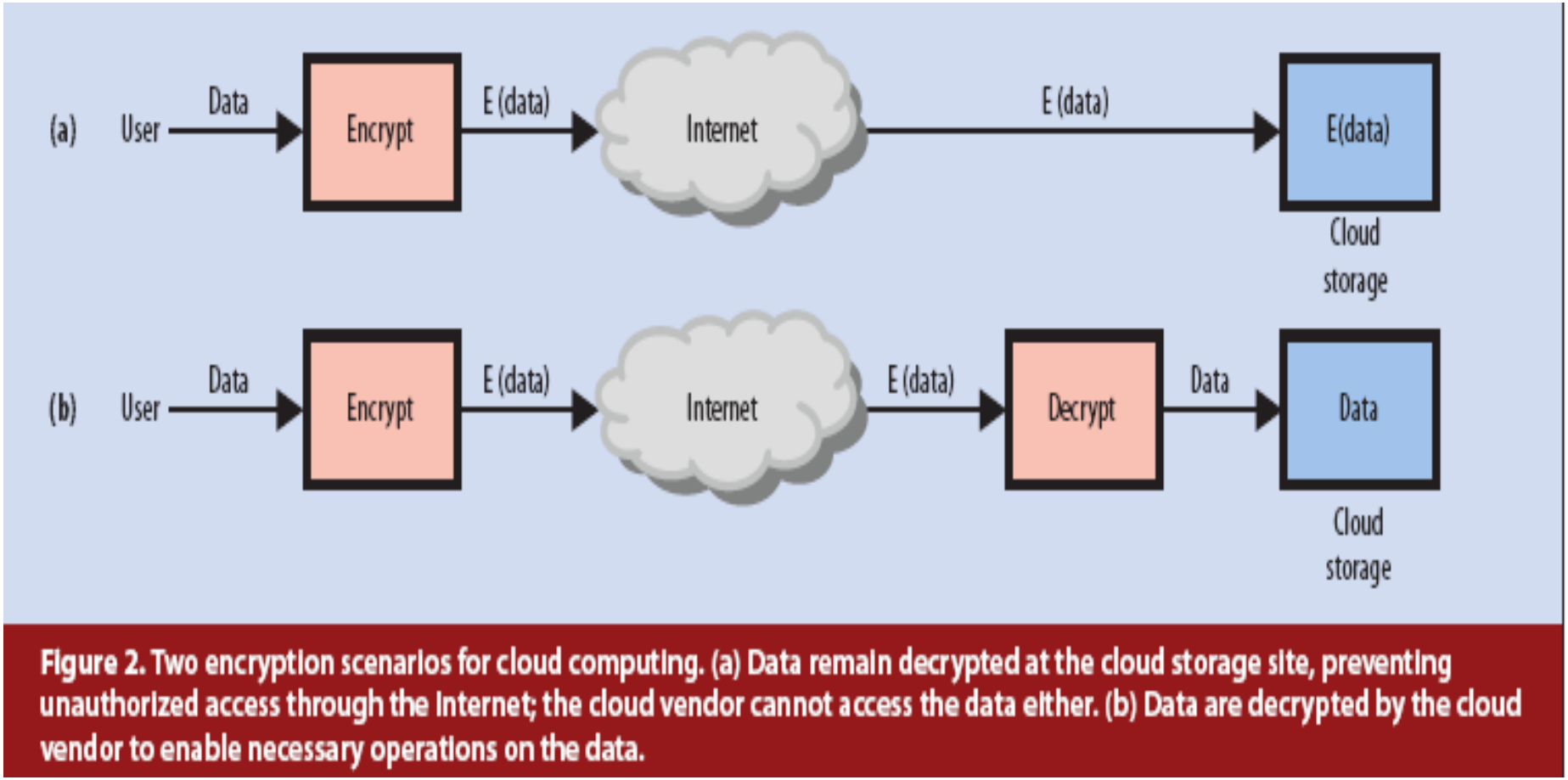


Figure 2. Two encryption scenarios for cloud computing. (a) Data remain decrypted at the cloud storage site, preventing unauthorized access through the Internet; the cloud vendor cannot access the data either. (b) Data are decrypted by the cloud vendor to enable necessary operations on the data.

When considering Privacy and security

Another possible privacy and security solution is to use a technique called steganography :

- Multimedia content like images and videos have significant redundancy. This makes it possible to hide data in multimedia using steganography.

- Steganographic techniques can be used to transform the data before storage so that operations can still be performed on the data.

When considering Privacy and security

Figure A. An example of steganography: images (a) and (c) look identical, but image (c) contains image (b) hidden in it. Applying appropriate transformations to image (c) can obtain image (b).

When considering Privacy and security

Performing encryption or steganographic techniques before sending data to the cloud requires some additional processing on the mobile system. So the formula become:

$$\frac{C}{M} \times \left(P_c - \frac{P_1}{F} \right) - P_{tr} \times \frac{D}{B} - P_c \times \frac{C_p}{M},$$

Conclusion

- cloud computing can potentially save energy for mobile users.
- not all applications are energy efficient when migrated to the cloud.
- cloud computing services would be significantly different from cloud services for desktops because they must offer energy savings.
- The services should consider the energy overhead for privacy, security, reliability, and data communication before offloading.

Conclusion

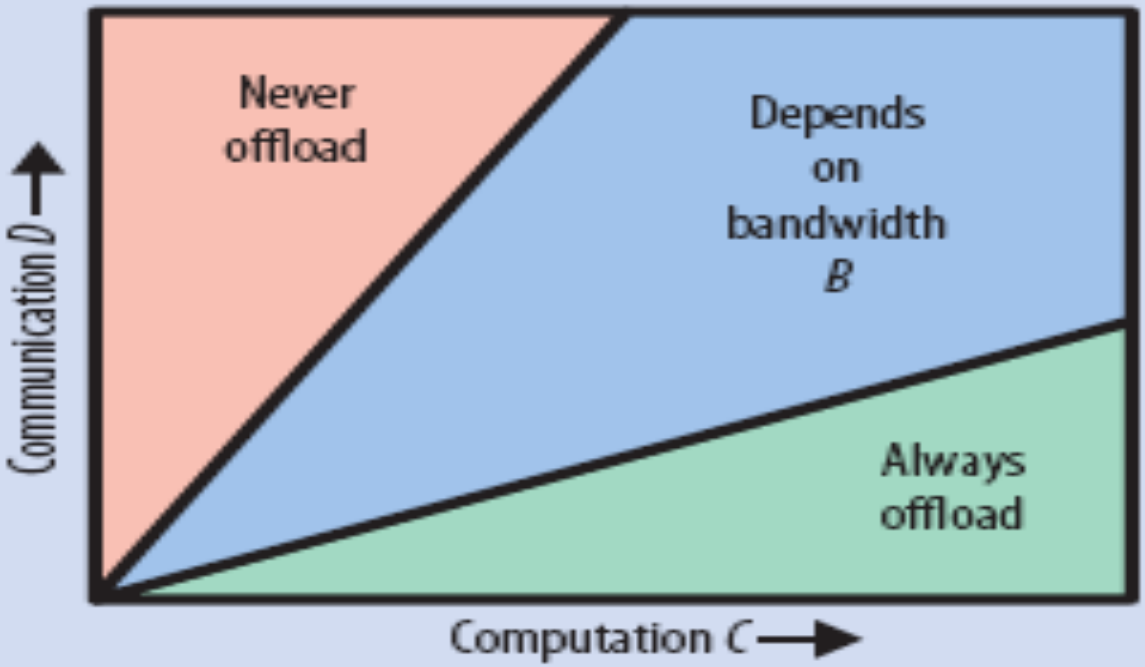


Figure 1. Offloading is beneficial when large amounts of computation *C* are needed with relatively small amounts of communication *D*.

Bandwidth Measurements for VMs in Cloud

MOTIVATION

- Many applications are being deployed in cloud to leverage the scalability provided by the cloud providers.
- Tools provided by the cloud providers do not give performance metrics from the network perspective.
- Network topology is not exposed to the cloud users and the applications consider all network links to be homogeneous.
- Metrics such as available bandwidth, latency etc. will be more useful to the cloud users.

Experimental Evaluation

- Set up
  - 19 EC2 small instances (US East)
  - 342 links between VMs
  - Ubuntu 10.04 server version

- Centralized Scheduler for starting Iperf clients
  - Predefined serialized schedule file at each VM instance.
  - Schedule file contains a time stamp along with the nodes that should communicate for a single reading.

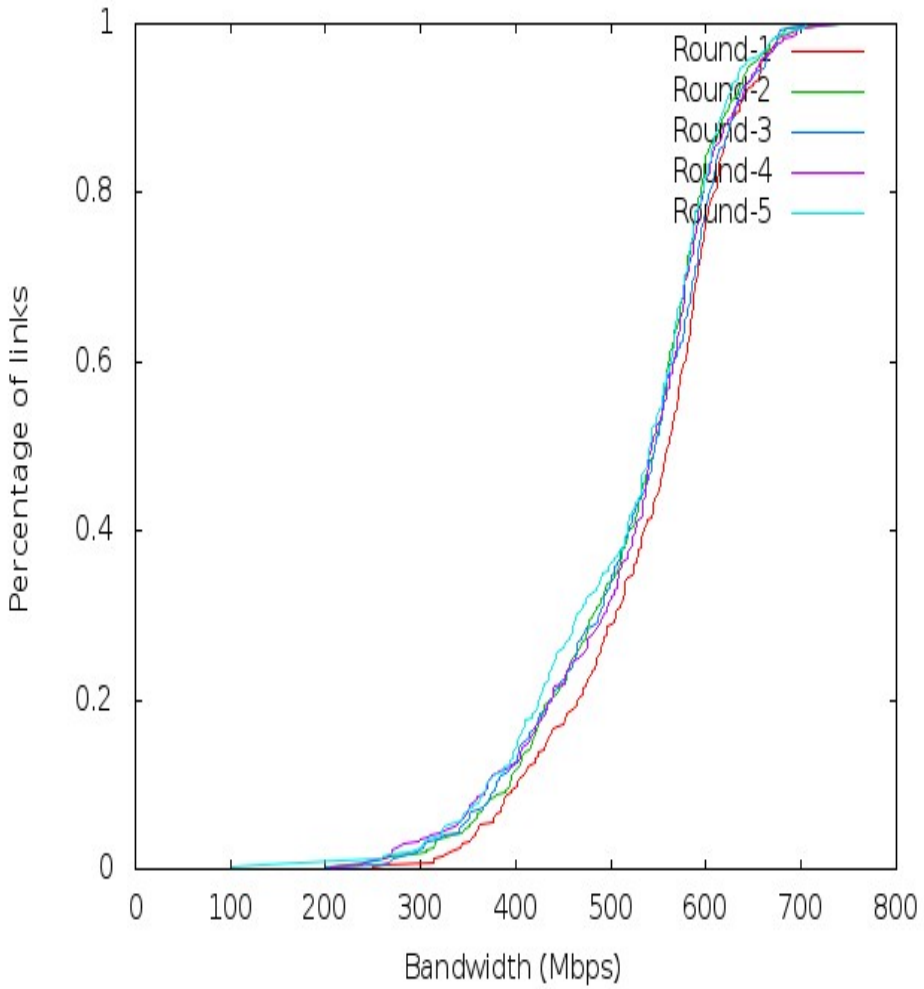
* Iperf - Network testing tool to measure the network throughput between end hosts.

