CADFEM

Ansys part of SYNOPSYS | APEX CHANNEL PARTNER

L'ingénierie en 2030

Les axes de développement des outils ANSYS

29/09/2025

Cabut Damien

Présentateurs du jour

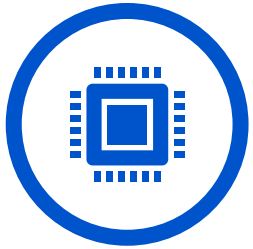
CABUT Damien

Ingénieur d‘application CFD & Optique CADFEM Europe (Lyon)

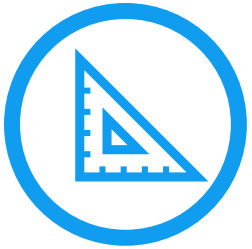
damien.cabut@cadfem.fr

Agenda

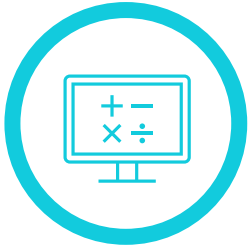
Introduction



Qu‘est ce qu‘une carte graphique (GPU)



L‘intégration au prédimensionnement



L‘accélération du process de validation

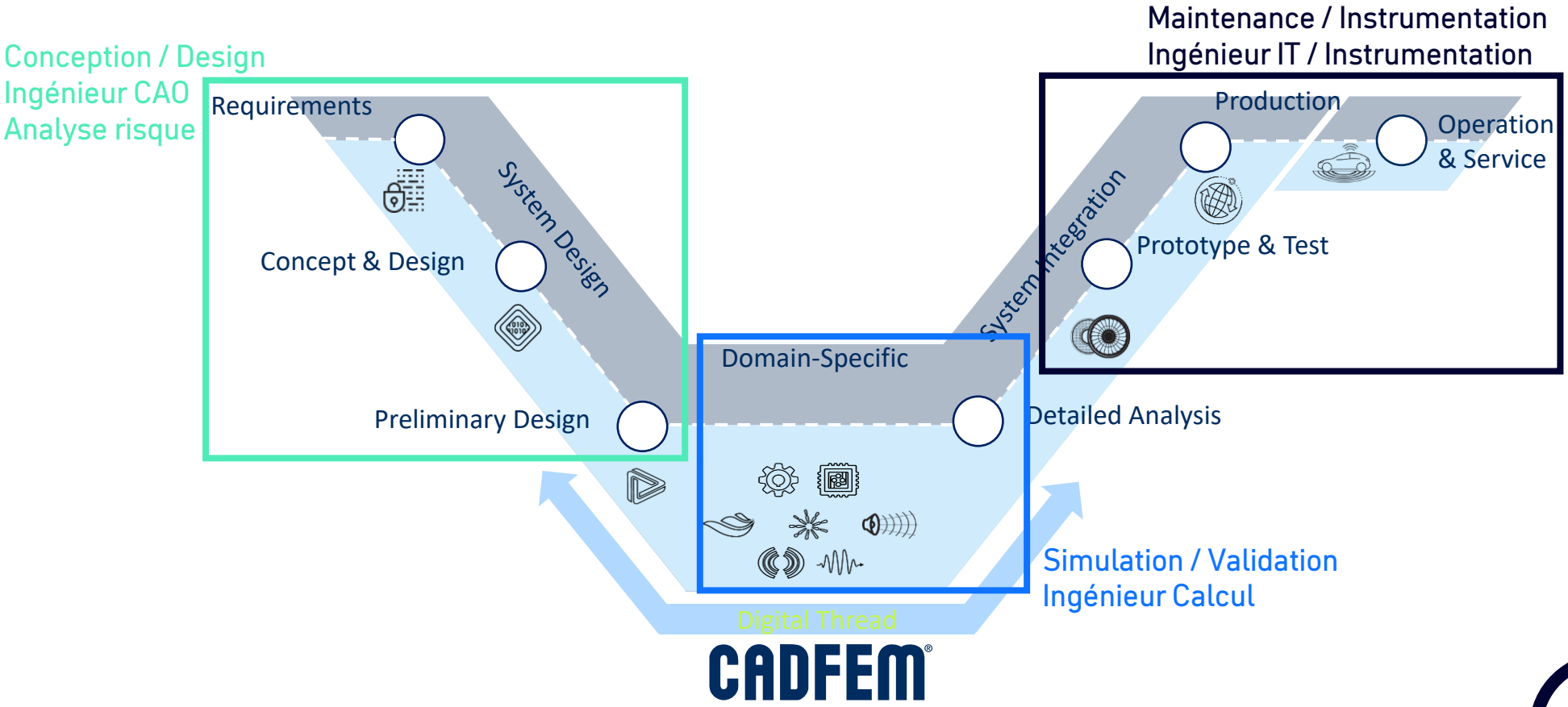
L‘optimisation et la prédiction

Conclusions

Introduction

La chaine de développement produit

Le cycle en V



Optimiser la chaine

Le cycle en V

Rendre le calcul accessible au prédimensionnement :

- **Donner accès au calcul plus tôt dans le process**
- **Calcul de première approche pour les idées de conception**
- **Limiter les calculs longs**

Accélérer les calculs de validations :

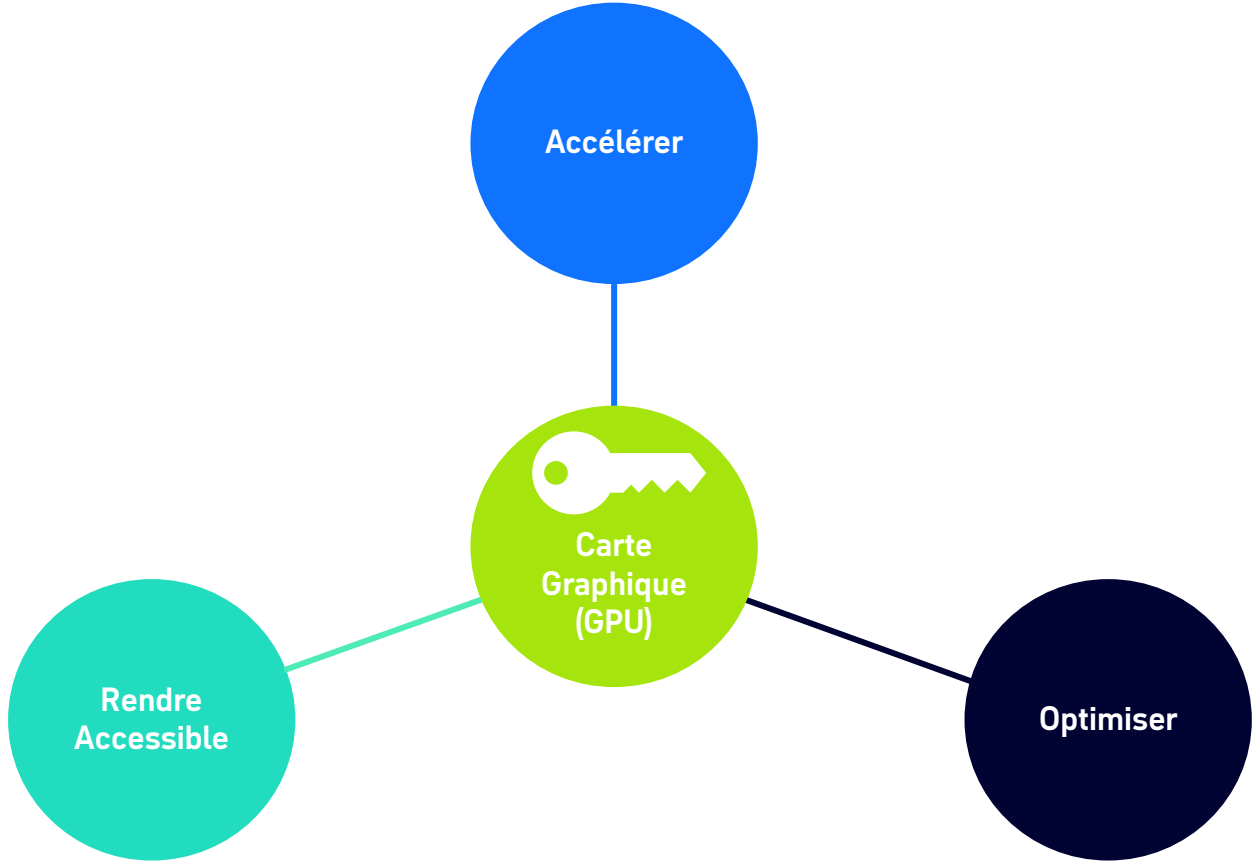
- **Validation avec les solveurs classique indispensable.**
- **Souvents très longs comme en CFD**

Optimiser le produit et le déployer :

