

Comité Technique HPC

Comité technique HPC

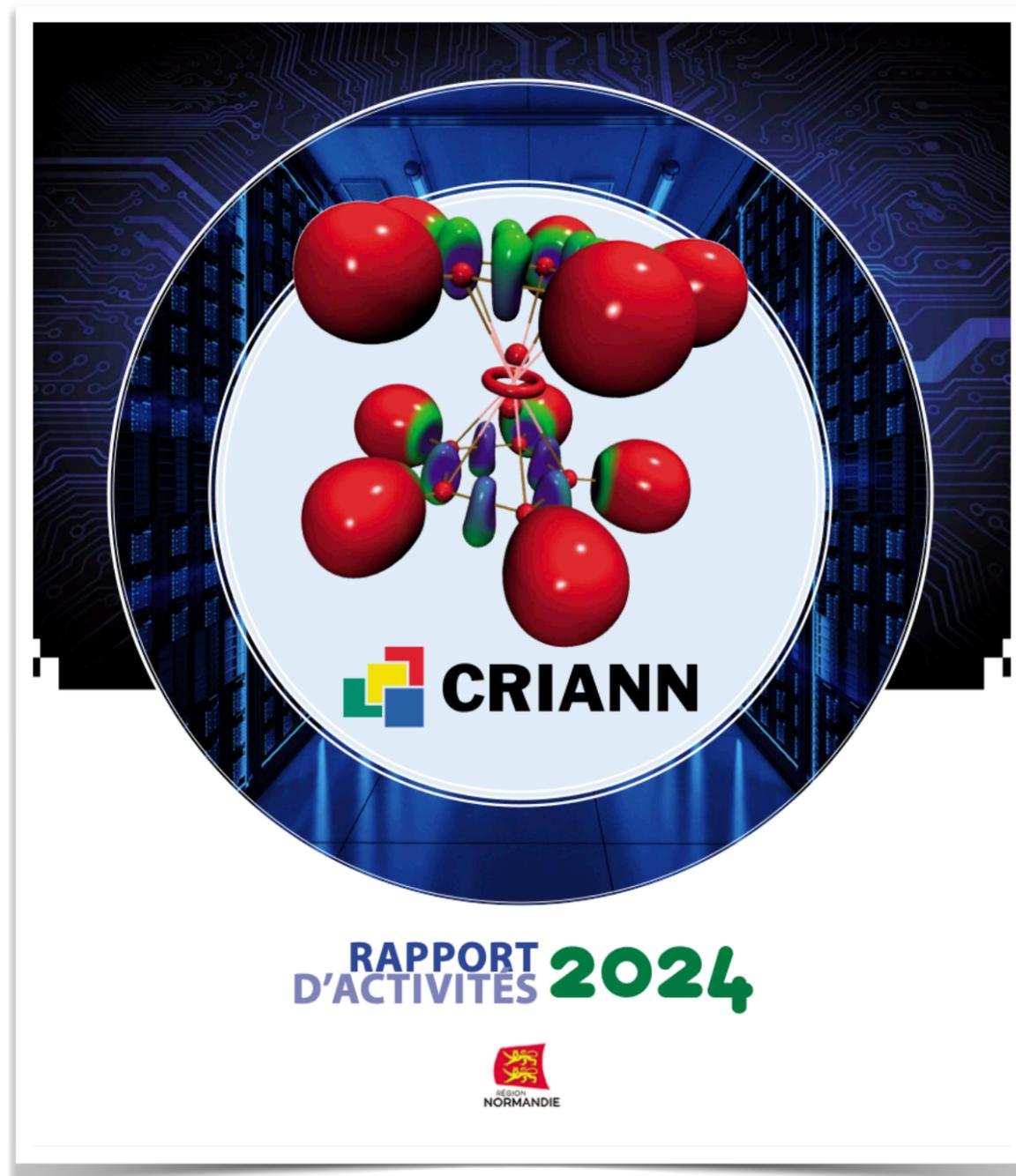
Ordre du jour

- RA 2024 / Attribution / production d'heures
- Bilan d'exploitation
- Focus sur les suites de la mise à jour de février
- Arrivée des nœuds H200 (MesoNET)
- Agenda
- Questions / Réponses

