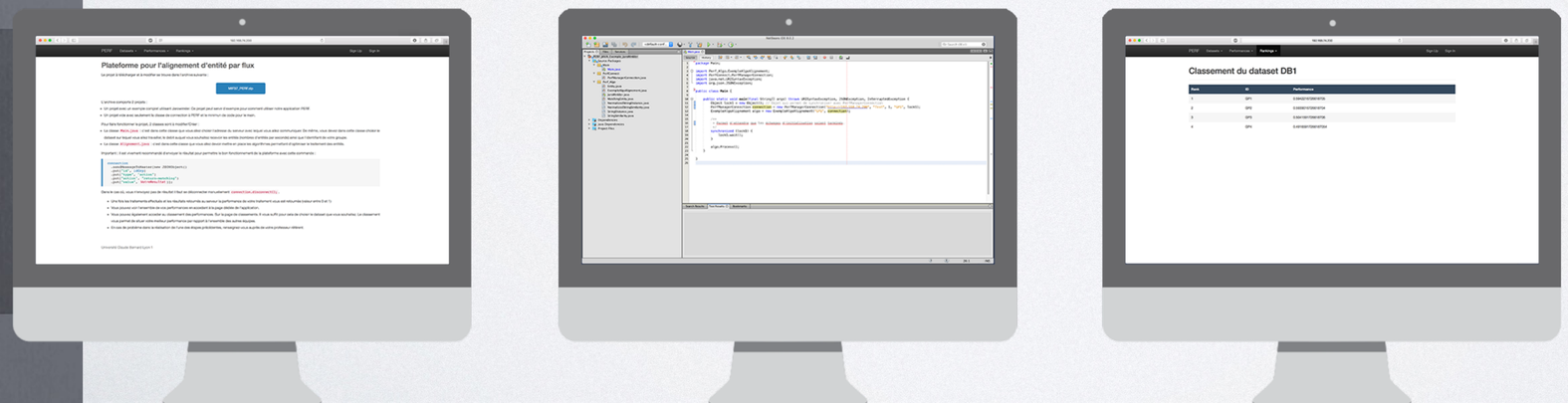
TI5 - PROJET PERF



Université Claude Bernard Lyon 1

PLAN

Présentation

Équipe

Organisation et travail en équipe

Conception

Qualité technique

Démonstration

Difficultés rencontrées

Bilan

PRÉSENTATION

CONTEXTE

Munir les enseignants de base de données d'un outil permettant de **distribuer des flux d'entités sémantique** à destination d'une promotion d'étudiants et d'**évaluer la qualité du travail effectué** par ces derniers.

PRÉSENTATION

OBJECTIFS

- Concevoir et réaliser une **plateforme** :
  - **facilement déployable** et configurable à petite ou grande échelle
  - **administrable** par les enseignants (édition de datasets)
  - facilitant la communication **client-serveur** pour les étudiants
  - **évaluant les résultats** des étudiants et **proposant un classement**

PRÉSENTATION
FONCTIONNALITÉS

- Administration des datasets
- Répartition des clients sur les workers
- Distribution de flux d'entités
- Calcul des performances et établissement d'un classement
- Mise à disposition de l'interface Java et de l'énoncé du TP

ÉQUIPE

	Murat Halat	Antoine Hintzy	Corentin Lonjarret	Felician Matinya	Alexandre Millot
Chef de projet / Correspondant encadrants					X
Scrum Master			X		
Concepteur base de données			X		X
Expert technique Web		X			
Designer		X			
Développeur	X	X	X	X	X
Concepteur	X			X	

Porteurs du projet : **Fabien Duchateau** et **Nicolas Lumineau**, enseignants chercheurs au sein du LIRIS, pôle Data Science, équipe Base de Données

