

Projet CIP, Calcul Intensif Pair à pair ANR-07-CIS7-011

Coordonnateur : Didier EL BAZ, LAAS-CNRS

<http://www2.laas.fr/CIS-CIP>

LAAS-CNRS
IRIT-ENSEEIH
Université de Franche-Comté, LIFC
Université de Picardie, MIS
EMD

Plan de l'exposé

1. Etat de l'art, objectifs
2. Projet CIP
3. Principes et innovations
4. Résultats
5. Retombées
6. Futurs développements

1. Etat de l'Art, Objectifs

- **Contexte :**

augmentation des performances des processeurs et des réseaux ;

développement des applications pair à pair : partage de fichiers (Gnutella, FreeNet), vidéo streaming.

calcul intensif pair à pair solution économique et attractive.

1. Etat de l'Art, Objectifs

- **Etat de l'art :**

Les solutions existantes sont tournées vers des **applications de type parallélisme de données** pour lesquelles les **tâches sont indépendantes** et la **communication entre pairs n'est pas nécessaire** (BOINC, OurGrid) ou vers le calcul global (XtremWeb).

Les **protocoles de communication sont peu adaptés à des applications HPC et non adaptatifs** ou tout au plus liés au réseau.

1. Etat de l'Art, Objectifs

- Objectifs :

Mettre en œuvre de manière décentralisée des applications HPC de type parallélisme de tâche sur réseaux de machines allant des grilles de calcul aux réseaux pair à pair.

Calcul intensif pair à pair, ex : **simulation numérique** et **optimisation**.

1. Etat de l'Art, Objectifs

- Objectifs :

Mettre en œuvre des méthodes parallèles nécessitant des **échanges fréquents et rapides** entre pairs, ex. méthodes itératives (schémas itératifs synchrones ou asynchrones) et méthodes récursives ;

faciliter la programmation, via l'utilisation de protocoles auto adaptatifs efficaces, un **ensemble réduit d'opérations de communications** et le **choix de schémas itératifs**.

1. Etat de l'Art, Objectifs

- Objectifs :

conception d'outils pour la mise en œuvre de calculs intensifs sur des réseaux pair à pair,

- **environnement décentralisé de calcul intensif ;**

- **outil de prédiction de performance ;**

- **démonstrateurs, série de codes en simulation numérique et optimisation, ex. mathématiques financières, génie des procédés et logistique.**

1. Etat de l'Art, Objectifs

- **Problématique :**

hétérogénéité ;

passage à l'échelle ;

volatilité des pairs ;

gestion des communications ;

protocoles dédiés aux applications HPC ;

environnements décentralisés ;

conception d'algorithmes parallèles efficaces.

2. Projet CIP

- **Trois sous-projets :**

P2PDC, environnement décentralisé pour le calcul intensif pair à pair, LAAS-CNRS ;

dPerf, Prédiction de performance d'applications de calcul distribué, LIFC.

P2PDEM, démonstrateurs, LAAS, IRIT, MIS, EMD.

3. Principe et Innovations

- P2PDC, LAAS-CNRS :

environnement décentralisé pour le calcul intensif pair à pair avec des fonctionnalités de sûreté de fonctionnement ;

basé sur le **protocole auto adaptatif** dédié au calcul intensif P2PSAP (extension de CTP utilisant les micro protocoles et CACTUS). **Choix dynamique** du mode de communication entre pairs en fonction de choix algorithmique (**couche application**) et éléments de contexte ex. topologie (**couche transport**).

3. Principe et Innovations

- P2PSAP, Protocole auto adaptatif

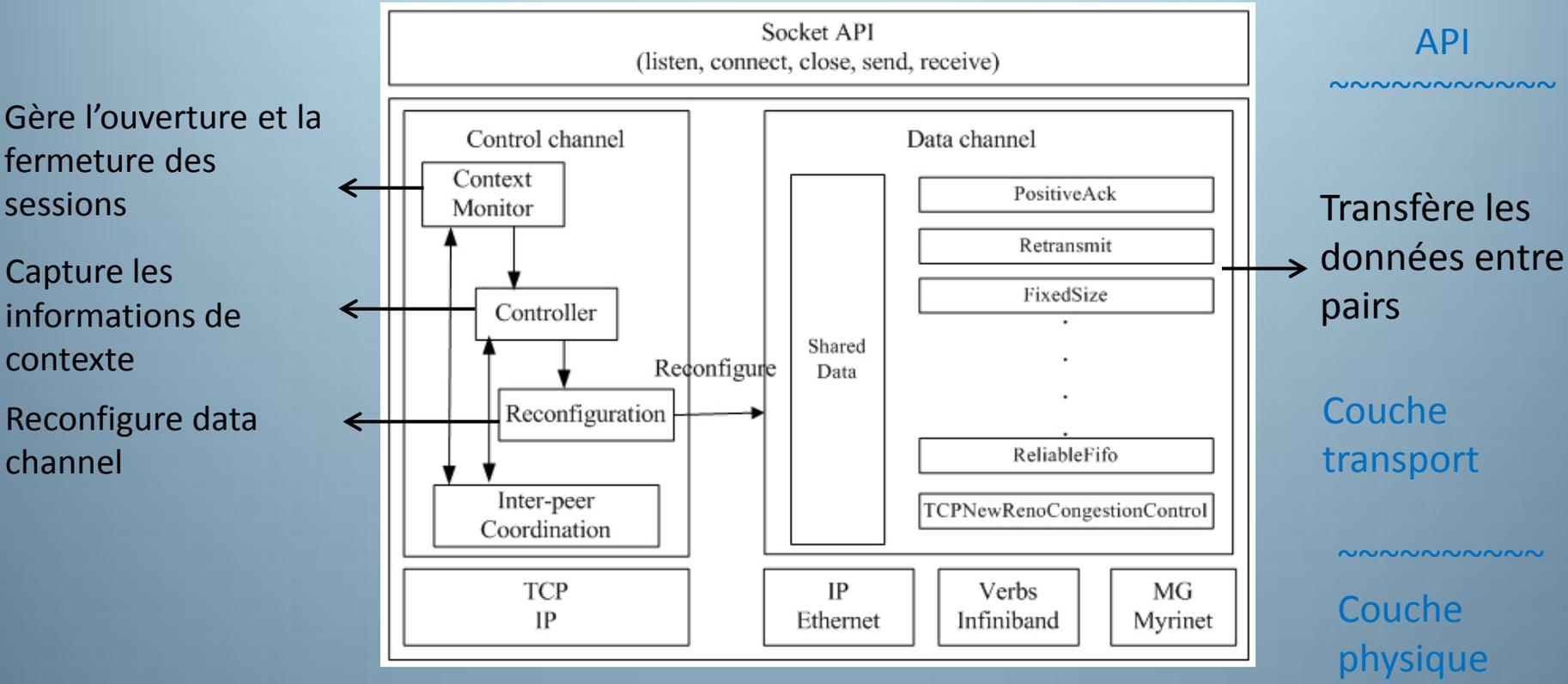


Fig. 1 Architecture du protocole P2PSAP

3. Principe et Innovations

- P2PSAP :

