
Parallélisme avancé en dynamique rapide des fluides et des structures

Vincent Faucher
(CEA Saclay, LaMSID)

H. Bung, P. Galon, F. Bliard, A. Beccantini (CEA)
F. Casadei (JRC Ispra)

Introduction

- Dynamique non-linéaire des fluides et des structures
- Intégration explicite en temps

Parallélisme à mémoire distribuée avec un schéma explicite

- Décomposition de domaine
- Nature des frontières
- Equilibrage de la charge
- Gestion de la stabilité conditionnelle
- Etat des lieux dans EUROPLEXUS

Connexions cinématiques

- Nature des connexions
- Gestion parallèle – 1. Ecriture des connexions
- Gestion parallèle – 2. Résolution

Vers un parallélisme hybride, collaboratif et générique

- Parallélisme SMP adaptatif
- Hybridation
- Utilisation des GPU

Crash et ruine des structures

- Grands déplacements, grandes rotations
- Contact et auto-contact
- Rupture et fragmentation (Eléments Finis, SPH pour solides et coques, Eléments Discrets)



Crash d'un corps creux à haute vitesse



Impact sur une poutre en béton armé

Explosions en eau et en air

- Ecoulements à topologie complexe (Eléments Finis, Volumes Finis d'ordre élevé)
- Ecoulements réactifs (Volumes Finis)
- Ecoulements multicomposants/suivi d'interface/changement de phase
- Interaction fluide-structure avec rupture et fragmentation



Explosion dans une cuve avec structures internes



Interaction d'une marche avec un écoulement supersonique



Pulsation d'une bulle comprimée dans un liquide

Conservation de la quantité de mouvement (Eléments Finis ou particules)

$$\mathbf{M}^{n+1} \ddot{\mathbf{U}}^{n+1} + \mathbf{F}_i \mathbf{U}^n, \dot{\mathbf{U}}^{n+1/2} + \mathbf{F}_{link}^{n+1} + \mathbf{F}_{transport}^{n+1} = \mathbf{F}_{ext}^{n+1}$$

$$\mathbf{C}^{n+1} \dot{\mathbf{U}}^{n+3/2} = \mathbf{B}^{n+1}$$

+ Equations de conservation au format Volume Fini (masse, énergie) pour les fluides

Intégration en temps

$$\mathbf{U}^{n+1} = \mathbf{U}^n + \Delta t \cdot \dot{\mathbf{U}}^{n+1/2} \quad (\text{structure seulement})$$

$$\dot{\mathbf{U}}^{n+1/2} = \dot{\mathbf{U}}^{n-1/2} + \Delta t \cdot \ddot{\mathbf{U}}^n$$

Caractéristiques principales

Pas d'inversion matricielle globale

Structure de données vectorielle

Localité des données en mémoire

Stabilité conditionnelle

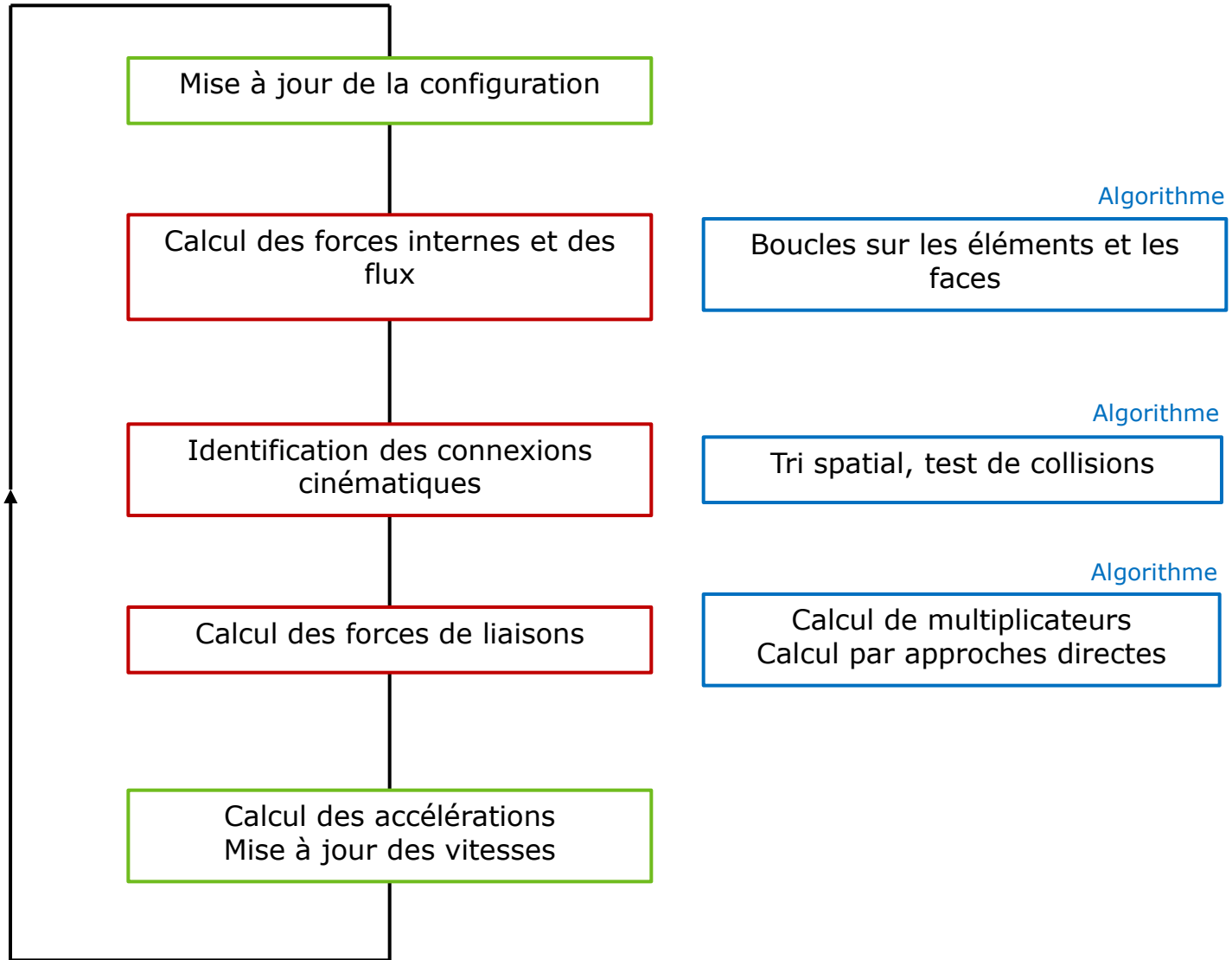
$$\Delta t < \frac{l_{\min}}{c} \quad \begin{array}{l} \text{Grand nombre de cycles} \\ \text{Problématique 4D lors du raffinement du maillage} \end{array}$$

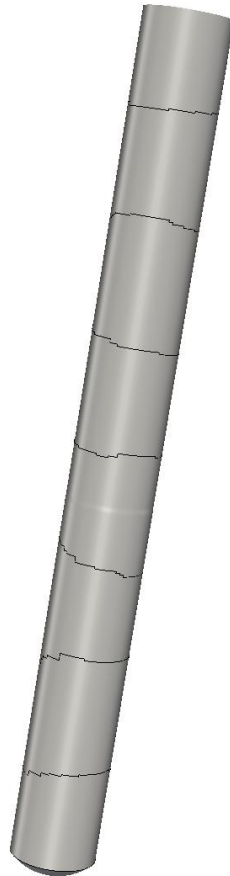
Gestion robuste des non-linéarités (bifurcations...)

Flambage, contact et frottement (...)



