University at Buffalo
The State University of New York

UbiSeC Lab
Ubiquitous Security & PrivaCy
Research Laboratory

Some Security Challenges of Cloud Computing

Kui Ren

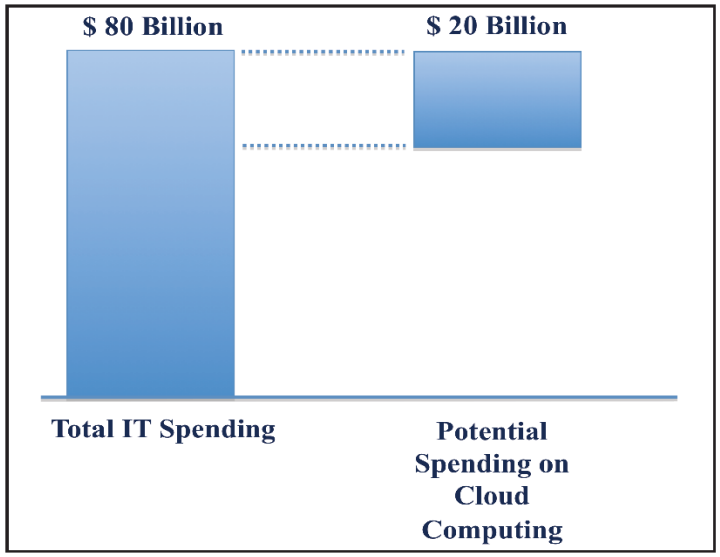
Associate Professor

Department of Computer Science and Engineering

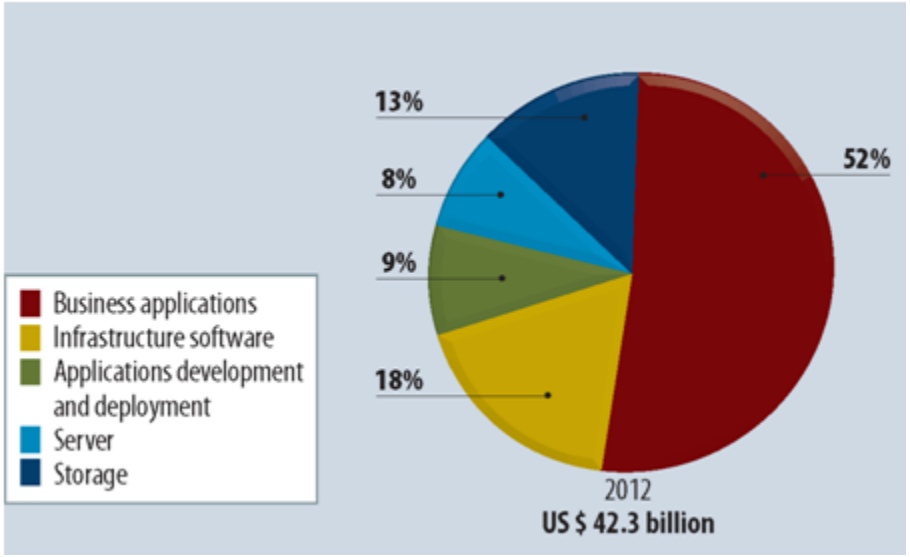
SUNY at Buffalo

Cloud Computing: the Next Big Thing

- Tremendous momentum ahead:

