Infrastructures de recherche

Collaborer, stocker, déposer, archiver

Université de Poitiers

Université de Poitiers

Enjeux des infrastructures autour de la donnée

La feuille de route de la DSIN ambitionne de répondre à plusieurs enjeux.

- **Un enjeu de flexibilité** en apportant des solutions agiles répondant à un besoin croissant en matière de stockage des données et de leurs traitements.
- **Un enjeu d'efficacité** qui se traduit par:
  - la mise en relation des ressources accessibles aux équipes scientifiques,
  - la mise à disposition de moyens à l'état de l'art technologique et des usages scientifiques,
  - la mutualisation des ressources pour générer des économies d'échelle sur les coûts d'équipements et de fonctionnement.
- **Un enjeu relatif à la cohérence** en proposant des services intégrés aux offres régionales ou nationales.

Nos infrastructures

Un réseau à Très Haut Débit RINA autour d'un datacenter (DC) labélisé

Proposer une interconnexion répondant aux besoins d'une latence la plus faible et à des débits importants.

Une offre de stockage Ceph/S3

Assurer un stockage capacitif répondant aux besoins de conservation et d'accessibilité des données de la recherche.

Des moyens de calcul locaux

Permettre un calcul local tremplin vers des moyens régionaux ou nationaux.

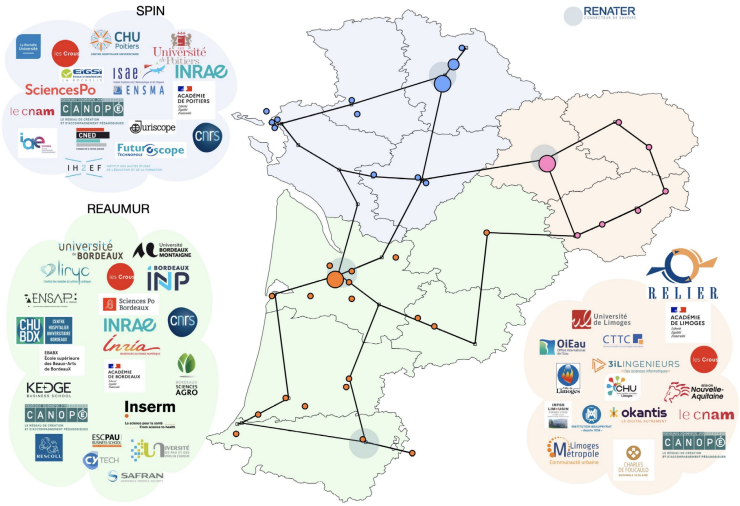
Des services pour accéder à nos ressources

Faciliter l'accès aux ressources de calcul autour des données et faciliter la diffusion des codes

Université de Poitiers

Connectivité et hébergements

Une modernisation des infrastructures techniques des réseaux d'accès et des établissements sur le CPER.



