

Evaluation des performances d'un portage ciblé sur GPUs dans un code de mécanique des fluides

1. Contexte et objectifs
2. Résultats

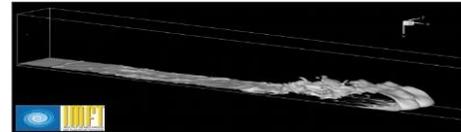
Elyakime P. - Ciavaldini Y. - Bonometti T.

JCAD 2022, 12 Octobre 2022

1.1 - JADIM

Code de calcul volumes finis du 2nd ordre @IMFT et @LGC

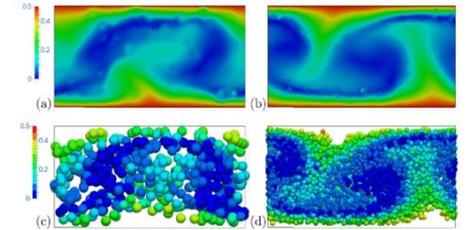
- Ecoulements instationnaires de fluides incompressibles
- Maillages structurés cartésiens ou curvilignes 2D-3D
- Fortran 90/95 parallélisé avec une méthode de décomposition de domaines (MPI)



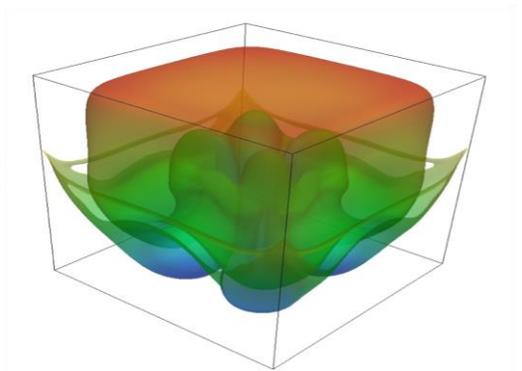
Gravity current. Y. Hallez

Modules physiques :

- Multiphasique par une approche de type Volume Of Fluid ou Level-set
- IBM : Ecoulement autour d'objets (Immersed Boundary Method)
- LES (Large Eddy Simulation)
- Transfert thermique
- Réaction chimique, ...



Inertial finite-size particles in turbulent Couette flow (Force Coupling Method). G. Wang.



Crustal polydiapirs
(Convection dans la croûte terrestre)
A. Louis-Napoléon

1.2 - Méthode de projection & temps de calcul

CPU

Actuellement
70% du temps
de calcul

- Calcul de l'avancement en temps (schéma de Runge-Kutta / Crank-Nicolson)

$$\frac{\hat{V}_i^{n+1} - V_i^n}{\Delta t} \mathfrak{G} = - \left(\frac{1}{\rho} \frac{\partial P}{\partial \xi_i} \right)^n \mathfrak{G} + \text{Gravité} + \text{Advection} + \text{Diffusion} + \text{Effets capillaires}$$

- Résolution d'un **système linéaire** (pseudo-équation de Poisson)

$$\frac{1}{\Delta t} \nabla \cdot \hat{V}^{n+1} = \nabla \cdot \left(\frac{1}{\rho} \nabla \Phi \right)$$

*Solveurs utilisés par JADIM sur CPU :
PETSc [Jacobi+PCG], Fourier*

- Mise à jour de la vitesse et de la pression

$$\frac{V_i^{n+1} - \hat{V}_i^{n+1}}{\Delta t} = - \frac{1}{\rho} \frac{\partial \Phi}{\partial \xi_i}$$

$$P^{n+1} = P^n + \Phi$$

1.3 - Stratégie d'optimisation du temps de calcul

Après



CPU

- Calcul de l'avancement en temps (schéma de Runge-Kutta / Crank-Nicolson)

$$\frac{\hat{V}_i^{n+1} - V_i^n}{\Delta t} \mathcal{G} = - \left(\frac{1}{\rho} \frac{\partial P}{\partial \xi_i} \right)^n \mathcal{G} + \text{Gravité} + \text{Advection} + \text{Diffusion} + \text{Effets capillaires}$$



GPU

?

- Résolution d'un **système linéaire** (pseudo-équation de Poisson)

$$\frac{1}{\Delta t} \nabla \cdot \hat{V}^{n+1} = \nabla \cdot \left(\frac{1}{\rho} \nabla \Phi \right)$$

Solveur sous GPU

Hackathon GENCI 2018 – A. Pedrono

- Mise à jour de la vitesse et de la pression

$$\frac{V_i^{n+1} - \hat{V}_i^{n+1}}{\Delta t} = - \frac{1}{\rho} \frac{\partial \Phi}{\partial \xi_i}$$