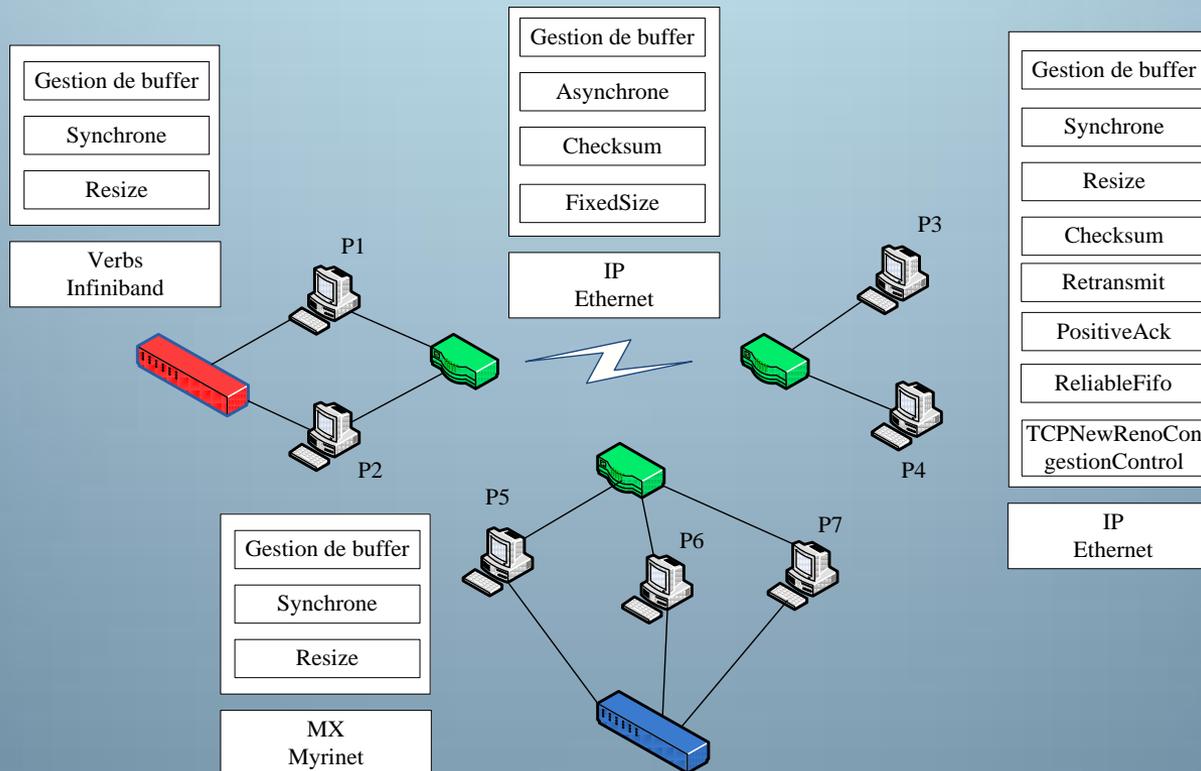


Fig. 2 Exemple de scénario P2PSAP

3. Principe et Innovations

- Environnement P2PDC

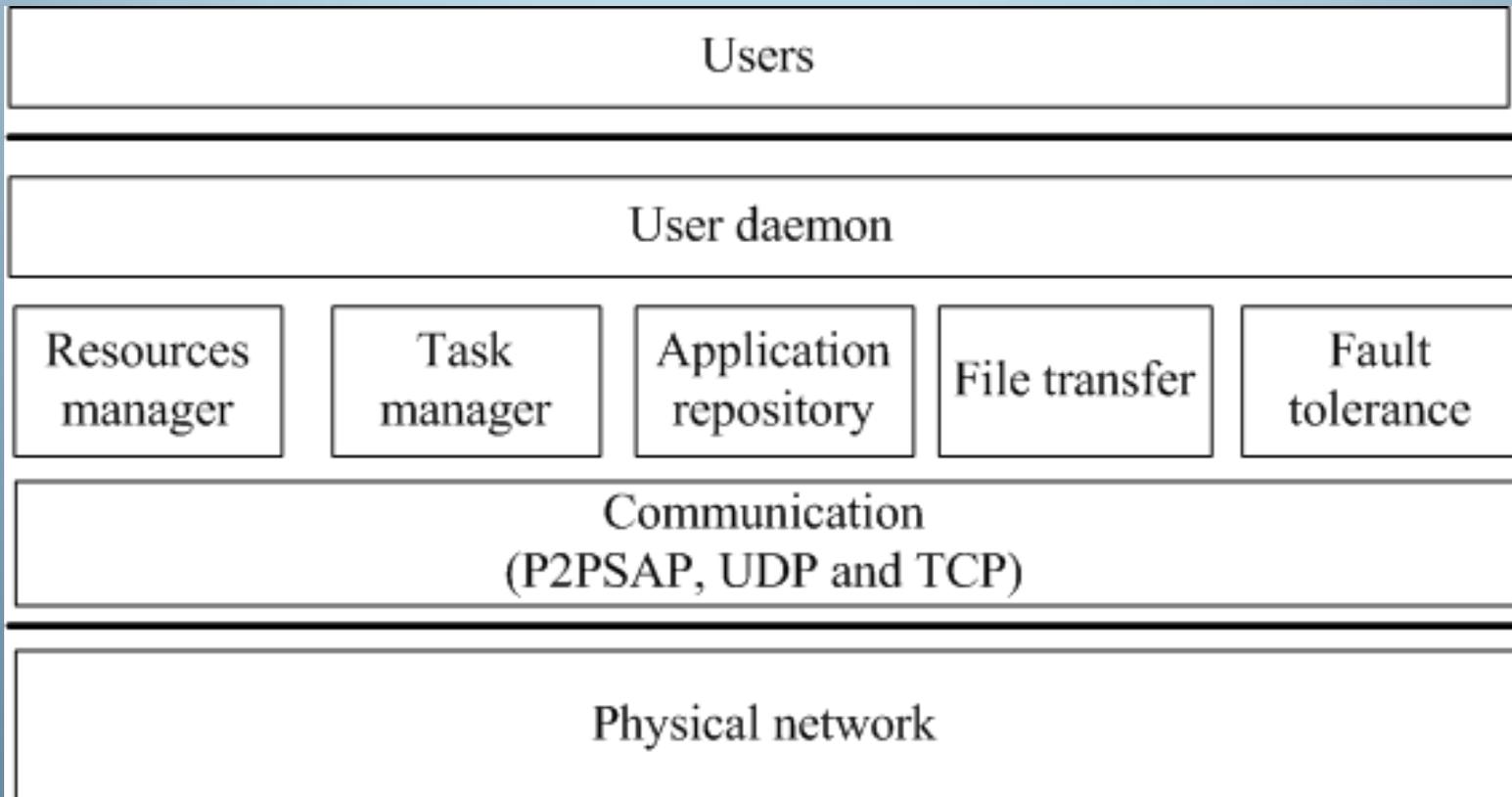


Fig. 3 Architecture de l'environnement

3. Principe et Innovations

- Topologie du réseau pair à pair et allocation de tâches :

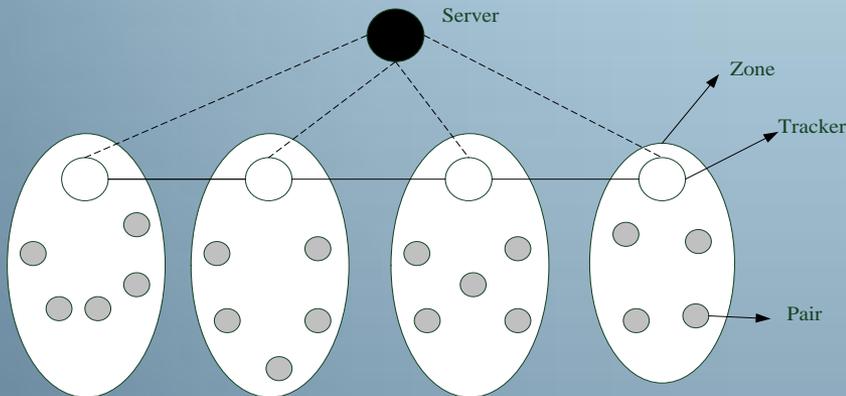


Fig. 4a Topologie du réseau P2P

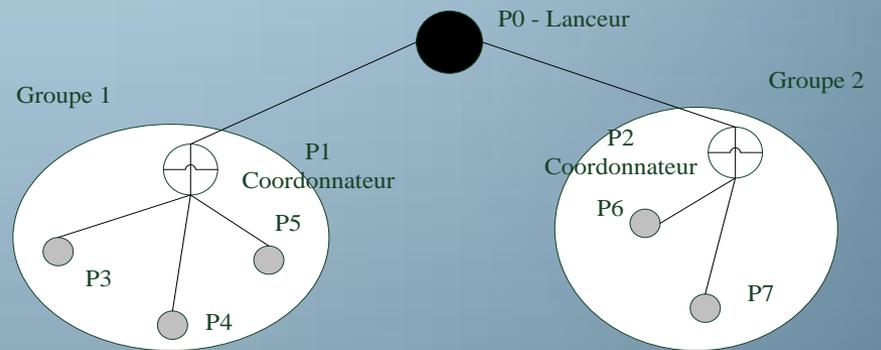


Fig. 4b Arbre de déploiement d'un calcul

3. Principe et Innovations

- Robustesse

workers: checkpoint-based rollback-recovery technique

