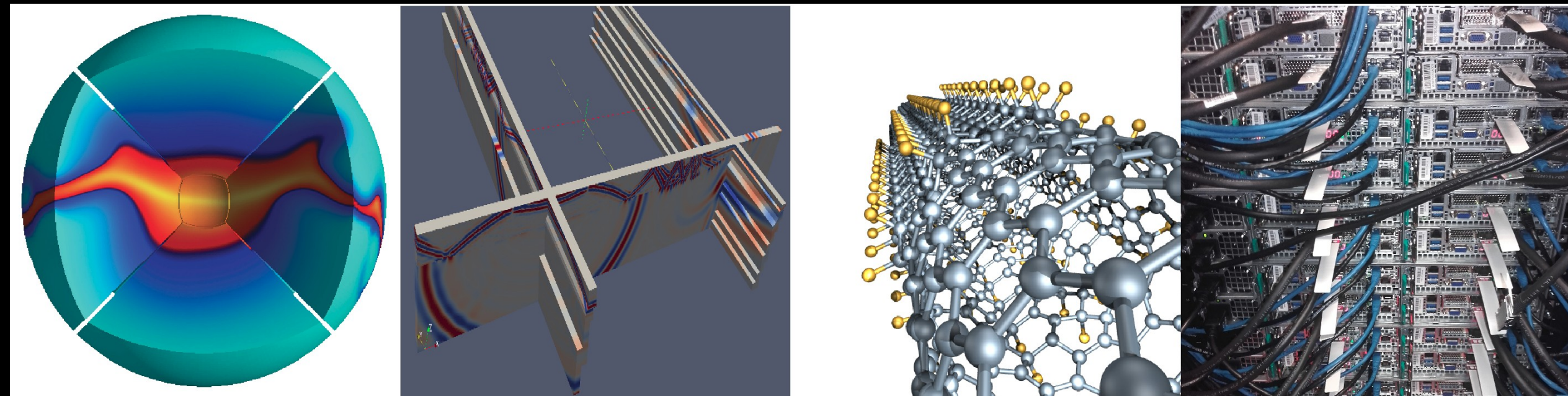




GLiCID



Présentation Générale



Plan



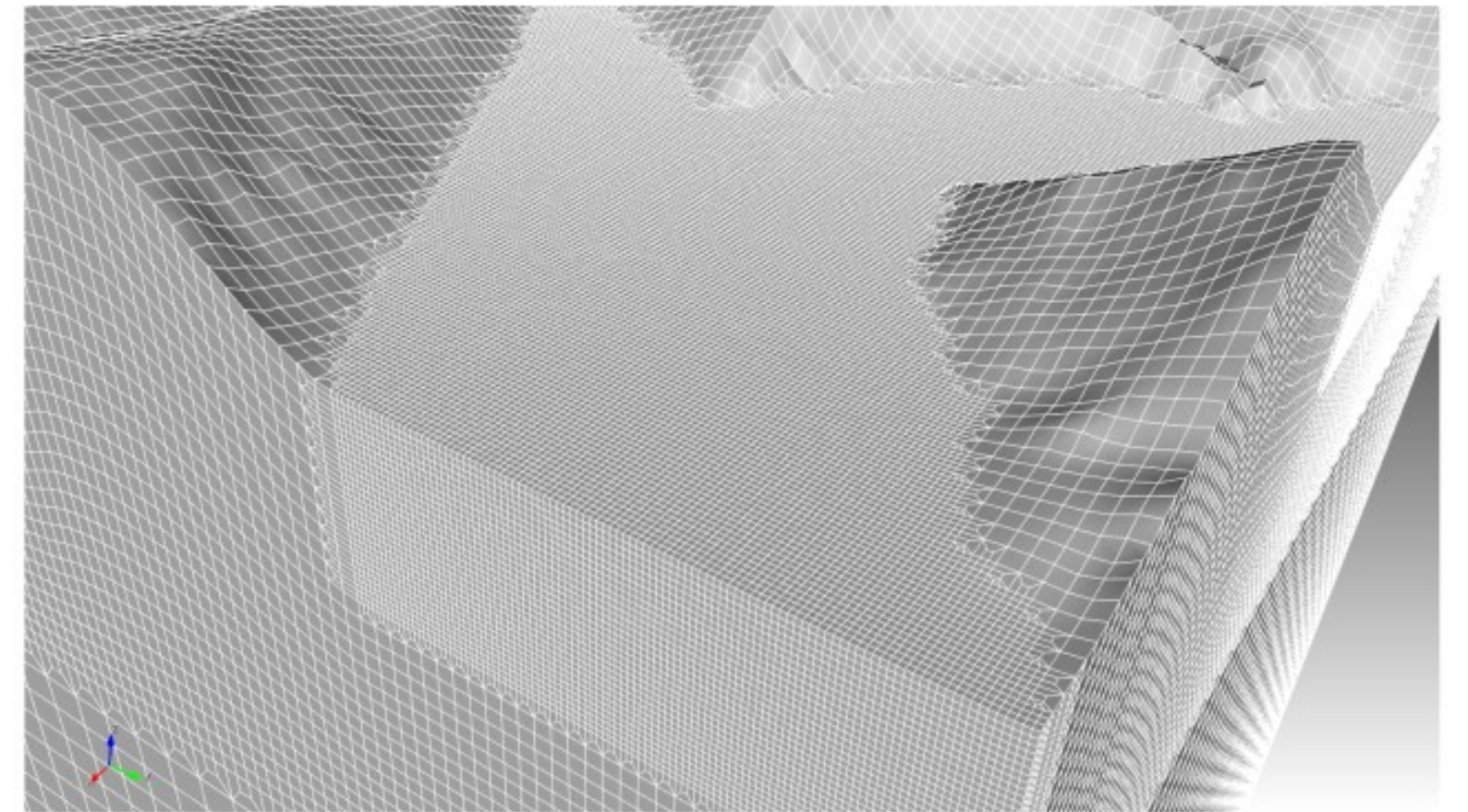
- Généralités
- Le projet GLiCID
- L'unité de Service GLiCID
- Les équipements
- Modèle économique
- Questions pratiques

Le calcul distribué : principe de base

Un problème à résoudre :

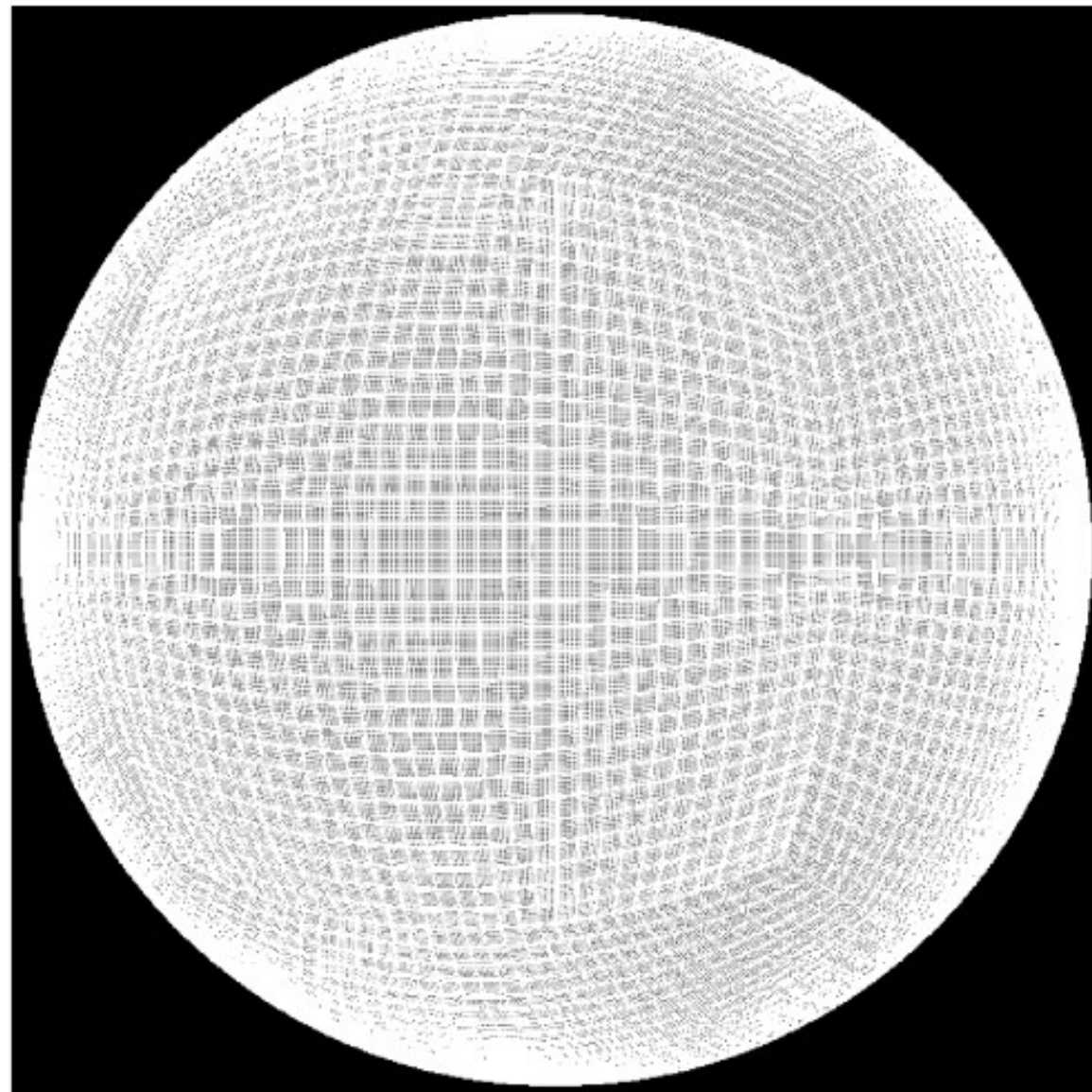
Équation des ondes

$$\rho \partial_{tt} \mathbf{u} - \nabla \cdot \boldsymbol{\sigma} = \mathbf{s}_s$$
$$\boldsymbol{\sigma} = \mathbf{c} : \boldsymbol{\epsilon}(\mathbf{u})$$

