



# Architecture of a Global Scale Alert Hub

---

2018 CAP Implementation Workshop - Wed 31st Oct  
Ian Ibbotson / e:[ian.ibbotson@k-int.com](mailto:ian.ibbotson@k-int.com) / t:@ianibbo

# Recent History (How we got started)

- Initially asked to participate and contribute specifically to work on real-time geo-spatial search of alerts and subscriptions.
- History of combining Text and Spatial Search
- Experience with Elasticsearch in a number of contexts
- Experience of microservice and serverless architectures
- Can we provide a subscription matching microservice, which given an alert shape (Polygon or Circle) can return any of the defined subscriptions whos shapes (Polygon or Circle) overlap with the alert.
  - <https://s3-eu-west-1.amazonaws.com/alert-hub-subscriptions/json>

# Elasticsearch - v1

- ES has a really great profile for cluster deployment, scale out and replication. Our target document set (Subscriptions) is not particularly dynamic, and query rates for the app are easily managed by a single node.
- We load the subscriptions
  - <https://s3-eu-west-1.amazonaws.com/alert-hub-subscriptions/json>
- Into ES and then search for them using the alert shapes
- Normally, ES is searched using URL parameters
  - `http://localhost:9200/alertssubscriptions/_search?q=sheffield`

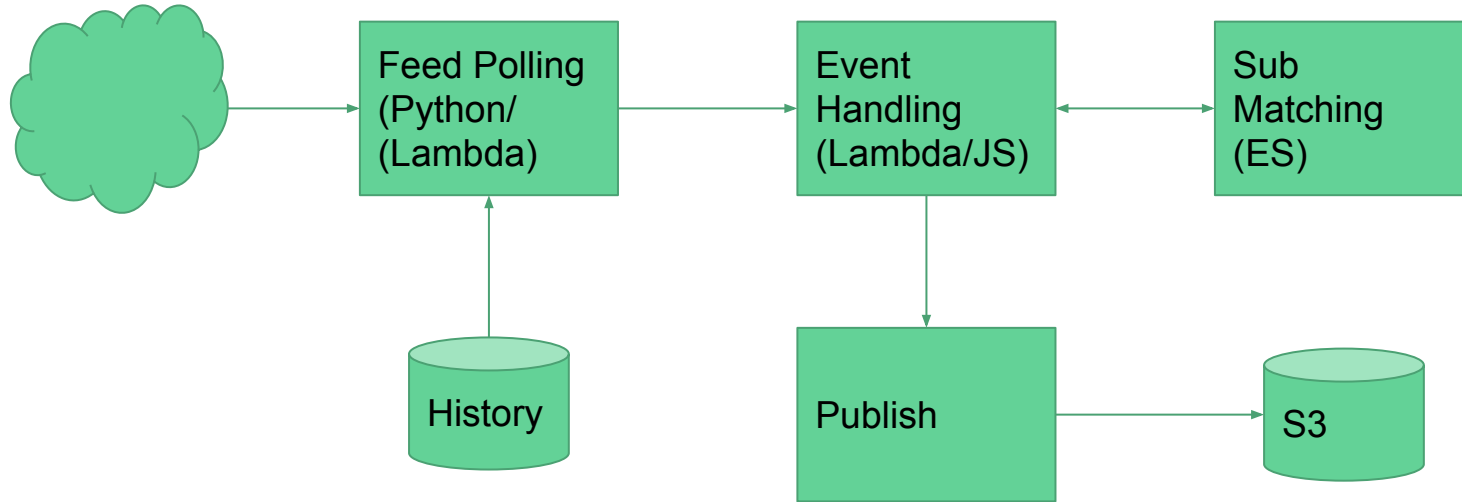
# Elasticsearch - v1

```
curl -X GET 'http://localhost:9200/alertssubscriptions/_search' -d '{
  "from":0,
  "size":1000,
  "query":{
    "bool": {
      "must": {
        "match_all": {}
      },
      "filter": {
        "geo_shape": {
          "subshape": {
            "shape": {
              "type":"circle",
              "coordinates":[-70.06,12.58],
              "radius": "20.8km"
            },
            "relation":"intersects"
          }
        }
      }
    }
  }
}
```

