THE DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING 計算機科學及工程學系

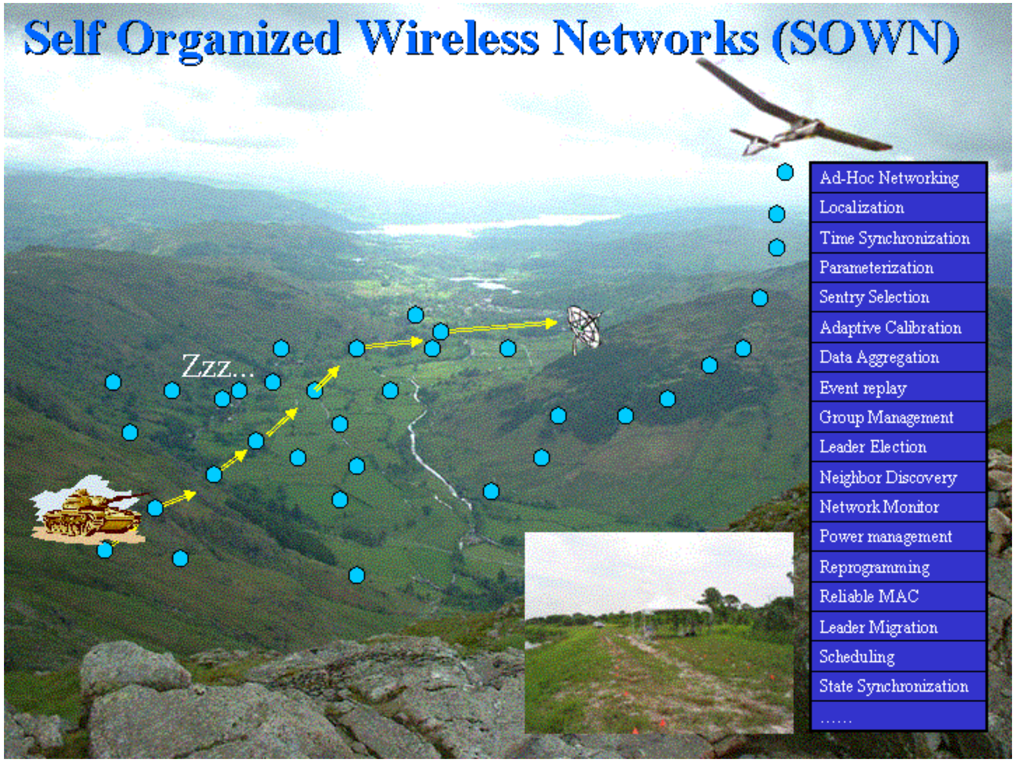
A Synergy of the Wireless Sensor Network and the Data Center System

Ke Hong, Shuo Yang, Zhiqiang Ma and Lin Gu

Department of Computer Science and Engineering
The Hong Kong University of Science and Technology

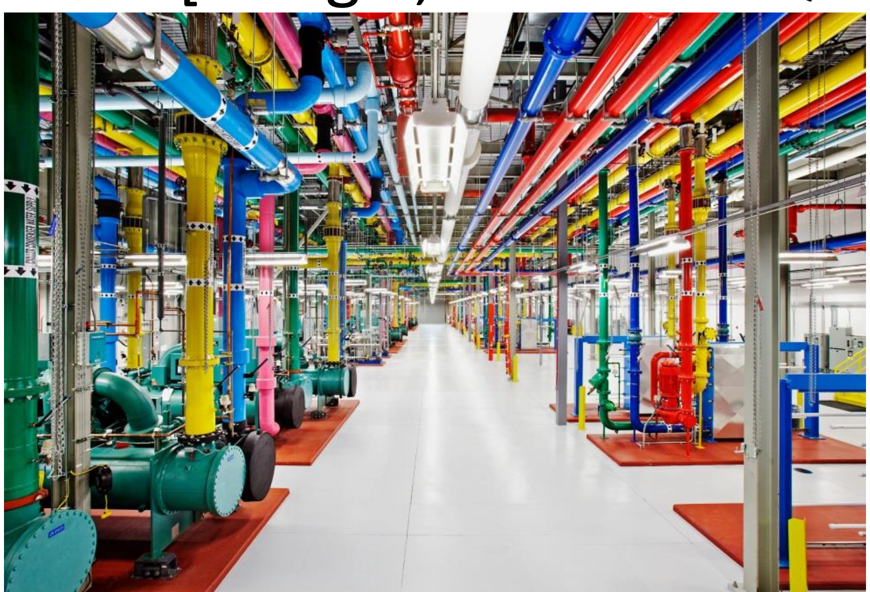
Data Center vs. Sensornet

- Both distributed, dense, scalable
  - 300 nodes in VigilNet, hundreds in GreenOrbs, 1000+ in ExScal



Data Center vs. Sensornet

- Both distributed, dense, scalable
  - 300 nodes in VigilNet, hundreds in GreenOrbs, 1000+ in ExScal
  - Thousands of compute servers organized in racks [Google, Microsoft Quincy]



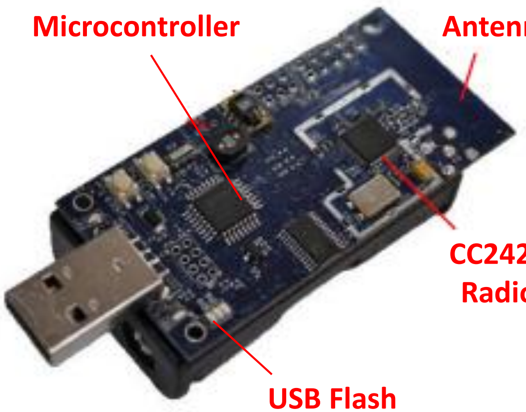
Data Center vs. Sensornet

- Both distributed, dense, scalable
  - 300 nodes in VigilNet, hundreds in GreenOrbs, 1000+ in ExScal
  - Thousands of compute servers organized in racks [Google, Microsoft Quincy]