Attribution / production d'heures

Rapport d'activité 2024

RA 2024

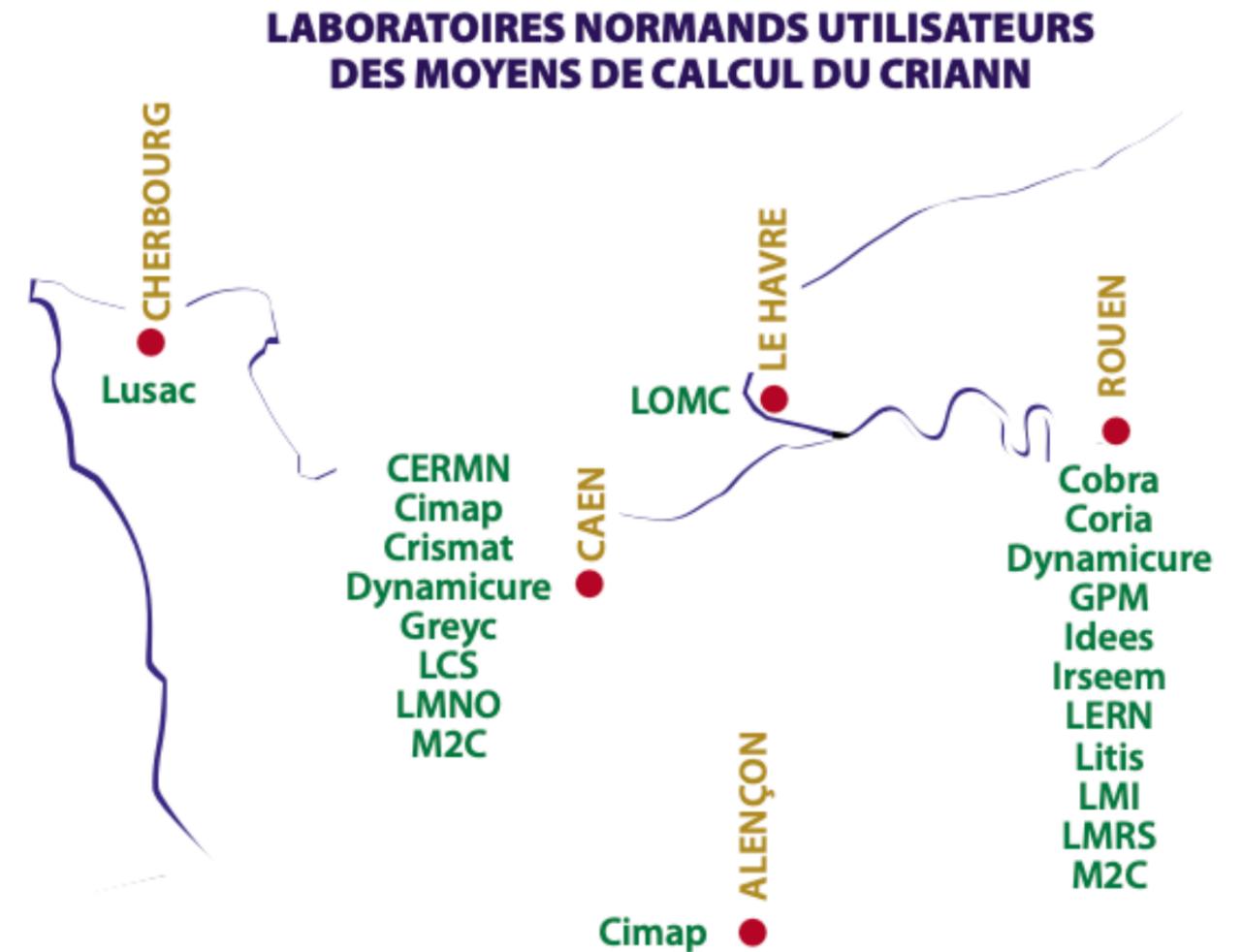
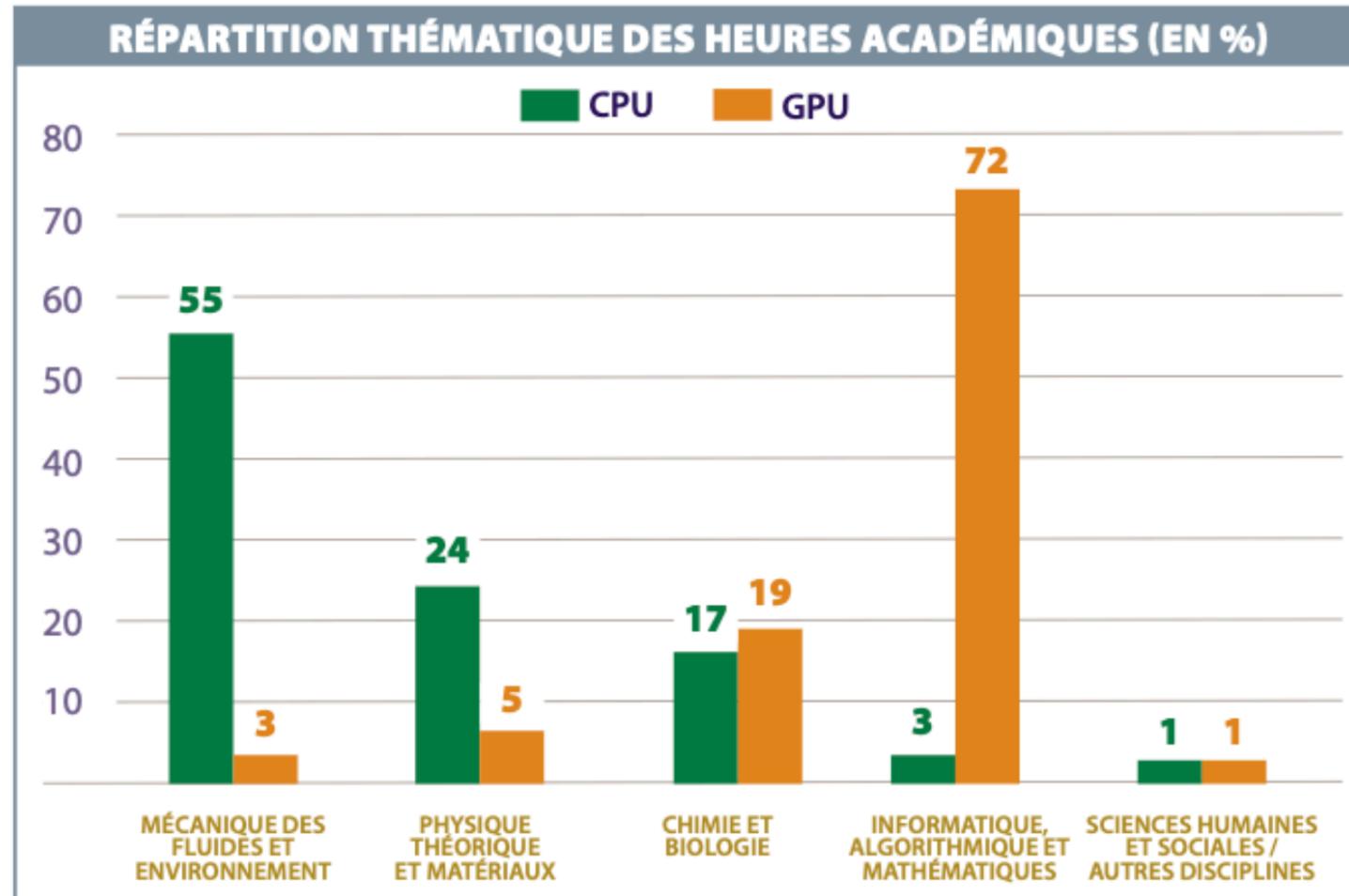


