IoT Naming
- some reflections

ICNRG @ IETF100, Singapore

Anders Lindgren
Bengt Ahlgren
Börje Ohlman

About this presentation

What this presentation tries to do:

- Understand requirements for IoT naming
- List problems/issues with IoT naming
- List a set of possible tools/concepts that can be used

What it does *not* try to do:

- Prescribe how IoT naming should be done or which tools should be used

Dimensions of naming for the IoT

- Application semantics
  - There is no single name hierarchy that will support all possible applications
  - IoT data publication should be independent of a particular application/service
  - Applications/services need a way to find available data sources
- Routability
  - There needs to be a way to route to a copy of the object
    - Location dependency
    - Finding off path caches
  - Tradeoff between “application semantics” and location-dependence
    - Location-dependence improves performance/scalability
    - Application semantics can give properties like: semantically meaningful names, manageability, name persistence
- Name Persistence
  - Is needed for archiving, application interoperation, auditing, data recovery, reproducibility (e.g. of experiments)
  - is in the general case in conflict with routability as routability implies location dependency
- Time
  - Naming time-series of IoT data - real-time and historic
  - Finding IoT data from a specific time

Shortcomings of Hierarchical Content Organization

Sometimes, a document can be part of multiple collections

- /personal/joe/2017/NYmarathon/time-location.gpx
- /personal/joe/marathons/NYmarathon/2017/time-location.gpx
- /organizer/NYmarathon/2017/ranking/1/time-location.gpx
- /organizer/NYmarathon/2017/joe/time-location.gpx

-> Different access patterns for a document require different ICN names.

Slide from the presentation of:
[Tschudin] Christian Tschudin and Claudio Marxer *Improved Content Addressability Through Relational Data Modeling and In-Network Processing Elements*, SIGCOMM ACM ICN-2017

Corollary

- No single naming convention fits all needs
  - E.g. the marathon example discussed in the ACM ICN-2017 paper [Tschudin] on using NFN to create database functionality for ICN.
- What to do? Use multiple names with different properties for the same data?
  - Names can be optimised for:
    - global uniqueness and persistency
    - time series (when applicable)
    - routability
    - manageability
      - E.g. hashing of object gives easily a persistent and unique object, but is hard to organise (time sequence, application semantics, etc.)
  - Do we need a way to map/translate between these alternative names?

[Tschudin] Christian Tschudin and Claudio Marxer *Improved Content Addressability Through Relational Data Modeling and In-Network Processing Elements*, SIGCOMM ACM ICN-2017

Time - prerequisites

- Time is important for IoT data!
  - IoT data is connected to a point in time
  - Time-series data is common in IoT applications
    - E.g., sensor values taken at regular time intervals
- Access to latest value (real time) as well as historic data
  - Time series values should be immutable and individually named
  - Caching should work as expected (e.g. for latest value)
  - Should be possible to request historical data by name
  - Names should be persistent (for consistency and auditing)

Naming time-series values - tradeoffs

- Using sequence numbers
  - P**ros:** Easy to generate, (normally) no “holes” in sequence number space, no prescribed time resolution, flexible mapping to application needs, possible to (binary) search the sequence space for particular values
  - **Cons:** sensor restart might result in restarting the sequence number, mapping needed to absolute time, non-trivial to find the current (latest) value, or the value at a particular time
- Using absolute time
  - **Pros:** easy to find value at a particular time, including latest
  - **Cons:** clients need to know time granularity - no possibility to search the name space for particular values, possibly a prescribed time resolution, sensor devices need good real time
- Neither of the above
  - Not included in name, only in data/metadata object, e.g. using hash of object as name
- Conclusion?
  - Seems difficult to always prescribe one method over the other
  - **Another case for multiple names for the same data!**