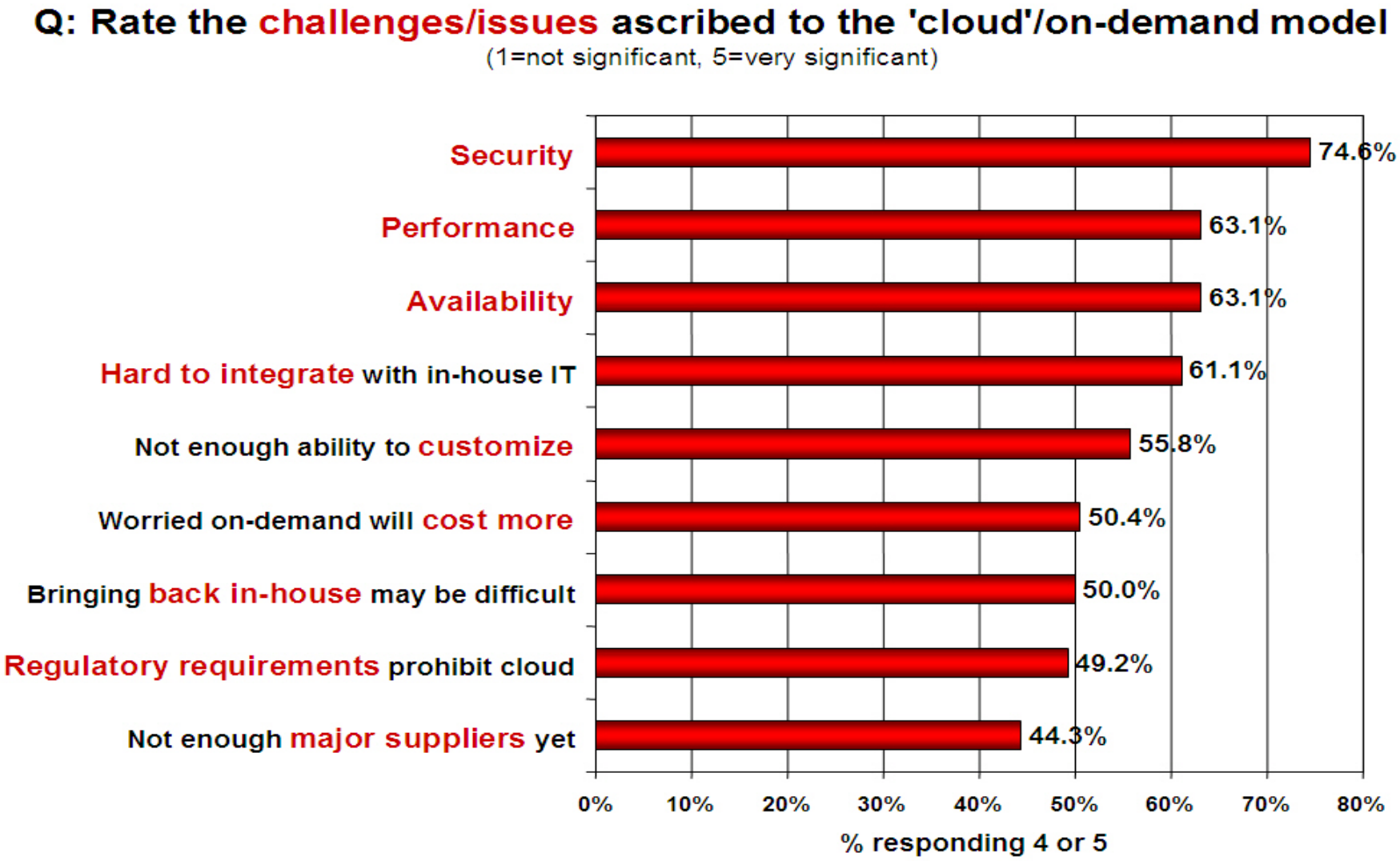


Prediction on Federal IT spendable to move to the cloud from US CIO.Gov in Feb. 2011.