- Disponible en ligne sur le site du CRIANN (www.criann.fr)
 - Rubrique Présentation → Documents → synthèse
- Chiffres clés 2024
 - 101 M.h.cœur CPU et 325 kh.GPU
 - 260 utilisateurs actifs
 - 104 projets (chercheurs et entreprises)
 - 31 laboratoires utilisateurs

<https://www.criann.fr/docs/0/ra/CRIANN-RA-2024.pdf>

Rapport d'activité 2024

Laboratoires utilisateurs en Normandie



Appel à projets scientifiques 2025

Bilan des attributions et des consommations

- Attributions du 1^{er} AAP 2025 (avec quelques projets au fil de l'eau du 1^{er} trimestre)
 - 71 M.h.c attribuées
 - 98 projets scientifiques
- Production académique sur le 1^{er} trimestre :
 - 19,4 M.h.c (et 93,5 k.h.GPU)
 - Soit 27 % des attributions
- Production des industriels : 5,6 M.h.c

Valorisation des heures de calcul

Publications

- Penser à mentionner l'utilisation des moyens de calcul Criann dans les publications
 - *Ce travail a bénéficié des moyens de calcul du mésocentre CRIANN (Centre Régional Informatique et d'Applications Numériques de Normandie).*
 - *Part of this work / The present work / was performed using computing resources of CRIANN (Normandy, France)*

