

PLMonyxia

Philippe Depouilly
Institut de Mathématiques de Bordeaux

20 mai 2025 - Limoges

- ▶ Onyxia est une plateforme de traitement de données open source développée par l'Insee pour répondre aux besoins grandissants des experts de la donnée dans le domaine de la statistique publique.¹
- ▶ Onyxia s'appuie sur les technologies Cloud Native (Kubernetes) afin de proposer un large catalogue d'applications autour du traitement de la donnée
 - ▶ Kubernetes pour l'orchestration des applications
 - ▶ Helm Charts pour le packaging des applications
 - ▶ Keycloak pour le SSO
 - ▶ Minio pour le Stockage S3 as a Service
 - ▶ Vault pour la gestion des secrets et préférences des utilisateurs

Onyxia - Datalab

Réduire

Accueil

Mon compte

Paramètres du projet

Catalogue de services

Mes services

Mes secrets

Mes fichiers


Explorateur de Données

Catalogue de services

Services for datascientists.


rechercher

Services interactifs Bases de données Automatisation Applications pour les Mathématiques (PLM)

 **Jupyter-python**


The JupyterLab IDE with Python and a collection of standard data science packages.

Lancer


 **Rstudio**

The RStudio IDE with a collection of standard data science packages.

Lancer


 **Vscode-python**

The Visual Studio Code IDE with P of standard data science packages

 **Vscode-tensorflow-gpu**


The VSCode IDE with Python and the deep-learning framework TensorFlow, with GPU support.

Lancer


 **Vscode-pyspark**

The Visual Studio Code IDE with PySpark, an interface to use Apache Spark from Python.

Lancer


 **Jupyter-pytorch**

The JupyterLab IDE with Python a framework PyTorch.

 **Rstudio-spark**


The RStudio IDE with a collection of standard data science packages. It includes SparkR, an R package that provides an interface to use Apache Spark from R.

Lancer

 **Vscode-r-python-julia**

The VSCode IDE with minimal installations of R, Python and Julia.

Lancer

 **Vscode-tensorflow**

The VSCode IDE with Python and TensorFlow.

Pourquoi ?

- ▶ la PLM propose déjà trois services JupyterHub ;
- ▶ se réduit à des Notebooks avec un paramétrage restreint ;
- ▶ dépendant du proxy JupyterHub ;
- ▶ investigation du groupe ShinyR autour de pouvoir déployer des Applications Shiny plus simplement que via le catalogue PLMshift.

Pourquoi ?

- ▶ proposer un catalogue d'applications orientées science
- ▶ largement paramétrables et simplement déployables
- ▶ et surtout autonomes (ne dépend pas d'un proxy authentifié)
- ▶ KubeApps amélioré avec une valeur ajoutée sur le Stockage en tant que service

Les utilisateurs sont guidés dans un catalogue d'applications avec la possibilité de lier leur instance à un dépôt GIT, à un stockage distant, à des versions différentes de conteneurs, etc.

Interface intuitive

Onyxia - Datalab

3. Help install shiny

Shiny Rinitialiser les configurations Copier l'URL de lancement automatique Enregistrer la configuration

Le chart **shiny** appartient au dépôt de charts Helm Applications pour les Mathématiques (PLM).

View personalized Version 10.0

Annuler Lancer

Formulaire Editeur de texte

Service

Resources Your service will have at least the requested resources and never more than its limits. No limit for a resource and you can consume everything left on the host machine.

CPU The amount of cpu guaranteed / The maximum amount of cpu.
 100 m Guaranteed ——— 1000 m Maximum

Memory The amount of memory guaranteed / The maximum amount of memory.
 1 Gi Guaranteed ——— 2 Gi Maximum

Configuration for persistence

Init init parameters

Environment variables environment variables available within your service

Role

Vault Configuration of vault client

S3 Configuration Configuration of temporary identity for AWS S3

GIT Git user configuration

Network access

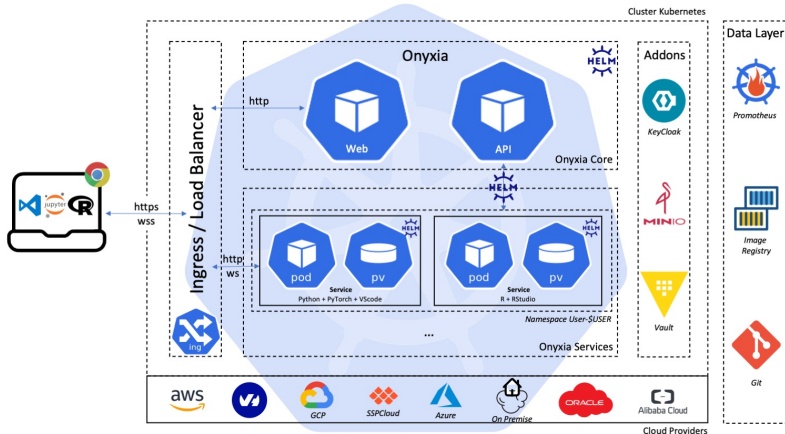
Security

Ingress Ingress parameters

Hostname mon-shiny-lab-math.cnrs.fr



Détails du déploiement



Authentification

- ▶ l'application s'authentifie via un service SSO OIDC ;
- ▶ les services Vault et S3 sont aussi Resources Providers du SSO ;
- ▶ peut s'appuyer sur un OIDC générique (Doorkeeper de la PLM), mais Doorkeeper n'est pas complètement fonctionnel avec l'authentification PKCE (via le navigateur) ;
- ▶ création d'un proxy SSO avec Keycloak.

Les services tiers

- ▶ Vault est déployé selon les règles de l'industrie (déverrouillage du coffre-fort avec *Vault Transit Engine* d'un Vault secondaire et hors site)
- ▶ Configuration ArgoCD Vault
- ▶ Minio déployé selon la documentation de Onyxia

Et maintenant ?

- ▶ Le proposer à l'évaluation auprès de la communauté mathématique
- ▶ Initier une dynamique pour
 - ▶ Accompagner les éventuels utilisateurs
 - ▶ Proposer des applications spécifiques pour les Mathématiques
- ▶ Evaluer l'intérêt et l'utilisation