ORGANISATION ET TRAVAIL EN ÉQUIPE

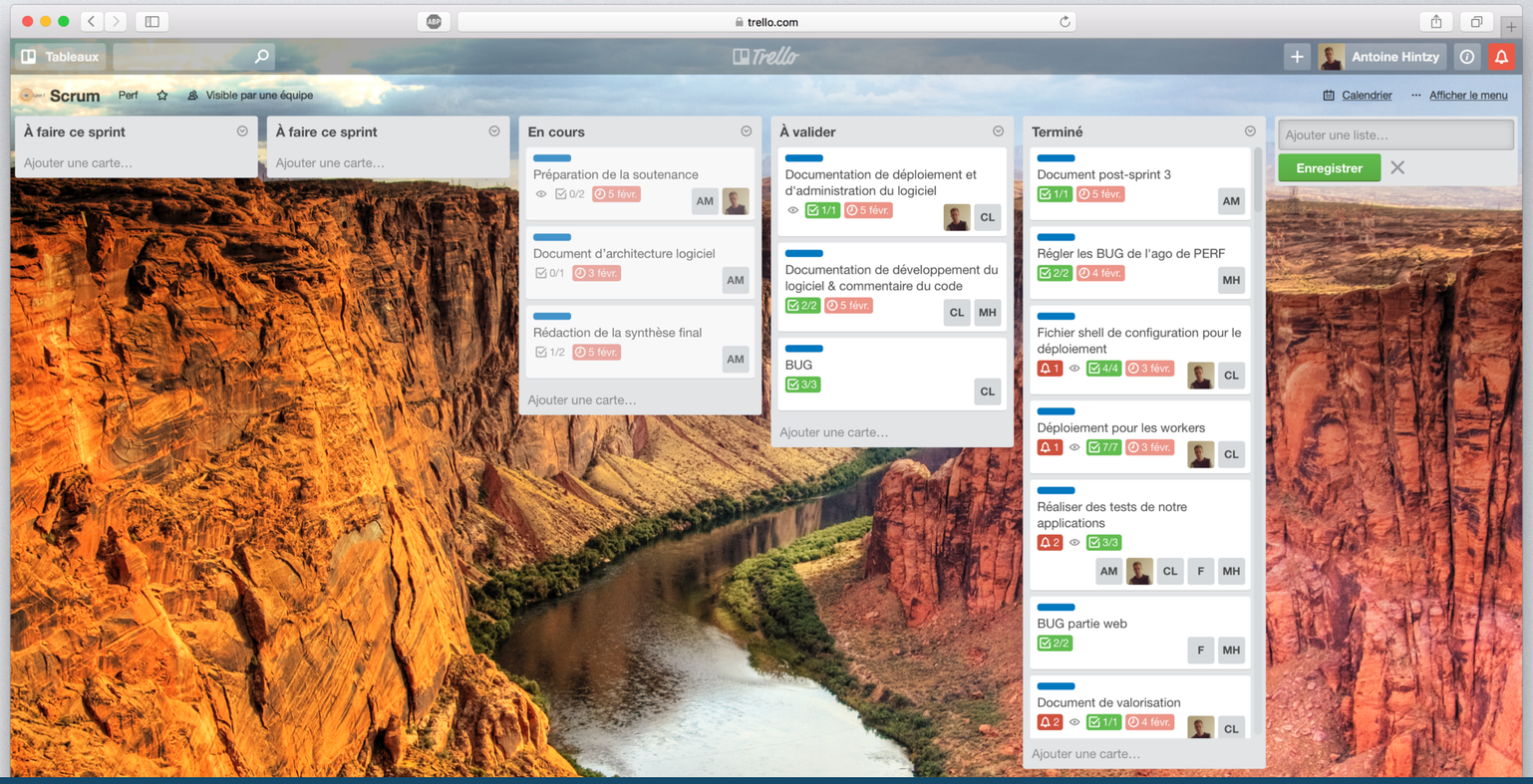
GESTION PRÉ-SPRINT ET POST-SPRINT

- Tâches à effectuer avant le sprint (configuration, installation, …)
    - éviter la perte de temps au début du sprint
- La plateforme est dans un état stable à la fin de chaque sprint

ORGANISATION ET TRAVAIL EN ÉQUIPE

MÉTHODE AGILE (SCRUM)

- À faire
  - En cours
    - À valider
      - Terminé



Trello
Organize anything, together.