Austral Volet exploitation

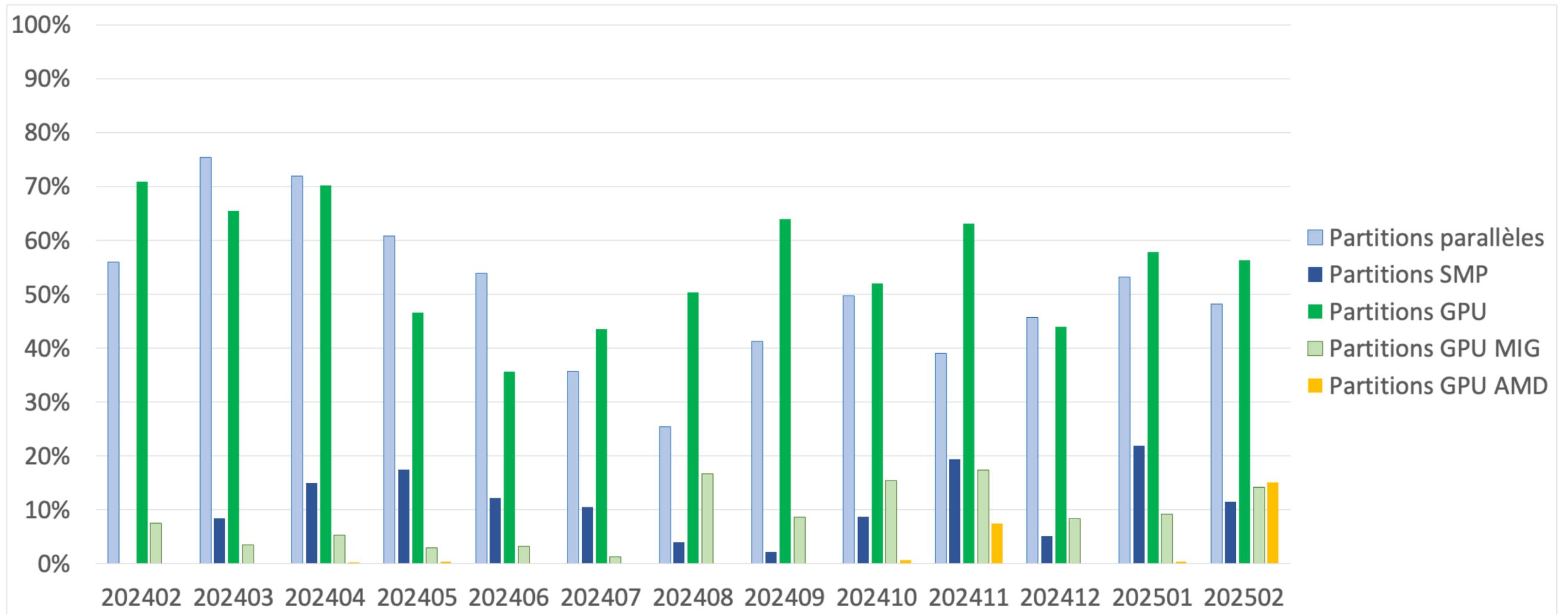
Arrêts de production / incidents

Depuis le dernier CT (08 octobre 2024)