$$P^{n+1} = P^n + \Phi$$



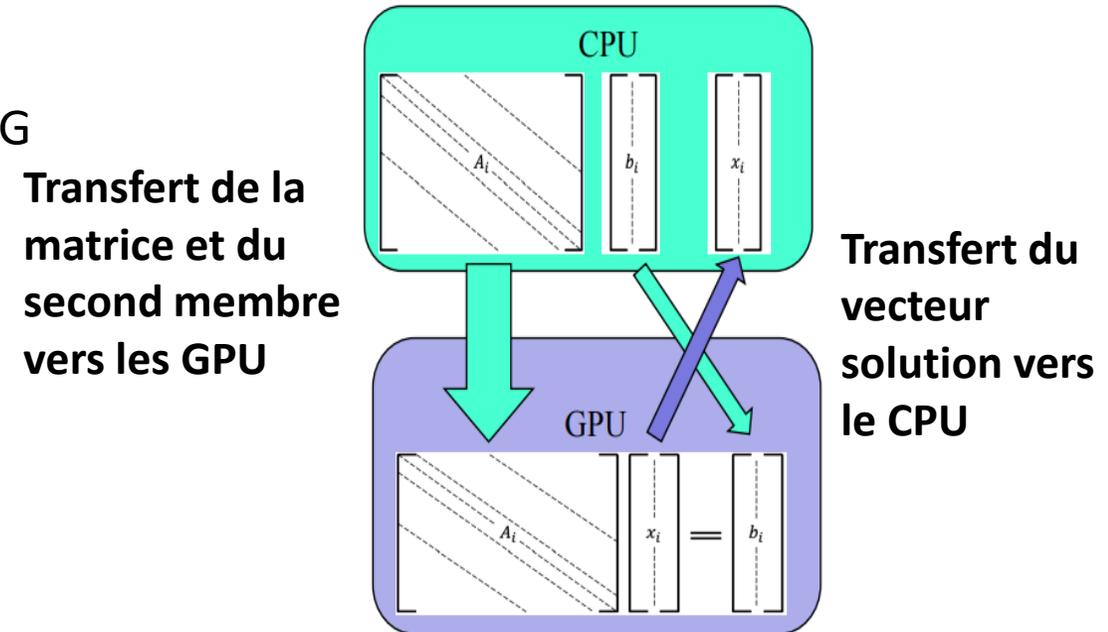
CPU

1.4 – La bibliothèque AmgX



- Bibliothèque **libre** de Nvidia pour la résolution de systèmes linéaires sur CPU et GPU
- **Flexible, Performante** et **large choix** pour la construction de solveurs linéaires sur GPU : toutes les méthodes peuvent apparaître comme un solveur, un préconditionneur ou un lisseur
 - Méthodes **itératives simples** : Gauss-Seidel, Jacobi
 - Méthodes de **Krylov** : GMRES, CG, BICGStab, ...
 - Méthodes **Multi Grille Algébrique** : Classical AMG, AGG AMG

- Implémentation dans JADIM et premiers résultats :
@JCAD 2019 A. Pedrono



Résolution du système linéaire par AmgX

1.5 – Objectifs de l'étude

- Définir les conditions optimales d'utilisation de cette parallélisation hybride MPI-GPU sur supercalculateurs en nous intéressant à l'influence :
 - i. des choix de configurations de la librairie AmgX
 - ii. des conditions aux limites
 - iii. du nombre de tâches MPI à distribuer par GPU à maillage fixé
- Evaluation des performances de la parallélisation hybride sur des maillages à 10 et 80 millions de mailles
- Valider le portage sur GPU avec AmgX pour les modules physiques de JADIM : IBM, VOF, Level-Set, LES, thermique, chimie, IBM+DEM, ...

2.1 – Les résultats

Présentation du supercalculateur Olympe de CALMIP



Supercalculateur OLYMPE

[site:<https://www.calmip.univ-toulouse.fr/>]

Deux partitions CPU

- 360 nœuds bi-socket (2 sockets de 18 cœurs)
 - 192 Go de mémoire par nœud
- 2 nœuds de large mémoire (1 socket de 18 cœurs)
 - 750 Go de mémoire par nœud

« nœud »

« mesca »

Une partition GP-GPU:

- 12 nœuds GPU (2 sockets de 18 cœurs + 4 cartes GPU)
 - 377 Go de mémoire par nœud

« volta »

2.2 – Configuration des tests

Configuration des tests :

Cas test : diphasique (VoF) - BULLE_3D

Maillage : 216x216x216

CPU : JADIM + **INTEL** + PETSc / **16 tâches MPI**

GPU Directe

GPU : JADIM + **PGI** + AmgX / **16 tâches MPI + 1 GPU**

Temps calculé sur **4 pas de temps** entre les itérations 4 et 7 sur un **nœud exclusif**

