Practical and Secure Machine-to-Machine Data Collection Protocol in Smart Grid

Suleyman Uludag[1], King-Shan Lui[2], Wenyu Ren[3], Klara Nahrstedt[3]

[1] University of Michigan – Flint

Department of Computer Science, Eng. and Physics

[2] University of Hong Kong

Department of Electrical and Electronic Engineering

[3] University of Illinois at Urbana-Champaign

Department of Computer Science

TCIPG

TRUSTWORTHY CYBER INFRASTRUCTURE FOR THE POWER GRID

M2MSec'14

with IEEE CNS 2014

October 29, 2014

San Francisco, CA, USA

First International Workshop on Security and Privacy in Machine-to-Machine Communications

(M2MSec'14)

Outline

- Background
  - TCIPG
  - Smart Grid
  - Smart Grid and M2M Communications
- Data collection model (PO, DC, MD)
- Related work
- Protocol Overview and System Parameters
- Shared Key Generation
- Data Collection
- Preliminary Experimental Results
- Conclusion

TCIPG -- ACknowledgement

- Trustworthy Cyber Infrastructure for the Power Grid (TCIPG)
- Takes on the challenges of making the electric power infrastructure continuously functioning
  - While enhancing security, reliability, and safety
- U. of Illinois at Urbana-Champaign, Dartmouth College, WA State U., UC-Davis
- Funded by the US Department of Energy and Department of Homeland Security

Smart Grid Background

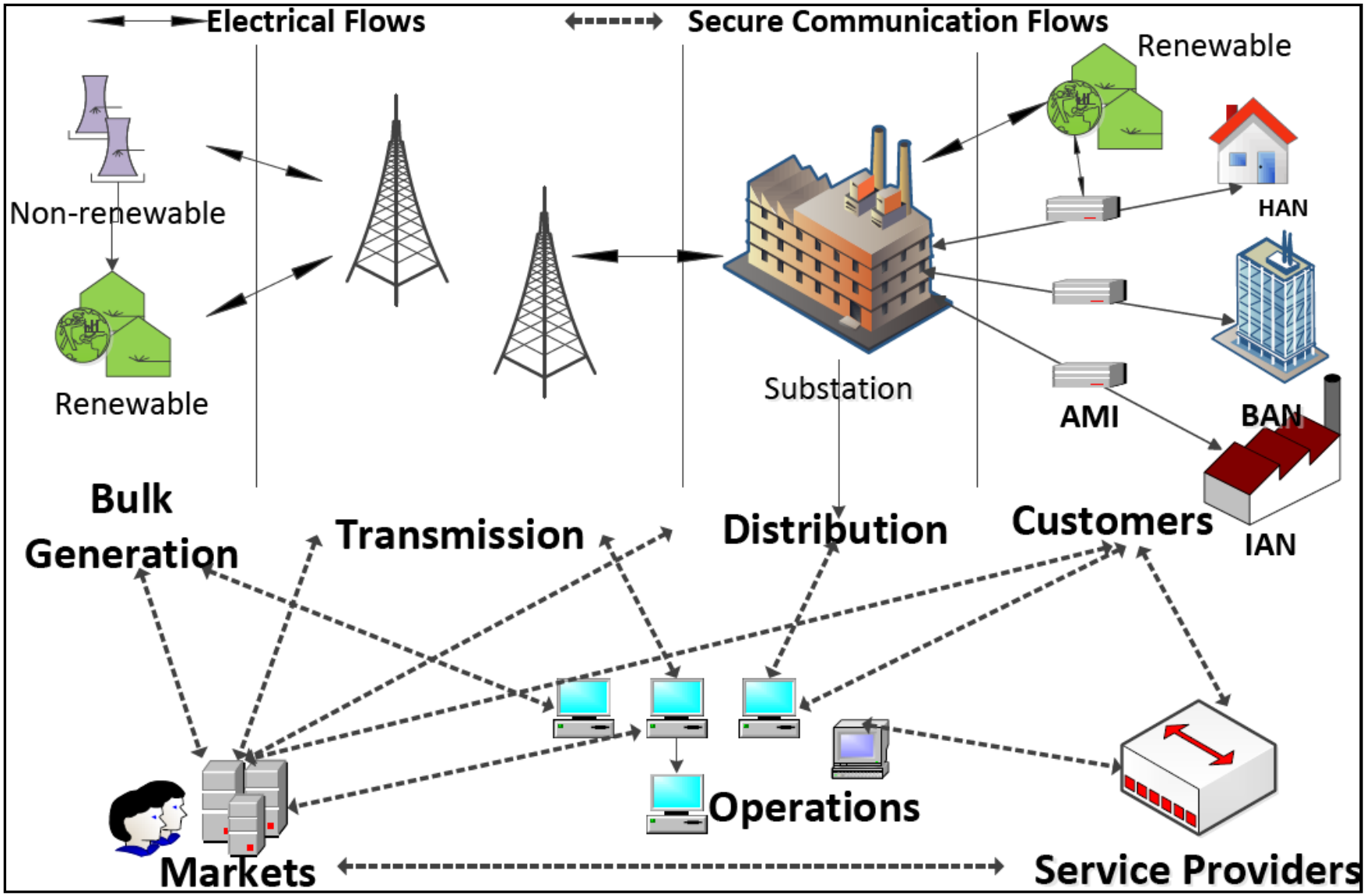
- Power Grid
  - Largest machine in the world
  - Centralized generation and control
  - Unidirectional power flow (from the grid)
  - Limited info flow (from the edge)
- Paradigm shift in energy production, transmission, distribution, consumption
- Key facilitator → **Smart Grid**

