ARTS Advanced Research and Technology Symposium 2018

The Future of Advanced (Secure) Computing

# Data-Centric Secure Computing

Dr. Emily Shen

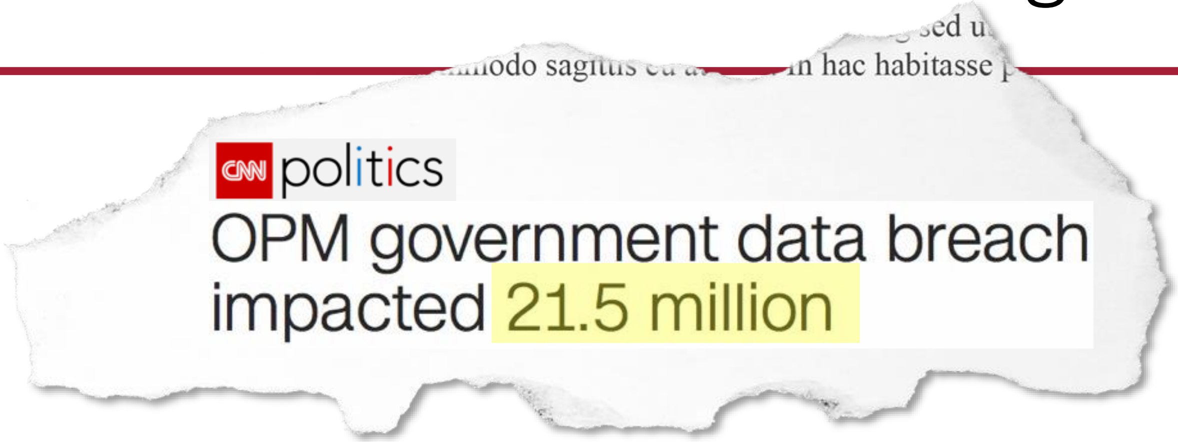
MIT Lincoln Laboratory

5 March 2018

This material is based upon work supported by the Assistant Secretary of Defense for Research and Engineering under Air Force Contract No. FA8721-05-C-0002 and/or FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Assistant Secretary of Defense for Research and Engineering.

LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

# Need for Secure Data Storage

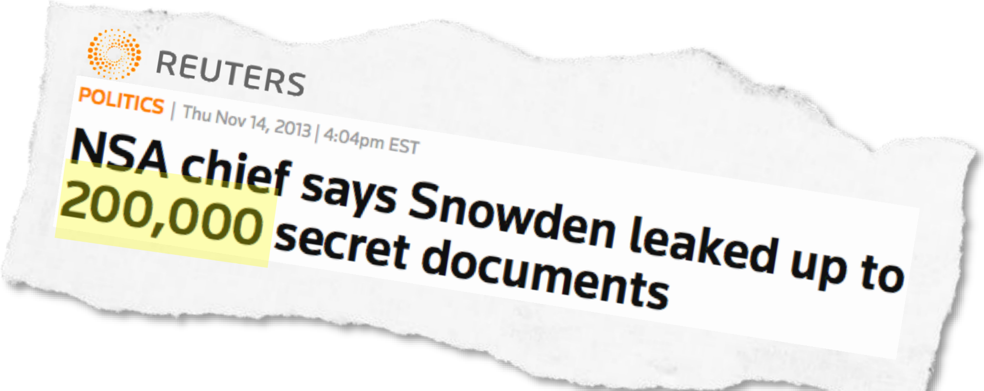


sed u

nodo sagittis eu a in hac habitasse p

CNN politics

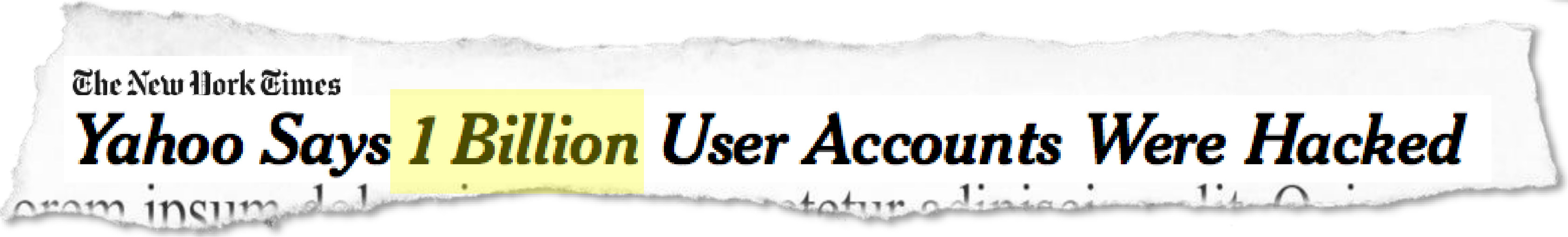
OPM government data breach impacted 21.5 million



REUTERS

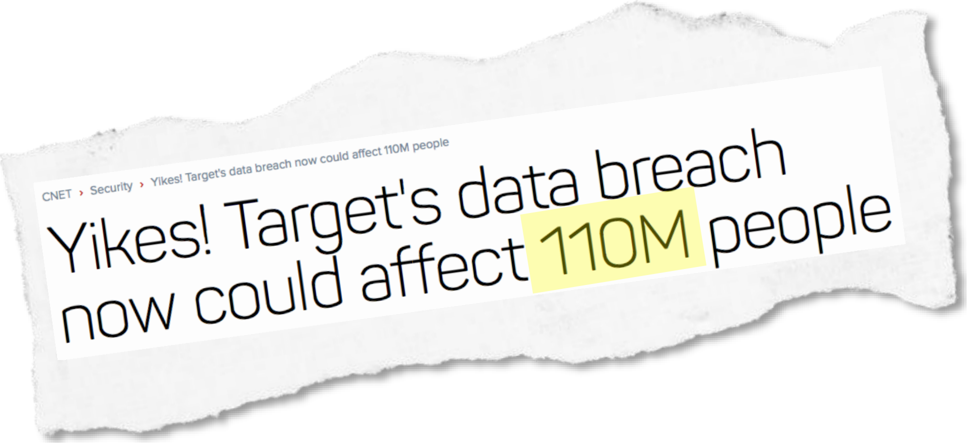
POLITICS | Thu Nov 14, 2013 | 4:04pm EST

NSA chief says Snowden leaked up to 200,000 secret documents



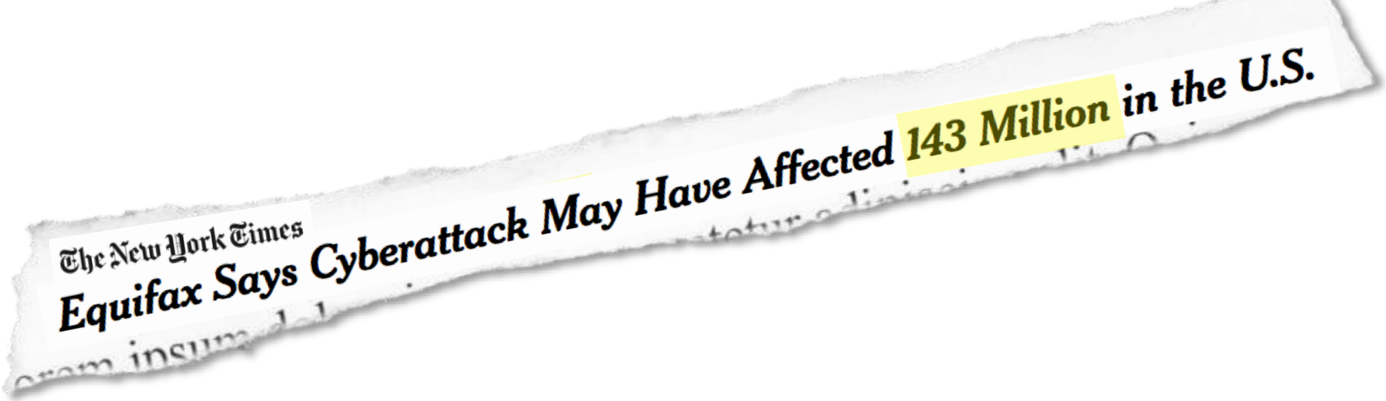
The New York Times

Yahoo Says 1 Billion User Accounts Were Hacked



CNET › Security › Yikes! Target's data breach now could affect 110M people

Yikes! Target's data breach now could affect 110M people



The New York Times

Equifax Says Cyberattack May Have Affected 143 Million in the U.S.