2.3 – Influence du choix de la configuration d'AmgX

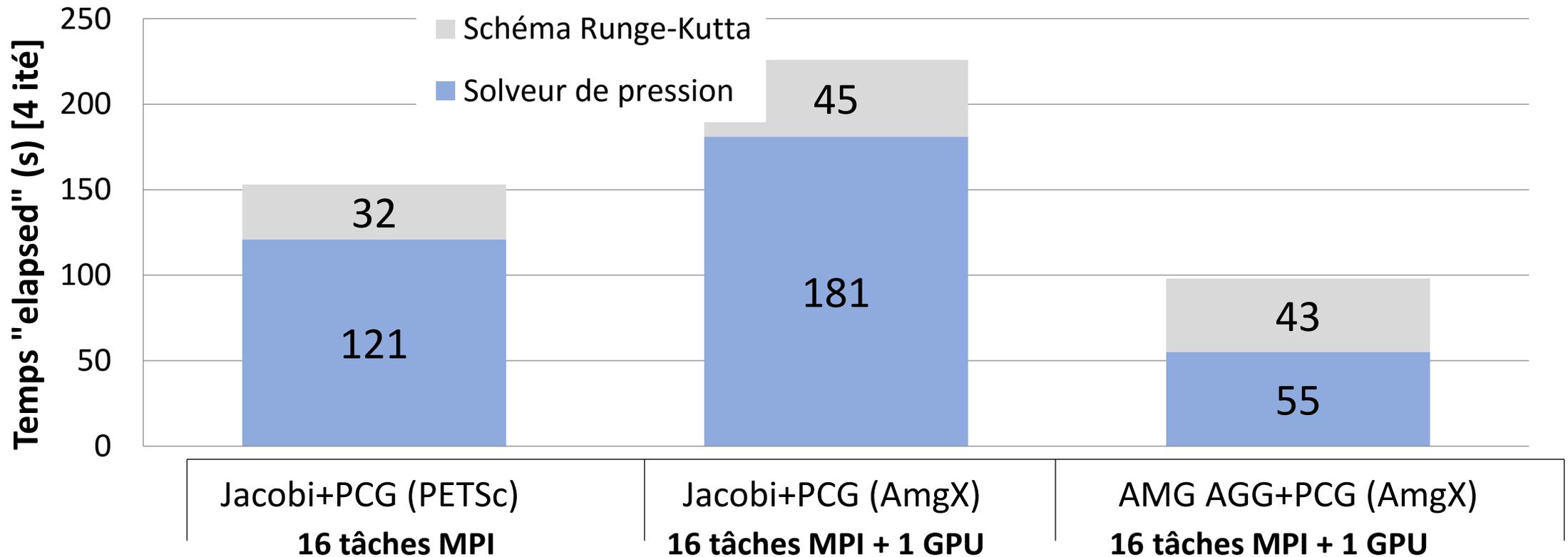
Tests sur les préconditionneurs + solveurs :

- Référence : **PETSc** avec **Jacobi + PCG** sur la **partition CPU**
- Challengeurs : **AmgX** avec **Jacobi + PCG** sur la **partition GPU**
AmgX avec **AMG-AGG + PCG** sur la **partition GPU**

2.3 – Influence du choix de la configuration d'AmgX

Tests sur les préconditionneurs + solveurs :

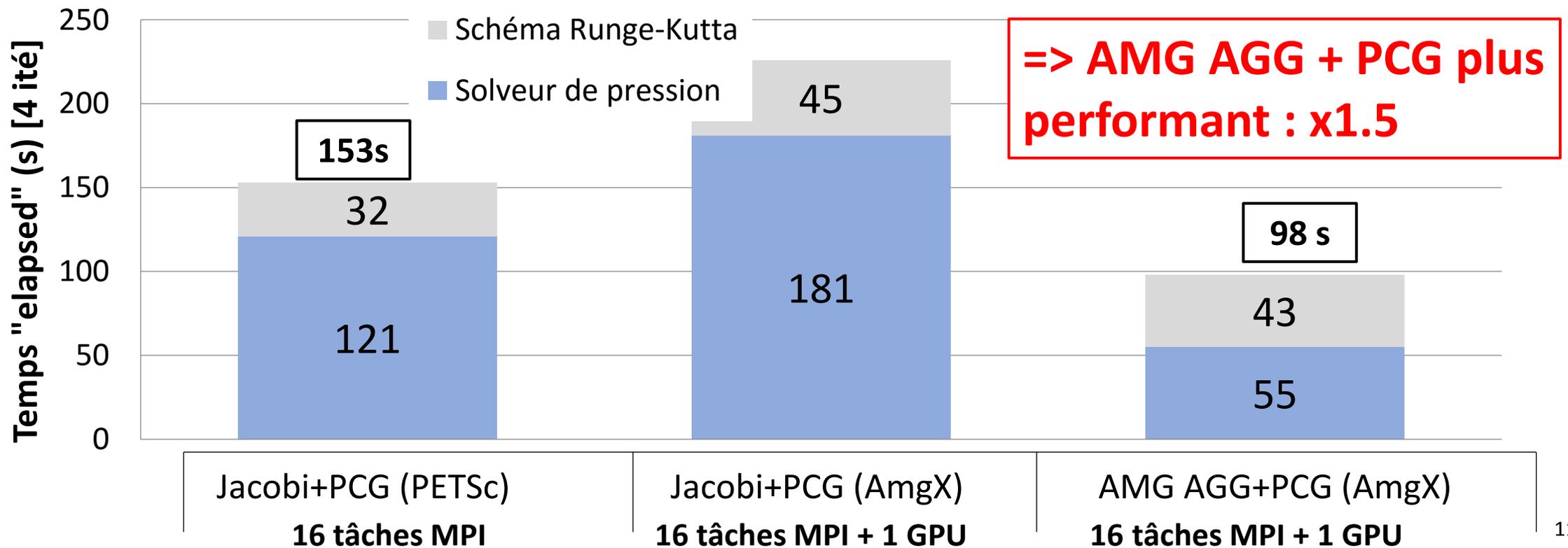
- Référence : PETSc avec Jacobi + PCG sur la **partition CPU**
- Challengeurs : AmgX avec **Jacobi + PCG** sur la **partition GPU**
AmgX avec **AMG-AGG + PCG** sur la **partition GPU**