A Rough SG Definition

- Smart Grid (SG) is a vision, and a system of systems
- More specifically, the SG can be regarded as an electric system that uses information, two-way cyber-secure communication technologies, and computational intelligence in an integrated fashion across electricity generation, transmission, substations, distribution and consumption to achieve a system that is clean, safe, secure, reliable, resilient, efficient, and sustainable.

Source: Fang et al. Smart Grid Survey, 2011, IEEE Comm. Surveys and Tutorials

NIST Smart Grid Conceptual Model



Based on http://www.oncor.com/images/content/pathwaytopower.gif

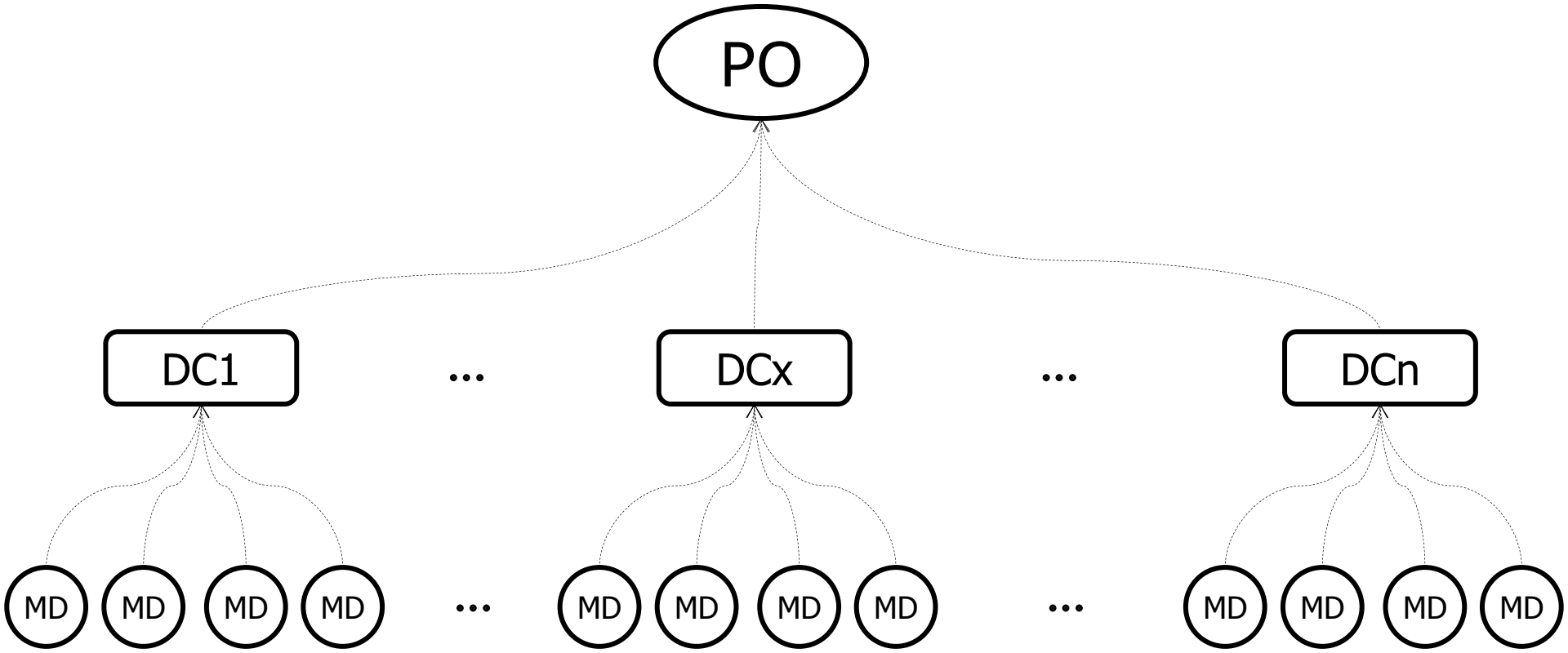
Smart Grid and M2M Communications

- Machine-to-Machine → a building block for SG
  - Especially for the wide-scale monitoring and control infrastructure
- SG → one of the most promising growth areas M2M adoption
  - Smart Cities, Smart Buildings, Vehicle-to-Grid (V2G), Grid-to-Vehicle (G2V) communications , Plug-in Hybrid Electric Vehicles (PHEV)
- Evolution of power distribution networks calling for sophisticated **Energy Management Systems (EMS)** for improved efficiency

Smart Grid M2M Applications

- Optimal line controls & loss minimization
- Fault-detection and management
- Load balancing
- Real-time state estimation
  - Phasor management units (PMU)
- AMI (Advanced Metering Infrastructure)
  - Smart meters
- Facilitate integration of distributed and/or renewable sources
- Distribution Automation (DA)
- Various other sensors, measurement devices, pole-top devices, etc.

