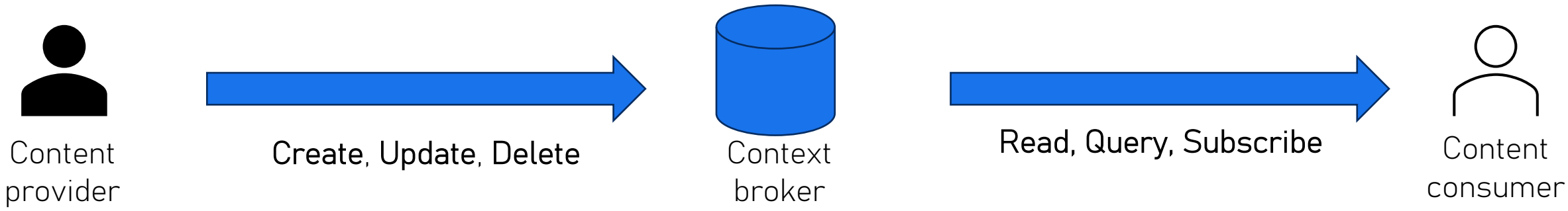
Enabling Semi-trusted Proxies for Data Spaces

Nikos Fotiou, George Xylomenos

Data Spaces

Data Spaces

- A collection of technologies that enables interoperable and secure data sharing
- Entities:
  - Context broker
  - Content provider
  - Content consumer



The ETSI NGSI-LD API – Item representation

- Data items are represented as JSON-LD objects, and they have:
  - An identifier
  - A type
  - Attributes and corresponding values
- The attributes that may exist for a type are defined in a context


```
{
    "id": "urn:ngsi-ld:Car:001",
    "type": "Car",
    "brand": {
        "type": "Property",
        "value": "BWM"
    },
    "dateVehicleFirstRegistered": {
        "type": "Property",
        "value": "2012"
    },
    "emissionsCO2": {
        "type": "Property",
        "value": "22"
    },
    "@context": [
        "https://example.com/data-models.context-ngsild.jsonld"
    ]
}
```

The ETSI NGSI-LD API – Data access

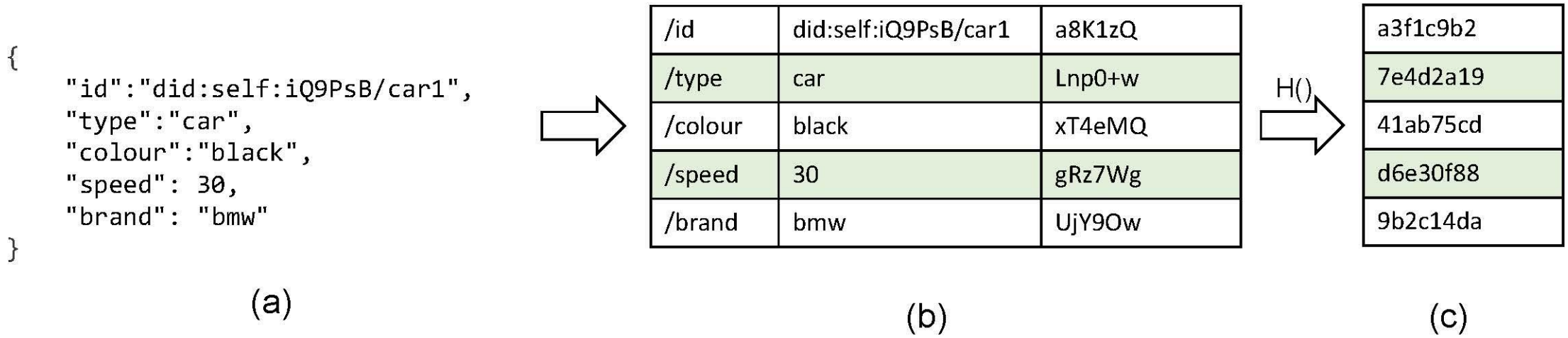
- Retrieve all/some attributes based on item id
- Retrieve all/some attributes of all items of a specific type
- Temporal queries
- Subscription to changes in attributes

Secure and Efficient Data Spaces - SeEDS

What SeEDS does better?