- cas synchrone: coordinated checkpointing,
- cas asynchrone: uncoordinated checkpointing ;

coordonnateurs: techniques de replication.

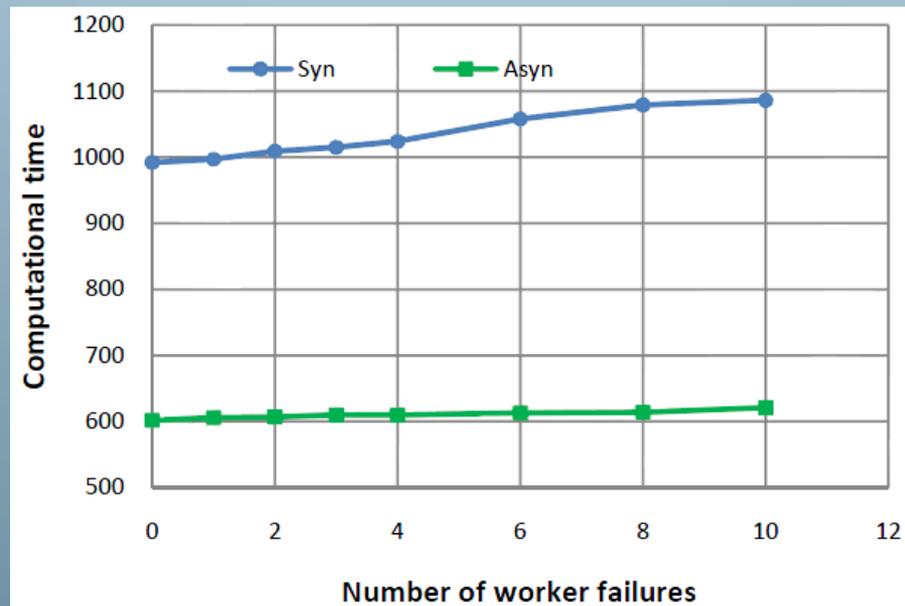


Fig. 4c Temps de calcul pour différentes fautes des workers

3. Principe et Innovations

- **dPerf, LIFC :**
 - **Outil de prédiction de performance d'applications distribuées.**
 - **Outil hybride.**
 - **Simulation basée sur des fichiers de trace compatibles SIMGRID.**
 - **Mesure de compteurs hardware (PAPI).**
 - **Précision adaptable.**
 - **Facilité du passage à l'échelle.**

3. Principe et Innovations

dPerf :

