

Contexte

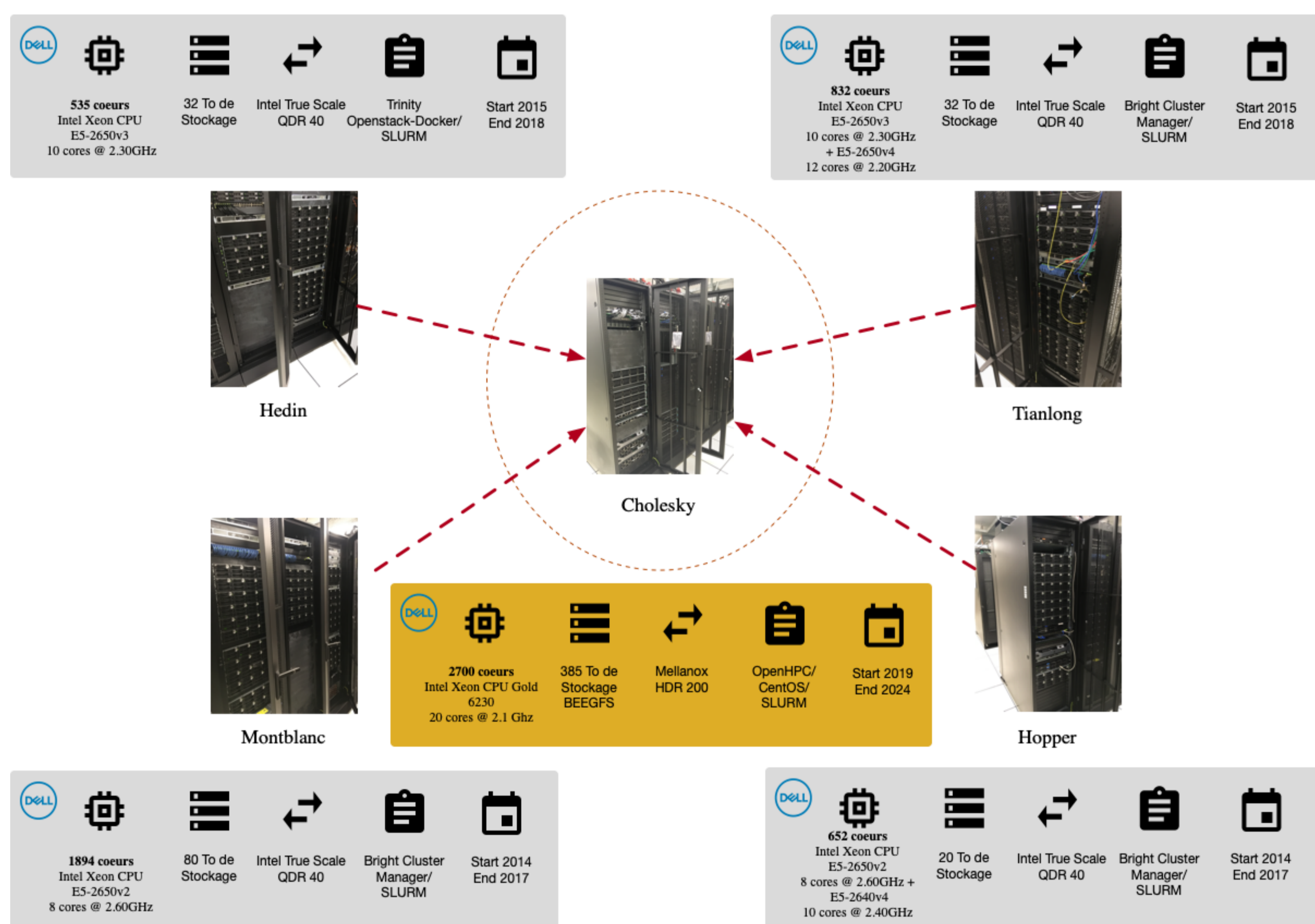
Le Centre de Physique Théorique est une UMR CNRS-École polytechnique (X) située à Palaiseau (91). Celle-ci réunit des chercheurs dont les activités couvrent un large spectre de la Physique aussi bien fondamentale qu'appliquée. Pour l'élaboration de théories nouvelles, les chercheurs s'appuient sur des clusters de calcul du CPHT, et un cluster mutualisé opéré par l'unité de service IDCS (Infrastructure-Données-Calcul-Scientifique) de l'X. Les clusters du CPHT financés par projets (ERC) sont arrivés en fin de support : **les systèmes d'exploitation et les bibliothèques sont obsolètes et ne sont plus maintenus.**

Face à la demande des chercheurs de continuer à utiliser ces ressources (processeurs encore d'actualité), l'équipe informatique accompagnée d'un prestataire externe a initié un **projet de jouvence** des clusters en les **intégrant au cluster mutualisé de l'unité IDCS.**