# Need for Secure Computing on Data

**Cloud Computing**

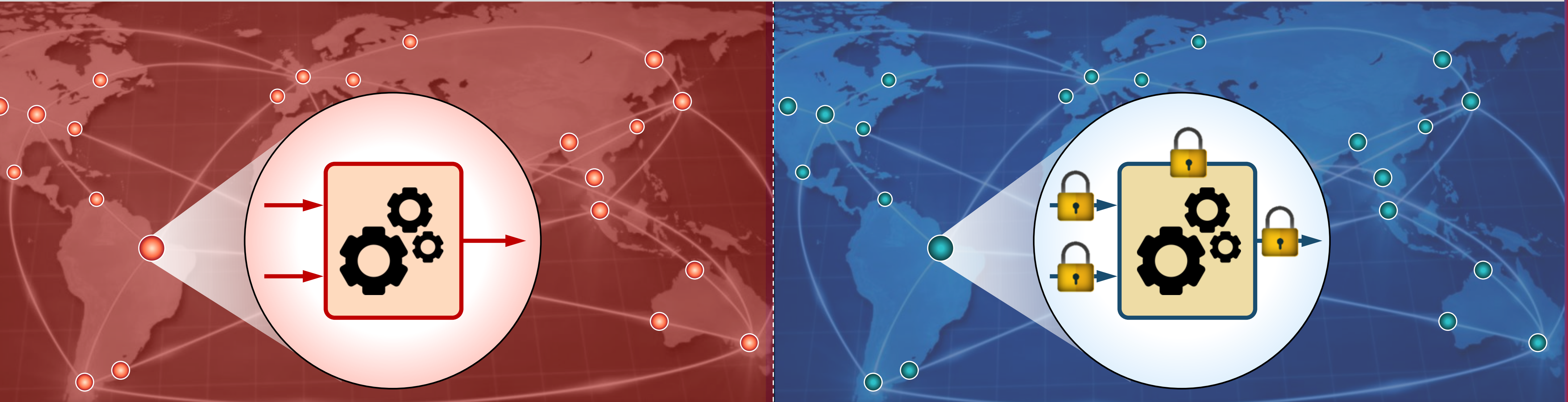
**Internet of Things**

**Cyber Threat Sharing**

**Medical Research**

# Data-Centric Secure Computing

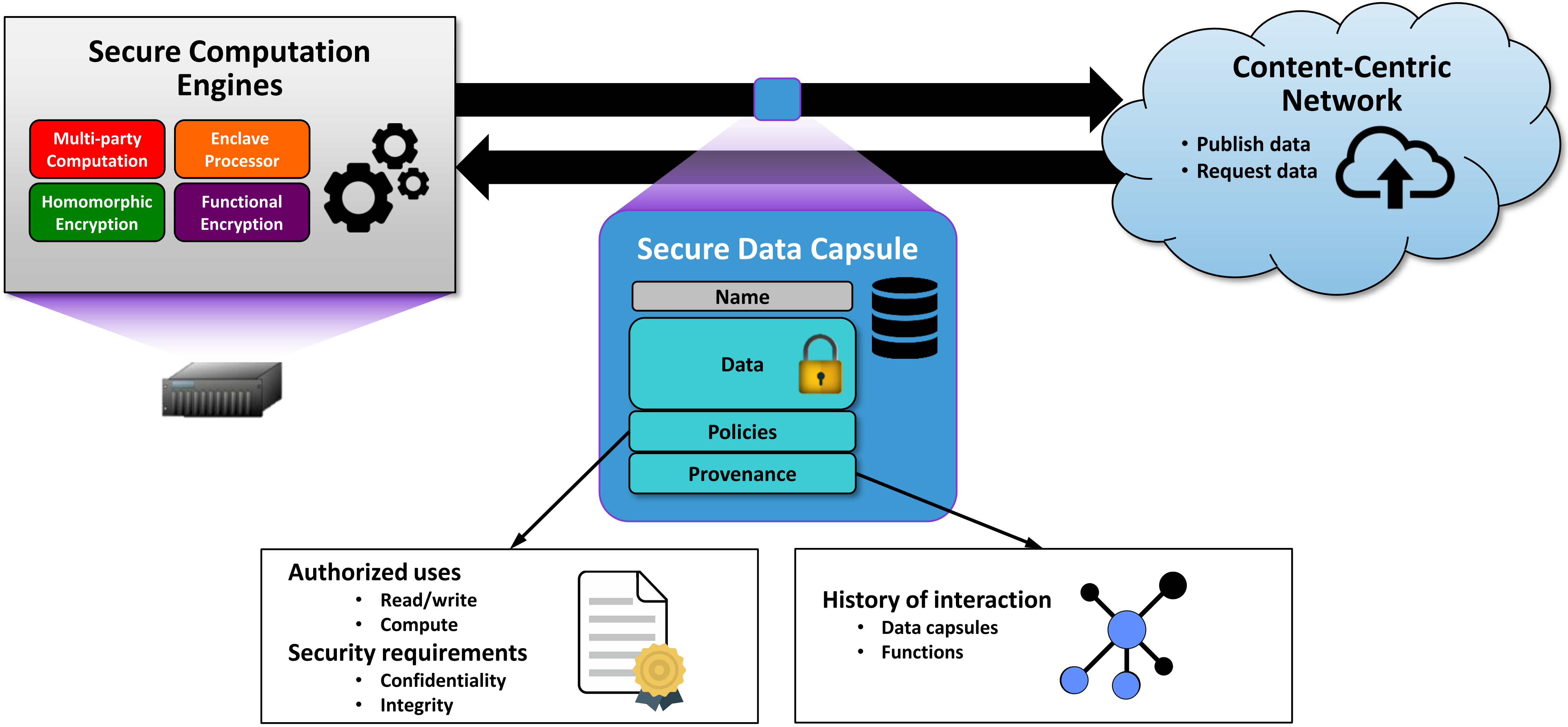
**Vision: Self-protecting data throughout data lifecycle in distributed systems**



| CURRENT: SYSTEM-CENTRIC | | |
|---|---|---|
| Storage (Data at Rest) | Computation (Data in Use) | Networking (Data in Transit) |
| Access control enforced by system; single point of failure | All computation allowed by system | Based on system identifiers (IP addresses) |

| FUTURE: DATA-CENTRIC | | |
|---|---|---|
| Storage (Data at Rest) | Computation (Data in Use) | Networking (Data in Transit) |
| Fine-grained access control self-enforced by data | Only authorized computation allowed by data | Based on data identifiers (content-centric networking) |

# Data-Centric Secure Computing Architecture



## Secure Computation Engines

- Multi-party Computation
- Enclave Processor
- Homomorphic Encryption
- Functional Encryption

