

ÉLÉMENT DE PORTFOLIO 02



Instrument de recherche

1 DÉFINITION DE CET ÉLÉMENT

Titre de l'élément : Évolution de la composante calcul de l'infrastructure de service du LIP6

2 MOTIVATIONS DU CHOIX DE CET ÉLÉMENT

La Direction de l'ingénierie Numérique du LIP6 s'appuie sur des équipements en constante évolution pour rendre tous les services qu'elle offre aux équipes de recherches du laboratoire. L'ensemble de ces infrastructures sont divisées en un bloc de production d'une part et une Infrastructure de Services (IS) d'autre part.

Cette dernière est composée de tous les outils conçus, créés et mis en place pour les équipes du laboratoire et pour lesquels d'autres unités ou partenaires ont manifestés un intérêt, ou qui ont été estimés comme pouvant servir la recherche au delà du seul périmètre du LIP6.

Nous avons regroupés tous nos dispositifs au sein d'une "infrastructure de service" (IS) balayant un spectre plus large et moins chronophage administrativement. Les ingénieurs maintiennent ainsi les équipements tels qu'ils l'auraient fait pour un service local, sans plus, et le financement et la jouvence de ses outils est assumé par l'ensemble des unités et services l'utilisant. La description de l'évolution de la composante calcul de l'infrastructure de service du LIP6 illustre ce mode de fonctionnement.

3 PRÉSENTATION DE CET ÉLÉMENT

La composante calcul de l'IS du LIP6 est composée de trois ressources principales :

- ▶ un cluster HPC (tier 3) de plus de 660 cœurs et 4To de RAM permettant aux équipes de recherche d'effectuer de petits calculs et de mettre au point les codes pour des exécutions sur des tiers supérieurs ;
- ▶ un cluster type "cloud computing" offrant les droits maximaux (super administrateur, ou "root") aux équipes de recherches sur des grappes de machines homogènes ;
- ▶ un "cluster GPU" offrant des ressources matérielles munies d'accélérateurs graphiques (GPU), adaptées aux besoins spécifiques des algorithmes d'intelligence artificielle.

Au cours des dernières années, l'infrastructure de service du LIP6 a assuré la rénovation des deux derniers clusters afin d'offrir à nos communauté des outils toujours plus performants tout en intégrant la problématique écologique oubliée jusqu'alors.

3.1 Jouvence du cluster "cloud computing"

Fin 2018, le cluster `small` du LIP6 était composé de 24 nœuds homogènes chacun à base de deux processeurs Intel Xeon E5645 à 2,40GHz (Nehalem Gulftown, série 5600, hexacore HyperThreadés), dans des chassis Dell R410 dotés de 8 Go de mémoire vive.

Côté logiciel, le cluster offrait à la communauté scientifique la possibilité de travailler avec tous les droits possibles sur n'importe quel système d'exploitation. En effet, les ingénieurs du LIP6 ont adapté la technologie "kadeploy", à l'origine conçue pour Grid 5000¹ à ce petit cluster. Il permettait donc aux équipes du LIP6 de réserver un ensemble de ressources et d'y installer tous les outils qu'elles désirent dans un environnement sécurisé et complètement maîtrisé.

La haute adaptabilité de `small` en faisait un outil idéal pour les phases d'expérimentation et de mises au point de codes avant le passage sur les moyens de calculs nationaux. Le code pouvait-être installé, testé et débogué

1. <https://www.grid5000.fr>

dans un environnement fait sur-mesure pour correspondre au mieux au calculateur sur lequel il serait déployé et exécuté in fine.

Début 2019, dans sa politique de rationalisation des ressources de l'unité, une baisse légère de l'utilisation de ce cluster a été constatée. Ce signal faible a mené le service informatique à enquêter auprès des utilisateurs du laboratoire et il a été constaté que l'architecture matérielle de `small` ne permettait plus de couvrir tous les besoins de notre communauté scientifique. Par exemple, les processeurs de notre cluster, mis sur le marché mi-mars 2010 n'intégraient plus certaines fonctionnalités attendues (NUMA², AVX-512³, etc.), excluant de faits de nombreux tests, enfin, la mémoire disponible sur chaque nœud du cluster était devenue largement insuffisante.

En se basant sur les besoins de ses équipes et des partenaires utilisant le cluster, le service informatique a défini le contour d'une jouvence de `small` qu'il a formalisé par la réponse à un appel à projet sur la campagne 2019 du soutien plateformes de INS2I.

Suite à un arbitrage favorable du CNRS, et à un co financement par le laboratoire, 150 000 € ont été levés pour cette opération d'envergure qui a permis au LIP6 de construire le cluster Tall composé de 20 bi processeurs Intel Xeon Silver 4216 (32 cœurs / 64 threads à 2.10 GHz) et 384 Go de RAM.