- Low-end and high-end of computation
  - Limited computing resource on each sensor node
  - Abundant computing resources on rack servers

Related Work

- Sensornet in data centers
  - “Cool” scheduling [USENIX ‘05]
  - RACNet [SenSys ‘09]
  - Thermocast [KDD ‘11]



- The combined computational and networking capability of a sensornet enables it to interact with compute clusters in a more sophisticated way

Cluster-Area Sensor Network

- CASN as a complementary solution
  - To improve the cluster management
  - To enhance the operational security

- Cluster-wide command dissemination

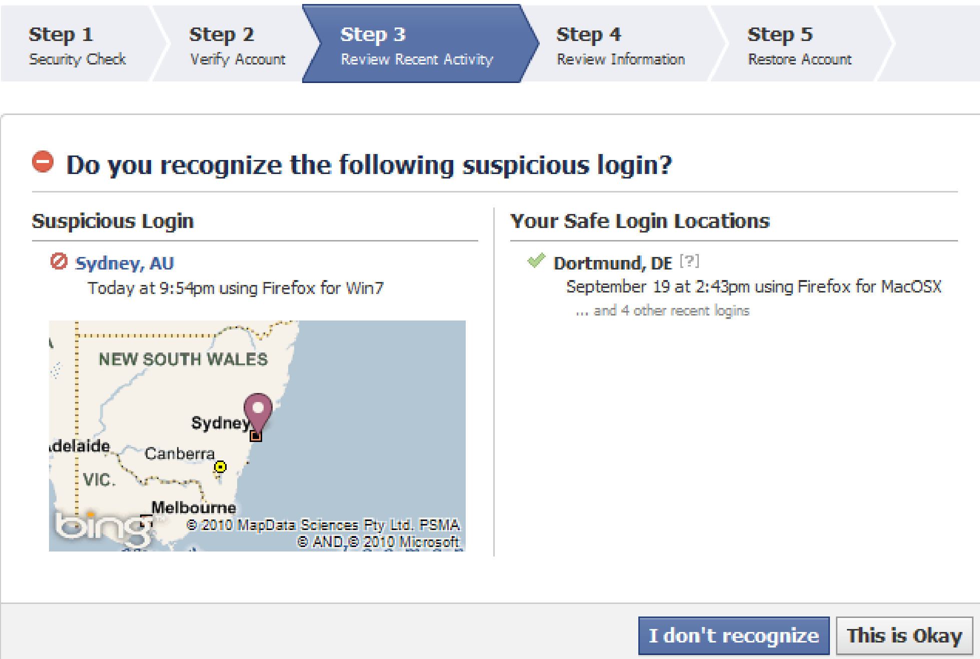
- Verification of server's physical presence

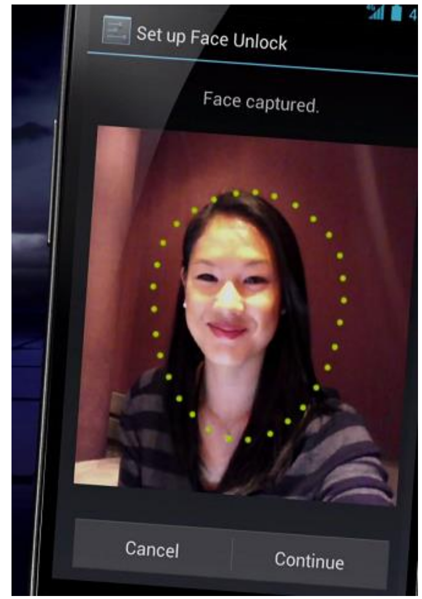
Management in Data Centers

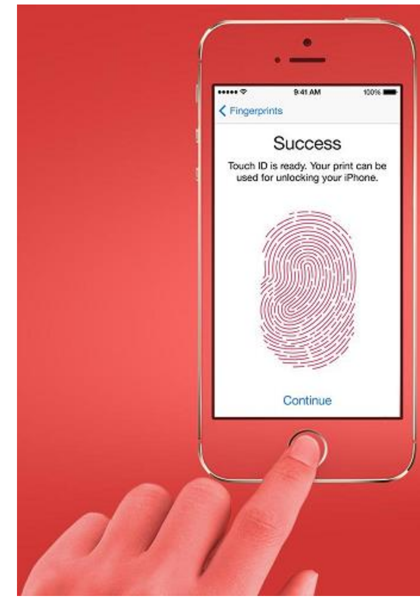
- Software reprogramming on compute servers
  - System settings, configuration files, software upgrade
  - Usually performed on a management station
  - Require certain manual operations

- Why not wirelessly broadcast commands and small files via a sensornet?
  - Wireless as a convenient and flexible broadcast medium

