



UNIVERSITY *of* LIMERICK
OLLSCOIL LUIM NIGH

Project Log

Spring Semester

Week 1

MEng Information & Network Security

Thomas Flynn

16117743

Project Supervisor: Sean McGrath

23/01/17- 29/01/17

1 Log Entries

1.1 Entry 24/01/17

Today I met with my supervisor. We discussed various aspects of the project. My supervisor advised me to focus on the back-end architecture for now as the scope of the project is too large.

1.2 Entry 25/01/17

Today I ran into some problems setting up the academic free trial. I had to email the Bluemix support team in order to resolve the issue.

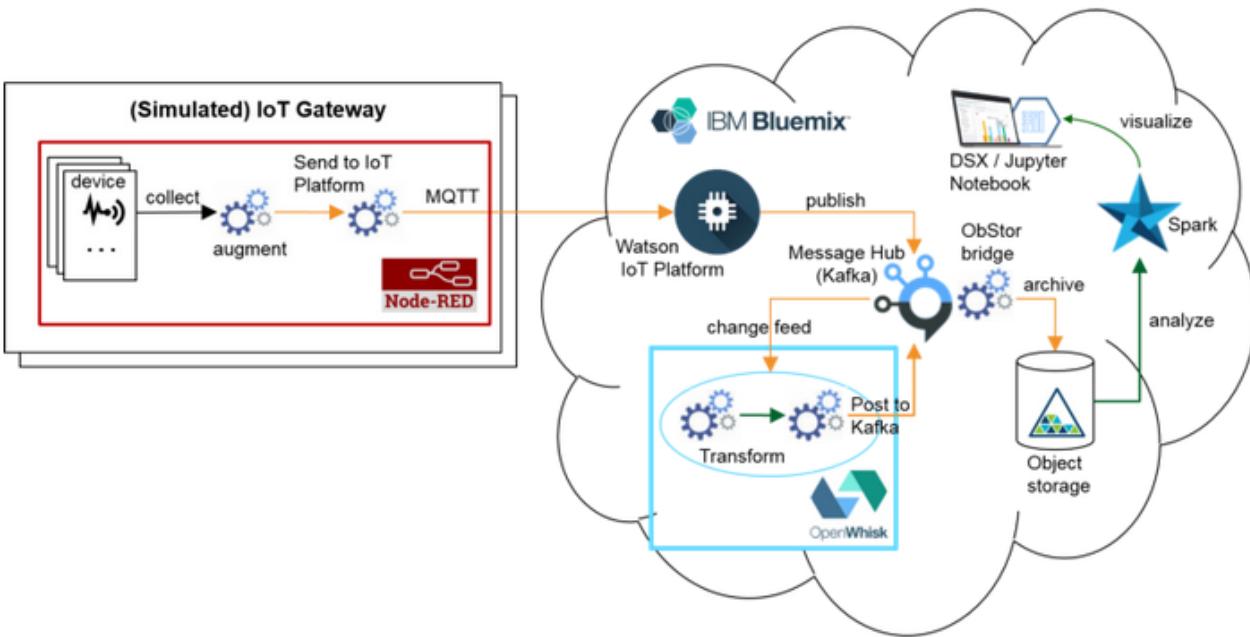
1.3 Entry 27/01/17

Today I received an email from the Bluemix sales team. They informed me that the reason I couldn't enter the Academic promo code was because I changed my account from a "free trial" to a "pay as you go" type. This was slightly frustrating as I had no choice but to convert to a "pay as you go" once my 30 day free trial ran out. When I logged into Bluemix, I was not able to access the dashboard as I was greeted with a pop up window that forced me to sign up for a "pay as you go" account. The issue should be resolved within a few days hopefully.

I carried out some extensive research today in the area of OpenWhisk and Node-Red. After a few hours of reading I came across the following link.