Boucle sur le temps

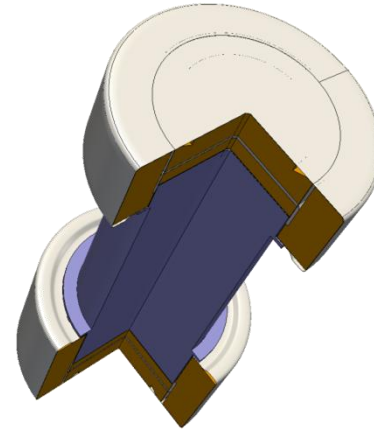




ParMetis

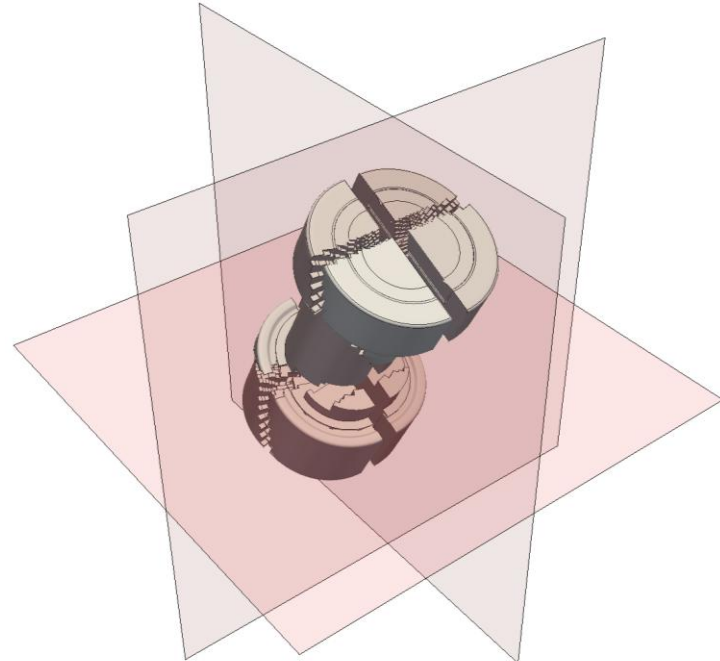
<http://glaros.dtc.umn.edu/gkhome/metis/parmetis/>

- ✓ Minimisation de la taille des frontières
- ✗ Recours exclusif à la connectivité



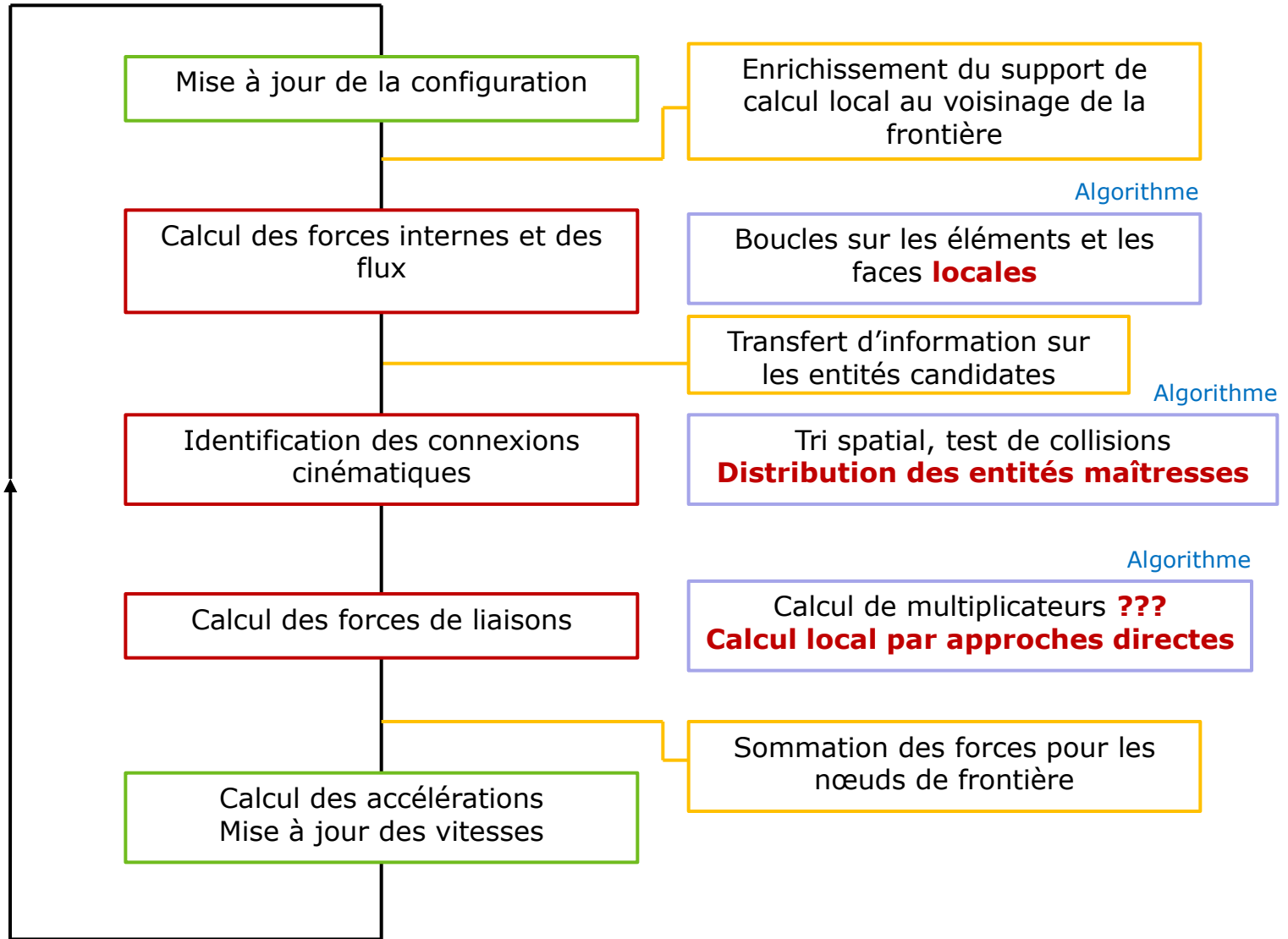
**Recursive
Orthogonal
Bisection**

- ✓ Meilleure localisation
- ✓ Découpage de tous les types de modélisation
- ✗ Pas de contrôle de la taille des frontières



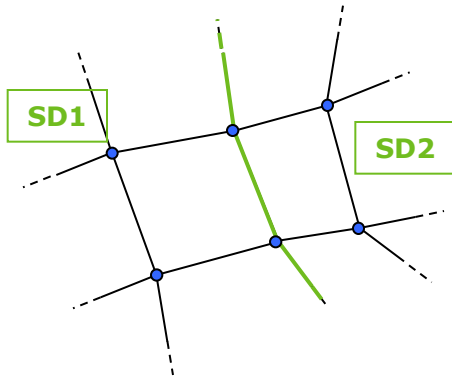


Boucle sur le temps

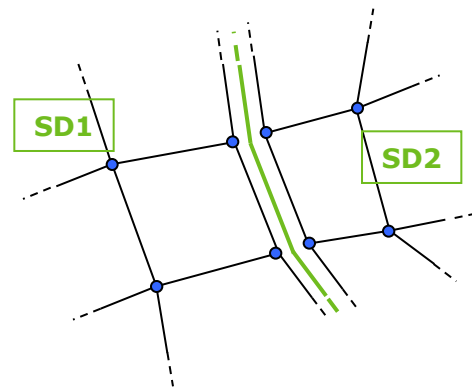




Frontière nodale

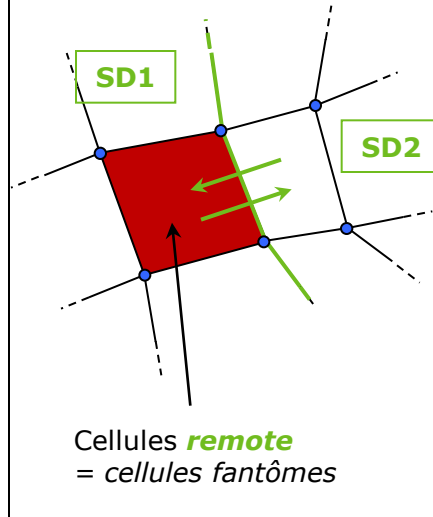


(a) Sommation des contributions nodales



(b) Liaisons de continuité à l'interface

Frontière par face (calcul de flux)