Experimental Evaluation

- Iperf takes 6 seconds to get a reading for a single link.
- Each round of measurement takes around 30 minutes for finding available bandwidth for all 342 links.
- Total 5 rounds in total
- Throughput matrix: Matrix containing estimated values for available bandwidth

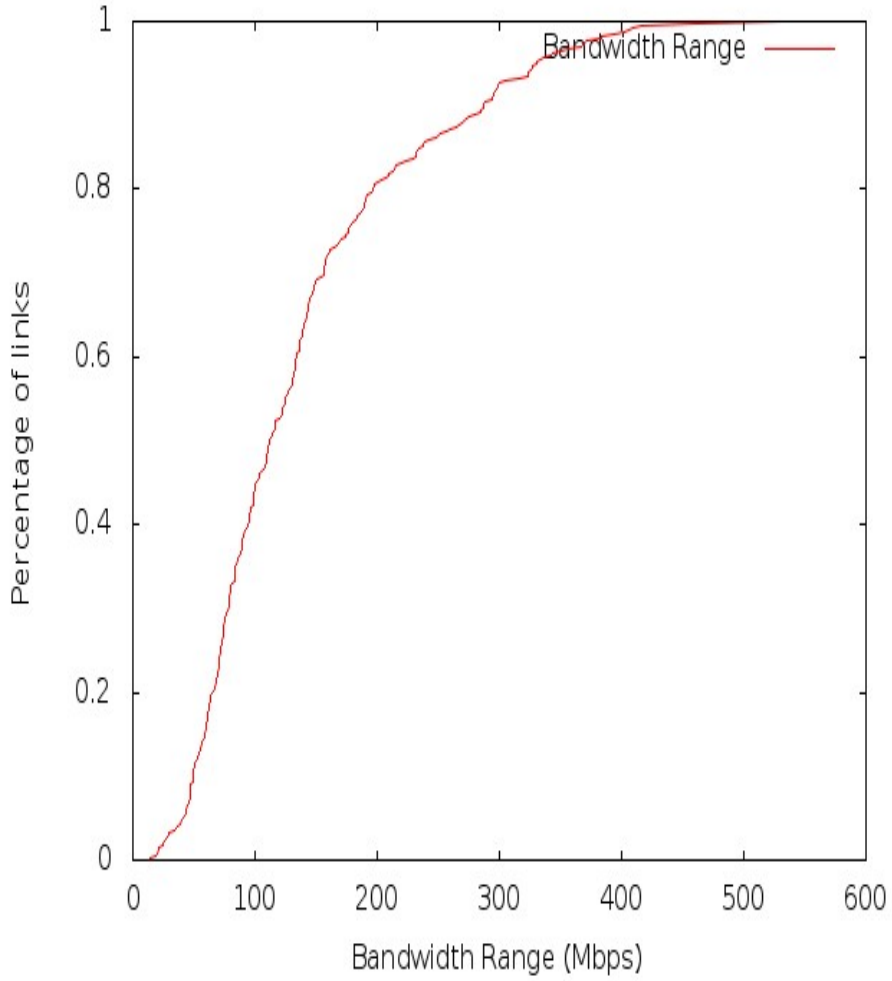
Bandwidth Estimation

- Shows the CDF of link bandwidth estimation for all the rounds.
- Used throughput matrix having estimated 342 values.
- All links in clouds are not homogeneous.
- Only 10% of the links have available bandwidth less than 400Mbps.



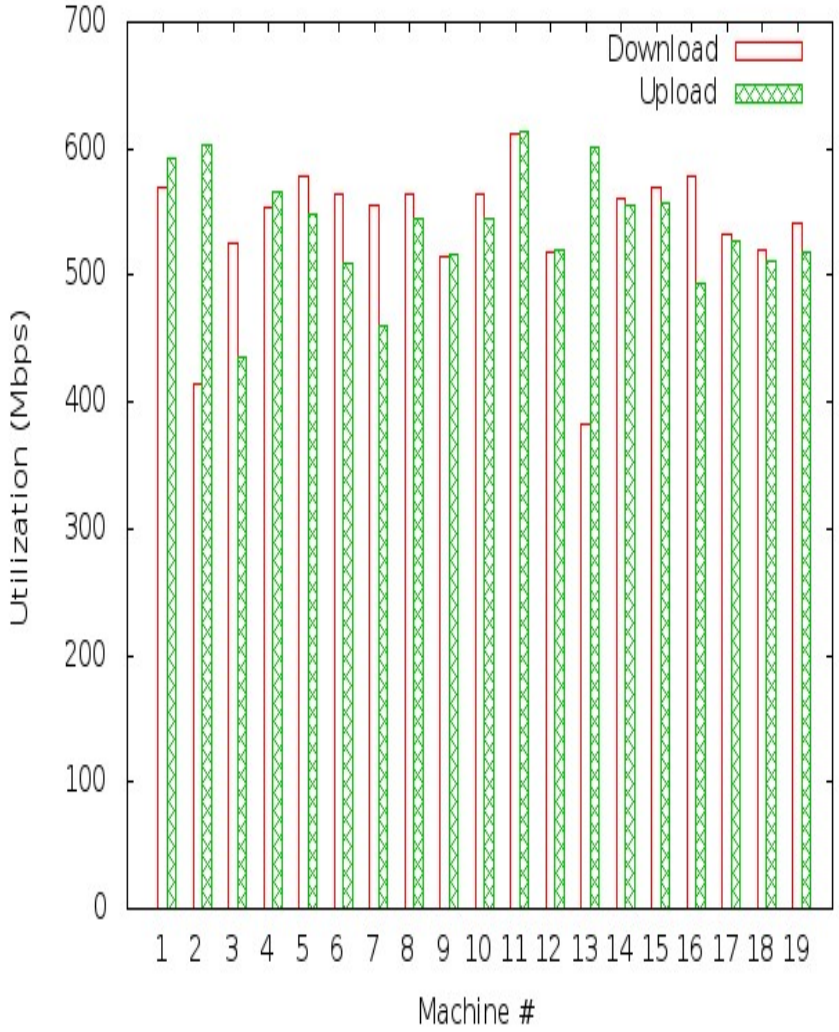
Bandwidth Variation Estimation

- Shows the CDF of link bandwidth variation across all the rounds.
- Bandwidth range of a link defined as the difference between the max and min value across all rounds.
- For most of the links, bandwidth is consistent across time. Only 20% links have variation of more than 200 Mbps.



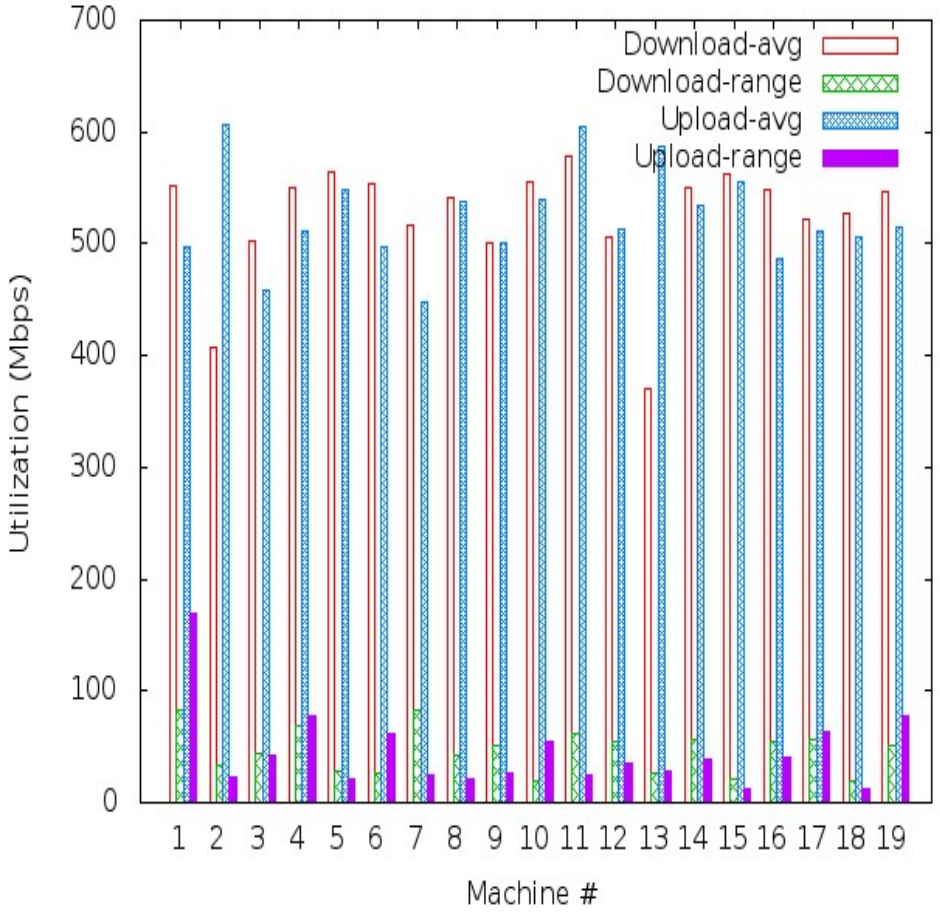
Virtual Machine Performance

- Shows the available download/upload bandwidth of all machines for a single round
- Almost all the machines have average available bandwidth more than 400 Mbps.



Virtual Machine Performance

- Shows the average available download/ upload bandwidth and its range for each machine across all rounds.
- Almost all the machines have average download/ upload bandwidth more than 400 Mbps.
- Some VMs (1, 4, 7) have large available bandwidth variation.



CONCLUSIONS

- Focussed on available bandwidth metric between each pair of VM instances.
- Amazon EC2 data center is optimally utilized with ample available bandwidth for almost all VMs.
- Some badly performing VMs can be pointed out based on the large variation in the available upload/download bandwidth and can be replaced with new VMs.

Future Work

- More performance metric such as latency etc. can be considered.
- These performance metrics can be used to improve the performance of applications running in the cloud.
- These performance metric tests can be run on large EC2 instances.