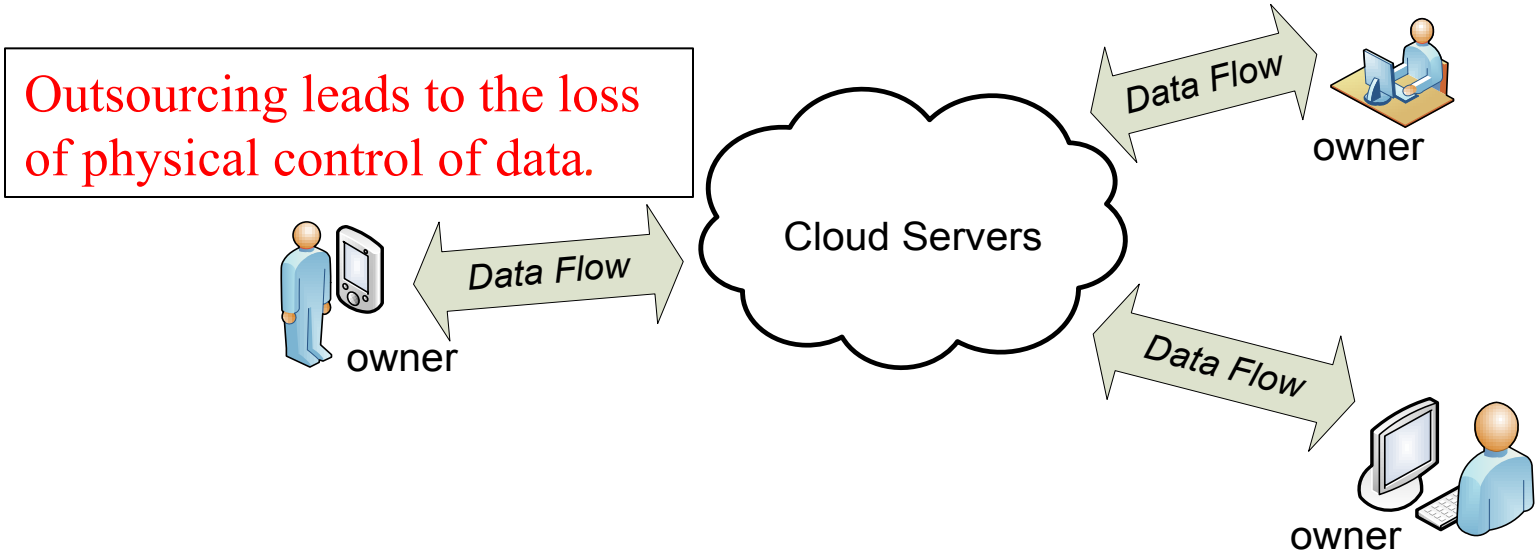
Prediction on cloud computing revenue in 2012 from Market-research firm IDC.

Challenges for Cloud Computing



Source: IDC Enterprise Panel, August 2008 n=244

I: Storage Outsourcing vs. Storage Security



- Cloud storage service allows owners to outsource their data to cloud servers for storage and maintenance.
  - Low capital costs on hardware and software, low management and maintenance overheads, universal on-demand data access, etc.
  - E.g., Amazon S3.
- However, data outsourcing also eliminates owners' ultimate control over their data.

