IEEE Cloud 2017

Privacy – Preserving Data Dissemination in Untrusted Cloud

Denis Ulybyshev,[1] Bharat Bhargava,[1] Donald Steiner,[2] Leon Li,[2] Jason Kobes,[2] Harry Halpin,[3] Miguel Villarreal – Vasquez,[1] Aala Alsalem,[1] Rohit Ranchal[4]

[1] **Computer Science, CERIAS, Purdue University**

[2] **Northrop Grumman**

[3] **MIT**

[4] **IBM**

Outline

- Problem Statement
- Related Work
- Core Design
- Thesis contributions
- Demonstrations and Experiments
- Future Work

ACKNOWLEDGMENT

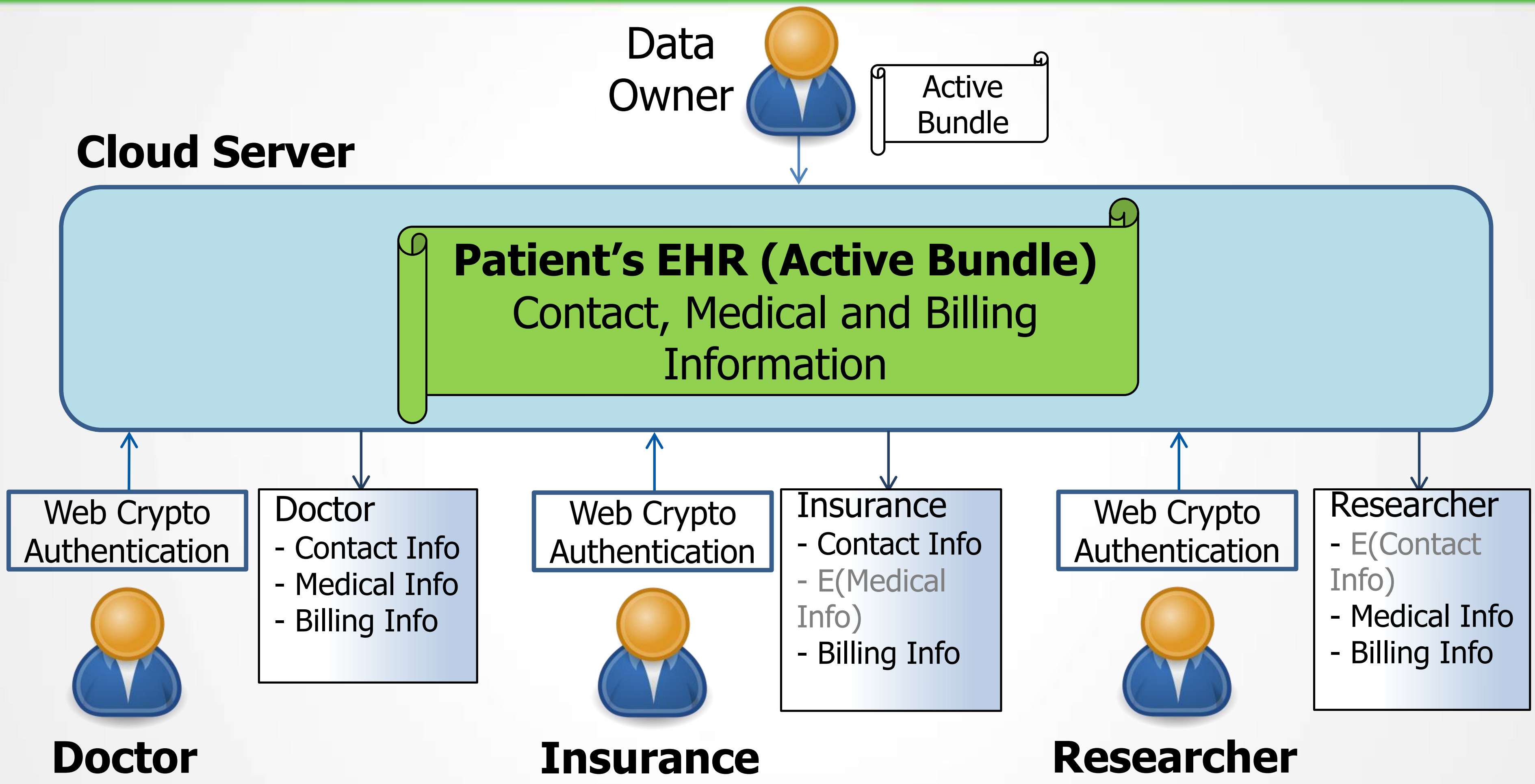
This work was funded by the Northrop Grumman Cybersecurity Research Consortium. Paper approved for public release by Northrop Grumman, Case #17-0995. The prototype was implemented in collaboration with Northrop Grumman and W3C / MIT and presented internally to Northrop Grumman in April, 2016. We are thankful to Prof. Leszek Lilien and Prof. Weichao Wang for their collaboration and valuable feedback.

Problem Statement

Privacy – preserving role – based and attribute – based data dissemination

- Authorized service can only access data items for which it is authorized
- Role – based data dissemination
- Attribute- and context – based data dissemination
- Periodic computation of trust level of services