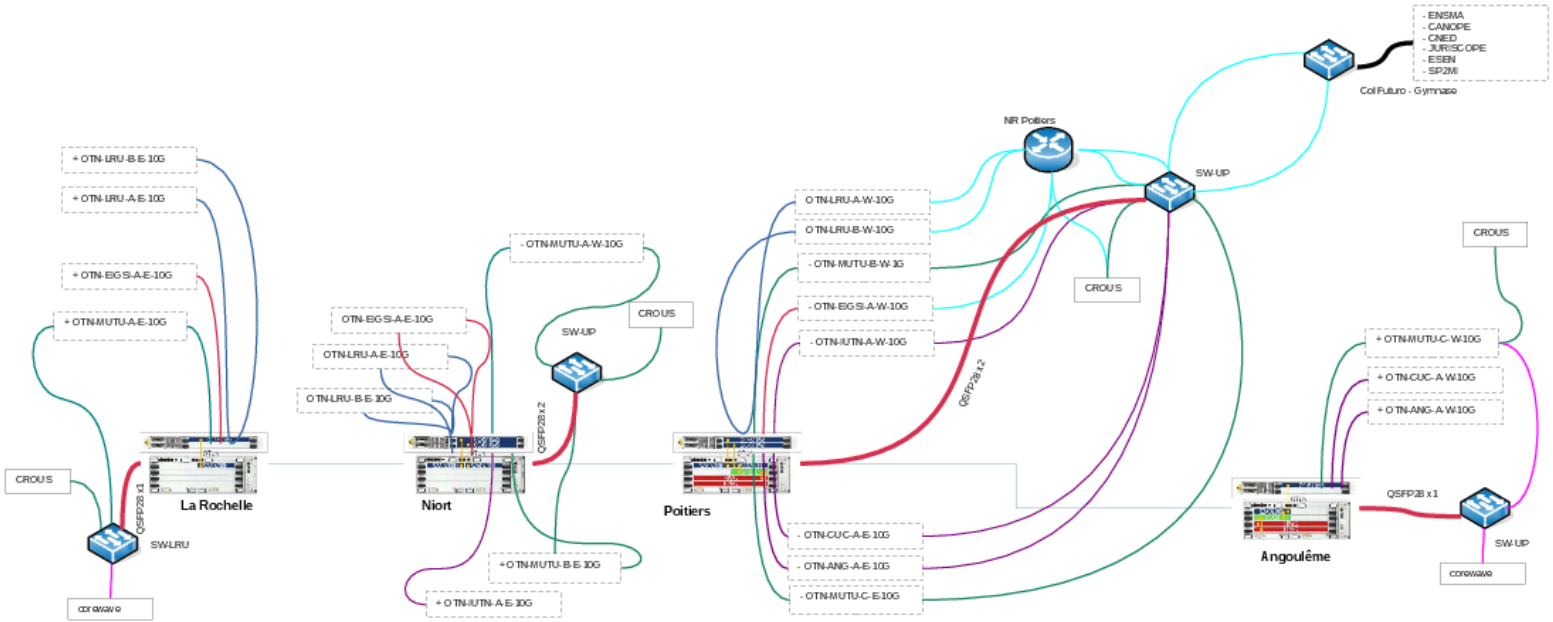
RINA

- IRU 15 ans
- 2200 km FON
- n x 100 Gb/s
- ESR / éducation / Territorial, Hospitalier

DC

- Salles mutualisées :
  - *Bordeaux*, *Limoges*, *Poitiers*, etc.
- Travaux en cours avec la région NA

Connectivité et hébergements



Connectivité des campus

Une mise à jour permanente des débits et de la connectivité :

- D'un coeur réseau à 4 x 10 Gb/s vers 100 Gb/s ;
- Modernisation des commutateurs sur tous nos campus ;
- Passage d'offres louées aux opérateurs vers des fibres optiques noires en propre ;
- Mise en place du FTTO pour répondre aux débits de la recherche ;
- Hébergement de nos serveurs de 10 Gb/s à 25 Gb/s.

Université de Poitiers

Une offre de stockage Ceph/S3

Ceph/S3, un service de stockage capacitif à la demande pour le stockage de très gros volumes de données ne nécessitant pas des accès trop performants.

- Bénéfices :
  - volumétrie de 2.1 Po bruts,
  - une protection des données sur 3 salles de notre établissement,
  - couplage au cloud par le standard S3 avec la possibilité de **migrer les données** facilement d'un système à un autre,
  - entrepôt de données **accessible en dehors du SI** en autonomie par les équipes de recherche,
  - augmentation incrémentale de la plateforme grace à un modèle économique
  - référencement et descriptions des données.

Université de Poitiers

Le principe du stockage objet 1/2

Le stockage objet est une méthode moderne de gestion de données qui a gagné en popularité en raison de sa flexibilité, de sa scalabilité et de sa facilité d'utilisation.