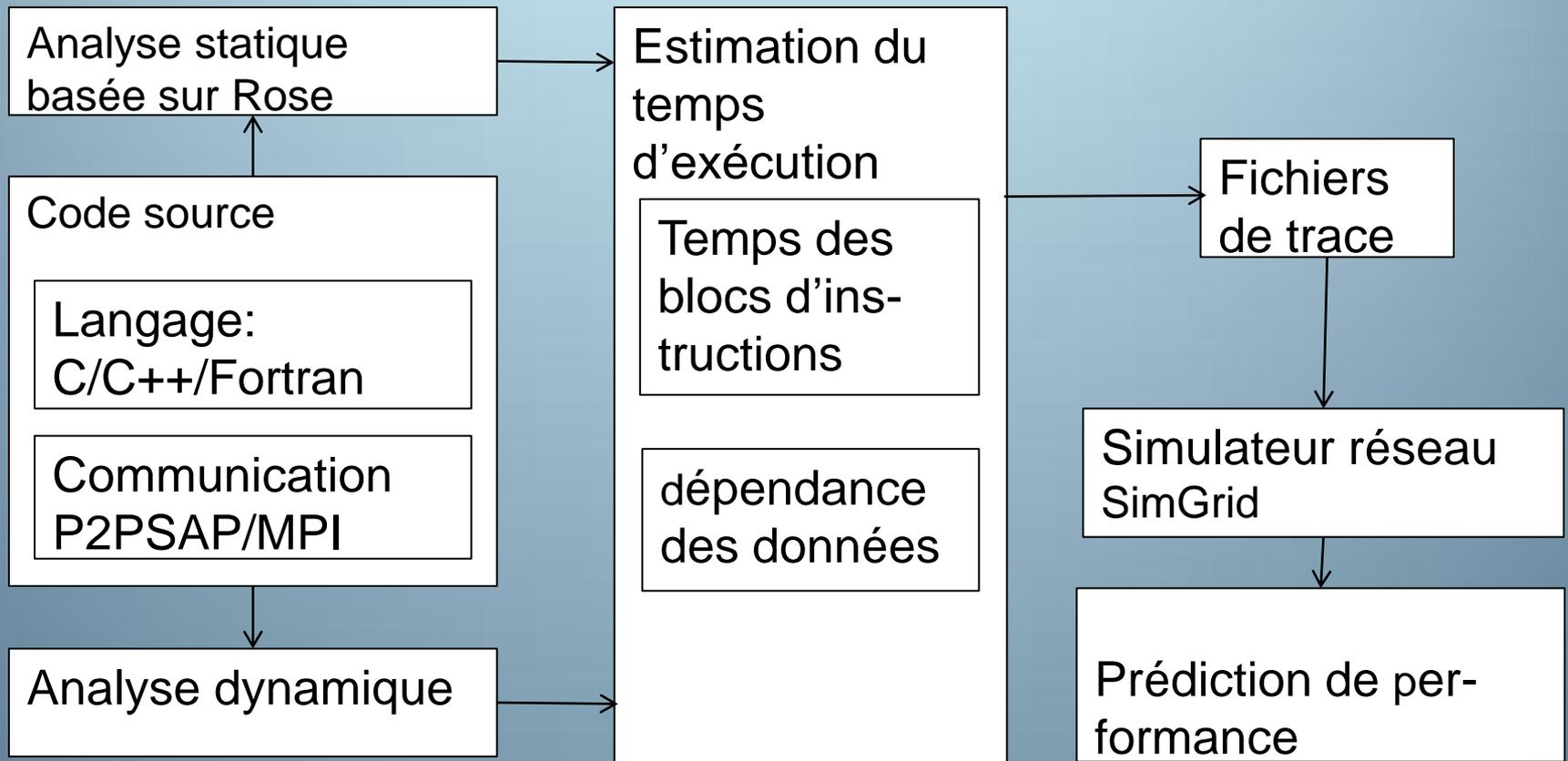


Fig. 5 Architecture globale de dPerf

3. Principe et Innovations

- Prédiction dans des conditions d'exécution réelles.
- Slowdown réduit.
- Possibilité d'extrapolation.
- Estimation rapide su temps d'exécution.

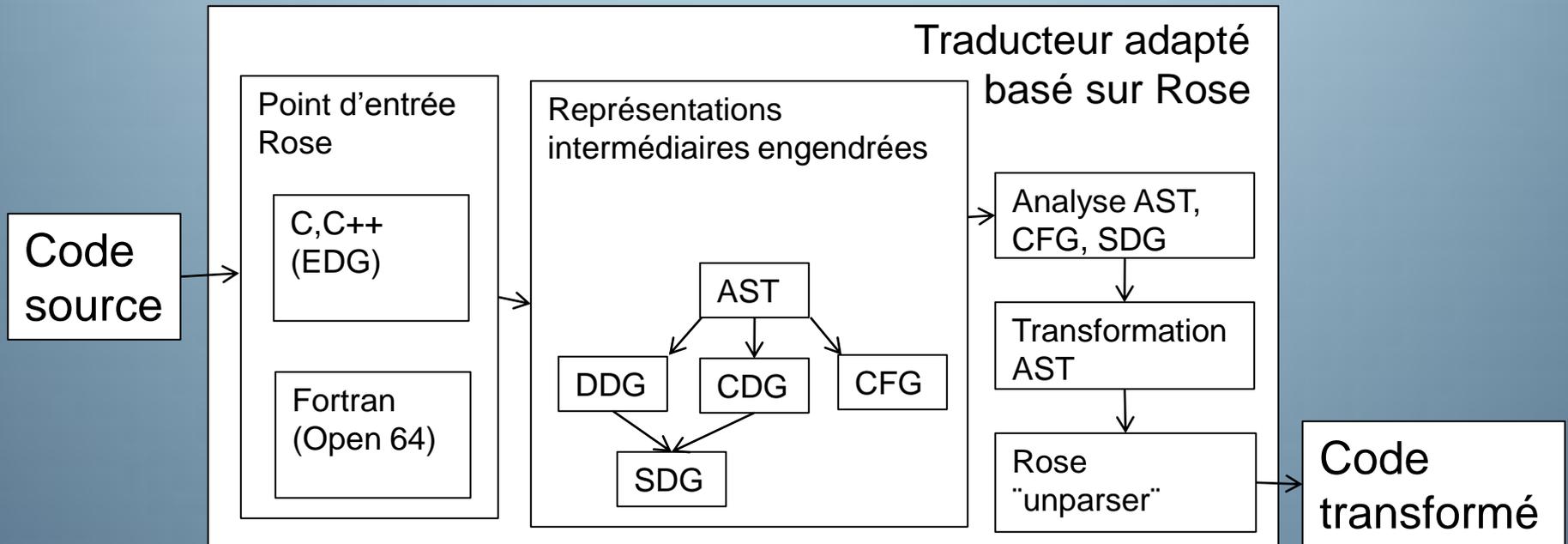


Fig. 6 Transformation du code source et représentations intermédiaires

3. Principe et Innovations

- P2PDEM, démonstrateurs et défis applicatifs :

maths financières, «option pricing» IRIT, LAAS ;

génie des procédés, IRIT-N7 ;

logistique, MIS et EMD.

