





Equilibrage de charge et distribution de données pour des simulations parallèles irrégulières couplées

Maître de stage

Olivier Coulaud ; Aurélien Esnard ; Jean Roman

Tuteur

François PELLEGRINI

Période du stage

09/02/2009 - 30/09/2009

Master Recherche

Mohamed Amine EL AFRIT

PLAN

- Introduction
- Modélisation
- Résultats
- Conclusion et perspectives

Cadre général







- Etude des propriétés des matériaux
- Couplage de code simulant à deux échelles différentes
- Avant le couplage était manuelle
- Maintenant LibMultiScale le fait automatiquement



Cadre général

MD

Une simulation parallèle sur plusieurs processeurs



Décomposition de la zone atomique (MD) en boîtes



Décomposition de la zone continue (FE) par partitionnement de graphe



Problème de partitionnement

partitionnement de graphe = diviser un graphe en morceaux

- → parties ont la même taille/poids
- → peu de liens entre les morceaux

[15] Teresco, Devine and Flaherty





2- Graphe du maillage



3- Partitionnement du graphe 4- distribution sur des proc

2 phases de calcul









Objectifs

1. Equilibrer la charge pendant les deux phases

deux phases = phase standard & phase de couplage

Objectifs

1 - Equilibrer la charge pendant les deux phases

deux phases = phase standard & phase de couplage

2 - Améliorer les communications FE-MD

FE = processeurs éléments finis *MD* = processeurs dynamique moléculaire

Objectifs



1 - Equilibrer la charge pendant les deux phases

deux phases = phase standard & phase de couplage

2 - Améliorer les communications FE-MD

FE = processeurs éléments finis MD = processeurs dynamique moléculaire

→ On se concentre sur le partitionnement des 2 phases du code FE (en supposant que le partitionnement du code MD fixe)

Notions préliminaires

Graphe vs Hypergraphe

l'Hypergraphe généralise la notion de Graphe. première fois par Claude Berge en 1960 [2]

Graphe	Hypergraphe		
deux sommets par arête	plusieurs sommets par hyper-arête		
$\begin{array}{c} \hline V6 \\ \hline V4 \\ \hline V5 \\ \hline V1 \\ \hline V3 \\ \hline V2 \\ \end{array}$			
$V = \{v_1, v_2, v_3, v_4, v_5\}$ $E = \{e_1, e_2, e_3, e_4, e_5\}$ $= \{\{v_1, v_5\}, \{v_2, v_5\}, \{v_5, v_4\}, \{v_4, v_3\}, \{v_4, v_6\}\}$	$V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$ $E = \{h_1, h_2, h_3, h_4\}$ $= \{\{v_1, v_2, v_3\}, \{v_2, v_3\}, \{v_3, v_5, v_6\}, \{v_4\}\}$		

Notions préliminaires

Graphe vs Hypergraphe

l'Hypergraphe généralise la notion de Graphe. première fois par Claude Berge en 1960 [2]



Avantage de l'hypergraphe

→ Regrouper plusieurs sommets ayant une propriété commune ensemble





→ Modélisation réelle des coûts de communications

[14] http://www.cs.sandia.gov/Zoltan

PLAN

- Introduction
- Modélisation
- Résultats
- Conclusion et perspectives

Modélisation



Modélisation



Avec modélisation en graphe



Avec modélisation en graphe



Avec modélisation en hypergraphe



Avec modélisation en graphe

Avec modélisation en hypergraphe



Avec modélisation en hypergraphe avec des hyper-arêtes de couplage



Avec modélisation en graphe

Avec modélisation en hypergraphe

Avec modélisation en hypergraphe avec des hyper-arêtes de couplage

Avec modélisation en un hypergraphe avec la technique de multi-contraintes







20

Avec modélisation en graphe

Avec modélisation en hypergraphe

Avec modélisation en hypergraphe avec des hyper-arêtes de couplage

Avec modélisation en un hypergraphe avec la technique de multi-contraintes

Avec modélisation en un hypergraphe avec technique des points fixes





21

Stratégies de partitionnement

Avec modélisation en graphe

Avec modélisation en hypergraphe

Avec modélisation en hypergraphe et des hyper-arêtes de couplage

Avec modélisation en un hypergraphe avec la technique de multi-contraintes

Avec modélisation en un hypergraphe avec technique des points fixes







Avec modélisation en un hypergraphe avec technique des points fixes Et hyper-arêtes de couplage