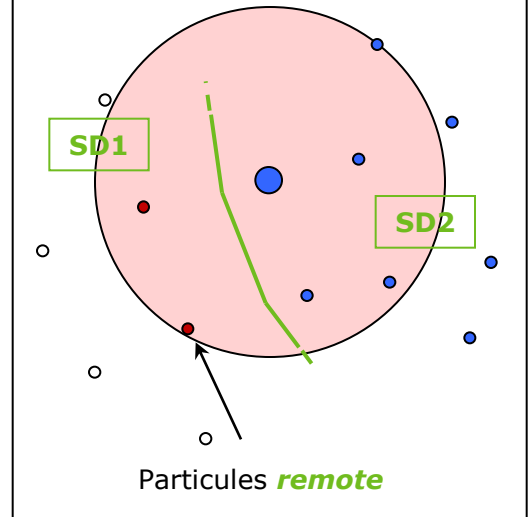


Entité *remote*

Entité n'appartenant pas à un sous-domaine nécessaire à l'évaluation correcte des grandeurs sur les entités du sous-domaine

Règle de base : on ne calcule pas sur les entités *remote*, [quoique...](#)

Frontière par voisinage (particules)



Principe

Pondération des éléments lors de la décomposition de domaine

Critères de pondération

- Formulation des éléments (nombre de points d'intégration, repère local...)
- Loi de comportement (non-linéarité, itérations internes...)
- ...

Calcul des poids

Calcul *a priori*

Analyse détaillée des opérations élémentaires et dans la loi de comportement

- ☑ Equilibrage correct sans intervention de l'utilisateur
- ☒ Difficile à mettre en œuvre pour l'intégralité des modélisations du programme
- ☒ Ne prend pas en compte les effets de structure de données et de programmation

Calcul *a posteriori*

Mesure des coûts lors de l'exécution du programme

- ☑ Analyse simplifiée
- ☑ Mesure indifférente des éléments et des matériaux
- ☒ Premiers *runs* avec une décomposition par défaut non-optimale

Constitution possible d'une base de données pour capitaliser les travaux analytiques ou les expériences numériques

Problème principal : gestion des **autres centres de coût** que les éléments

Objectif

Eviter la **limitation globale du pas de temps** par le plus petit élément du maillage

2 approches

Un pas de temps par sous-domaine

[Gravouil & Combescure 2000]

[Bousquet 2000][Herry, Faucher 2001]

Ecriture généralisée pour le couplage de schémas d'intégration

- ✓ Utilisé dans le cadre d'études qualifiées
- ✗ Manque de généralité pour les connexions cinématiques inter-domaines
- ✗ Force l'utilisation de multiplicateurs de Lagrange pour les connexions de frontière nodales
- ✗ Non-fonctionnel pour le fluide
- ✗ Utilisation restreinte pour le couplage de schémas explicites seuls
- ✗ Optimisation difficile de la décomposition de domaine

Partitionnement des éléments

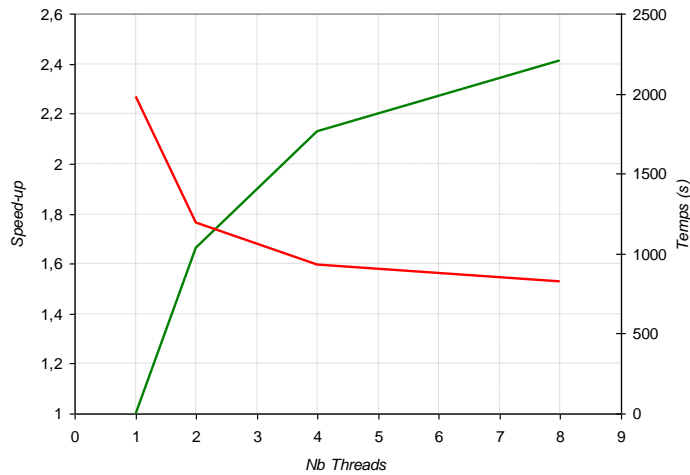
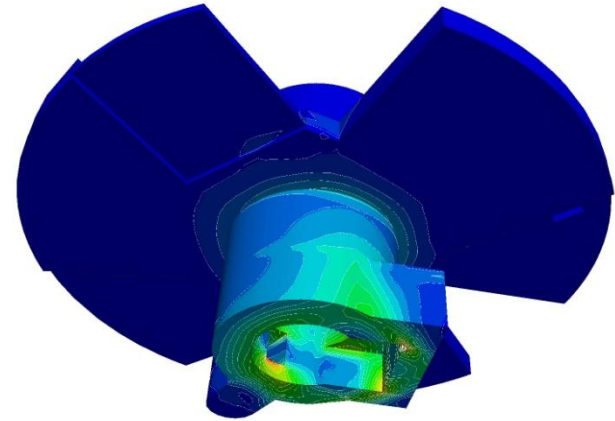
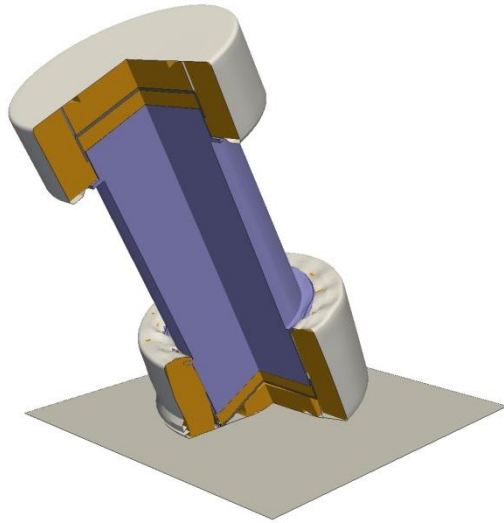
[Halleux, Casadei 2006]

Répartition des éléments dans une hiérarchie de classes en fonction de leur pas de stabilité propre

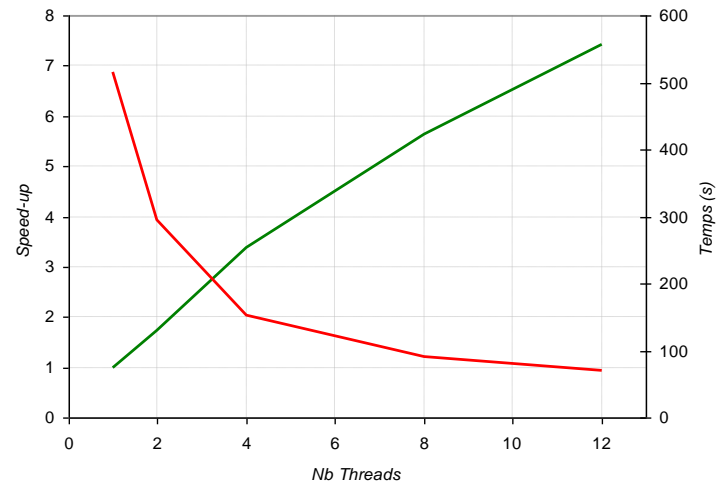
En descendant la hiérarchie, chaque classe est traitée deux fois plus souvent que la précédente

- ✓ Développement fonctionnel avec la majorité des connexions cinématiques inter-domaines ou non
- ✓ Fonctionnel avec le fluide EF
- ✓ Ajout d'un critère local aux éléments pour la pondération de la décomposition de domaine
- ✗ Peu éprouvé en contexte industriel
- ✗ Fluide VF et particules non-prises en compte

Extensibilités caractéristiques (1)