Data Collection Model



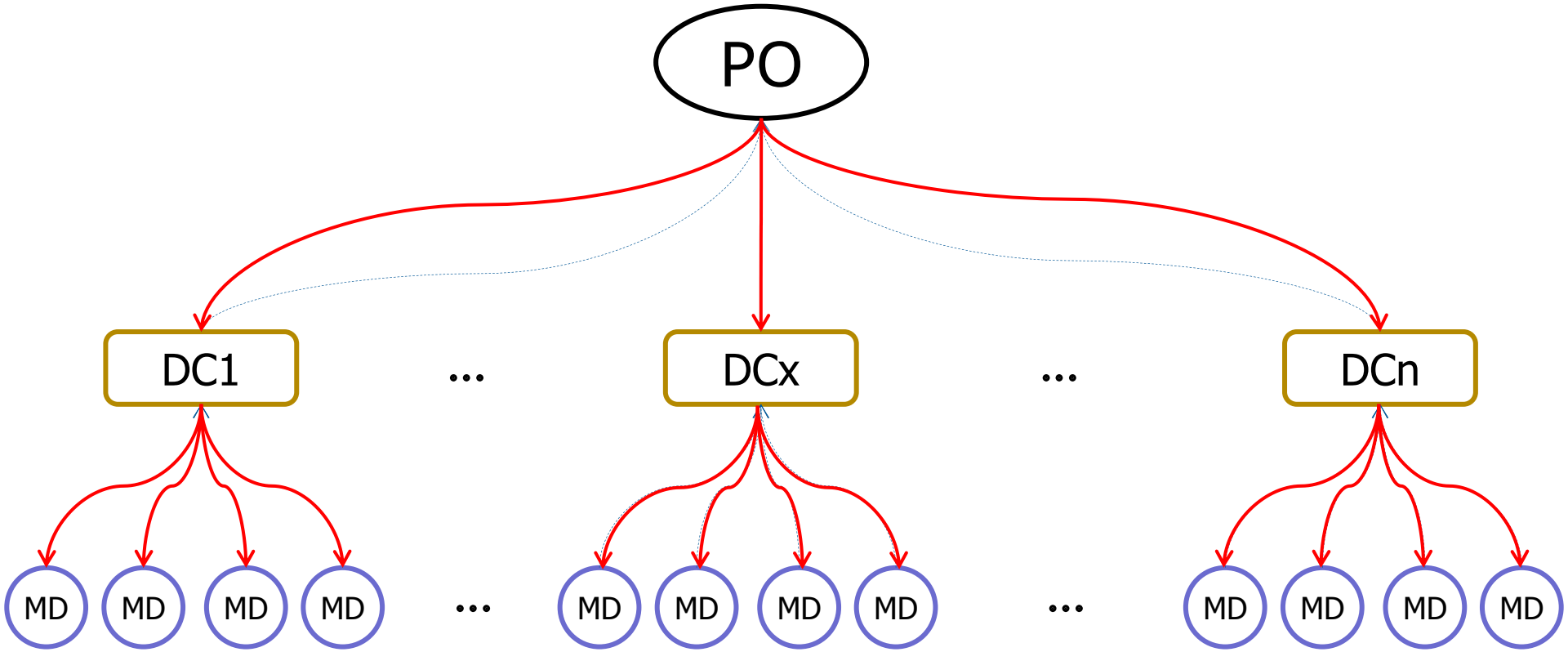
- Measurement Devices (MD)
- Data Collectors (DC)
- Power Operator (PO)

Related Work

- End-to-End data protection studied extensively for the Internet
  - TLS/SSL assume abundant power/memory
- Distributed Network Protocol 3 (DNP3) for SCADA
  - Assumes operating a security perimeter of the operator
- IEC 61850 does not have such a secure provisioning
- Overall existing protocols cannot address a hierarchical data collection model where PO and MDs cannot establish direct connection

System & Protocol Overview

- PO initiates the data collection process periodically or otherwise



Protocol specs

- Goal is to protect against
  - Eavesdropping
  - Impersonation
  - Message tampering
- PO → Always trustworthy
- DCs → Honest-but-curious
- MDs :
  - Trustworthy at the time of installation
  - Can then be attacked
- Our protocol cannot identify whether data is legit if a legit key is used
  - But, MDs cannot impersonate other MDs

