

DE LA RECHERCHE À L'INDUSTRIE

cea



www.cea.fr

Une nouvelle stratégie de glissement de données pour les caches élastiques dans les architectures many-coeurs

Journée Logiciels Embarqués et Architectures Matérielles du GDR SoC-SiP | [Safae](#)
[DAHMANI](#) Loïc CUDENNEC

15/11/2012

list

CONTEXTE GENERAL

Vers les many-coeurs

- Le marché des applications multimédia, jeux vidéo exige plus de performances.

→ Intégrer des cœurs complexes et puissants: **Contraintes de coût/énergie.**

→ Multiplier le nombre de cœurs: **Many-coeurs.**

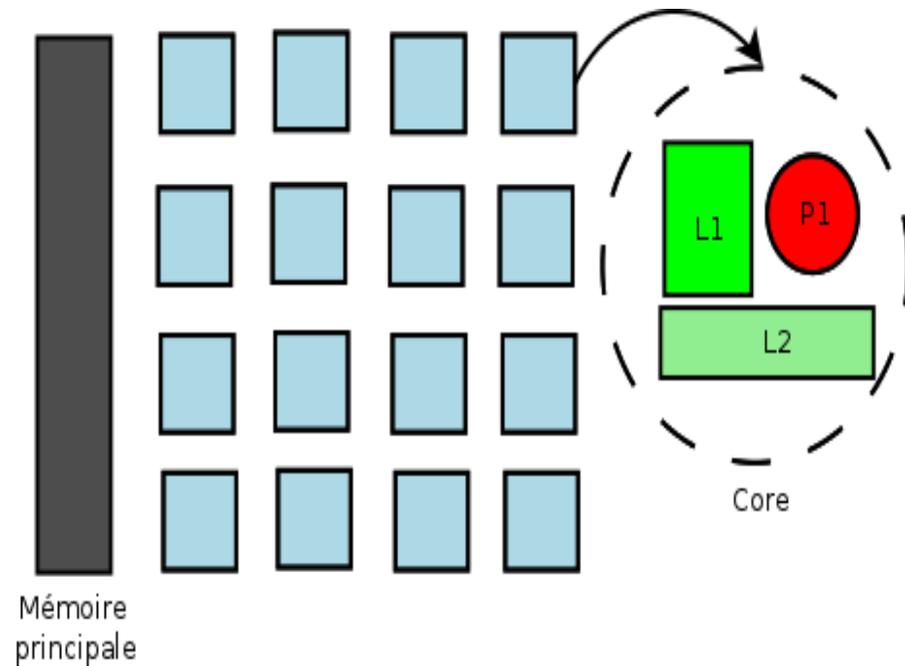


Figure 1 : Grille de 16 unités de calcul structurées en maille

Hiérarchie des caches

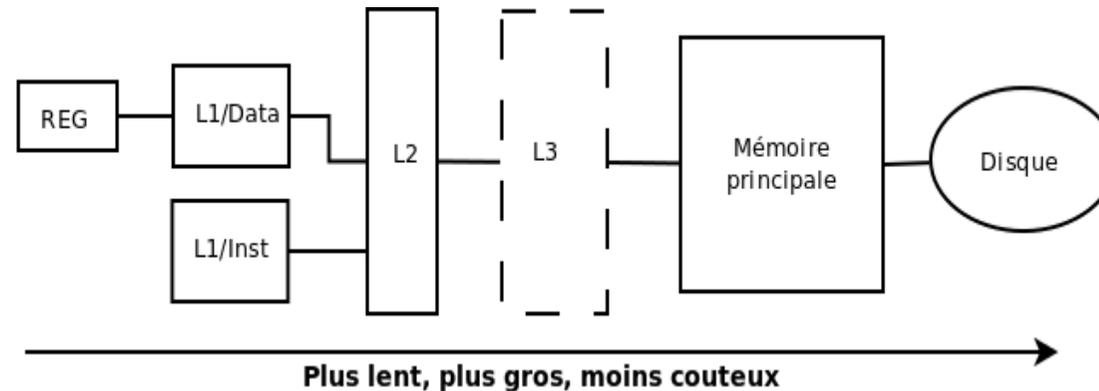


Figure 2 : Différents niveaux de mémoire

L'objectif de la hiérarchie des caches est de **minimiser la latence** d'accès au données fréquemment utilisées.