Storage Outsourcing vs. Storage Security

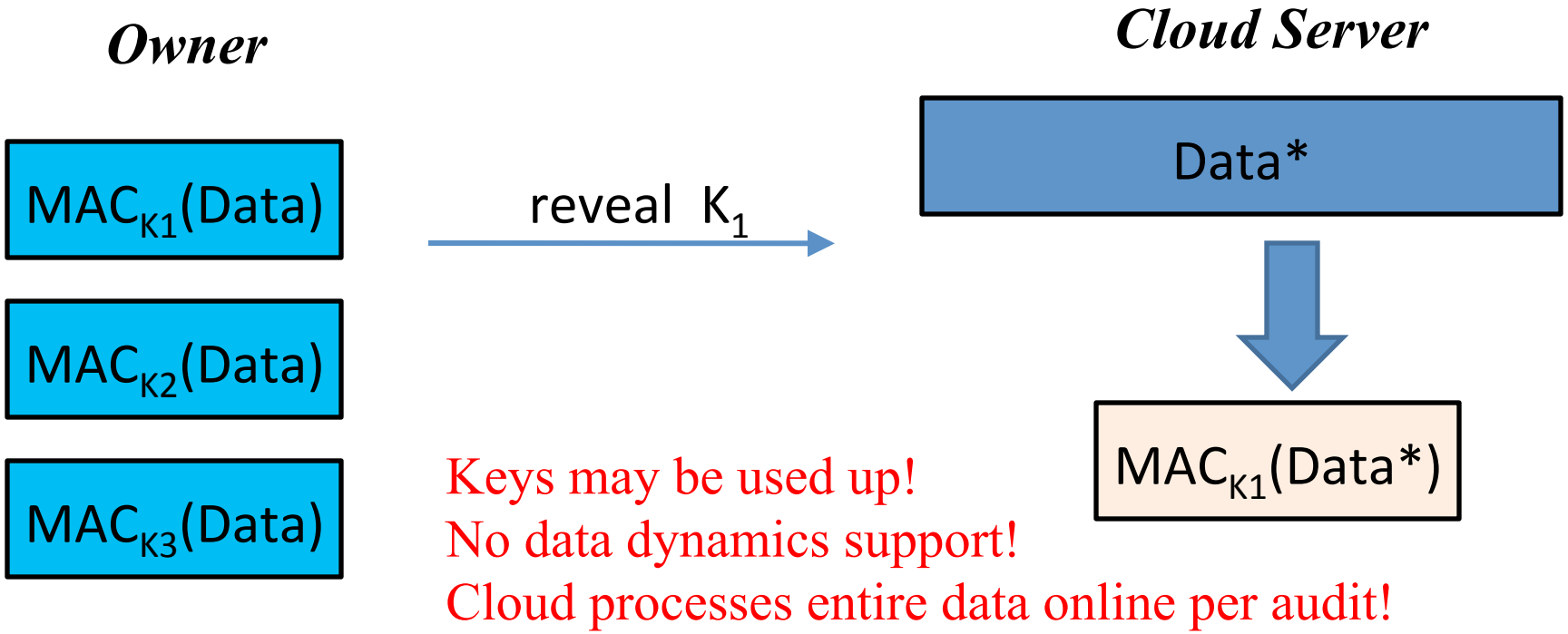
- Cloud currently offers no guarantee:
  - Amazon S3: not liable to any data damages or data loss.
- Broad range of threats for data integrity do exist:
  - Internal: Byzantine failure, management errors, software bugs, etc.
  - External: malicious malware, economically motivated attacks, etc.
  - E.g., Amazon S3 - Feb., Jul. 2008; Gmail - Dec. 2006, Mar. 2011; Apple MobileMe - Jul. 2008, Hotmail – Dec. 2010, …
- Cloud servers might behave unfaithfully:
  - Discard rarely accessed data for monetary reason
  - Hide data loss incidents for reputation
- Data owners demands continuous storage correctness assurance for their data in the cloud.

Challenges for Storage Integrity Auditing

- How to enable data owners to audit cloud for correctly storing their outsourced data (through their mobile devices).
  - Function effectively without requiring local data copies.
    - Traditional methods for storage security can not be directly adopted.
    - Retrieving massive data for checking is impractical, i.e., overwhelming bandwidth cost.
  - Cope with cloud data dynamic updates
    - Changes due to the underlying applications
  - Communication and computation efficiency
    - allows mobile devices to perform the auditing tasks