ORGANISATION ET TRAVAIL EN ÉQUIPE

MÉTHODES DE GESTION DE PROJET

Bitbucket

- Sauvegarde et versionnage du code source avec **Git**
- Méthode **Git Flow** :

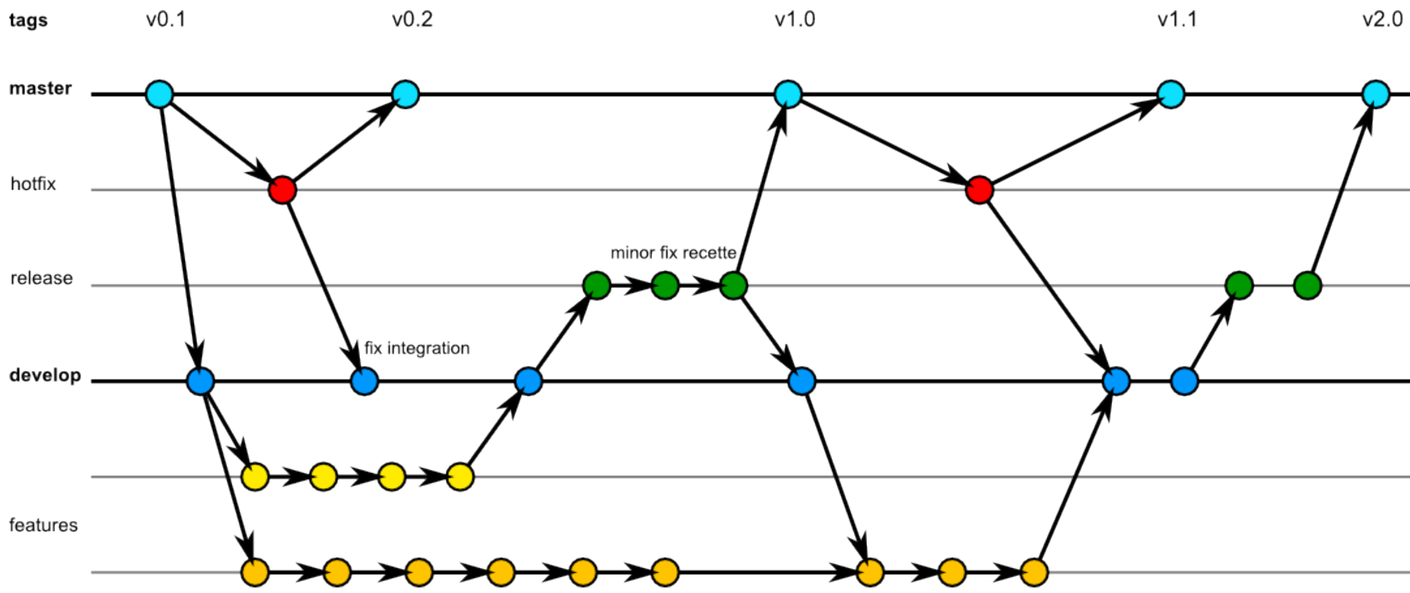


Schéma : http://blog.techniv.fr/2015/workflow-git-au-sein-dune-entreprise/

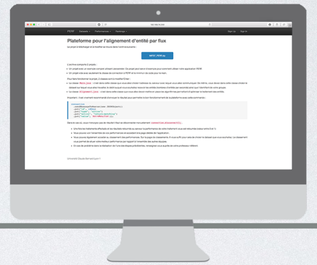
CONCEPTION
TECHNOLOGIES

Déploiement

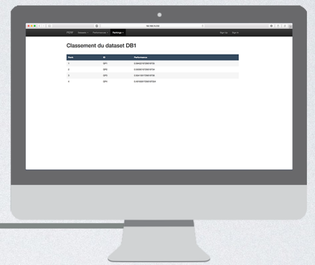
Serveur