Objets et Métadonnées :

- Les données sont stockées sous forme d'objets ;
- Chaque objet est accompagné de métadonnées descriptives ; informations supplémentaires (tailles ou autres infos. . .).

Adressage via Clé :

- Chaque objet est identifié de manière unique par une clé ;
- adressage via clé permet un accès rapide et efficace aux objets.

Le principe du stockage objet 2/2

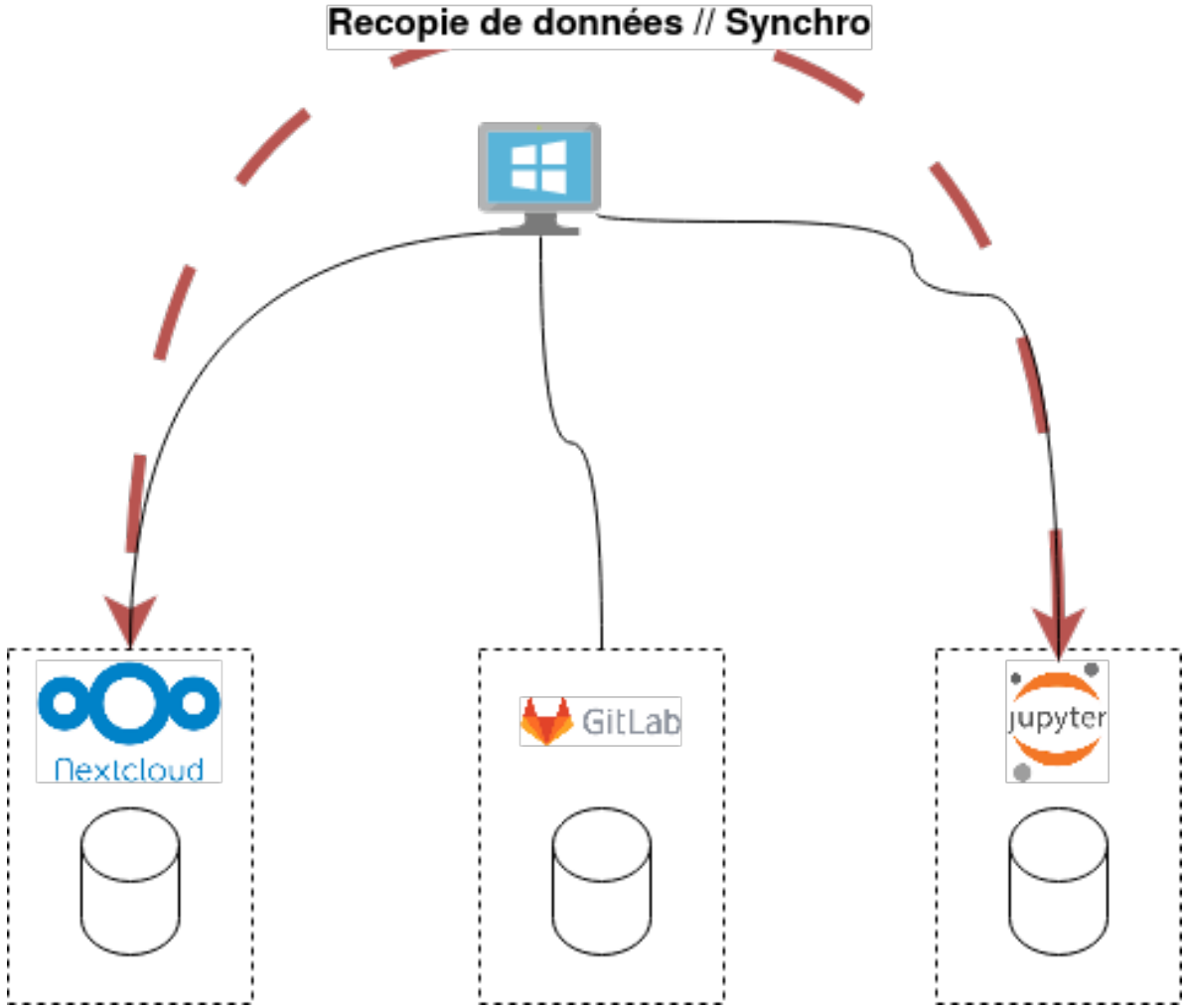
Scalabilité Horizontale :