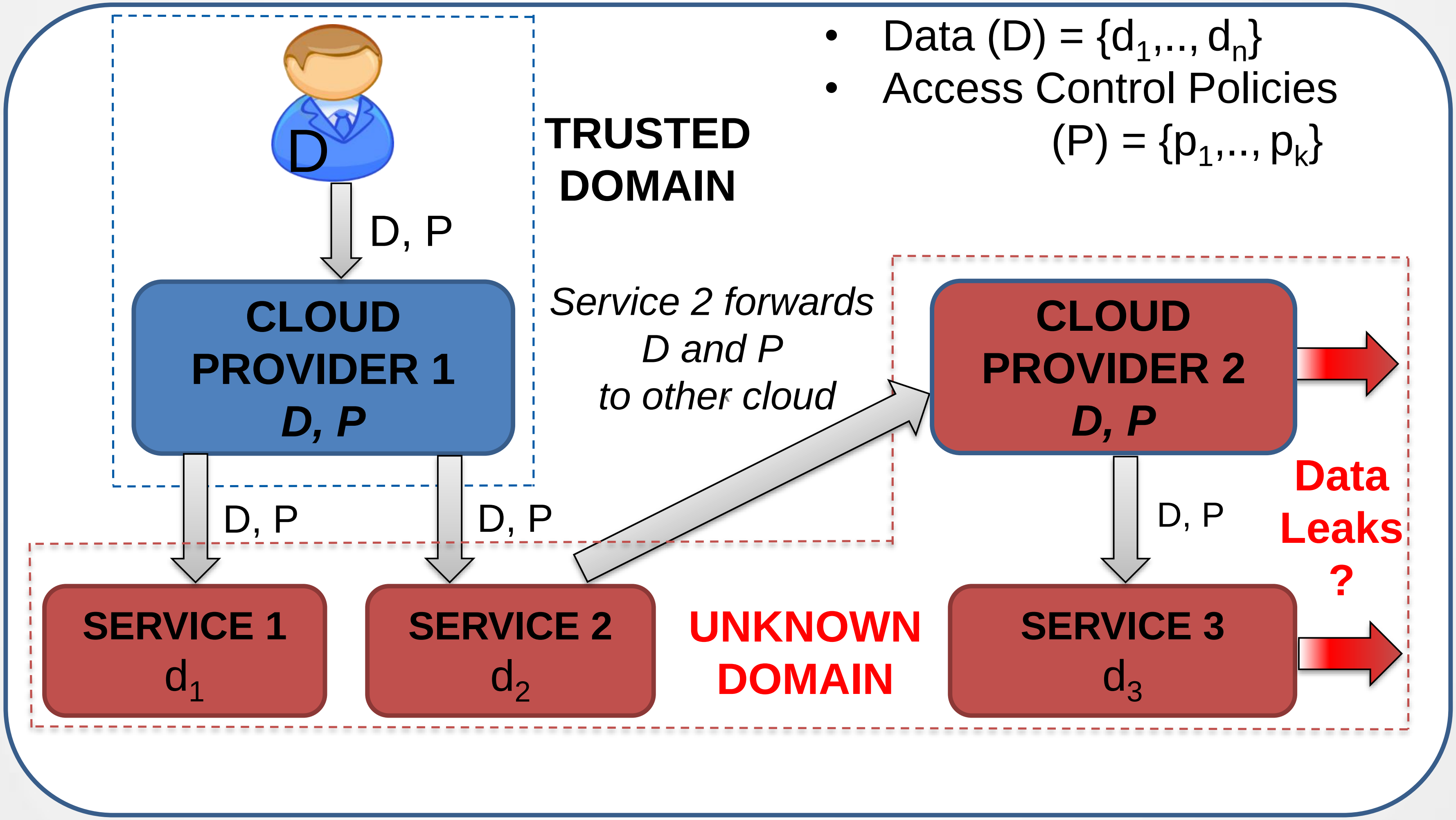
Problem Statement



Scenario of EHR Dissemination in Cloud (suggested by Dr. Leon Li, NGC)

Problem Statement

Data dissemination in SOA



Research Solutions

Privacy – Preserving Data Dissemination based on:

- Active Bundles with policies and policy enforcement engine

- Central Monitor constantly re-computing trust level of services

- Secure Browser with detection of its cryptographic capabilities

R. Ranchal, D. Ulybyshev, P. Angin, and B. Bhargava. "Policy-based Distributed Data Dissemination," *CERIAS Security Symposium, April 2015* ***(Best poster award, 1 out of 43)***

Research Solutions

Features:

- Is independent from TTP
- Data owner’s availability is not required
- Dissemination considers client's attributes
  - Crypto capabilities of a browser
  - Trust level (which is constantly recomputed)
  - Authentication method
  - Type of client’s device
- On-the-fly data updates are supported
- Secure key generation scheme

Related Work

Policy-based Data Dissemination

- Policy enforcement at browser's side [8] *(Prof. Matteo Maffei, Saarland University, Germany)*
  - Micro-policies specified in terms of tags, used to label URLs, network connections, cookies, etc and a transfer function
  - Transfer function defines permitted operations by the browser based on tags.
  - Trust level of clients is not constantly monitored and recalculated in the data dissemination model
  - Requires browser’s code modification
  - Implemented as a Chrome plugin (MiChrome [9])