Réalisation

- Module développé en C
- Utilisation de partitionneurs de graphes et d'hypergraphes PATOH ; METIS ; HMETIS
- Disponible sur la gforge de l'INRIA http://gforge.inria.fr/projects/mpicpl/



PLAN

- Introduction
- Modélisation
- Résultats
- Conclusion et perspectives

Résultats

Cas test géométrie - 2D



Coût sur les sommets

Calculs théoriques : surcoût de couplage négligeable devant le coût normales
 → Coût uniforme sur tous les sommets
 Coût sur les hyper-arêtes de calcul
 → Constant uniforme
 Coût sur les hyper-arêtes de calcul

➔ Déterminé expérimentalement

Résultats

Avant



Après



Modélisant par un graphe

Modélisant par un hypergraphe points fixes et hyper-arêtes de couplages

Phase standard Phase de couplage com FE-MD 12,5 (0) 12,5 (13,5) [0 0 2 4 3 0 2 0] Phase standard Phase de couplage com FE-MD 12.5 (0.04) 12.5 (0.14) [1 2 2 2 1 1 2 2]

Résultats

Avant



Modélisant par un graphe

Après



Modélisant par un hypergraphe points fixes et hyper-arêtes de couplages

Phase standard Phase de couplage com FE-MD 12,5 (0.22) 12,5 (12.66) [0 4 4 0 0 4 0 3] Phase standard Phase de couplage com FE-MD 12.5 (0.08) 12.5 (0.1) [2 3 2 2 2 2 3 2]

nV1_nHe1_bV11_bHe11_cHe110

PLAN

- Introduction
- Modélisation
- Résultats
- Conclusion et perspectives

Conclusion et perspectives

- Meilleur stratégie = technique des points fixes + hyper-arêtes de couplage
 → technique des points fixes → équilibrage entre les 2 phases → objectif 1
 → hyper-arêtes de couplage → optimisation des communications→ objectif 2
- Autres stratégies (sans points fixes)

→ utilisation graphes ou hypergraphes est équivalente

 Coûts théoriques de couplage sont négligeables devant les autres coûts réalité

→ Communications extérieures (FE-MD) non prises en compte encore

- Reste à intégrer avec LibMultiScale (travail en cours)
- Cas test considérés = géométrie en 2D
 → structure de donnée est générique et évolutif vers des cas en 3D
- Code source disponible en ligne sur la gforge de l'INRIA : http://gforge.inria.fr/projects/mpicpl/