Straightforward Approaches

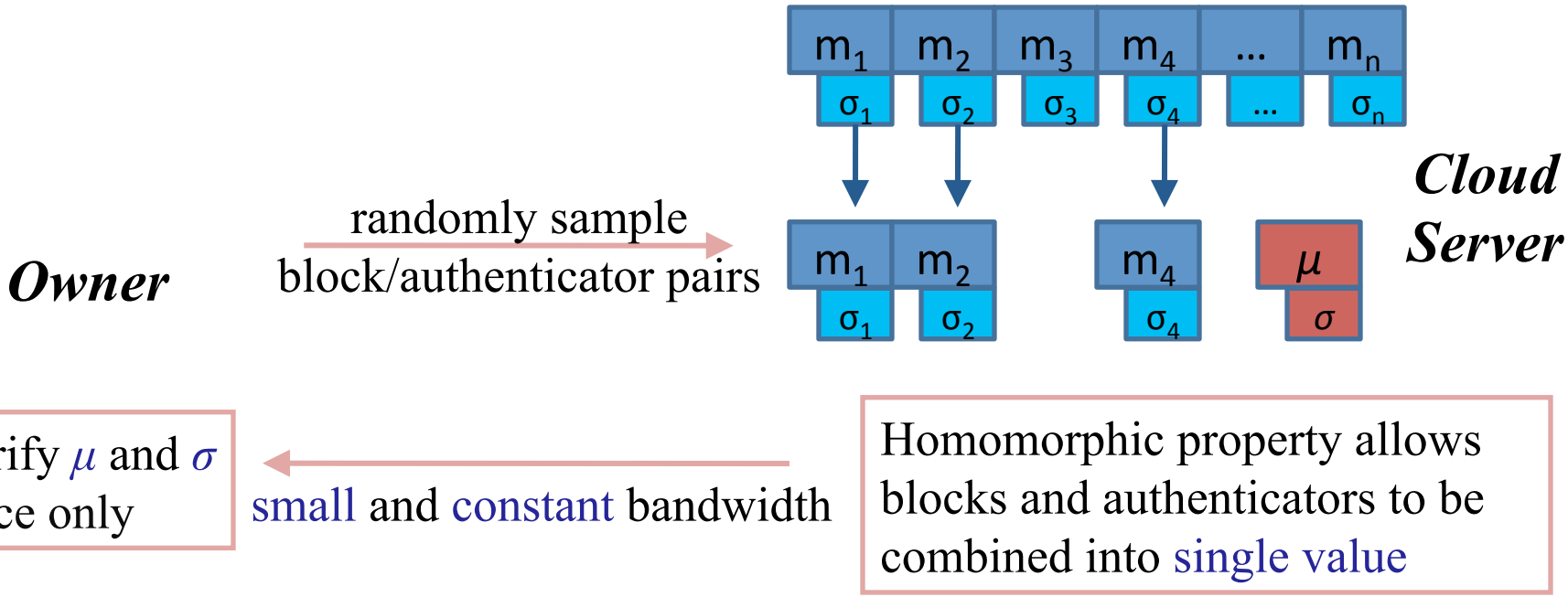
- The traditional approach is not applicable.
  - Owner pre-computes MACs for the data.



equal?

The Solution Direction

- Audit the aggregated block and authenticator for the constant bandwidth cost and much saved computational cost.



More storage security challenges

- Storage integrity audition
  - Support for data dynamic updates
  - Support for public audition
  - Efficiency optimization

