

Une solution PaaS sur plusieurs IaaS pour une communauté scientifique ESR

Sylvain Allemand¹, Philippe Depouilly¹, Damien Ferney²,
Henri Massias³, Romain Théron⁴

¹Institut de Mathématiques de Bordeaux
Université de Bordeaux

²Laboratoire de mathématiques Blaise Pascal
Université Clermont Auvergne

³XLIM
Université de Limoges

⁴Institut Denis Poisson
Université d'Orléans

12 décembre 2024



Le contexte

- ▶ le Réseau Thématique Mathrice opère à travers la PLM¹ un ensemble de services destinés à la communauté mathématique française et ses interactions (> 7 000 personnes) ;
- ▶ l'équipe en charge de l'administration de ces services, la PLMTeam (une quinzaine de personnes), est composée d'informaticiens de laboratoires qui consacrent un temps sur ces services en complément de leur activité de site ;



¹PLM : Plateforme en Ligne pour les Mathématiques



Le contexte

- ▶ le Réseau Thématique Mathrice opère à travers la PLM¹ un ensemble de services destinés à la communauté mathématique française et ses interactions (> 7 000 personnes) ;
- ▶ l'équipe en charge de l'administration de ces services, la PLMTeam (une quinzaine de personnes), est composée d'informaticiens de laboratoires qui consacrent un temps sur ces services en complément de leur activité de site ;
- ▶ ces services sont essentiellement collaboratifs et s'élargissent en 2020, avec l'arrivée de Kubernetes, à un périmètre plus scientifique et expérimental ;
- ▶ une adhésion rapide à la technologie permet de déployer simplement des applications web, basées sur l'authentification OpenIDConnect de la PLM à travers la Fédération RENATER (dès 2011).



¹PLM : Plateforme en Ligne pour les Mathématiques

Nos plateformes accueillent des applications de différents types :

- ▶ self service à la demande
 - ▶ applications scientifiques : JupyterHub, NoteBooks, ShinyR, Python, onyxia, etc.
 - ▶ sites Wordpress
- ▶ hébergements Web applicatifs
 - ▶ Django, NodeJS, PHP, etc.
 - ▶ sites web institutionnels
- ▶ services nationaux
 - ▶ dans le périmètre CNRS Mathématiques (INSMI) : plmlab/gitlab, plmbox/seafile, SSO pour le labo, etc.
 - ▶ ouvert à la Fédération d'identité : codimd, plmkan, jupytercloud, plmlatex (21 400 utilisateurs)

Les instances Kubernetes de la PLM

Actuellement, nous avons quatre plateformes :

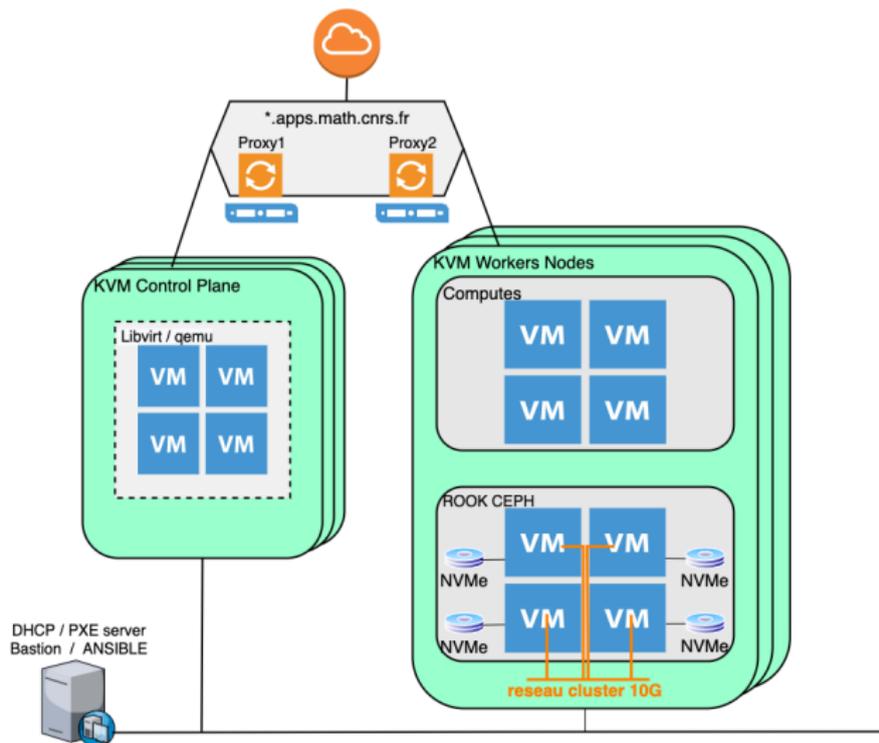
Nom	Infra	Type	CPUs	Caractéristiques	Depuis
PLMshift	Bare Metal	Manuel	380 CPUs	vGPUs et Big Mem pour l'expérimentation	2019
PLMapps	Openstack	IaC	120 vCPUs	Applis nationales et institutionnelles	2022
PLMopp	Openstack	IaC	110 vCPUs	Projet S3 Labo PRA PLMapps	2023
PLMvshift	Openstack	IaC	120 vCPUs	Evolution PLMshift	2024