2.3 – Influence du choix de la configuration d'AmgX

Tests sur les préconditionneurs + solveurs :

- Référence : PETSc avec Jacobi + PCG sur la **partition CPU**
- Challengeurs : AmgX avec **Jacobi + PCG** sur la **partition GPU**
AmgX avec **AMG-AGG + PCG** sur la **partition GPU**



2.4 – Influence des conditions aux limites

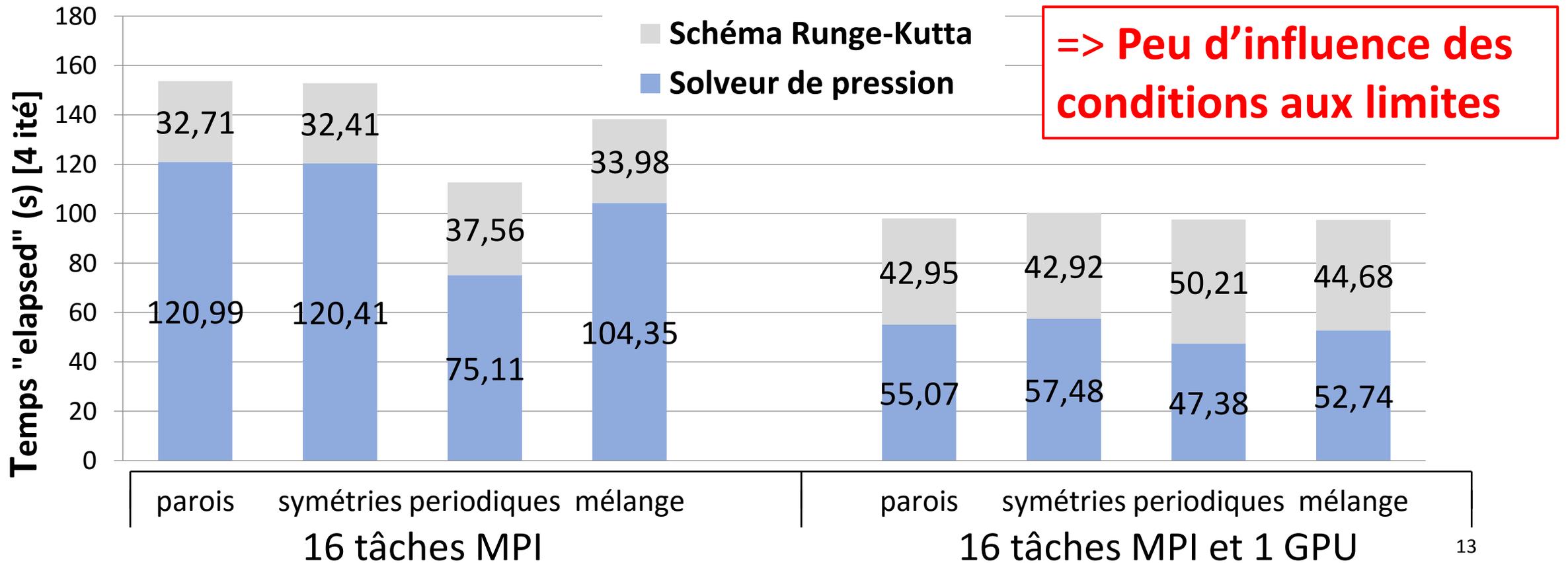
- **Paroi** : vitesse nulle au niveau de la frontière
- **Périodique** : vitesse identique sur ces frontières
- **Symétrique** : gradient normal à la vitesse nul
- **Mélange** : Périodique suivant X, Paroi suivant Y et Symétrique suivant Z

Configuration AmgX : AMG AGG + PCG

2.4 – Influence des conditions aux limites

- **Paroi** : vitesse nulle au niveau de la frontière
- **Périodique** : vitesse identique sur ces frontières
- **Symétrique** : gradient normal à la vitesse nul
- **Mélange** : Périodique suivant X, Paroi suivant Y et Symétrique suivant Z

Configuration AmgX : AMG AGG + PCG



2.5 – Influence du nombre de tâches MPI par GPU

Configuration du test :