- **Scripting/Automatisation**
- **IA**

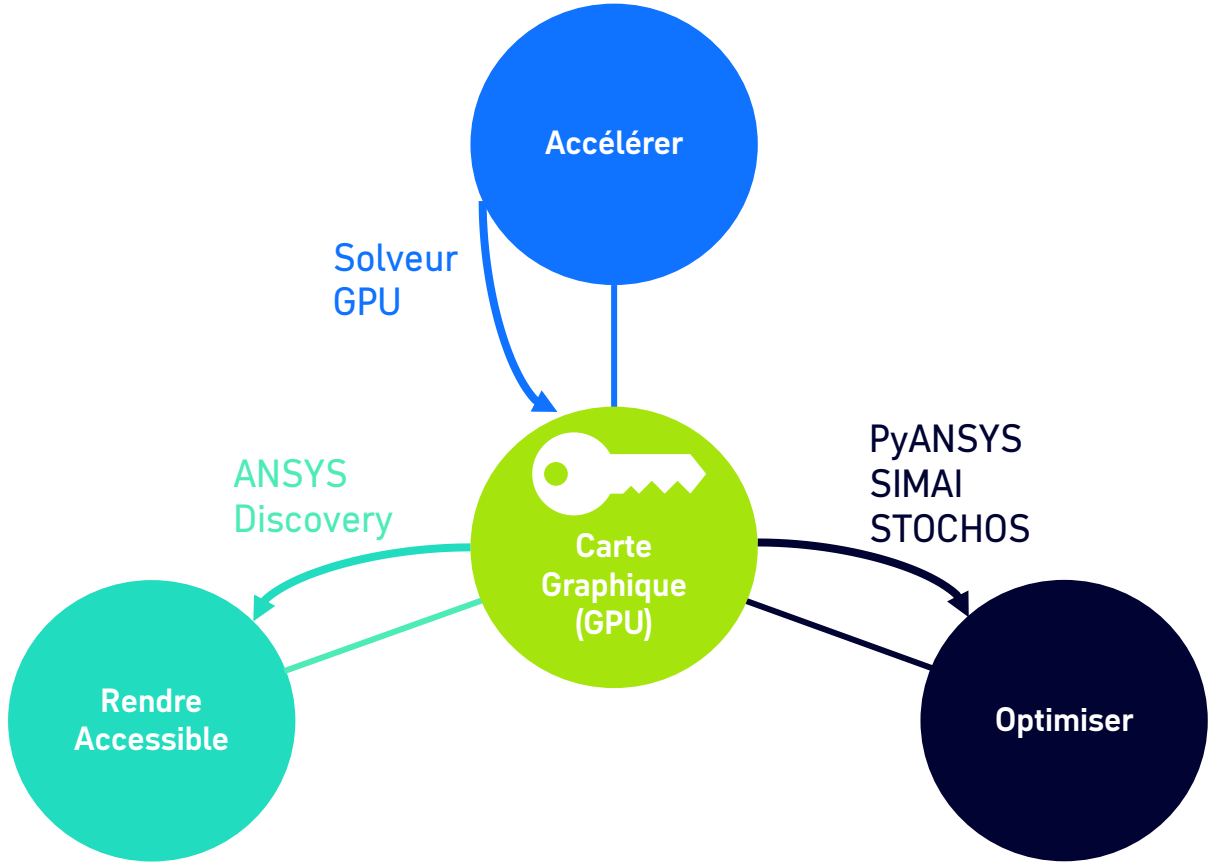
Optimiser la chaine

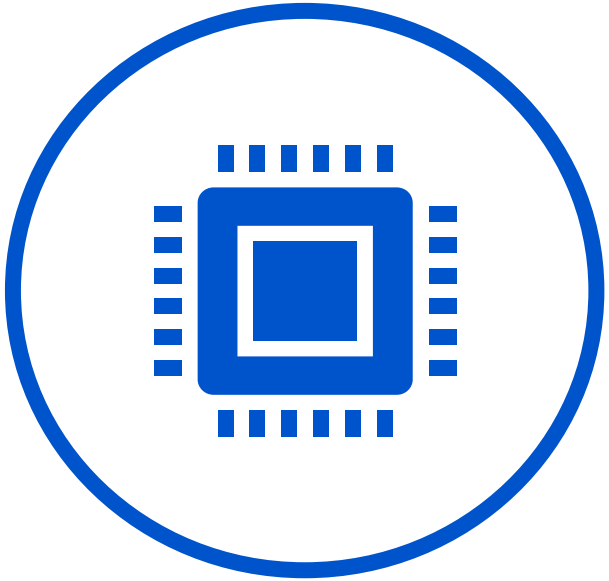
Le cycle en V



Optimiser la chaine

Le cycle en V



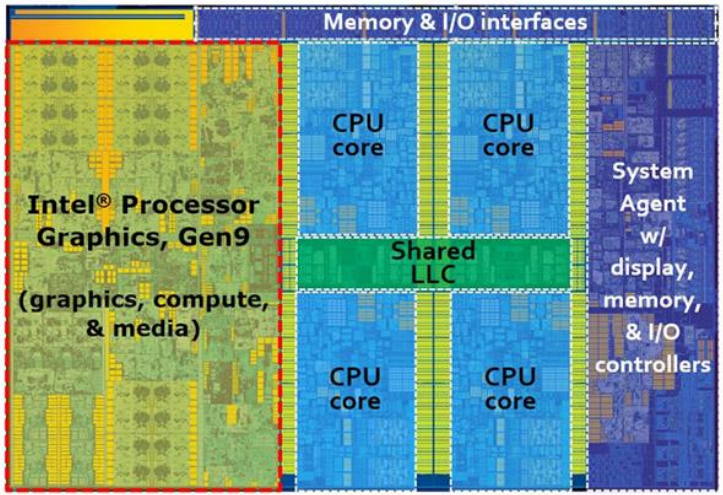


Qu‘est ce qu‘une carte Graphique

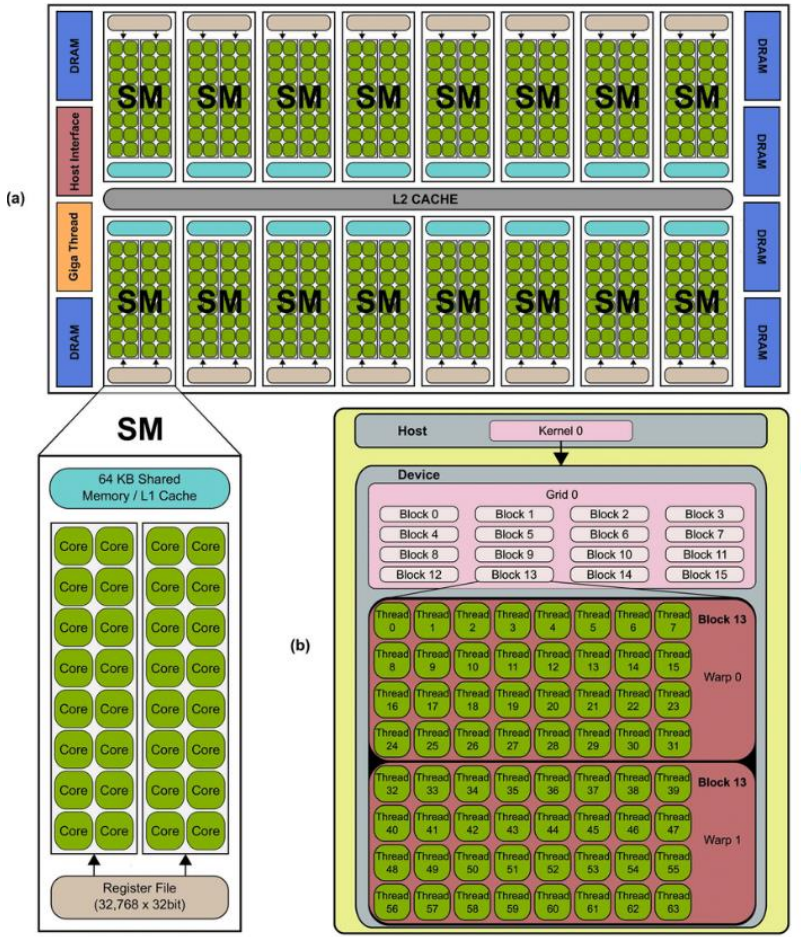
La Carte Graphique

Pleins de processeurs moins performants

Processeur



Carte Graphique



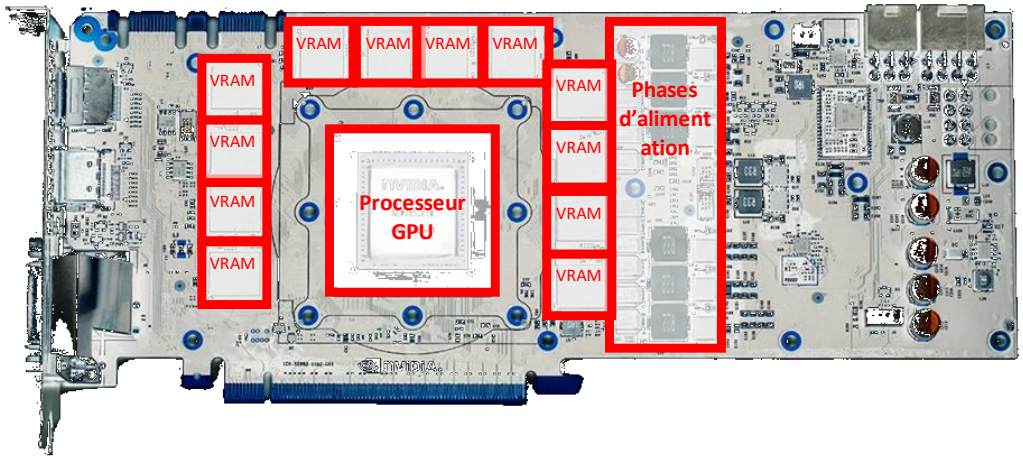
Hernandez Fernandez, Moises & Guerrero, Ginés & Cecilia, José & García, José & Inuggi, Alberto & Jbabdi, Saad & Behrens, Timothy & Sotiropoulos, Stamatios. (2015). Erratum: Accelerating fibre orientation estimation from diffusion weighted magnetic resonance imaging using GPUs (PLoS ONE (2015) 10:6 (e0130915) 10.1371/journal.pone.0130915). PLoS ONE. 10. 10.1371/journal.pone.0130915.