# Elasticsearch - v1

- .... And all was well with the world
  - This the spatial search completes in 1ms against the full set of subscriptions

# V1 Architecture then (As I recall it)



# V1 Retrospective - Overview

- Feed Polling was difficult to get right
- Feed Polling is “Bursty”
- Need to feed-back to feed providers / Error Checking / Reporting / Debugging are critically important - and not terribly easy to debug from cloudwatch logs.
- Event Handling worked OK - sometimes long delays, not clear why
- Again, debugging not so easy
- Polyglot environment is really cute!
- Sub matching seems to work OK
- Publishing seems to block - Debugging proved difficult

# V1 - Retrospective - #1 Lesson

- If you aren't in complete control of your sources, the ability to report on, log and debug your serverless handlers is really crucial.
- This is still a slightly high friction process for a small development team.
- Need diagnostic tools.



# V2 - evolution

We've evolved three projects

- FeedFacade - <https://github.com/SemwebServices/PubSubHubBubFacade>
- CAP Collator - <https://github.com/SemwebServices/CAPCollator>
- Devops - <https://github.com/SemwebServices/SWCapAlertHubDevops>

# FeedFacade

- Uses PostgreSQL to coordinate feed fetcher threads
- Capable of running on multiple nodes
- Substantial new structure to describe and track feeds and alerts
  - Each feed can have individual polling intervals
- A lightweight abstraction over RSS and ATOM that converts pull based feeds into a reactive event stream.
- Messaging substrate agnostic, but implemented with RabbitMQ
- Does nothing apart from listen for new “Item” entries in source feeds and emits an event containing that item.
- Supports HTTP HEAD, is the place we implement any special efficiency or behavior to deal with non-standard servers.

# FeedFacade

Registered Source Feeds - Mozilla Firefox

https://demo.semweb.co/feedFacade/sourceFeed/index

FeedFacade 1.0.4 Home Alert Sources Live Log Entries Topics Admin - admin -

## Registered Feeds

Text input  Has Errors Search

1 2 3 4 5 6 7 8 9 10 11 Next

Name	Enabled	Feed Status	Topics	Tags	Fetcher Status	CAP Status	Base Url	Last Completed	Poll Interval	Next Due	HTTP Last Modified	HTTP Expires	Last Error Message																																																																																																										
ai-dma-en	Yes toggle		<ul style="list-style-type: none"> <li>ai-dma-en</li> <li>allfeeds</li> <li>ai</li> <li>dma</li> </ul>	<ul style="list-style-type: none"> <li>sourceIsOfficial: true</li> <li>sourceLanguage: null</li> <li>authorityCountry: ai</li> <li>authorityAbbrev: dma</li> <li>registerUrl: http://alerting.worldweather.org/authorities.php?recid=253</li> <li>logoUrl: null</li> <li>author: null</li> <li>guid: urn:oid:2.49.0.0.660.0</li> </ul>	paused	operating	https://axa-primary.capews.com/capews/public/atom?type=cap	2018 10 29 14:31:22.057 (-754)	60000	2018 10 29 14:32:22.057	0	0																																																																																																											
<table border="1"> <thead> <tr> <th>Hour</th> <th>13</th> <th>14</th> <th>15</th> <th>16</th> <th>17</th> <th>18</th> <th>19</th> <th>20</th> <th>21</th> <th>22</th> <th>23</th> <th>0</th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> <th>6</th> <th>7</th> <th>8</th> <th>9</th> <th>10</th> <th>11</th> <th>12</th> <th>13</th> <th>14</th> </tr> </thead> <tbody> <tr> <td>Error Count</td> <td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td> </tr> <tr> <td>Success Count</td> <td>30</td><td>30</td><td>30</td><td>30</td><td>30</td><td>30</td><td>30</td><td>30</td><td>29</td><td>30</td><td>30</td><td>30</td><td>30</td><td>30</td><td>30</td><td>30</td><td>29</td><td>30</td><td>30</td><td>30</td><td>29</td><td>30</td><td>30</td><td>30</td><td>16</td> </tr> <tr> <td>New Entries</td> <td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td> </tr> </tbody> </table>														Hour	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Error Count	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Success Count	30	30	30	30	30	30	30	30	29	30	30	30	30	30	30	30	29	30	30	30	29	30	30	30	16	New Entries	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hour	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14																																																																																													
Error Count	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																																																																													
Success Count	30	30	30	30	30	30	30	30	29	30	30	30	30	30	30	30	29	30	30	30	29	30	30	30	16																																																																																														
New Entries	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																																																																																														
ag-ms-en	Yes toggle		<ul style="list-style-type: none"> <li>allfeeds</li> <li>ag-ms-en</li> <li>ag</li> <li>ms</li> </ul>	<ul style="list-style-type: none"> <li>sourceIsOfficial: true</li> <li>sourceLanguage: null</li> <li>authorityCountry: ag</li> <li>authorityAbbrev: ms</li> <li>registerUrl: http://alerting.worldweather.org/authorities.php?recid=204</li> <li>logoUrl: null</li> <li>author: null</li> </ul>	paused	operating	https://alert.metoffice.gov.ag/capfeed.php	2018 10 29 14:31:21.615 (-312)	60000	2018 10 29 14:32:21.615		javax.net.ssl.SSLHandshakeException: security.validator.ValidatorException: PKIX path validation failed: java.security.cert.CertPathValidator: validity check failed																																																																																																											