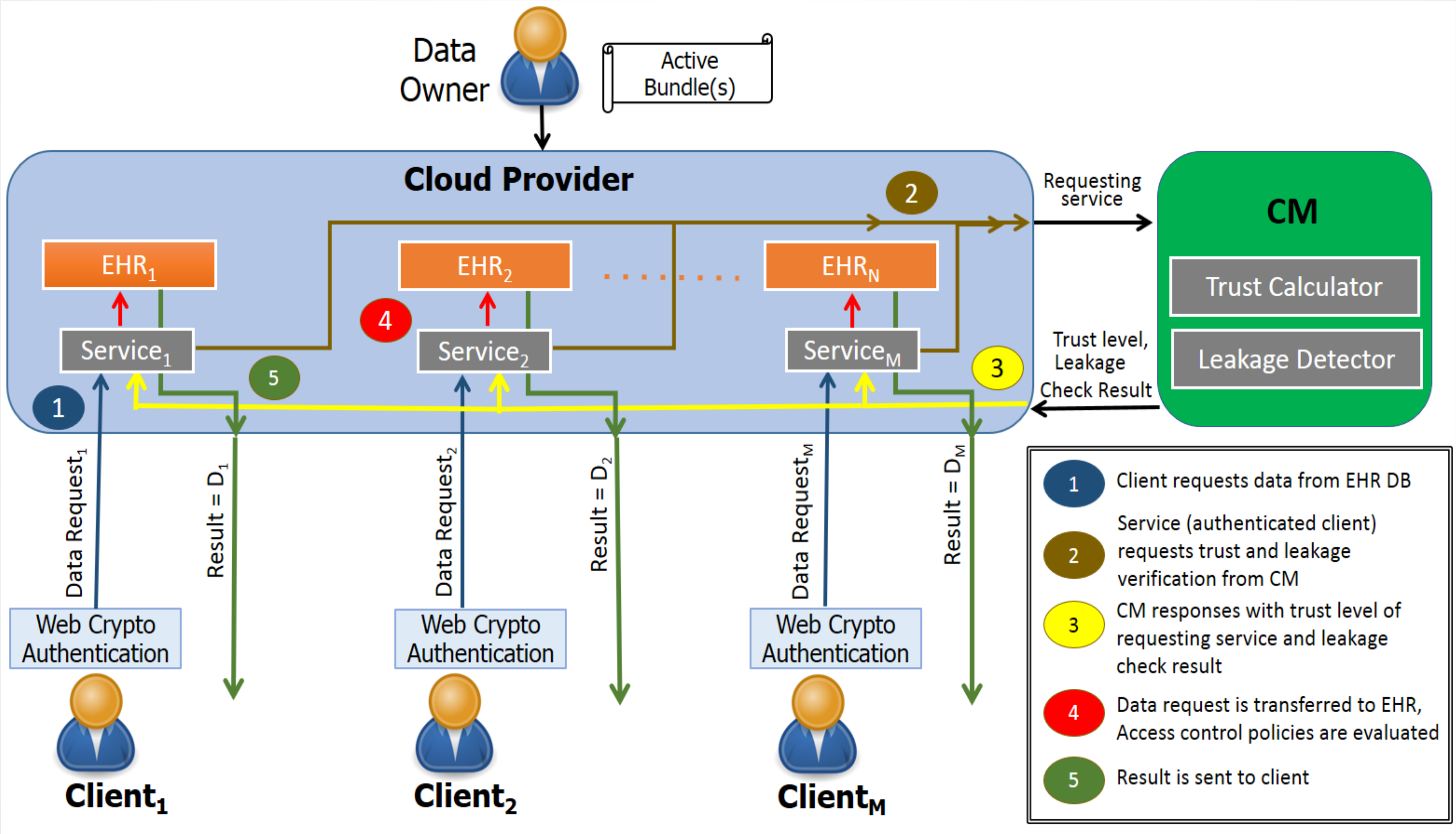
Related Work

Policy-based Data Dissemination

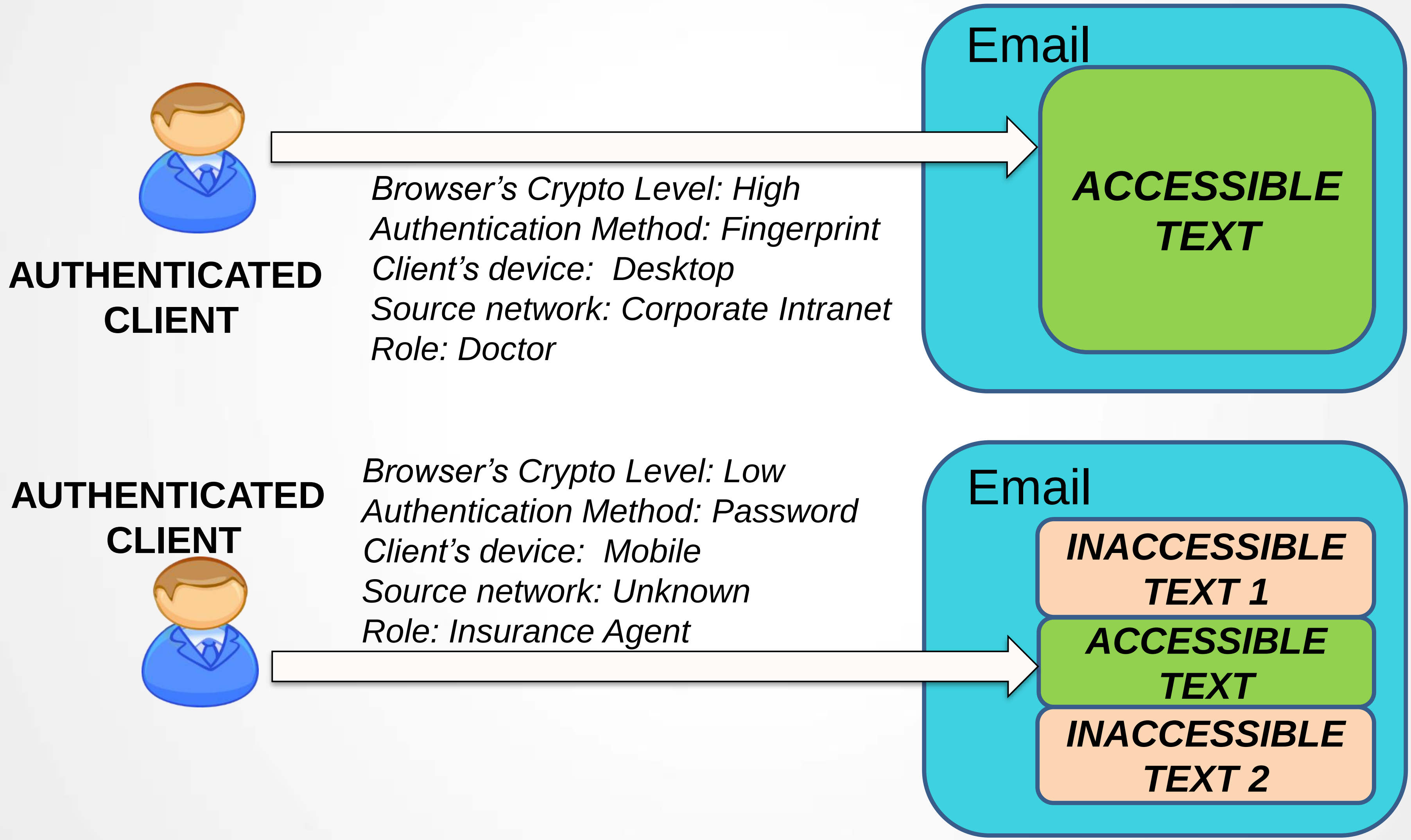
- *“Encore”* (sticky policies) system [7]
  - Policies and data are made inseparable
  - Policies are enforced by TTP
  - Policies are prone to tamper attacks from malicious recipients
  - Prone to Trusted Third Party (TTP)-related issues

- *Privacy – preserving information brokering (PPIB)* [6]
  - Divides processing among multiple brokers, no single component has enough control to make a meaningful inference from data disclosed to it
  - Prone to centralized TTP (manages keys, metadata) issues

Framework Architecture



Attribute and Role–based Data Dissemination

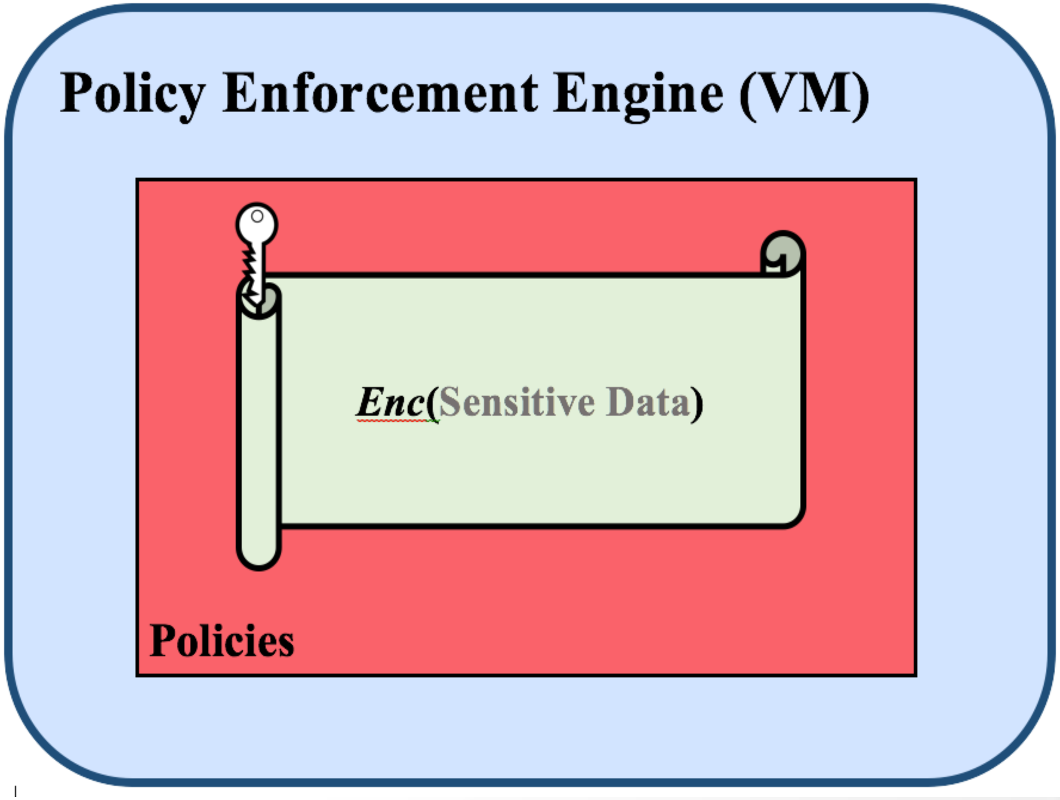


AB Core Design

Active Bundle (AB) parts [10], [11]