La Carte Graphique

Pleins de processeurs moins performants

La carte graphique est coposée de milliers de coeurs sur des disaines de micro processeurs

- Fréquences plus faibles vitesse des cœurs divisée par 2 environ.
- Opérations plus simples (moins de transistors par cœur)
- Créée pour le rendu graphique (jeu vidéo par exemple) tirs de rayons => Opérations simples avec peu de floating point (nombre de chiffres après la virgule) et peu de consommation de RAM
- Pilotée par le processeur



Simple précision ou FP32 : $\Pi = 3{,}141592653$

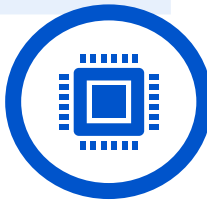
Double présision ou FP64 : $\Pi = 3{,}141592653589793$

La Carte Graphique

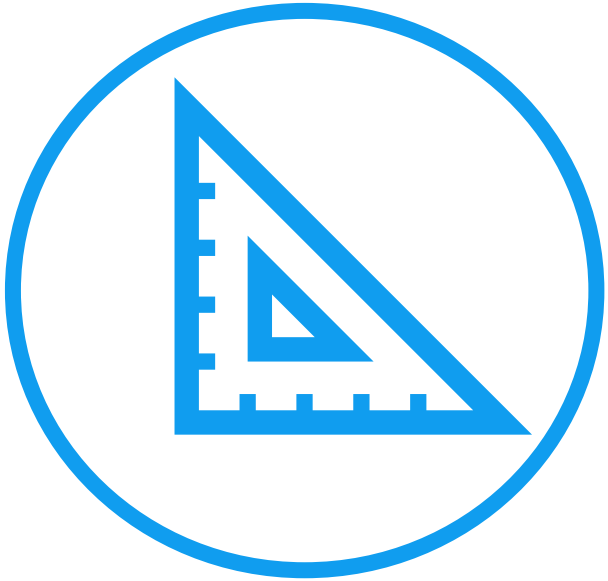
Pleins de processeurs moins performants

L‘essort de l‘IA moteur de développement

	Processeur CPU (Xeon Gold 6430)	Workstation (2023): Nvidia RTX 4000 Ada	Workstation (2025): Nvidia RTX 4000 Blackwell	Workstation (2023): Nvidia RTX 6000 Ada	Workstation (2025): Nvidia RTX 6000 Blackwell	Server (2023): Nvidia H100	Server (2025): Nvidia H200 PCIe
Performance Single Prec.	5,33 TFLOPS	26,7 TFLOPS	46,9 TFLOPS	91,06 TFLOPS	**125 TFLOPS**	51 TFLOPS	60 TFLOPS
Performance Double Prec.	2,66 TFLOPS	0,418 TFLOPS	0.733 TFLOPS	1,423 TFLOPS	1.968 TFLOPS	26 TFLOPS	30 TFLOPS
Memory Bandwidth	281,6 GB/s	360,0 GB/s	672.0 GB/s	960.0 GB/s	1.79 TB/s	1.94 TB/s	4,89 TB/s
Memory	Jusqu‘à plusieurs TB	20 GB	24 GB	48 GB	**96 GB**	80 GB	141 GB