3. Principe et Innovations

- Algorithmes itératifs asynchrones, LAAS-CNRS, IRIT-ENSEEIH
- Schémas itératifs hybrides, LAAS-CNRS

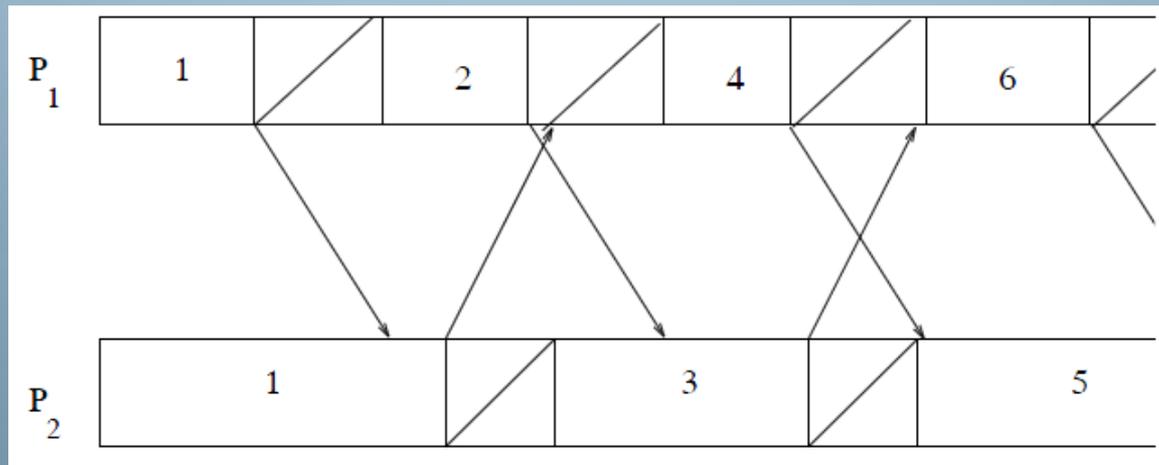


Fig. 7 Schéma itératif asynchrone

4. Résultats

Test sur Grid 5000, 5 sites, 256 machines, problème d'obstacle, 17.10^6 variables.

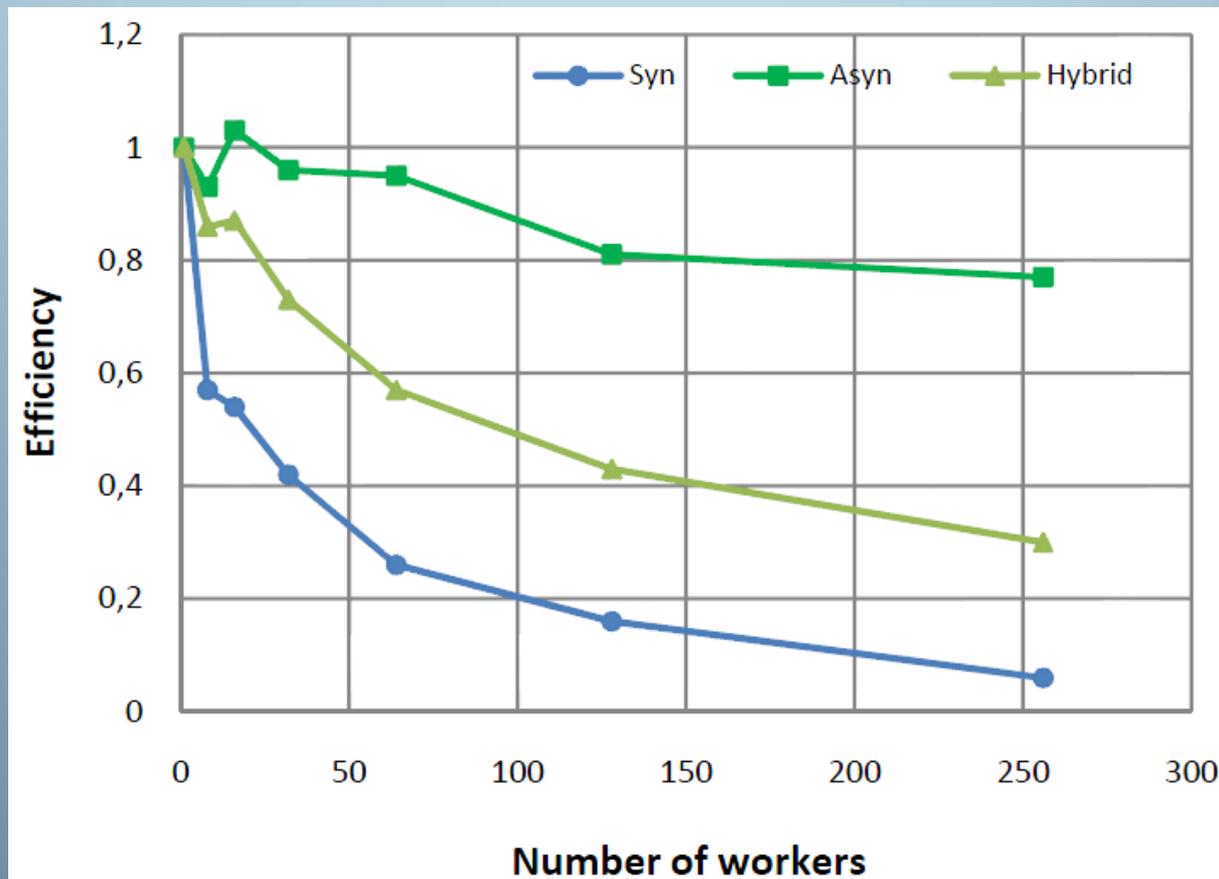


Fig. 8 Problème d'obstacle, Grid 5000

4. Résultats

Test sur la plateforme NICTA Sidney, 24 machines, problème d'obstacle, $3 \cdot 10^6$ variables.