<https://medium.com/openwhisk/transit-flexible-pipeline-for-iot-data-with-bluemix-and-openwhisk-4824cf20f1e0#.vxa7c3upd>

This was an important discovery because the link describes the architecture for a project that is similar to my own. The key difference being that it uses Message Bus, Object Storage and Spark Service in order to track taxis. A summary of the link is given below.



1. **Node-RED** (simulated device): the ‘inject node’ triggers the flow at given intervals (e.g., every second). A sample event JSON is generated (by a custom ‘function node’). It is then aggregated using the ‘join’ node, compressed (using a custom function node), and sent to the Watson IoT Platform (using MQTT with properly configured orgId, device type, event type, device id and security token).
2. **Watson IoT Platform** (with historical data storage extension to Message Hub): receives MQTT events from devices (after proper auth validation), augments them (e.g., adding timestamp and context metadata) and publishes in small batches to a Kafka topic (Message Hub).
3. **OpenWhisk** (with custom transformation action): the Message Hub feed provider in OpenWhisk subscribes to messages in the iotp Kafka topic, aggregates small batches of messages and sends them as a payload to the OpenWhisk trigger, which in turn triggers a sequence of actions performing the message transformation and publishing the results to the transformed Kafka topic.
4. **Message Hub** (bridge to Object Storage): aggregates messages published to the transformed topic according to a specified policy (e.g., files of up to 1 hour or 1MB of data), and uploads the files to Object Storage according to the specified layout (e.g., applying date-based partitioning).
5. **Spark Service** (part of Data Science Experience platform): retrieves historical data from Object Storage, applying SQL or other interfaces—according to the needs of the particular batch analytics job.

1.4 Entry 28/01/17

Today I refreshed my understanding of serverless architecture. I researched couchDB, Graph and Cloudant databases. I also looked at the GeoJSON format.

2 Tasks completed

- Meet with supervisor

- Report Template

- Research MQTT

- Research Kafka

- Find similar project

- Research databases

3 GIT Repositories

3.1 INS-Thesis-Documentation

4 Trello boards

4.1 Updated labels



4.2 Board at the start of the week

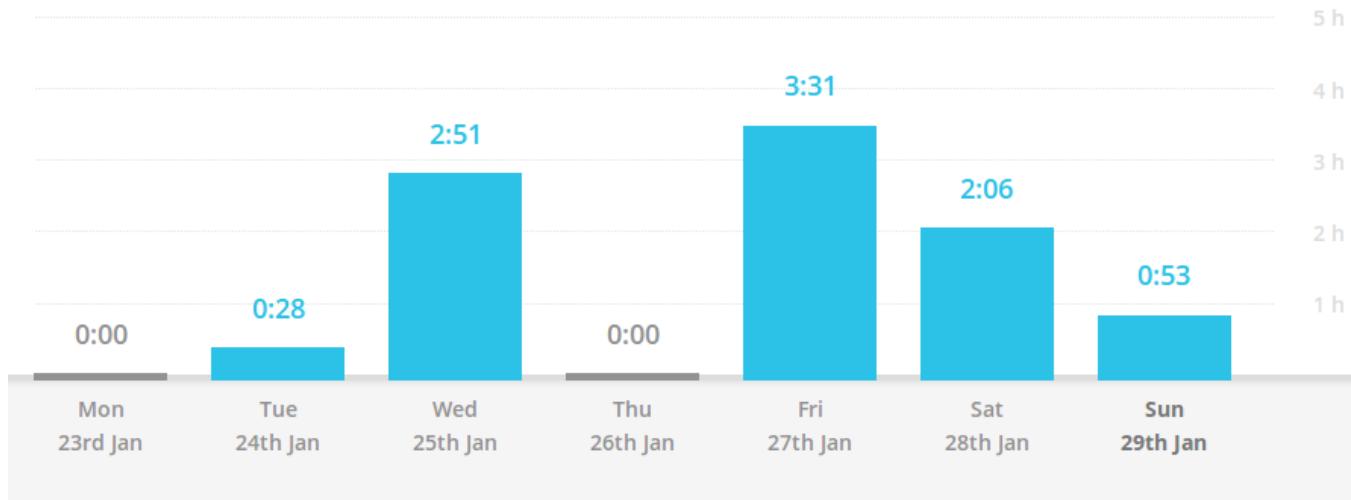
Later	This Week	Today	next action	Done
Architectural Diagram	Meet with supervisor	Add a card...	Add a card...	Add a card...
Detailed synopsis	Academic 6 month trial			
Best case scenario	Report Template			
Worst case scenario	Research Kafka			
Detailed plan	Research MQTT			
Week 2 Log	Find Similar Project			
Research ELK Stack	Research databases			
Setup GitHub Repositories	Week 1 log			