Security Hints







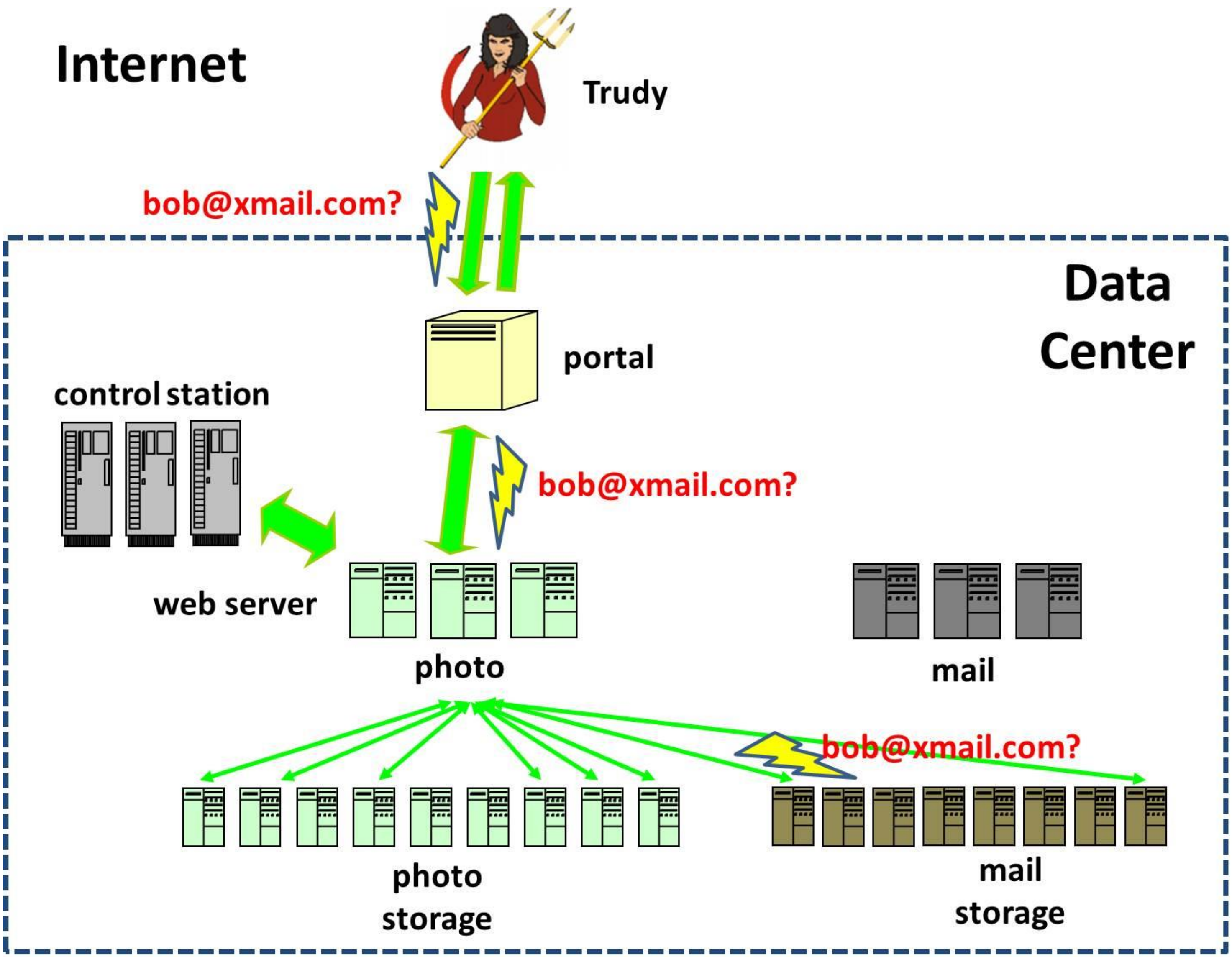
Two-step verification adds an extra layer of protection to your account. Whenever you sign in to the Dropbox website or link a new device, you'll need to enter both your password and a security code sent to your mobile phone.

Security in Data Centers

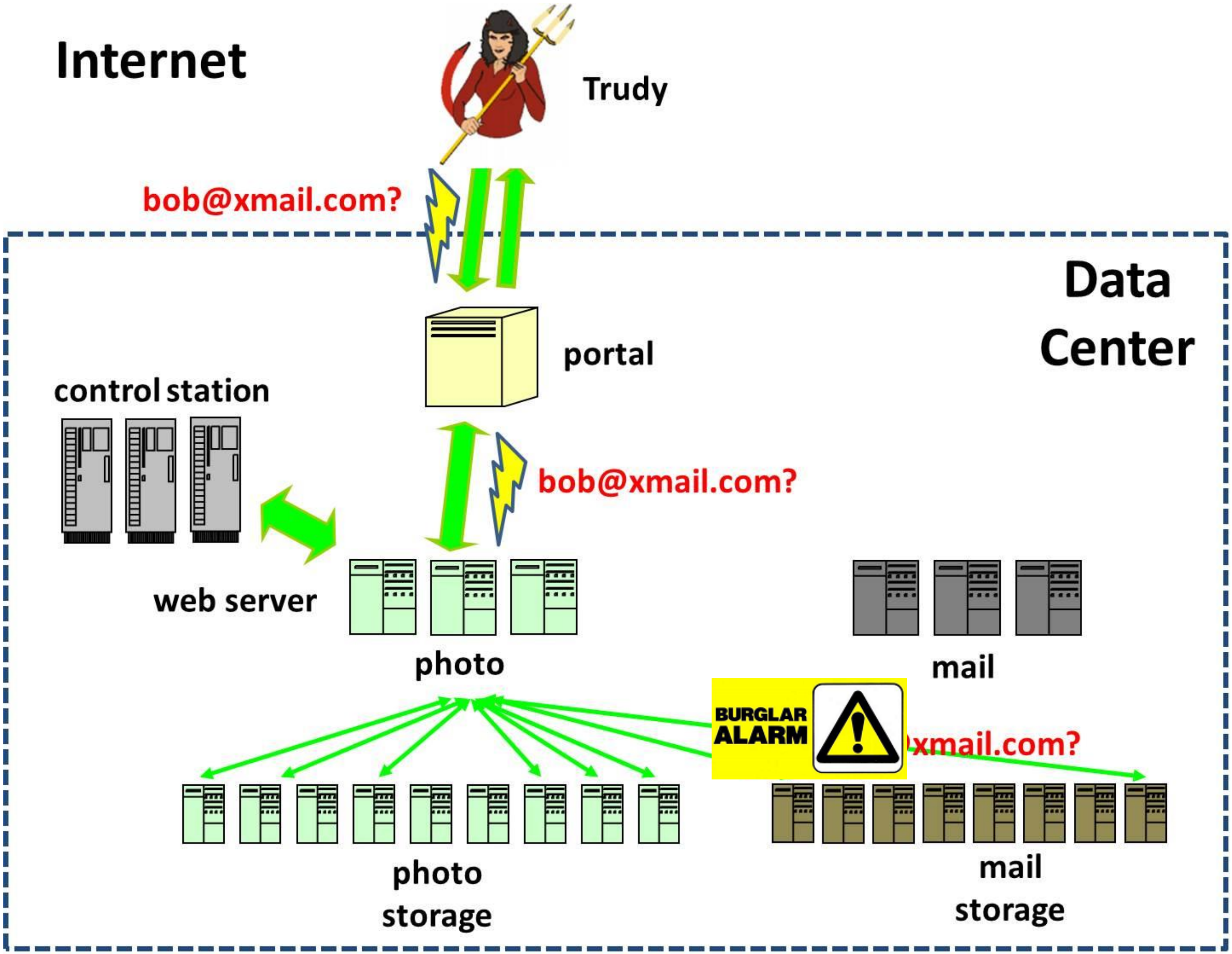
- Existing cryptologic methods do not entirely ensure the operational security of data centers
  - User account leakage at *Yahoo!*, *Sony PlayStation Network* and *Qriocity*
  - Need additional measures for security monitoring