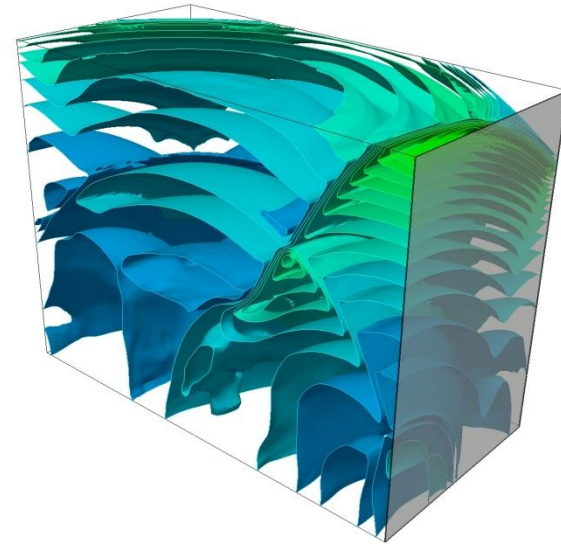
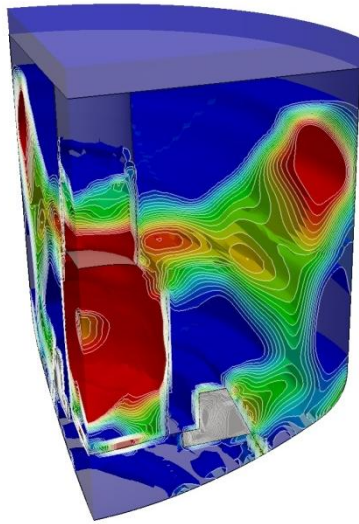


(a) Chute oblique d'un château de transport



(b) Mise en pression du puits de cuve d'un REP

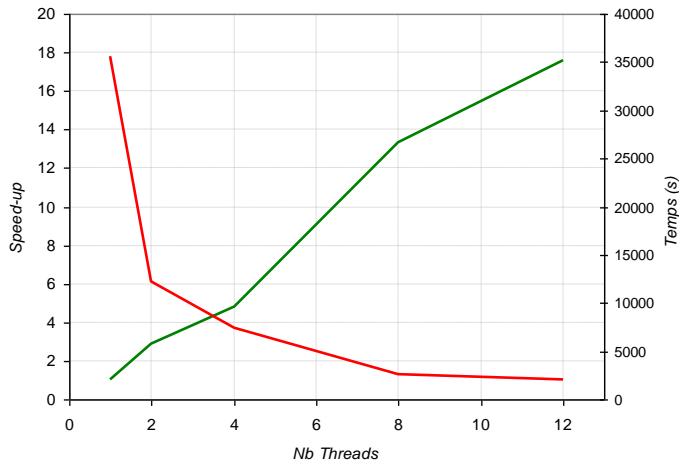
Extensibilités caractéristiques (2)



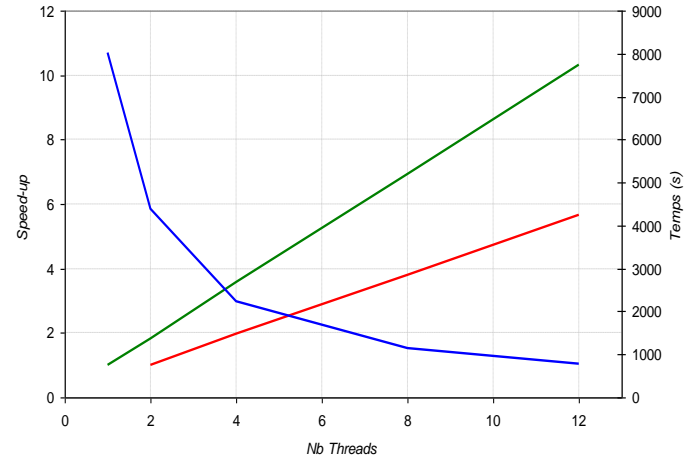
Instabilité RM 2D



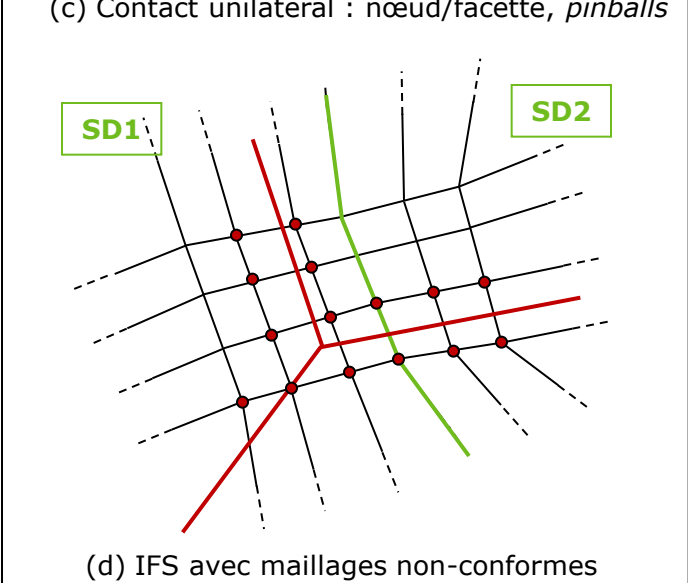
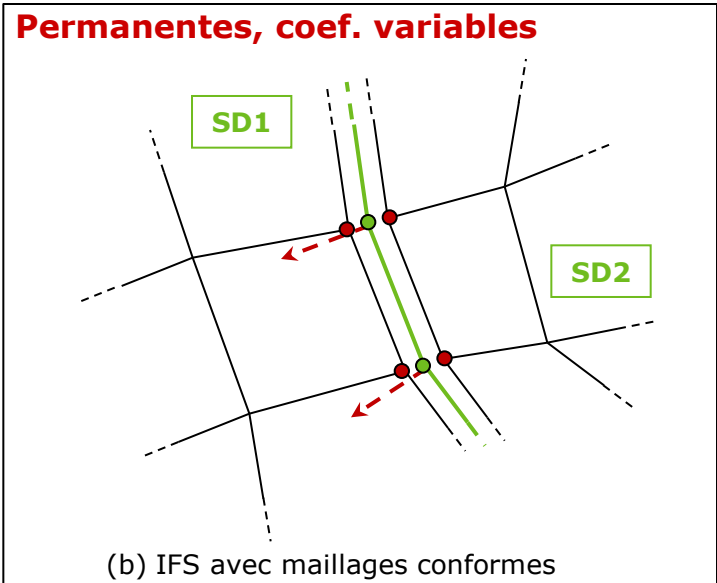
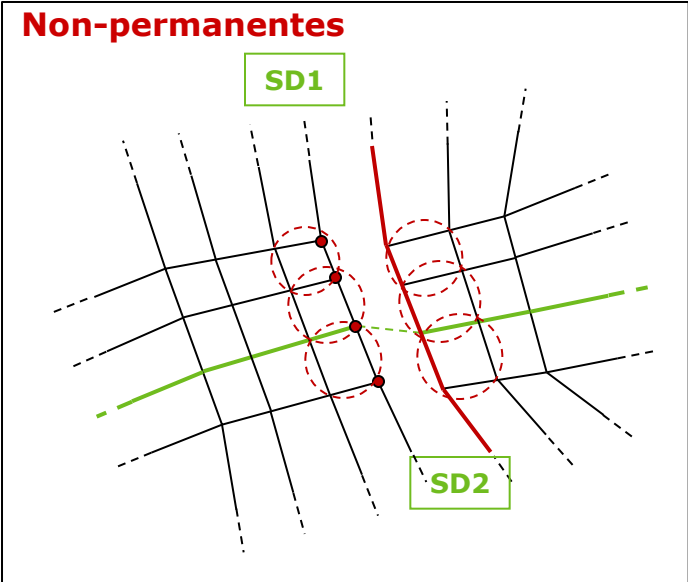
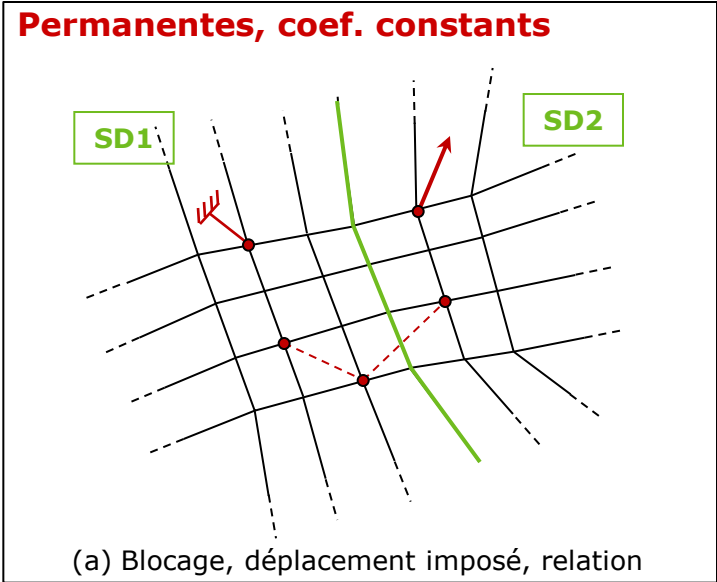
Instabilité RM 3D