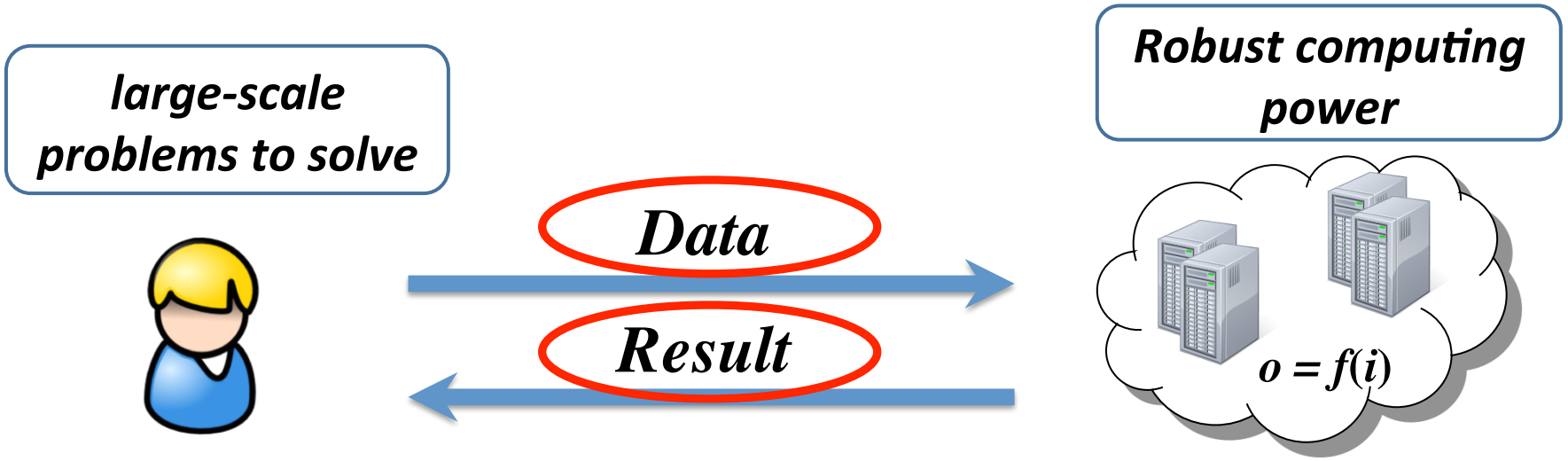
More storage security challenges

- Storage integrity audition
  - Support for data dynamic updates
  - Support for public audition
  - Efficiency optimization
- Proof of storage geolocation
  - Within geographic boundaries
  - On the same physical machines
- Assured data deletion
- Proof of ownership
- Proof of encryption
- Secure data deduplication
- …

II: Computation Outsourcing vs. Data Security

- Cloud provides robust and elastic computational power at a reduced cost.
- Computationally weak end-users can leverage the abundant cloud resources to solve large-scale problems.
  - Massive computational power easily utilized in a pay-per-use manner.
  - Large-scale optimization problems
  - Genomic computation problems
  - Program execution problems
- Given the inherent security obstacles, how could this computation outsourcing paradigm become practical?

Computation Outsourcing vs. Security



- End-users rely on cloud to perform computations over their data.
- But sensitive data protection can be a must

Computation Outsourcing vs. Security

Result correctness has no assurance.

- Cloud may not be fully trusted for computation result correctness.
  - — Software bugs, hardware failures, or outsider attacks
  - — Intentionally being lazy to save cost for monetary reasons

Towards Secure Computation Outsourcing

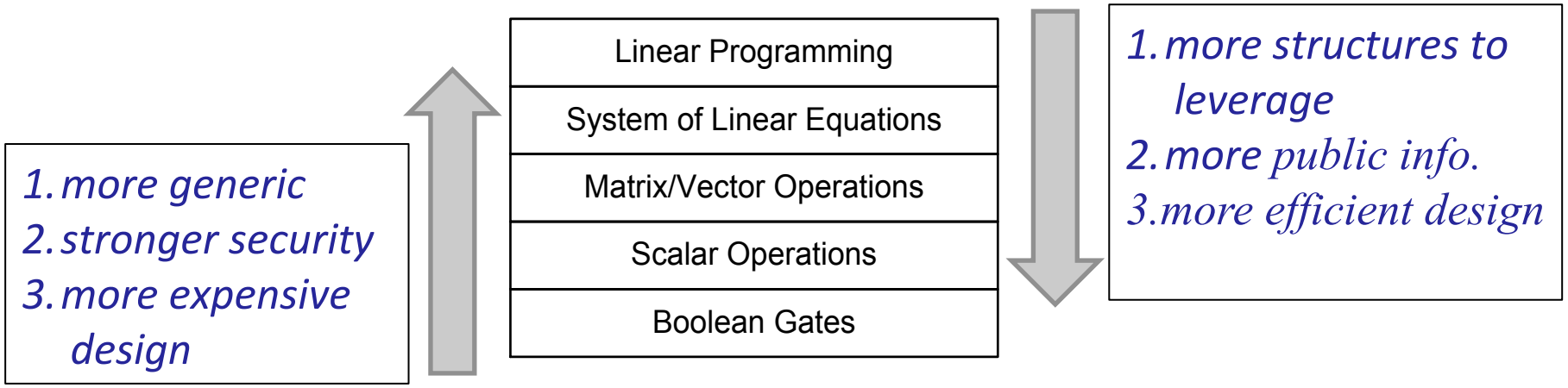
- Enabling a trustworthy computing environment
  - e.g, trusted hypervisor boot/VM launch, strong isolation, continuous/dynamic attestation of platform, …
    - Making interesting progress, but critical challenges remain.
- Computing over encrypted data
  - bottom-up approach
  - top-down approach