A Mobile-Cloud Collaborative Approach for Context-Aware Blind Navigation

Outline

- Problem Statement
- Goals
- Challenges
- Context-aware Navigation Components
- Existing Blind Navigation Aids
- Proposed System Architecture
- Advantages of Mobile-Cloud Approach
- Traffic Lights Detection
  - Related Work
  - System Developed
  - Experiments
- Work In Progress

Problem Statement

- Indoor and outdoor navigation is becoming a harder task for blind and visually impaired people in the increasingly complex urban world
- Advances in technology are causing the blind to fall behind, sometimes even putting their lives at risk
- Technology available for context-aware navigation of the blind is not sufficiently accessible; some devices rely heavily on infrastructural requirements

Demographics

- 314 million visually impaired people in the world today
- 45 million blind
- More than 82% of the visually impaired population is age 50 or older
- The old population forms a group with diverse range of abilities
- The disabled are seldom seen using the street alone or public transportation

Goals

- ***Make a difference***

  Bring mobile technology in the daily lives of blind and visually impaired people to help achieve a higher standard of life
- Take a major step in context-aware navigation of the blind and visually impaired
- Bridge the gap between the needs and available technology
- Guide users in a non-overwhelming way
- Protect user privacy

Challenges

- Real-time guidance
- Portability
- Power limitations
- Appropriate interface
- Privacy preservation
- Continuous availability
- No dependence on infrastructure
- Low-cost solution
- Minimal training

Discussions

- Cary Supalo: Founder of Independence Science LLC (http://www.independencescience.com/)
- T.V. Raman: Researcher at Google, leader of Eyes-Free project (speech enabled Android applications)
- American Council of the Blind of Indiana State Convention, 31 October 2009
- Miami Lighthouse Organization

Mobility Requirements

- Being able to avoid obstacles
- Walking in the right direction
- Safely crossing the road
- Knowing when you have reached a destination
- Knowing which is the right bus/train
- Knowing when to get off the bus/train

All require SIGHT as primary sense

Context-Aware Navigation Components

- Outdoor Navigation (finding curbs -including in snow, using public transportation, interpreting traffic patterns/signal lights…)
- Indoor Navigation (finding stairs/elevator, specific offices, restrooms in unfamiliar buildings, finding the cheapest TV at a store…)
- Obstacle Avoidance (both overhanging and low obstacles…)
- Object Recognition (being able to reach objects needed, recognizing people who are in the immediate neighborhood…)

Existing Blind Navigation Aids – Outdoor Navigation

- Loadstone GPS (http://www.loadstone-gps.com/)
- Wayfinder Access (http://www.wayfinderaccess.com/)
- BrailleNote GPS (www.humanware.com)
- Trekker (www.humanware.com)
- StreetTalk (www.freedomscientific.com)
- DRISHTI [1]
- …

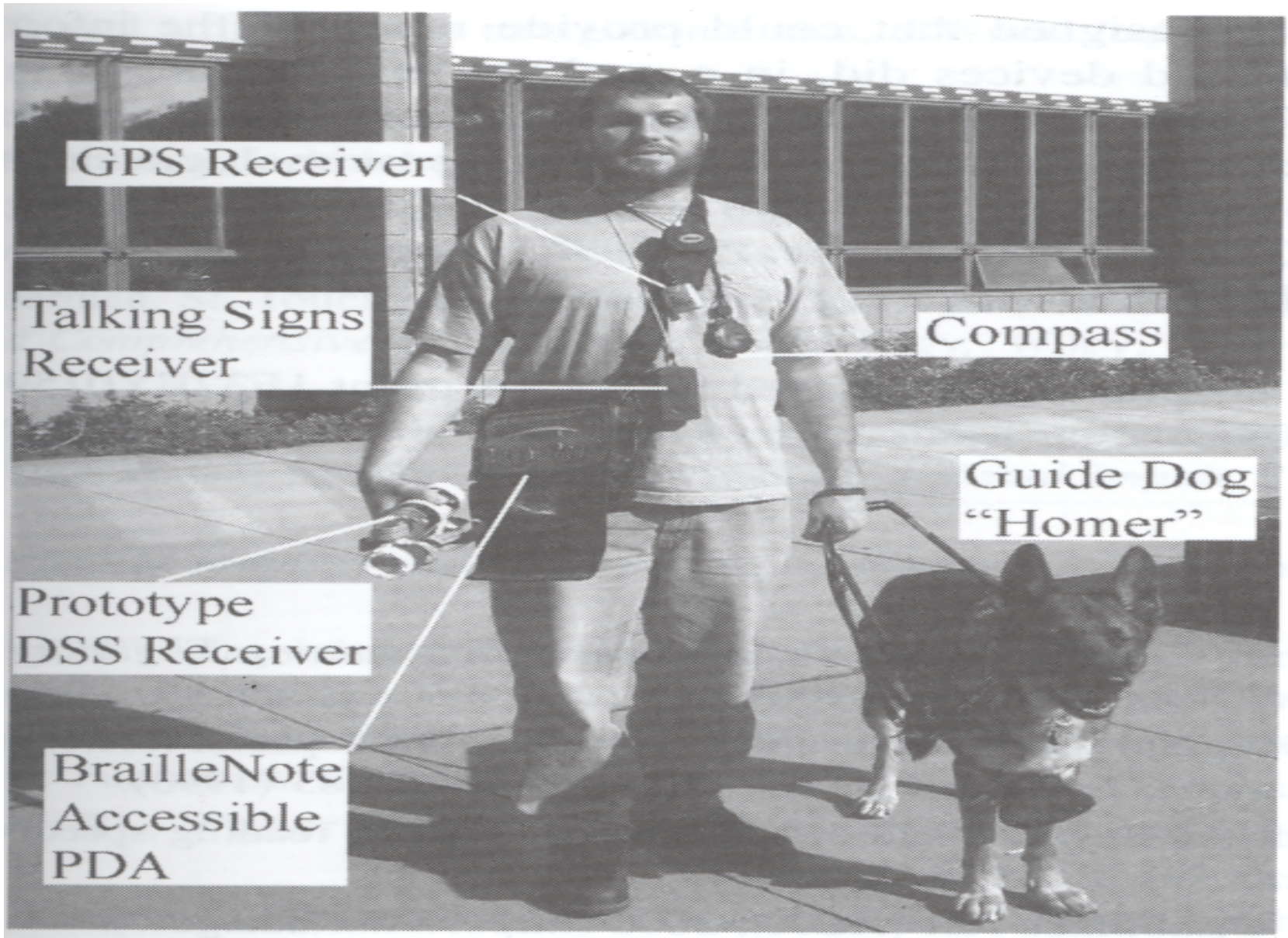
Existing Blind Navigation Aids – Indoor Navigation

- InfoGrid (based on RFID) [2]
- Jerusalem College of Technology system (based on local infrared beams) [3]
- Talking Signs (www.talkingsigns.com) (audio signals sent by invisible infrared light beams)
- SWAN (audio interface guiding user along path, announcing important features) [4]
- ShopTalk (for grocery shopping) [5]

Existing Blind Navigation Aids – Obstacle Avoidance

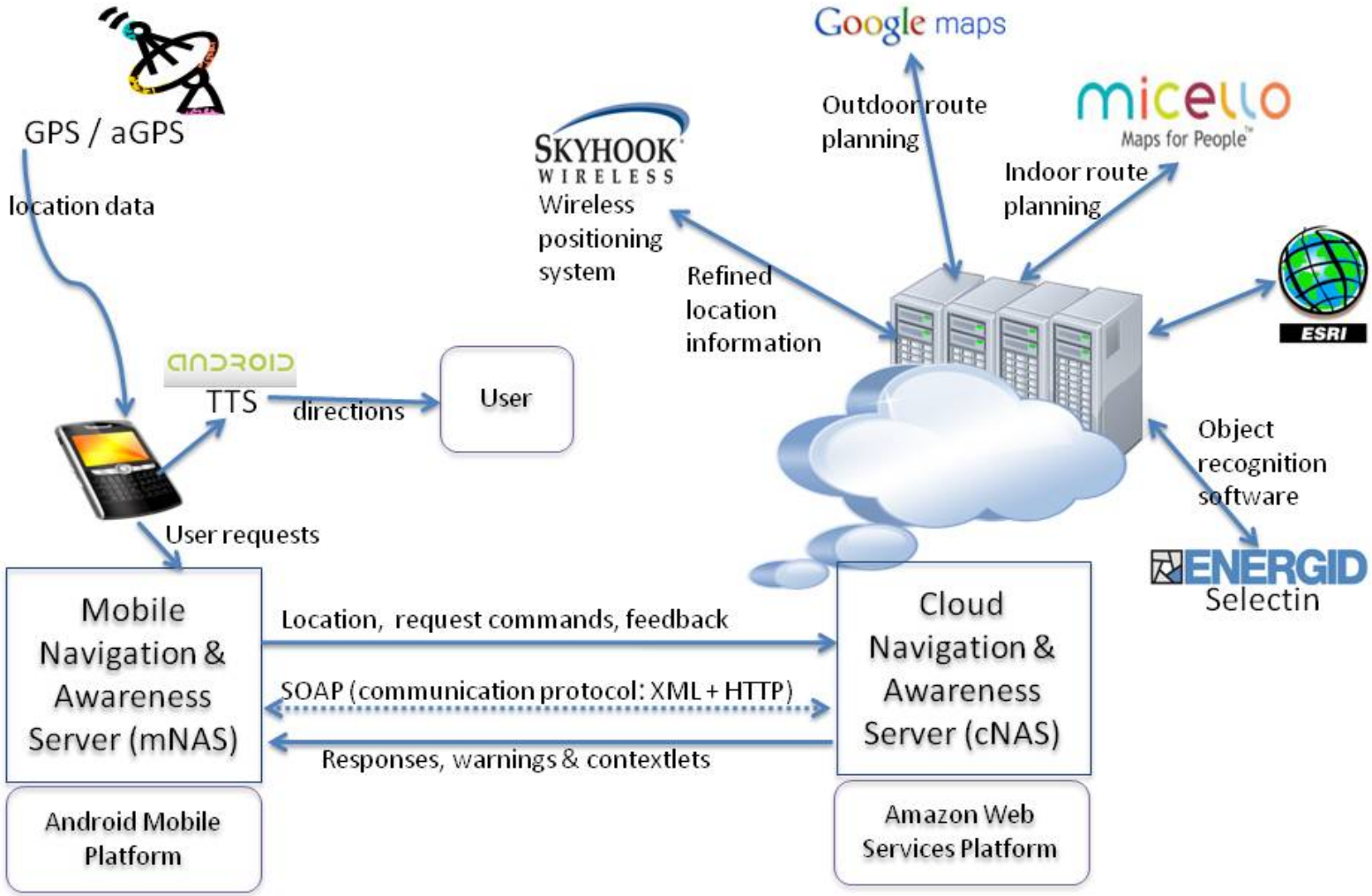
- RADAR/LIDAR
- Kay's Sonic glasses (audio for 3D representation of environment) (www.batforblind.co.nz)
- Sonic Pathfinder (www.sonicpathfinder.org) (notes of musical scale to warn of obstacles)
- MiniGuide (www.gdp-research.com.au/) (vibration to indicate object distance)
- VOICE (www.seeingwithsound.com) (images into sounds heard from 3D auditory display)
- Tactile tongue display [6]
- …

Putting all together…



Gill, J. Assistive Devices for People with Visual Impairments.
In A. Helal, M. Mokhtari and B. Abdulrazak, ed., *The Engineering Handbook of Smart Technology for Aging, Disability and Indeper*
John Wiley & Sons, Hoboken, New Jersey, 2008.