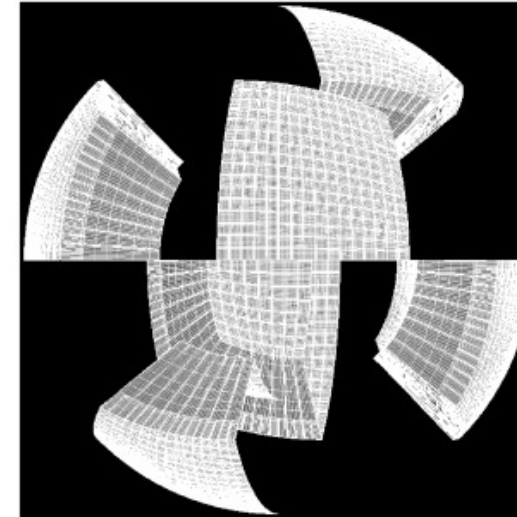


Le calcul distribué

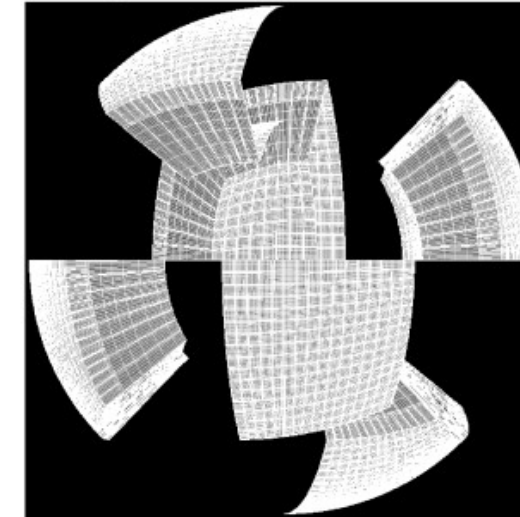
Un problème à résoudre :



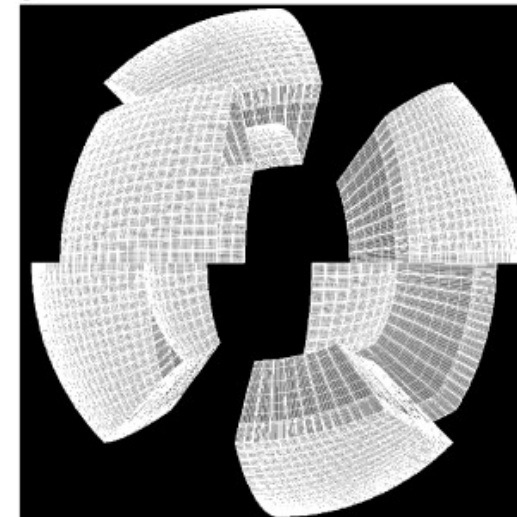
processeur 0



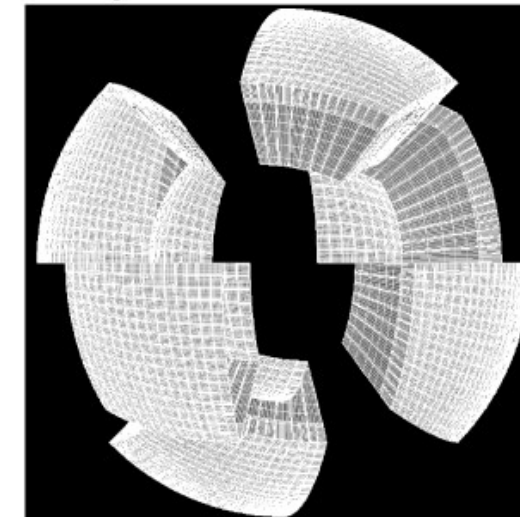
processeur 1



processeur 2



processeur 3



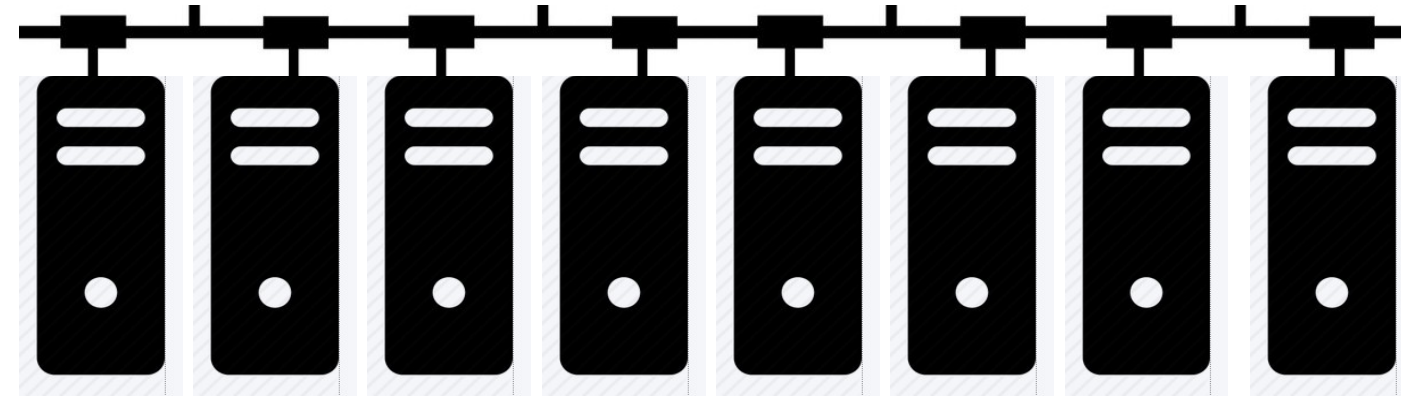
Le calcul distribué

1 ordinateur

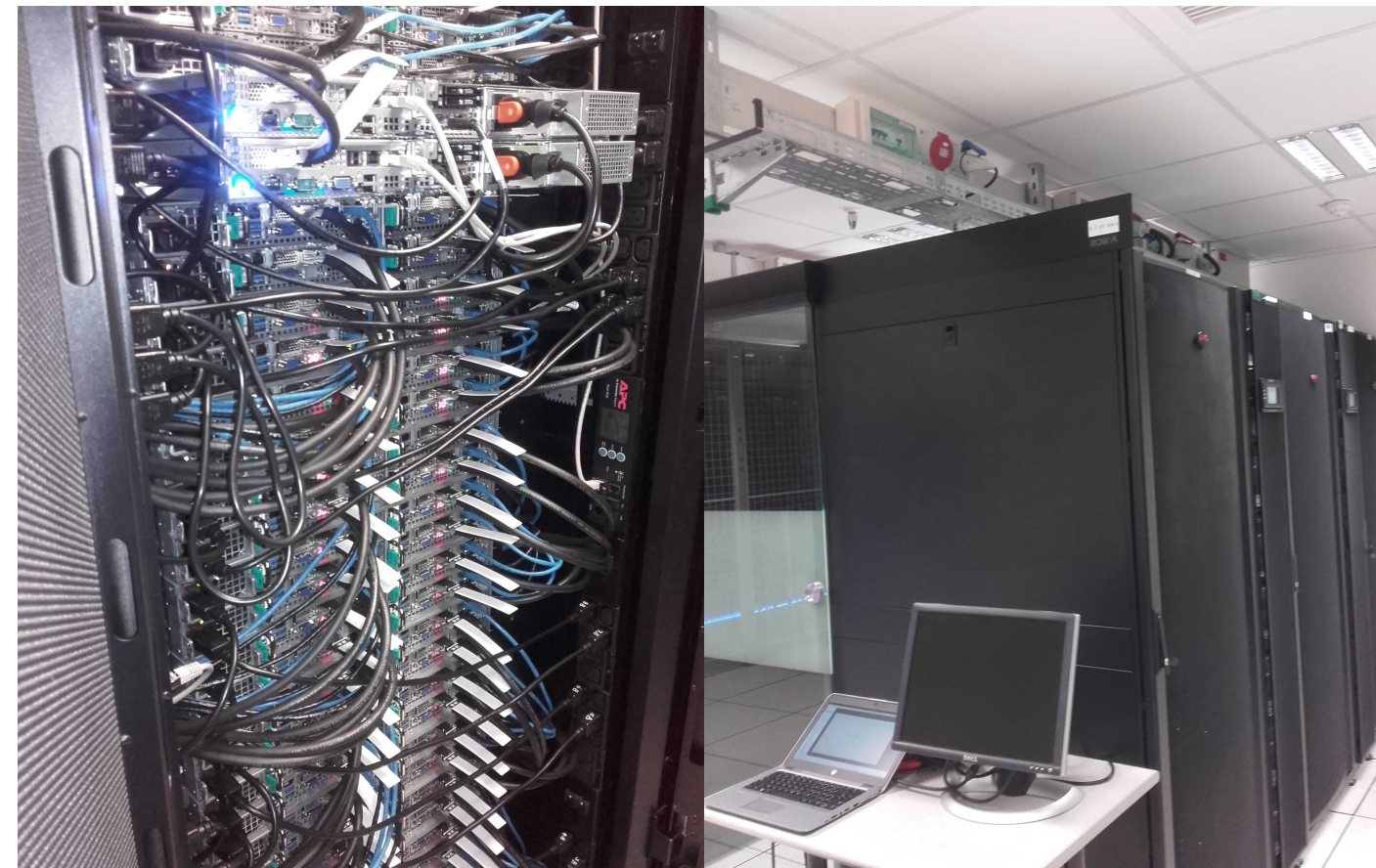


Beowulf (~2000)