- *Sensitive data*:
  - Encrypted data items

- *Metadata*: describe AB and its access control policies
  - Policies [14], [15] manage AB interaction with services and hosts



- *Policy Engine* [18]: enforces policies specified in AB
  - Provides tamper-resistance of AB [1]

AB Example

Key-value pair stored in the Active Bundle:

*{ "ab.patientID" : "***Enc***(0123456789)" }*

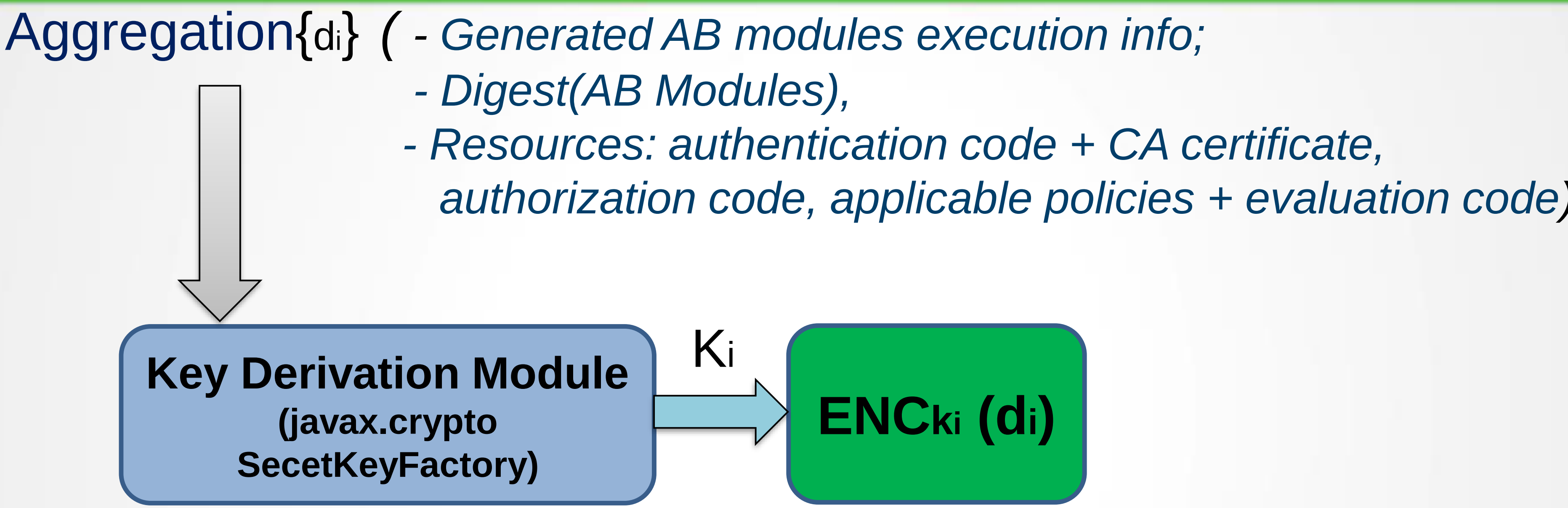
*{ "ab.name" : "***Enc***('Monica Latte')" }*

Policy Examples:

ALLOW	
Resource	patientID
Subject's Role	Doctor, Insurance, Researcher
Action	Read

ALLOW	
Resource	name
Subject's Role	Doctor, Insurance
Action	Read

Key Generation



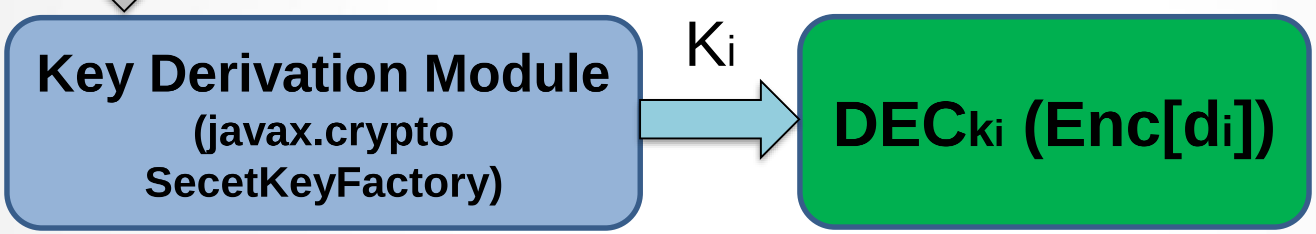
- AB Template [1] used to generate new ABs with data and policies (specified by data owner)
- AB Template includes implementation of invariant parts (monitor) and placeholders for customized parts (data and policies)
- AB Template is executed to simulate interaction between AB and service requesting access to each data item of AB

Key Generation (Cont.)

- Info generated during the execution and digest (modules) and AB resources are collected into a single value
- Value for each data item is input into a Key Derivation module (such as ***SecretKeyFactory, PBEKeySpec, SecretKeySpec*** from ***javax.crypto*** library)
- Key Derivation module outputs the specific key relevant to the data item
- This key is used to encrypt the related data item [1]

Key Derivation

Aggregation{d_i} (*- Generated AB modules execution info;*
- Digest(AB Modules),
- Resources: authentication code + CA certificate, authorization code, applicable policies + evaluation code)



- AB receives data item request from a service
- AB authenticates the service and authorizes its request (evaluates access control policies) [1]

1. "Cross-Domain Data Dissemination and Policy Enforcement", R. Ranchal, PhD Thesis, Purdue University, Jun. 2015.

Key Derivation (Cont.)

- Info generated during the AB modules execution in interaction with service, and digest (AB modules) and AB resources are aggregated into a single value for each data item [1]
- Value for each data item is input into the Key Derivation module
- Key Derivation module outputs specific key relevant to data item
- This key is used decrypt the requested data item
- If any module fails (i.e. service is not authentic or the request is not authorized) or is tampered, the derived key is incorrect and the data is not decrypted

Other Key Distribution Methods

- Centralized Key Management Service
  - TTP used for key storage and distribution
  - TTP is a single point of failure
- Key included inside AB
  - Prone to attacks!