- New security hint: **servers' physical presence**
  - Servers in data centers usually serve different roles (i.e. management, web agent, mail agent, storage)
  - Alarm triggered upon request from strange roles

Access Path Verification

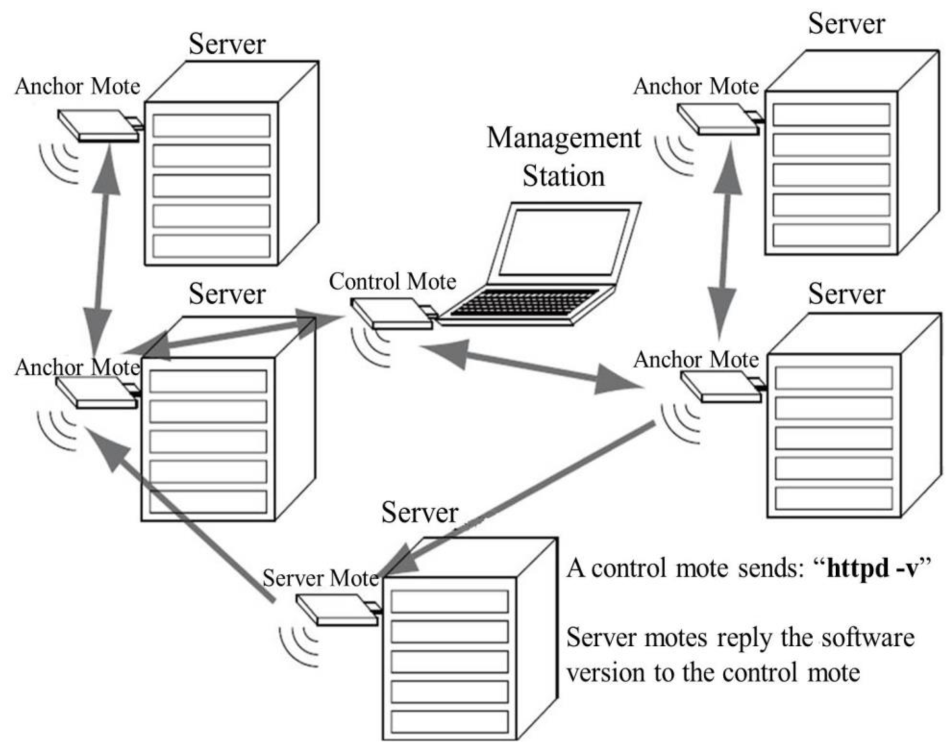


Access Path Verification



CASN Architecture

- System components
  – Sensor network
  – Compute servers
- Three types of motes
  – Control motes
  – Anchor motes
  – Server motes

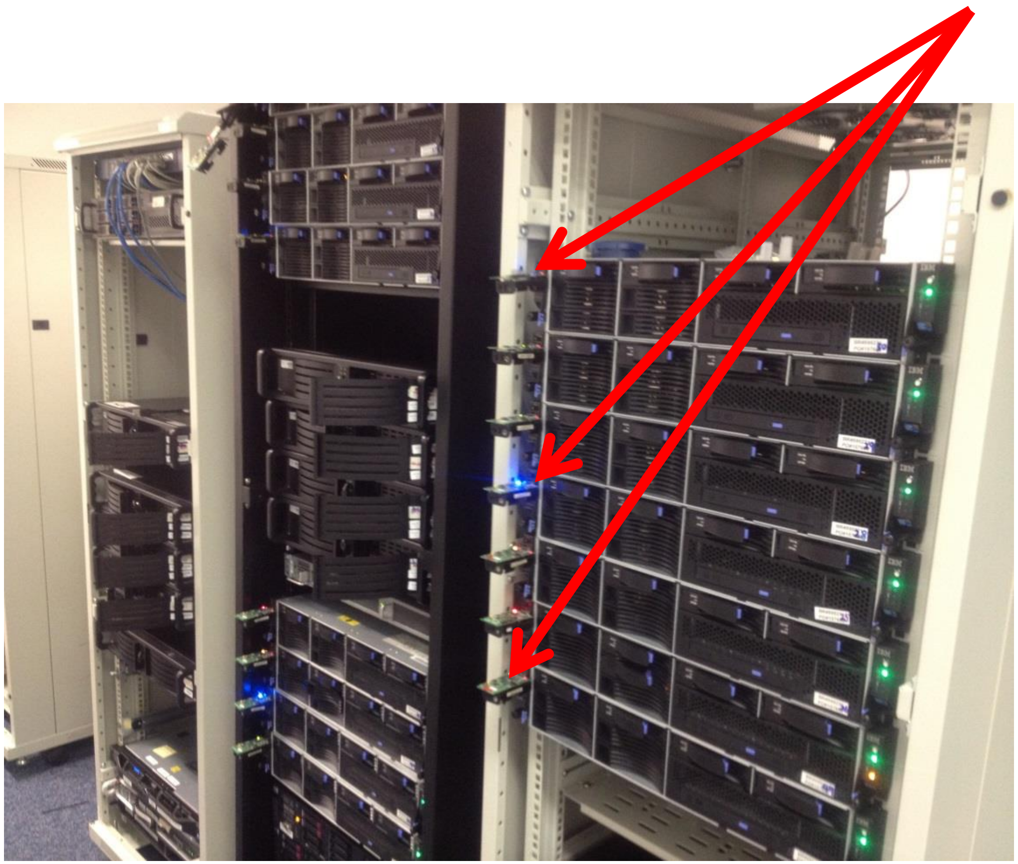


Prototype Implementation

A prototype of CASN consisting of 1 control mote and 4 anchor motes (Telos B) in a research cluster