Parallélisation : ordinateurs + réseau



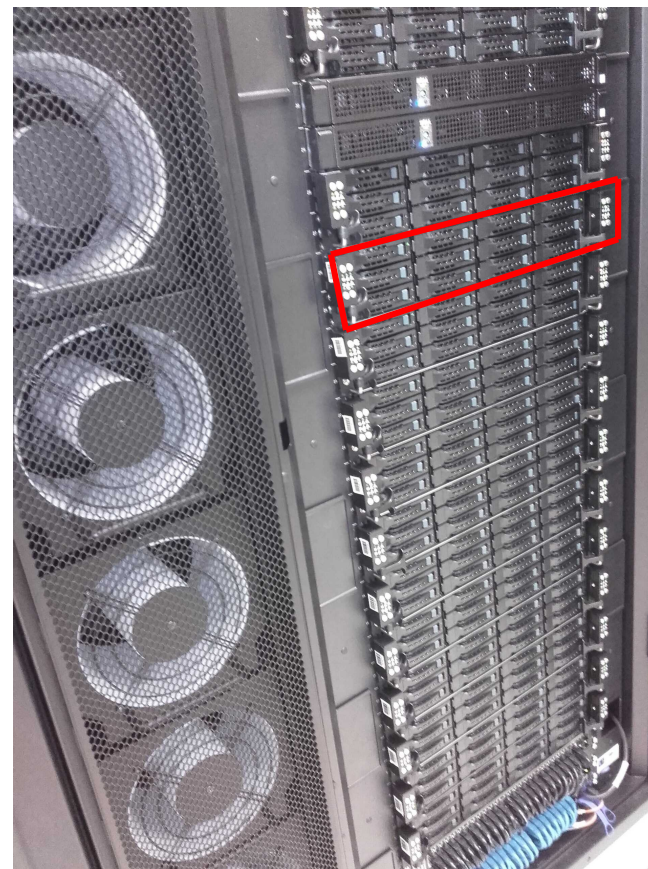
Maintenant (en rack)



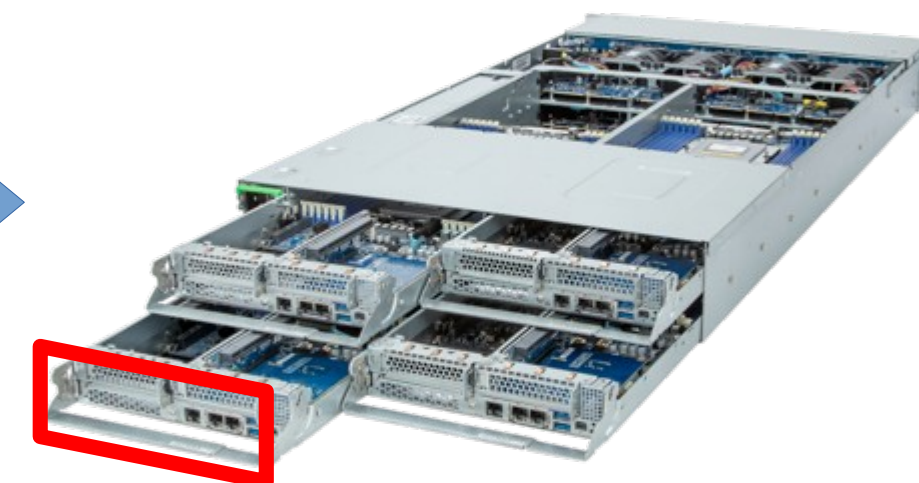
HPC clusters



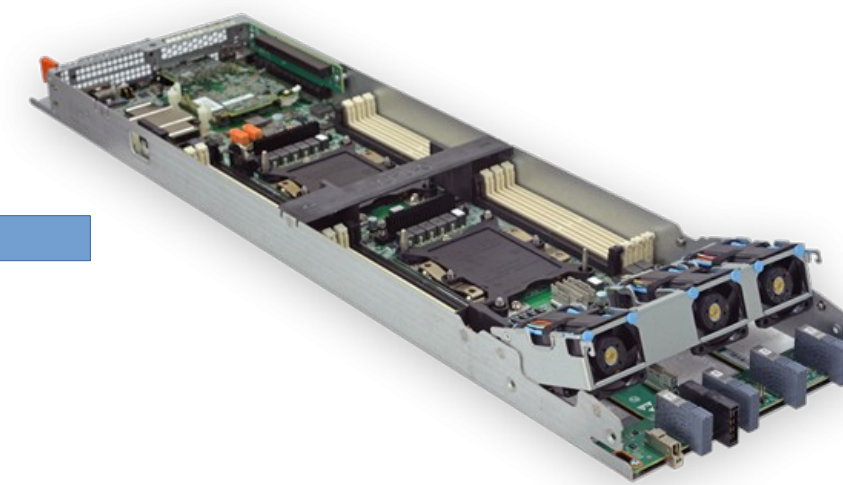
Rack (42U)



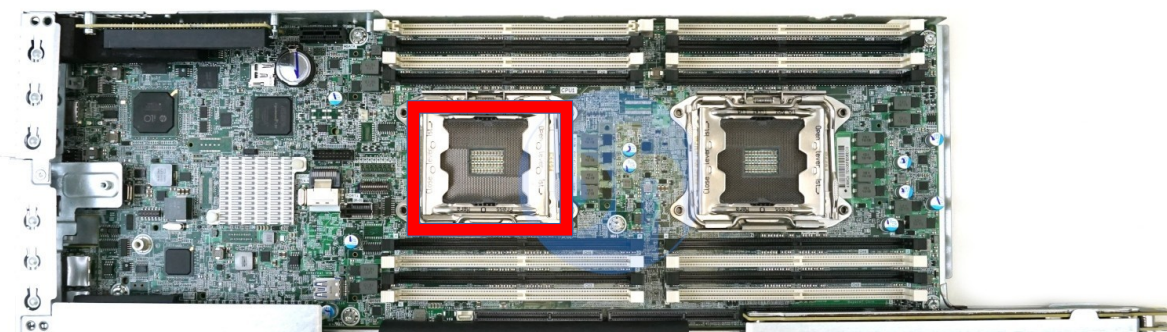
Server (2U, 4 nodes)



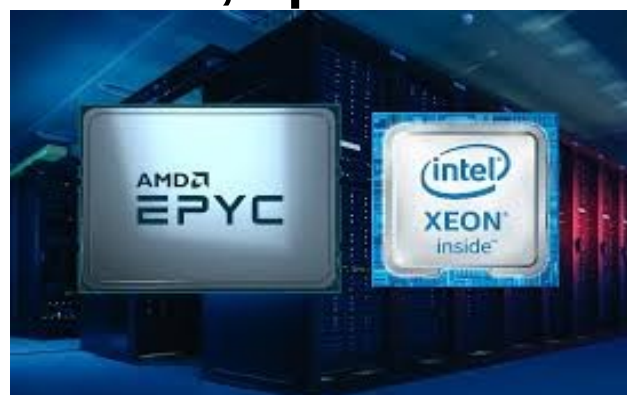
1 node (2 CPU)