Tamper Resistance of AB

- Key is not stored inside AB [2]
- Separate symmetric key is used for each separate data value
- Ensure protection against tampering attacks

Aggregation{d_i} (*Execution info; Digest(AB Modules); Resources*) → **Key Derivation Module** $\xrightarrow{K_i}$ **DEC$_{ki}$ (d_i)** $\xrightarrow{d_i}$

Aggregation{d_i} (***Tampered (*** *Execution info; Digest(AB Modules); Resources* ***)***)

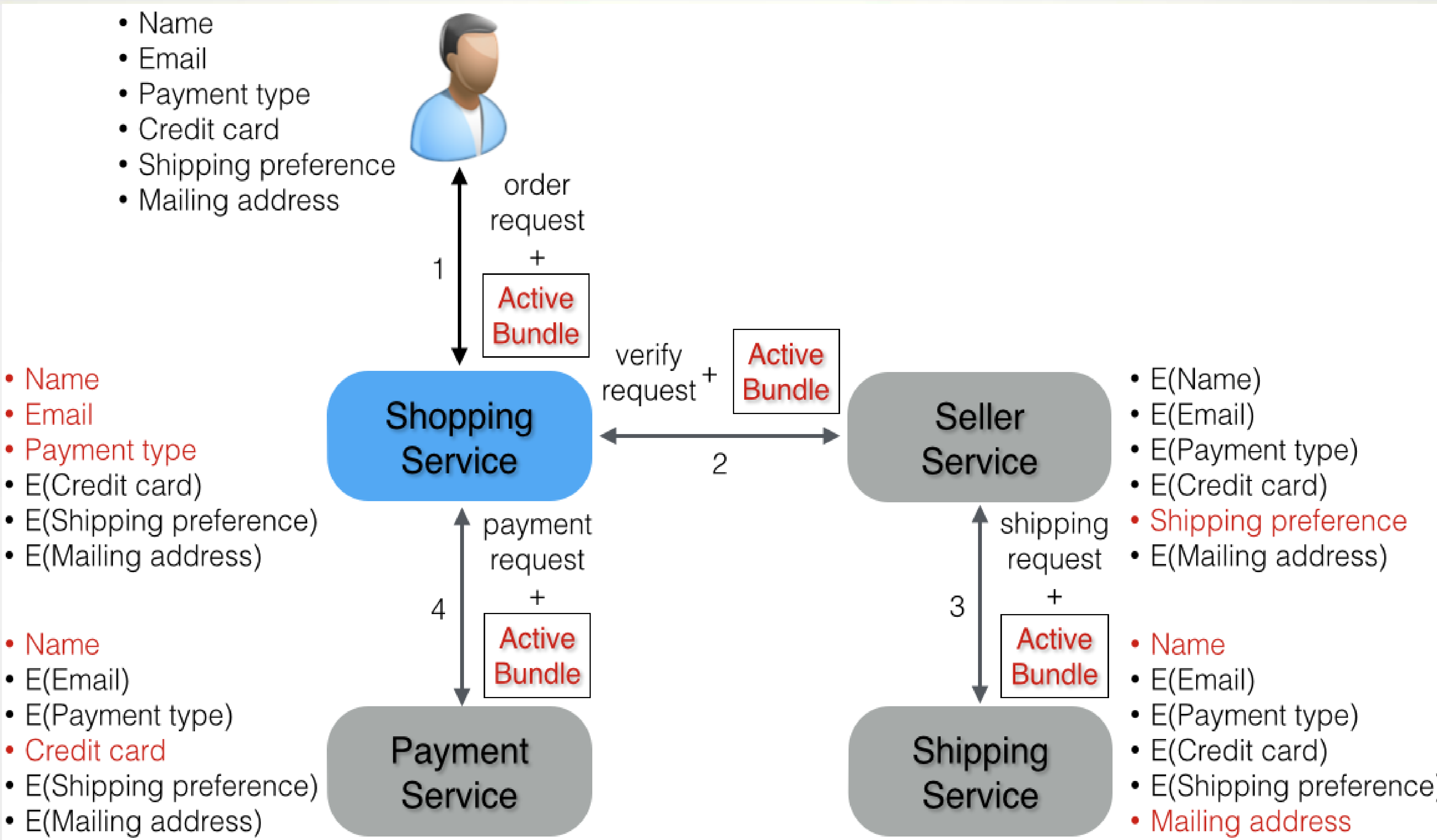
→ **Key Derivation Module** $\xrightarrow{K'_i}$ **DEC$_{k'i}$ (d_i)** $\xrightarrow{\text{wrong } d_i}$