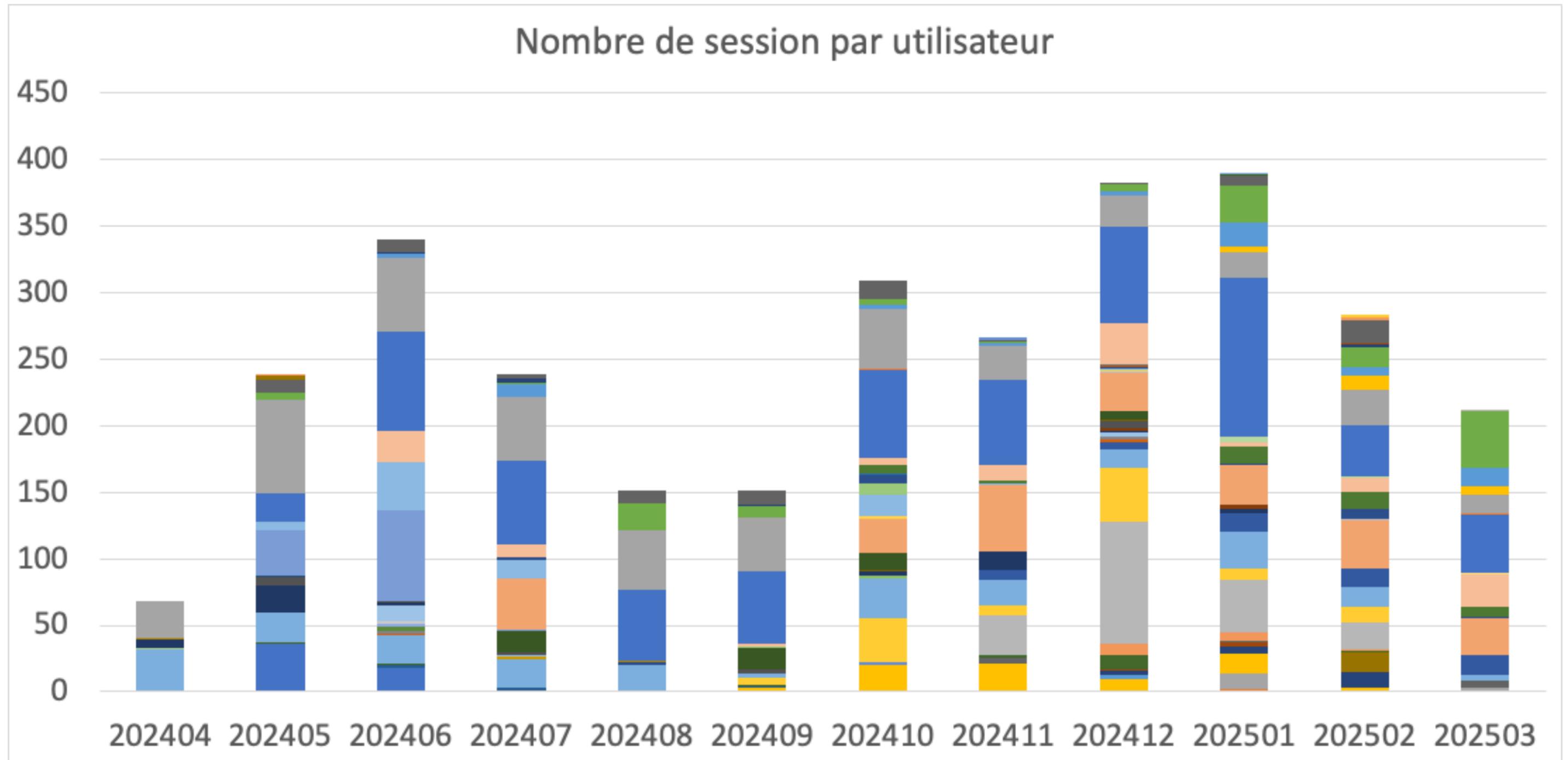
- Arrêt de production planifié
 - Mise à jour du 10 au 17 février 2025
 - 8 jours comme initialement prévu
- Arrêt pour incident
 - 6 janvier 2025 (1,5 jour)
 - Coupure électrique d'une durée de quelques minutes le dimanche matin. Accès aux frontales et aux données maintenu, retour en production des nœuds de calcul le lundi en fin de matinée.

Charge d'occupation d'Austral

Pourcentage d'occupation des ressources (hors jours de maintenance)
CPU (partitions parallèles et smp) et GPU (partitions GPU)
GPU (partitions GPU MIG et GPU AMD)



Sessions Jupyter - utilisation mensuelle



Actions mises en place suite au dernier CT

Partitions GPU

- Test d'une partition `gpu_smt` pour évaluer l'impact de l'activation du multi-threading
 - ok, cf. MaJ de février
- Activation du mode `semaine/week-end` pour les partitions GPU (`gpu`, `gpu_all`, `hpda`)
 - ok, cf MaJ de février

Focus technique :
suites de la Mise à Jour

Mise à jour de février 2025

Bilan

- Durée prévue de l'arrêt de service : 8j
- Durée effective
 - pour les serveurs Fin + GPU/HPDA + VISU : 8j
 - pour le serveur LARGE : 9j
 - pour les serveurs Alt : 8j mais environnement non finalisé pour 1 code

Mise à jour de février 2025

Éléments mis à jour

- Système d'exploitation : RHEL 8.8 -> RHEL 8.10
- Brique de gestion du cluster : HPCM 1.12
- Slurm 23.02 -> 24.05
- Drivers NVIDIA : 560.35.03 -> 565.57.01
 - versions CUDA compatibles jusqu'à 12.7
- Environnement de compilation CPE : ajout de la version 24.07

Mise à jour de février 2025

Partitions GPU

- Multi-threading généralisé sur tous les nœuds GPU Nvidia
 - Partition `gpu_smt` supprimée
- Mode semaine/week-end pour les partitions GPU (`gpu`, `gpu_all`, `hpda`)
 - Valeurs limites à retenir :
 - jusqu'à 16 coeurs par GPU
 - limite en semaine : `gpu=16` et `cœurs=256`
 - limite le week-end : `gpu=32` et `cœurs=512`

Mise à jour de février 2025

Service MLDE

- Service MLDE pour l'IA non fonctionnel suite à la MaJ
 - Arrêt du produit annoncé par HPE pour 01/26
 - Arrêt du service sur Austral
- Maintien du service JupyterHUB
 - cf. nombre de sessions / utilisateurs