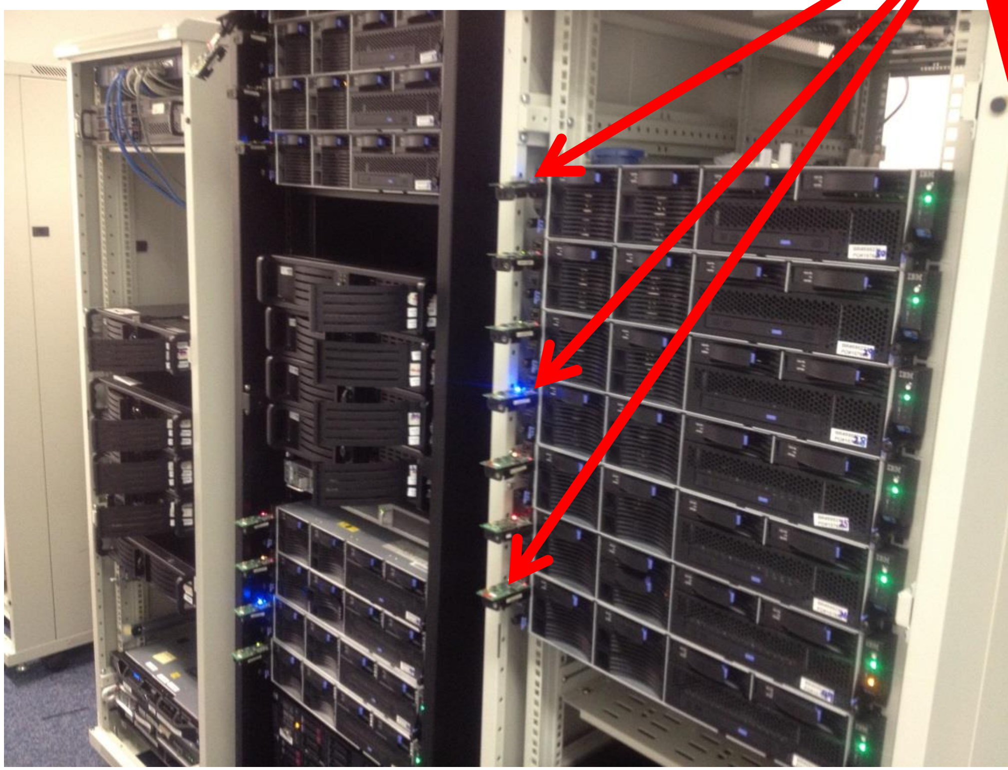
Prototype Implementation

motes attached to servers via USB interfaces



Prototype Implementation

motes attached to servers via USB interfaces

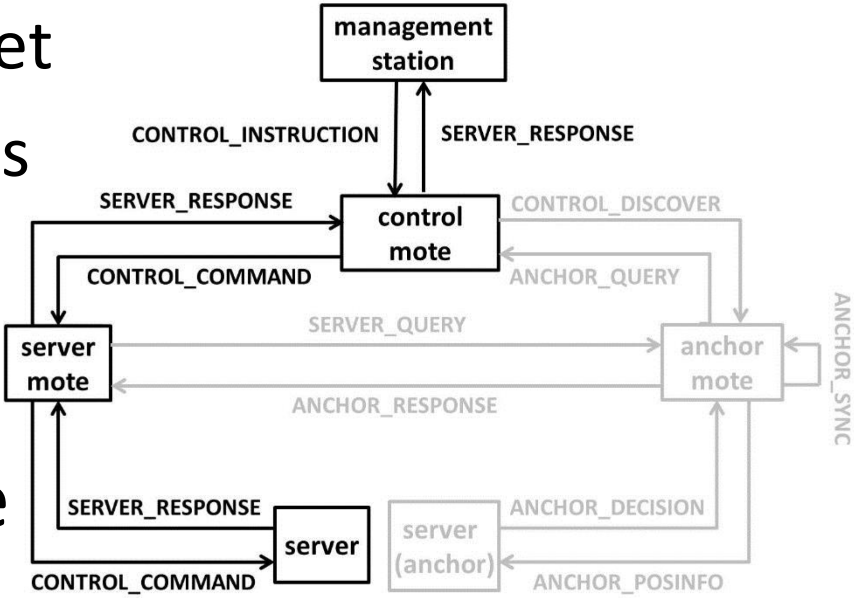


Cluster-Area Sensor Network

- CASN as a complementary solution
  - To improve the cluster management
  - To enhance the operational security
- Cluster-wide command dissemination
- Verification of server's physical presence

Command Dissemination

- Workflow of command dissemination
  – Issued from the management station
  – Forwarded to the control mote
  – Broadcasted via sensornet
  – Received by server motes
  – Executed on servers