4 GPUs / Variation des tâches MPI : 4 / 8 / 16 / 32 / 36 – sur nœuds GPU

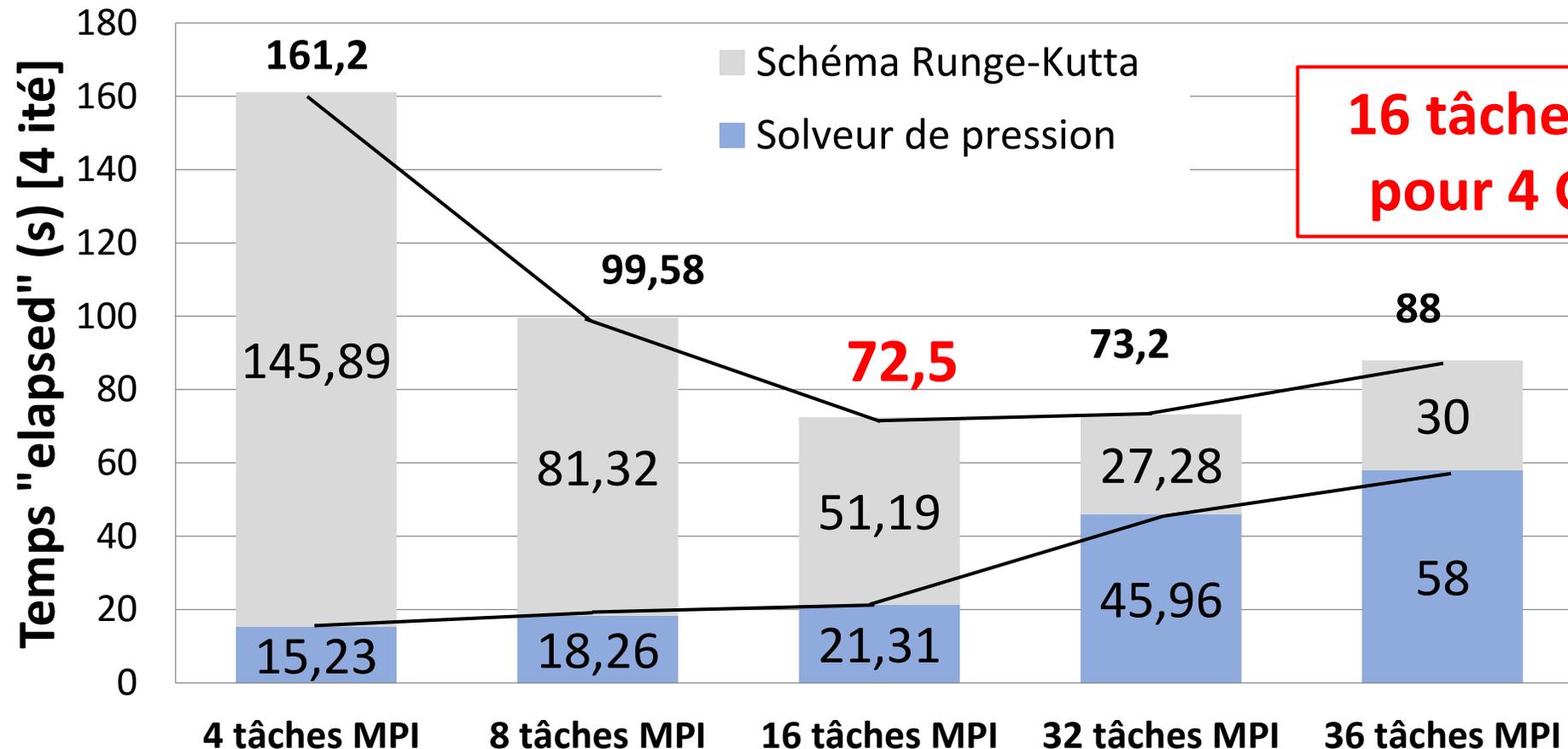
AmgX : AMG - AGG + PCG - Conditions aux limites : parois

2.5 – Influence du nombre de tâches MPI par GPU

Configuration du test :

4 GPUs / Variation des tâches MPI : 4 / 8 / 16 / 32 / 36 – sur nœuds GPU

AmgX : AMG - AGG + PCG - Conditions aux limites : parois



2.6 – Performance – time is money

Configuration du test :

Cas tests utilisés :

BULLE_3D : cas diphasique (méthode VoF)

IBM_SEDIMENTATION : cas monophasique avec frontière immergé (méthode IBM)

Maillage :