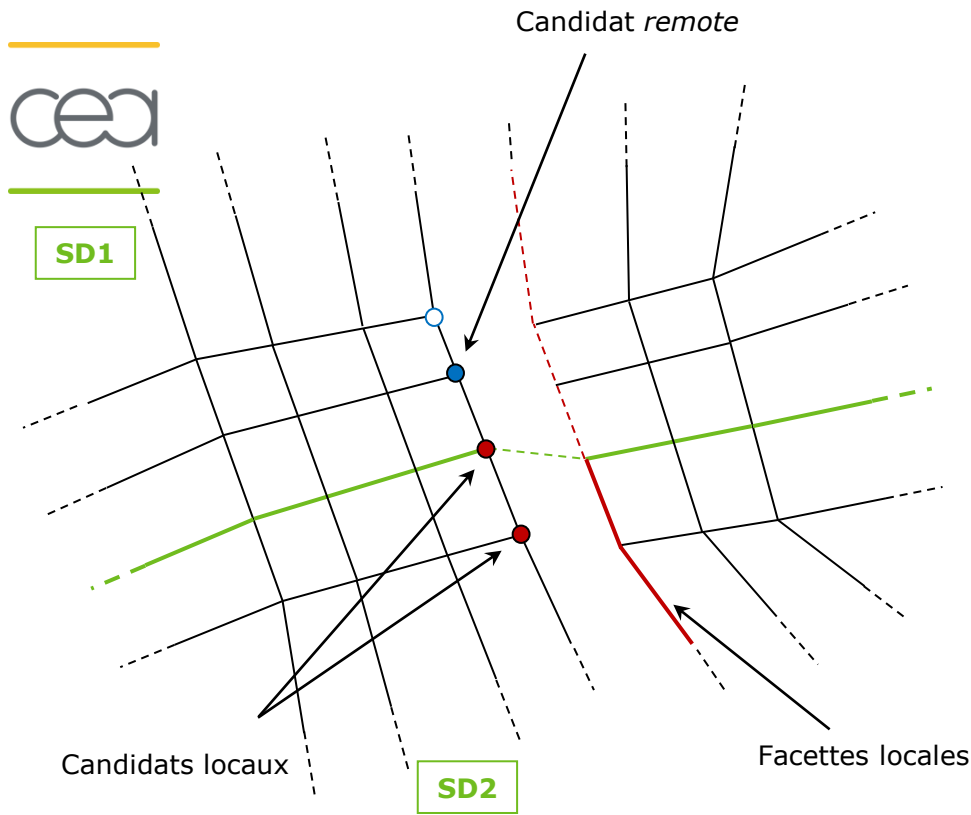


(c) Essai MARA 10

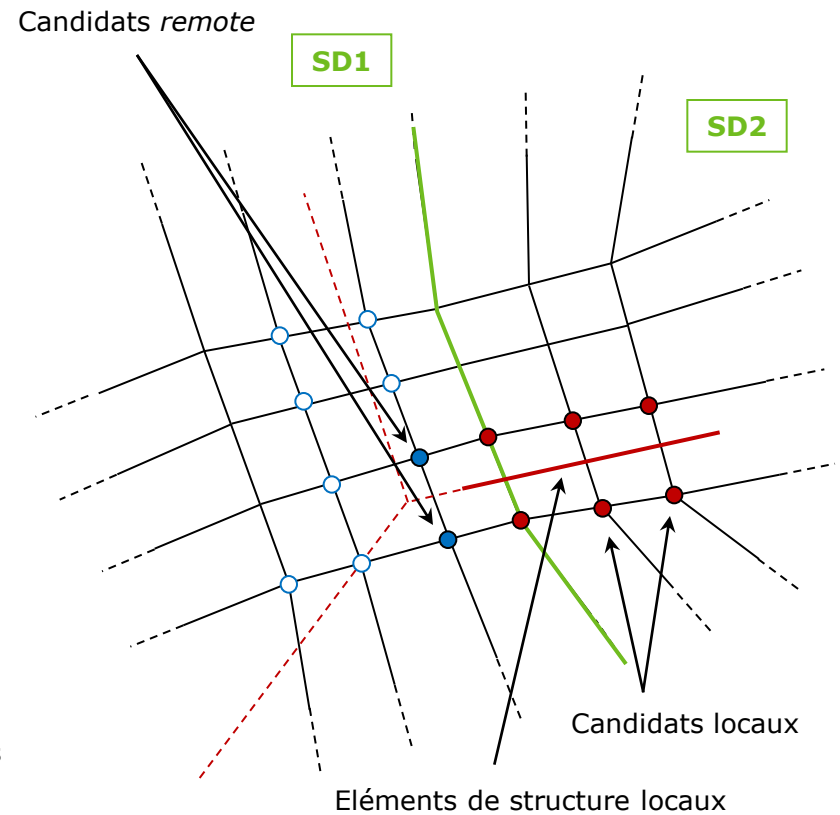


(b) VFCC - Explosion à proximité d'un mur rigide





(a) Contact unilatéral



(b) IFS avec maillages non-conformes

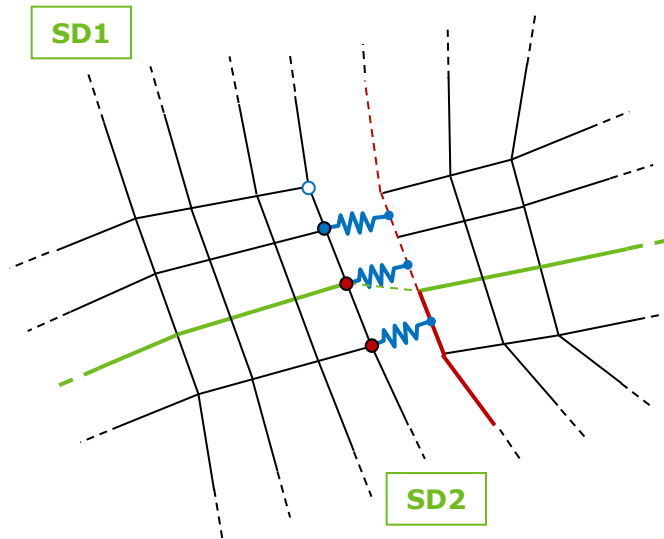
Accélération parallèle

Distribution des entités maîtresses

Filtrage indispensable des candidats *remote* \Leftrightarrow frontière par voisinage



Pénalisation



Forces de connexion analogue aux forces élémentaires

- ✓ Pas de système à résoudre
- ✓ Traitement parallèle = simple sommation sur les nœuds *remote*
- ✗ Choix des coefficients de pénalité

