Review SME Continous Call (F4Fp – SME) FRIDA – Stage 2

Celestino Alvarez

Teiichi Capital Tech.

Virtual FEC 9

The World, June 2nd 2021


Sep-11 Oct-11 Nov-11 Dec-11 Jan-12 Feb-12
2013/10 2013/11 2013/12 2014/01 2014/02 2014/03

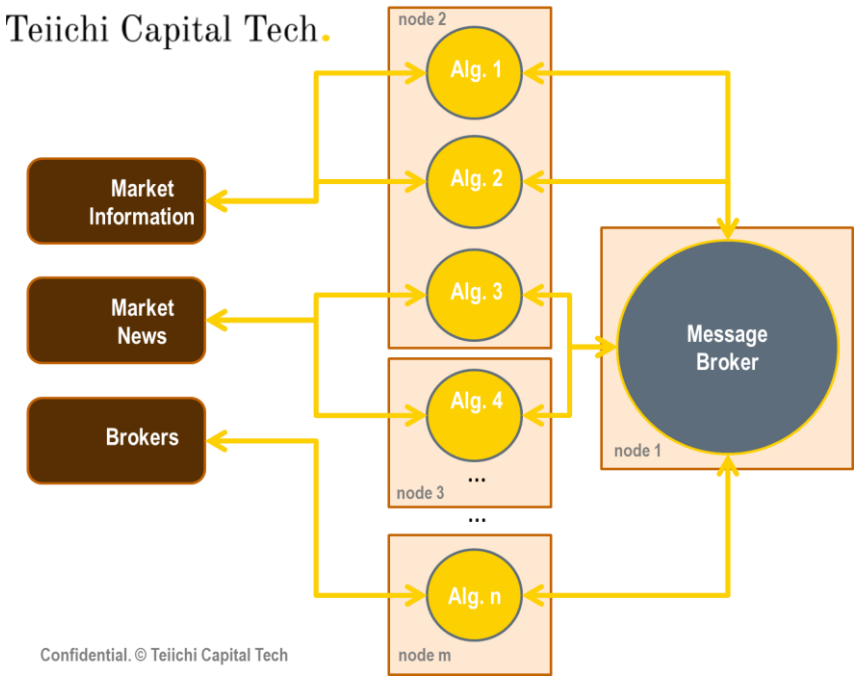
Outline

- **Experiment description (max. 4 slides)**
  - Concept and objectives
  - Background and motivation
  - Experiment set-up
- **Project results (max. 3 slides)**
  - Measurements
  - Lessons learned
- **Business impact (min. 4 slides)**
  - Impact on your business, .. how did Fed4FIRE helped you ?
  - Value perceived, .. why did you come to Fed4FIRE ?
- **Feedback (min. 4 slides)**
  - Used resources and tools
  - Added value of Fed4FIRE

Experiment Description

FINANCIAL RESEARCH INVOLVING DATA ANALYSIS – FRIDA – STAGE 2

Concept and Objectives



Objectives

Stage 1

O1. Measure and compare the performance of the message broker and network while scaling up the volume of messages processed

Stage 2

O2. Improve and automate the machine learning workflow for training, evaluating and deploying new algorithms

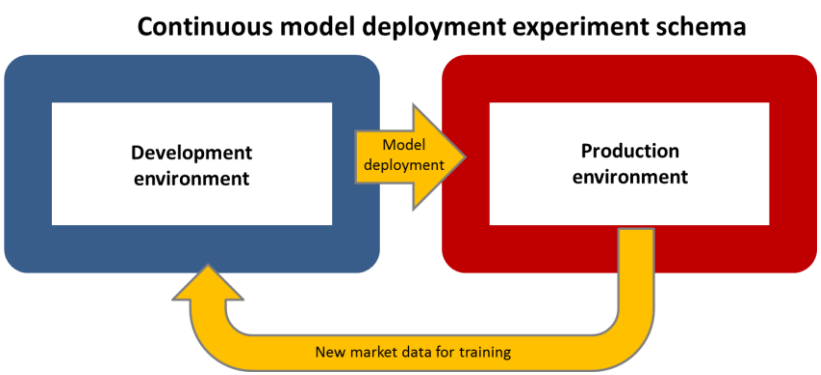
O3. Dimension the infrastructure required for scaling up