Agenda



L‘intégration au prédimensionnement

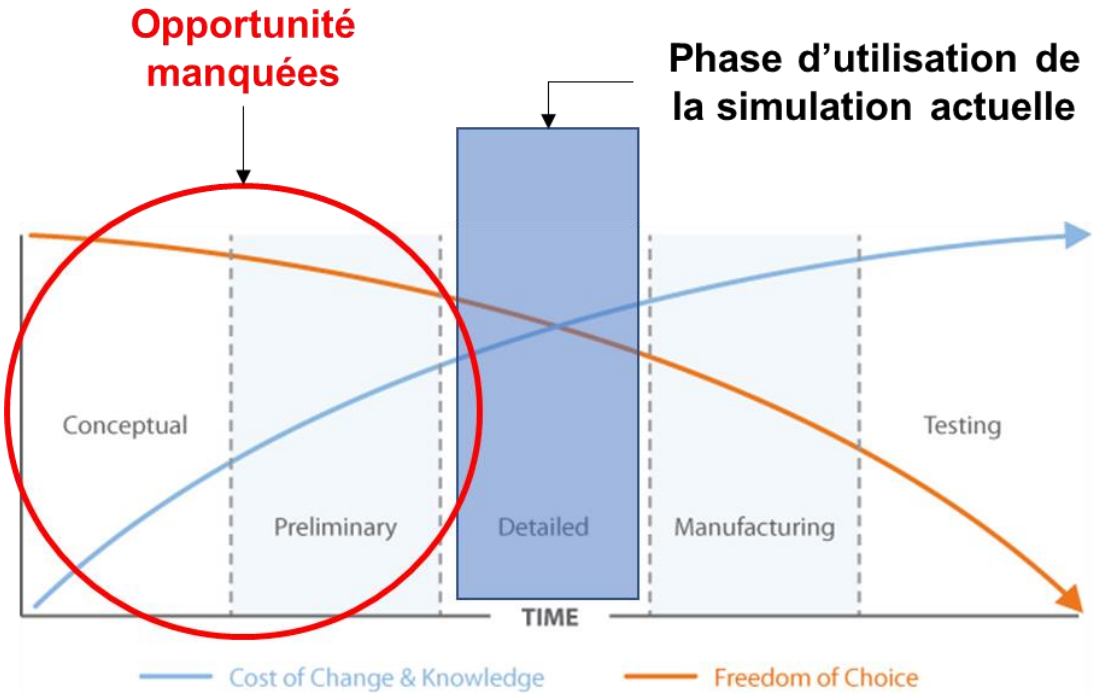
avec ANSYS Discovery

Des conceptions guidées par la physique

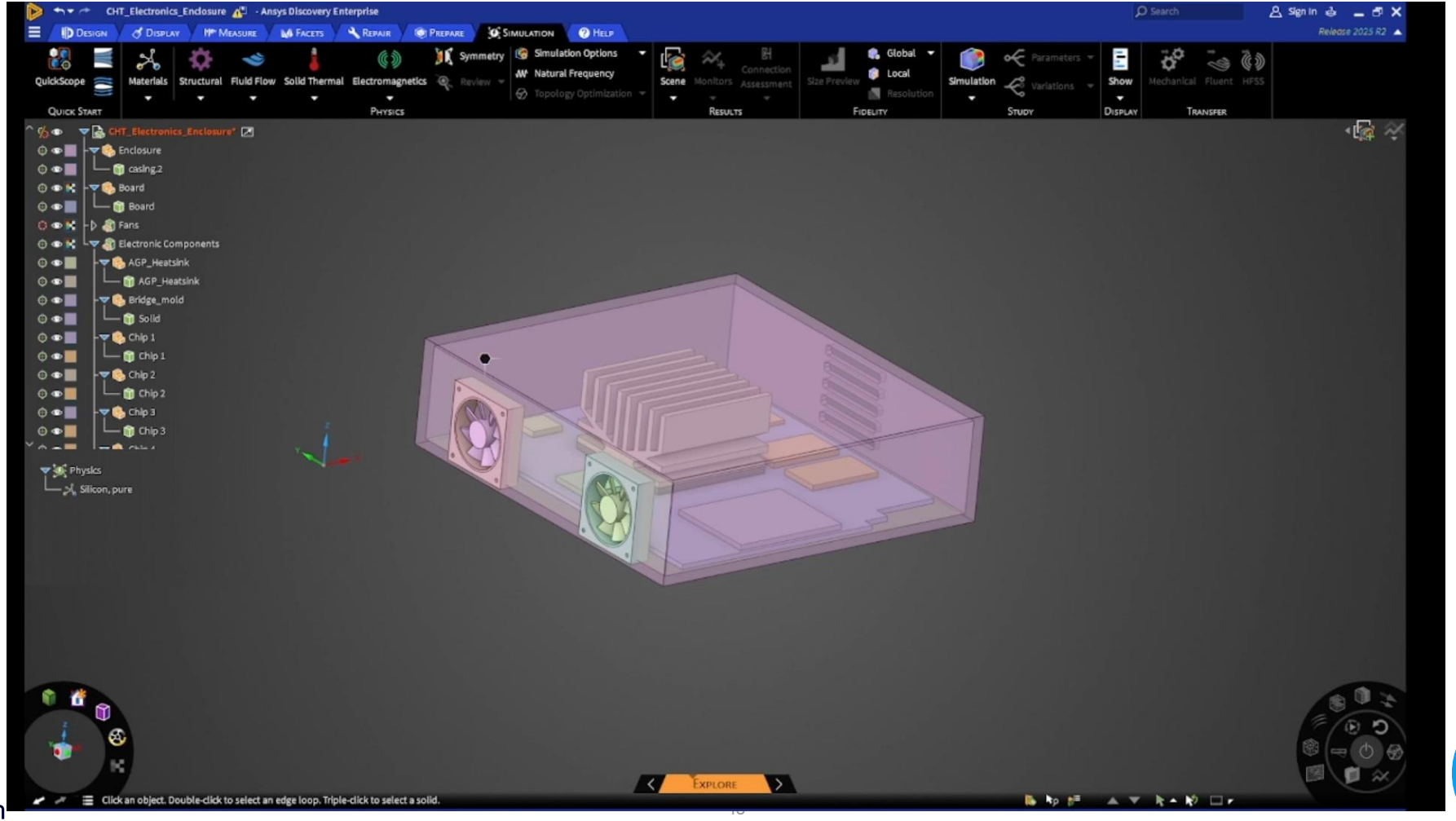
L'aide au choix proche de la CAO

Objectif

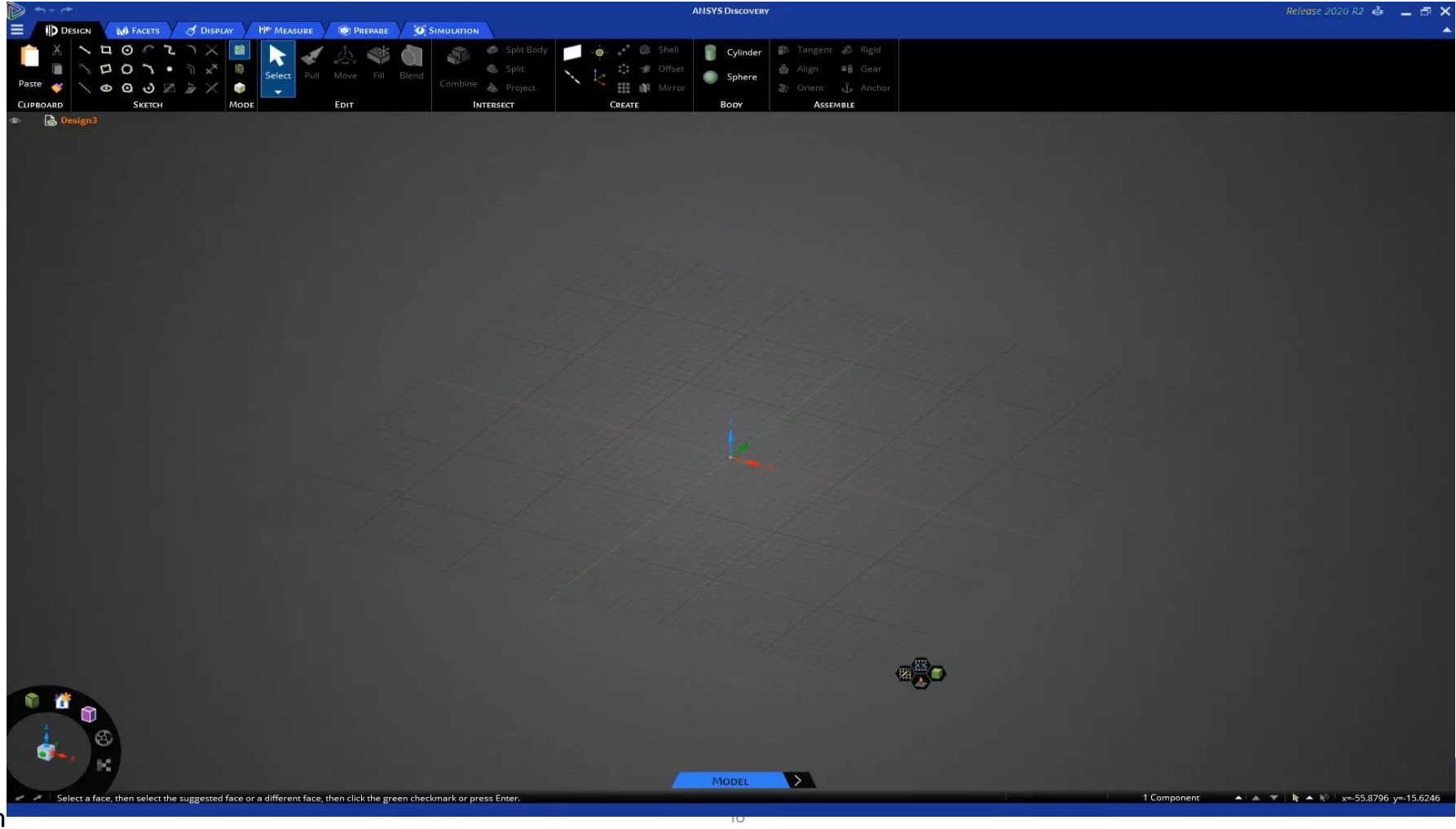
- Trier les bonnes et mauvaises idées le plus tôt possible
- Ne pas valider la performance mais donner des informations physiques
- Résultats en direct sur un design
- Logiciel simple sans calibration complexe de modèle ou de maillage



Ansys Discovery



Ansys Discovery

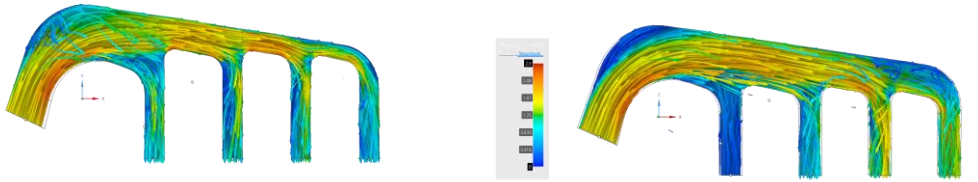


Des conceptions guidées par la physique

L'aide au choix proche de la CAO

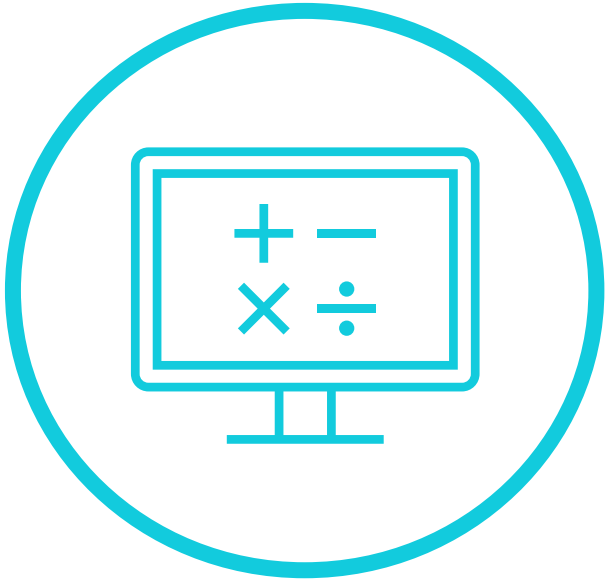
☑ Analyse des paramètres influents

- Premier niveau d'étude paramétrique sur des choix forts de design
- Essayer de trouver la base la plus pertinente et les paramètres influents
- Limiter les allers retour avec l'équipe calcul
- Gagner du temps (résultats « instantanés » grâce au calcul GPU live sur Voxels)



	Name	Diameter	Flow Inlet 1	Pressure Dr...	Max. Velocity	Avg. Total P...	Avg. Total P...
37	DV37	75 mm	0.5 m/s	74 Pa	1.18 m/s	206 Pa	132 Pa
38	DV38	75 mm	0.75 m/s	149 Pa	1.77 m/s	445 Pa	295 Pa
39	DV39	75 mm	1 m/s	246 Pa	2.36 m/s	770 Pa	523 Pa
40	DV40	75 mm	1.25 m/s	364 Pa	2.95 m/s	1180 Pa	816 Pa
41	DV41	75 mm	1.5 m/s	502 Pa	3.54 m/s	1670 Pa	1170 Pa
42	DV42	75 mm	1.75 m/s	659 Pa	4.13 m/s	2250 Pa	1590 Pa