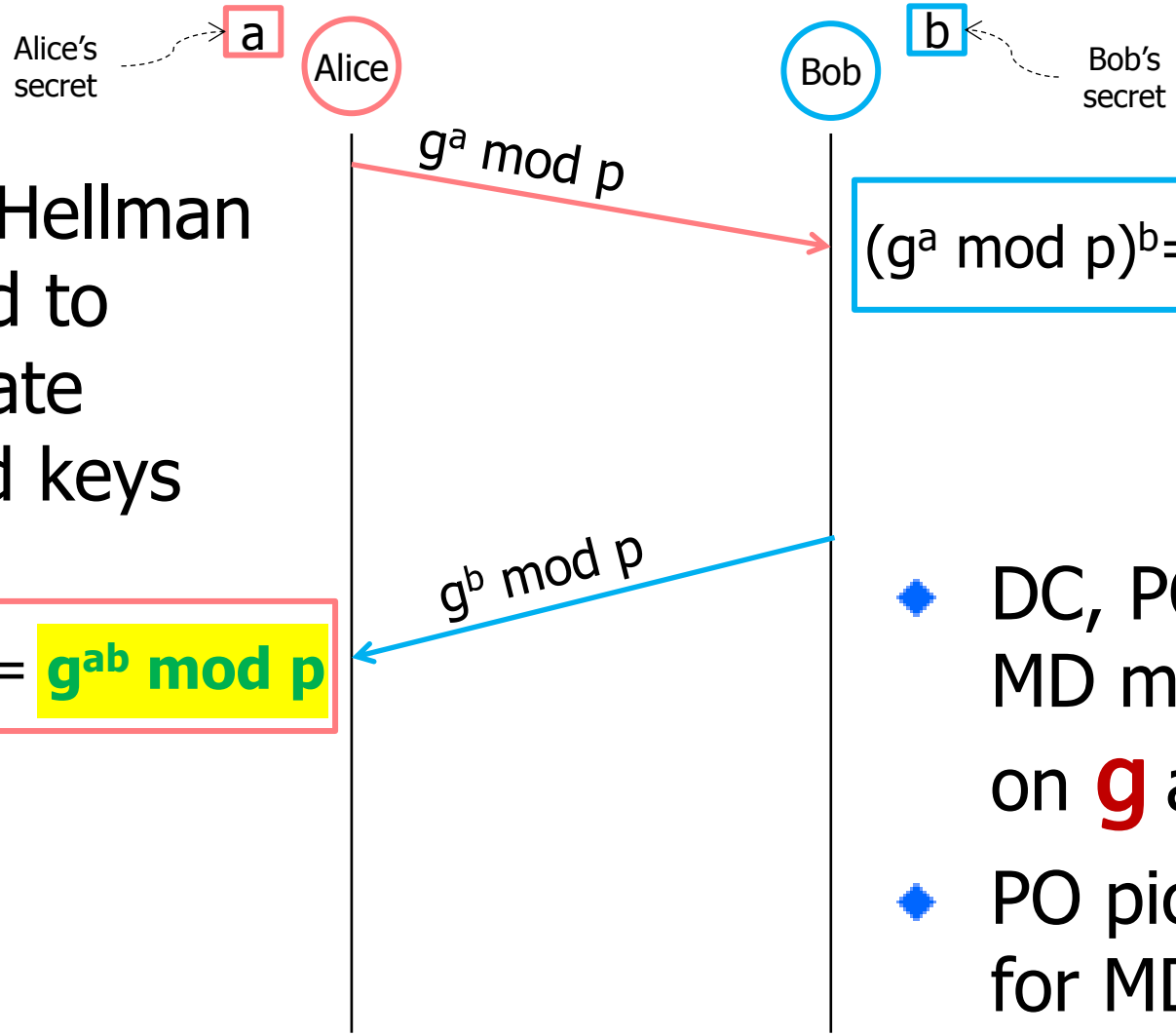
System Parameters

- PO, DCs, and MDs are assumed to have been equipped in advance with
  1. Long-term secrets,
  2. A set of system parameters, and
  3. The required cryptographic functions.
- A DC or MD should have all parameters and functions configured before it is deployed/installed in the field
- Notation:
  - Public key of node **A** → $\mathbf{A}^{+}$
  - Private key of node **A** → $\mathbf{A}^{-}$

1. Long-Term Secrets

- Key server generates private/public keys for all
  - Burned into DCs/MDs before installations
- PO keeps own as well as keys of all the MDs/DCs
  - PO does not make any keys available to outside
  - Our protocol is secure even if the public keys are known