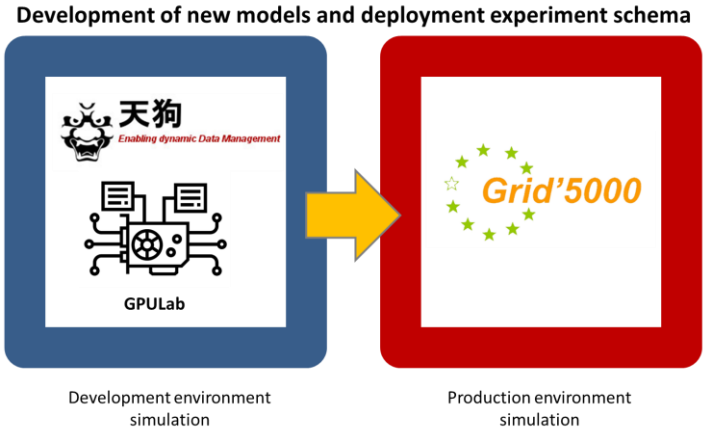
Background and Motivation

Teiichi Capital Tech is a fintech that has developed a distributed cloud platform for processing financial data, taking decisions and executing operations in real time, fully automated, using cutting edge AI algorithms. The platform is currently oriented to ingest data from stock markets, seeking deals to obtain +20% monthly ROI.

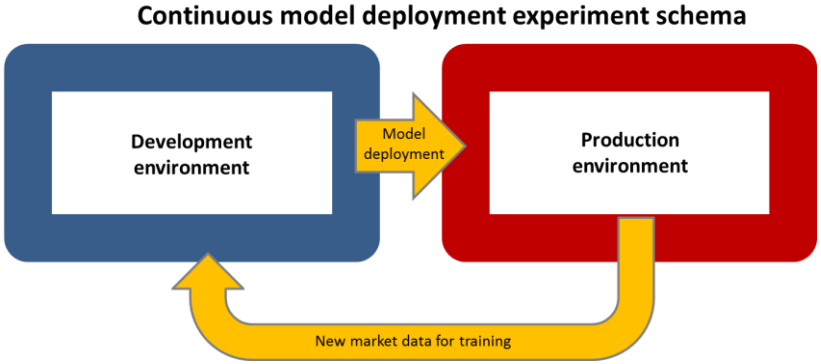
The company intends to **scale the platform** to process up to +16.000 financial instruments, so it is needed to **optimize and validate the evolution of the platform**, which has a critical impact in our business.

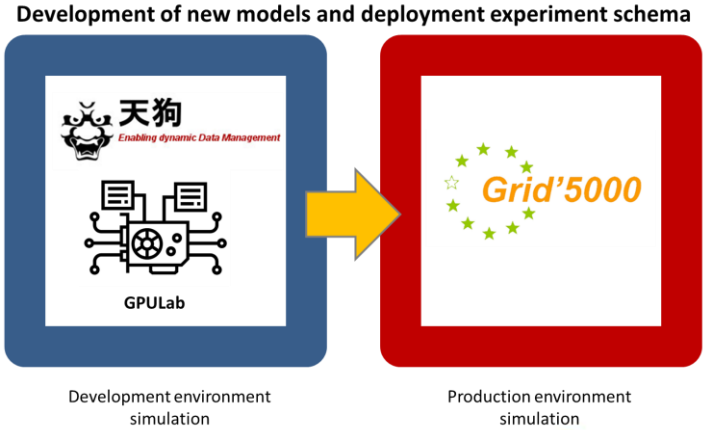
Experiment Setup





Experiment Setup



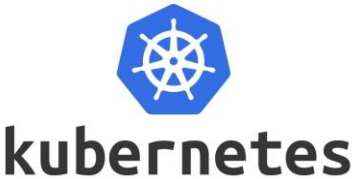


"Current research pointed out the lambda architecture as a system doesn't always deliver on its promises"

Give a shot to kubernetes

Experiment Setup

Experiment Setup



- Officially, it is a container orchestrator
- In real life is a game changer on how applications are designed, deployed and operated
- Need to rethought existing aplications to convert to a microservice architecture
- The promise is: *«Just throw me more resources (CPU, RAM and Storage), and I will handle the scalability and reliability of the application»*
- This is ok with stateless services, but how to deal with persistent storage to be distibuted and reliable?

Experiment Setup

- S3 Object Storage compatible with Amazon S3
- Minimum recommended setting is 4 nodes with 4 disks each
- It offers bucket replication and reconstruction on the fly
- No block device supported by default
- It is posible to use s3fs or s3backer toprovide HA filesystem

  (https://honeyguide.eu/posts/minio-fuse/)
- Integration with Docker Private Repository to store Docker images

Experiment Setup

- Cloud-native distributed persistent block storage for Kubernetes
- Define persistent volumes to be used by K8 pods, and it will take care of replication
- Backups to a S3 storage
- Relying on Longhorn to store tablespaces for MariaDB and Influx DB