Proposed System Architecture

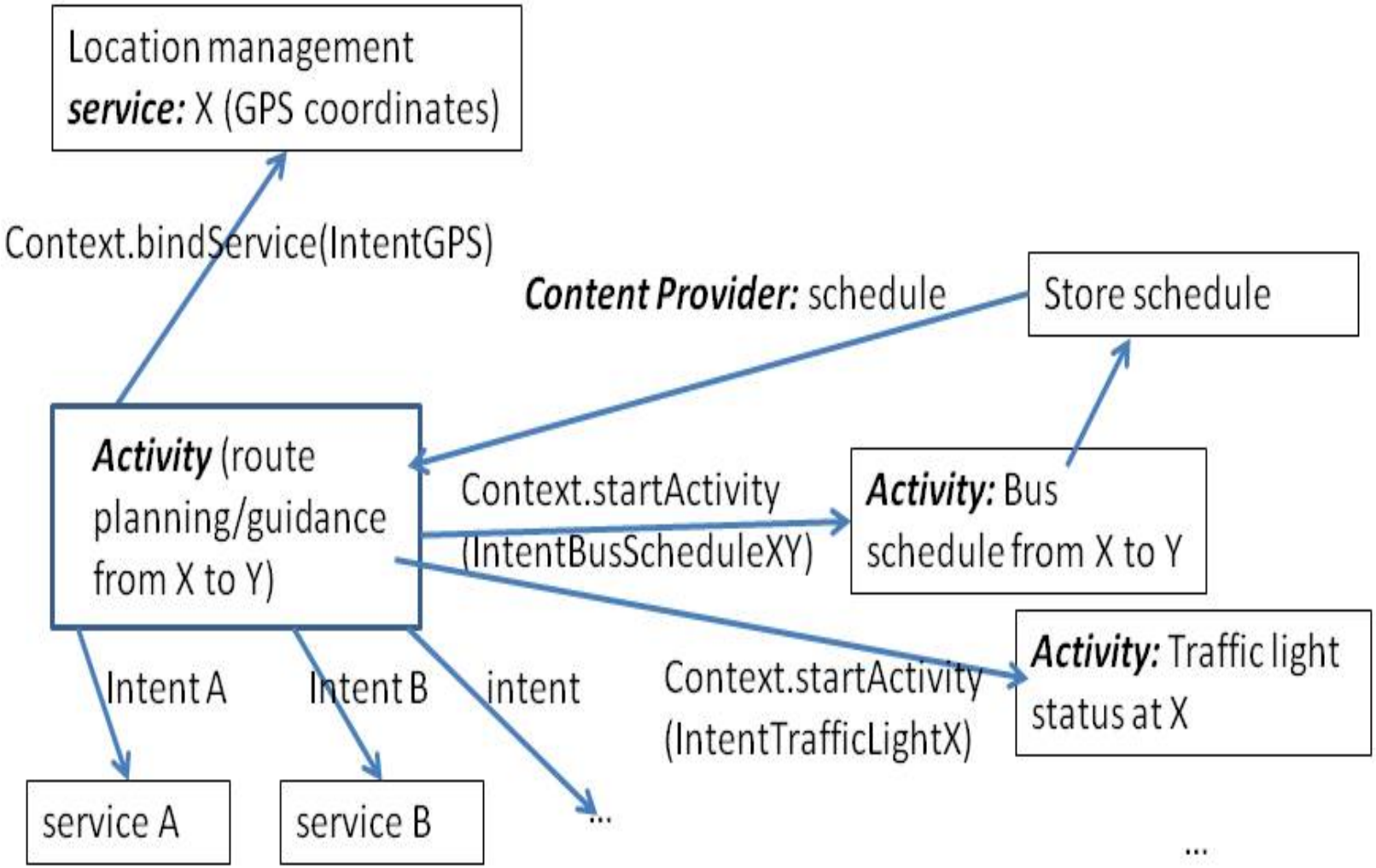


Proposed System Architecture

Services:

- Google Maps (outdoor navigation, pedestrian mode)
- Micello (indoor location-based service for mobile devices)
- Object recognition (Selectin software etc)
- Traffic assistance
- Obstacle avoidance (Time-of-flight camera technology)
- Speech interface (Android text-to-speech + speech recognition servers)
- Remote vision
- Obstacle minimized route planning

Use of the Android Platform



Advantages of a Mobile-Cloud Collaborative Approach

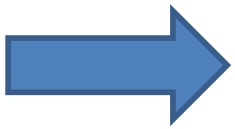
- Open architecture
- Extensibility
- Computational power
- Battery life
- Light weight
- Wealth of context-relevant information resources
- Interface options
- Minimal reliance on infrastructural requirements

Traffic Lights Status Detection Problem

- Ability to detect status of traffic lights accurately is an important aspect of safe navigation
  - Color blind
  - Autonomous ground vehicles
  - Careless drivers
- Inherent difficulty: Fast image processing required for locating and detecting the lights status → demanding in terms of computational resources
- Mobile devices with limited resources fall short alone

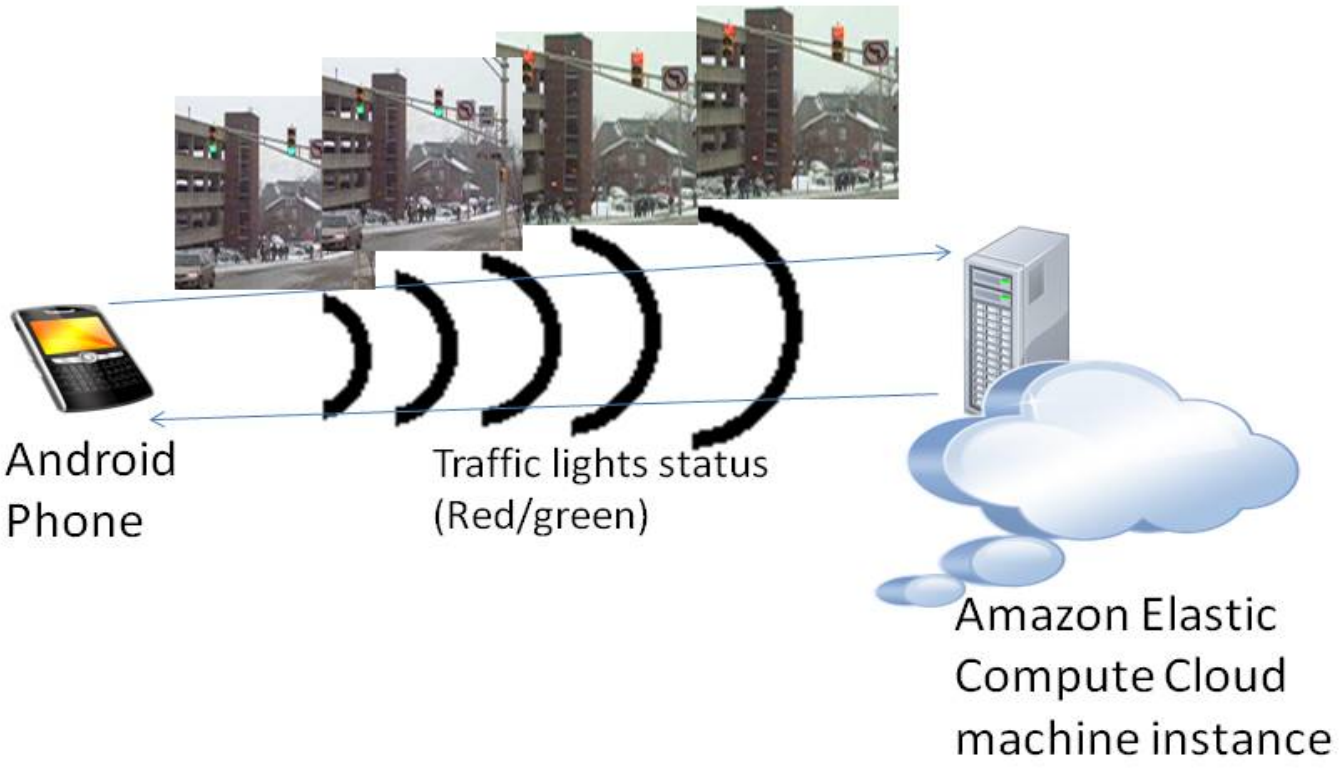
Attempts to Solve the Traffic Lights Detection Problem

- Kim et al: Digital camera + portable PC analyzing video frames captured by the camera [7]
- Charette et al: 2.9 GHz desktop computer to process video frames in real time[8]
- Ess et al: Detect generic moving objects with 400 ms video processing time on dual core 2.66 GHz computer[9]