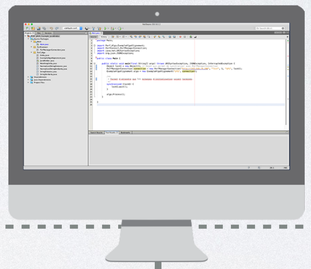
CONCEPTION
ARCHITECTURE

Déploiement

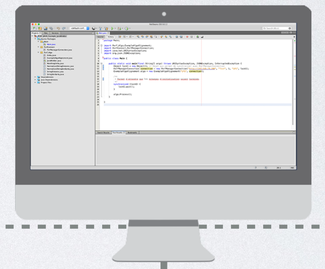


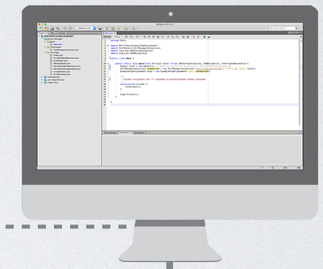
Énoncé
Administration





Réception
Traitement





VM
Docker
Base de Données
Master

VM
Docker
Worker

VM
Docker
Worker

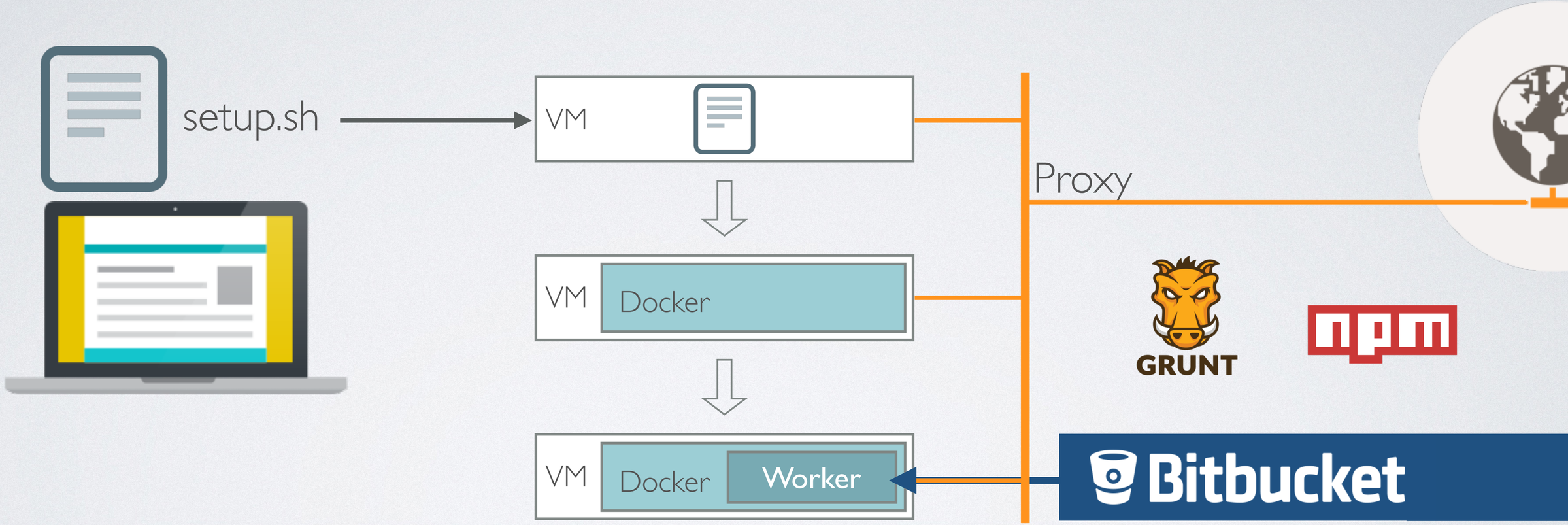
VM
Docker
Worker

Résultats

Flux d'entités

1 master
n workers