Cartésien orthogonal régulier

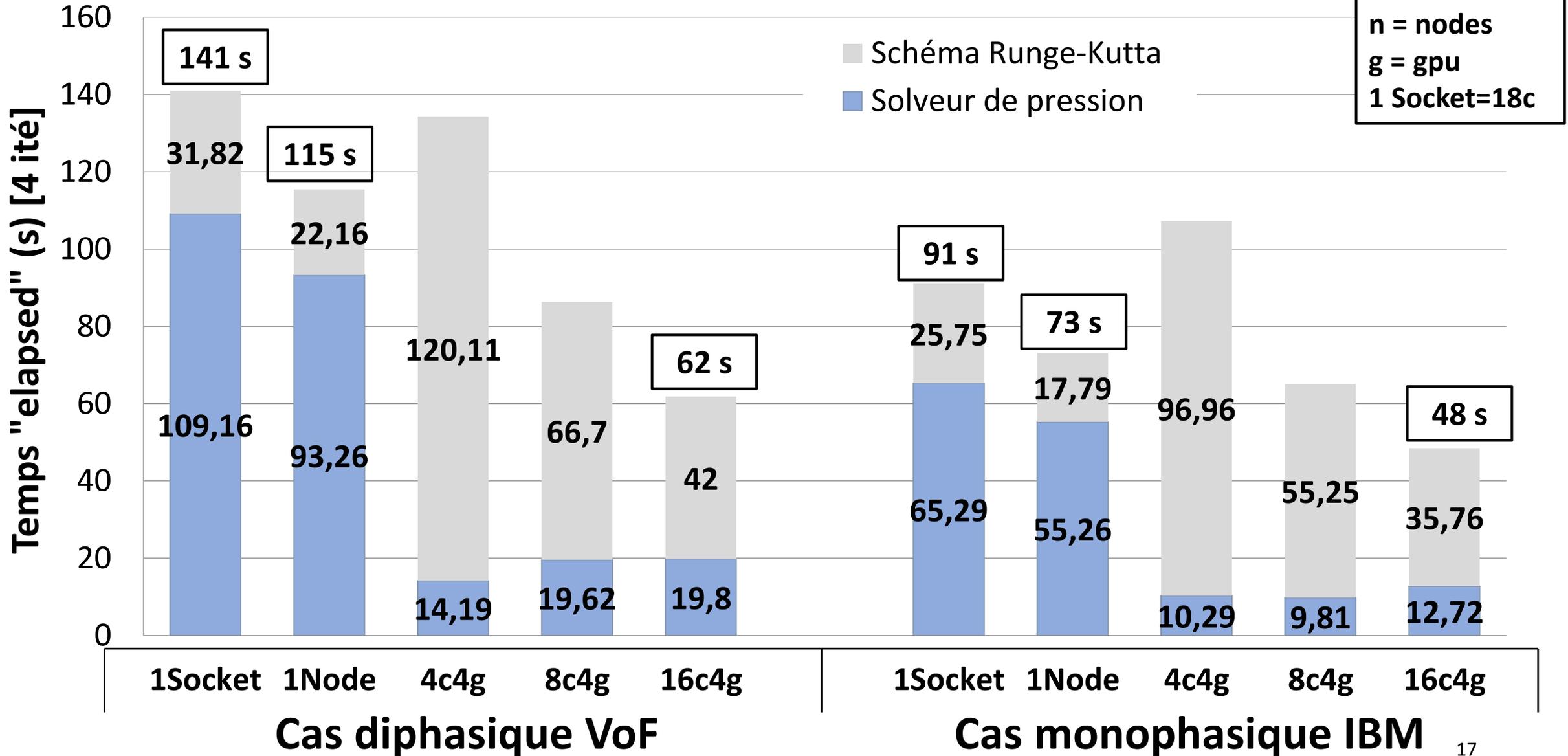
$216^3 / 432^3$

AmgX : AMG-AGG + PCG

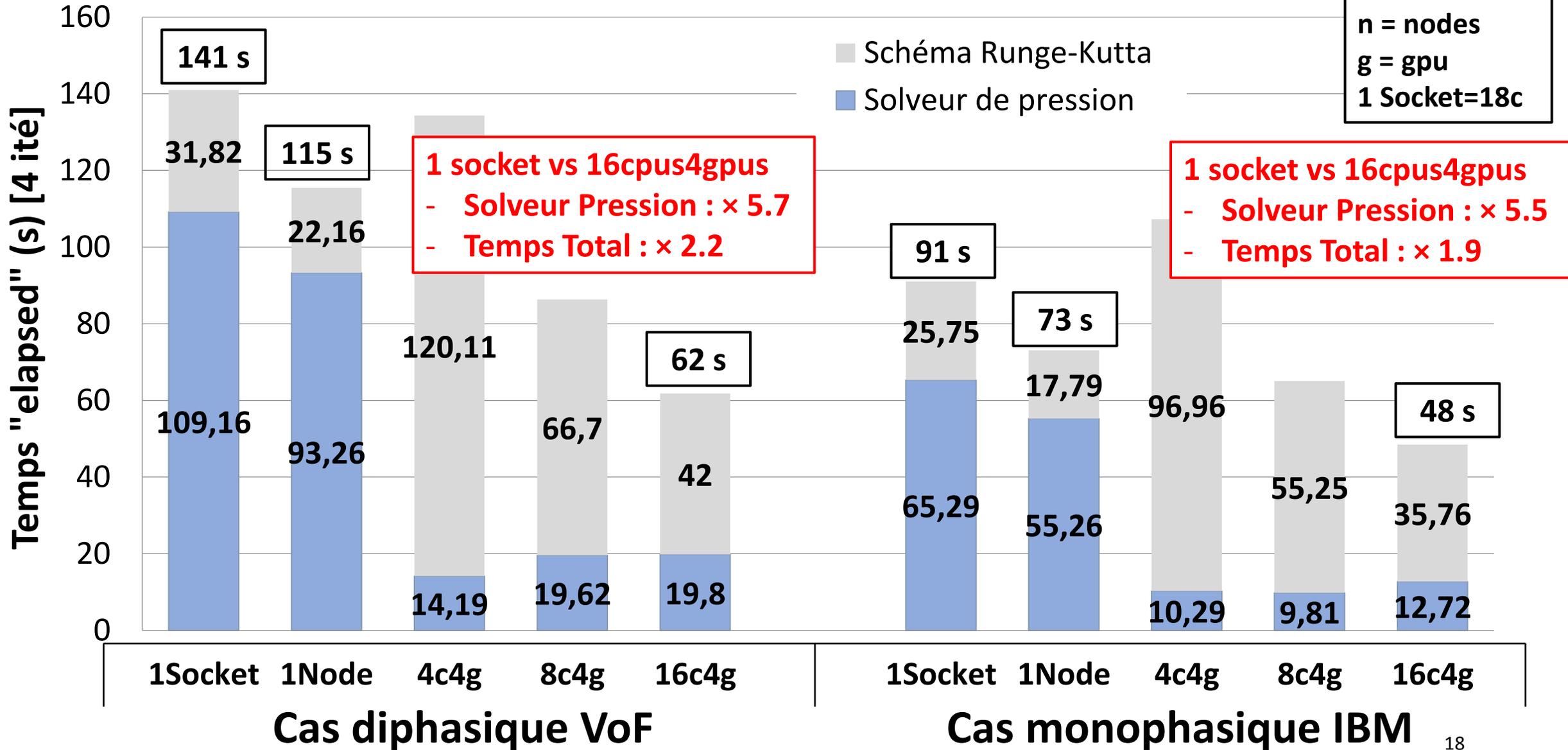
Conditions aux limites : Parois