4.3 Board at the end of the week

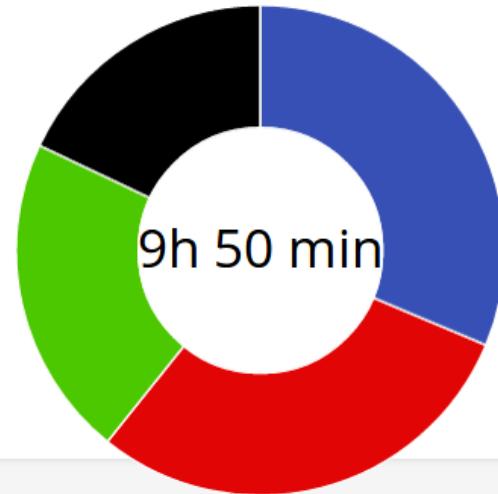
Later	This Week	Today	next action	Done
Week 2 Log	Academic 6 month trial	Add a card...	Add a card...	Week 1 log
Research ELK Stack	Architectural Diagram			Research databases
Setup GitHub Repositories	Detailed synopsis			Research Kafka
Add a card...	Best case scenario			Find Similar Project
	Worst case scenario			Research MQTT
	Detailed plan			Report Template
	Add a card...			Meet with supervisor

5 Toggl Time Logs

5.1 Weekly time Log bar chart



● Bluemix	3:04:57
● Database	2:53:58
● Back end	2:05:35
● General	1:46:09

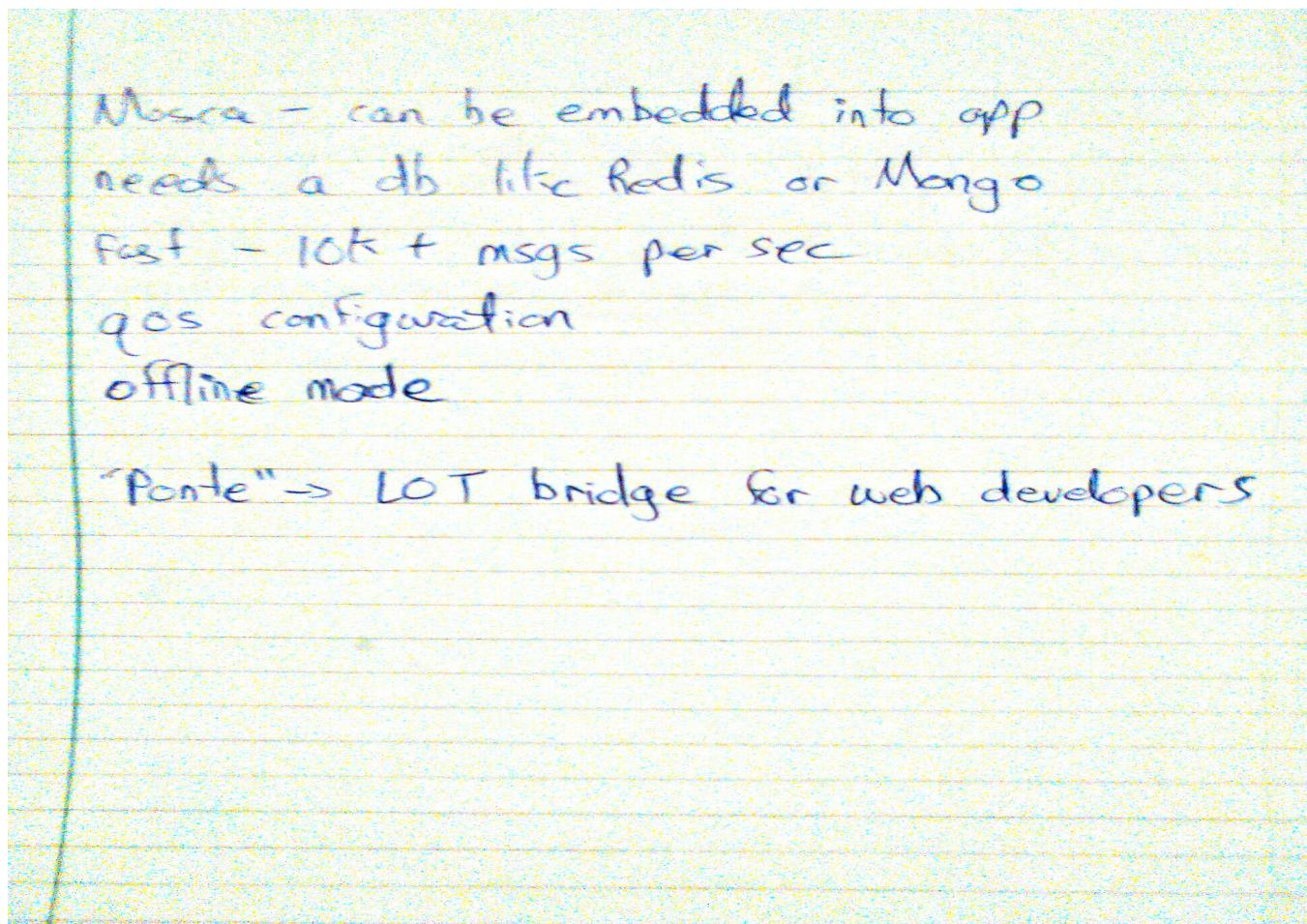


5.2 Weekly time log

Today			0:53:16
Week 1 log	● General	6:45 PM - 7:38 PM	0:53:16
Yesterday			2:06:39
GeoJSON research	● Database	5:47 PM - 6:16 PM	0:28:48
couchDB research	● Database	4:36 PM - 5:15 PM	0:38:59
Graph db research	● Database	3:31 PM - 3:56 PM	0:25:11
project synopsis	● General	2:35 PM - 3:00 PM	0:24:51
serverless research	● Back end	2:15 PM - 2:24 PM	0:08:50
Fri, 27 Jan			3:31:22
Kafka research	● Back end	5:29 PM - 6:34 PM	1:04:40
OpenWhisk research	● Bluemix	3:09 PM - 5:22 PM	1:34:37
MQTT research	● Back end	12:51 PM - 1:43 PM	0:52:05
Wed, 25 Jan			2:51:20
Redis research	● Database		0:36:26
RabbitMQ research	● Database		0:44:34
BLuemix research	● Bluemix		0:37:06
email bluemix support	● Bluemix		0:12:04
setting up academic account bluemix	● Bluemix		0:41:10
Tue, 24 Jan			0:28:02
meeting with supervisor	● General	10:35 PM - 11:03 PM	0:28:02

6 Pictures

6.1 Handwritten notes



Message Hubs - aggregates msgs published to
the "transformed" topic
& places files to Object Storage.

Optimised for time ordered reads & can handle
vast amounts of throughput.

Serverless

Servers are where software runs

They com with one another

Have IP addresses that need to be discovered.

Servers have local state not visible to other servers.

"The essence of the serverless trend is the absence of the server during SW development!"

- Write code that implements business logic.
- Durable state needs to be externalised.
- Local storage mechanisms are for optimization only.

What happened to my Servers?

Instead of supporting a specific application, clusters of servers provide a generic execution environment for any number of applications.