Bibliographie

- [1] G. ANCIAUX. Simulation multi-échelles des solides par une approche couplée dynamique moléculaire/éléments finis. De la modélisation à la simulation haute performance. PhD thesis, Université de Bordeaux 1 (2007)
- [2] G. ANCIAUX. LibMultiScale, AMEL User Guide. Université Bordeaux 1, septembre 2007.
- [3] C. Berge, Introduction à la théorie des hypergraphes, Université de Montréal (1973).
- [4] Metis serial graph partitioning. Technical report. glaros.dtc.umn.edu/gkhome/
- [5] Hmetis serial hypergraph partitioning Technical report. glaros.dtc.umn.edu/gkhome/
- [6] SANDIA. Zoltan, http://www.cs.sandia.gov/zoltan/. http://www.cs.sandia.gov/Zoltan/.
- [7] Cevdet Aykanat Umit V. Catalyurek. PaToH : Partitioning Tool for Hypergraphs. The Ohio State University Department of Biomedical Informatics, Bilkent University Computer Engineering Department, 1999.
- [8] Umit V. Catalyurek al. Hypergraph-based dynamic load balancing for adaptive scientific computations.
- [9] Jean-François Remacle Christophe Geuzaine. Gmsh Reference Manual, 2009.
- [10] O.C. Zienkiewicz and R.L. Taylor. The nite element method, volume 1. 4 edition.
- [11] Karen D. Devine, Rob H. Bisseling, Parallel Hypergraph Partitioning for Scientific Computing, Sandia National Laboratories
- [12] Umit V. Catalyurek al. Hypergraph-based Dynamic Load Balancing for Adaptive Scientific Computations, Sandia National Laboratories (2007)
- [13] Umit V. Catalyurek , A Repartitioning Hypergraph Model for Dynamic Load Balancing, Sandia National Laboratories (2008)
- [14] Karen D. Devine, al. Balancing for Emerging Parallel Applications and Architectures, Sandia National Laboratories
- [15] James D. Teresco, al. Partitioning and Dynamic Load Balancing for the Numerical Solution of Partial Differential Equations, Sandia National Laboratories

Questions ?

ANNEXE A état de l'art

Comment il ont fait ? [11 ; 12 ; 13 ; 14 ; 15]	Qu'est ce qu'on veut faire ? Et comment adapter ?	
 Modélisation par hypergraphe → Estimation précise des coûts de calcul 	 Estimation précise des coûts → Modélisation par hypergraphe 	
 Équilibrage dynamique de la charge dans le même code, Répartition entre plusieurs processeurs Repartitionnement dynamique de l'hypergraphe 	 Équilibrage statique de charge entre deux codes Répartition entre plusieurs processeurs Un seul partitionnement de l'hypergraphe à l'initiation 	
 coût de migration des données 	 coût de couplage 	

ANNEXE A état de l'art

Équilibrage de charge dynamique



[13] Umit V. Catalyurek al.

ANNEXE C

simplification = considérer uniquement les premiers voisins



en considérant les premiers voisins.

```
nb hyperedges = 9542
total weights = 9542
edgecut = 445
% weights = [12 13 13 13 13 12 13 12]
```

En considérant les premiers et seconds voisins.

nb hyperedges = 9542 total weights = 9542 edgecut = 959 % weights = [13 13 13 13 12 12 12 12]

➔ dans le second cas, les hyper-arêtes se chevauchent davantage, mais ça n'a pas d'influence à priori pour le partitionneur

ANNEXE D

Avantages du modèle



Avant

Comm FE-MD : [3 4 2] % répart ph1 : [34 33 33] % répart ph2 : [32 40 28]



Après

Equilibrer la charge pendant les deux phases → partitionnement avec points fixes

% répart ph1 : [33 33 33] % répart ph2 : [33 33 33]





Améliorer les communications FE-MD

→ partitionnement avec hyper-arêtes de couplage
 Comm FE-MD : [2 2 2]
 % répart ph1 : [33 33 33]
 % répart ph2 : [33 33 33]

ANNEXE D

Limitations du modèle



nous ne pouvons pas assurer que les parties accordées à un même processeur soient contiguës sur le maillage.

Avec un coût uniforme pour les sommets et les hyper-arêtes et avec des hyperarêtes de couplages.

coût H	oût He poids v			5 V
cHe	Ν	В	Ν	В
50	1	1	1	1
% répa	art ph1 :	[12 12 13	3 13 12 1	3 13 13]
% répa	art ph2 :	[13 13 13	3 12 13 1	2 12 13]
com F	E-FE :	[3 3 3	4336	3]
com F	E-MD :	[1 3 2	2312	1]

Equilibrage du poids ph1 ph2 12.5 ; 0.1 12.5 ; 0.12

ANNEXE D

Limitations du modèle





ANNEXE E détermination expérimentale du coût

Calculs des coûts en fonction des complexités et des dépendances, Mais on a mis des coût artificiels pour mieux guider le partitionneur.