Temps calculé sur **4 pas de temps** entre les itérations 4 et 7 sur un **nœud exclusif**

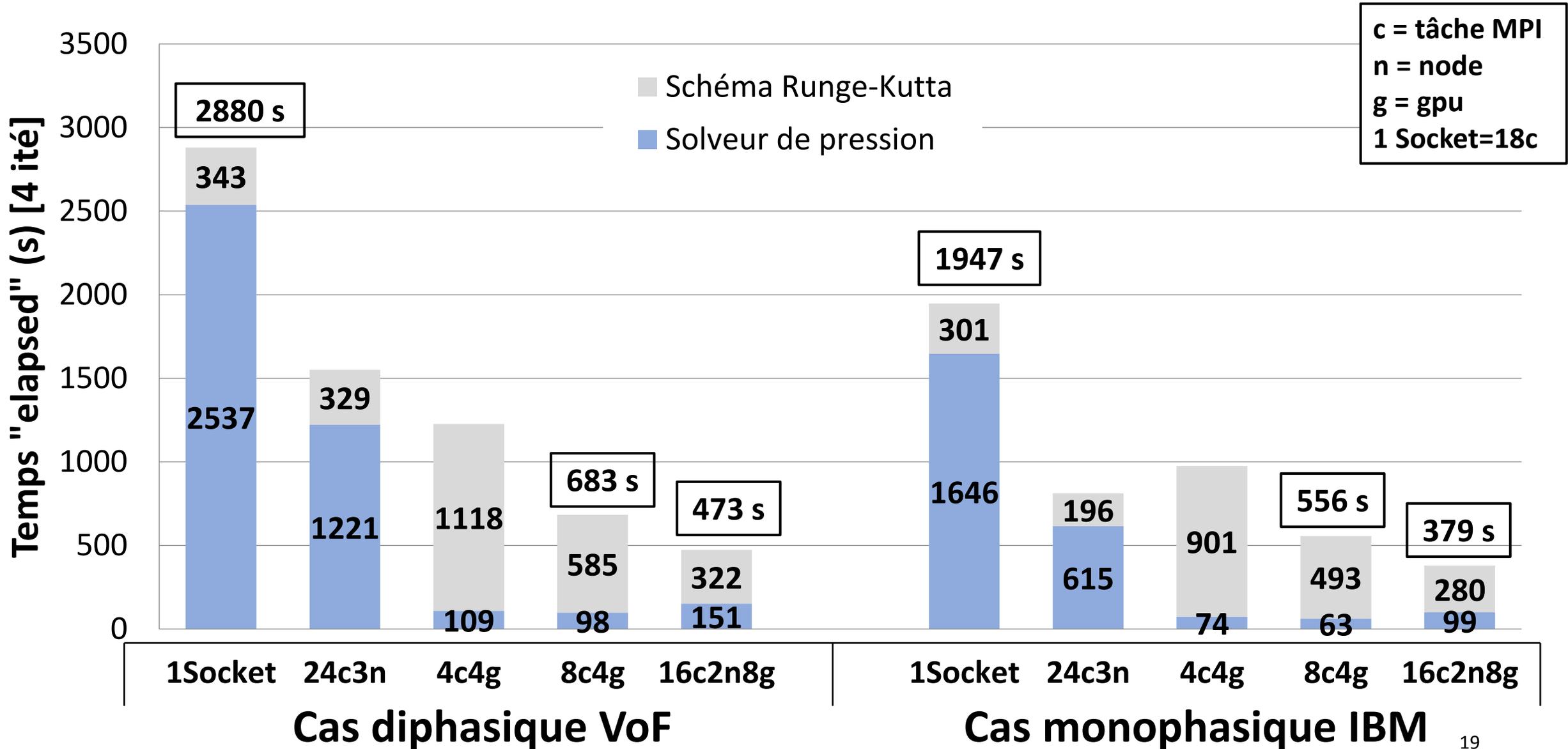
Maillage : 10 Millions de cellules - 216^3