Serverless is to operations what IaaS was to HW

Message Hub - bridge from IoT Platform

Provision an instance of the message hub service.
name it "kafka".

- Create 2 kafka topics
"iotp" and "transformed"
- Configure to publish historical data to the "kafka" service
set default topic to "iotp"

OpenWhisk

Configure to perform transformation, subscribing
to the "iotp" topic &
publishing to the "transformed" topic

Node.js app as a conduit for OpenWhisk requests.

OW JS client library

```
var openwhisk = require('openwhisk');
```

Couch DB → eventual consistency

schema:

pouch DB → clientside javascript library

saving data to cloud → "database replication capability"

Geo JSON format

Mapbox.js → visualize

Polygon → UL

"Geofences"

mosquito broker OR IoT service on BM

6.2 GeoJSON format

```
{  
  "type": "Feature",  
  "geometry": {  
    "type": "Point",  
    "coordinates": [0, 0]  
  },  
  "properties": {  
    "name": "null island"  
  }  
}
```

6.3 RabbitMQ research

Answer by [Simen@bouvet.no \(1\)](#) | Aug 07, 2014 at 01:46 PM

 I see there is a possibility to setup a rabbitmq instance in bluemix. Rabbitmq can implement MQTT if you enable it (<http://www.rabbitmq.com/mqtt.html>), but is that option available in this configuration?

That aside, I was a tad surprised that BlueMix doesn't provide MQTT from before of - MQTT being a brainchild of IBM and the first protocol supported in Node-RED

 0  Hide 1 Share

 knolleary (2549) Aug 07, 2014 at 07:40 PM 0 Like

This is limitation of the current Cloud Foundry architecture that Bluemix is built on. It only supports routing HTTP(s) traffic to the hosted applications. Even if you ran rabbitmq with MQTT enabled within Bluemix, you wouldn't be able to access it with a regular MQTT client.

6.4 Current state of Bluemix account

The screenshot shows the Bluemix account dashboard. On the left is a sidebar with the following menu items:

- Profile
- Notifications
- UNIVERSITY OF LIMERICK —
- Team Directory
- Invite Team Members
- Usage Dashboard
- Billing
- Manage Organizations

The "Billing" item is highlighted with a blue background. The main content area is titled "Billing" and displays the message "Cancellation Request Submitted". Below this, a smaller message states: "Your request for cancellation was received and your account will be canceled shortly."

7 Bookmarks

https://github.com/jthomas/openwhisk_mqtt_feed

<http://www.slideshare.net/AnimeshSingh/how-to-build-a-distributed-serverless-polyglot-microservices-iot-platform-using-docker-and-openwhisk>

<https://medium.com/openwhisk/transit-flexible-pipeline-for-iot-data-with-bluemix-and-openwhisk-4824cf20f1e0#.ylmms651u>

<https://github.com/glikson/transit>

<https://www.ibm.com/blogs/bluemix/2017/01/managing-iot-devices-with-kafka-and-mqtt/>

<https://hub.jazz.net/project/wprichar/data-analytics-transportation/overview>

<https://developer.ibm.com/recipes/tutorials/integrating-watson-iot-platform-with-message-hubkafka/>

<http://jamesthom.as/blog/2016/10/31/serverless-logs-with-elasticsearch/>

<https://github.com/mcollina/mosca/wiki/Docker-support>

<https://github.com/Kitard/node-red-geospatial-node>

<https://www.ibm.com/blogs/bluemix/2014/12/find-bluemix-geospatial-analytics/>

<http://www.macwright.org/2015/03/23/geojson-second-bite.html>

<https://github.com/tmcw/awesome-geojson>

<https://tools.ietf.org/html/rfc7946#section-10>

<https://developer.ibm.com/clouddataservices/2016/11/15/7-databases-7-days-ibm-graph/>

<https://www.ibm.com/developerworks/cloud/library/cl-openwhisk-node-bluemix-user-facing-app/index.html>

<http://heidloff.net/article/how-to-trigger-openwhisk-cloudant>