Dualisation

Résolution générale

$$\mathbf{M}^{n+1} \ddot{\mathbf{U}}^{n+1} + \mathbf{F}_i \mathbf{U}^n, \dot{\mathbf{U}}^{n+1/2} + \mathbf{C}^{n+1T} \boldsymbol{\Lambda}^{n+1} = \mathbf{F}_{ext}^{n+1}$$

$$\mathbf{C}^{n+1} \dot{\mathbf{U}}^{n+3/2} = \mathbf{B}^{n+1}$$

⇒

$$\mathbf{C}^{n+1} \mathbf{M}^{n+1^{-1}} \mathbf{C}^{n+1T} \boldsymbol{\Lambda}^{n+1} = \mathbf{C}^{n+1} \mathbf{M}^{n+1^{-1}} \mathbf{F}_{ext}^{n+1} - \mathbf{F}_i^{n+1} - \frac{1}{\Delta t} \mathbf{B}^{n+1} - \mathbf{C}^{n+1} \dot{\mathbf{U}}^{n+1/2}$$

Système condensé sur les multiplicateurs

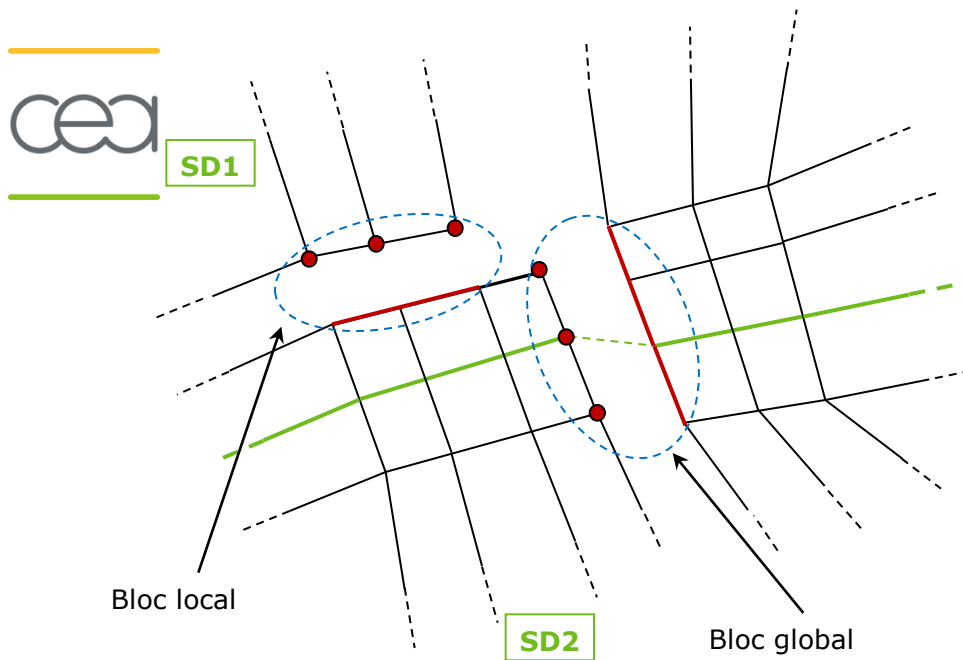
- ✓ Pas de paramètres additionnels
- ✓ Résolution exacte des équations de connexion
- ✗ Système matriciel à résoudre

Autres approches directes

IFS faible pour les VF et EF...

- ✓ Traitement économique sans paramètre additionnel
- ✗ Invalide si couplage entre connexions

Blocs dans l'opérateur de condensation



Structure de l'opérateur (taille, remplissage, coefficients) **variable** dans le temps

↓
Persistence... (1)

Stratégies avec assemblage du problème condensé

Identification *a priori* des blocs diagonaux

Structure de données dynamique pour l'ajout/suppression de connexions et l'identification des couplages

Listes chaînées à deux niveaux

1 par sous-domaine
1 liste pour les blocs globaux

- ✓ Largeur de bande minimale pour l'opérateur de condensation
- ✓ Parallélisme direct pour les blocs locaux
- ✗ Localité des données en cache
- ✗ Traitement séquentiel de la liste globale
- ✗ Coût de la tâche d'identification des couplages

Assemblage d'un opérateur global

Identification des blocs diagonaux reportée au niveau du solveur

- ✓ Mise en œuvre possible d'un solveur distribué
- ✗ Extensibilité restreinte à celle du solveur

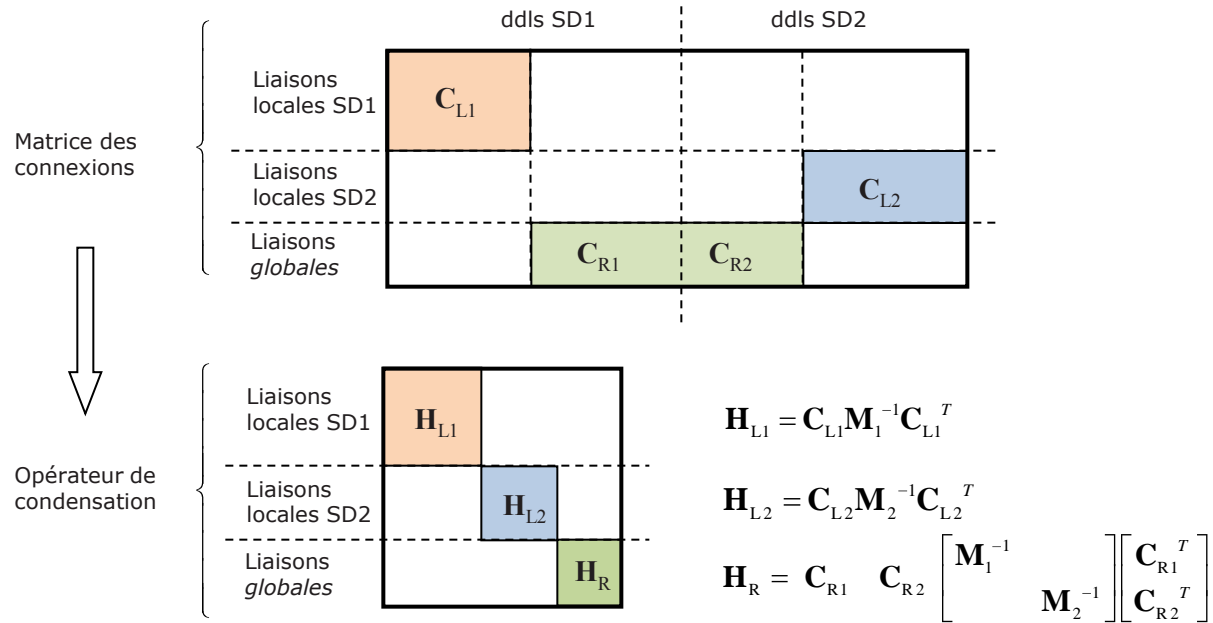
Approche mixte

1 bloc par sous-domaine + 1 bloc global

En cours de test
ANR REPDYN

Résolution sans assemblage du problème condensé

1. Organisation des connexions et des inconnues



2. Résolution **itérative du bloc global** sans assemblage

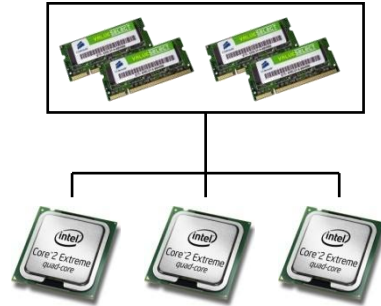
Préconditionnement : **résolution locale en négligeant les couplages au niveau global**

$$\begin{bmatrix} C_{R1} & C_{R2} \end{bmatrix} = \begin{bmatrix} C_{R1}^1 & C_{R2}^1 \\ C_{R1}^2 & C_{R2}^2 \end{bmatrix} \quad \text{Séparation des connexions globales écrites sur SD1 et SD2}$$

$$H_R = \begin{bmatrix} C_{R1}^1 & C_{R2}^1 \\ C_{R1}^2 & C_{R2}^2 \end{bmatrix} \begin{bmatrix} M_1^{-1} & \\ & M_2^{-1} \end{bmatrix} \begin{bmatrix} C_{R1}^1 T & C_{R2}^1 T \\ C_{R1}^2 T & C_{R2}^2 T \end{bmatrix} = \begin{bmatrix} H_R^1 & H_{RC} \\ H_{RC}^T & H_R^2 \end{bmatrix} \Rightarrow \tilde{H}_R = \begin{bmatrix} H_R^1 & 0 \\ 0 & H_R^2 \end{bmatrix} \quad \text{Assemblage local sur SD1 et SD2}$$