2. System Params. (Diffie-Hellman)



- Diffie-Hellman is used to generate shared keys

- DC, PO, and MD must agree on **g** and **p**
- PO picks these for MDs/DCs

3. Cryptographic Functions

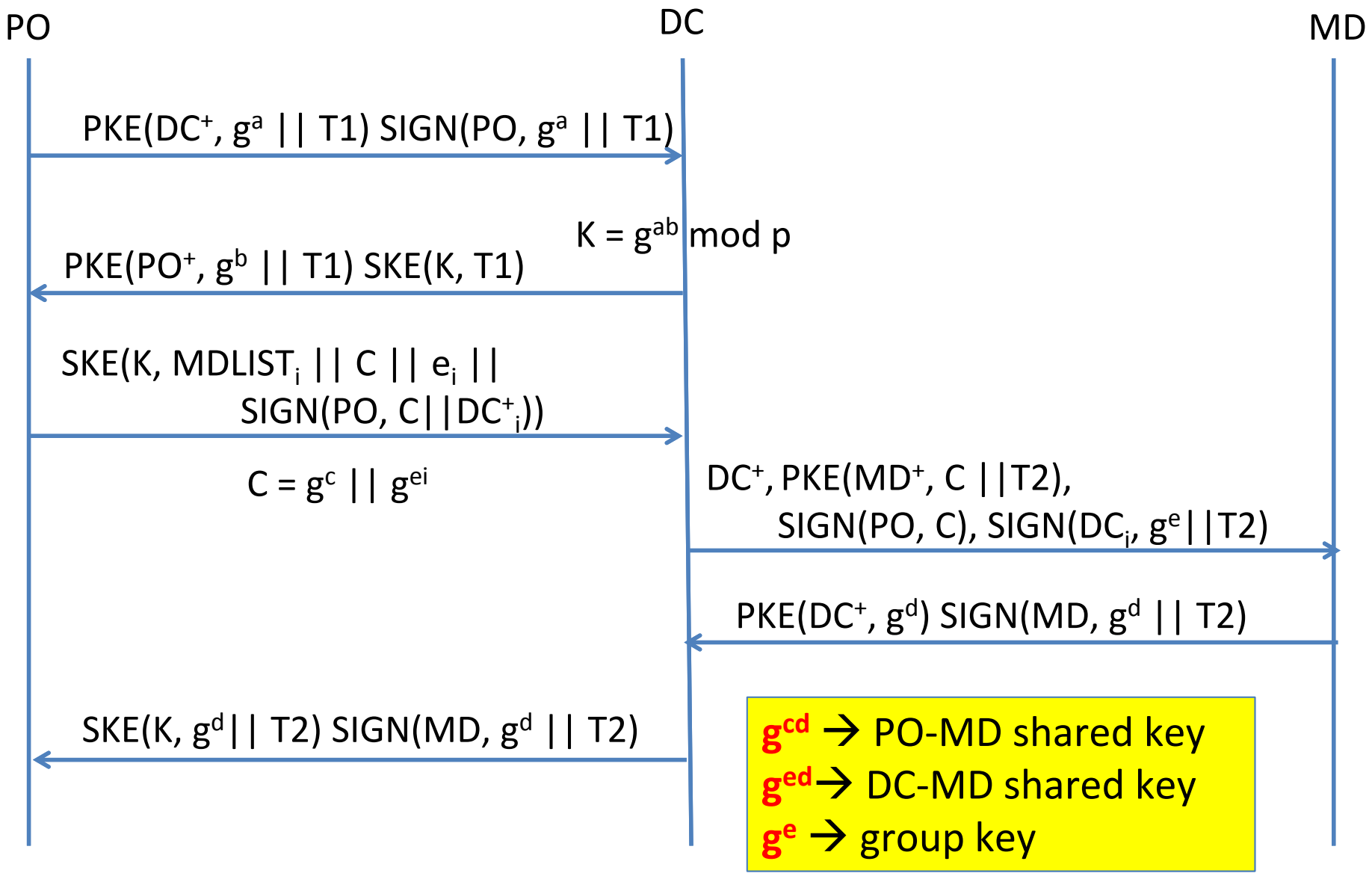
- Encryption, hashing, or signing
  - Needed for authentication, confidentiality, integrity
- PO selects appropriate cryptographic functions to be pre-installed in DCs/MDs

Name	Description
PKE(K,M)	apply public key encryption on M using K
SKE(K,M)	apply symmetric key encryption on M using K
SIGN(A,M)	The signature of M by A (created using A^-)
HASH(K,M)	compute the keyed-hash of M using key K
GENKEY(X,Y)	generate a key based on X and Y