## Content-Centric Network

- Publish data
- Request data

## Secure Data Capsule

- Name
- Data
- Policies
- Provenance

**Authorized uses**

- Read/write
- Compute

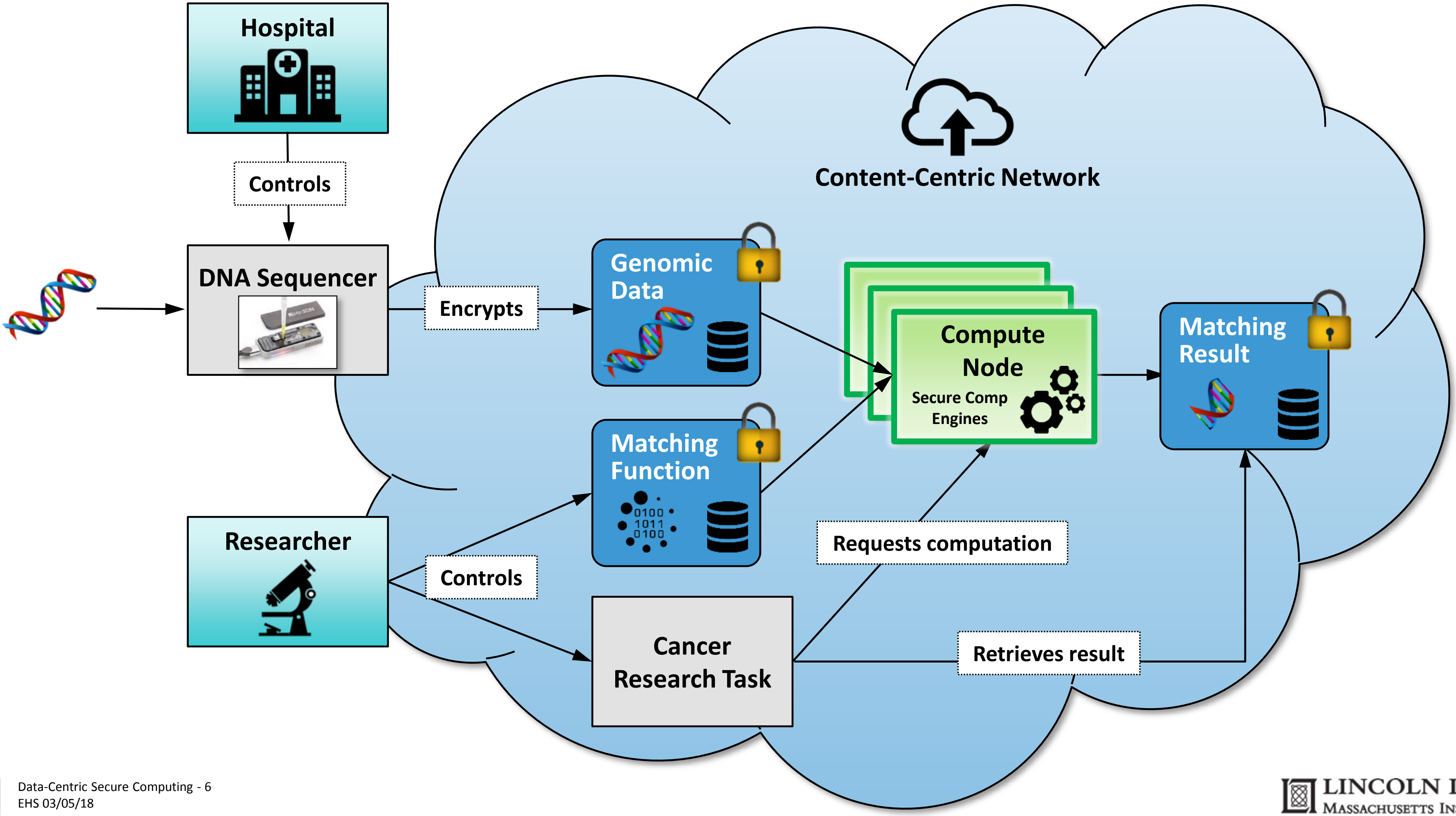
**Security requirements**

- Confidentiality
- Integrity

**History of interaction**

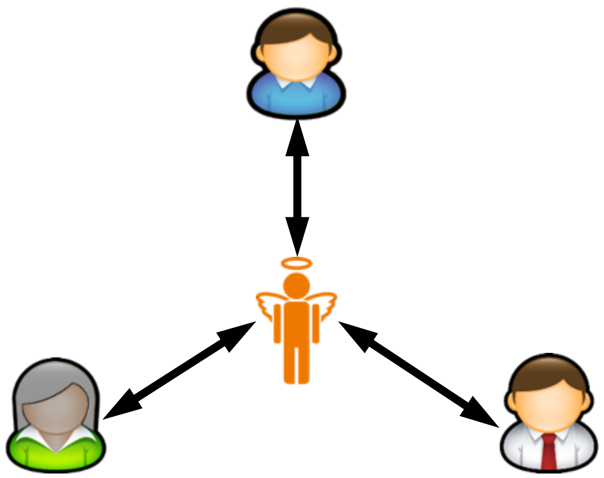
- Data capsules
- Functions

# Data-Centric Secure Computing for Medical Research

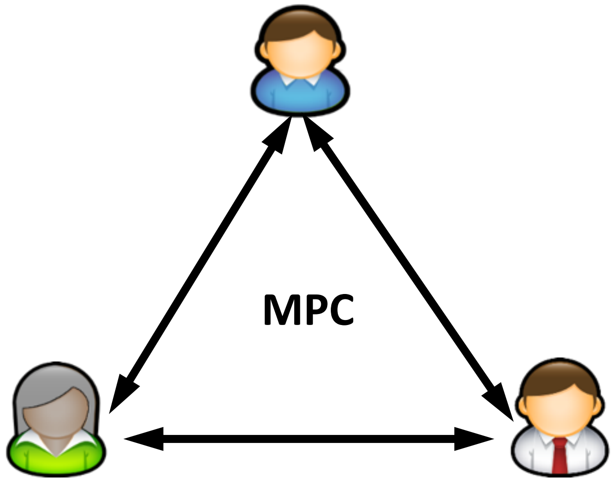


# Secure Computation Example: Multi-Party Computation (MPC)

**Ideal World**



**Real World**

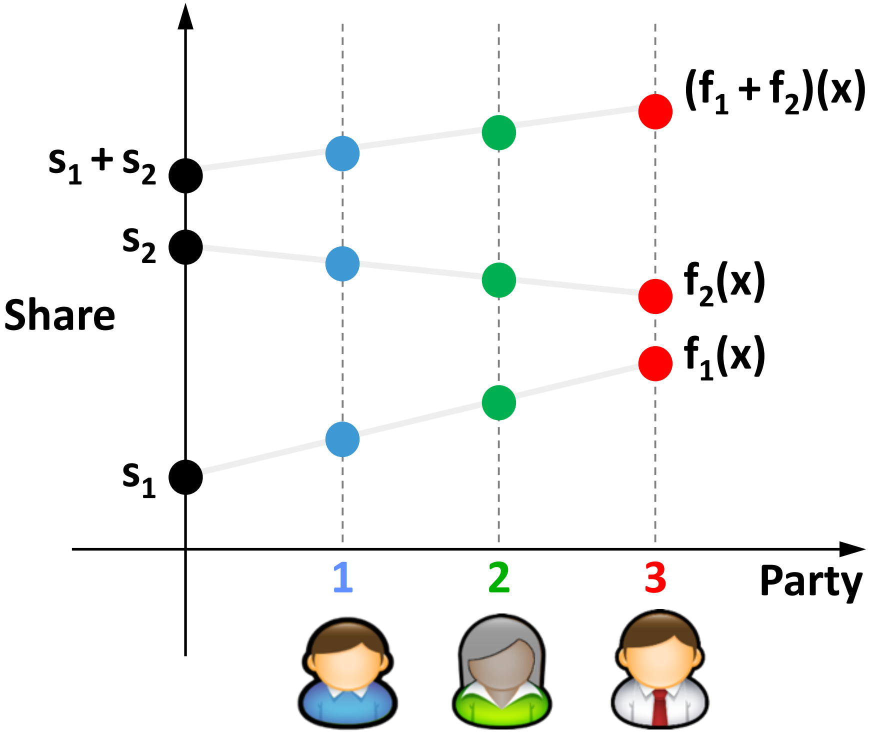


- MPC uses cryptography to emulate functionality and security of a trusted party
  - Confidentiality of inputs and outputs
  - Correctness of computation
  - Resilience to communication/party failures

# MPC Protocols

1. Secret share inputs
   - Each party encodes private data, sends a share to each party
   - Shares completely hiding unless more than t shares are combined
2. Compute on secret shares
   - Addition uses only local computation
   - Multiplication requires communication
3. Open output: Combine final shares to learn result