Agenda



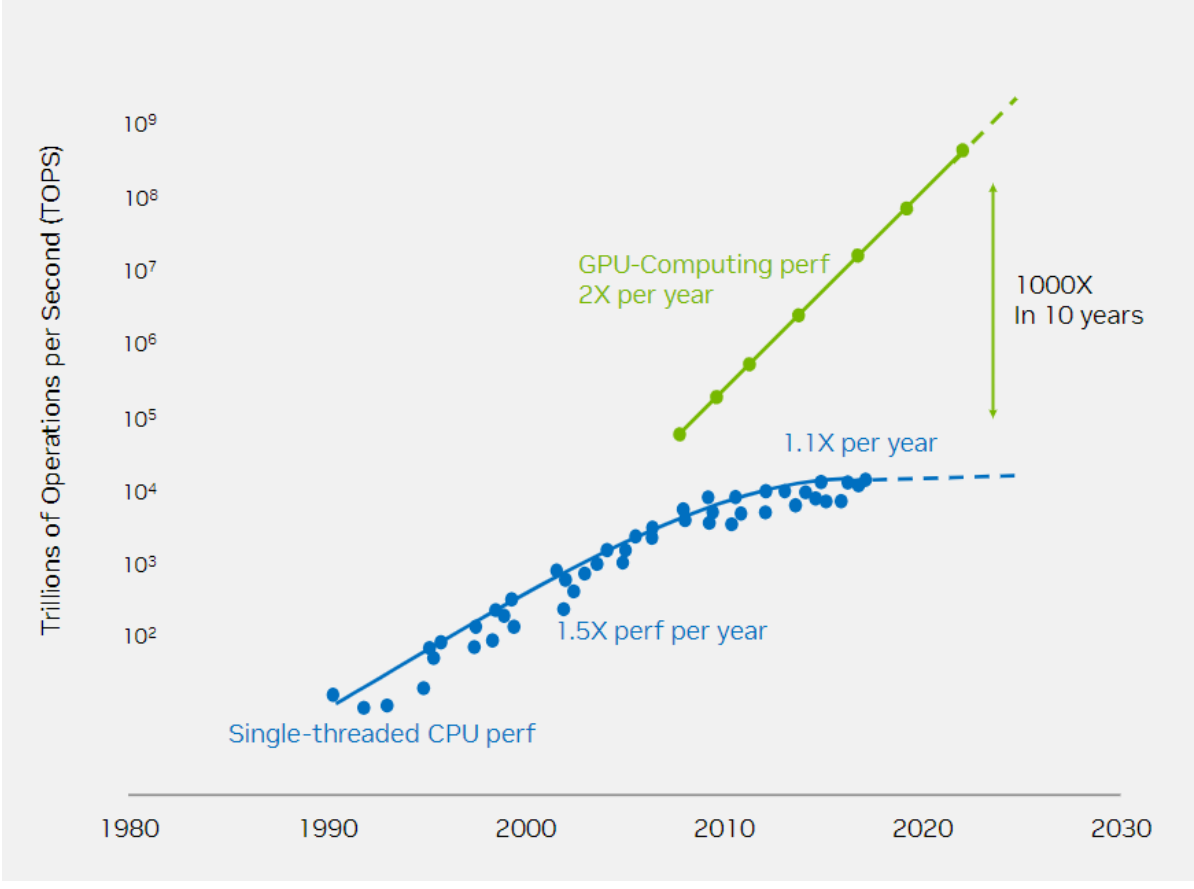
Accélération du calcul GPU

La fin de la loi de MOORE et son après

Le GPU nouveau paradigme

Loi de Moore

- Gordon E. Moore 1965
- Doublement du nombre de transistors dans les puces tous les 2 ans.
- Ralentissement depuis 2010.
- Architectures CPU incluent des complexités (petite puce, grosse concentration => Chauffe par exemple)
- Le GPU propose une architecture différente permettant un nombre massif de transistors plus lents et mieux répartis



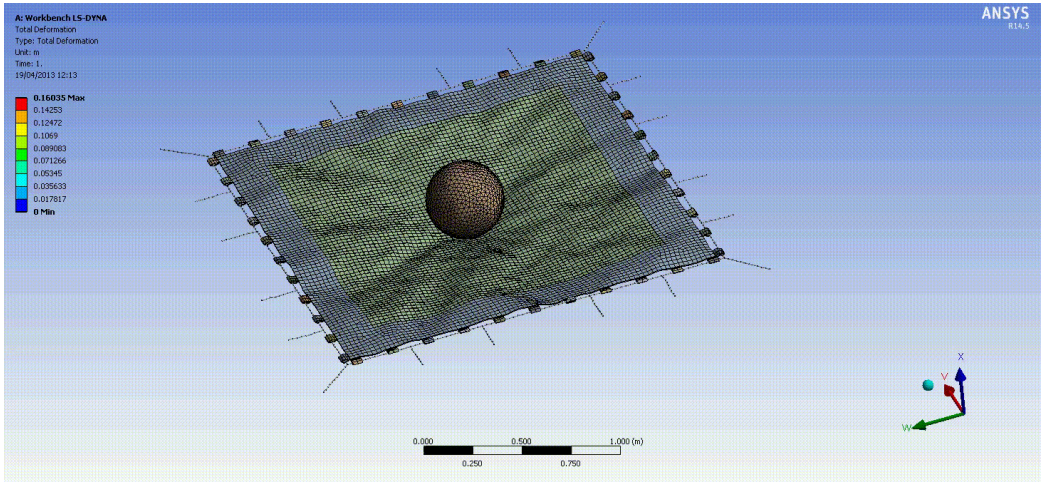
Source : Bloomberg

Les Différentes méthodes numériques

Les éléments finis

Le maillage et la méthode

- Les valeurs sont calculées aux nœuds qui sont les sommets des cellules.
- Les éléments relient les nœuds et « transmettent » l'énergie d'un nœud à l'autre tels un système masse ressort.
- => Très approprié pour les équations d'ondes ou la mécanique



Explicite

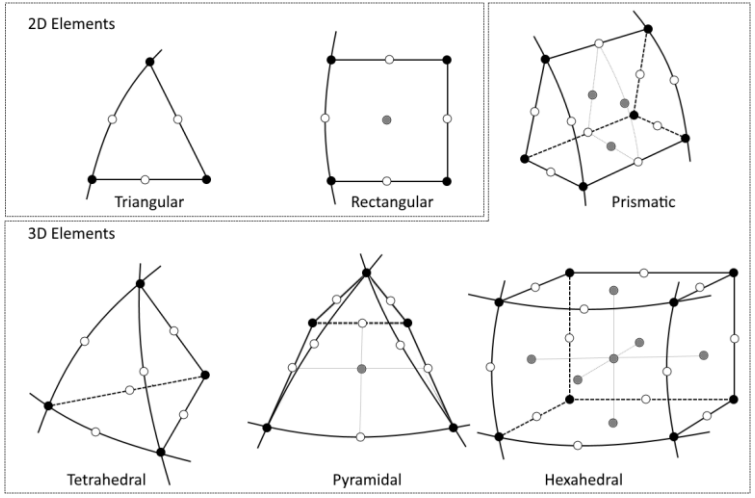
$$\frac{\partial \phi}{\partial t} = \frac{\phi_i^{n+1} - \phi_i^n}{\delta t} = G(\phi_i^n)$$

- Conditionnellement stable
- Contrainte sur le pas de temps
- Formulation plus simple
- Matrice moins dense

Implicite

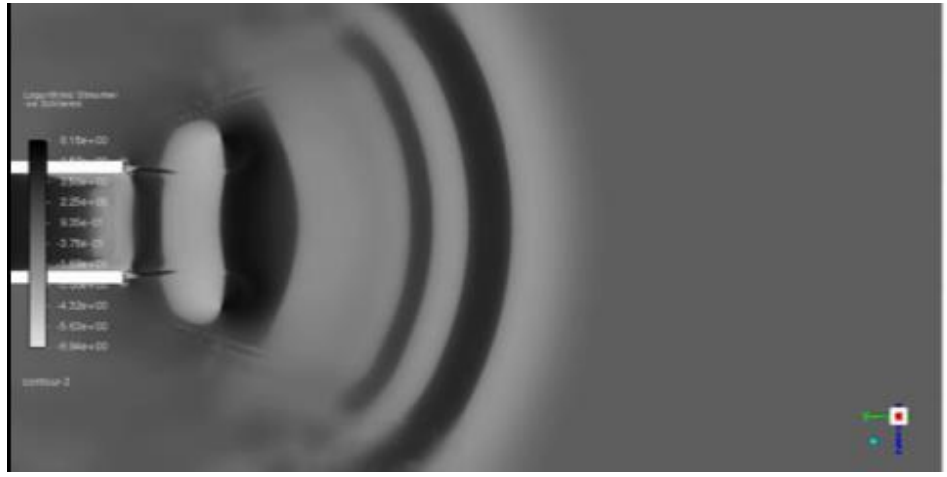
$$\frac{\partial \phi}{\partial t} = \frac{\phi_i^{n+1} - \phi_i^n}{\delta t} = G(\phi_i^{n+1})$$

- Inconditionnellement stable
- Formulation Complexe
- Matrice très dense
- Gros besoin de RAM