- Context broker is centralized ➡️ single point of failure, scalability issues
  - Information-Centric Networking architecture as an in-network, distributed “broker”
- Application layer API ➡️ Subscriptions for same event handled using multiple unicasts
  - Information-Centric networking allows ubiquitous content caching
- Context broker can manipulate data ➡️ broker should be trusted
  - Data-centric security solution based on self-certifying data
  - Selective content revelation scheme
  - Trusted and semi-trusted brokers

Selective revelation

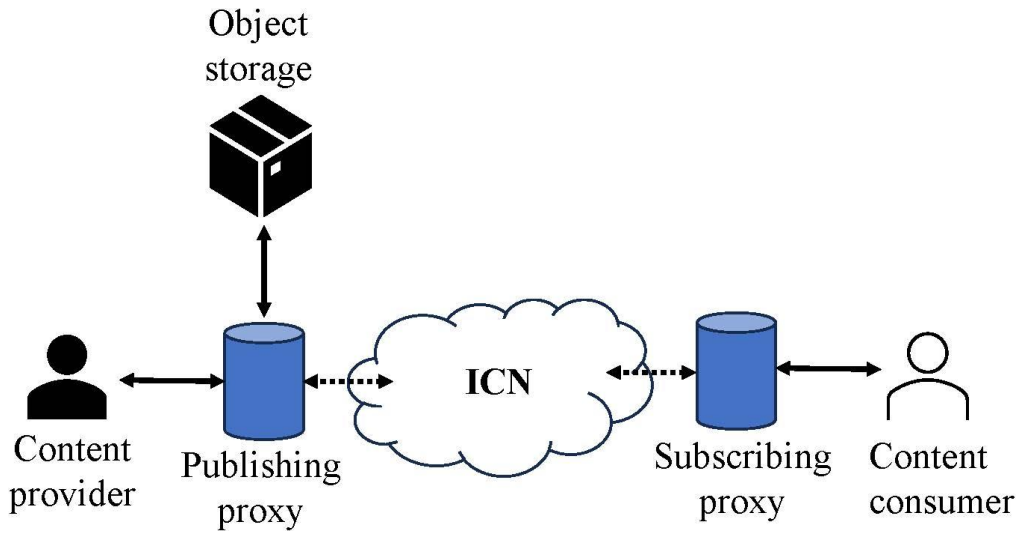


- Content producer signs entire item
  - Each attribute+salt is hashed
  - Then, list of hashes is signed
  - Broker reveals some attributes+salt and hashes of hidden attributes

Semi-trusted proxies

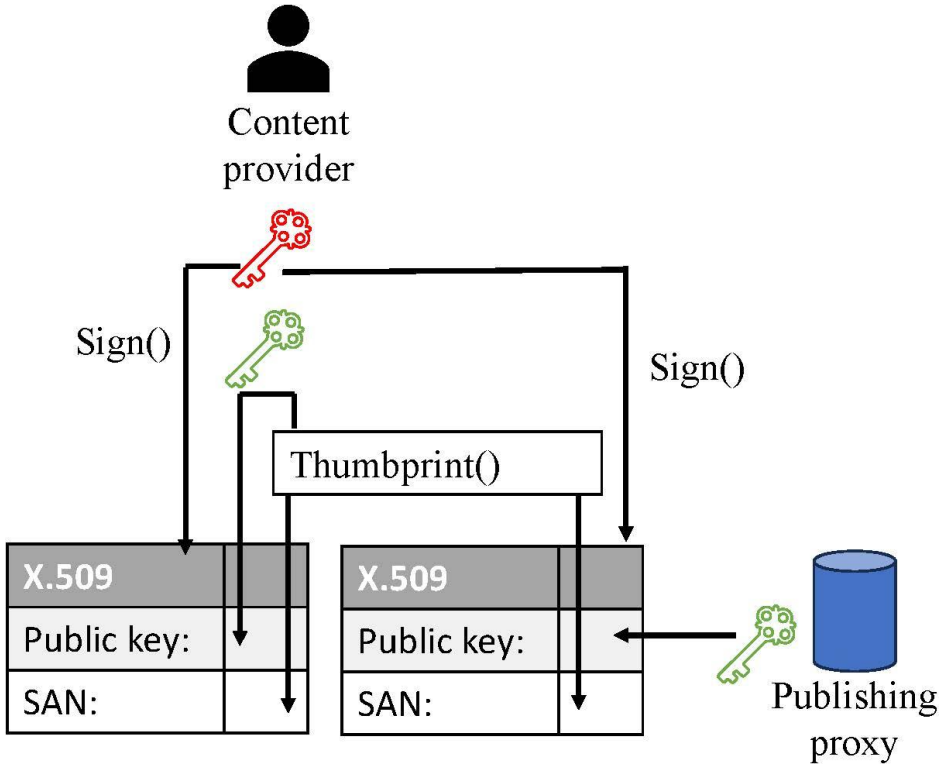
Overview

- Proxies mediate between entities
  - Translating NGSI-LD to ICN messages
- Publishing proxy: fully trusted by publisher
  - Explicitly trusted to sign publisher's data
  - Certificate binds proxy's key to provider's
- Subscribing proxy: semi-trusted by consumers
  - Reveals only selected attributes
  - Only trusted to filter content