Cette machine, s'appuie sur les concepts de souplesse, de robustesse et de sécurité appliqués aux systèmes d'information du laboratoire et offre aux membres du laboratoire un outil performant, largement utilisé et plébiscité.

3.2 Le cluster GPU Convergence@Paris

En 2019, le LIP6 était doté d'un cluster GPU, hébergé au laboratoire et maintenu par le service informatique mais dont l'usage était exclusivement réservé à l'équipe qui avait financé les nœuds de cet outil. Face à l'augmentation de la demande en temps de calcul GPU, le service informatique du LIP6, en se basant sur l'expérience acquise via le cluster d'équipe, a commencé à définir les contours d'un cluster GPU moderne et mutualisé pour ses équipes.

Au même moment, nos partenaires du LIMICS, Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances en e-Santé, sur le campus de Sorbonne Université, disposaient dans le cadre de projets de financements pour acquérir un GPU, mais ils ne disposaient plus du personnel en interne pour les conseiller et gérer cette ressource. Les besoins de nos unités étant assez différents, il a été décidé de mettre en place un POC⁴ afin de délimiter précisément les besoins et de définir une potentielle nouvelle machine. Cet outil, composé de deux nœuds de calcul en bi Intel Xeon Silver 4216 (32 cœurs / 64 threads @ 2.10 GHz), offrant 8 cartes GPU NVIDIA Titan RTX, a été déployé en 2020. Hébergée au LIP6 il a rapidement suscité l'adhésion et de nombreuses demandes de comptes pour des utilisateurs en dehors de notre unité ont afflué.

Bientôt, d'autres unités telle que le LCQB (Laboratory of Computational and Quantitative Biology) expérimentant la solution proposée, sont venues financer ce projet et fin 2020, l'IRIF, laboratoire en informatique de l'Université Paris Cité, a rejoint le LIP6 dans un projet de cofinancement de deux nœuds complémentaires combinant à la fois CPU, GPU et grande quantité de mémoire vive. Grâce aux retours des usagers du POC, constitués en un comité de pilotage, la machine a donc évolué d'une structure pure GPU vers une architecture "convergente"⁵.

Cette évolution a considérablement augmenté le coût d'un nœud, c'est pourquoi nous avons déposé un projet de co financement auprès de Sorbonne Université et un second auprès du DIM RFSI⁶ afin de monter une structure de taille suffisante pour répondre aux besoins des acteurs partenaires en île-de-France.

Ces projets ont reçu un accueil enthousiaste de la communauté scientifique et au second trimestre 2021, 650 000 € ont été consacrés à l'acquisition de la machine Convergence@Paris.

Malgré de nombreuses péripéties (changement de marché matériel ayant entraîné des impossibilités de commandes pendant plusieurs mois ; changement de gamme sur le matériel évalué lors des premiers essais ; blocage des commandes sur plus de six mois ; crises sanitaires ; problèmes d'approvisionnements en composants, retards dans les travaux sur Sorbonne Université, etc.), la machine Convergence@Paris a intégré les locaux du LIP6 fin décembre 2022 et est en cours de câblage.

La configuration de Convergence@Paris permettra bientôt aux membres du projet d'accéder aux ressources fournies par :

- ▶ 9 nœuds bi-Intel® Xeon® Gold 6330 2G, 28C/56, 1To de RAM et 4 GPU Nvidia A100 - 80G dans des chassis DELL R750xa ;

2. Non uniform memory access

3. Advanced Vector Extensions

4. Preuve de concept

5. Il s'agit de nœuds regroupant à la fois des CPU performants et des cartes GPU.

6. Domaine d'Intérêt Majeur, Réseau francilien en sciences informatiques

- ▶ 1 nœuds bi-AMD EPYC 7543 2.8GHz, 32C/64, 1To de RAM et 4 GPU NVIDIA A100 SXM4 - 80G dans un chassis DELL XE8545 ;
- ▶ 250To de stockage redondé.

L'ensemble de la structure et interconnectée en interne par un réseau Mellanox à 100Gbits/s et par une connexion 10Gbits/s vers la structure du LIP6.

Des contributions d'entités diverses, exploitant encore le POC GPU du LIP6, ont déjà permis, fin 2021, de faire évoluer cette machine et 50K€ complémentaires ont permis de doubler la quantité de mémoire vive de la moitié des nœuds leur permettant d'attendre les 2To/nœuds.

Convergence@Paris devrait être opérationnelle et en pleine exploitation vers la mi-2023.