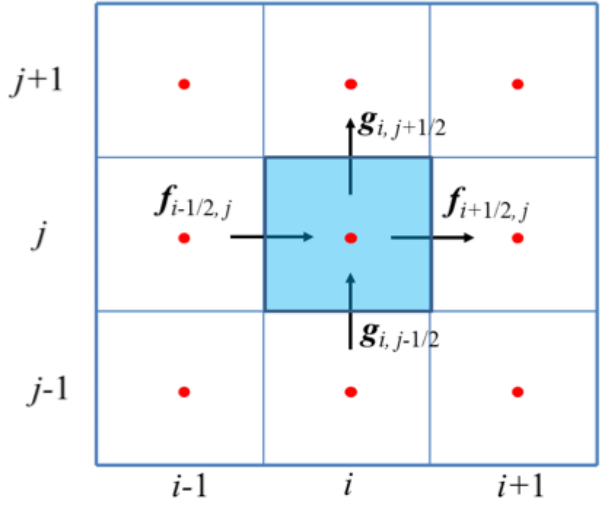
Les Différentes méthodes numériques

Les Volumes finis



La méthode et ses spécificités

- Les valeurs sont calculées aux centres des mailles. (pour fluent)
- Chaque cellule correspond à un volume qui reçoit et transfert de l'information par intégration du flux à chaque face de maille.
- Presque toujours Implicite sous Ansys
- Plusieurs méthodes permettent d'intégrer le flux aux faces d'une cellule (least square cell based, green gauss node based ou cell based). La valeur est interpolée à partir des cellules voisines.
- Matrice moins Dense => Moins de RAM consommée
- Cell centered => Matrice plus petites (Fluent)



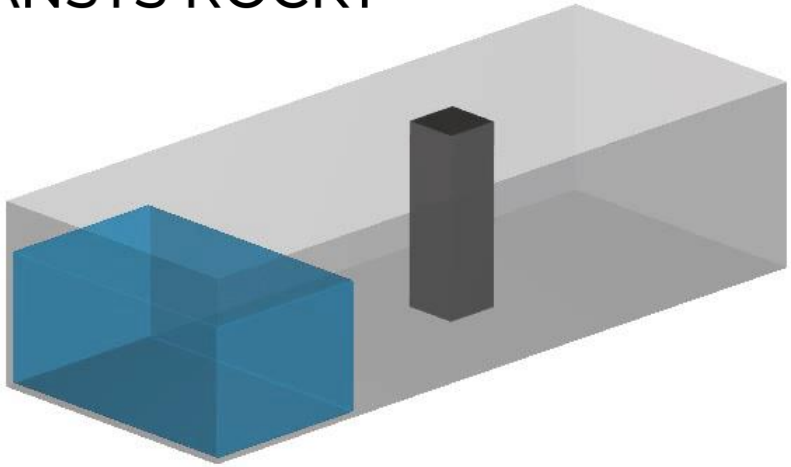
Les Différentes méthodes numériques

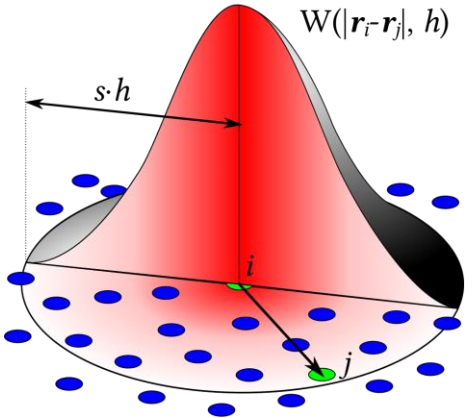
Les Particules

Methodes à particules

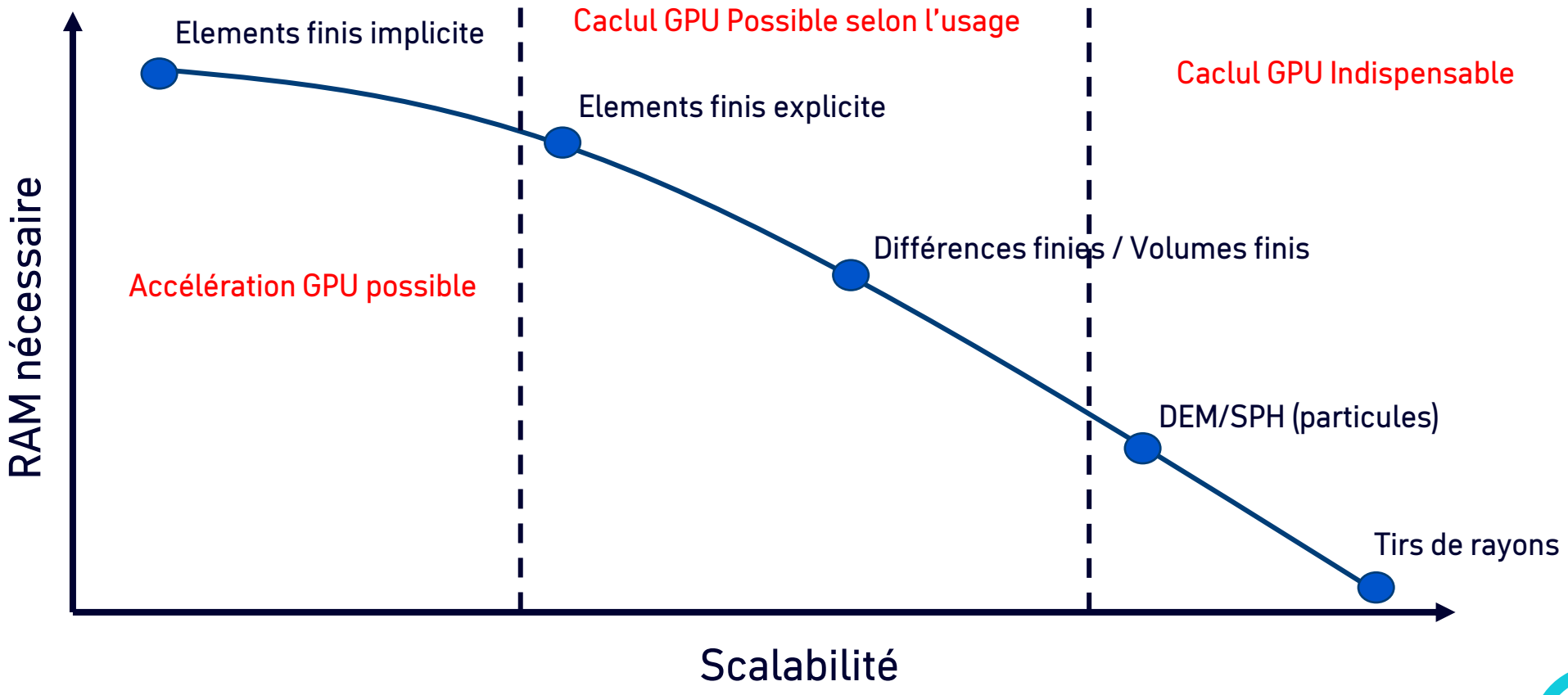
- Résolution de la 2nd loi de newton sur un point.
- Les particules sont comme des billes qui interagissent via des forces avec leur voisines.
- Les particules sont « quasi » indépendantes. DEM pour le calcul de particules solides. SPH pour le calcul d'écoulements fluides.
- Indépendance réduisant considérablement le besoin de RAM mais pouvant nécessiter beaucoup de particules pour une bonne résolution
- Solveur Explicite