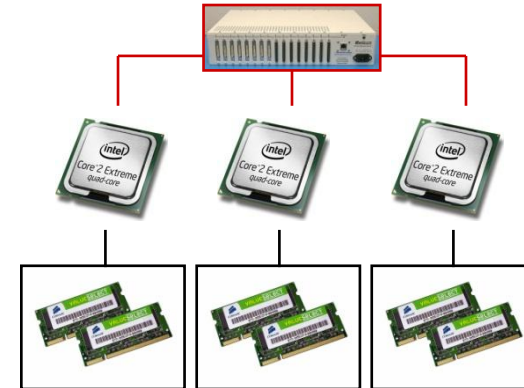


Machines à mémoire partagée (SMP)

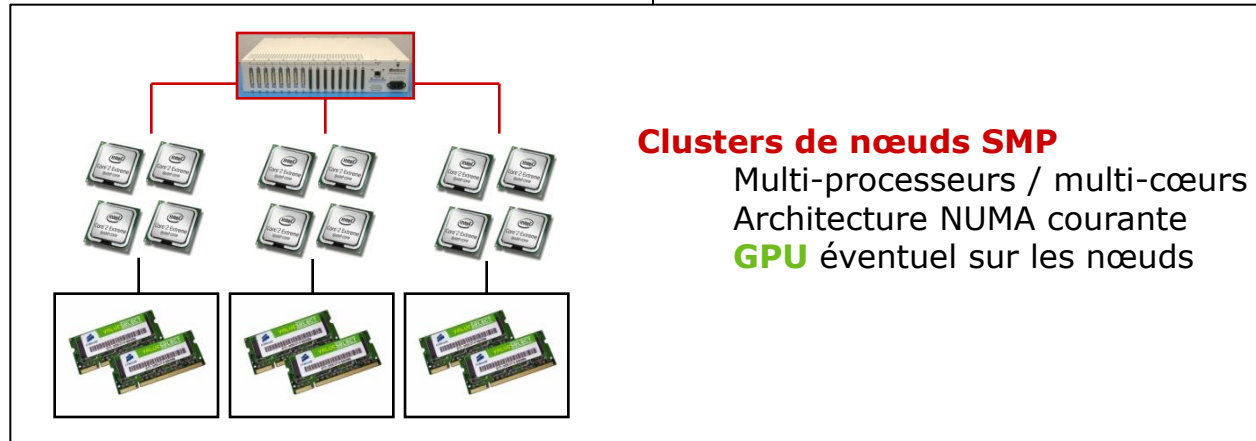


- ✓ Facilité de programmation (OpenMP)
- ✗ Coût et taille limite des machines
- ✗ Performances variables pour les accès mémoire
- ✗ Programmation parfois moins aboutie

Machines à mémoire distribuée



- ✓ Agglomérats de grande taille (*clusters*)
- ✓ Accès toujours local à la mémoire
- ✓ Programmation parallèle poussée...
... mais spécifique
- ✗ Définition *manuelle* des communications



Clusters de nœuds SMP

- Multi-processeurs / multi-cœurs
- Architecture NUMA courante
- GPU éventuel sur les nœuds

Recours à une librairie externe

X-KAAPI (INRIA LIG/MOAIIS)

[T. Gautier, X. Besson & L. Pigeon.
KAAPI : a thread scheduling runtime system for data flow computations on cluster of multi-processors.
In Parallel Symbolic Computation'07 (PASCO'07),
 number, 15-23, London, Ontario, Canada, 2007]
<http://kaapi.gforge.inria.fr>

Identification dans le programme de tâches et de leurs dépendances

- Persistence... (2)
- Processus de déclaration des tâches en cours de définition
 Stage MR2 D. Leone - INRIA/ENSIMAG
 - Dépendances identifiées automatiquement par X-KAAPI en fonction des accès aux variables

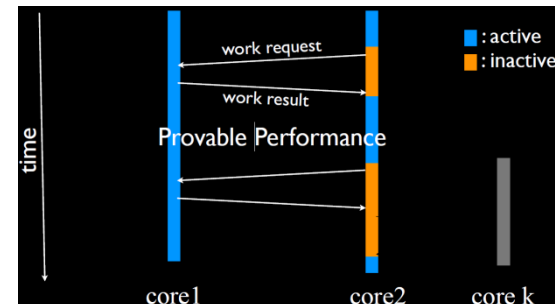
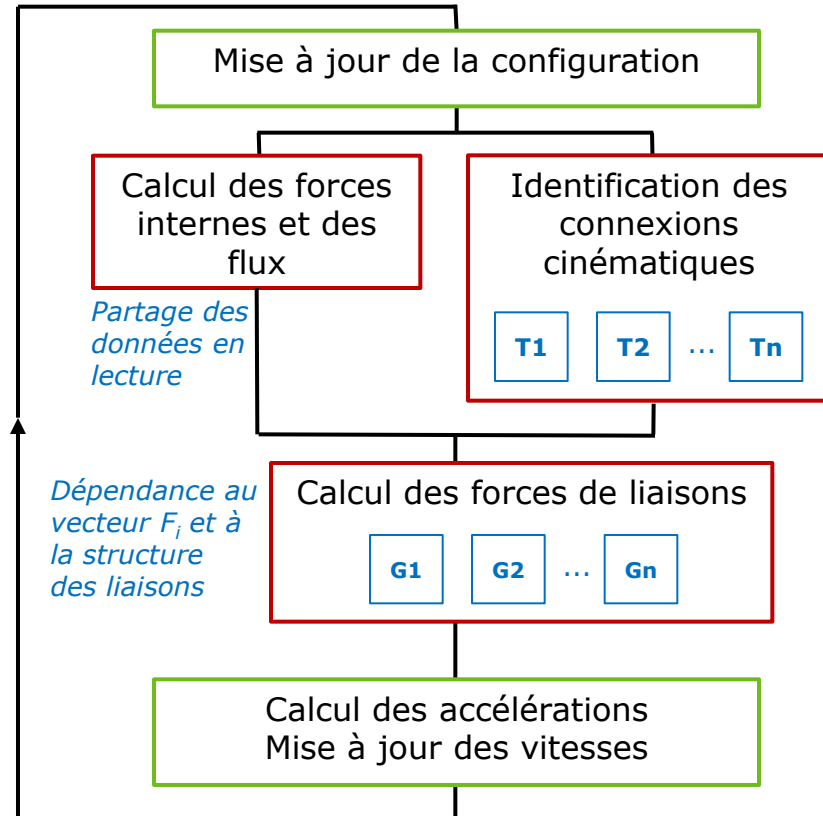
Distribution des tâches sur les ressources

gérées par la librairie
 Gestion des affinités (NUMA...)
Vol de travail pour l'équilibrage dynamique