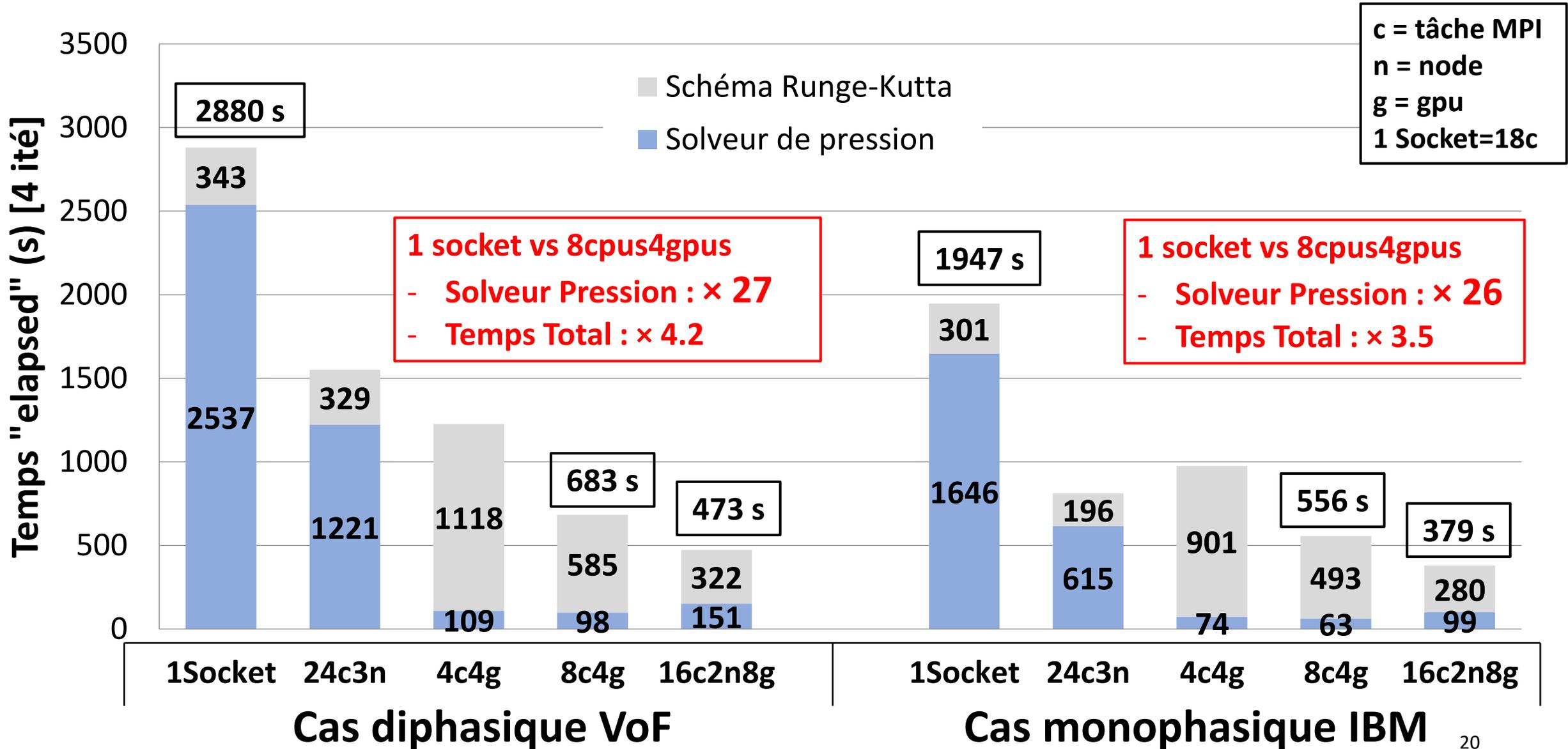
Maillage : 10 Millions de cellules - 216^3



Maillage : 80 Millions de cellules - 432^3



Maillage : 80 Millions de cellules - 432^3



2.7 – Validation des modules physiques de JADIM avec AmgX

Comparaison des solutions physiques CPU vs CPU/GPU :

JADIM+Intel+PETSc vs JADIM+PGI+AmgX

Module(s) associé(s)	Cas testés	Maillage	Solveur de Pression sous CPU
IBM	IBM_SEDIMENTATION	32^3	PETSc
VOF	BULLE_3D	32^3	PETSc
Level-set	PARASITE_3D	32^3	PETSc
Chimie / LES	BULLES_TDM_COUPLAGE	12x12x10	Fourier
LES	CANAL	52x66x66	PETSc
Chimie	CHIMIE-2PHASES-1ordre	32x62x1	PETSc
Thermique	LS_BULLE_3D_SURFACTANT	152x60x60	PETSc
Thermique	LC80	100x56x34	Fourier
IBM+DEM	PARASITE_3D	216x216x216	PETSc

Conclusion

- Valider l'implémentation d'un solveur de la bibliothèque AmgX sur GPU pour de nombreux modules physiques du code JADIM
- Définit les conditions optimales d'utilisation de cette parallélisation hybride MPI/GPU sur supercalculateur
- Evaluer les performances de la parallélisation hybride MPI-GPU avec un **gain de temps observé jusqu'à 27** sur un maillage à 80 millions de cellules (432^3) pour le solveur de pression

Merci @Team CALMIP pour les heures de calculs sur le projet p19046

Merci @NVIDIA pour l'aide au portage sur GPU avec AmgX

Merci @ATOS pour la compilation d'AmgX2.2 et ses meilleures performances