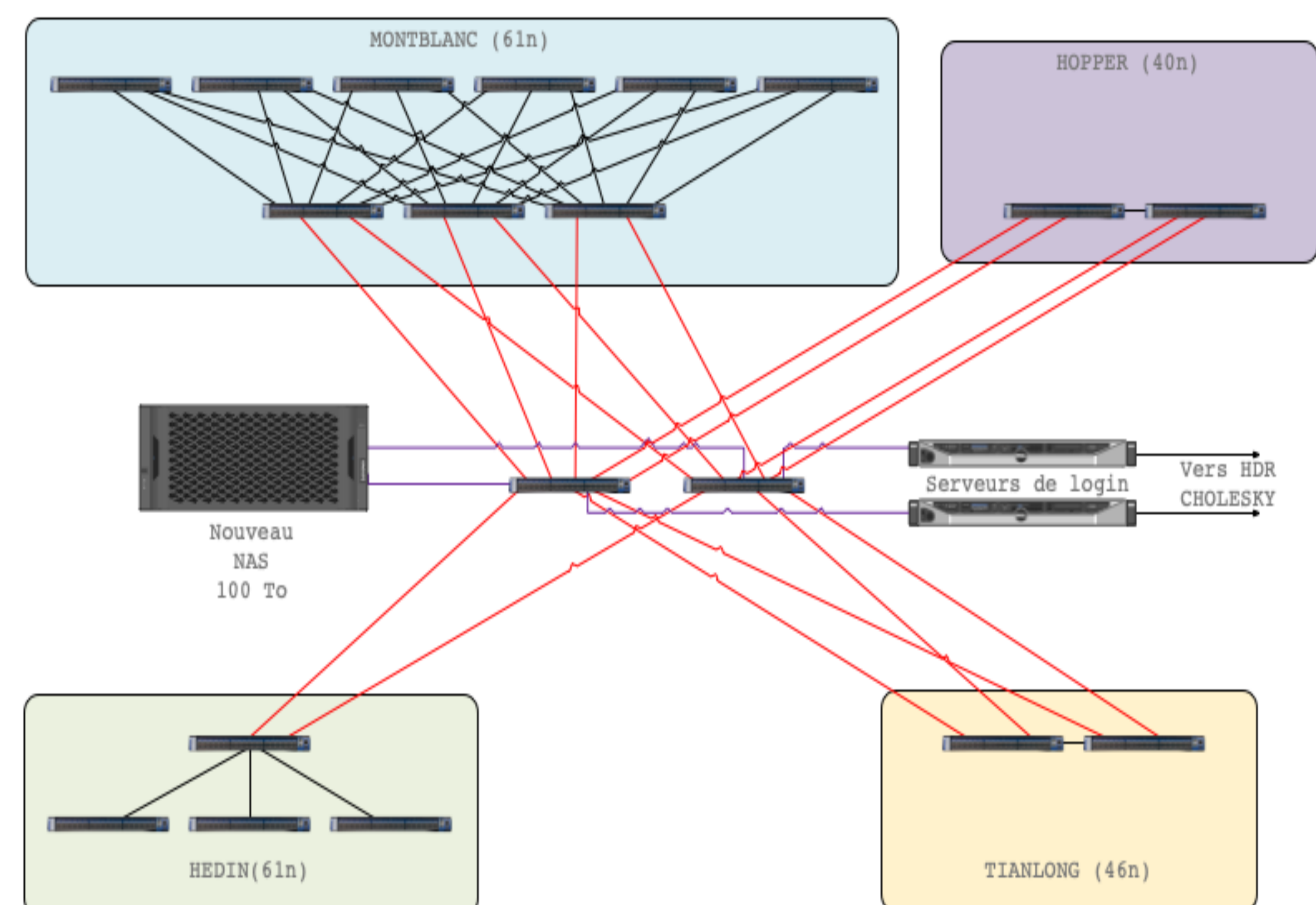
Ce projet a nécessité **peu de moyens matériels et s'appuie sur les ressources matérielles et logicielles existantes.** L'architecture finale se veut extrêmement simple et facile à maintenir : 2 réseaux « Infiniband » indépendants (HDR et QDR), un espace disque « scratch » (workdir) par utilisateur sur chaque réseau IB (soit 1 espace de stockage de type NAS par réseau IB), un seul système d'exploitation et une seule arborescence de bibliothèques pour l'ensemble du cluster.

Ce projet a été financé sur des ressources propres de 3 unités CPHT, CMAP (Centre de Mathématiques Appliquées), et IDCS pour un **budget d'environ 40 KE.**

Intégration



Architecture IB QDR



Avantages

- Architecture logicielle homogène pour l'ensemble des nœuds :
 - un système d'exploitation
 - une arborescence de modules (bibliothèques) LMOD
 - un annuaire LDAP de gestion des comptes utilisateurs
 - une base d'accounting SLURM
 - un système de déploiement d'images des nœuds XCAT
- Un guichet d'entrée unique (serveurs de login) pour accéder à l'ensemble des ressources de calcul et soumettre des jobs

Limites

- Nécessite de recompiler la bibliothèque OpenMPI pour la prise en charge des 2 couches transport MTL IB : **PSM 1.x** (Intel TrueScale Fabric) et **UCX** (Mellanox)
- Couche transport PSM 1.x dépréciée (non maintenue) dans OpenMPI 5.x et Intel MPI (>2019) : il faut utiliser la bibliothèque tierce « libfabric » (OFI)