QUALITÉ TECHNIQUE
DÉPLOIEMENT

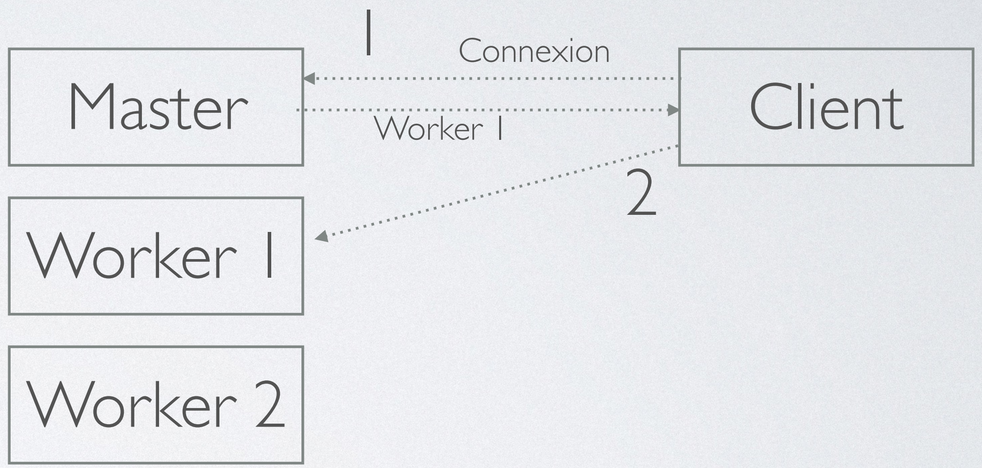


QUALITÉ TECHNIQUE
APPLICATION WEB

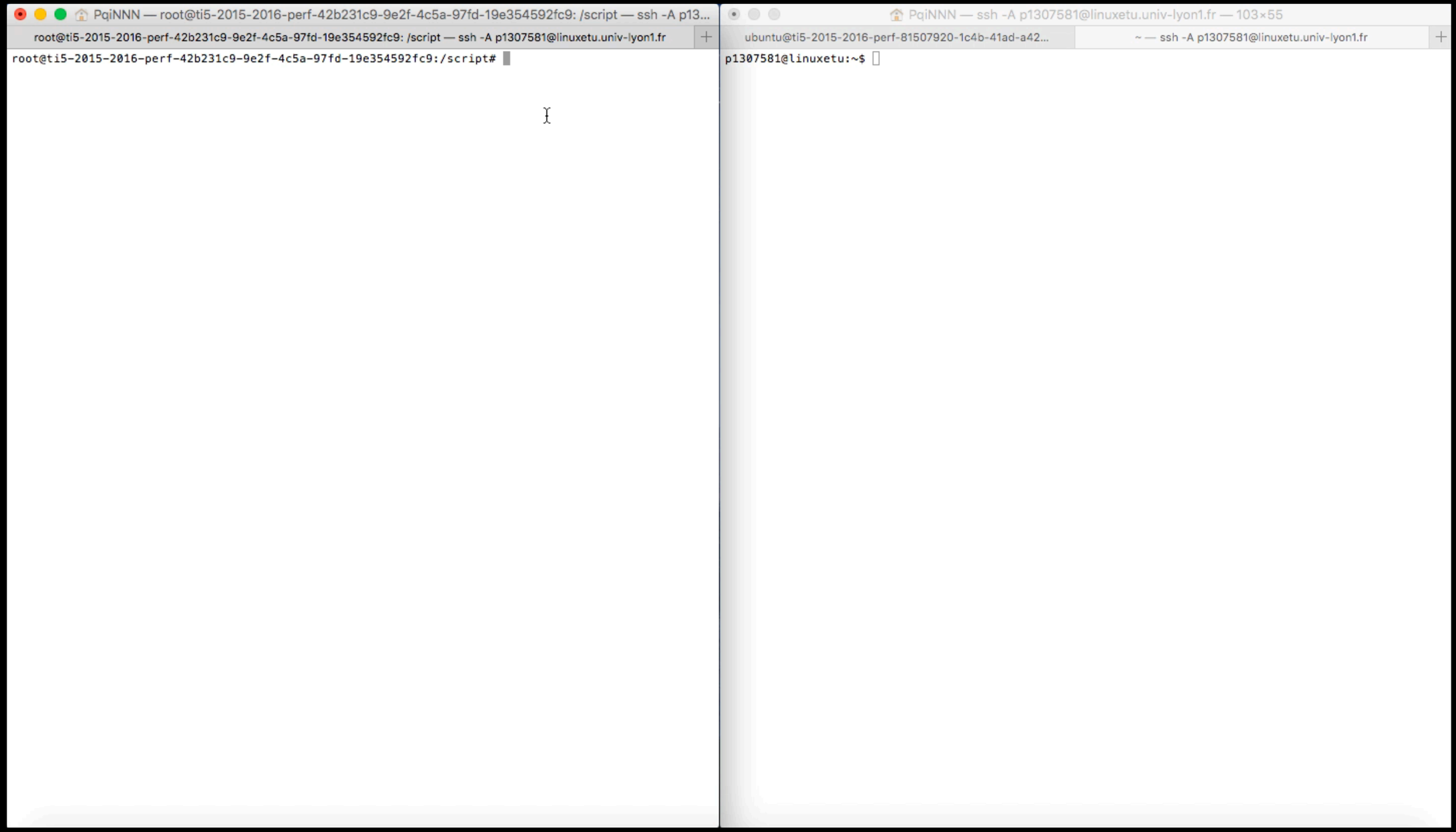
- 2 types d'utilisateur
    - admin, authentification requise (professeur)
    - client (étudiant)
- Permet d'administrer la plateforme
    - création de dataset
    - ajout de worker
- Pages d'aide à la prise en main du système
- Visualisation des classements