Motherboard, 2 CPU



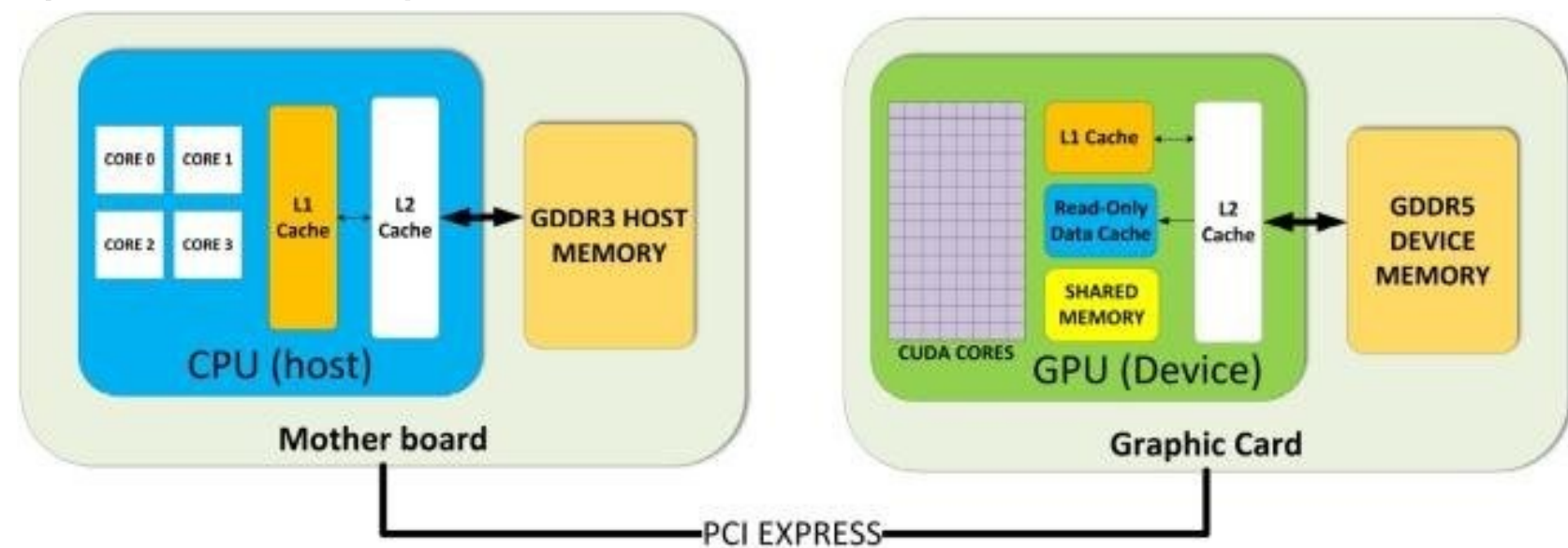
One CPU, up to 128 cores



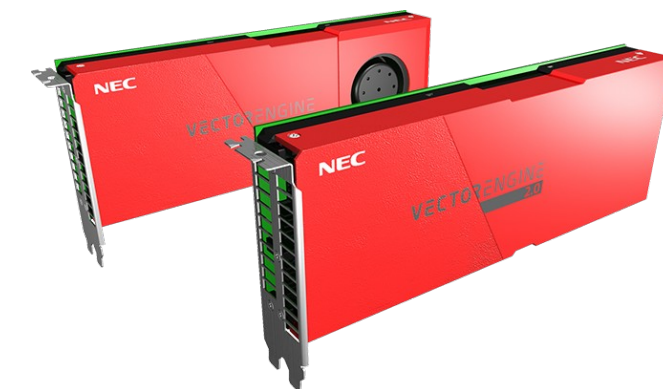
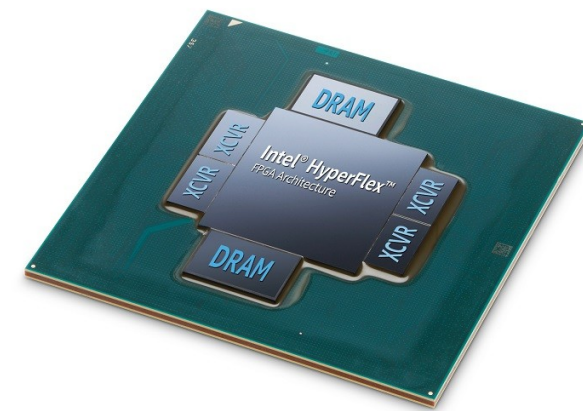
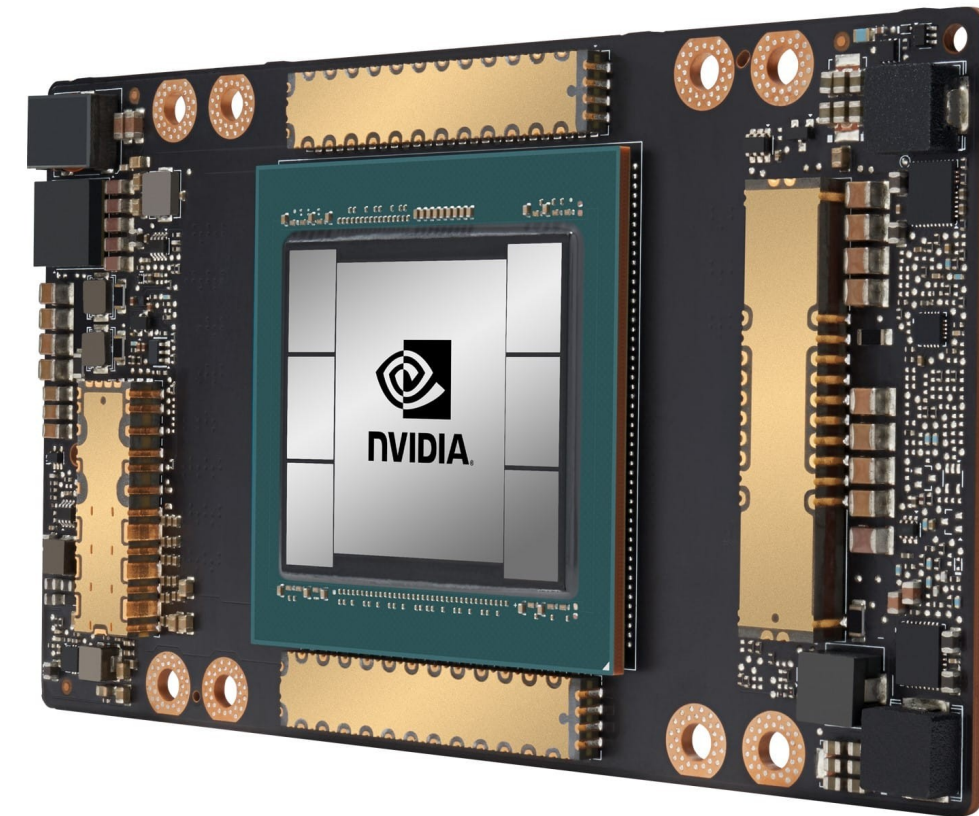
GPU et Accélérateurs

Graphic Processors Unit :

- Multicœur (MI250 and A100 : ~ 10000 cœurs)
- Frequence plus lente (~1 GHz)
- Mémoire partagée plus grande (plus de 140Go pour un H200)



Mais aussi : FPGA, Vector Unit (NEC)

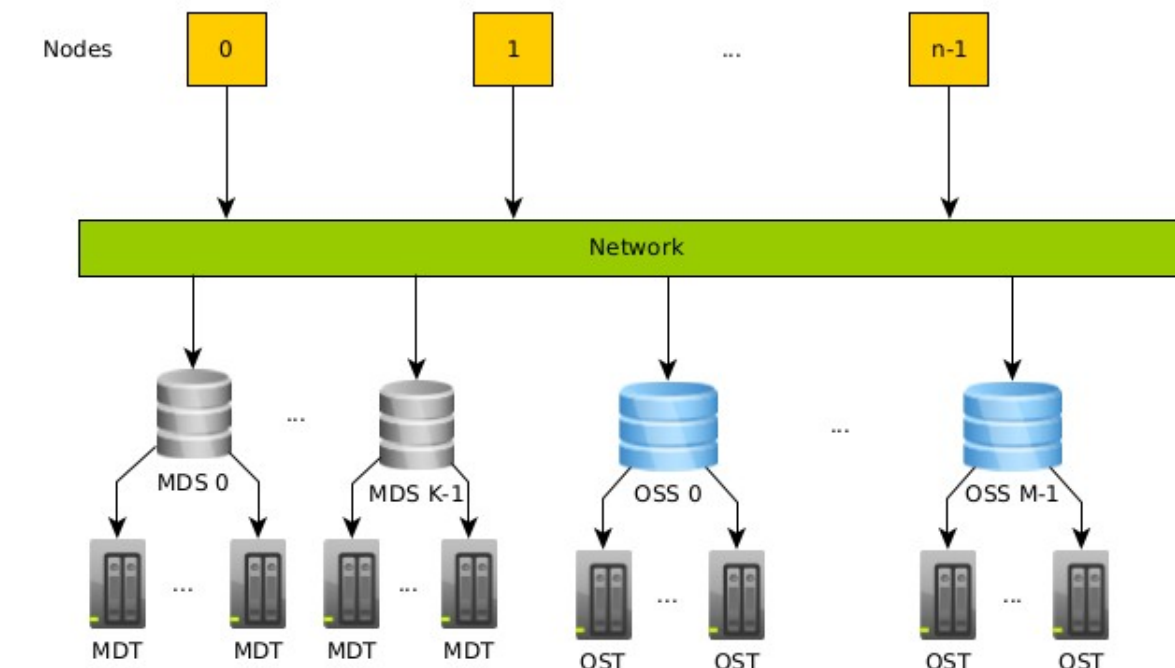
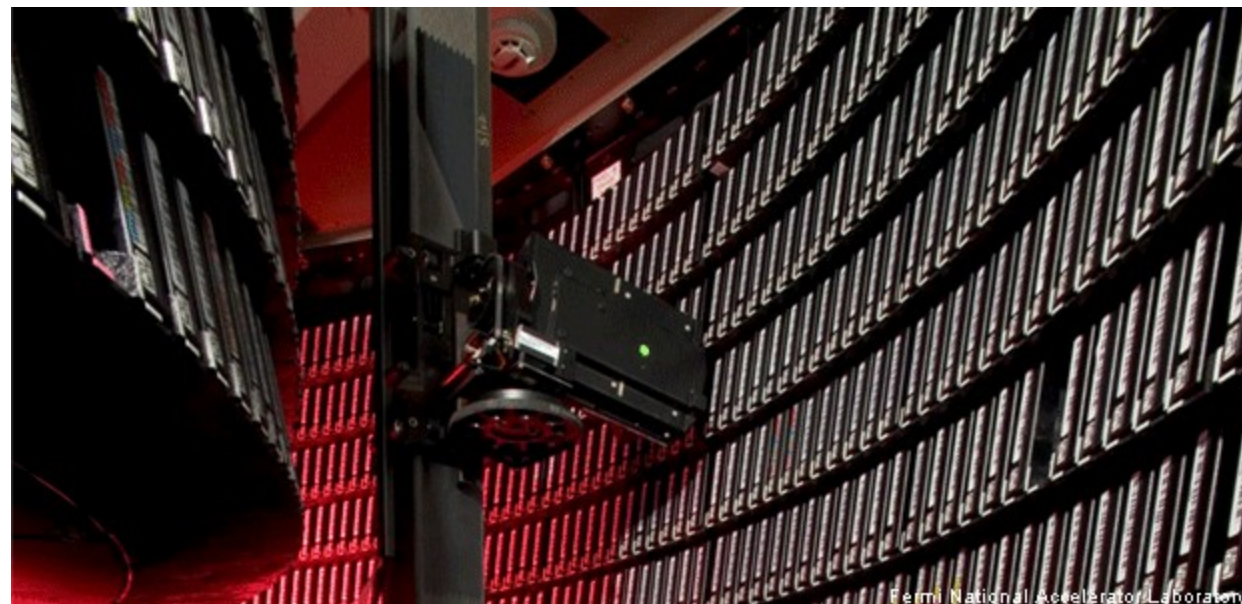
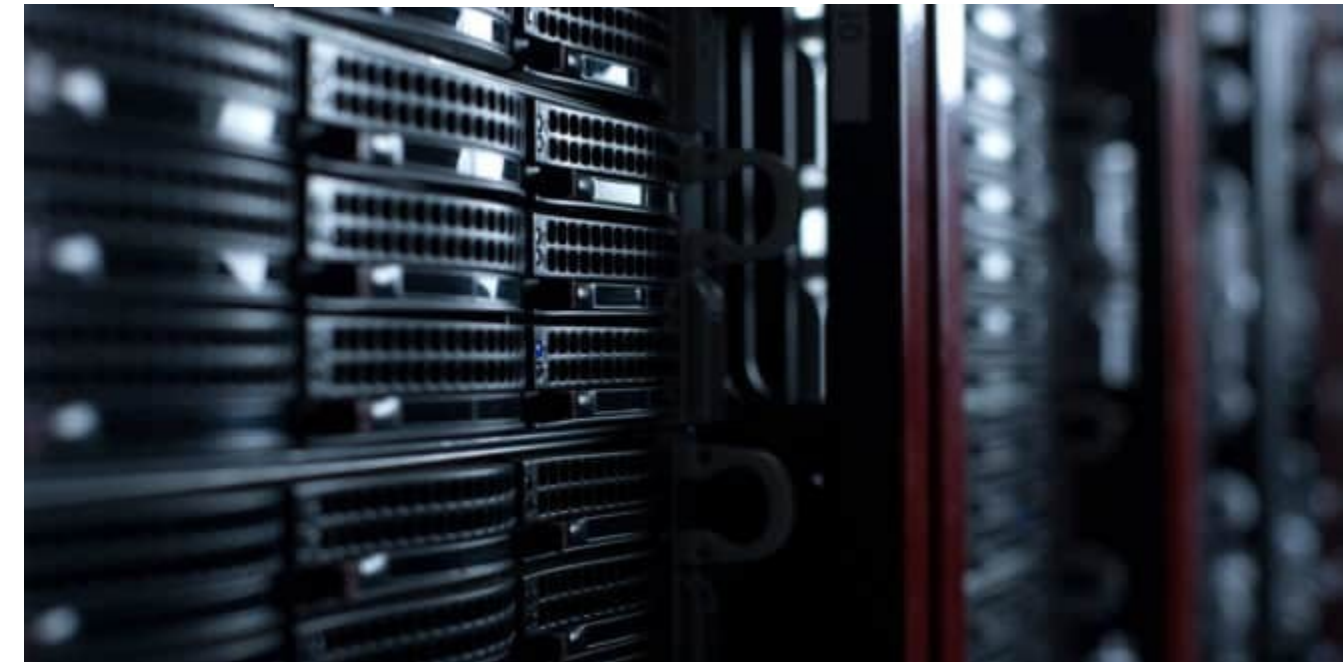