Protocol Overview

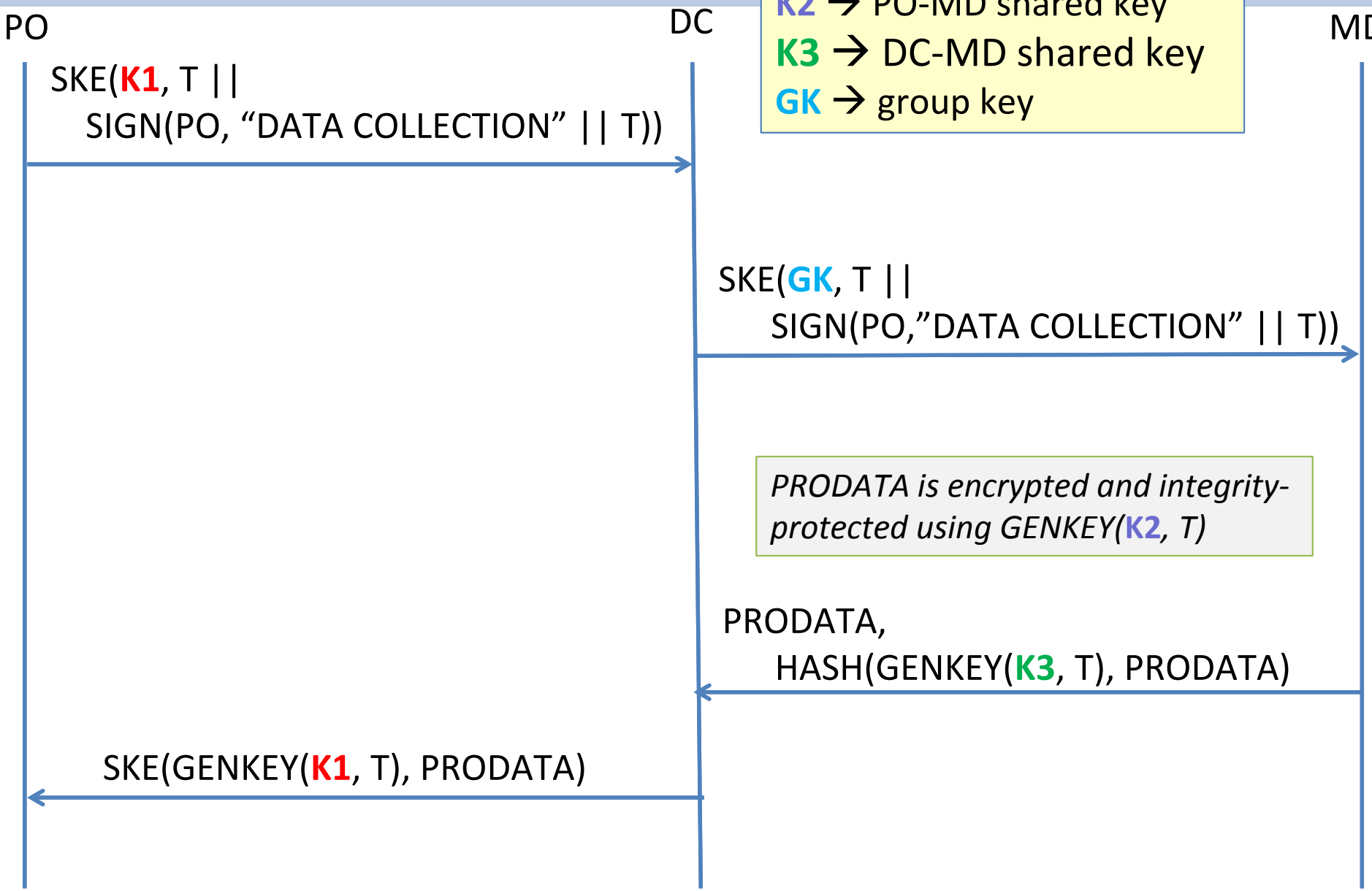
- Public-key infrastructure is NOT used for data
  - Expensive and not a good practice
- PO, DC, and MD use long-term secrets to establish shared keys
  - **Pairwise shared key**: A key only known by two parties
    1. PO-MD pairwise shared key for data
    2. DC-MD pairwise shared key (also known by PO)
    3. Group Keys among the PO, a certain DC and its MDs
- PO initiates the process to generate shared keys using the Diffie-Hellman exchange
  - DH half keys are authenticated using the long-term keys to prevent man-in-the-middle attacks
  - DH is expensive, so timestamps are used to avoid it every time