QUALITÉ TECHNIQUE
ÉCHANGES CLIENT-SERVEUR

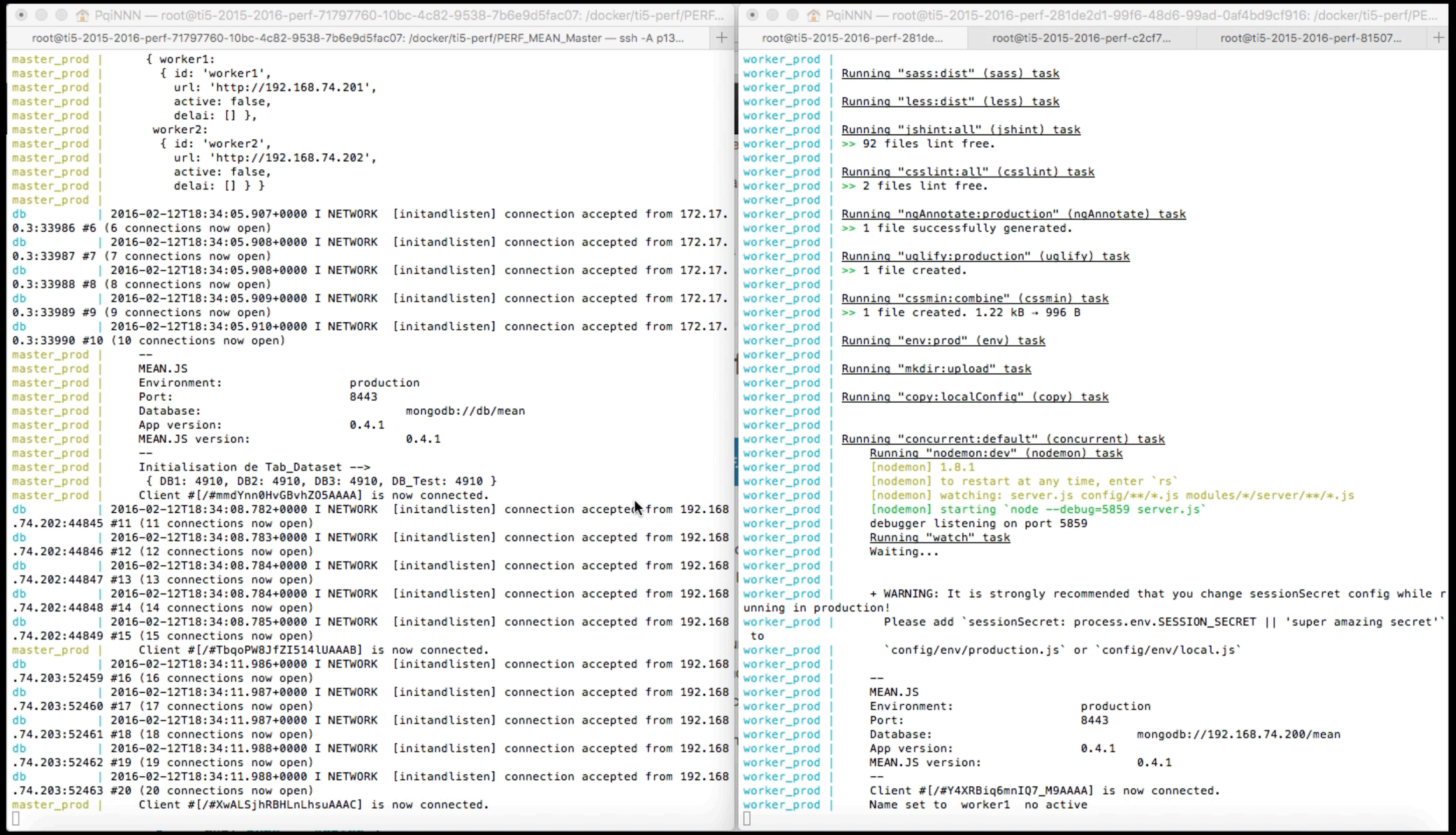
- 2 sockets de connexion client-serveur
  (client-master et client-worker)
- Echange de messages au format JSON (entités)
- Calcul de performances