Stockage des données

Hot, Warm, Cold



- Hot: Disque dur interne du nœud : invisible aux autres nœuds
- Hot: stockage storage (RAID): typiquement pour le /home
- Hot: stockage parallèle (Lustre, GPFS, BeeGFS, ...) : pour le /scratch
- Hot / Warm: stockage parallèle moins contraint (CEPH, NetApp ...): pour le /data
- Cold : stockage à bandes magnétiques



Refroidissement et puissance

Consommation électriques des éléments dans le HPC :

- CPU:
 - Intel Sapphire rapid: 350W-420W
 - AMD Genoa: 260W-400W
- GPU
 - GPU A100-80Go: 400W
 - GPU AMD MI250: 500W
- Nœuds:
 - Un nœud AMD Genoa avec 4 GPU A100: 2700W
- Un Rack HPC: 25 to 60kW
- Configurations complètes:
 - Un mésocentre tier2: 100 to 500kW
 - Un centre national: 1 to 5MW
 - Un centre européen: up to 20MW

➔ **Une grosse puissance électrique qui doit être refroidie**

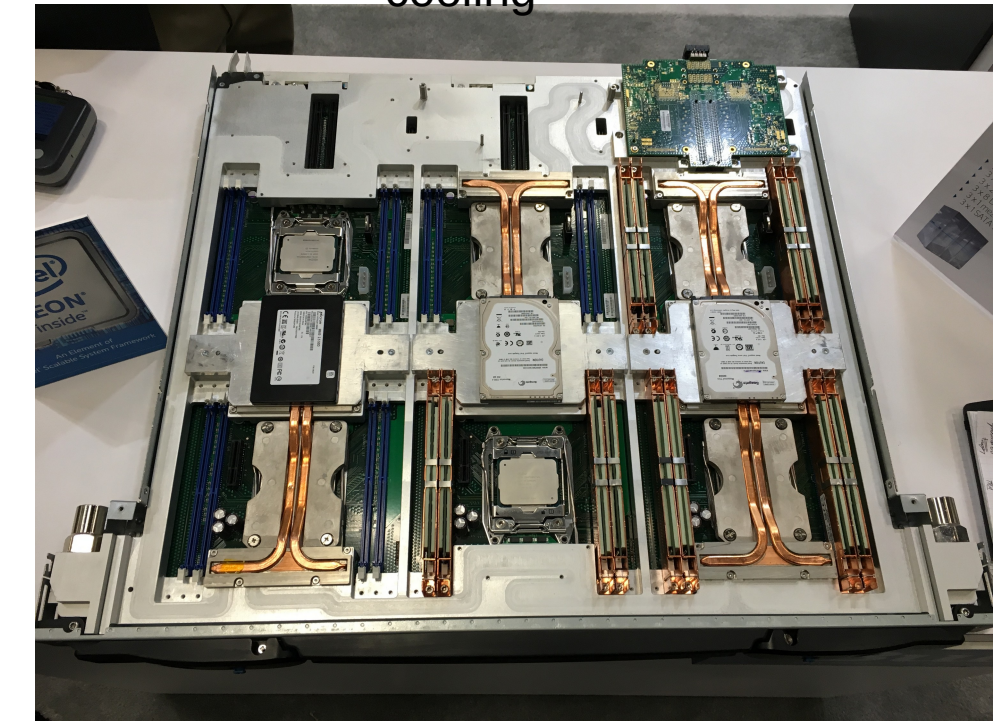
Une maison de 6 kVA: 6kW

Une ville de 70 000 personnes: 18 MW

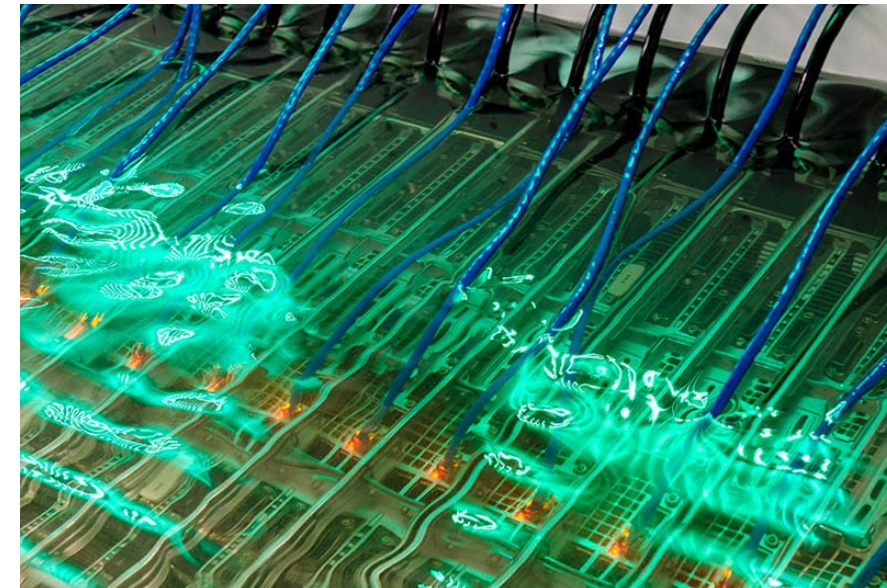
Air cooling



Water cooling



Oil bathing cooling



DC AC :



Quelques banalités sur le calcul scientifique

Le calcul scientifique concerne :

- Historiquement, principalement les domaines comme la Physique, la Mécanique, les Géosciences, l'Astrophysique, la Chimie...
- Depuis une dizaine d'années, on constate de nouveaux acteurs comme la Biologie, l'Économie, les Sciences Sociales, les Langues... toutes les communautés scientifiques sont concernées

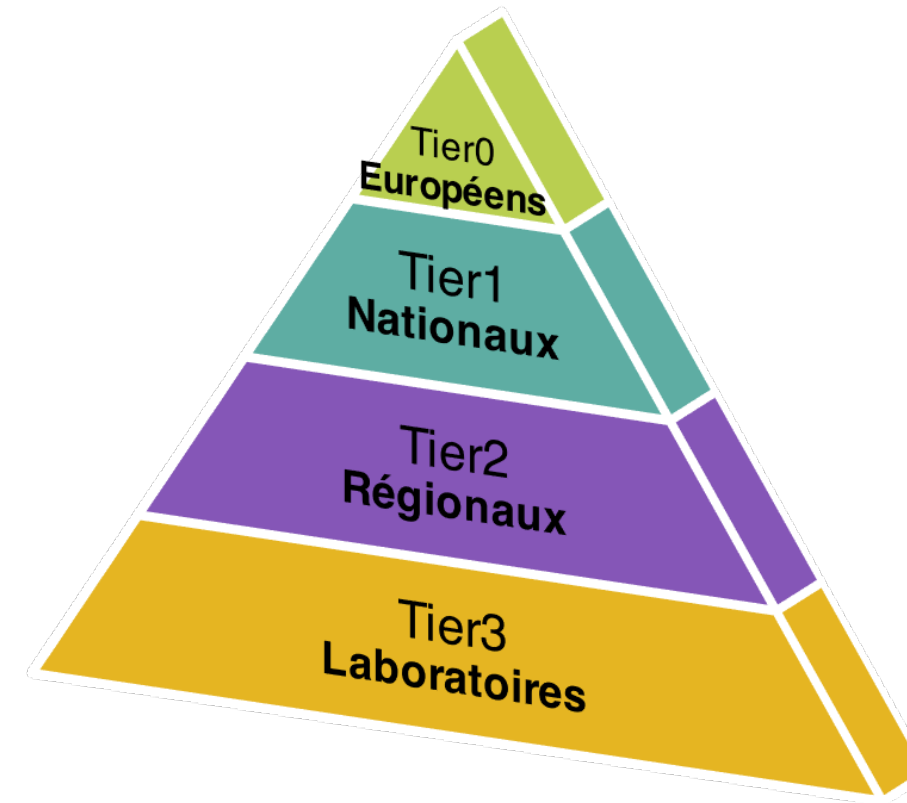
Les besoins explosent :