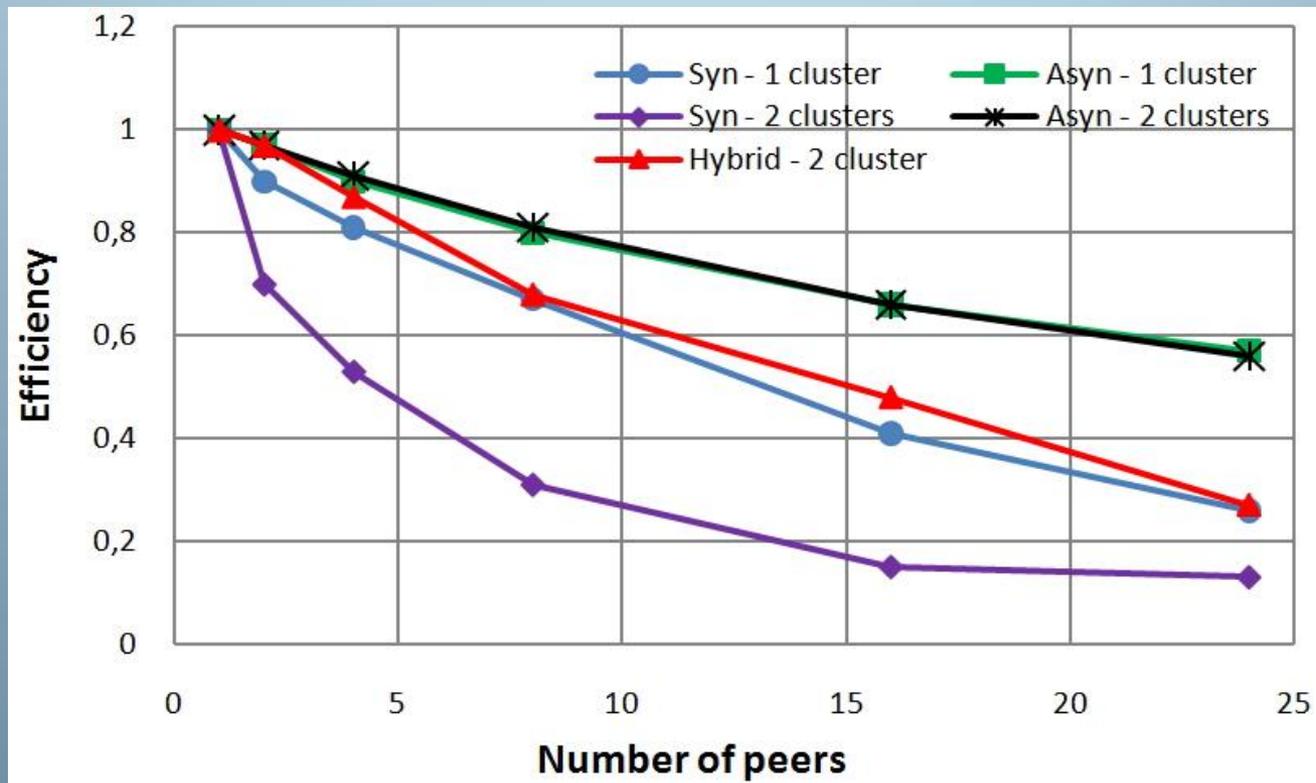


Fig. 9 Problème d'obstacle, plateforme NICTA

4. Résultats

Test sur la Grid 5000, 2 sites, 128 machines, problème d'options européennes, 17.10^6 variables

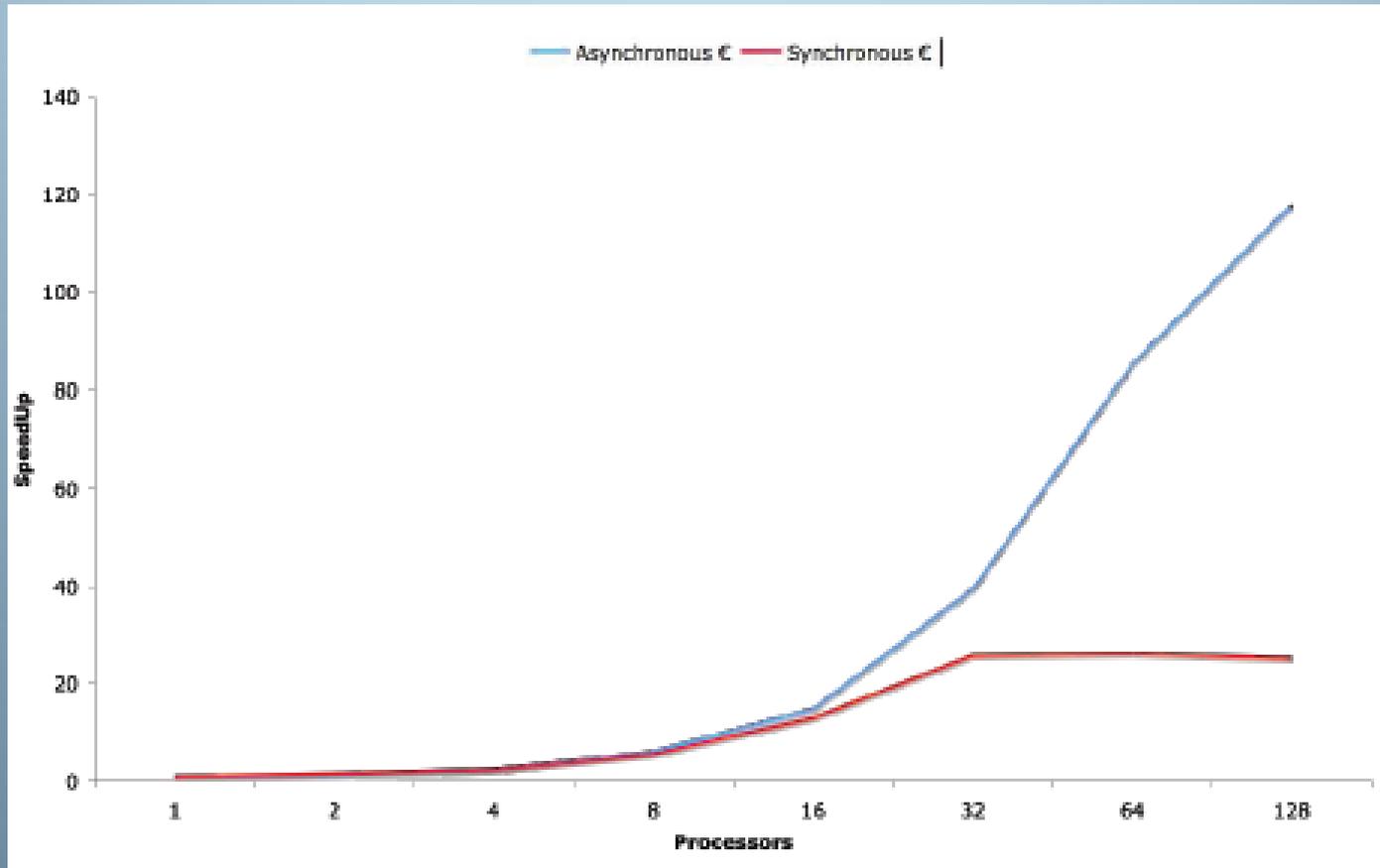


Fig. 10 Problème d'options européennes, Grid 5000

4. Résultats

Test sur Grid 5000, 2 sites, 32 machines, problème d'électrophorèse, $4.5 \cdot 10^6$ variables