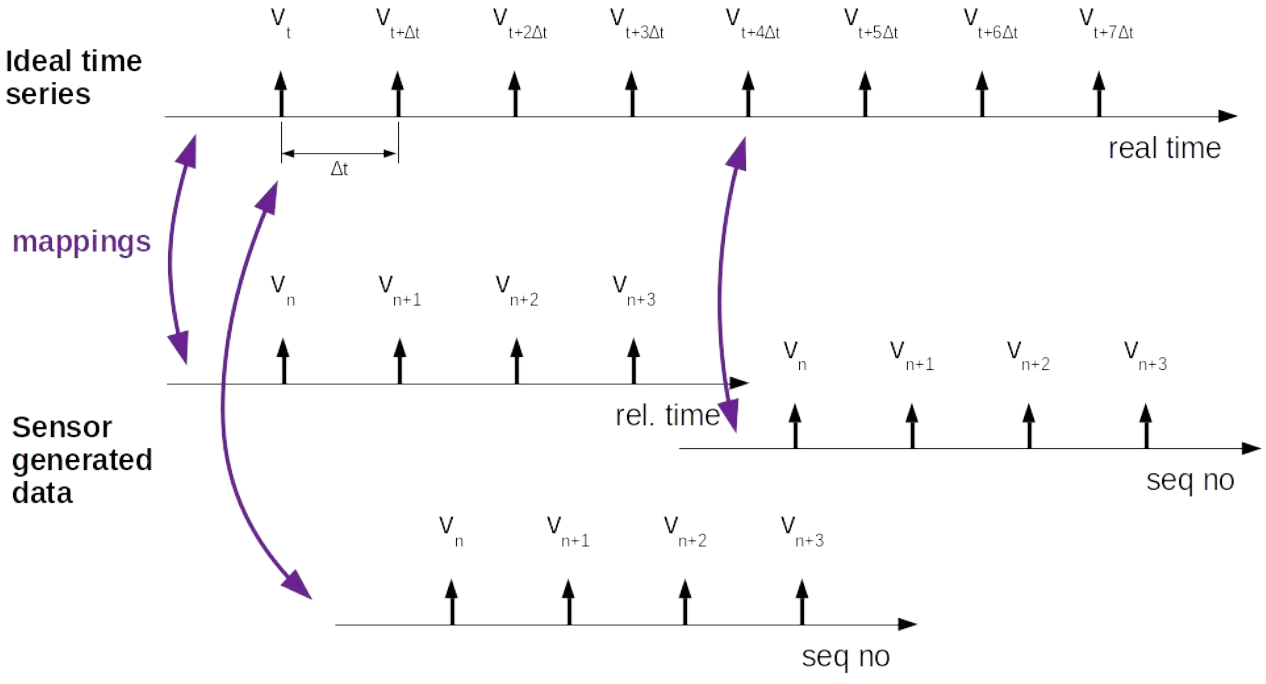
Need for time-series mapping

Sensors generate imperfect time series

- Multiple devices in same location and/or replacement devices
- May have sequence numbers due to real time unavailability
- Restarting sequence numbers / relative time due to no persistent storage

Applications/services desire “ideal” time series

- **Need name mappings between sensor-specific time series “snippets” and ideal time series!**



How to create and learn about names

- Creating names/prefixes at sensor device, alternatives:
  - Either the sensor gets a name prefix (possibly, but not necessarily unique) from some registration/configuration service,
  - Or, it creates its prefix by itself, if so this prefix needs to be made known to other entities, e.g. through some directory service
  - Could publish under multiple names
- Ways for clients/applications/services learn about names:
  - Service registers
  - General search (e.g. Google)
  - Offline (someone telling you)

Finding available data sources - directory service

- A directory service needs to be trustable, scalable and findable.
- Hierarchical NRSes with one directory service per administrative organisation
  - For example **/ericsson.com/NRS**, **/sics.se/NRS**, etc
    - If having routing entries for the NRS of each publisher, they could potentially be aggregated under ICN service providers, e.g. **telia/ericsson.com/NRS**, **telia/sics.se/NRS**
  - How is the NRS name suffix found? Is this just a naming convention?
  - How does an application know which NRS:es to contact, that is, under which admin domains?
    - Each service could have its own “bootstrapping” service with a specific name
      - E.g., /ericsson.com/GreenIoT-testbed through which that application gets a list with participating NRS:es

- Consider using OMA LWM2M for sensor registration/configuration
  - See draft-lindgren-icnrg-lwm2m4icn-00

How to find a specific sensor reading?

- There are a number of different needs for how we want to find a sensor value:
  - Finding the current value?
  - Finding a value for a specific time in a sequence?
    - finding event-based sensor values when the update frequency is not constant?
  - Finding a value for a specific location (not caring about specific sensor)?
- Alternative ways a sensor reading can be requested:
  - Directly by name of sensor reading
  - Use of NFN
    - Time can be included as one attribute in NFN relational datas
    - For time critical applications data can not be packed in a relational database that will give to slow access
  - Smart use of manifests
    - E.g. Express a query in the interest specifying a time range for sensor values, then the manifest lists the object names that matches the request
  - Ask for algorithm describing how readings are named

Finally:
Some discussion items on IoT naming and time

- How to handle event-based sensor values when the update frequency is not constant?
- If we need to give sensor values globally unique names, using a timestamp in the name is one obvious way to achieve that
  - Timestamp could be explicit part of the name
  - Timestamp could be part of what is hashed to get a unique name, useful e.g. for binary (e.g. on/off) sensor readings
  - Resource constrained devices might not be able to deal with very long names, especially for small data
- Values can be published under multiple names, e.g. both timestamped and sequence number
- Time sequences of sensor data generated by multiple sensors might represent the same physical phenomenon (and the same time)

Backup slides

Additional IoT naming and time issues

- How many ways to name IoT data do we need/want?
- Does the sensor give the IoT data the persistent name?
- One can have one unique component/prefix of the name that the device generates when it starts a series of sensor values and then append with some type of sequence number. The prefix needs to make it possible to map against some absolute time.
- All IoT-generated time series are finite with a defined start and end!
  - They may not be continuous
  - end time might be “now” for time series being generated
  -
-