Extension GPU H200

Extension GPU H200

Contexte

- Acquisition dans le cadre de MesoNET
 - Ressources « code-formation » du projet
 - Développement d'applications
 - Formation (de niveau Master)
 - Choix d'une intégration à Austral
 - Modalités en cours de définition / finalisation
 - Envisagé :
 - H200 accessibles à tous les utilisateurs d'Austral
 - une partie plus large d'Austral (appelée Arctic) accessible via MesoNET pour le volet « code-formation »
 - Portail de demandes de ressources MesoNET : [gramc-meso](https://gramc-meso.org)

Serveurs à GPU H200

Architecture

- 2 nœuds de calcul dotés chacun de :
 - bi-sockets 48 cœurs - Xeon 8558 (96 cœurs physiques, 192 avec SMT), 2.1 GHz
 - 8 GPUs NVIDIA H200-SXM, 141 GB de mémoire globale par GPU
 - Interconnexion NVLink
 - 2 TB de RAM DDR5
 - 4 cartes d'interface réseau Slingshot (200 Gb/s par carte)
 - 15 TB de disques internes
 - potentiellement disponibles pour les applications (usage en lecture envisagé)

Performances : théorique

GPU H200 vs A100

Modèle	A100-SXM-80GB	H200-SMX-141GB	Facteur
TFlops FP64	9,7	34	3,5
TFlops FP64 Tensor core / FP32	19,5	67	3,4
TFlops FP32 Tensor core	312	989	3,17
Memory bandwidth (TB/s)	2	4,8	2,4

Performances : benchmarks

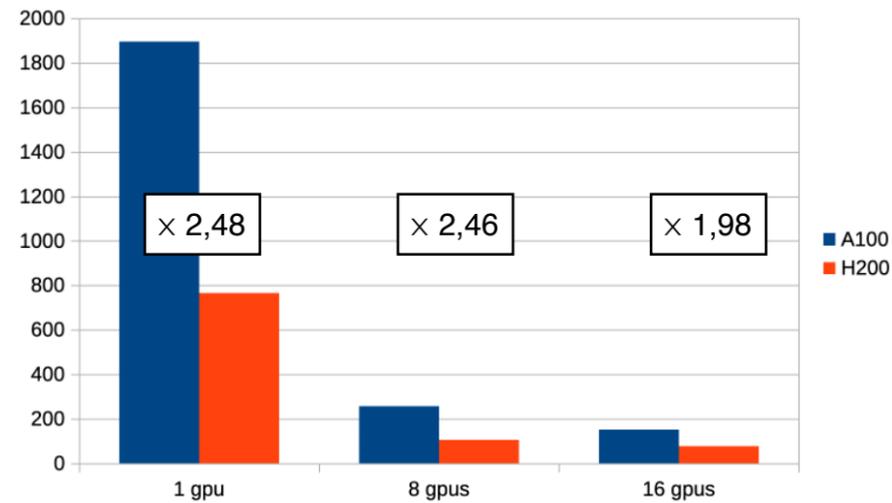
H200

- HPCG (High Performance Conjugate Gradient), 1 / 8 / 16 GPUs : 775 / 580 / 11072 GFlops
- HPL (High Performance Linpack, FP64 tensor core), **1** / 8 / 16 GPUs : **44.5** / 341 / 661 TFlops
 - (Pic théorique 1 GPU : 67 TFlops)
- **Stream : 4.3 TB/s** (théorique : 4.8 TB/s)
- Codes de simulation : accélération d'un facteur 2 constatée sur applications testées, par rapport aux GPU A100