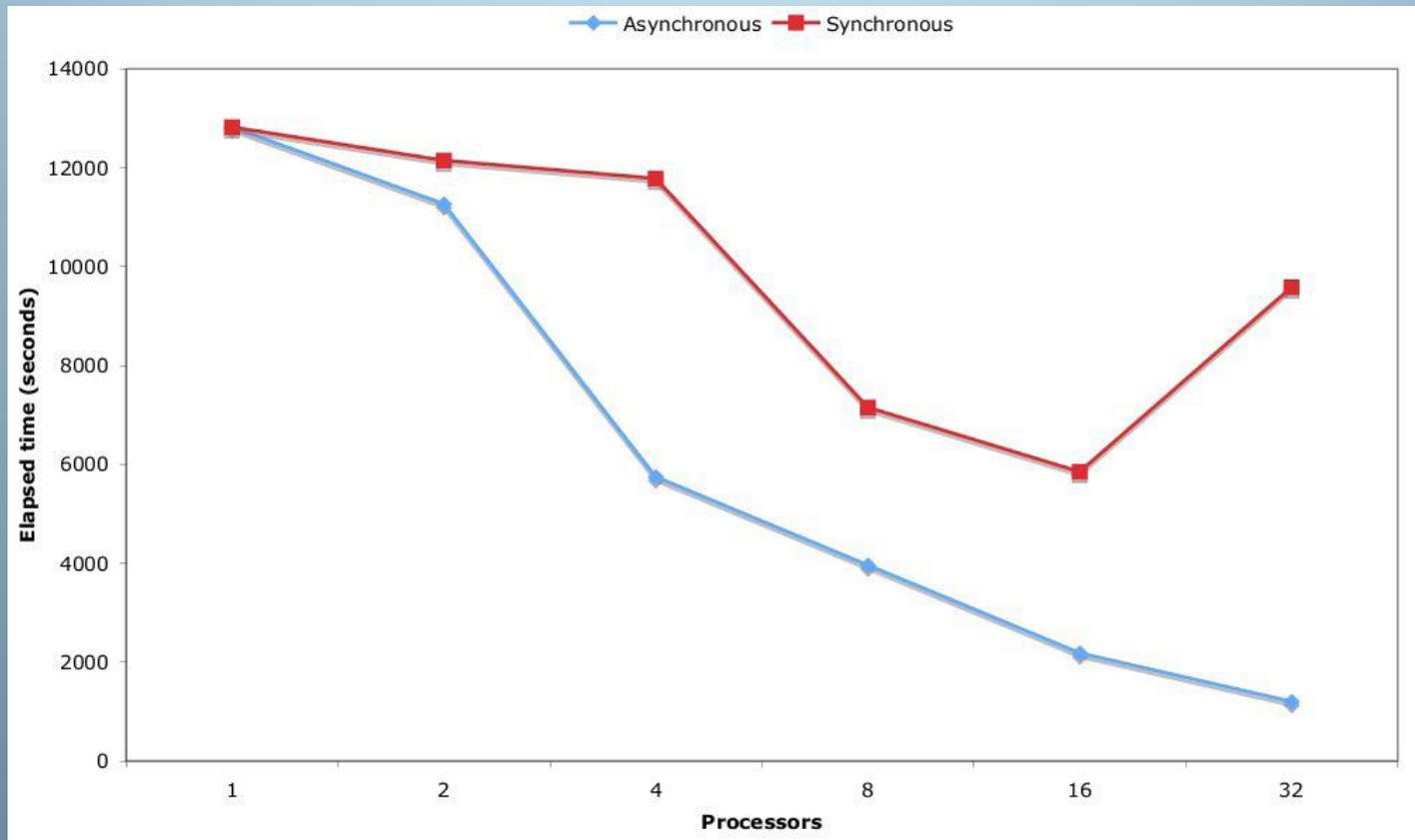


Fig. 11 Problème d'électrophorèse, Grid 5000

4. Résultats

- P2PDEM, MIS

Logistique, «Guillotine cutting problems»