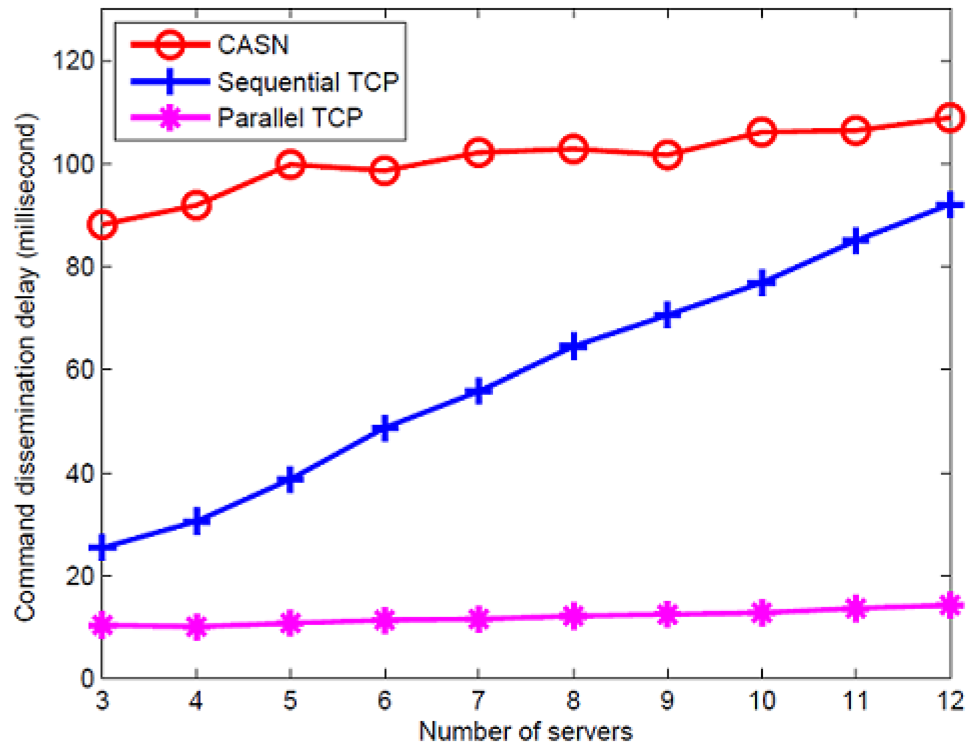


- Command-line interface

Command Dissemination Delay

- To evaluate the round-trip delay of command dissemination to a number of servers across three racks

- Results
  - Scalable broadcast via sensornet
  - Stable delay



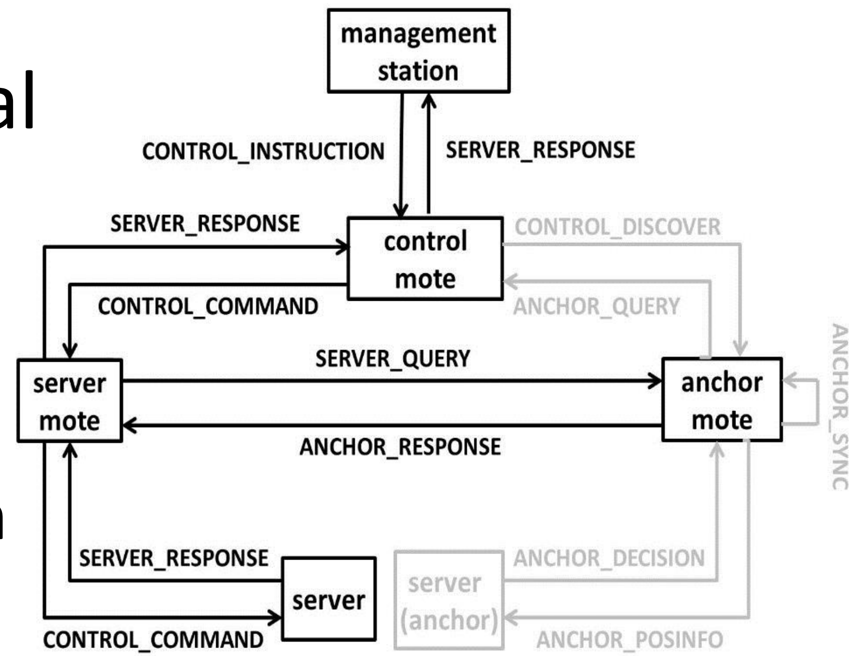
Cluster-Area Sensor Network

- CASN as a complementary solution
  - To improve the cluster management
  - To enhance the operational security
- Cluster-wide command dissemination
- Verification of server’s physical presence

Verification of Physical Presence

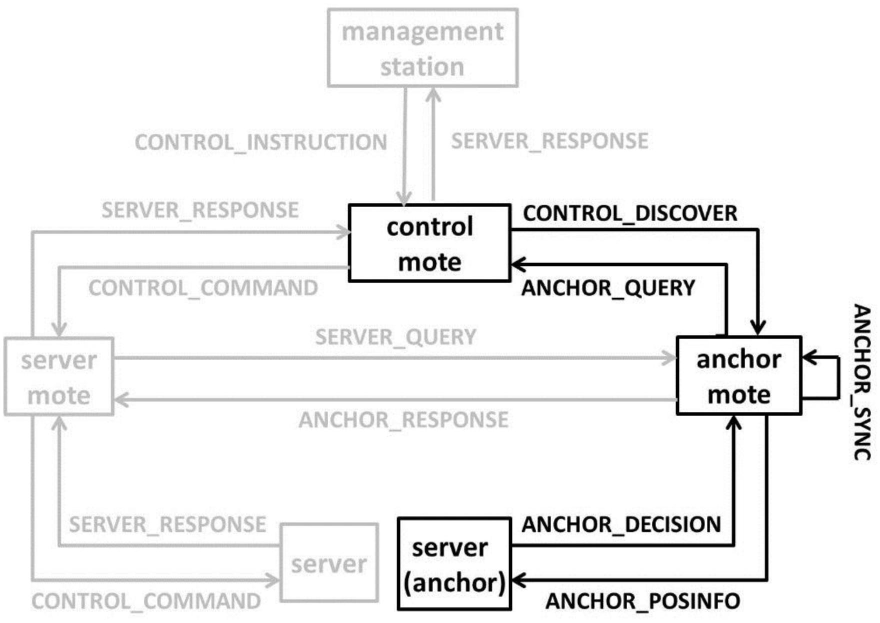
- Operations in data center are yet to be secure
  - Example: impersonating the management station

- Example: verify the physical location of a control mote
  - Before execution, server motes query anchor motes for the legitimacy of certain control mote