Shared Key Generation



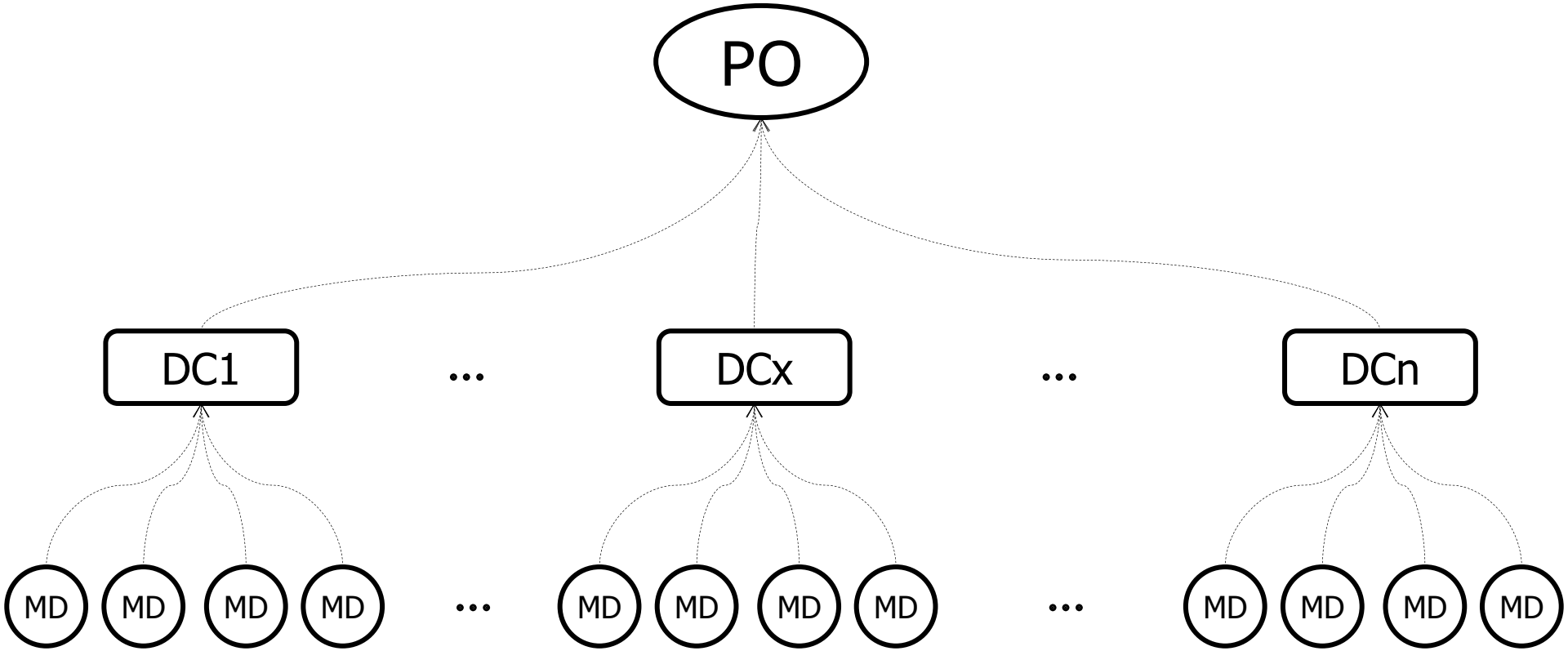
Data Collection

K1 → PO-DC shared key
K2 → PO-MD shared key
K3 → DC-MD shared key
GK → group key



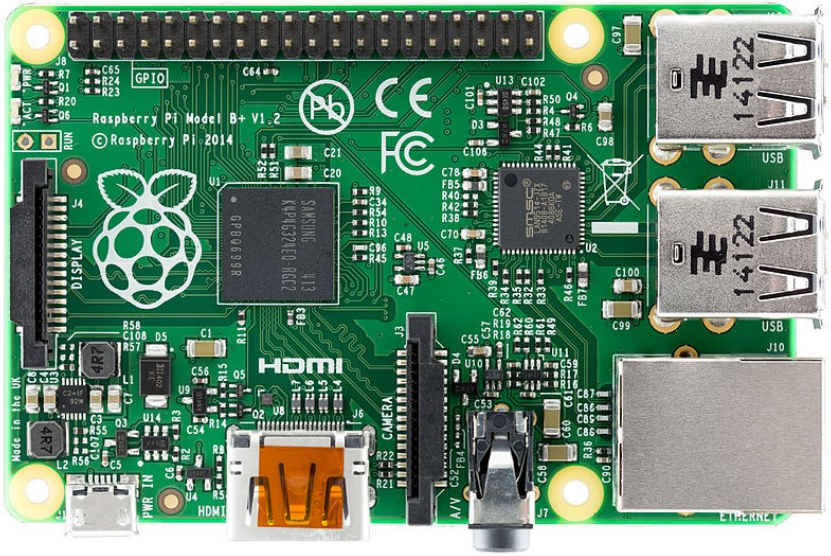
Preliminary Experimental Results

- Since the bottleneck of the protocol is on MD side, our focus is on MD-DC part
  - Message generation, communication, and message processing



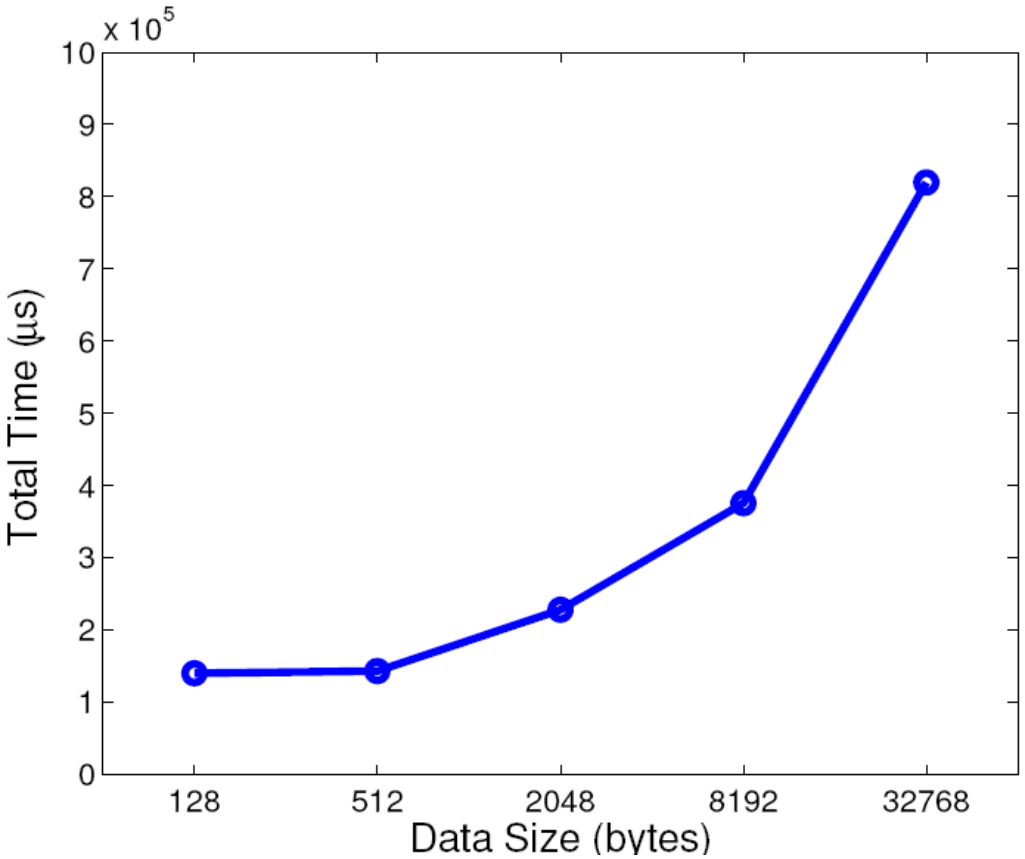
Preliminary Experimental Results

- A small testbed of data collection
- Resource-constrained MDs → by Raspberry Pi
  - Raspberry Pi: Credit-card sized, inexpensive, single-board computers