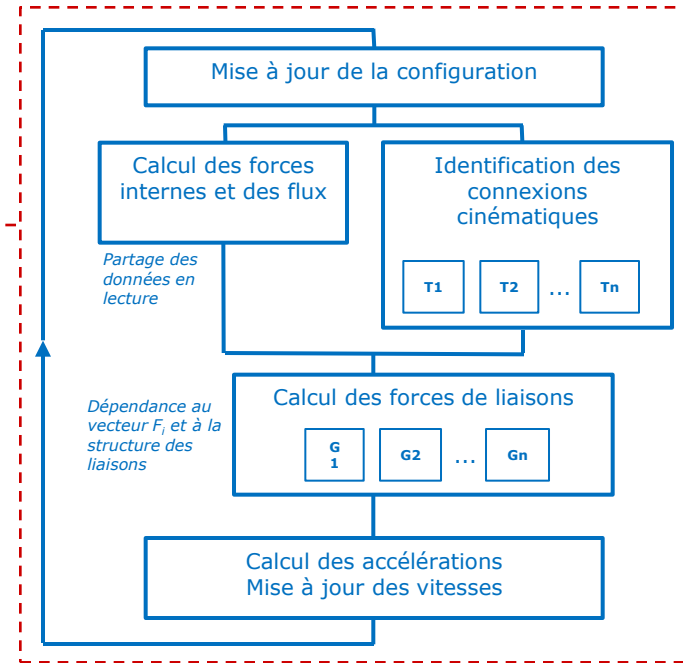
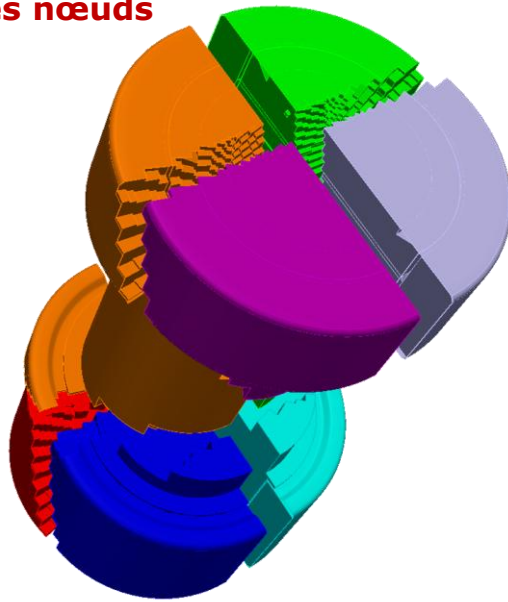
Test de démonstration : crash d'un tube

Concurrence entre découpage de la boucle élémentaire et traitement des liaisons d'auto-contact



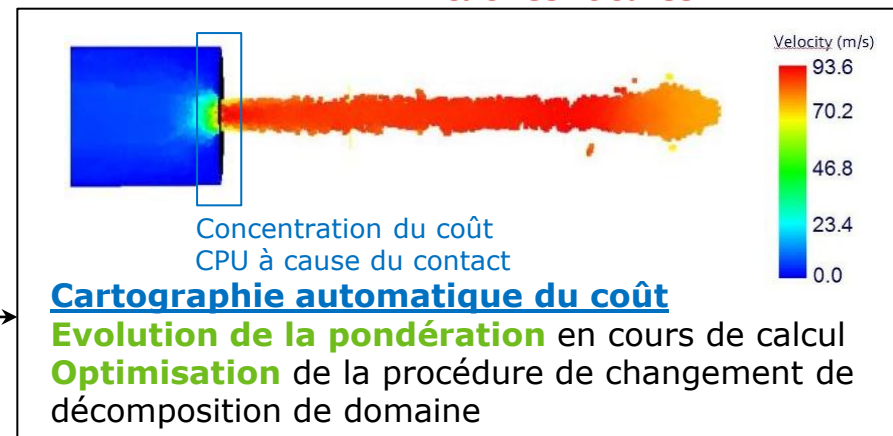
KAAPI v4
 T. Gautier
 B. Raffin

Décomposition de domaine entre les nœuds



Utilisation de X-KAAPI pour les tâches locales

- ✓ **Moins de sous-domaines à nombre de CPUs constant**
- ✓ Utilisation optimale des architectures
- ✓ **Progressions relativement orthogonales** des deux sources de parallélisme
- ✗ Stratégie d'**équilibrage de la charge entre les sous-domaines** à améliorer



- ✓ **Accélérations spectaculaires pour certaines tâches**
- ✓ C'est la mode...
- ✗ **Programmation spécifique complexe**
- ✗ **Pérennité** des formats et du concept ?
- ✗ Manque d'expérience et de recul

Stratégie pour EUROPLEXUS

Test de **faisabilité/pertinence** sur une modélisation particulière : **Eléments Discrets**

ANR Vulcain

<http://vulcain.ujf-grenoble.fr/>

Programmation confiée à une **unité informatique** : INRIA/LIG

Thèse de M. Durand

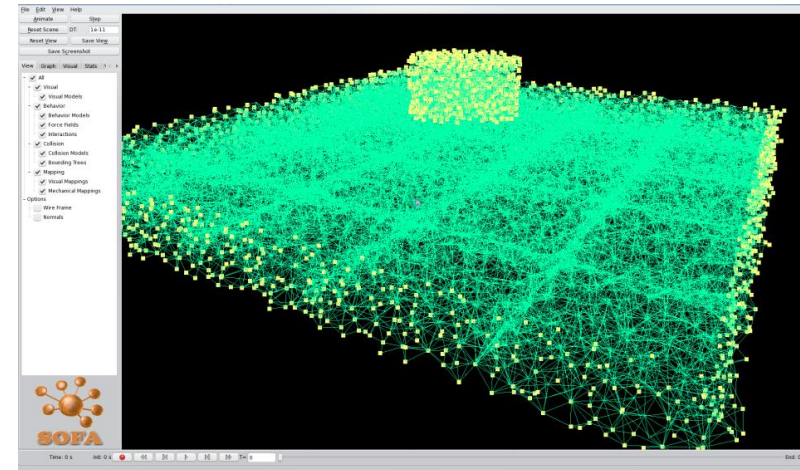
Financement ANR REPDYN

Codage dans une **librairie OpenSource** liée à EUROPLEXUS : **SOFA**

<http://www.sofa-framework.org/>

Interaction EUROPLEXUS MPI/SOFA/X-KAAPI

- MPI/SOFA : appel local à SOFA sur un sous-domaine
 - ⚠ **Frontière évolutive**
- Triple implémentation Eléments Discrets
 - CPU EUROPLEXUS
 - CPU SOFA
 - GPU SOFA
- Gestion par X-KAAPI des implémentations multiples et des affinités avec les ressources



Implémentation SOFA/X-KAAPI fonctionnelle et travaux de recherche en cours

Première phase de développement en cours de finalisation

- Parallélisme à mémoire distribuée générique dans EUROPLEXUS
- Mise en œuvre accessible
- Extensibilité hétérogène en fonction des applications
- Maintenance/adaptation aux nouveaux algorithmes encore difficiles
 - Adaptation de maillage (!!!!)
 - Nouveaux algorithmes d'IFS avec Volumes Finis
 - XFEM

Gros œuvre d'optimisation en cours (ANR REPDYN <http://www.repdyn.fr>)

- Seconde phase de développement
- **Directions de progrès clairement identifiées** et panel important de solutions potentielles
- 2 objectifs stricts
 1. Préserver la spécificité algorithmique d'EUROPLEXUS (**minimum de paramètres non-physiques**)
 2. Proposer un **parallélisme générique**
- **Résultats quantitatifs à venir** REPDYN = [2010-2012]

Tests de passage à l'échelle à mettre en œuvre dès à présent

- Sélection des **cas proposant une extensibilité compatible** avec les supercalculateurs massivement parallèles
 - Ecoulement explosif avec solveurs volumes finis haute-précision
 - Calculs particuliers à grand nombre d'éléments
 - ...
- Mise en œuvre et gestion des **difficultés spécifiques** : distribution de la mémoire en particulier
- Demande d'un *Preparatory Access* sur CCRT/Curie