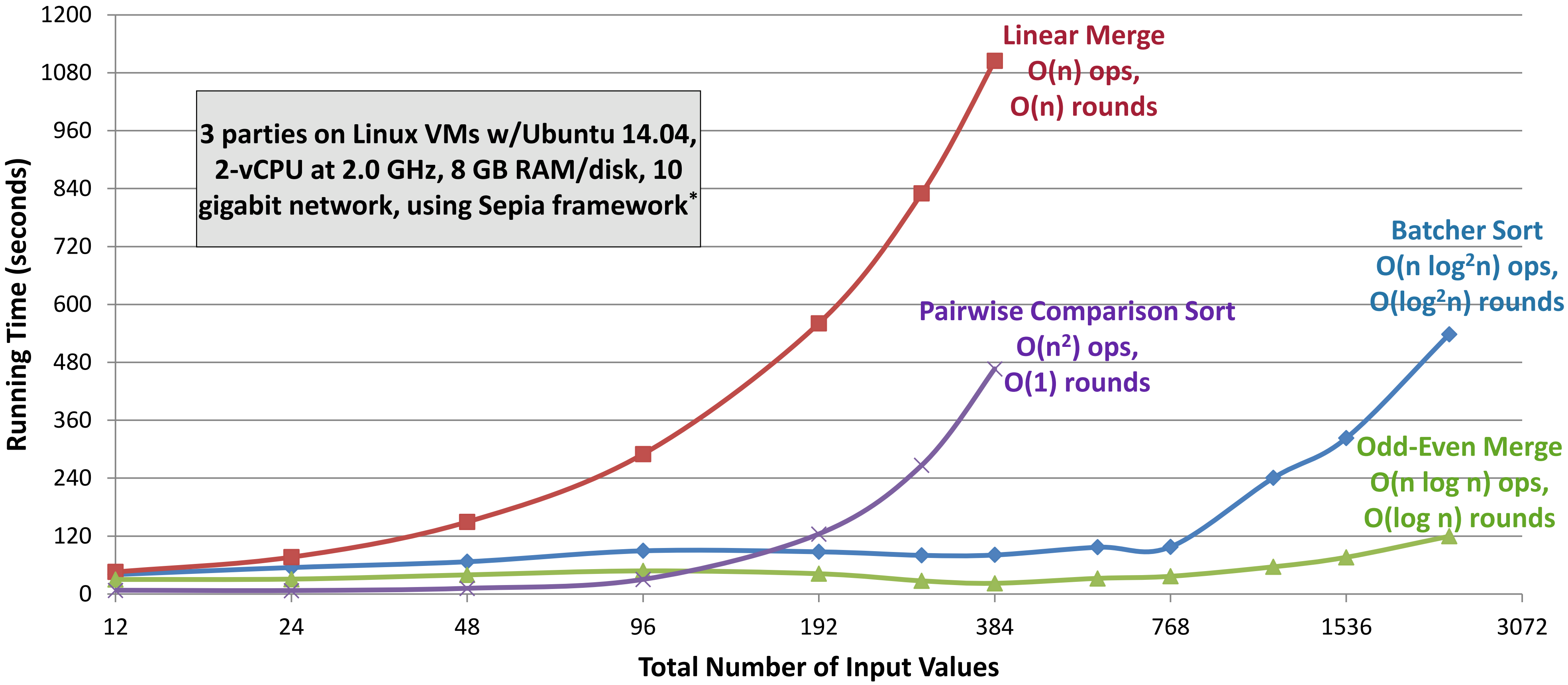
Secret sharing for threshold t = 1



**MPC can compute any arbitrary function securely, can be optimized for specific applications**

BGW – M. Ben-or, S. Goldwasser, A. Wigderson. Completeness Theorems for Non-Cryptographic Fault-Tolerant Distributed Computation. STOC 1988

# Example: Optimizing MPC Sorting Protocols



3 parties on Linux VMs w/Ubuntu 14.04, 2-vCPU at 2.0 GHz, 8 GB RAM/disk, 10 gigabit network, using Sepia framework*

**Optimal MPC sorting protocol depends on preconditions and number of inputs**

# Research Challenges

### Secure Data Capsule

- Transformation of data to match protections specified by policy
- Integration with policy and provenance

### Secure Computation

- Automatic selection and composition of techniques
- Integration with policy and provenance

### Security Policies

- Rich policy representation formats
- Combining policies on data from multiple owners

### Data Provenance

- Truncation-resistant provenance store
- Provenance analytics

### Content-Centric Networking

- Secure resource discovery
- Resilience against malicious nodes

# Summary

- Data-centric secure computing shifts paradigm from protecting large systems to protecting data
- Data protected at rest, in transit, and in use with respect to expressive policies
- Vision requires integrated architecture and component technologies: cryptographically secure storage and computation, policy, data provenance, content-centric networking
- Interested in your ideas for applications and collaboration