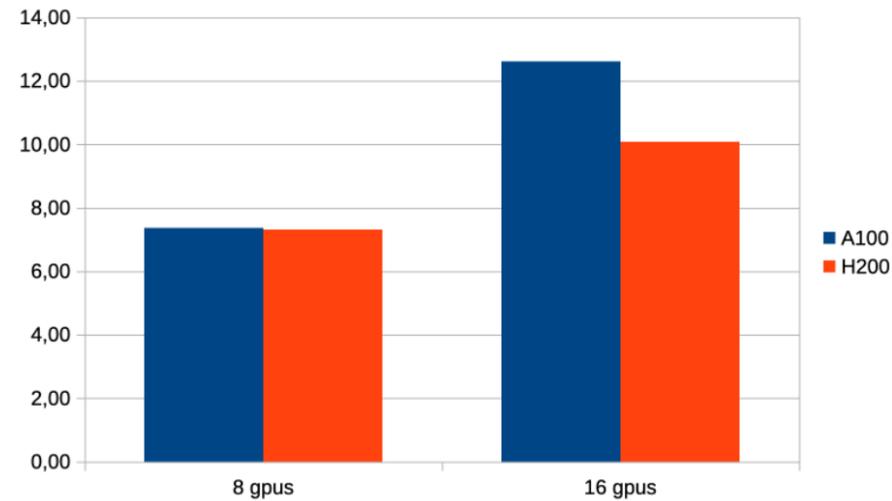
H200

Tests et bench IA

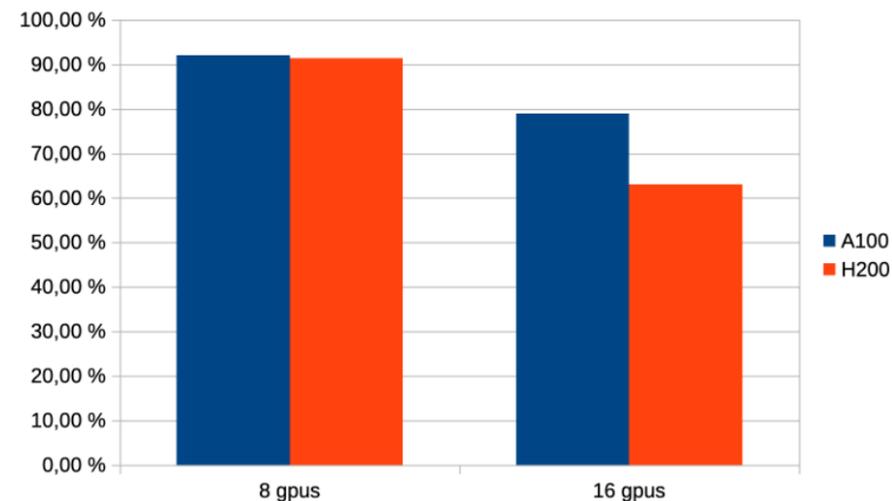
- Pytorch uniquement pour le moment (/!\ Nécessite $\geq 2.6.0$)
- Benchs Criann
 - Modèle gpt-2-XL Community (1,5B) 10 000 items de Wikitext Graphcore
 - 2,5 \times plus rapide
- Beta test utilisateurs en cours
 - 1^{ers} retours : 1,6 à 2,8 \times plus rapide
 - Mémoire 141 Go => augmentation de la taille des modèles ou des batches



Performance A100 vs H200 : 1, 8 et 16 gpus (2 nœuds)



Speedup A100 et H200 8 et 16 gpus



Efficacité A100 et H200 1 et 2 nœuds

Partitions GPU Nvidia A100

État actuel pour discussion d'évolutions

Partition	Durée max	Nœuds disponibles	Limites par calculs
gpu	48h	<ul style="list-style-type: none">• 6 nœuds AMD Milan SMT (128c)• 5x8 GPUs Ampere A100, 80GB• NVLink4	<ul style="list-style-type: none">• 2 nœuds :• 16 GPU• 256 cœurs CPU• 2x480000 MB de mémoire
hpda	72h	<ul style="list-style-type: none">• 4 nœuds AMD Milan SMT (128c)• 4x8 GPUs Ampere A100, 80GB• NVLink4	<ul style="list-style-type: none">• 1 nœud• 8 GPU• 128 cœurs CPU• 480000 MB de mémoire
gpu_all	48h	Les nœuds des partitions gpu et hpda	<ul style="list-style-type: none">• 2 nœuds :• 16 GPU,• 256 cœurs CPU• 2x480000 MB de mémoire
hpda_mig	72h	<ul style="list-style-type: none">• 1 nœud AMD Milan SMT (128c)• 8 GPUs Ampere A100 80GB• NVLink4	

Logithèque

THÉMATIQUE SCIENTIFIQUE	NOM DU LOGICIEL	THÉMATIQUE SCIENTIFIQUE	NOM DU LOGICIEL
SIMULATION ATOMISTIQUE ET OUTILS CONNEXES	CHARMM	MODÉLISATION ATMOSPHERIQUE, CLIMATOLOGIE, OCÉANOGRAPHIE	WRF - WPS
	GROMACS		NCL
	NAMD		WGRIB
	MOLPRO		GEOS
	VASP		GDAL
	PSI4		CDO
	DALTON		R_TERRA
	AMF		NCO
	LAMMPS		SIRANE
	QCORE		Code ASTER
	ASE	CAST3M	
	GAMESS	HYPERWORKS	
	GAUSSIAN	LS-DYNA	
	Quantum Espresso	NASTRAN	
	Koopmans	SALOME-MECA	
	SIESTA	COMSOL	
	ORCA	FREEFEM ++	
	Open Babel	OCTAVE	
	VESTA	SCILAB	
	VMD	R	
BIOLOGIE	Augustus	DEEP LEARNING	Python/dask
	Guppy		Python/pandas
	Dorado		PyTorch
	openRDP		TensorFlow/Keras
MÉCANIQUE DES FLUIDES	Star CCM+	MAILLAGES	Horovod
	SWASH		Scikit-learn
	MODULEF	COUPLEURS	GMSH
	DUALPHYSICS		Oasis
	YADE	VISUALISATION	Precice
	TELEMAC-MASCARET		Paraview
	OPENFOAM		Ferret
	FOAM-EXTEND		Xmgrace
	FLUENT		ChimeraX
	CODE_SATURNE		Molden
			Ncview
			VTK