Sacrifice portability for real-time, accurate detection

Mobile-Cloud Collaborative Traffic Lights Detector

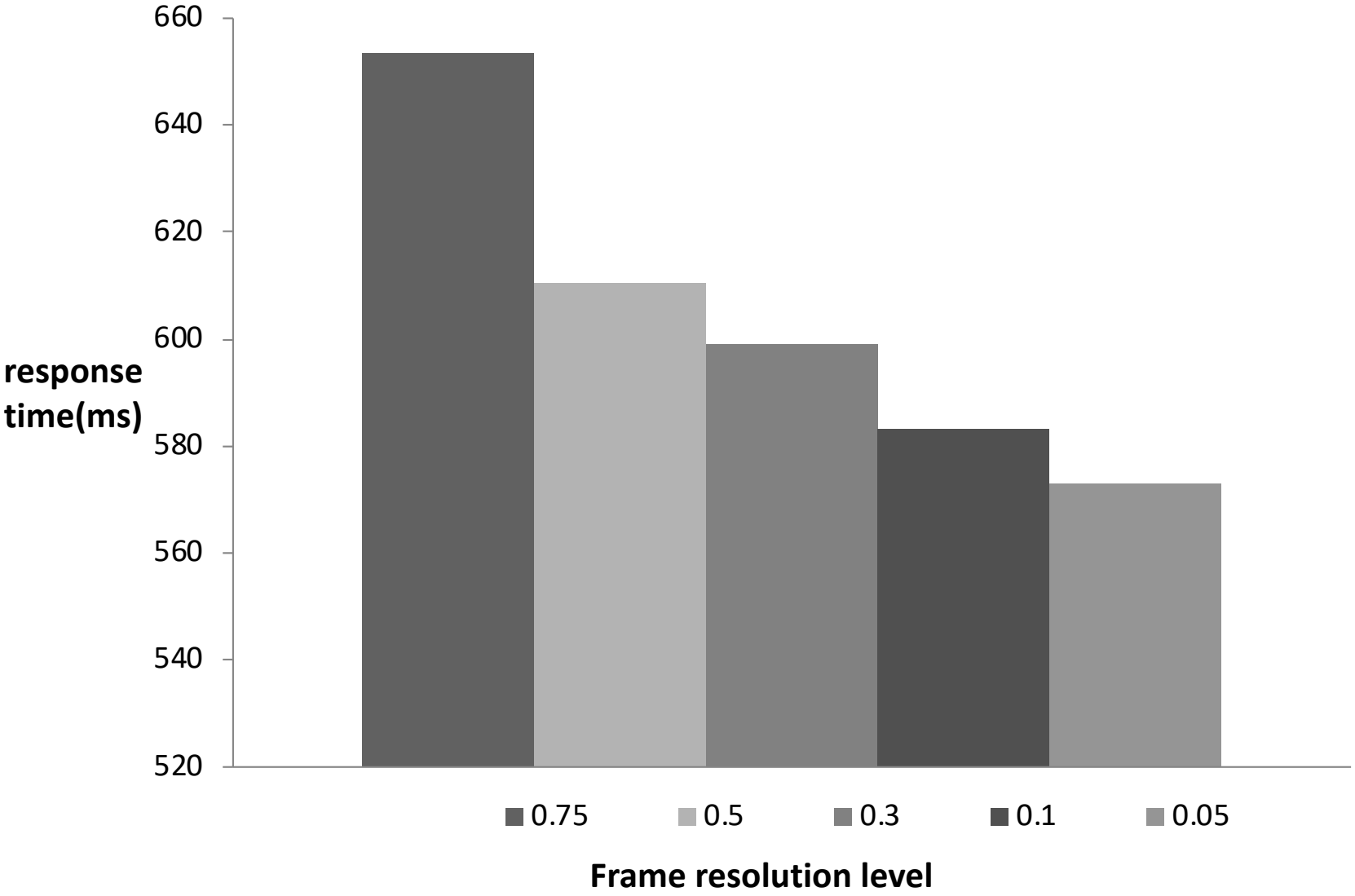


Adaboost Object Detector

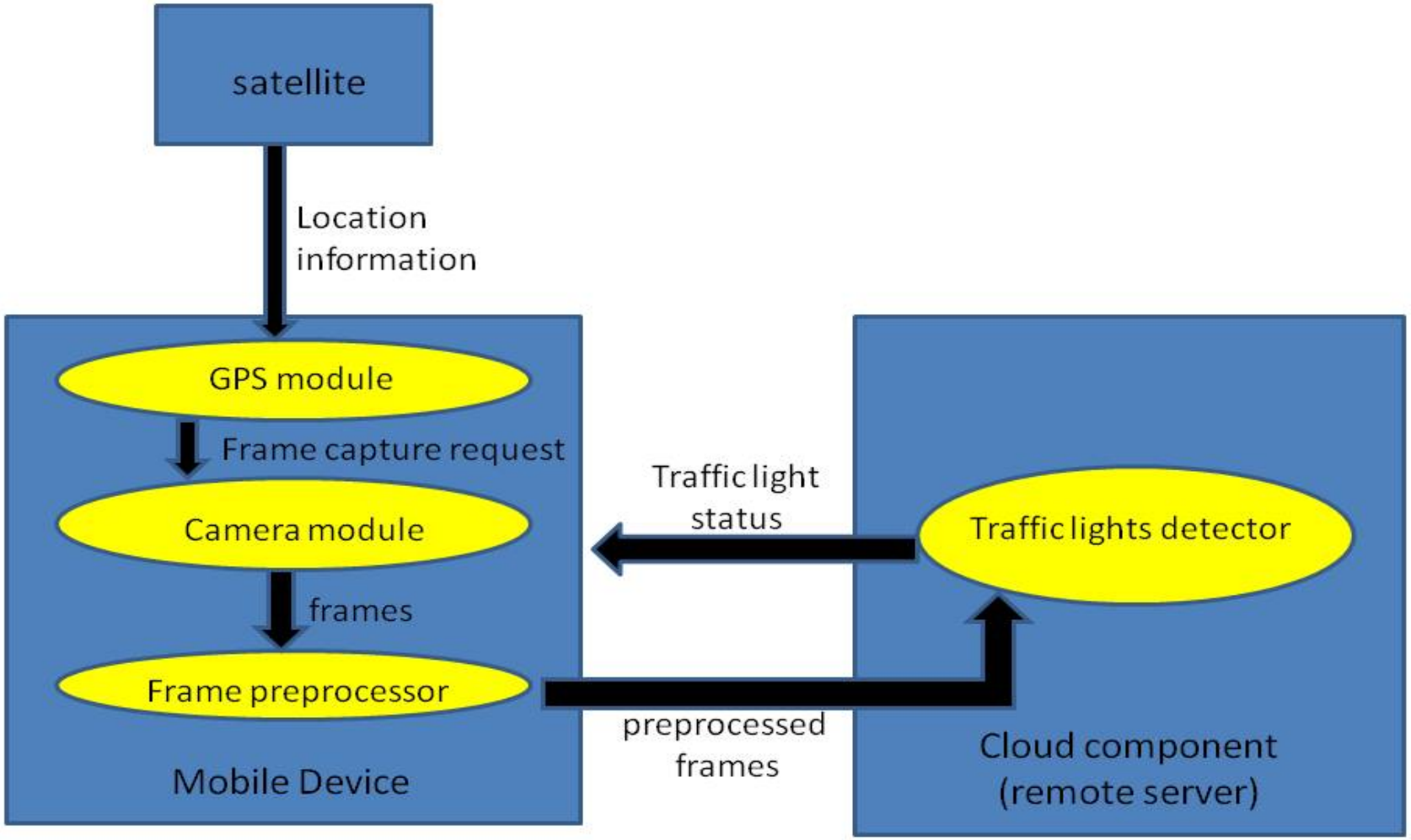
- Adaboost: Adaptive Machine Learning algorithm used commonly in real-time object recognition
- Based on rounds of calls to weak classifiers to focus more on incorrectly classified samples at each stage
- Traffic lights detector: trained on 219 images of traffic lights (Google Images)
- OpenCV library implementation

Experiments: Detector Output

Experiments: Response time



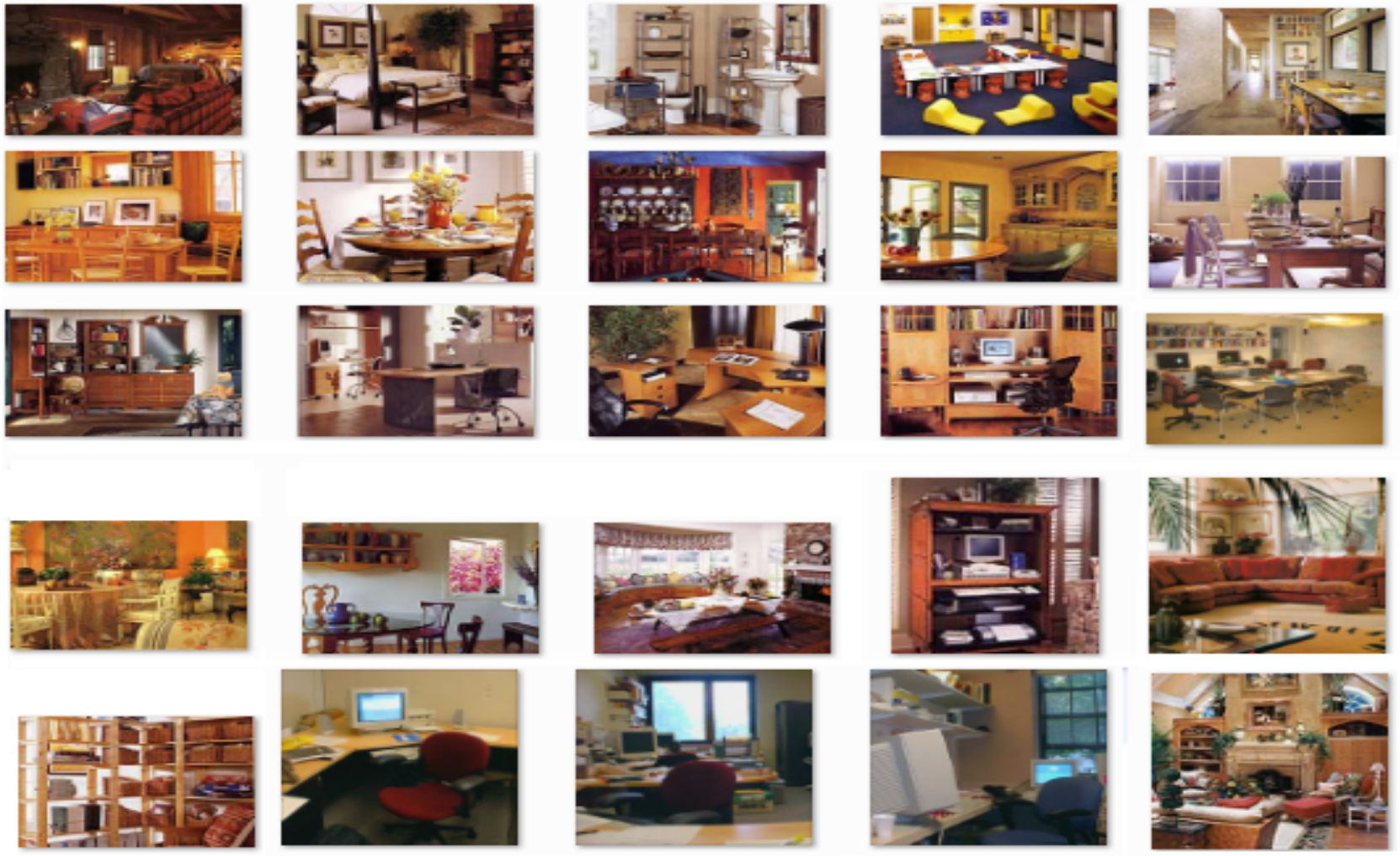
Enhanced Detection Schema



Work In Progress

- Develop fully context-aware navigation system with speech/tactile interface
- Develop robust object/obstacle recognition algorithms
- Investigate mobile-cloud privacy and security issues (minimal data disclosure principle) [10]
- Investigate options for mounting of the camera

Collective Object Classification in Complex Scenes



LabelMe Dataset (http://labelme.csail.mit.edu)

Relational Learning with Multiple Boosted Detectors for Object Categorization

- Modeling relational dependencies between different object categories
- Multiple detectors running in parallel
- Class label fixing based on confidence
- More accurate classification than AdaBoost alone
- Higher recall than classic collective classification
- Minimal decrease in recall for different classes of objects

Object Classification Experiments

Object category/Model	Boosting Only	Full Joint Collective	Conf. ranked iterative
chair	0.43	0.19	0.25
lamp	0.33	1.00	0.45
table	0.13	0.23	0.19
monitor	0.33	0.97	0.47
keyboard	0.20	1.00	0.40
sink	0.19	0.95	0.36
bed	0.32	1.00	0.52
faucet	0.07	0.92	0.13
cupboard	0.19	0.75	0.32
mouse	0.12	0.89	0.30
plant	0.18	0.88	0.31
vase	0.04	0.00	0.05

Table 1: Precision values for different classification models.

Object category/Model	Boosting Only	Full Joint Collective	Conf. ranked iterative
chair	0.25	0.98	0.58
lamp	0.26	0.06	0.25
table	0.08	0.01	0.08
monitor	0.59	0.08	0.60
keyboard	0.33	0.08	0.38
sink	0.34	0.12	0.30
bed	0.58	0.12	0.51
faucet	0.17	0.05	0.13
cupboard	0.16	0.00	0.40
mouse	0.16	0.02	0.39
plant	0.21	0.04	0.17
vase	0.08	0.00	0.05

Table 2: Recall values for different classification models.

Identity-Based Authentication for Cloud Computing

Hongwei Li, Yuanshun Dai, Ling Tian, and Haomiao Yang

CloudCom '09

What did they do?

- Proposed identity-based authentication for cloud computing, based on the identity-based hierarchical model for cloud computing (IBHMCC) and corresponding encryption and signature schemes
- Being certificate-free, the authentication protocol aligned well with demands of cloud computing

Identity-Based Hierarchical Model for Cloud Computing (IBHMCC)

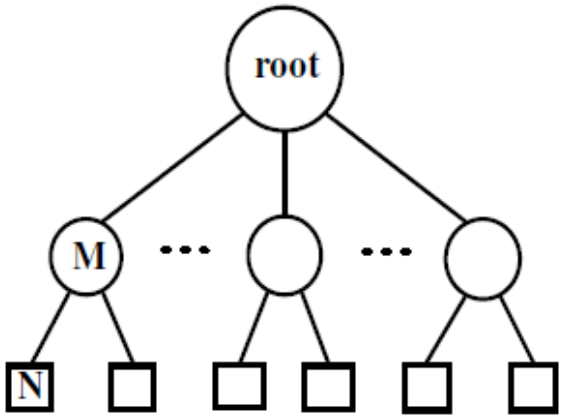


Fig. 1. IBHM for cloud computing

- Define the identity of node is the DN string from the root node to the current node itself.
- The identity of entity N is

ID_N = DN_0 || DN_M || DN_N

Deployment of IBHMCC

- Root PKG setup and Low-level setup

Root PKG setup: Root PKG acts as follows:

1. Generate group G_1, G_2 of some prime order q and an admissible pairing $\hat{e}: G_1 \times G_1 \to G_2$;
2. Choose an arbitrary generator $P \in G_1$;
3. Choose cryptography hash functions $H_1: \{0,1\}^* \to G_1, H_2: G_2 \to \{0,1\}^n$ for some n;
4. Pick a random $\alpha \in \mathbb{Z}_q^*$ and set $Q_0 = \alpha P$, $P_0 = H_1(DN_0), S_0 = \alpha P_0$. The root PKG's master key is S_0 and the system parameters are $< G_1, G_2, \hat{e}, Q_0, P, P_0, H_1, H_2 >$.