Proxy authorization

- Provider creates public/private key pair
  - DID_p of provider is thumbprint of public key
  - Self-signs certificate including DID_p
- Publishing proxy creates public/private key pair
  - Provider issues certificate for proxy's key
  - Includes namespace allocated to proxy
- Consumers can easily verify authorization
  - Certificates may be sent with the content



Filtering & Retrieval

- A consumer asks for a partial item
  - Only some attributes are requested
  - Either due to authorization or preference
- Publishing proxy sends the entire item
  - Item can be cached
  - Can be used to satisfy different requests
- Subscribing proxy filters the attributes
- Consumer verifies signature over attributes and hashes

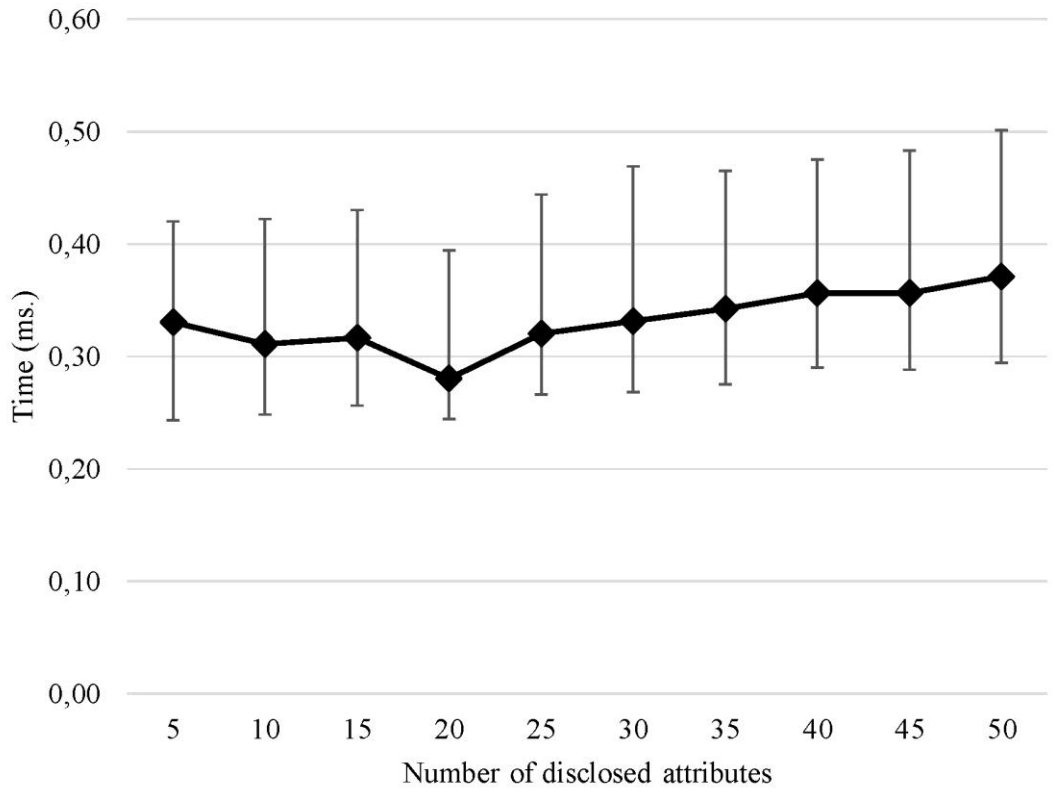
Evaluation

Performance: proof generation

- Prototype implementation
  - Python implementation of did:self
  - Jwcrypto library, NIST P-256 elliptic curve
  - Apple M4 CPU and 16 GB of RAM
- Proof generation
  - 1000 random JSON objects with 100 attributes
  - Proof generation: 2.433 ms average (6.711 max)

Performance: proof verification

- For each generated object
  - Randomly select 5-50 attributes
  - Reveal only those, verify content
  - Proof verification: <0.5 ms



Security

- Attacker manipulating network traffic
  - Any unauthorized modification is detected
- Malicious subscribing proxy
  - May suppress attributes (only reveal hashes)
  - Can mitigate by also signing attribute list (not values)
- Malicious publishing proxy
  - Wants to impersonate a trusted proxy
  - Cannot include its key in publisher's certificate

Summary

Conclusion & Future work

- SeEDS is an ICN-based Data Space implementation
  - Distributed brokers
  - Caching and multicast
  - End-to-end security
  - Selective content revelation
  - Plus: subscriptions, temporal filters, filter migration, ….
- SeEDS prototype nearing completion
  - Currently tested in NDN testbed
  - Will be released in October 2025

Thank you

xgeorge@aueb.gr