Toutes les plateformes sont configurées via ArgoCD, l'ensemble des configurations est publique

<https://plmlab.math.cnrs.fr/plmteam/okd-clusters>



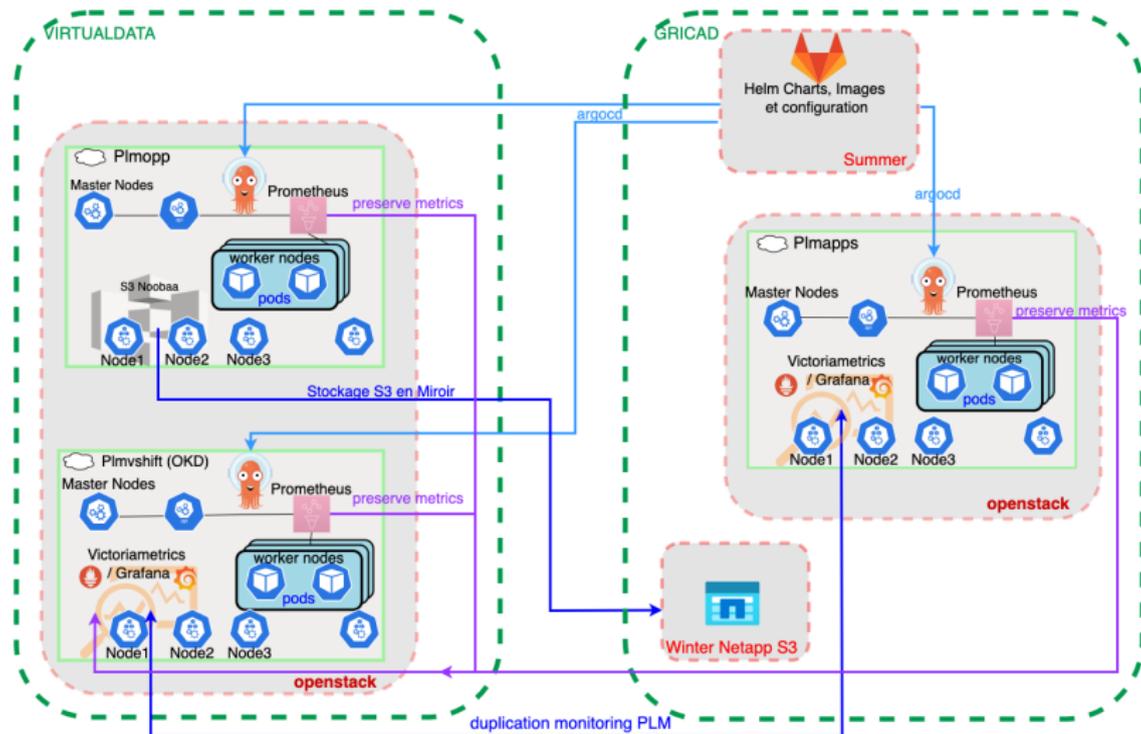
OKD Bare Metal



OPENSIFT mode Hyperconvergé



OKD / Openshift sur les IaaS



Bare Metal

- ▶ historiquement, Openshit Origin (futur OKD) ne s'installait que via Ansible et surtout sur Bare Metal ou VM. PLMshift est déployé sous cette forme² ;
- ▶ maîtrise des coûts d'infrastructure ;
- ▶ cohabitation des architectures matérielles, pure Bare Metal, VMs, GPUs, réseaux, etc.