Deployment of IBHMCC (cont.)

Lower-level setup

1. Assume there are m nodes in the level-1. For each node, the root PKG acts as follows (let X be an arbitrary node in the m nodes):
2. Compute the public key of node X : $P_X = H_1(ID_X)$, where $ID_X = DN_0 \| DN_X$;
3. Pick the secret point $\rho_X \in \mathbb{Z}_q^*$ for node X . ρ_X is only known by node X and its parent node;
4. Set the secret key of node X : $S_X = S_0 + \rho_X P_X$;
5. Define the Q-value: $Q_{ID_X \|} = \rho_X P$. $Q_{ID_X \|}$ is public.

- After that, all nodes in the level-1 get and securely keep their secret keys and the secret points.
- The public key and the Q-value are publicized.
- Then, Each node in the level-1 similarly repeats the above steps (2-5).

Identity-Based Encryption

Encryption: Assume E_1 and E_2 are two entities in the cloud computing. The identity of entity E_2 is $ID_{E_2} = DN_0 \| DN_1 \| DN_2$. To encrypt message m with ID_{E_2}, E_1 acts as follows:

1. Compute

$$P_1 = H_1(DN_0 \| DN_1) \tag{1}$$

$$P_2 = H_1(DN_0 \| DN_1 \| DN_2) \tag{2}$$

2. Choose a random $r \in \mathbb{Z}_q^*$;
3. Output the ciphertext

$$C = < rP, rP_1, rP_2, H_2(g^r) \oplus m > \tag{3}$$

where $g = \hat{e}(Q_0, P_0)$ which can be pre-computed.

Identity-Based Encryption (cont.)

Decryption: After receiving the ciphertext $C = < U_0, U_1, U_2, V >$, entity E_2 can decrypt C using its secret key $S_{E_2} = S_0 + \rho_1 P_1 + \rho_2 P_2$, where ρ_1 is the secret point of node $DN_0 \| DN_1$, ρ_2 is the secret point of node $DN_0 \| DN_1 \| DN_2$:

1. Compute

$$d = \frac{\hat{e}(U_0, S_{E_2})}{\prod_{i=1}^{2} \hat{e}(Q_{ID_{E_2}|i}, U_i)} \tag{4}$$

where $Q_{ID_{E_2}|1} = \rho_1 P, Q_{ID_{E_2}|2} = \rho_2 P$;

2. Output the message $m = H_2(d) \oplus V$.

Identity-Based Signature

Signature: To sign message m, entity E_2 acts as follows:

1. Compute $P_m = H_1(DN_0 \| DN_1 \| DN_2 \| m)$;
2. Compute $\delta = S_{E_2} + \rho_2 P_m$, where ρ_2 is the secret point of entity E_2;
3. Output the signature $< \delta, P_m, Q_{ID_{E_2}\|}, Q_{ID_{E_2}|2} >$.

Verification: Other Entities can verify the signature by acting as follows: Confirm

$$\hat{e}(P,\delta) = \hat{e}(P, \rho_2 P_m)\ \hat{e}(Q_0, P_0) \prod_{i=1}^{2} \hat{e}(Q_{ID_{E_2}|i}, P_i) \qquad (5)$$

if the equation is true, the signature is validated.

Identity-Based Authentication for Cloud Computing

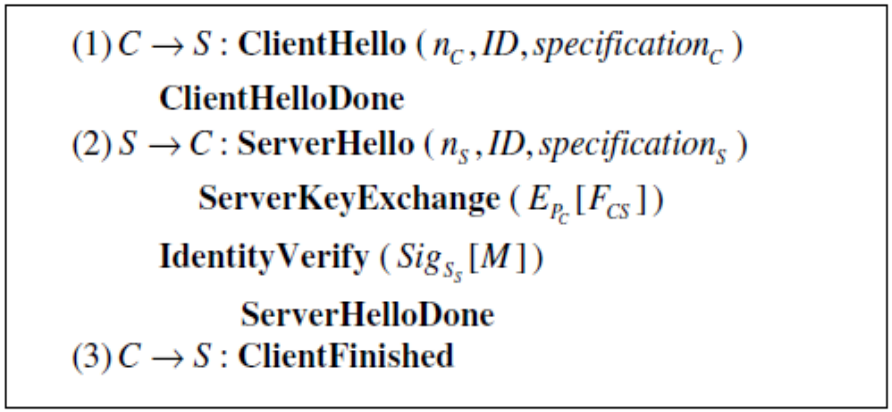


Fig. 2. Identity-based Authentication Protocol

where

n_C, n_S : the fresh random number

ID : the session identifier

$specification_C$: the cipher specification of C

$specification_S$: the cipher specification of S

F_{CS} : a pre-master secret used to generate the shared key

$E_{P_C}[F_{CS}]$: encrypt F_{CS} with the public key P_C of entity C using the encryption algorithm of IBE

M : all handshake messages since the ClientHello message

$Sig_{S_S}[M]$: sign M with the private key S_S of entity S using the signature algorithm of IBS

.Extends from TLS to handle the IBE and IBS schemes

A Simple Technique for Securing Data at Rest Stored in a Computing Cloud

Jeff Sedayao, Steven Su, Xiaohao Ma, Minghao Jiang, and Kai Miao

CloudCom '09

What did they do?

- Simple technique implemented with Open Source software solves the confidentiality of data stored on Cloud Computing Infrastructure by using public key encryption to render stored data at rest unreadable by unauthorized personnel, including system administrators of the cloud computing service on which the data is stored
- Validated their approach on a network measurement system implemented on PlanetLab
- Used it on a service where confidentiality is critical – a scanning application that validates external firewall implementations

Problem Scope

- Goal is to ensure the confidentiality of data at rest
- "Data at rest" means that the data that is stored in a readable form on a Cloud Computing service, whether in a storage product like S3 or in a virtual machine instance as in EC2

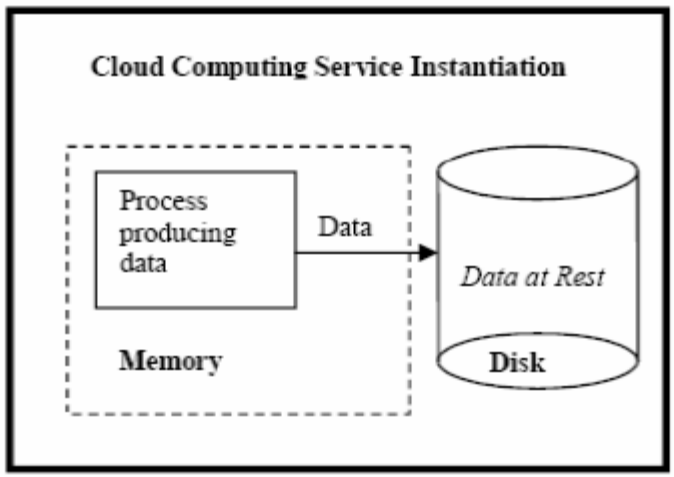


Fig. 1. Process in a Cloud Computing Infrastructure producing Data at Rest

Problem Scope (cont.)

- To protect data at rest, they want to prevent other users in the cloud infrastructure who might have access to the same storage from reading the data our process has stored
- They also want to prevent system administrators who run the cloud computing service from reading the data.
- They assume that it is unlikely for an adversary to snoop on the contents of memory.
  - If the adversary had that capability, it is unlikely that we could trust the confidentiality of any of the data that we generated there.

Problem Scope (cont.)

- While the administrative staff of the cloud computing service could theoretically monitor the data moving in memory before it is stored in disk, we believe that administrative and legal controls should prevent this from happening.
- They also do not guard against the modification of the data at rest, although we are likely to be able to detect this.

Solution Design

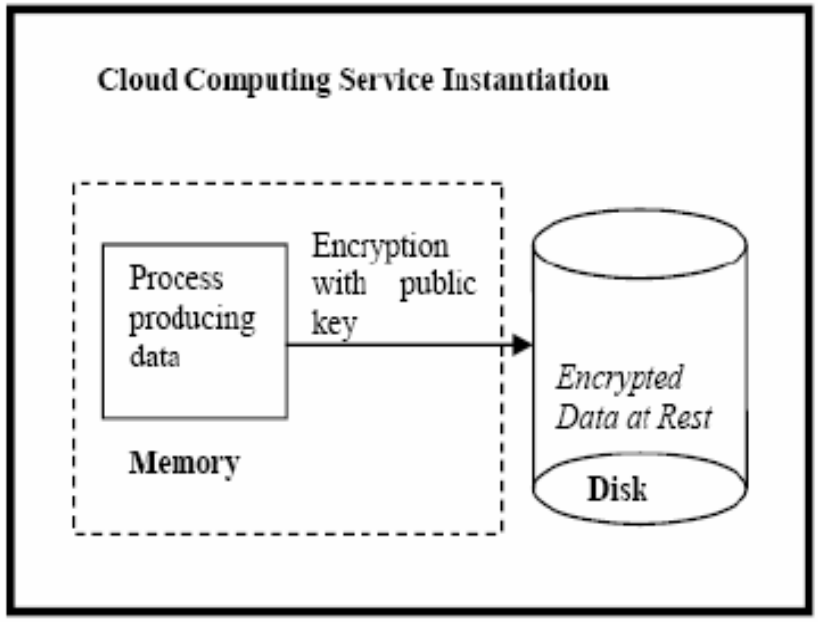


Fig. 2. Process in a Cloud Computing Infrastructure producing Encrypted Data at Rest

Solution Design (cont.)

- On a trusted host, collect the encrypted data, as shown in Figure 3, and decrypt it with the collection agent's private key which stays on that host. Note that in this case, we are in exclusive control of the private key, which the cloud service provider has no view or control over.
- They w[illegible] ter.