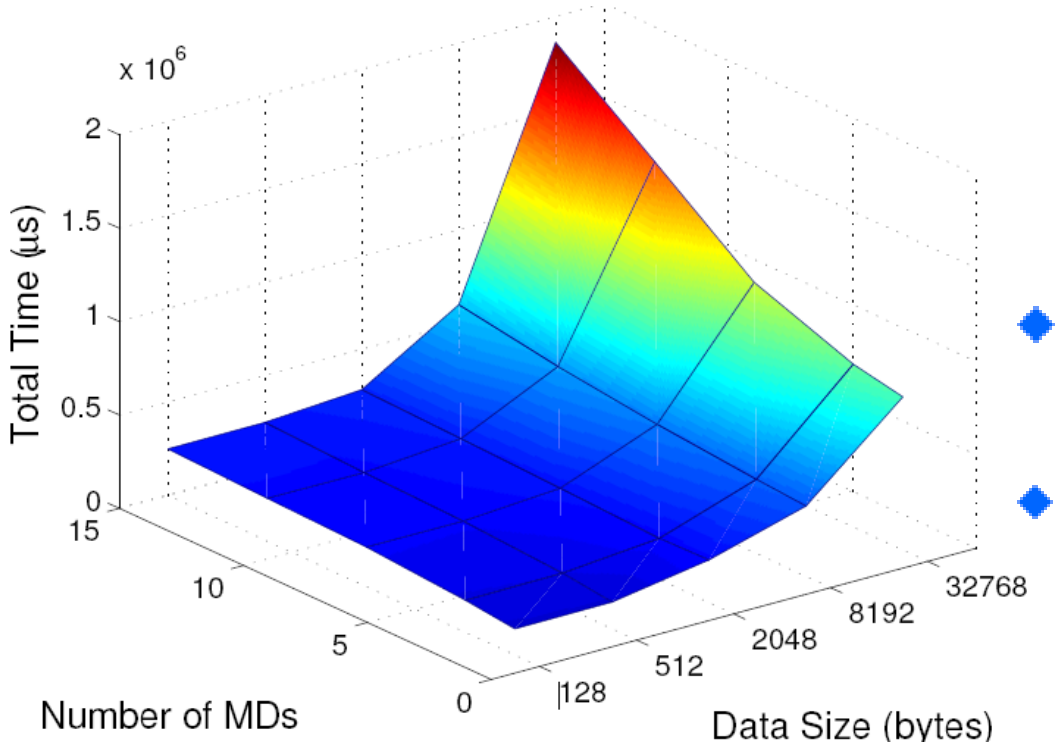
  - 700MHz Broadcom System-on-a-Chip (SoC)
    - Equivalent to late 1990s Pentium CPUs
  - 256/512 MB RAM
- 15 MDs
- DC → laptop

Time for data collection



- Total time
  - DC msg generation and transmission to MD
  - MD processing and transmission time
  - DC final processing time
- DH key size : 1024 bits
- RSA key size: 3072 bits
- AES key size: 256 bits

Collection Time, Size, # of MDs



- Total data collection time by one DC
  - DH key size : 1024 bits
  - RSA key size: 3072 bits
  - AES key size: 256 bits
- Total time rises with more MDs and data size
- Association of MDs with DCs emerges as a critical problem, which we will be studying as part of our future work.

Conclusion

- Data collection is critical for the success of the Smart Grid.
- Vastly increased data calls for mechanisms with security, reliability, and privacy
  - Automation becomes inevitable and thus M2M
- Proposed: a secure communications protocol data collection in a practical, scalable, and efficient manner under a hierarchical model
- 3rd party service providers enabled as envisioned by NIST's Smart Grid Conceptual Model

Thank You!

Questions

or

Comments?

Backup Slides…

Renewable Sources – Strong Push

- President Obama's **25by25** goal : 25% renewable sources (Wind turbines, solar panel, etc.) by 2025
  - Now < 5%.
- European Union's ambitious climate and energy targets for 2020
  - These targets, known as the "**20-20-20**" targets, set three key objectives for 2020:
    - A 20% reduction in EU greenhouse gas emissions from 1990 levels;
    - Raising the share of EU energy consumption produced from renewable resources to 20%;
    - A 20% improvement in the EU's energy efficiency.

SG and M2M Communications???

Advanced Metering Infrastructure

- ??????
- Example scenarios (AMI, sensors measurement devices)