- Le stockage objet est conçu pour être hautement scalable ;

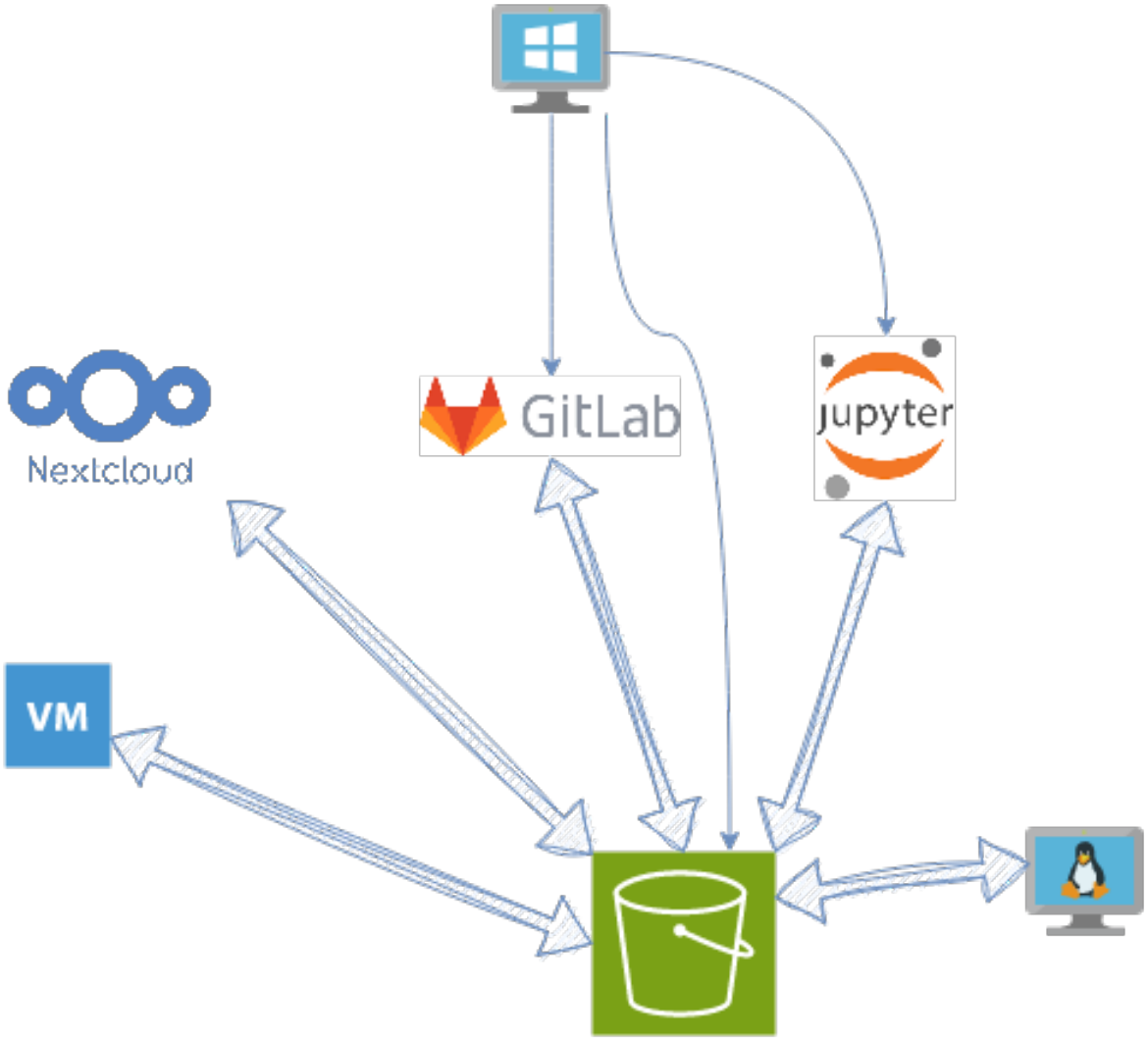
Redondance et Fiabilité :