- **L'organisation des caches :**
passante.
- **La politique de gestion des caches :** Accessibilité, allocation, éjection des données, stratégie de cohérence , cache inclusif/ exclusif.

Organisation des caches

■ Cache privé :

- + Latence d'accès réduite, non uniforme,
- + Meilleure scalibilité,
- Capacité de stockage réduite.

■ Cache partagé:

- + Augmente la capacité effective (pas de duplication de données partagées),
- Une donnée placée arbitrairement dans le cache peut être éloignée du cœur qui l'utilise,
- Passage à l'échelle limité.

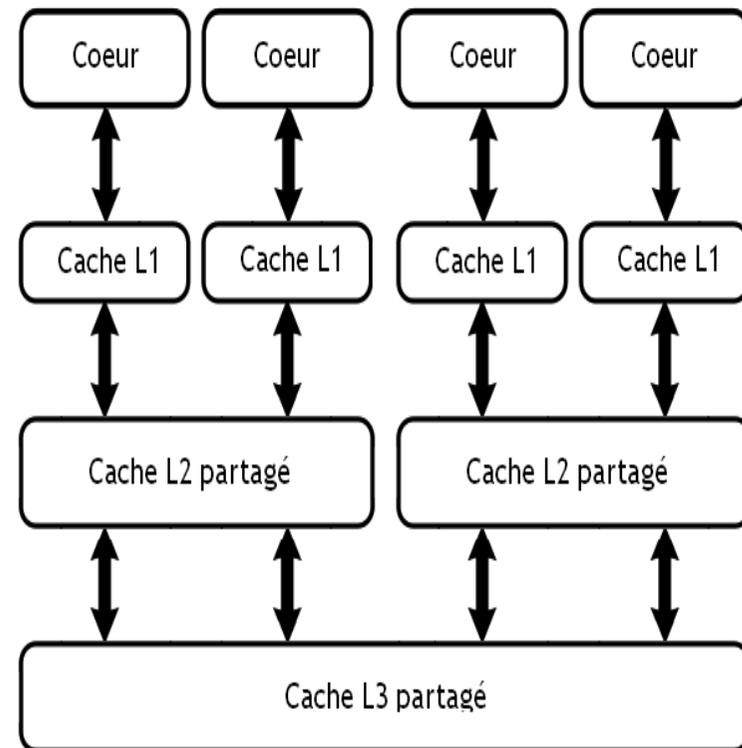


Figure 3: Options Privé/Partagé des caches hiérarchiques

**Quel serait le meilleur compromis
PRIVE/PARTAGE?**

MODELES ADAPTATIFS DE LA HIEARCHIE DES CACHES

Cache Coopératif

Principe : créer un espace de stockage commun.
Chaque nœud a le droit d'utiliser l'espace libre d'un autre.

Performances :

- Augmente la capacité de stockage (partage logique)
- Réduit le taux des défauts de caches
- Réduit la latence d'accès

Limitations :

- Contrôle de cohérence centralisé (CCE) : Consommation, Contention du réseau, Passage à l'échelle limité.

Perspectives :

- Optimisation de l'espace de stockage, autonomie de contrôle.