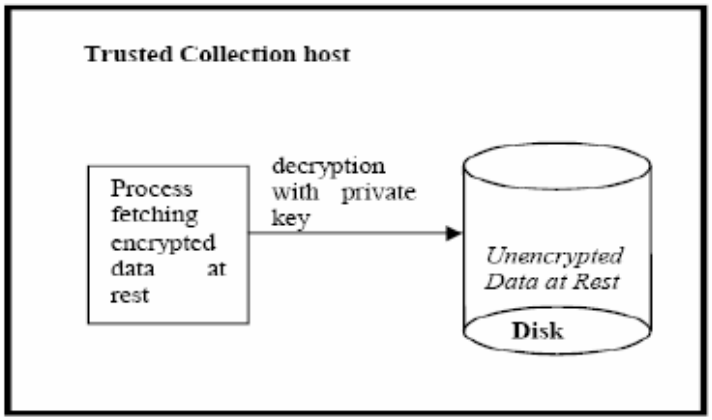


Fig. 3. Process in a Cloud Computing Infrastructure producing Encrypted Data at Rest

Implementation Experiences

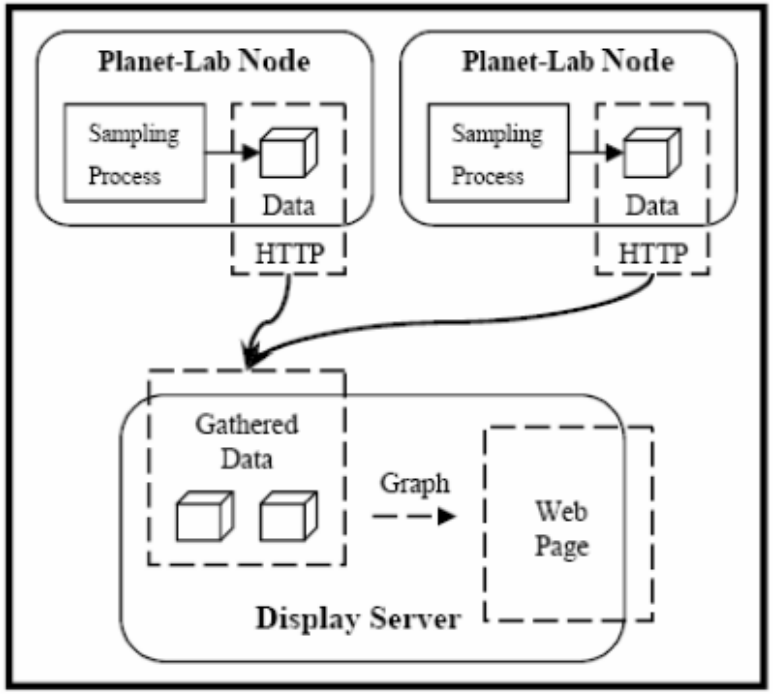


Fig. 4. Web performance data gathering and display methodology

Implementation Experiences (cont.)

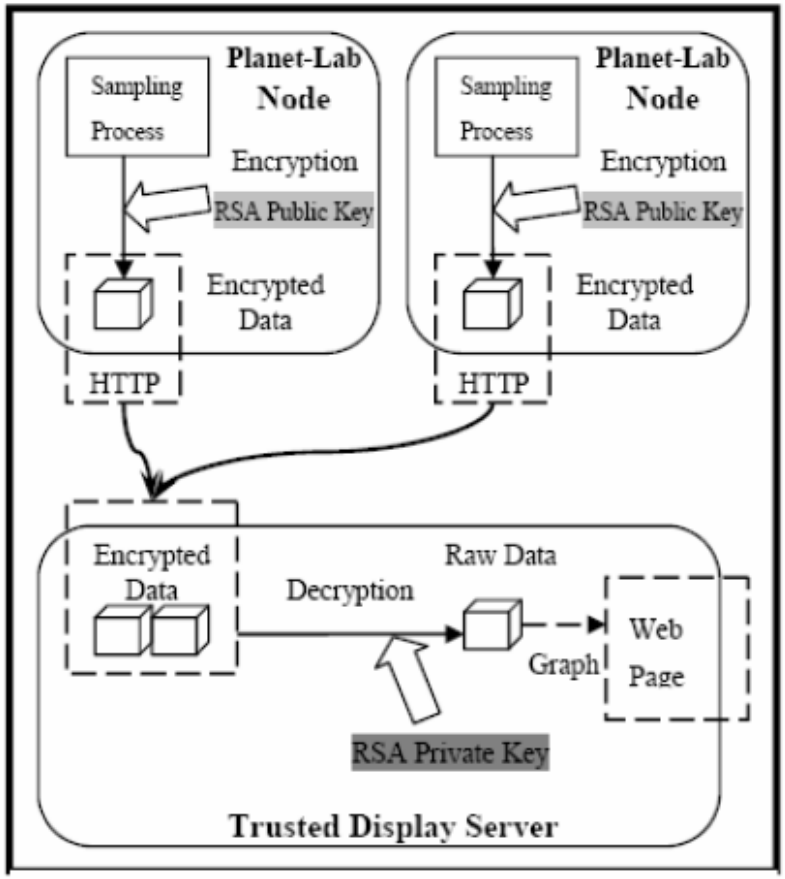


Fig. 5. Secured Web Performance Monitoring Application with Data Encryption and Decryption

Privacy in a Semantic Cloud: What's Trust Got to Do with It?

Åsmund Ahlmann Nyre and Martin Gilje Jaatun

CloudCom'09

What did they do?

- A brief survey on recent work on privacy and trust for the semantic web, and sketch a middleware solution for privacy protection that leverages probabilistic methods for automated trust and privacy management for the semantic web

Trust Management

- Definition of trust
  - The willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor and control that other party.

Trust Management (cont.)

- Trust Models
  - Mayer, R., Davis, J., Schoorman, F.: An integrative model of organizational trust. Academy of Management Review
    - The main factors of trustworthiness were identified as ability, benevolence and integrity.
    - On the trustor's part, disposition to trust and perceived risk were identified as the most influential factors with regards to trust.
    - Furthermore, the outcome of a trust relation (experience) is assumed to influence one or more of the trustworthiness factors and hence the trustworthiness of the trustee.

Trust Management (cont.)

- Trust Models
  - The complexity of several proposed models does not necessarily give better trust assessments
  - Conrad, M., French, T., Huang, W., Maple, C.: A lightweight model of trust propagation in a multi-client network environment: to what extent does experience matter?
    - Proposed a lightweight model for trust propagation. The parameters self confidence, experience, hearsay and prejudice are used to model and assess trust. This computational model also allows agents to compute a trust value to automatically perform trust decisions.

Trust Management (cont.)

- Trust Models
  - Gil, Y., Artz, D.: Towards content trust of web resources
    - The idea is to arrive at content trust, where the information itself is used for trust calculation.
    - This allows for a whole new range of parameters (such as bias, criticality, appearance, etc.) to be used when assessing trust in resources.
    - The problem of such parameters is that they require user input, which conflicts with the assumption of agents conducting the assessment autonomously.

Trust Management (cont.)

- Trust Propagation
  - Golbeck, J., Hendler, J.: Accuracy of metrics for inferring trust and reputation in semantic web-based social networks
    - Inferring trust and reputation in social networks when entities are not connected directly by a trust relationship.
    - Done by computing the weighted distance from the source to the sink.
    - Any distrusted entity is not included in the computation since the trust assessments done by such entities are worthless.

Trust Management (cont.)

- Trust Propagation
  - Guha, R., Kumar, R., Raghavan, P., Tomkins, A.: Propagation of trust and distrust
    - Introduce the notion of distrust to address the problem of expressing explicit distrust as a contrast to the absence of trust.
    - Absence of trust may come from lack of information to conduct a proper trust assessment, while distrust expresses that a proper assessment have been conducted and that the entity should not be trusted.
    - Furthermore, they argue that distrust could also be propagated and proposes several propagation models in addition to trust transitivity, including co-citation, which is extensively used for web searches.

Trust Management (cont.)

- Trust Propagation
  - Huang, J., Fox, M.S.: An ontology of trust: formal semantics and transitivity
    - claim that not all kinds of trust can be assumed to be transitive.
    - They note that trust based on performance, i.e. an entity performing as expected repeatedly, is not necessarily transitive, while trust based on a belief that the entity will perform as expected often is.

Probabilistic Privacy Policy Enforcement

- A probabilistic approach to policy enforcement, where users are given a probability that their requirements will be respected and polices enforced.
- Thus when interacting with websites who are known to be less trustworthy, policy adherence is given by a probability metric that the website will actually enforce its own policies.
- This enforcement model does not include a privacy or trust model
    - i.e. it is only occupied with how to handle uncertainty in enforcement and provide a tool for interacting with non-conforming entities while minimising the risks involved.

Probabilistic Privacy Policy Enforcement (cont.)

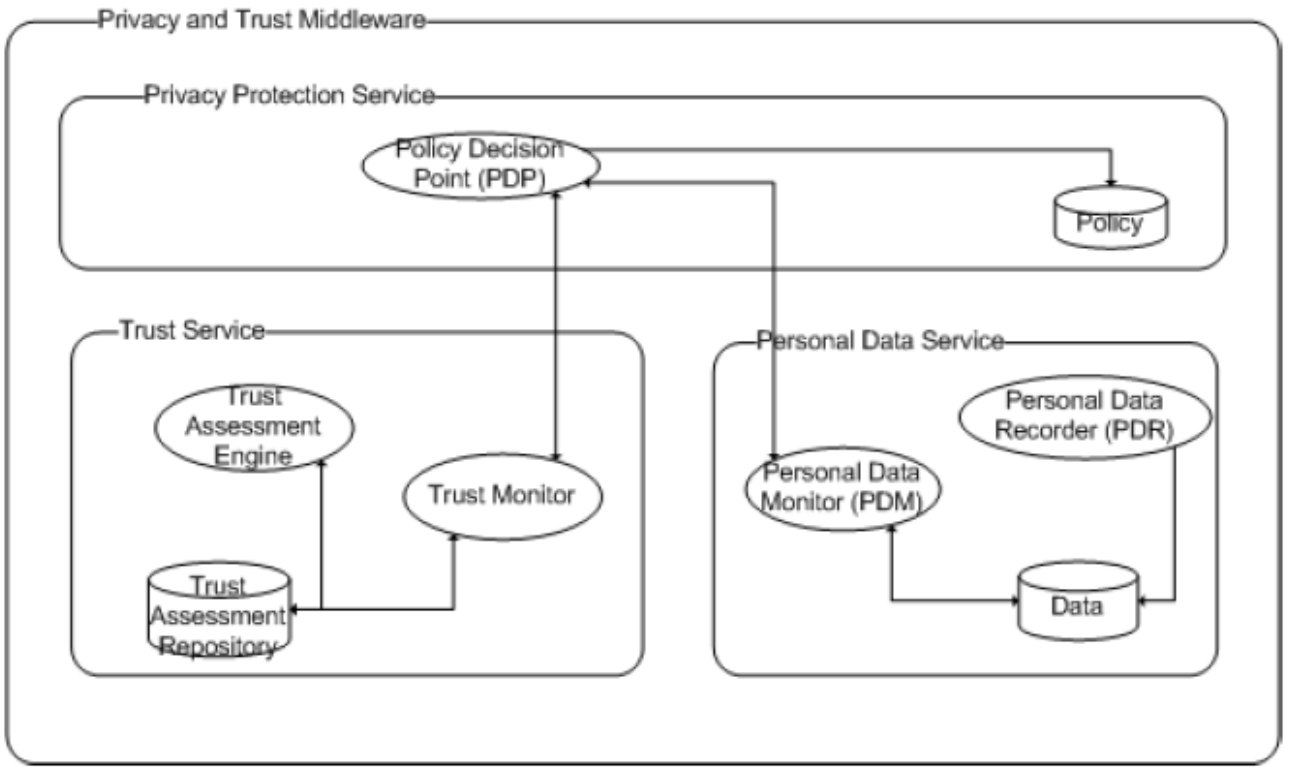


Fig. 1. Middleware architecture for probabilistic privacy management

Probabilistic Privacy Policy Enforcement (cont.)

- Personal Data Recorder (PDR)
  - Protecting users from this kind of aggregation requires complete control of what information has been distributed and to whom.
  - Records what data is transmitted to which receivers.
    - Example: Consider the situation where a user wanting to stay unidentified has provided his postal code and anonymous e-mail address to a website. Later he also provides age and given name (not the full name) and the anonymous e-mail address. Now, the website is able to combine the data (postal code, age and given name) to identify the anonymous user
      - The second interaction with the website should have been blocked, since it enables the website to reveal the user's identity. The PDR allows the user to view

Probabilistic Privacy Policy Enforcement (cont.)