²<https://plmteam.pages.math.cnrs.fr/docs/okd>

³<https://replay.jres.org/videos/watch/47e4a3e4-031c-46f3-b5e9-ca34b6615127>



Bare Metal / IaaS

Bare Metal

- ▶ historiquement, Openshit Origin (futur OKD) ne s'installait que via Ansible et surtout sur Bare Metal ou VM. PLMshift est déployé sous cette forme² ;
- ▶ maîtrise des coûts d'infrastructure ;
- ▶ cohabitation des architectures matérielles, pure Bare Metal, VMs, GPUs, réseaux, etc.

IaaS / Openstack

- ▶ *openshift-installer* permet de déployer sur IaaS un cluster complet, administré et supervisé en 20 à 40 minutes, avant la configuration ArgoCD³ ;
- ▶ délégation de l'exploitation de l'infrastructure ;
- ▶ abstraction des spécificités d'architecture (par la virtualisation).

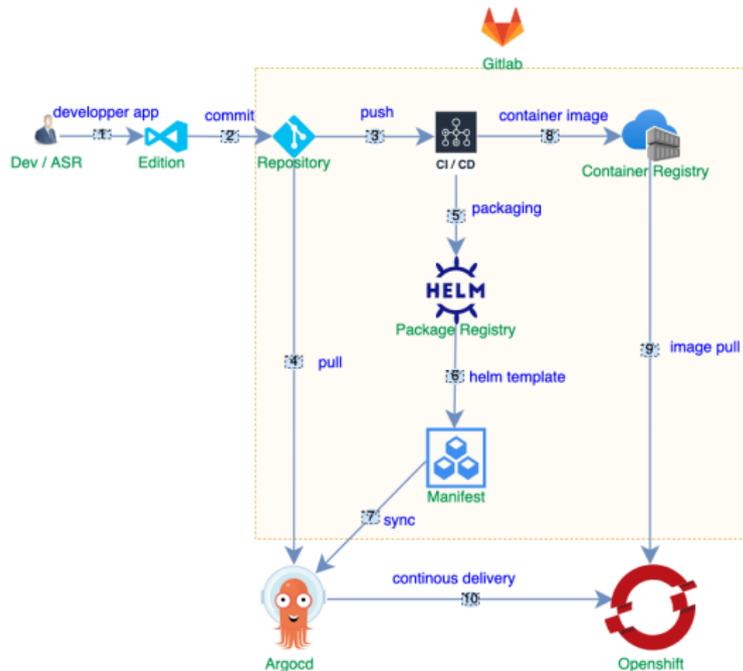


²<https://plmteam.pages.math.cnrs.fr/docs/okd>

³<https://replay.jres.org/videos/watch/47e4a3e4-031c-46f3-b5e9-ca34b6615127>



Administrer aujourd'hui des clusters K8s



<https://charts.math.cnrs.fr>

<https://plmlab.math.cnrs.fr/plmshift>

<https://plmlab.math.cnrs.fr/plmteam/services>

Retour d'expérience sur les IaaS

Les choix techniques sont stratégiques

- ▶ stockage bloc performant (I/O intensif) ;
- ▶ gabarits adaptés (rapport RAM / CPU favorable pour la RAM) ;
- ▶ pas de surcharge des hyperviseurs (overprovisionning).

Retour d'expérience sur les IaaS

Les choix techniques sont stratégiques

- ▶ stockage bloc performant (I/O intensif) ;
- ▶ gabarits adaptés (rapport RAM / CPU favorable pour la RAM) ;
- ▶ pas de surcharge des hyperviseurs (overprovisionning).