Experiment Setup

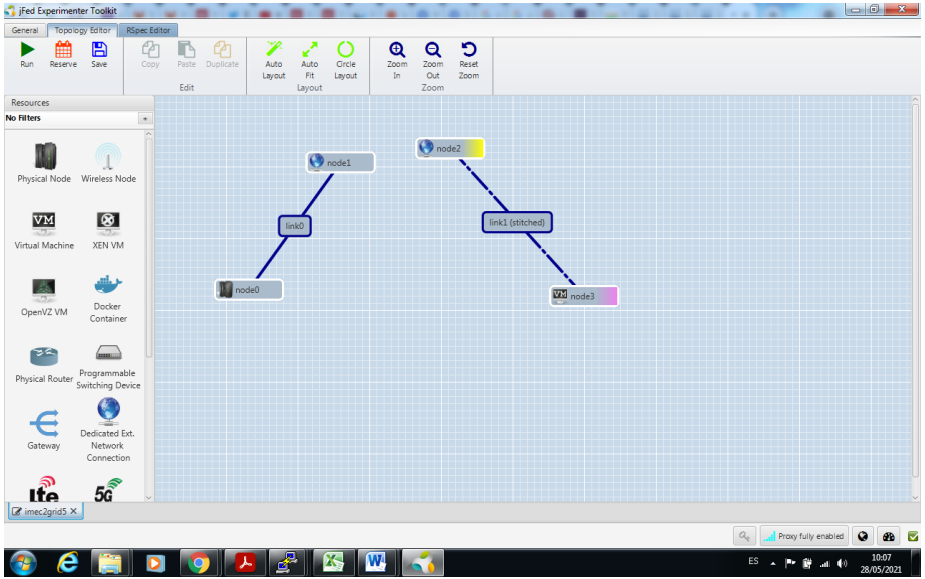
- Deployed in a HA configuration
- Using its own mechanism for replication
- Storing Tablespaces in Lonhgorn with its own replication mechanims
- Two replication mechanism are going on here at the same time… Need further research…

Experiment Setup

- It is Git-repository that covers the full DevOps cycle
- It includes CI/CD life cycle
- It handles different K8 clusters (development, staging, production…)
- It allows the automation for training new models, and create the container image with a new micro-service with the new model

Experiment Setup

- Haven’t been able to connect Grid5000 and imec!
- Even if we had achieved it, it wouldn’t be posible to connect Tengu’s pre-reserve nodes using jFed
- Decided to move forward and create the cluster in Grid5000



Experiment Setup

- We played arround with GPULab and concluded that a similar frontend using Jupyter is useful for development, but it can just be added to the Kubernetes cluster

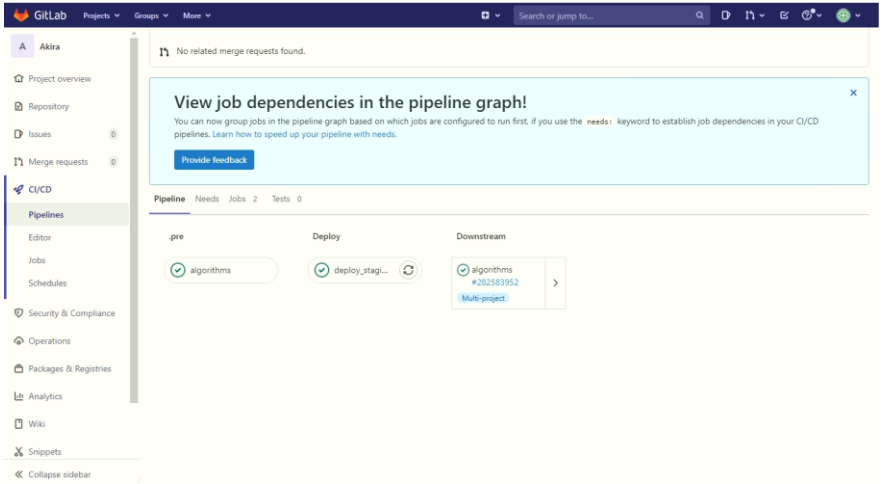
Project Results

FINANCIAL RESEARCH INVOLVING DATA ANALYSIS – FRIDA – STAGE 2

Measurements

- Written scripts for creating a Kubernetes cluster in Grid5000 with any number of nodes
- Automated training and deployment of new models

```
root@uvb-9:/home/malvarez/stage2#  kubectl get nodes
NAME                      STATUS   ROLES                  AGE     VERSION
uvb-12.sophia.grid5000.fr   Ready    <none>                 84s     v1.21.1
uvb-26.sophia.grid5000.fr   Ready    <none>                 7m35s   v1.21.1
uvb-28.sophia.grid5000.fr   Ready    <none>                 7m24s   v1.21.1
uvb-37.sophia.grid5000.fr   Ready    <none>                 101s    v1.21.1
uvb-41.sophia.grid5000.fr   Ready    <none>                 7m30s   v1.21.1
uvb-42.sophia.grid5000.fr   Ready    <none>                 7m24s   v1.21.1
uvb-44.sophia.grid5000.fr   Ready    <none>                 7m17s   v1.21.1
uvb-5.sophia.grid5000.fr    Ready    <none>                 7m17s   v1.21.1
uvb-6.sophia.grid5000.fr    Ready    <none>                 7m18s   v1.21.1
uvb-8.sophia.grid5000.fr    Ready    <none>                 7m15s   v1.21.1
uvb-9.sophia.grid5000.fr    Ready    control-plane,master   9m23s   v1.21.1
root@uvb-9:/home/malvarez/stage2#
```