AB Use Cases

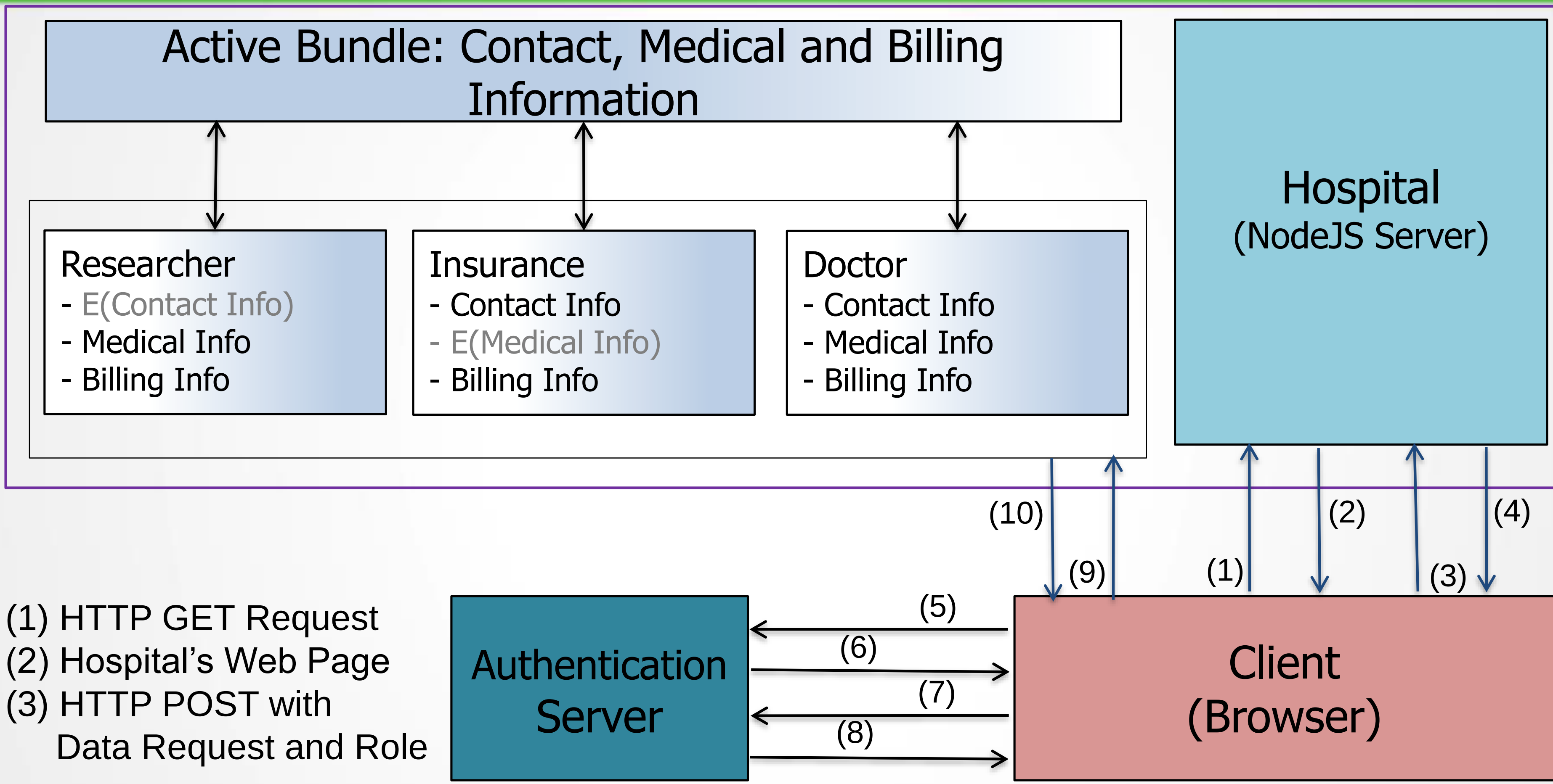
- **Hospital Information System (collection of EHRs)**
  - Doctor, Researcher and Insurance are authorized for different parts of patient's EHR [5]
  - Database of EHRs is hosted by untrusted cloud provider
- **Secure Email**
  - Email is AB
  - Entire email can be sent to the whole mailing list
  - Recipients are authorized for different fragments of email
  - It is guaranteed for the sender that each recipient will only see those email fragments it is authorized for
  - No need for multiple mailing lists for different authorization levels
- **Online shopping**
  - Decentralized data accesses: data can travel across the services

C. Qu, D. Ulybyshev, B. Bhargava, R. Rohit, and L. Lilien. "Secure Dissemination of Video Data in Vehicle-to-Vehicle Systems." 6th Intl. Workshop on Dependable Network Computing and Mobile Systems (DNCMS2015), Sep. 2015

AB Use Cases: Online Shopping



NGC TechFest'16 Demo: Electronic Health Record Dissemination in Cloud



(1) HTTP GET Request
(2) Hospital's Web Page
(3) HTTP POST with
Data Request and Role

(4) HTTP 302 with AB Request and Role
(5) HTTP Get Request
(6) AS Web Page
(7) HTTP POST with Credentials
(8) HTTP 302 with Ticket
(9) HTTP Get Request with Ticket
(10) Data provided by AB

Data dissemination features

Data Dissemination based on:

- Access control policies
- Trust level of a subject (service, user)
- Context (e.g. emergency vs. normal)
- Security level of client’s browser (crypto capabilities) [16], [17]
- Authentication method (password-based, fingerprint, etc)
- Source network (secure intranet vs. unknown network)
- Type of client’s device: desktop vs. mobile (detected by Authentication Server)

Lightweight encryption

- Can be used in Active Bundle instead of regular AES [1]

Cipher	Key size [bits]	Block size [bits]	Throughput at 4 MHz [kbit/sec]	Relative Throughput (% of AES)
Hardware-oriented block ciphers				
DES	56	64	29.6	38.4
DESXL	184	64	30.4	39.3
Hight	128	64	80.3	104.2
Software-oriented block ciphers				
AES	128	128	77.1	100.0
IDEA	128	64	94.8	123