Des équipes à l'écoute

- ▶ très bonnes relations avec les Ops des IaaS ;
- ▶ se libérer du travail d'infra ne doit pas faire oublier que d'autres le font (pas de pression sur les Ops des IaaS) ;
- ▶ relativiser les temps d'indisponibilité ;
- ▶ équipes preneuses du retour d'expérience pour faire évoluer les infrastructures.

Un immense merci aux équipes de Gricad et Virtualdata

Accompagner les utilisateurs

Dès le début, nous avons fait le choix d'ouvrir la plateforme aux utilisateurs :

- ▶ prise en main non triviale et montée en compétence nécessaire ;
- ▶ en 2021, les conteneurs et K8S ne sont pas très connus dans le milieu des utilisateurs des unités ;
- ▶ beaucoup de documentation, de formation, d'accompagnement pas à pas ;
- ▶ explosion de l'utilisation du Chat Mattermost de la PLM ;
- ▶ sollicitation beaucoup plus interactive entre la PLMTeam et les utilisateurs ;
- ▶ implication beaucoup plus importante des utilisateurs, notamment des chercheurs et ingénieurs calcul scientifique dans le développement et la maintenance des applications.



Retour d'expérience, les plus

Est-ce que l'expérience est concluante ?

- ▶ dix fois plus de services déployés en 4 ans qu'en 15 ans ;
- ▶ nous suivons beaucoup mieux le cycle de vie et l'état des applications qui sont rassemblées dans un même cluster ;
- ▶ nous sommes soulagés d'une administration historique des infrastructures et des OS ;
- ▶ mise à disposition d'un outil qui permet de maquetter très rapidement de nouveaux services ;



Retour d'expérience, les plus

Est-ce que l'expérience est concluante ?

- ▶ dix fois plus de services déployés en 4 ans qu'en 15 ans ;
- ▶ nous suivons beaucoup mieux le cycle de vie et l'état des applications qui sont rassemblées dans un même cluster ;
- ▶ nous sommes soulagés d'une administration historique des infrastructures et des OS ;
- ▶ mise à disposition d'un outil qui permet de maquetter très rapidement de nouveaux services ;
- ▶ les utilisateurs avertis ont accès à un service avancé ;
- ▶ une telle plateforme rapproche la PLMTeam des membres des unités ;



Retour d'expérience, les plus

Est-ce que l'expérience est concluante ?

- ▶ dix fois plus de services déployés en 4 ans qu'en 15 ans ;
- ▶ nous suivons beaucoup mieux le cycle de vie et l'état des applications qui sont rassemblées dans un même cluster ;
- ▶ nous sommes soulagés d'une administration historique des infrastructures et des OS ;
- ▶ mise à disposition d'un outil qui permet de maquetter très rapidement de nouveaux services ;
- ▶ les utilisateurs avertis ont accès à un service avancé ;
- ▶ une telle plateforme rapproche la PLMTeam des membres des unités ;
- ▶ nous bénéficions d'une meilleure visibilité de notre action et d'une valorisation de notre implication dans le projet collectif.



Retour d'expérience, les moins

- ▶ Expérience très positive mais plus complexe que nous aurions pu l'imaginer en 2018...



Retour d'expérience, les moins

- ▶ Expérience très positive mais plus complexe que nous aurions pu l'imaginer en 2018...
- ▶ L'observabilité est un noeud gordien : l'explosion des données collectées nécessite la mise en place d'outils adaptés.
On supervise beaucoup plus... supervisons-nous beaucoup mieux ?
- ▶ Pas de plateforme unique, compensée par une utilisation intensive du modèle GitOps.

Retour d'expérience, les moins

- ▶ Expérience très positive mais plus complexe que nous aurions pu l'imaginer en 2018...
- ▶ L'observabilité est un noeud gordien : l'explosion des données collectées nécessite la mise en place d'outils adaptés.
On supervise beaucoup plus... supervisons-nous beaucoup mieux ?
- ▶ Pas de plateforme unique, compensée par une utilisation intensive du modèle GitOps.
- ▶ Les opérateurs de Kubernetes masquent la complexité des technologies installées mais ne nous exemptent pas de la montée en compétence nécessaire.