Localizing Control Motes

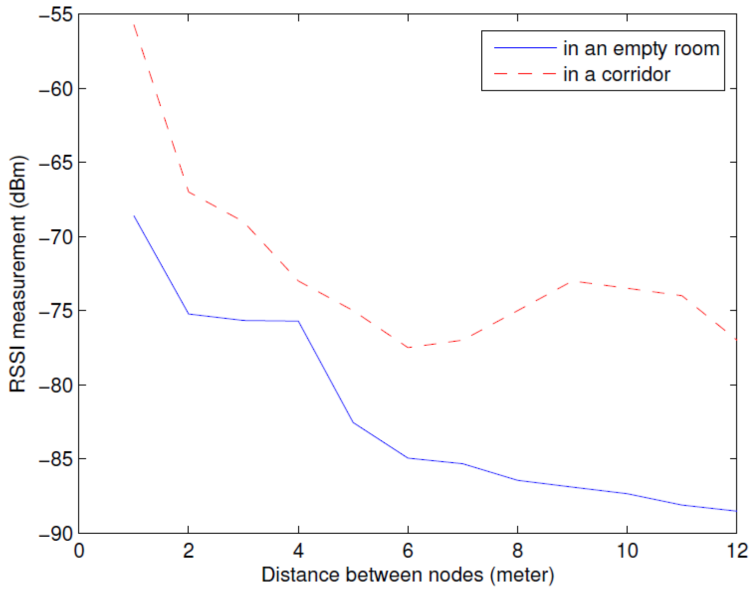
- Workflow of physical localization
  - Passive discovery: anchor motes periodically query the location of control motes
  - Active discovery: control mote initiates discovery upon its arrival
  - Anchor motes together localize a control mote to determine its legitimacy
- Suffice with 4 anchors



Radio-based Localization

- Coarse-grained radio-based localization
  - Suffice even at 5-meter precision
  - Inefficacy of RSSI-based ranging approach

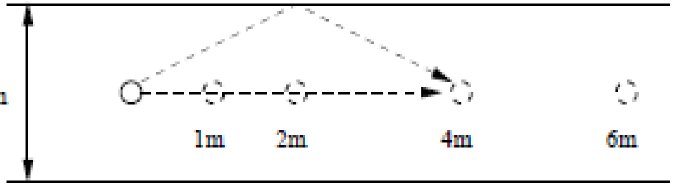
$$P(d) = P(d_0) - 10n \log(\frac{d}{d_0})$$



- Necessity for empirical RSSI modeling in a data center environment

Empirical Localization Model

- Cope with the multipath effect by considering indirect signals



$$P(d) = P(d_0) - 10n\log\left(\frac{\sum_{i=1}^{k} r_i d_i}{d_0}\right)$$

  - $\boldsymbol{R} = [r_1\ r_2\ \ldots\ r_k]$ as the amplitude coefficients of signal components
  - $\boldsymbol{D} = [d_1\ d_2\ \ldots\ d_k]$ as discretized distances of signal components

- Rician distribution used to model amplitudes of indirect signals

$$R(x|\gamma,\sigma) = \frac{x}{\sigma} e^{\frac{-(x^2+\sigma^2)}{2\sigma^2}} I_0\left(\frac{x\gamma}{\sigma^2}\right)$$

Probabilistic Ranging

- Solving ***R*** in $\boldsymbol{R} * \boldsymbol{D} = d_0 * 10^{\frac{P(d_0)-P(d)}{10n}}$
  - Consider only the 5 shortest reflected signals
  - d_{AB} as the distance between the transmitter A and receiver B (i.e. 2 meters)

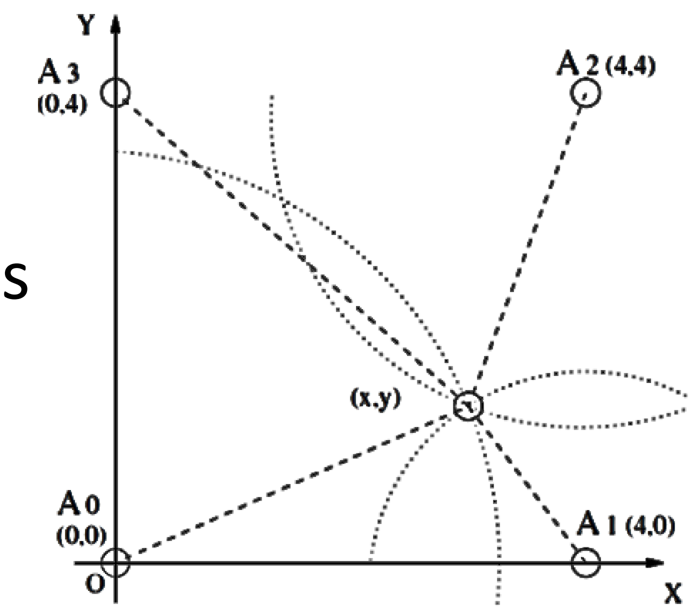
$$r_i = \begin{cases} 0 & if\ d_i < d_{AB}\ or\ d_i - d_{AB} \geq 2 \\ 1 & if\ d_i = d_{AB} \\ a_i * R(d_i - d_{AB}) & if\ d_i > d_{AB}\ and\ d_i - d_{AB} < 2 \end{cases}$$

- Localization: after obtaining the probabilistic ranging results, compute the most plausible location using trilateration

Reduce Computation Cost

- Computationally costly for all possible cases
  - In total K^H cases for *H* RSSI measurements per transmission, give that each maps to *K* **R**s

