Hierarchical Topic Aggregation for Geospatial Applications through Processing Extensions in Message Brokers

Josef Spillner
josef.spillner@zhaw.ch
Zurich University of Applied Sciences
Winterthur, Switzerland

Abstract
Publish/subscribe (pub/sub) message brokers are typically operated with multiple tenants and multiple topics per tenant. Messages sent by a producer to a topic are received by consumers who subscribed to the same topic. In many applications, topic hierarchies are a desired property. This is especially the case for geospatial applications in which regions are composed of smaller-scale regions. However, most message brokers do not support hierarchies natively. HTABroker demonstrates how a light-weight processing function deployed to an extensible message broker overcomes this limitation for several geospatial index systems.

ACM Reference Format:

1 Problem Statement
Digitalised societies and democratic participation of e-citizens, smart cities and regions, as well as autonomous driving and navigation are application fields which benefit from strong geospatial computing. In this computing paradigm, data is tagged with location information, and therefore subject to implicit relations such as is-nearby, is-neighbour-of, or consists-of. The paradigm implies aggregation functionality. For instance, each state of a country may report case numbers in case of an epidemy, and the country-wide numbers are calculated as the sum of the cases per state. Building such system for scalable use in real-time requires a powerful messaging infrastructure, and at the core of it, distributed message brokers that support propagation relations.

Existing brokers available to software engineers today do not possess these capabilities. They implement many features around reliable, scalable and low-latency messaging [2], but are not sufficiently treating geospatial information as first-class citizen. This contrasts database management systems that have possessed such capabilities for many years (e.g. PostGIS, MongoDB, Redis, Tile38) [1]. As a generalisation, the support for general hierarchical topics, channels or queues is also of concern. While hierarchical naming is typically supported to enable filtering (e.g. main.*, *), propagation and aggregation across these hierarchies is not. Table 1 summarises the capabilities of a selection of existing brokers.

Table 1. Message brokers overview; *: denotes extensibility

<table>
<thead>
<tr>
<th>Broker</th>
<th>Geospatial</th>
<th>Hierarchical</th>
</tr>
</thead>
<tbody>
<tr>
<td>NATS</td>
<td>none</td>
<td>only multi-sub via regexp</td>
</tr>
<tr>
<td>RabbitMQ</td>
<td>none</td>
<td>only multi-sub via regexp</td>
</tr>
<tr>
<td>Kafka</td>
<td>none</td>
<td>only multi-sub via list/regexp</td>
</tr>
<tr>
<td>VerneMQ</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Pulsar *</td>
<td>none</td>
<td>only multi-sub via list/regexp</td>
</tr>
</tbody>
</table>

2 Research and Technical Approach
The solution consists of (1) defining appropriate hierarchical geospatial models, (2) defining propagation and aggregation operations, and (3) implementing the solution with user-defined processing logic (functions) atop existing message brokers to benefit from the existing scalability and resilience.

2.1 Models
HTABroker supports several geospatial models. Apart from computational models whose data points are calculated on demand, further data-driven models are expressed in tree-structured data files. The current version of HTABroker integrates with four models, also called index systems:
Computational: Geohash. Each frame between two longitude/latitude pairs is represented by one or multiple geohashes which are alphabetical codes. With arbitrary precision $n \geq 1$, this results in $36^n$ entries.

Data-driven: NUTS. The Nomenclature des unités territoriales statistiques is used in the EU and associated countries. It subdivides countries into regions and districts. Apart from Local Administrative Units (LAUs), there are currently around 1500 entries defined.

Data-driven: Communes. More localised than NUTS, the Swiss Communes model exemplifies the relation between a country, its subdivisions (cantons, districts) and municipalities. It contains 2200 municipalities.

Hybrid: H3. Used by Uber navigation, regions are divided into hexagons. It contains around 8700 entries.

2.2 Messaging

Apart from simple publish/subscribe message delivery, the hierarchical topic aggregation foresees three underlying propagation directions that are configurable per application.

- Upward propagation. An event is produced in a sub-entity (e.g., geo region) and propagated to its super-entity. For instance, the registration of a foreigner as state resident leads to changed country-wide statistics.
- Downward propagation. For instance, a new country-wide alert is replicated on municipal apps that mix hierarchy levels for the benefit of the users.
- Sideward propagation. Events are propagated to neighbouring entities. For instance, a traffic jam in a city district affects rerouting decisions in nearby places.

Fig. 1 shows an example of the three propagation types with subdivisions of a country in the NUTS model, with producers (P) and consumers (C) of messages.

![Propogation directions in a NUTS scheme](image)

3 Related Work

Pallatella et al. investigated aggregation of MQTT topics using multi-topic subscription patterns which is a closely related concept but without any geospatial bindings [5]. For wide-area IoT, Livaja et al. proposed geospatial pub/sub conceptually, but do not introduce a working solution [4]. Kassab et al. built a prototype for real-time emergency that processes GPS locations, without propagation directions [3].

4 Conclusions

HTABroker is a versatile real-time messaging middleware for geospatial applications. The software is available as open source at https://github.com/serviceprototypinglab/htabroker.

Acknowledgments

This work is supported by a DIZH Fellowship under grant Smart Cities and Regions Services Enablement (SCReSE).

References