- en calcul distribué classique (CPU, cœurs)
- en calcul avec accélérateurs (GPU)
- en stockage de données associées

→ Il y a eu beaucoup de changements dans le domaine du HPC, en particulier sur le public concerné

Le calcul scientifique en Europe

- Organisé en Tier
- Quelques exemples
 - Tier0 : Leonardo (Italie) ~ 110 000 cœurs, 14000 GPU A100
 - Tier1 : IDRIS ~85 000 cœurs ; ~ 3000 GPU
 - Tier2 : CALMIP (Toulouse) 13000 cœurs, ~60 GPU
- Modèle économique dominant: accès sur projet (gratuit pour les utilisateurs) ; quelques Tier2 payants à l'heure (<10 %)
- Missions des Tier2 : développement, petite production, formation, mutualisation, proximité, flexibilité



Calcul scientifique : état des lieux en Pays de la Loire en 2020

5 acteurs principaux :

- CCIPL (UN)
- BiRD (INSERM, UN, Biologie)
- ICI (ECN)
- MathStic (UA)
- INFRALAB (LMU)



	Utilisateurs actifs	CPU cœurs	GPU	CPU.h (10 ⁶ h)
CCIPL	155	5500	27	28.4
BiRD	142	450	9	1.5
ICI	130	6300	18	19
UA	30	1000	10	8.4
LMU	~50	1900	90	1

Contexte: labellisation INFRANUM et DACAS

- **Origine** : volonté du MESRI depuis 2016 de labelliser **un** datacenter par Région pour les équipements ESR
- **Objectifs** : rationalisation (fermer 90 % des salles informatiques ESR)
- **Contrainte** : l'ensemble des projets d'équipements financés ou co-financés par des fonds publiques devront être hébergés dans ces datacenters mutualisés (issus de tout projet - PIA, ANR, FUI, Horizon, ...)
- **Réponse** pour les Pays de La Loire : le projet **DACAS-RRTHD (Datacentre et Calcul Scientifique)**
- Le DC UN **et** le projet DACAS ont été labellisés INFRANUM le 10/11/2020

DACAS-RRTHD

Projet soumis au CPER 2021-2027

Porteurs initiaux : Yann Capdeville, Nicolas Wendling / Stéphane Amiard et Thierry Oger

3 volets :

- Datacenter (DACAS, budget 10M€)
- Calcul scientifique (GLiCID 6M€)
- Réseau (RRTHD, 4M€)

Partenaires:

- Datacenter et Réseau : Universités de Nantes, du Mans et d'Angers ; création du Service Inter Établissement Numérique (SIEN) voté par les 3 CA le 4/11/2021
- GLiCID : Universités de Nantes, du Mans et d'Angers et l'École Centrale

L'ensemble du projet a été financé par le CPER pour un montant total de 20M€

L'ensemble des équipements de GLiCID seront localisés dans le DC DaCaS à sa mise en service (juin 2026)

Projet GLiCID

- Origines :
 - partie « Calcul scientifique » du projet DaCaS
 - Volonté de rationalisation des équipements
 - Volonté des partenaires de mutualisation de moyens
- Objectifs principaux du projet :
 - n'avoir qu'un seul acteur calcul Tier2 en Région
 - avoir plus de ressources et de services pour un coût équivalent
 - Mutualiser les ressources et les RH

Unité d'Appui à la recherche GLiCID

GLiCID en tant qu'Unité d'Appui à la Recherche a été créée le 15/03/2022

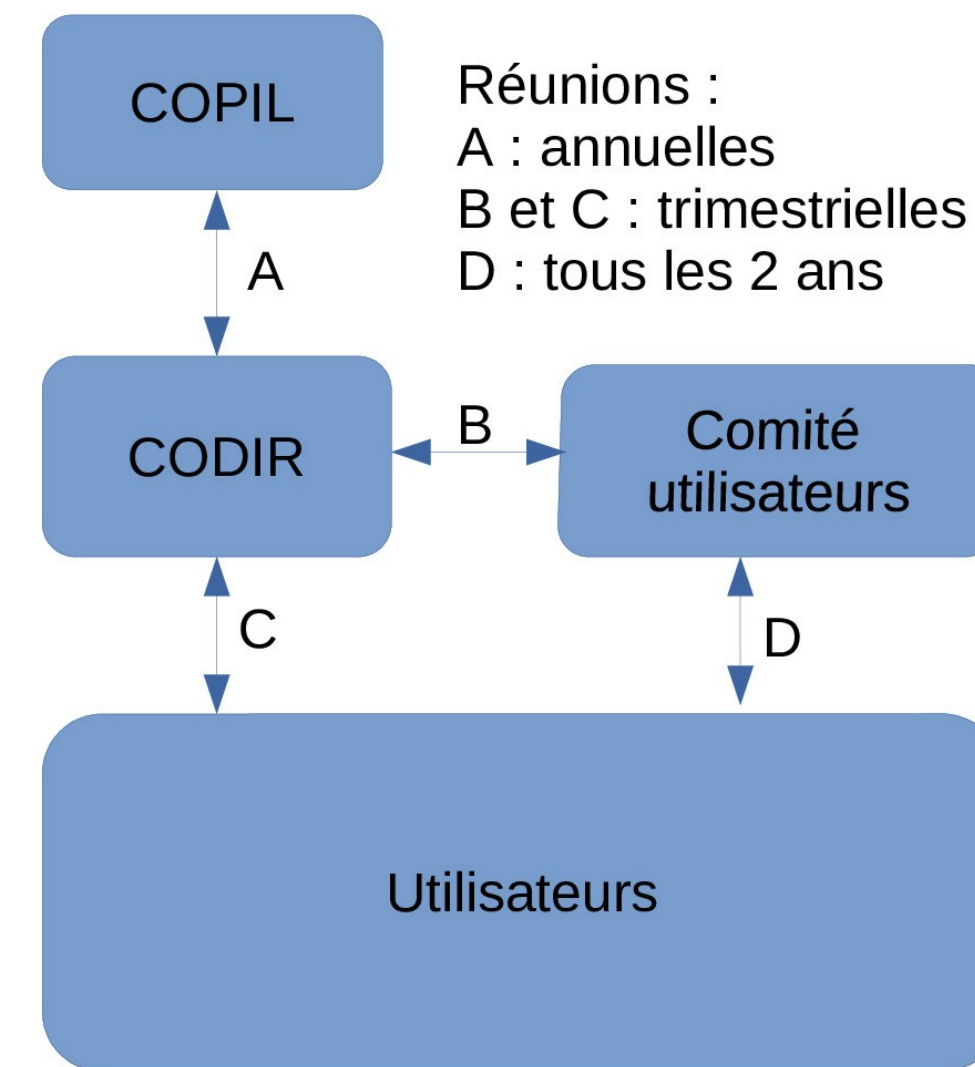
4 tutelles : UA,LMU,ECN,NU

Site web : www.glicid.fr

- Directeur : Yann Capdeville (CNRS-NU)
- Co-directrices.eurs :
 - Luisa Rocha da Silva (ICI, ECN), Audrey Bihouée (UN),
 - Frédéric Saubion (UA), Sylvain Meigner (LMU).

Instances :

- Comité de Pilotage (COPIL)
- Comité de Direction (CODIR)
- Comité des utilisateurs (CU)



Unité d'Appui à la recherche GLiCID

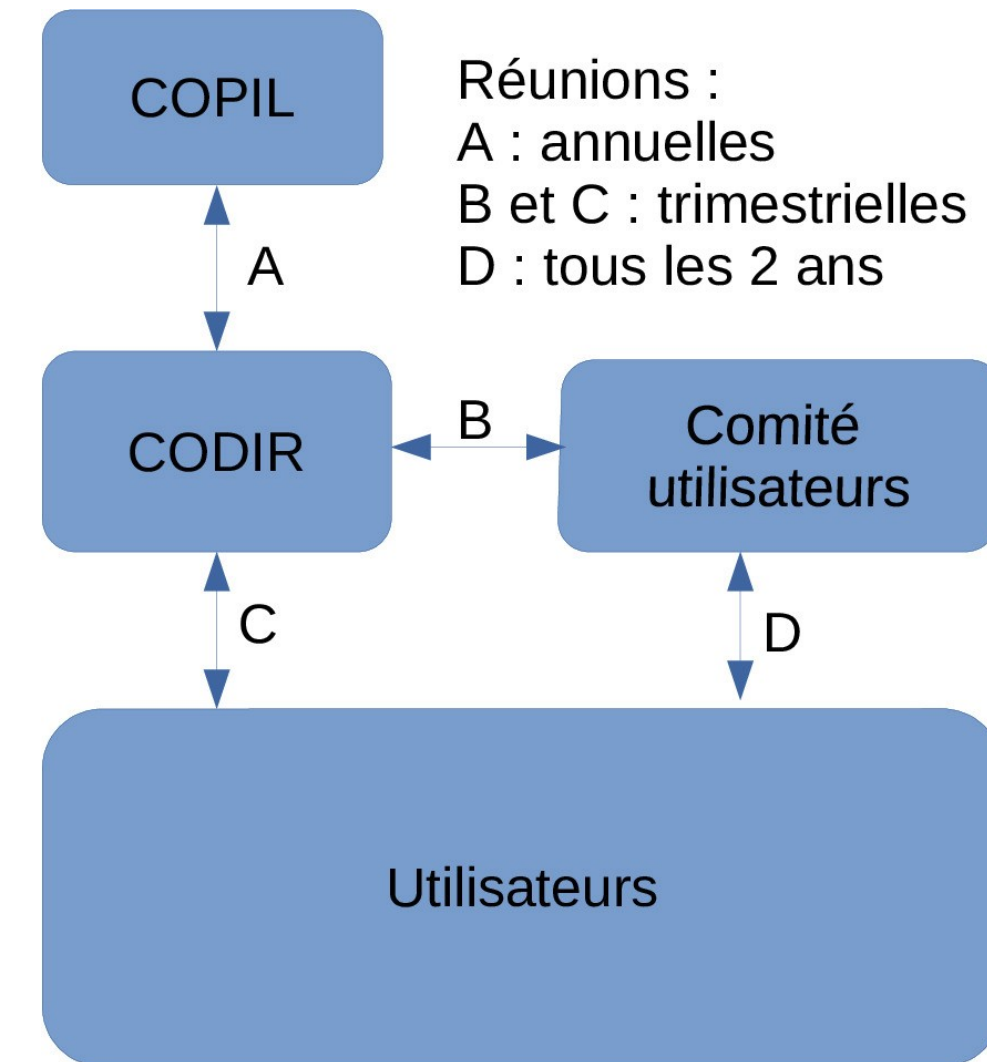
Instances :