Notes

1. Assumption: hardware and OS are trusted

2. Data is extracted from Active Bundle at a server side and send to client via https

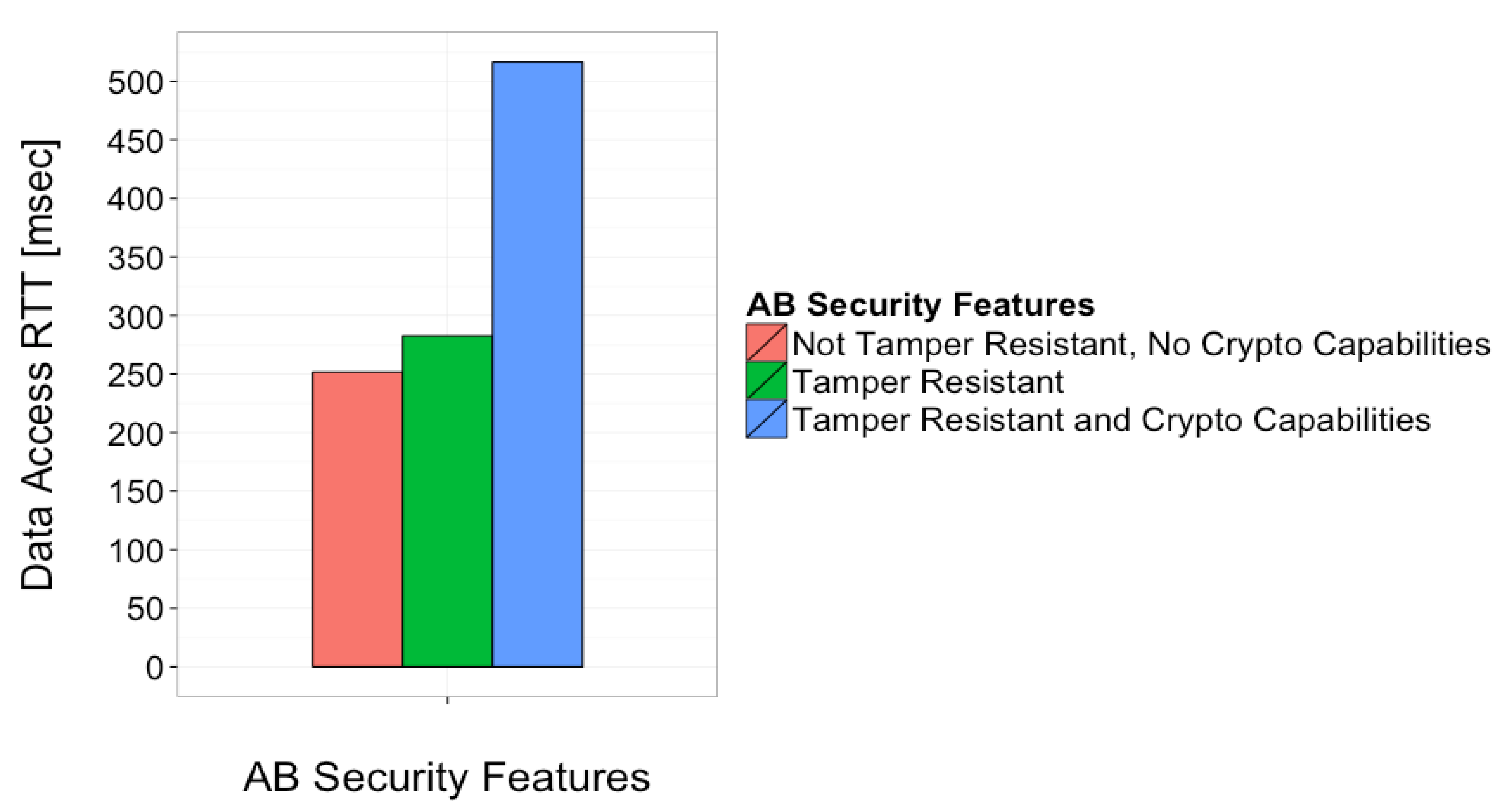
    - Data can't be tampered

Contributions

Contributes to Data Privacy, Integrity and Confidentiality

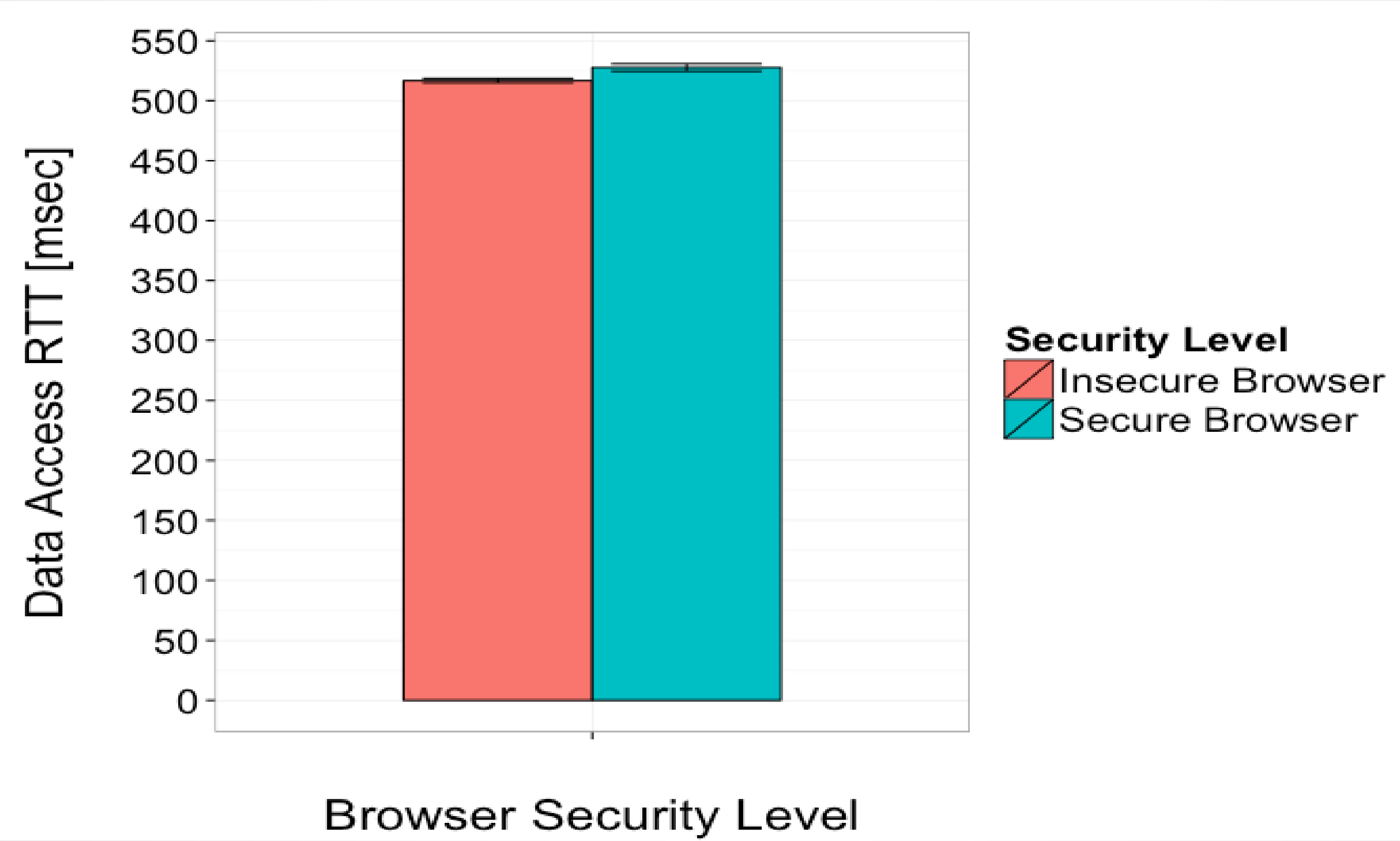
- Dissemination does not require data owner’s availability
- TTP-independent for recipient’s key generation
- Trust level of subjects is constantly recalculated
- On-the-fly key generation
- Supports data updates for multiple subjects
- Agnostic to policy language and evaluation engine
- Tamper-resistance: data and policies integrity is provided
- Compatible with industry-standard SOA / cloud frameworks