Retour d'expérience, les moins

- ▶ Expérience très positive mais plus complexe que nous aurions pu l'imaginer en 2018...
- ▶ L'observabilité est un noeud gordien : l'explosion des données collectées nécessite la mise en place d'outils adaptés.
On supervise beaucoup plus... supervisons-nous beaucoup mieux ?
- ▶ Pas de plateforme unique, compensée par une utilisation intensive du modèle GitOps.
- ▶ Les opérateurs de Kubernetes masquent la complexité des technologies installées mais ne nous exemptent pas de la montée en compétence nécessaire.
- ▶ Est-ce que nous avons une utilisation efficiente des ressources matérielles ? Hum ! Nous avons encore une marge de progression...

Retour d'expérience, les moins

- ▶ Expérience très positive mais plus complexe que nous aurions pu l'imaginer en 2018...
- ▶ L'observabilité est un noeud gordien : l'explosion des données collectées nécessite la mise en place d'outils adaptés.
On supervise beaucoup plus... supervisons-nous beaucoup mieux ?
- ▶ Pas de plateforme unique, compensée par une utilisation intensive du modèle GitOps.
- ▶ Les opérateurs de Kubernetes masquent la complexité des technologies installées mais ne nous exemptent pas de la montée en compétence nécessaire.
- ▶ Est-ce que nous avons une utilisation efficace des ressources matérielles ? Hum ! Nous avons encore une marge de progression...
- ▶ Et enfin, l'augmentation de la complexité est-elle soutenable ?

Nos projets

- ▶ Rationaliser les plateformes
 - ▶ migrer la plateforme expérimentale Bare Metal sur IaaS, avec intégration des nœuds spécifiques (vGPUs, Big Mem) sur les IaaS ou des nœuds Edge sur MesoNet ;
 - ▶ garder une plateforme pour les services institutionnels et un PRA ;
 - ▶ disposer d'un PRA fournissant un service S3 à la demande ;
 - ▶ améliorer le monitoring ;

Nos projets

- ▶ Rationaliser les plateformes
 - ▶ migrer la plateforme expérimentale Bare Metal sur IaaS, avec intégration des nœuds spécifiques (vGPUs, Big Mem) sur les IaaS ou des nœuds Edge sur MesoNet ;
 - ▶ garder une plateforme pour les services institutionnels et un PRA ;
 - ▶ disposer d'un PRA fournissant un service S3 à la demande ;
 - ▶ améliorer le monitoring ;
- ▶ Des services pour les unités
 - ▶ avec Victoria-Metrics : un monitoring pour les laboratoires ;
 - ▶ avec Noobaa : un service S3 à la demande multi-site ;

Nos projets

- ▶ Rationaliser les plateformes
 - ▶ migrer la plateforme expérimentale Bare Metal sur IaaS, avec intégration des nœuds spécifiques (vGPUs, Big Mem) sur les IaaS ou des nœuds Edge sur MesoNet ;
 - ▶ garder une plateforme pour les services institutionnels et un PRA ;
 - ▶ disposer d'un PRA fournissant un service S3 à la demande ;
 - ▶ améliorer le monitoring ;
- ▶ Des services pour les unités
 - ▶ avec Victoria-Metrics : un monitoring pour les laboratoires ;
 - ▶ avec Noobaa : un service S3 à la demande multi-site ;
- ▶ Poursuivre les échanges avec les équipes d'infrastructure
 - ▶ afin de faire converger les spécifications requises pour assurer le bon fonctionnement de ces plateformes.

Nos conclusions

- ▶ imaginons une offre de solutions PaaS pour les unités et les utilisateurs finaux par communauté, par site ou au niveau national ;
- ▶ de notre point de vue, Kubernetes est suffisamment mature et offre un véritable plus dans le déploiement et la maintenance d'applications ;
- ▶ cela nécessite bien sûr une évolution des compétences et des usages, par la formation et en resserrant les liens entre Devs et Ops.