ANSYS ROCKY





La scalabilité et la consommation de RAM



3 Cas de figures

3 Salles 3 ambiances

Accélération GPU

Elements finis :
ANSYS MECHANICAL

Solveur GPU méthode discrète

Différences finies / Volumes finis :
ANSYS FLUENT
ANSYS LUMERICAL

Solveur GPU Lancer de rayons / Particules

Méthodes à particules ou tirs de rayons :
ANSYS ROCKY
ANSYS FREEFLOW
ANSYS SPEOS

L'accélération GPU

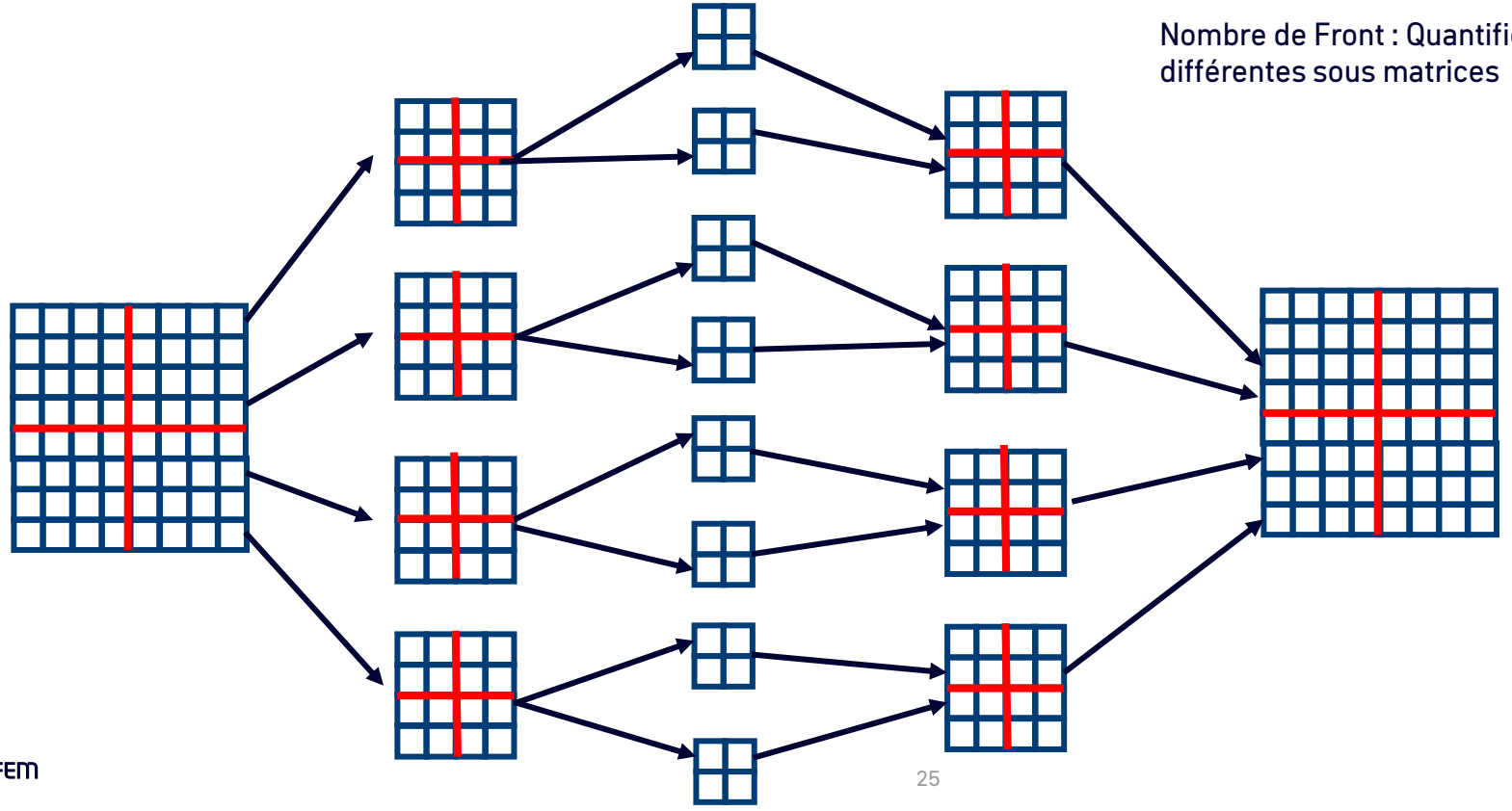
Ansys Mechanical

Accélération GPU

Parallélisation du solveur Direct

Découpage du problème en sous matrices inversables pour la parallélisation initiale

Nombre de Front : Quantifie la densité des différentes sous matrices



L‘accélération GPU
Ansys Mechanical

Accélération GPU

L‘accélération en mécanique

Un petite portion des partitions correspond à la majorité du temps de calcul (sous matrices les plus denses). Le offloading consiste à déléguer le calcul de ces sous matrices denses au GPU.

- Besoin de VRAM pour faire tenir l‘ensemble à résoudre et maximiser le offload.
- Solveur Direct ou Mixed uniquement
- Meilleurs performances à faibles nombres de coeurs
- Performances simple précision souvent suffisantes.

Solveur GPU méthodes discrètes

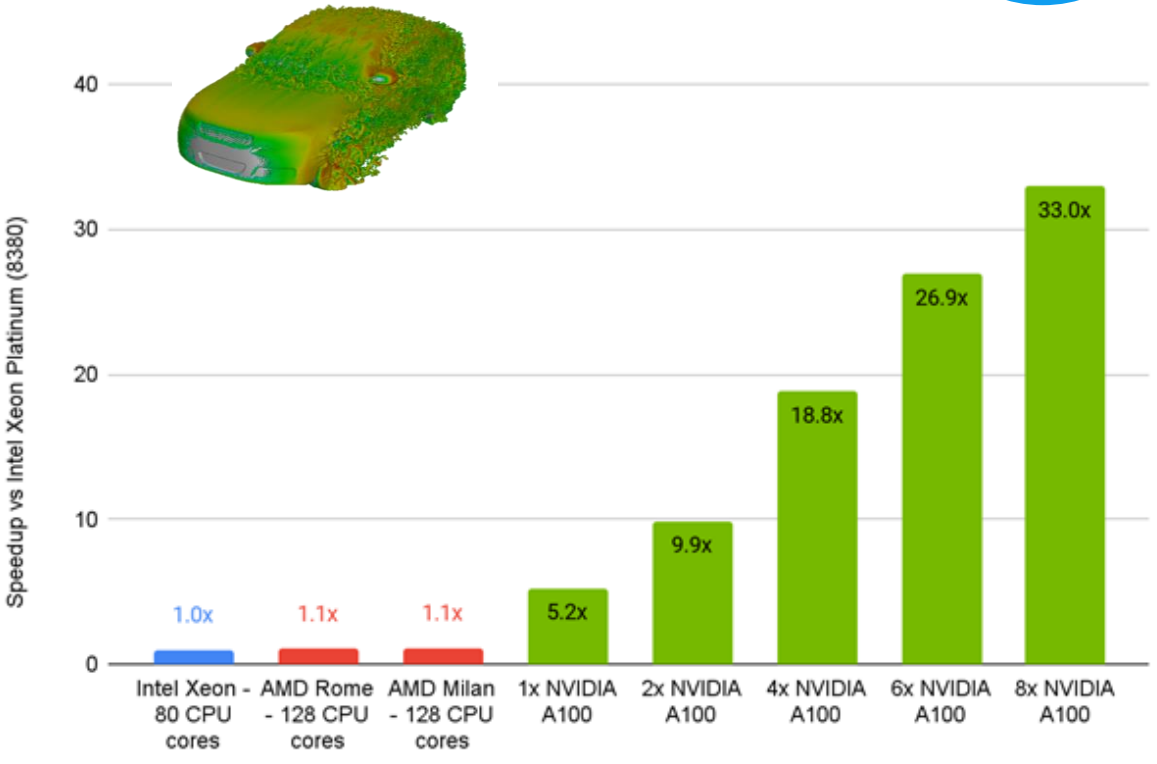
Ansys Fluent

Solveur GPU méthode discrète

Solveur parallélisé complètement sur GPU

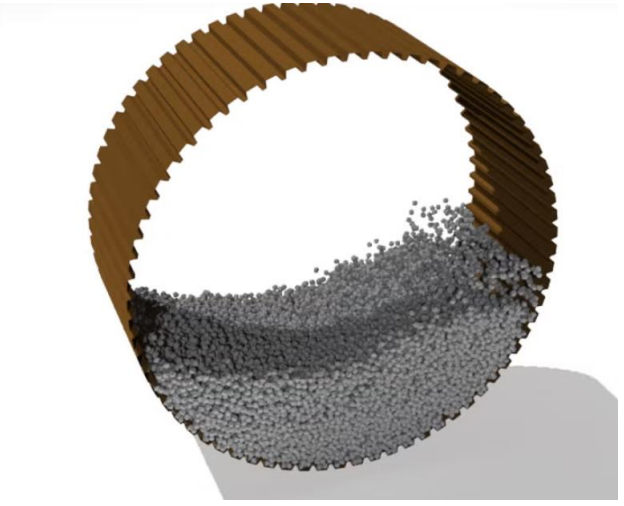
Solveur GPU complet. Attention communication avec la RAM du PC longue => Obligation de tenir dans la VRAM du GPU

1. Vérifier la RAM de vos modèles habituels (considérer un besoin légèrement plus grand en GPU)
2. Quantifier la réduction de RAM envisageable (Passer en solveur Segregated par exemple)
3. Definir le besoin (Multi-GPU ou une seule GPU)
4. TESTER !!!!! (CADFEM et FRA-SYS peuvent vous proposer des tests)