Austral

Développement de la logithèque

- Visualisation
 - VTK 9.4.1 (Visualisation Toolkit)
 - ChimeraX 1.9.1 (privé, modélisation moléculaire)
- CFD : OpenFOAM v2412
- Simulation atomistique
 - GAMESS 2024R2P1, LAMMPS 29Aug24
 - GROMACS 2024.2 (ajout version GPU AMD) et 2025.1
- Biologie
 - Dorado 0.9.0 et 0.9.1, DSRC 2.0.2
- Bibliothèque FFT
 - cFFFT 2.14.0, OpenCL / compilation pour GPU NVIDIA
- AI-Deep Learning
 - Pytorch 2.4.1, 2.5.1 et 2.6.0 (× 2 python 3.12 et 3.13)

Datasets IA

Datasets IA

Systeme de fichier Lustre

- Rappel
 - Lustre non adapté aux bases nombreux petits fichiers => perte de performance pour les accès aux données
 - Foisonnement de petits fichiers => saturation des metadata ; dépassement du quota inode
- Solutions (CT 9 avril 2024) :
 - Utiliser des archives non compressés pour de telles bases :
 - Montage squashfs
 - Utilisation de bibliothèques optimisées (webdataset)
 - <https://services.criann.fr/services/hpc/cluster-austral/guide/ia-deep-learning/#datasets>
 - Utiliser des formats de fichiers adaptés

Datasets IA

Mutualisation des datasets

- Demandes utilisateurs :
 - CT 8 octobre 2024 : disposer des datasets « standards » dans des dossiers accessibles à tous, (solution Idris)
 - Demandes régulières d'augmentation de quota pour des bases de plus en plus grosses (de quelques 100aines de Go à plusieurs To).
- Retour du Criann
 - Téléchargement des datasets standards laborieux et potentiellement complexe (adaptation des formats de fichiers, licence de diffusion, ...)
 - A faire en collaboration avec les utilisateurs
 - Proposition d'une réunion spécifique sur le sujet : CT thématique IA (juin 2025)

Agenda

Formations & actualités diverses

Agenda des formations

À venir sur 1^{er} semestre

- Programmation parallèle avec MPI
 - 22 et 23 mai, en visio (ouverture nationale)
- Python pour le HPC : 12 & 13 juin (2 jours)

- Inscriptions sur le site
 - <https://indico.criann.fr/category/3/>

- En cours de préparation : initiation à cmake

Agenda

À venir

- Dans le cadre de EuroCC (CC-FR)
 - Le Criann sera partenaire de la Gray Scott School, programmation sur architecture hétérogène
 - du 23 juin au 4 juillet
 - 1^{ère} semaine : CPU - 2^{ème} semaine GPU
 - Cemracs : école d'été sur le calcul quantique
 - du 15 au 19 juillet à Marseille (et en visio)
 - suivi d'un hackathon du 21 juillet au 22 août
- JCAD 2025, Lille, du 15 au 17 septembre
 - Appel à contributions à venir

Agenda

À venir

- **Journée scientifique des utilisateurs du Criann**
 - 3 juin 2025 au Madrillet
 - Focus :
 - environnement et risques associés
 - volet technologique : machine Adastra du Cines avec retour d'expérience d'utilisateur
- Envisagé avec Nvidia : une demi-journée de présentation orientée IA
 - LLM et AI for science
 - le 4 juin (à confirmer)
 - avec pour perspective des journées de formation avancée sur les 2 sujets

Agenda 2025

- Prochain Comité Technique
 - Spécifique IA semaine du 16 juin
 - Généraliste courant septembre

- Prochaine MaJ d'Austral
 - Semaine du 27 octobre
 - (arrêt de service)

Le plateau de calcul intensif du Criann est cofinancé par la Région Normandie, l'État français et l'Union européenne (Fonds Feder).
MesoNET bénéficie d'un financement de l'Agence nationale de la recherche au titre des Investissements d'avenir.
Le Centre de Compétence EuroCC français est cofinancé par l'Union européenne et par l'État français.
Le réseau régional Syvik est cofinancé par la Région Normandie et par l'Union européenne (fonds Feder).
Le fonctionnement du Criann bénéficie du soutien de la Région Normandie.



Centre Régional Informatique et d'Applications Numériques de Normandie
www.criann.fr