Lessons Learned

- Learn about Kubernetes (quite a lot)
- Learn about distributed pesistance storage (Minio, Longhorn, Rook, Ceph…)
- Managing the development cycle with Gitlab CI/CD can sinplify DevOps (or not!)

- K8 is a very powerful and complex tool since it abstracts every single component in a cluster
- Do not subestimate the effort needed to manage kubernetes
- Every single app deployed has its own complexity
- Fully testing distributed storage requires specific harware config (i.e. min. 4 disks in min. 4 nodes)
- Migrating from traditional applications to K8 requires to rethink the whole architecture
- Jupyter Notebooks could be included in our development loop

Business Impact

FINANCIAL RESEARCH INVOLVING DATA ANALYSIS – FRIDA – STAGE 2

Impact on your business

- Scaling the business depends on scaling the technology
    - Need to go from 4-5 assets to +16.000 assets
- Early technology decission have a strong impact on the future business, but **how to take this decissions being an SME with limited resources**?
    - Take for granted the product documentation and cross your fingers
    - Use Fed4FIRE infrastructure to validate your use case

Impact on your business

- After the experiment we have obtained hands-on experience on how to developed automated ML proceses using Gitlab and Kubernetes, and how to use Kubernetes to scale
- We were mostly wrong in our assumption about lambda architectures and are pretty confidnet that Kubernetes is the right technology to move to.
- We expect to be able to migrate our platform to Kubernetes in the next 6 months.

Value Perceived

- Expertise and advice from Tengu
- Hands on experience from messing up with GPULab and a bunch of ideas to incorporate in our environment
- Availability of a large amount of resources
    - G5000 is a great infrastructure that can be easily used to test different software possibilities
- Funding helps to lower the cost of learning curve of the platform
    - But furthermore, it allows to dedicate time to experiment, which is important but not urgent in an SME, where costs requiere to be assigned to billable projects

Value Perceived

- Being backed by Fed4FIRE, it is possible to target more ambitious projects
    - We have spent quite a lot of time experimenting with technologies that otherwise would result very difficult (both because the resources needed but also for the time needed).

Feedback

Used resources and tools

- Testbed: **Tengu**
    - Finally no lambda architecture but an initial Kubernetes cluster to mess up accesible through a VPN to imec. Valuable expertise from the support team. (Thanks Gregory!)
- Testbed: **GPULab**
    - Used to run some scripts and train networks. It is not alwasy easy to reserve the nodes. Using Jupyter Notebooks as part of the development environment is a great idea for us. This could be “easily” replicated in a Kubernetes cluster.

Used resources and tools

- Tool: **jFed**
  - Used to reserve a node in G5000 and a node in Virtual wall connected through a VLAN. We got the reservation of the resources but we did not achieve visibility between nodes (neither ICMP or TCP).
    - Tried different network addresses without success
      - Grid5000 side looked ok (thanks Luke!)
      - Imec side cannot tell (issue posted on Google groups)

Our feeling is that jFed is becoming outdated

- Desktop application
- Some features do not work (i.e. reporting bugs)
- Not really a one stop shop
- Would suggest to migrate to a web app

Used resources and tools

- Testbed: **Grid5000**
  - We have used Gemini cluster (NVIDIA DGX-1 nodes), Hercule and Nova in Lyon site for training models and uvb on Sophia site to create the Kubernetes cluster.
  - Although Tutorials on using Terraform to create a Kubernetes cluster in Grid5000 was available, we did prefer to use our own scripts, so we be able to replicate in our environment in the future.

Added value of Fed4FIRE

- Having such amount of resources ready to be used is the most appreciated value. Also the good documentation to get hands on quickly is a plus.
- After completing Stage 2, there are a bunch of open question that we couldn’t face this time like:
    - Stress test of Longhorn
    - Best practices regarding the replication of DB (ie. MariaDB HA vs Longhorn storage)

We expect to have answers on this questions with new experiments on Fed4FIRE testbeds.


Co-funded by the European Union


Co-funded by the Swiss Confederation

This project has received funding from the European Union's Horizon 2020 research and innovation programme, which is co-funded by the European Commission and the Swiss State Secretariat for Education, Research and Innovation, under grant agreement No 732638.

WWW.FED4FIRE.EU