- Comité de Pilotage (COPIL) : constitué des VP

Recherche des tutelles, il décide des grandes orientations

- Comité de Direction (CODIR) : constitué de 5 co-dirigeants de GLiCID, il propose des orientations au COPIL

- Comité des utilisateurs (CU) : 30 représentants des Utilisateurs. Il oriente le CODIR dans ses choix, en particulier les choix d'équipements



GliCID : équipe

Pierre-Emmanuel Guérin	IGR, ECN	1 ETP CDI	Ingénieur système et réseau
Pablo Bondia-Luttiau	IGR, ECN	1 ETP CDD	Ingénieur système et réseau
Yann Dupont	IGR, NU	0.8 ETP	Ingénieur système et réseau
Jean-Francois Guillaume	IGE, BiRD-SFR, NU	1 ETP CDI	Ingénieur système et réseau
Guy Moebs	IGR, CNRS-LPG	0.5 ETP	Ingénieur calcul scientifique
Aymeric Blondel	IGE, CNRS-CEISAM	0.4 ETP	Ingénieur calcul scientifique
Jérôme Coatanéa	IGE, NU	0.1 ETP	Ingénieur système et réseau
Jean-Christian Feufeu	IGE, NU	0.1 ETP	Ingénieur système et réseau
Damien Fligiel	IGE, CNRS-OSUNA, NU	0.1 ETP	Ingénieur système et réseau
A recruter	IGR, MesoNET, ECN	1 ETP CDD	Ingénieur calcul scientifique
Hugues Digonnet	MdC, ECN	0.2 ETP	calcul scientifique
à recruter	IGR, UA	0.7 ETP	Ingénieur système et réseau
Julien Drouin	IGR, LMU	0.5 ETP	Ingénieur système et réseau
Benoit Seignovert	IGR, CNRS, ONSUA	0.1 ETP	Ingénieur Data
Corine Varipatis	Gestionnaire, ECN	0.5 ETP	Gestionnaire

Unité de service GLiCID

Missions de GLiCID :

- Fournir des moyens de calcul aux chercheurs de la Région et leurs partenaires sans se limiter aux « gros » calculs
- Proposer un volume de stockage de données associées au calcul suffisant
- Apporter un support au développement et une veille technologique
- Proposer des formations au HPC et aux outils s'y rapportant
- Garder une forte proximité avec les utilisateurs
- Participer aux projets nationaux

GliCID : Le modèle économique

Objectifs : garder le modèle accessible à tous tout en finançant au moins les postes GLiCID sur fonds propres.

- Heures de calculs
 - Au moins 60 % des ressources sur GLiCID accessible (sur projet) sans contre partie financière
 - Une partie des ressources achetée par les projets financés de recherche sous forme d'heures prioritaires
 - Une partie des ressources vendue aux établissements externes et au privé (max 20%)
- Stockage de données :
 - Pour les ayants droits : 3To de base sans contre partie financière
 - Payant au To par année au-delà. Tarifs différents entre les ayants droits, le privé etc.

Il est important pour la pérennité de GLiCID que ses utilisateurs prévoient des financements dans leurs projets pour les ressources de calcul et de stockage.

GliCID : Les ressources

Calcul HPC :

Cluster	Nœuds	Total cœurs	Total RAM	GPU	Réseau	• Refroidissement
Nautilus (2023)	AMD Genoa	5376	28To	16 A100-80G 8 A40	Infiniband 100	• eau
Phileas (2023) MesoNET	Intel Sapphire rappids	3072	16To	0	Infiniband 100	• eau
Waves (2016)	Intel Xeon AMD	6712	66To	22 A100-40Go 8 A40 20 T4	OPA 100 RoCE 100	• air
Bird (2016)	Intel et AMD	1102	8To	9 T4	Infiniband	• air

Stockage :

Nom	Type	Volumétrie	Technologie
scratch/Nautilus	Chaud	450To	NetApp
Scratch/Waves	chaud	350To	CEPH-ssd
scratch/BiRD	chaud	437To	BeeGFS
LabDATA	Tiède	5000To	CEPH

Le CPER : équipements et déploiement de GLiCID

- Le financement 6M€ est scindé en 3 tranches de 2M€

La 1ere tranche est est production depuis septembre 2023

La 2eme tranche est conditionnée à l'arrivé du DC DaCaS
Scénario initial

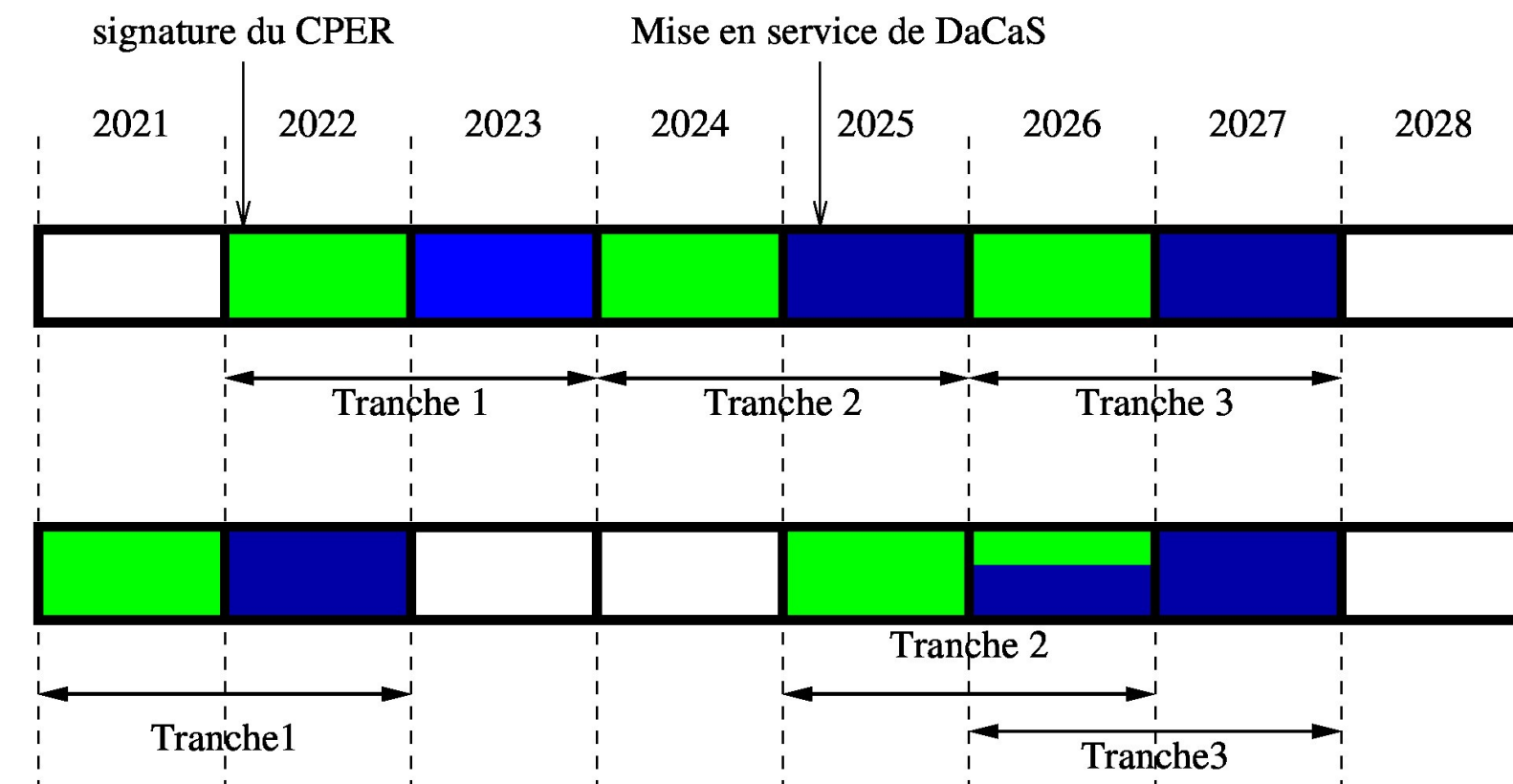
- GLiCID :

- Migration des services, des équipements et des utilisateurs par tranches depuis septembre 2023