- Personal Data Monitor (PDM)
  - Computing and assessing policies and behaviour, and to update the personal data recorder with inferred knowledge.
  - Determine the likelihood that the personal information distributed to the receiver will also reach other.
    - Example: sending an e-mail with a business proposition to a specific employee of a company, it is likely that other employees in that company also will receive the e-mail (e.g. his superior).
    - PDM is responsible for inferring other recipients and to include such information in the Personal Information Base.
    - Hence, any interaction later on should consider this information when assessing the kind of information to

Probabilistic Privacy Policy Enforcement (cont.)

- Trust Assessment Engine (TAE)
  - Calculating trust values of different entities in order to determine their trustworthiness.
  - The TAE is focused solely on assessing communicating parties and does not take into account risk willingness, vulnerability and criticality.

Probabilistic Privacy Policy Enforcement (cont.)

- Trust Monitor (TM)
  - Detecting events that might affect the perceived trustworthiness and the willingness to take risks.
  - Calculating and deciding on what is an acceptable trust level, given the circumstances.
  - Any computed trust value and feedback received from cooperating entities is stored in the trust assessment repository

Probabilistic Privacy Policy Enforcement (cont.)

- Policy Decision Point (PDP)
  - The final decision on whether to engage in information exchange and if so; under what conditions.
  - Collects the views of both the TM and the PDM and compares their calculations to the policies and requirements found in the policy repository.
  - The decision is reported back to the TM and PDM to allow recalculation in case the decision alters the calculated trust values or distribution of personal information

Towards an Approach of Semantic Access Control for Cloud Computing

Luokai Hu, Shi Ying, Xiangyang Jia, and Kai Zhao

CloudCom'09

What did they do?

- Analysis existing access control methods and present a new Semantic Access Control Policy Language (SACPL) for describing Access Control Policies (ACPs) in cloud computing environment.
- Access Control Oriented Ontology System (ACOOS) is designed as the semantic basis of SACPL.
- Ontology-based SACPL language can effectively solve the interoperability issue of distributed ACPs.

Access Control Oriented Ontology System (ACOOS)

- Provide the common understandable semantic basis for access control in cloud computing environments.
- Divided into four parts, Subject Ontology, Object Ontology, Action Ontology and Attribute Ontology
- Web Ontology Language (OWL) is selected as the modeling language of ACOOS.
  - Ontology is helpful to construct authorization policy within the scope of whole cloud computing environment based on policy definition elements with determined semantics.

Access Control Oriented Ontology System (ACOOS)

- Subject Ontology
  - Subject is the entity that has a number of action permissions over object.
    - e.g., a user, a user group, an organization, a role, a process, a service
  - Attribute of a subject is described by the data property
  - The role in subject ontology represents the capability of a subject to implement a task.
  - Access permission of resources can be encapsulated in the role.
    - If a subject is assigned to a role, it can access the resources indirectly.

Access Control Oriented Ontology System (ACOOS)

- Object Ontology
  - Object is the entity as receptor of action and is need for protection.
    - e.g., data, documents, services and other resources.
  - Attribute of an object is described by the data property and object property of OWL with hasObjectDataAttribute and hasObjectAttribute respectively.
  - Object group can also be used to define the rule to organize objects.
    - Each object group in fact establishes a new object concept, all object individuals of the object concept have object attribute values of the object group.

Access Control Oriented Ontology System (ACOOS)

- Action Ontology
  - With the cloud computing technology, usually a large number of subjects and objects but only a relatively small number of actions could be found
    - e.g., such as reading, writing and execution
  - Action also has properties, known as the ActionAttribute, which describes various information of action for authorization and management.
  - Action group can be defined with helpful for the definition of rules.
    - The definition of action group, nearly the same with the object group, will not repeat it again.

Access Control Oriented Ontology System (ACOOS)

- Attribute Ontology
  - Attribute types are defined in the attribute ontology, can be used to define the attribute of almost all entities, including the subject, object and action.
  - The attribute value of entities is often needed to determine whether meet the Permit conditions or Deny ones.

Semantic Access Control Policy Language (SACPL)

- Policy markup language, such as XACML, supports description and management of distributed policies.
- The ACP of an object (resource) may be completed by a number of departments even organizations, such as information systems department, human resources and financial department.
- The same ACP may be applied to the internal network protection, e-mail system, remote access systems, or a cloud computing platform.
- As a result, in cloud computing environment, the issue of interoperability among policies is more important than ever before.

References

1. NIST (Authors: P. Mell and T. Grance), "The NIST Definition of Cloud Computing (ver. 15)," National Institute of Standards and Technology, Information Technology Laboratory (October 7 2009).
2. J. McDermott, (2009) "Security Requirements for Virtualization in Cloud Computing," presented at the ACSAC Cloud Security Workshop, Honolulu, Hawaii, USA, 2009.
3. J. Camp. (2001), "Trust and Risk in Internet Commerce," MIT Press
4. T. Ristenpart et al. (2009) "Hey You Get Off My Cloud," Proceedings of the 16th ACM conference on Computer and communications security, Chicago, Illinois, USA
5. Security and Privacy in Cloud Computing, Dept. of CS at Johns Hopkins University. www.cs.jhu.edu/~ragib/sp10/cs412
6. Cloud Security and Privacy: An Enterprise Perspective on Risks and Compliance by Tim Mather and Subra Kumaraswamy
7. Afraid of outside cloud attacks? You're missing the real threat. http://www.infoworld.com/d/cloud-computing/afraid-outside-cloud-attacks-youre-missing-real-threat-894
8. **Amazon downplays report highlighting vulnerabilities in its cloud service. http://www.computerworld.com/s/article/9140074/Amazon_downplays_report_highlighting_vulnerabilities_in_its_cloud_service**
9. **Targeted Attacks Possible in the Cloud, Researchers Warn. http://www.cio.com/article/506136/Targeted_Attacks_Possible_in_the_Cloud_Researchers_Warn**
10. **Vulnerability Seen in Amazon's Cloud-Computing by David Talbot. http://www.cs.sunysb.edu/~sion/research/sion2009mitTR.pdf**
11. **Cloud Computing Security Considerations by Roger Halbheer and Doug Cavit. January 2010. http://blogs.technet.com/b/rhalbheer/archive/2010/01/30/cloud-security-paper-looking-for-feedback.aspx**
12. **Security in Cloud Computing Overview.http://www.halbheer.info/security/2010/01/30/cloud-security-paper-looking-for-feedback**
13. **Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds by T. Ristenpart, E. Tromer, H. Shacham and Stefan Savage. CCS'09**
14. **Cloud Computing Security. http://www.exforsys.com/tutorials/cloud-computing/cloud-computing-security.html**
15. **Update From Amazon Regarding Friday's S3 Downtime by Allen Stern. Feb. 16, 2008. http://www.centernetworks.com/amazon-s3-downtime-update**
16. R. Ranchal, B. Bhargava, L.B. Othmane, L. Lilien, A. Kim, M. Kang, "Protection of Identity Information in Cloud Computing without Trusted Third Party," Third International Workshop on Dependable Network Computing and Mobile Systems (DNCMS) in conjunction with 29th IEEE Symposium on Reliable Distributed System (SRDS) 2010
17. P. Angin, B. Bhargava, R. Ranchal, N. Singh, L. Lilien, L.B. Othmane, "A User-Centric Approach for Privacy and Identity Management in Cloud Computing," 29th IEEE Symposium on Reliable Distributed System (SRDS) 2010
18. H. Khandelwal, *et al.*, "Cloud Monitoring Framework," Purdue University. Dec 2010.

Other References for Cloud Security

- M. Armbrust, *et al.*, "Above the Clouds: A Berkeley View of Cloud Computing," UC Berkeley Reliable Adaptive Distributed Systems LaboratoryFebruary 10 2009.
- Cloud Security Alliance, "Security Guidance for Critical Areas of Focus in Cloud Computing, ver. 2.1," 2009.
- M. Jensen, *et al.*, "On Technical Security Issues in Cloud Computing," presented at the 2009 IEEE International Conference on Cloud Computing, Bangalore, India 2009.
- P. Mell and T. Grance, "Effectively and Securely Using the Cloud Computing Paradigm," ed: National Institute of Standards and Technology, Information Technology Laboratory, 2009.
- N. Santos, *et al.*, "Towards Trusted Cloud Computing," in *Usenix 09 Hot Cloud Workshop*, San Diego, CA, 2009.
- R. G. Lennon, *et al.*, "Best practices in cloud computing: designing for the cloud," presented at the Proceeding of the 24th ACM SIGPLAN conference companion on Object oriented programming systems languages and applications, Orlando, Florida, USA, 2009.
- P. Mell and T. Grance, "The NIST Definition of Cloud Computing (ver. 15)," National Institute of Standards and Technology, Information Technology LaboratoryOctober 7 2009.
- C. Cachin, *et al.*, "Trusting the cloud," *SIGACT News,* vol. 40, pp. 81-86, 2009.
- J. Heiser and M. Nicolett, "Assessing the Security Risks of Cloud Computing," Gartner 2008.
- A. Joch. (2009, June 18) Cloud Computing: Is It Secure Enough? *Federal Computer Week*.
- AWS Amazon EC2: http://aws.amazon.com/ec2/
- Amazon CloudWatch: http://aws.amazon.com/cloudwatch/
- Iperf: http://iperf.sourceforge.net/

Cloud Blind References

- L. Ran, A. Helal, and S. Moore, "Drishti: An Integrated Indoor/Outdoor Blind Navigation System and Service," 2nd IEEE Pervasive Computing Conference (PerCom 04).
- S.Willis, and A. Helal, "RFID Information Grid and Wearable Computing Solution to the Problem of Wayfinding for the Blind User in a Campus Environment," IEEE International Symposium on Wearable Computers (ISWC 05).
- Y. Sonnenblick. "An Indoor Navigation System for Blind Individuals," Proceedings of the 13th Annual Conference on Technology and Persons with Disabilities, 1998.
- J. Wilson, B. N. Walker, J. Lindsay, C. Cambias, F. Dellaert. "SWAN: System for Wearable Audio Navigation," 11th IEEE International Symposium on Wearable Computers, 2007.
- J. Nicholson, V. Kulyukin, D. Coster, "ShopTalk: Independent Blind Shopping Through Verbal Route Directions and Barcode Scans," The Open Rehabilitation Journal, vol. 2, 2009, pp. 11-23.
- Bach-y-Rita, P., M.E. Tyler and K.A. Kaczmarek. "Seeing with the Brain," International Journal of Human-Computer Interaction, vol 15, issue 2, 2003, pp 285-295.
- Y.K. Kim, K.W. Kim, and X.Yang, "Real Time Traffic Light Recognition System for Color Vision Deficiencies," IEEE International Conference on Mechatronics and Automation (ICMA 07).
- R. Charette, and F. Nashashibi, "Real Time Visual Traffic Lights Recognition Based on Spot Light Detection and Adaptive Traffic Lights Templates," World Congress and Exhibition on Intelligent Transport Systems and Services (ITS 09).
- A.Ess, B. Leibe, K. Schindler, and L. van Gool, "Moving Obstacle Detection in Highly Dynamic Scenes," IEEE International Conference on Robotics and Automation (ICRA 09).
- P. Angin, B. Bhargava, R. Ranchal, N. Singh, L. Lilien, L. B. Othmane, "A User-centric Approach for Privacy and Identity Management in Cloud Computing," submitted to SRDS 2010.