# CAPCollator

- Listens out for new events from feedFacade and reacts to them
  - Run many collators, each will remove items from the queue
- If a CAP link is found in an Item emits a CAP event
  - Event picked up by any listening collator
- Listens for CAP events
  - Fetches the source feed, parses and validates
  - Matches subscriptions
  - Emits an event for each matched subscription, Submits an Alert Indexing Record
- Listens for Subscription Match events
  - Updates a static RSS feed and publish to AWS S3 bucket
- Net effect - A loosely coupled service that converts CAP events to an index of alerts tagged by subscription, a static RSS file updated and published on S3.

# CAPCollator

CAPCollator - Emergency Alert: Aggregation Middleware - Mozilla Firefox

https://demo.semweb.co/CAPCollator/subscriptions/details/country-ad-lang-en

CAPCollator 1.0.11 Home Nearby About Admin Subscriptions Unfiltered admin

## country-ad-lang-en / Official Public alerts in English for Andorra


Subscription Info

URL	<a href="https://alert-feeds.s3.amazonaws.com/country-ad-lang-en/rss.xml">https://alert-feeds.s3.amazonaws.com/country-ad-lang-en/rss.xml</a>	Filter Type	polygon
Alert Count	13 (10 shown)	Filter Geometry	[[[1.416667, 42.433333], [1.783333, 42.433333], [1.783333, 42.65], [1.416667, 42.65], [1.416667, 42.433333]]]