Comment les déterminer ?

➔ une étude exhaustive en faisant varier les coûts et voir leur influence sur les résultats de partitionnement



Cas d'une modélisation avec un graphe



Influence de la variation du poids des sommets de la zone de recouvrement du graphe sur l'équilibrage de charge. Influence de la variation du poids des sommets de la zone de recouvrement du graphe sur les communications.

ANNEXE F perspective désynchronisation







•Synchronisation nécessaire à chaque pas de temps.

•Possibilité de fusionner la dernière étape avec l'étape de calcul de la force F



Partitionnement du maillage en le modélisant par un graphe



Coût uniforme



Surcoût pour les sommets

1 1 1 1 12,5 ; 0 12,5 ; 13,5

% répart ph1 : [12 13 13 13 12 13 13 13] % répart ph2 : [0 0 17 36 26 0 21 0] com FE-FE : [3 6 4 4 3 3 5 4] com FE-MD : [0 0 2 4 3 0 2 0] 1 1 1 10 12.5 ; 0,01 12,5 ; 3,8

% répart ph1 : [13 12 12 12 12 13 12 13] % répart ph2 : [15 5 15 15 10 15 9 15] com FE-FE : [3 5 5 4 4 2 5 4] com FE-MD : [2 3 3 2 2 2 3 4]

Partitionnement du maillage en le modélisant par un hypergraphe

Avant



Coût uniforme



Surcoût pour les sommets et les hyper-arêtes

1 1 1 1 12,5 ; 0 12,5 ; 13,5

% répart ph1 : [12 13 13 13 12 13 13 13] % répart ph2 : [0 0 17 36 26 0 21 0] com FE-FE : [3 6 4 4 3 3 5 4] com FE-MD : [0 0 2 4 3 0 2 0]

0 1 10 1 10 12.5 ; 0.02 12.5 ; 2.8

% répart ph1 : [12 12 12 12 13 13 13 13] % répart ph2 : [15 7 15 11 15 10 15 11] com FE-FE : [2 3 1 2 2 4 1 3] com FE-MD : [2 2 2 2 3 2 2 2] 41

Partitionnement du maillage en le modélisant par un hypergraphe avec des hyper-arêtes de couplage



Coût uniforme



Coût uniforme

1 1 1 1 12,5 ; 0 12,5 ; 13,5

% répart ph1 : [12 13 13 13 12 13 13 13] % répart ph2 : [0 0 17 36 26 0 21 0] com FE-FE : [3 6 4 4 3 3 5 4] com FE-MD : [0 0 2 4 3 0 2 0]

100 1 1 1 1 12.5 ; 0.13 12.5 ; 15.26

% répart ph1 : [12 13 13 12 12 13 12 13] % répart ph2 : [12 0 38 0 0 0 13 37] com FE-FE : [5 3 2 3 2 5 3 3] com FE-MD : [1 0 3 0 0 0 1 3]

Partitionnement du maillage en le modélisant par un hypergraphe avec la technique de multi-contraintes



Coût uniforme



Technique multi-contraintes

1 1 1 1 12,5 ; 0 12,5 ; 13,5

% répart ph1 : [12 13 13 13 12 13 13 13] % répart ph2 : [0 0 17 36 26 0 21 0] com FE-FE : [3 6 4 4 3 3 5 4] com FE-MD : [0 0 2 4 3 0 2 0] 12.5; 0.22 12.5; 0.27

% répart ph1 : [12 12 12 12 13 12 13 13] % répart ph2 : [12 12 12 13 12 12 13 13] com FE-FE : [4 5 6 5 5 4 5 6] com FE-MD : [2 4 4 3 2 3 6 3]

ANNEXE H modélisation



ANNEXE I Réalisation

Environnement d'exécution de notre module