- Actuellement 560 utilisateurs ont migré sur GLiCID

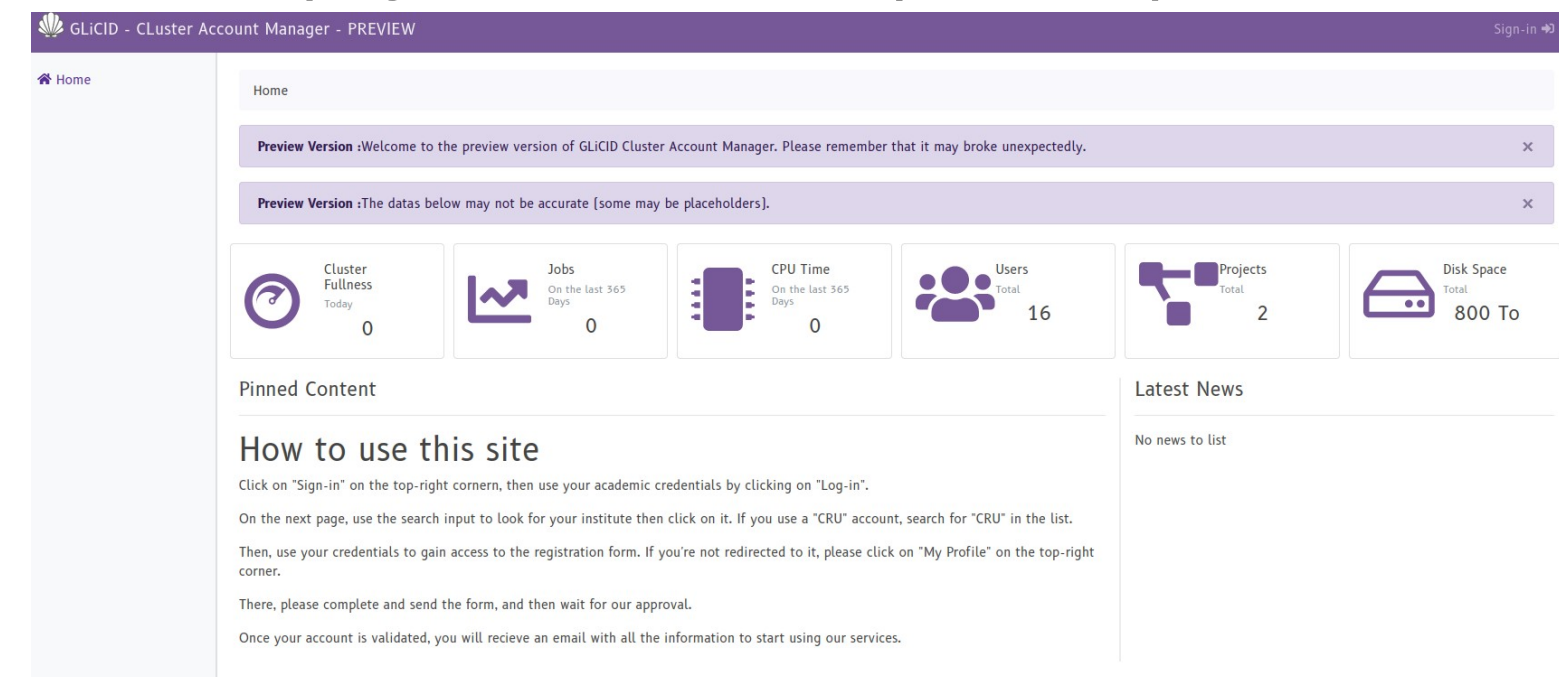
- Quelques mots clés systèmes et technos :

- - **Systeme** : guix, proxmox, blue banquise, redhat8, Zabbix, gitlab ;
- - **Stockage** : ceph, ceph-ssd, beegfs, GPFS, netapp, S3 ;
- - **Réseaux** : switchs Sonic (Rare) ; HPC: ROCE, infiniband, OPA



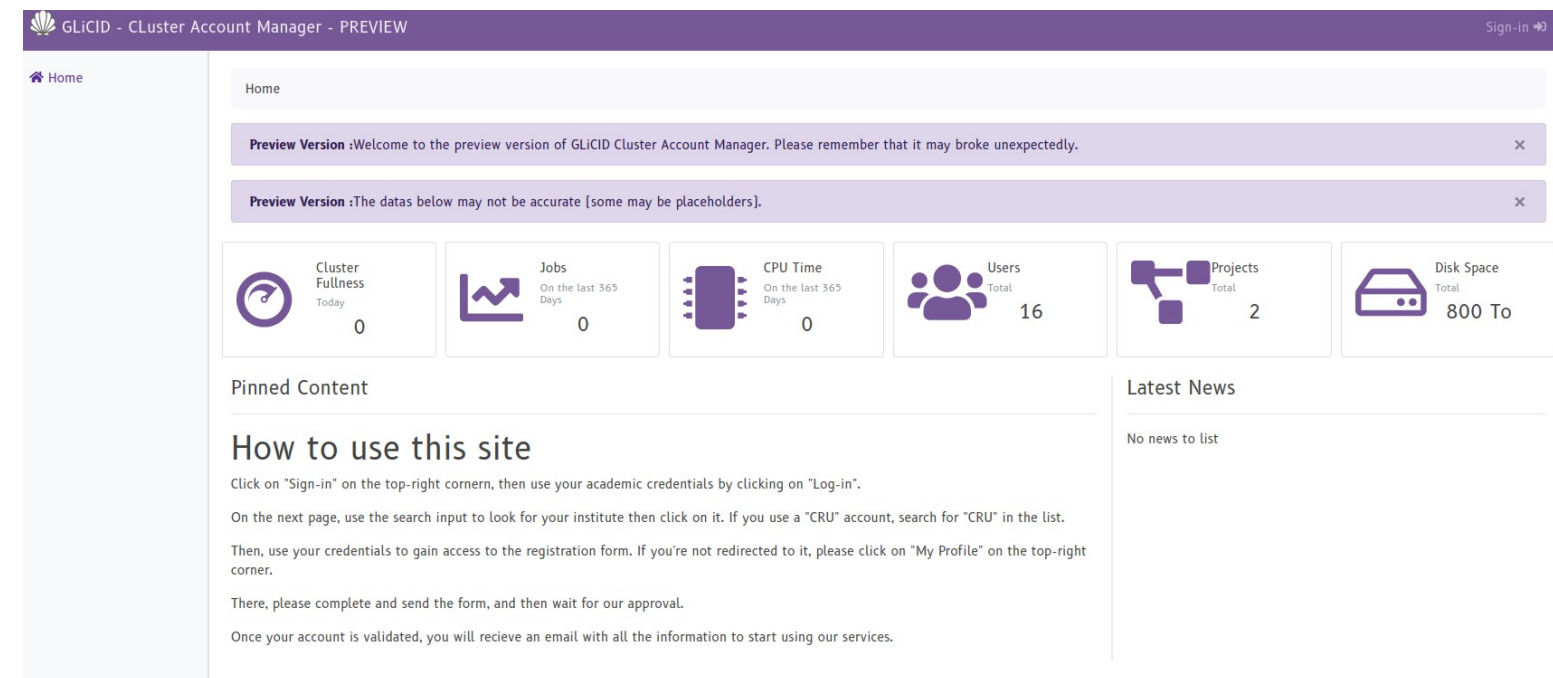
Aspects pratiques : Les services

- Cluster:
 - Un seul Slurm pour l'ensemble des équipements
 - Partition unifiée. Accès aux équipements spécifiques par « constraint » (e.g. --constraint=cpu_amd)
- Création des comptes gérés par l'interface « CLAM » (gestion des demandes, des clés ssh, du stockage, etc.)
- Services en prévision :
 - Marketplace Openstack de machines environnées
 - JupyterHub et JupyterLite
 - Visualisation déportée
 - Grande volumétrie de stockage et backup (payant)
 - IA Generative (en test)



Aspects pratiques :

- Demande de compte en ligne sur CLAM. Validation en quelques jours (normalement ...)
- Création de projets sur CLAM
- Accès par ssh avec clé
- Logiciels via modules ou guix
- Soumissions via slurm, jupyter ou openstack
- Plusieurs espaces de stockage :
 - /home
 - /scratch (plusieurs)
 - /LAB-DATA
- Les utilisateurs doivent :
 - Demander des financements pour GLiCID dans le projets (2cc heures.coeur prioritaire, 54€ To/an pour le stockage)
 - Remercier GLiCID dans leurs publications



Fin

Le CPER : 1ère tranche

- Le lauréat est BULL-ATOS. La mise en production des équipements est prévue en juin 2023
- Équipements prévus :
 - 40 nœuds 2x AMD Genoa 48 cœurs (3.3GHz) 384Go de RAM DDR5
 - 8 nœuds 2x AMD Genoa 48 cœurs (3.3GHz) 768Go de RAM DDR5
 - Interconnexion des nœuds en Infiniband HDR100
 - 4 nœuds GPU avec 4 cartes GPU Nvidia A100-80Go chacun (Genoa 48 cœurs, 768Go de RAM)
 - 4 nœuds de visu 2 cartes GPU Nvidia A40-48Go chacun (Genoa 48 cœurs, 768Go de RAM)
 - Stockage scratch Spectrum Scale 0.5Po
 - Refroidissement à eau
 - 1.5 Po de CEPH (HPE)

Le CPER : 1ère tranche

- Localisation initiale : salle informatique de l'ECN, à la place de Liger. Déménagement prévu vers DaCaS à son ouverture
- Pour mémoire, les autres ressources existantes seront toujours accessibles (~6000cœurs, 22GPU A100, 8A40, 24T4, 3 Po de stockage etc)

Les projets PIA3

Deux projets PIA3 sont en cours :

- MesoNET (porté par l'ECN)
- MUDiS4LS (porté par BiRD)

MesoNET devrait mutualiser entre 2000 et 3000 cœurs Intel Sapphire Rapids dans GLiCID. Le projet permettra aussi l'accès, sur d'autres sites, pour les utilisateurs de GLiCID

- à 3 architectures prototype : ARM, vectorielle, quantique
- à des ressources plus classiques pour la formation (CPU/GPU/OpenStack)

MUDiS4LS devrait mutualiser 8 nœuds pour Openstack (64C, 512Go RAM), 4 nœuds « fat » (4To RAM) et aussi des nœuds pour du HPC.

Projet GLiCID

Dans le cadre du CPER 2021-2027

- GLiCID est la partie calcul HPC du projet DACAS+réseau
- Établissements : UA, LMU, NU, Centrale Nantes, Inserm
- Porteur : Yann Capdeville
- Responsables par partenaire :
 - Yann Capdeville (CCIPL)
 - Luisa Rocha da Silva (ICI, ECN)
 - Audrey Bihouée (BirD/SFR F.Bonamy, INSERM, UN)
 - Frédéric Saubion (UA)
 - Sylvain Meigner (LMU).
- Budget et équipements demandés : 6M€ (pour >12000 cœurs, >100 GPU type A100, >8Peta octets de stockage)