Towards Secure Computation Outsourcing

- Computing over encrypted data
  - Bottom-up approach: interpret all computations as low-level circuit evaluations + FHE
    - Theoretically feasible for arbitrary computations, but impractical
      - e.g., cost of a Google search with FHE ~= **10^{12} ×** cost of today.
  - Top-down approach: treat different computations individually + leverage their resp. characteristics
    - achieve desirable tradeoffs for security, efficiency, functionality
      - e.g., data mining, data search, engineering computation, …

Design Methodology

- Systematically exploit security/efficiency tradeoffs
  - interpret optimization computations at different abstraction levels organized in a hierarchy



one exemplary hierarchy of computations

e.g., Gennaro et al. (CRYPTO'10)
Yao's garbled circuits + Gentry's FHE

Secure Cloud Computation: The Case of Linear Programming

- Linear Programming is fundamental to engineering computing/optimizations.
  - widely used in scheduling, assignment, system design..

$$\begin{aligned} \text{minimize} \quad & 2x_1+3x_2+x_3+x_4 \\ \text{subject to} \quad & x_1 + x_2 - x_3 = 9 \\ & x_2 + 2x_4 = 10 \\ & x_1 + x_3 \geq 0 \\ & x_1 + x_2 + x_4 \geq 0 \end{aligned}$$

vector notation

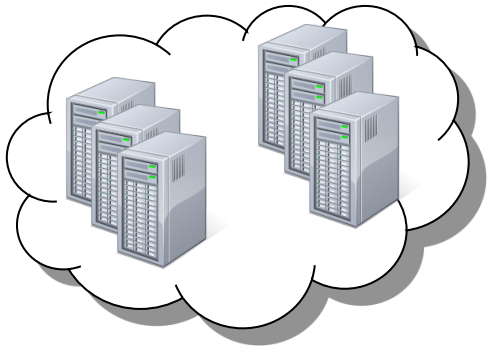
$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{B}\mathbf{x} \geq \mathbf{0}. \end{aligned}$$

Problem Formulation: Secure LP Outsourcing

minimize $\mathbf{c}^T\mathbf{x}$
subject to $\mathbf{Ax} = \mathbf{b}$
$\mathbf{Bx} \geq \mathbf{0}$