Evaluation



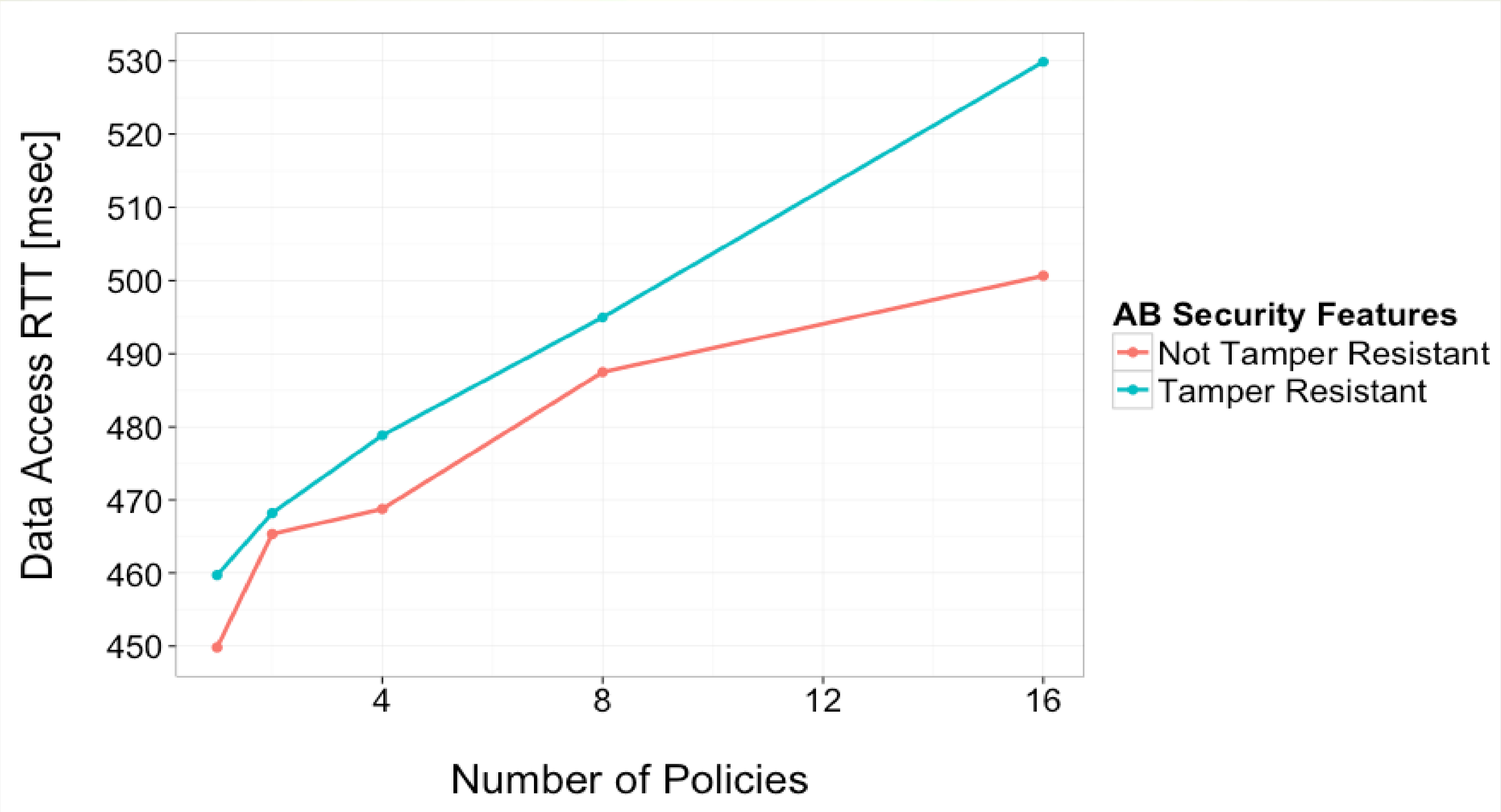
Performance overhead of Active Bundle with detection of browser's crypto capabilities on / off

Evaluation



Performance overhead of Active Bundle for data request from insecure / secure browser

Evaluation



Performance overhead of Active Bundle, hosted by Google Cloud

Deliverables

- **Prototype implementation:**
  - Privacy – Preserving Data Dissemination Prototype
  - Active Bundle Module
    - AB implementation as an executable JAR file
    - AB API implementation using Apache Thrift RPC framework
    - Policy specification in JSON and evaluation using WSO2 Balana

Source code: http://github.com/Denis-Ulybysh/absoa16

- **Documentation:**
  - Deployment and user manual
  - Demo video [13] *"Data dissemination/provenance in untrusted cloud"*

Future Work

- Lightweight encryption schemes in Active Bundle instead of AES
- Isolated AB Execution (Linux Docker Containers)
- Data Leakage Detection
- Encrypted Search over Database of Active Bundles

References

1. R. Ranchal, “Cross-domain data dissemination and policy enforcement,” PhD Thesis, Purdue University, Jun. 2015
2. C. Qu, D. Ulybyshev, B. Bhargava, R. Rohit, and L. Lilien. "Secure Dissemination of Video Data in Vehicle-to-Vehicle Systems.” 6th Intl. Workshop on Dependable Network Computing and Mobile Systems (DNCMS2015), Sep. 2015
3. L. Lilien and B. Bhargava, ”A scheme for privacy-preserving data dissemination,” IEEE Trans. on Systems, Man and Cybernetics, Part A: Systems and Humans, vol. 36(3), May 2006, pp. 503-506.
4. R. Ranchal, D. Ulybyshev, P. Angin, and B. Bhargava. “Policy-based Distributed Data Dissemination,” *CERIAS Security Symposium, April 2015* ***(Best poster award)***
5. D. Ulybyshev, B. Bhargava, L. Li, J. Kobes, D. Steiner, H. Halpin, B. An, M. Villarreal, R. Ranchal. “Authentication of User’s Device and Browser for Data Access in Untrusted Cloud,” *CERIAS Security Symposium, April 2016*.
6. F. Li, B. Luo, P. Liu, D. Lee, and C.-H. Chu, “Enforcing secure and privacy- preserving information brokering in distributed information sharing,” IEEE Transactions on Information Forensics and Security, vol. 8, no. 6, pp. 888–900, 2013.
7. S. Pearson and M. C. Mont, “Sticky policies: an approach for managing privacy across multiple parties,” IEEE Computer, no. 9, pp. 60–68, 2011.
8. S. Calzavara, R. Focardi, N. Grimm, M. Maffei, “Micro-policies for web session security”. *Computer Security Foundations Symp. (CSF), 2016 IEEE 29th* (pp. 179-193), Jun. 2016
9. Anonymus, “Micro-policies for web session security,” 2016, available at https://sites.google.com/site/micropolwebsese, accessed: Feb.2017

References

10. L. Ben Othmane and L. Lilien, “Protecting privacy in sensitive data dissemination with active bundles,” 7-th Annual Conf. on Privacy, Security and Trust (PST 2009), Saint John, New Brunswick, Canada, Aug. 2009, pp. 202-213

11. L. Lilien and B. Bhargava, ”A scheme for privacy-preserving data dissemination,” IEEE Trans. on Systems, Man and Cybernetics, Part A: Systems and Humans, vol. 36(3), May 2006, pp. 503-506.

12. B. Bhargava, “Secure/resilient systems and data dissemination/provenance,” NGCRC Project Proposal, CERIAS, Purdue University, Aug.2016

13. D. Ulybyshev, B.Bhargava, “Secure dissemination of EHR,” demo video https://www.dropbox.com/s/30scw1srqsmyq6d/BhargavaTeam_DemoVideo_Spring16.wmv?dl=0

14. “Lightweight data-interchange format JSON,” http://json.org/ , accessed: Oct.2016

15. “eXtensible access control markup language (XACML) version 3.0,” http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-os-en.html, accessed: Oct. 2016

16. “W3C Web Cryptography API,”https://www.w3.org/TR/WebCryptoAPI/, accessed: Oct.2016

17. “Web authentication: an API for accessing scoped credentials,” http://www.w3.org/TR/webauthn , accessed: Oct.2016
http://www.regularexpressions.info/creditcard.html , accessed: Oct.2016

18. “WSO2 Balana Implementation,” https://github.com/wso2/balana , accessed: Oct.2016