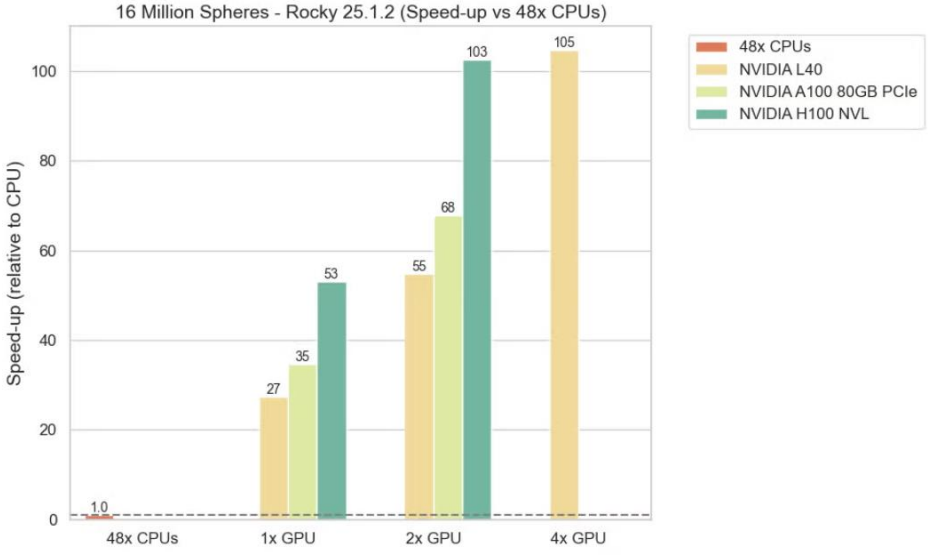
Particules
Ansys Rocky/Freeflow

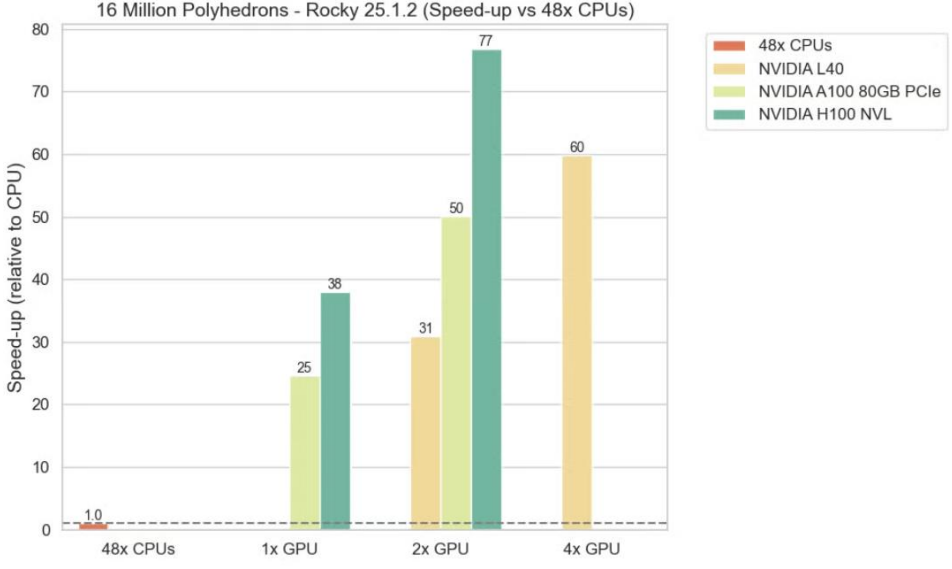
Solveur GPU
Lancer de rayons / Particules



La question du GPU ne se pose PRESQUE pas :

Toute GPU est bonne à prendre.
Attention tout de même à la RAM pour les polyhères
Surtout en double précision privilégier tout de même les versions solveur





Source : Rocky GPU Buying Guide | Ansys Knowledge

Agenda

L‘optimisation

L‘IA et ses mystères
Les possibilités

Les réseaux de neurones

- L’essors des cartes graphiques (GPU) et leurs changements de conceptions (augmentation de la mémoire) est piloté par un besoin de l’IA sur des modèles de réseaux de neurones toujours plus profonds.
- Plus le réseau est profond plus on a de paramètres et de capacité à faire varier les sorties et capter les non-linéarités
- Super ! Mais a-t-on vraiment toujours besoin d’autant ?
- Modèles de réseaux de neurones plus simples à objectifs fonctionnels => temps d’entrainement plus faible et nombre de données plus faibles.
- Une régression linéaire ? C’est un moyen de prédiction (IA). Les métamodèles ou surfaces de réponses donnent souvent l’information nécessaire

3 possibilités / combinaisons

3 Salles 3 ambiances

Optimisation (métamodèle)

Optislang

Réseaux de neurones légers

STOCHOS (CADFEM AI)
SIM AI

Scripting et automatisation

PyANSYS

Optimisation
ANSYS Optislang

Optimisation (métamodèle)

Process integration

Automation

Enable a holistic view in optimisation

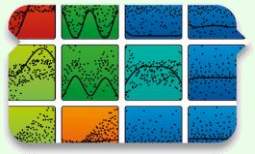
Parametric Variation Analysis

Automated Workflows

CAE expert

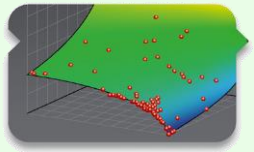
Easy to build and reuse repetitive workflows

Sensitivity
Design Understanding



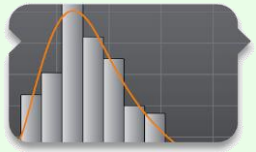
Investigate parameter sensitivities, reduce complexity and generate best possible metamodels

Optimization
Design Improvement



Optimize design performance

Robustness
Design Quality



Ensure design robustness and reliability

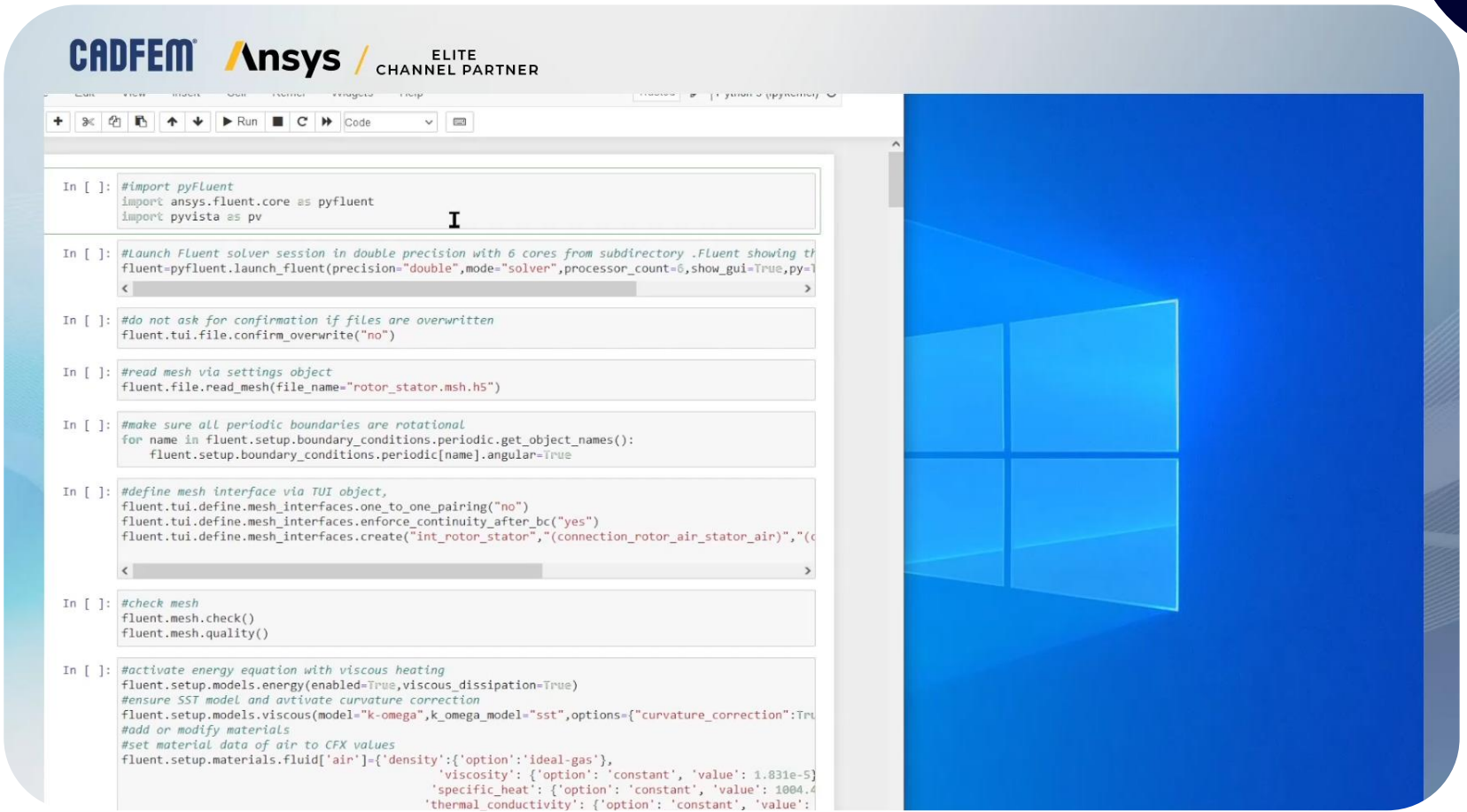
IA
STOCHOS

Réseaux de neurones légers

		aka	Engineering Value
Understand Unsupervised learning	Root cause analysis Fault Detection	SENSI / ROB	Robustness (predictive maintenance)
Predict Supervised learning	DIM-GP Prediction 17.5 s, 32 training simulations training: 21s Crash simulation 0.5h on 20 CPUs DIM-GP Prediction 30 s, 34 training simulations CFD simulation: 5h on 1 low-end GPU	MOP / ROM	Accelerate (time to market)
Decide Reinforcement/ active learning	Automatic Face Selection: Mixing Optimisation:	OPTI	Accuracy (quality, automation)
Create Pre-trained Generative AI	Ansys GPT Stochos GEN-AI	Copilot	Know-How/ Improve (innovation)

PyAnsys
Facilitateur d'interactions

Scripting et automatisation



Agenda

Conclusion

En conclusion

L’avenir du calcul se passe sur GPU !

- L’ouverture au calcul live en première approche dès la conception
- L’accélération des calculs solveurs flagships
- Ouverture à l’IA et l’optimisation.
- Facilité par l’ouverture au scripting Python

Commercial aircraft external aerodynamics in high lift configuration;
1B cells,
28 mins to mesh on 512 CPUs,
30 hrs to solve on 40 L40 GPUs

CADFEM

CADFEM Europe
148 avenue Jean Jaurès,
69007, Lyon

contact@cadfem.fr

CADFEM.net

YouTube

Ansys part of Synopsys APEX CHANNEL PARTNER