$(\mathbf{A}, \mathbf{B}, \mathbf{b}, \mathbf{c})$

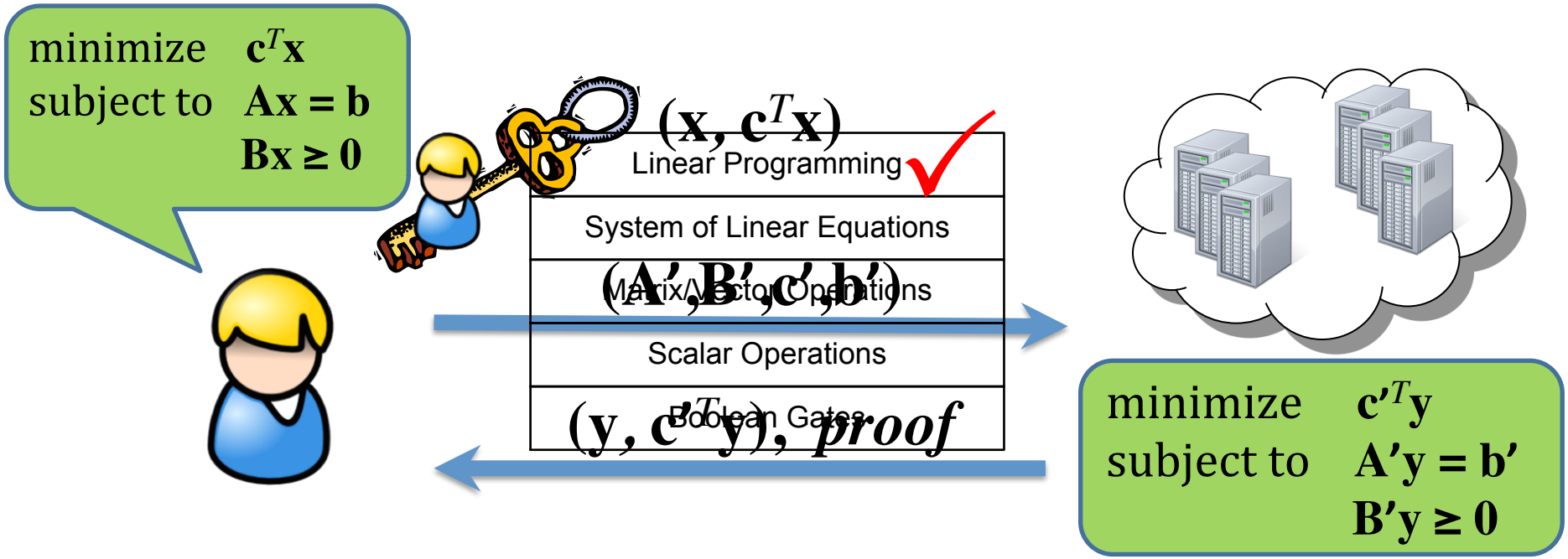
$(\mathbf{x}, \mathbf{c}^T\mathbf{x})$



A	m × n matrix, m < n, with rank m.
B	n × n non-singular matrix
b	m × 1 column vector
c	n × 1 column vector
x	n × 1 vector of decision variables

How to leverage cloud to faithfully solve LP without revealing data/result?

Problem Formulation: Design Overview



- To develop secure and efficient LP transformation techniques
  - Protect original LP via randomly transformed LP problem
  - Cloud solves transformed LP without knowing **A, B, b, c** and **x**.

III: Utility Computing vs. Trustworthy Metering

- In cloud computing, users are charged based on the resources consumed.
  - e.g., Google's AppEngine charges by the number of CPU cycles consumed by a user's application.
- However, the opaqueness of cloud raises concerns on the trustworthiness of the resource metering.
  - Existing research shows possibility for dishonest service providers to easily cheat/over-charge cloud users.
- How do service providers prove or users verify the actual resource consumptions in cloud?

IV: Security Overhead vs. Cloud Benefits

- Security burden may be an overkill to cloud computing.
  - Researchers estimate that using FHE for encrypted Google search would increase the amount of computing time by about 1 trillion.
- The security design also incurs overhead on the end-users, which may conflict with their aims of using the cloud.
  - Cost of many security aspects could offset the economically appealing cloud benefits.
- Can we quantitatively explore the tradeoffs between security overhead and cloud benefits?

Thank you!
Questions?

Research supported by
NSF CAREER and CSR Grants