- Pour garantir la disponibilité et la résilience, le système de stockage utilise la redondance des copies ;
- stockage en 1:3 (chaque objet à 2 copies).

Problématiques d'un stockage non centralisé



Centralisation du stockage aux services



Université de Poitiers

Moyens dédiés à la virtualisation

Un service de virtualisation basé sur 7 serveurs PROXMOX :

- AMD EPYC 7713 64-Core Processor (2 Sockets) soit un total de 256 cores ;
- RAM : 1.9 To ;
- Stockage : 550 Go utiles cache NVMe, 20 To sur des SSD locaux.
- 2 Cartes graphiques dédiées NVIDIA A40

Pour répondre aux besoins en hébergement de machines virtuelles dédiées à la pédagogie, à la recherche et aux déploiements de nouveaux services.

Université de Poitiers

Moyens de calcul GPU

Cette machine de calcul GPU est conçue pour des applications spécifiques qui bénéficient de la puissance de calcul parallèle offerte par les GPU, telles que l'apprentissage automatique (machine learning), l'apprentissage profond (deep learning), la simulation scientifique.

Une machine Apollo 6500 HPE intégrant :

- 8 Cartes NVIDIA A100 ;
- AMD EPYC 7543 32-Core Processor (128 cores) ;
- RAM 512 Go
- Stockage SSD et NVMe.

Université de Poitiers

Moyens de calcul locaux

Architecture NUMA (Non-Uniform Memory Access) permet à tous les processeurs d'avoir accès à la mémoire à des vitesses différentes en fonction de leur position dans la topologie.

- Bénéfices :
  - simplicité de mise en oeuvre OpenMP vs MPI,
  - intégration aux applications,
  - forte adaptation aux environnements conteneurisés ou virtualisés,
  - déploiement de portails de travaux ou de solutions type cloud.

JupyterHub

JupyterHub est une application web open-source qui permet de créer et de gérer des serveurs Jupyter Notebook de manière centralisée.

- Bénéfices :
  - **partage et reproductibilité** $\rightarrow$ les notebooks Jupyter devenus un standard, partage simple, favorise la transparence et la reproductibilité de la recherche,
  - **intégration de divers langages de programmation** $\rightarrow$ plusieurs langages disponibles comme Python, R, Julia mais aussi intégration de OCAML ou Fortran,
  - **intégration de moyens de calcul et de stockage** $\rightarrow$ calculs déportés sur machine NUMA, GPU ou dédiée à la visualisation et accessibilité des buckets Ceph/S3.

Université de Poitiers

NextCloud

Nextcloud est une plateforme open source de services de cloud computing qui permet aux utilisateurs de stocker et de partager des fichiers, ainsi que d'accéder à diverses applications et fonctionnalités collaboratives.

- Bénéfices :
  - abstraction du stockage S3,
  - stockage de fichiers,
  - partage des fichiers,
  - synchronisation.

Université de Poitiers

Autres services internes à la DSI

D'autres services sont déployés à la DSI pour un usage interne :

- GitLab : gestion de codes et intégration continue ;
- Elasticsearch : indexation et recherche de données ;
- Différents types de SGBD SQL ou NoSQL.

Ces services ou recettes d'ingénieries peuvent avoir vocation à être déployés ou utilisés pour de nouveaux services.

Université de Poitiers

Merci pour votre attention

Questions ?

Université de Poitiers