DÉMONSTRATION



Déploiement de l'application



Communication client-serveur

PERF

192.168.74.200

Rechercher

PERF Datasets Performances Rankings S'enregistrer S'authentifier

L'archive comporte 2 projets :

- Un projet avec un exemple complet utilisant Jarowinkler. Ce projet peut servir d'exemple pour la façon d'utiliser notre application PERF.
- Un projet vide avec seulement la classe de connexion à PERF et le minimun de code pour le main.

Pour faire fonctionner le projet, 2 classes sont à modifier/créer :

- La classe `Main.java` : c'est dans cette classe que vous allez choisir l'adresse du serveur avec lequel vous allez communiquer. De même, vous devez dans cette classe choisir le dataset sur lequel vous allez travailler, le débit auquel vous souhaitez recevoir les entités (nombres d'entités par seconde) ainsi que l'identifiant de votre groupe.
- La classe `Alignement.java` : c'est dans cette classe que vous allez devoir mettre en place les algorithmes permettant d'optimiser le traitement des entités.

Important : il est vivement recommandé d'envoyer le résultat pour permettre le bon fonctionnement de la plateforme avec cette commande :

```
connexion
  .sendMessageToMaster(new JSONObject()
  .put("id", idGrp)
  .put("type", "action")
  .put("action", "return-matching")
  .put("value", VotreResultat ));
```

Dans le cas où vous n'envoyez pas de résultat, il faut se déconnecter manuellement `connection.disconnect();`.

- Une fois les traitements effectués et les résultats retournés au serveur, la performance de votre traitement vous est retournée (valeur entre 0 et 1)
- Vous pouvez voir l'ensemble de vos performances en accédant à la page dédiée de l'application.
- Vous pouvez également accéder au classement des performance sur la page de classements. Il vous suffit pour cela de choisir le dataset que vous souhaitez. Le classement vous permet de situer votre meilleur performance par rapport à l'ensemble des autres équipes.
- En cas de problème dans la réalisation de l'une des étapes précédentes, renseignez-vous auprès de votre professeur référent.

Université Claude Bernard Lyon 1

Interface client - administration

DIFFICULTÉS RENCONTRÉES

- Déploiement
  - proxy
  - connexion à MongoDB
  - mise en production (modification des variables d'environnement)
- Interropérabilité Node.js/Java
  - sockets de Node.js trop sécurisées (nécessité de socket ID)
- Calcul de performances sur le master (blocage des autres fonctionnalités)

BILAN

Projet fonctionnel

Déployable à différentes échelles, mais un worker par client conseillé pour de meilleurs résultats

Administrable par les enseignants via une interface graphique

Interface Java toute faite pour la communication avec le serveur

Propose un classement des résultats obtenus

POURSUITE DE PROJET

Gestion de nouveaux types de sources de données

Mise en place de nouveaux modes : projet, test, …

Gestion des utilisateurs (étudiants) via LDAP

Algorithme de recherche du meilleur worker

Faire un Sharding et de la réplication de la base MongoDB

IMPRESSIONS PERSONNELLES

Projet complet, proche d'un travail en entreprise

Acquisition de compétences en déploiement d'application, en interropérabilité et dans l'utilisation de MEAN

REMERCIEMENTS

Nous tenons à remercier :

Fabien Rico et **Emmanuel Coquery**
pour leur aide précieuse

Nicolas Lumineau, **Fabien Duchateau** et **Romuald Thion**
pour l'encadrement de ce projet