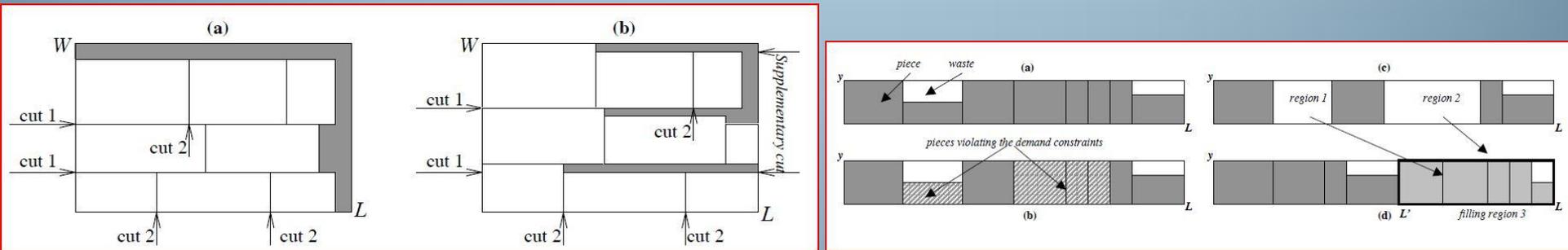


Fig. 12 Problème de découpe 2D

4. Résultats

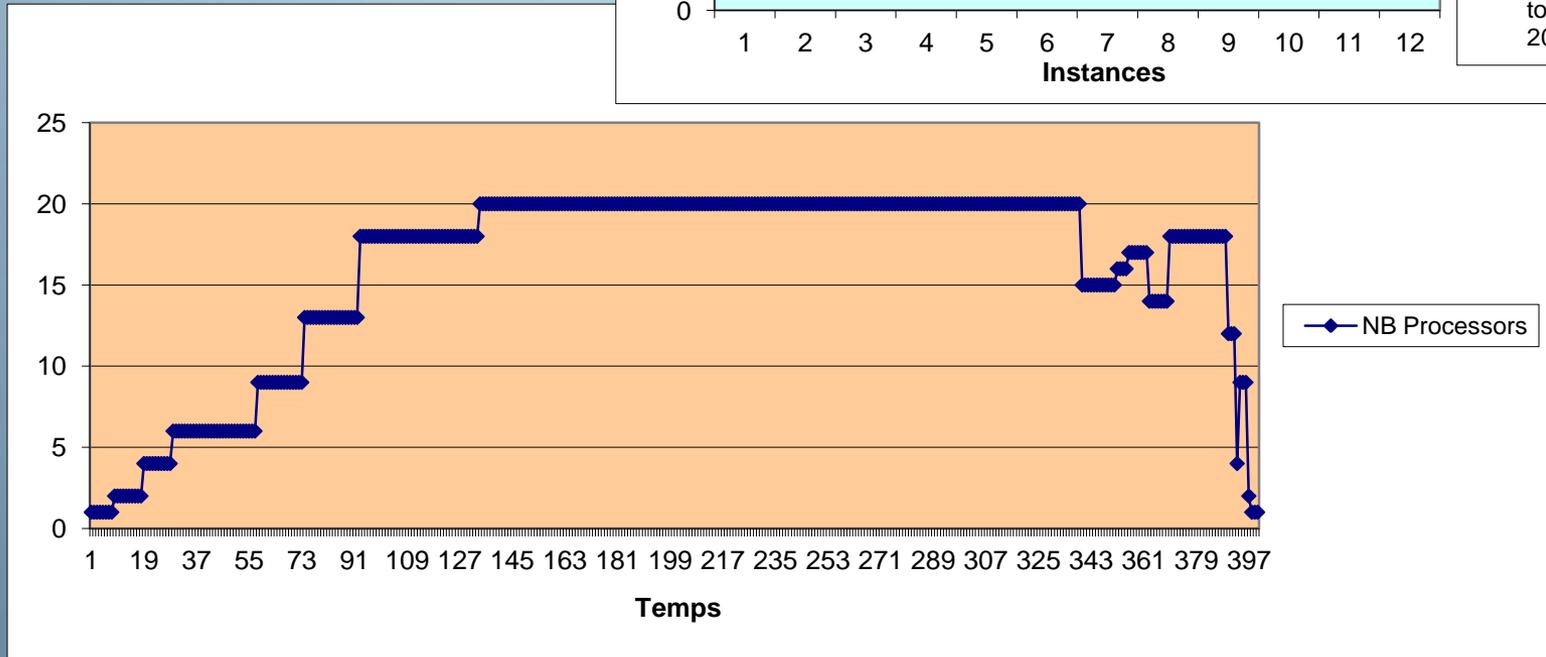
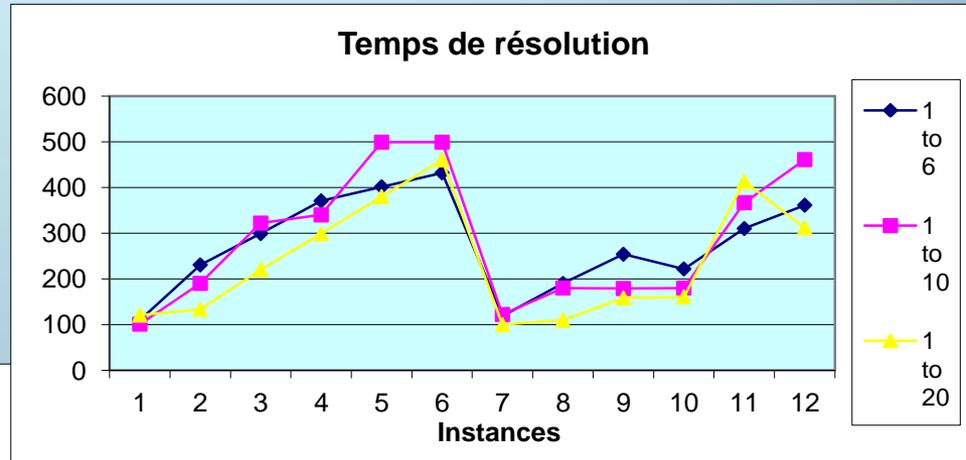


Fig. 13 Problème de découpe, Grid 5000

4. Résultats

Instance	Best	Cplex.Gap	PAR.Gap	SBS Algorithm				P2PCGBS Algorithm			
				$\beta=2$		$\beta=4$		Gap	1 to 6 CPU	1 to 10 CPU	1 to 20 CPU
				Gap	CPU	Gap	CPU				
First Cut Horizontal											
UL1	887393	-10711	-294	0	138,56	-187359	Δ	1993	110,09	101,15	120,96
UL2	906237	-5314	0	0	310,13	-216139	Δ	1629}	230,81	190,27	133,66
UL3	1017575	-8544	0	0	347,89	-18810	Δ	0	299,16	322,17	220,39
WL1	688840	-1054	0	0	410,09	-168781	Δ	0	370,82	340,06	299,01
WL2	782566	-4538	0	0	454,76	-125136	Δ	1562	401,55	499,14	380,18
WL3	899890	-2517	0	0	488,76	-486401	Δ	0	432,12	499,05	460,53
First Cut Vertical											
UL1	891516	-2754	0	0	120,87	-316084	Δ	8583	118,81	122,19	99,91
UL2	908982	-2382	0	0	210,41	-230231	Δ	1039	190,07	180,15	109,77
UL3	1022483	-3595	-3595	0	320,87	-124832	Δ	0	254,08	179,08	158,31
WL1	694657	0	0	0	340,89	-177948	Δ	0	221,95	180,11	120,65
WL2	783836	-1687	-1687	0	410,78	-293071	Δ	5985	310,35	366,88	413,09
WL3	881931	0	0	0	456,81	-272056	Δ	0	361,44	460,99	312,07
AV_Gap		-3591,33	-464,67	0,00		218070,67		1742,00		286,77	322,38
AV_CPU					334,24				275,10	286,77	322,38

5. Retombées

- Mise en ligne de dPerf, P2PDC.
- 2 thèses soutenues.
- Publications IPDPS, HPCC, PDP, ICCSA, GENI Engineering Conference, CIE 39, COSI'2009.
- Organisation de Workshops internationaux sur le calcul pair à pair et le calcul parallèle :
 - MSOP2P 2008-2012,
 - HOTP2P 2010-2011,
 - PCO 2011-2012.

6. Futurs Développements

- Collaboration avec le NICTA Sidney

Déploiement d'applications pair à pair à grande échelle sur plateforme réseaux grâce à OML.

- NVIDIA, Academic Partnership

Couplage calcul pair à pair calcul sur GPU.

6. Futurs Développements

- Collaboration avec Emory University, Atlanta

Prédiction de performance de plateformes de cloud computing.