Text input  Search

Showing alerts 1 to 10 of 13

1 2 Next



Map Satellite

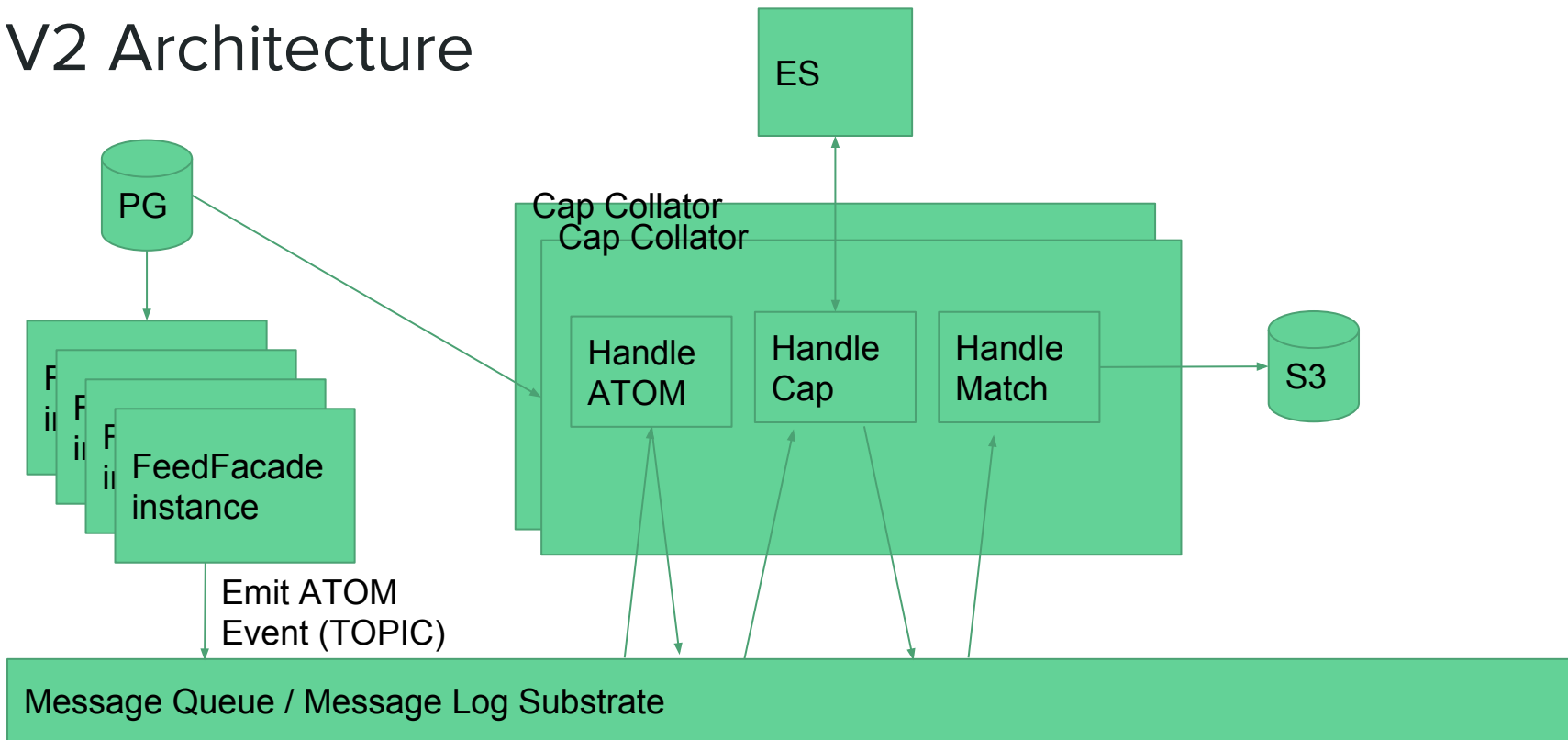
Yellow Low temperature Warning issued from for España - Girona

Alert Identifier	2.49.0.0.724.0.ES.181029102355.137139330		
Alert Sender	<a href="http://www.aemet.es/es/eltiempo/prediccion/avisos">http://www.aemet.es/es/eltiempo/prediccion/avisos</a>		
Times:	Date on alert	2018-10-29T10:23:55+00:00	
	CC-ATOM-notified	2018-10-29 10:38:55.399 UTC	
	CC-HTTP Get completed	2018-10-29 10:39:07.190 UTC	
	CC-parse complete	2018-10-29 10:39:07.191 UTC	
	CC-emit CAP event	2018-10-29 10:39:07.193 UTC	
	CC-spatial-processing-complete	2018-10-29 10:39:07.605 UTC	
Source	<a href="http://meteoalarm.eu/CAP/POLY/ES_29102018_137139330.cap.xml">http://meteoalarm.eu/CAP/POLY/ES_29102018_137139330.cap.xml</a>		

# Devops Project

- Devops project gives us a docker container architecture with all components
  - `docker-compose -f ./docker-compose-dev-setup.yml up`
  - Gives us a system and
  - [http://localhost:9200/alertssubscriptions/\\_search?q=\\*](http://localhost:9200/alertssubscriptions/_search?q=*)
  - Gives us alters indexed

# V2 Architecture



# Recent Issues / Current Problems

- Replace static feed writer with a delayed write queue
- Date/Time formats
- Occasional Lag Spikes



# Future Developments..

- Try Kafka
- Alert Publishing App to Drive Integration tests and Mock Feeds
- Feed Owner Alerting - Broken Feeds, 404s, invalid polygons, etc.