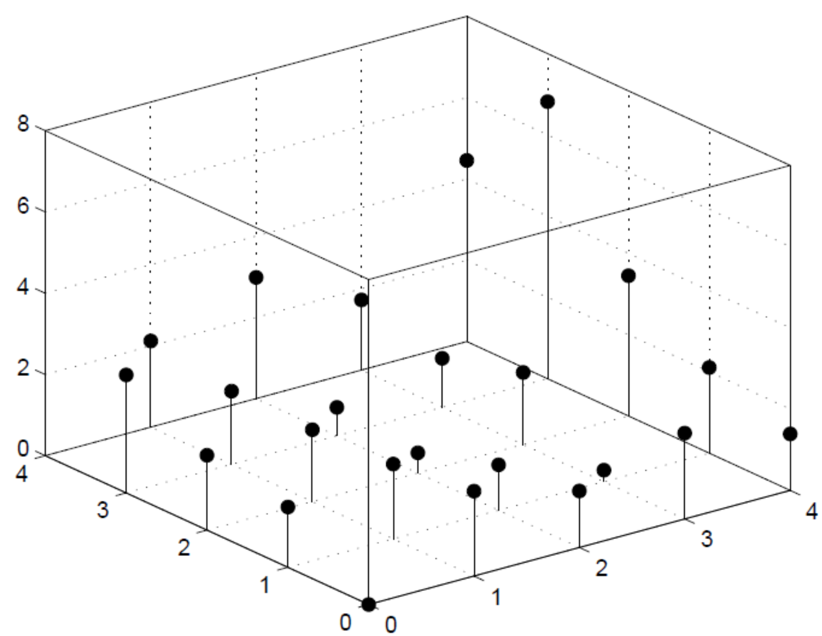
- Reduce computation cost by
  - Narrowing down distances by applying geometric constraints
  - Utilizing the known distances between anchors



Localization Accuracy

- Evaluate the localization accuracy in a 4m x 4m square field defined by A_0, A_1, A_2, A_3
  - Localization error *e* = $\sqrt{(x' - x)^2 + (y' - y)^2}$

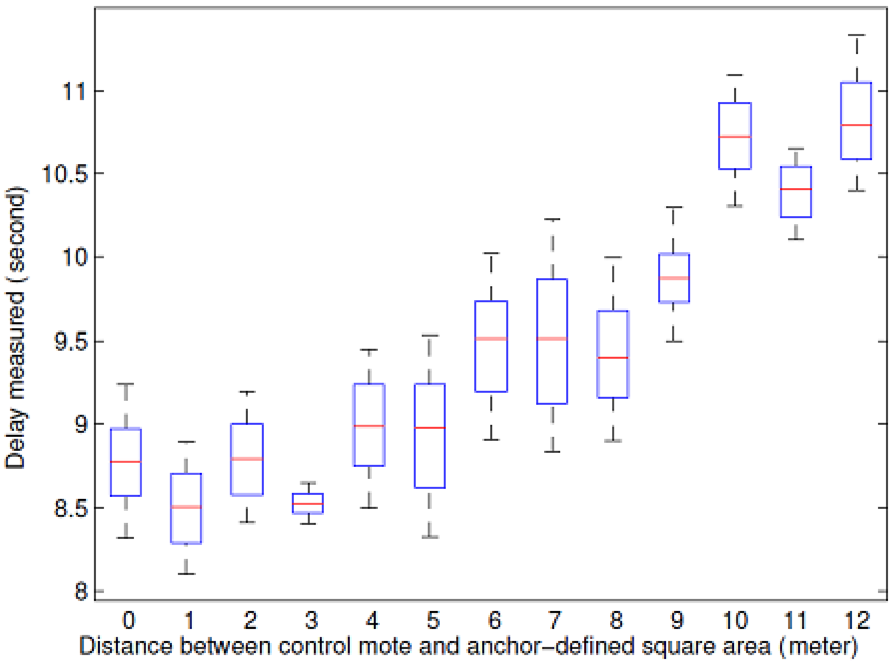
- Results
  - 88% of localization errors within 5 meters
  - Errors for positions inside the square within 2 meters



Localization Delay

- To evaluate the localization delay by varying the distance between a control and an anchor mote inside the 4m x 4m square field

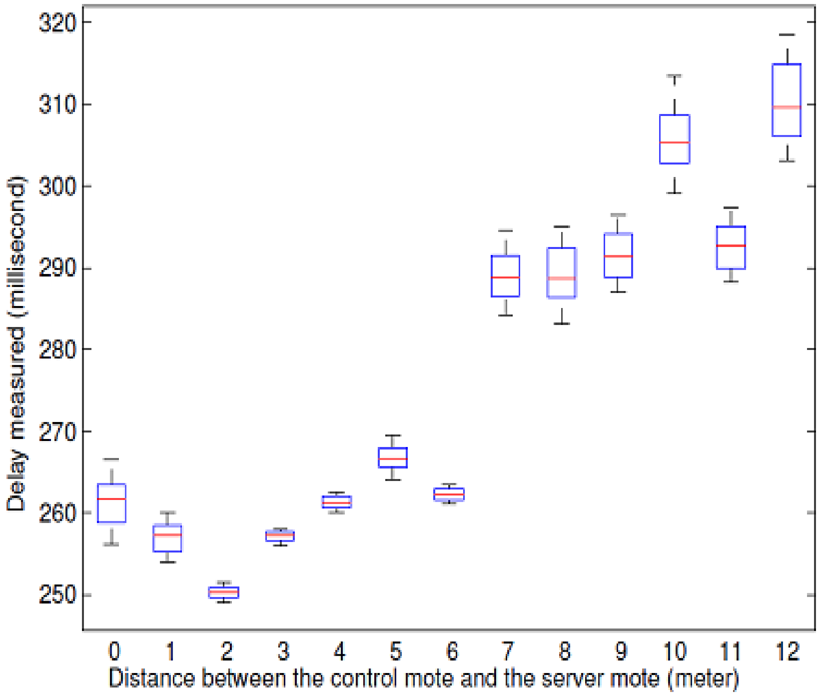
- Results
  - Overall 8-12 seconds
  - Small variation
  - Acceptable with 30-sec localizing period



Reprogramming Delay of CASN

- Reprogramming delay: command dissemination delay + physical verification delay

- Results
  - Less than 300 milliseconds with distance closer than 10 meters
  - Low enough for effective command dissemination



Summary

- We design and implement a cluster-area sensor network in a data center
  - Wireless cluster-wide command dissemination
  - Empirical localization for verification of server’s physical presence
- Future work
  - CASN with fingerprint-based localization
  - CASN in geographically distributed data centers

Q&A

Thank You!

Ke Hong
khongaa@ust.hk