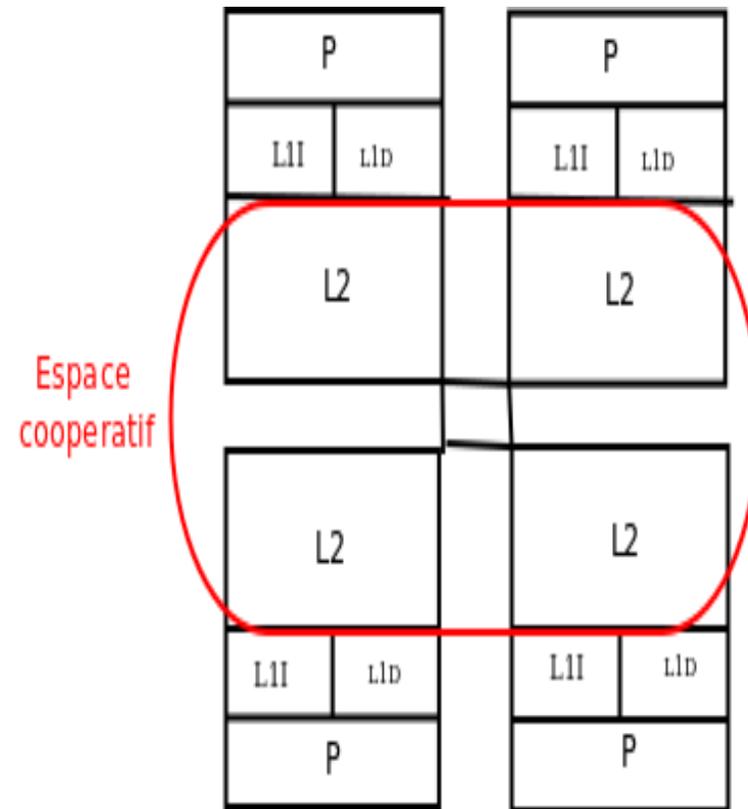


Figure 4: Zone coopérative

Cache Coopératif Élastique

■ Motivations:

Proposer un système à cache coopératif, dynamique, autonome.

S'adapter aux besoins applicatifs et au passage à l'échelle (DCE).

■ Principe:

Le cache L2 est logiquement divisé en zones privée/partagée.

Chaque nœud est capable d'aller stocker des données chez son voisin(1 Chance Forwarding).

■ Deux unités de contrôle :

- Unité de partitionnement
- Unité d'allocation de bloc

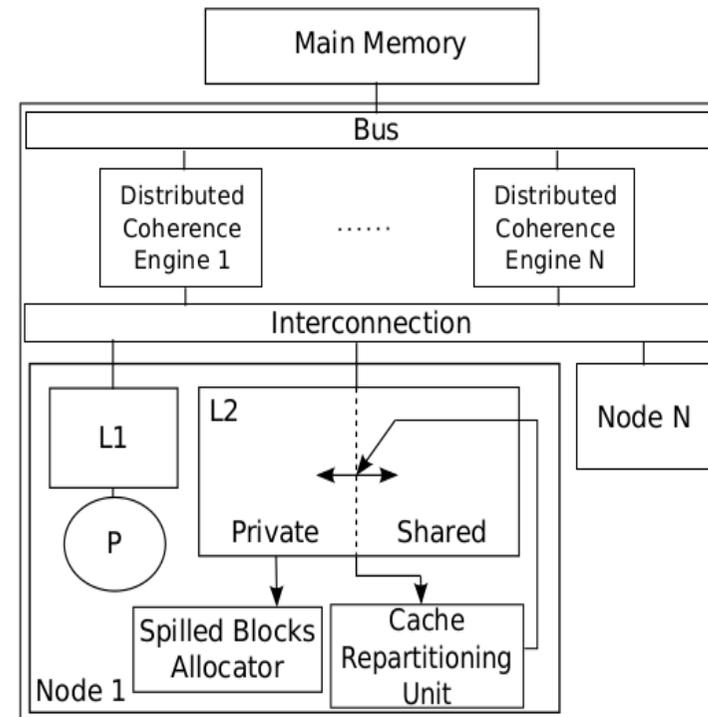


Figure 5: Structure d'un noeud à cache coopératif élastique

**LIMITATION DU MODELE COOPERATIF
ELASTIQUE**

Cas de voisinage stressé

En cas de voisinage fortement stressé:

1. Le compteur stagne autour de 0: éjection des données.
2. Le compteur oscille entre des valeurs extrêmes: comportement instable de l'unité de partitionnement.
3. La stratégie **Round Robin**: les nœuds avec plus de blocks partagés sont les plus sollicités. Nécessité de transmission des données de partitionnement à chaque mise à jour.

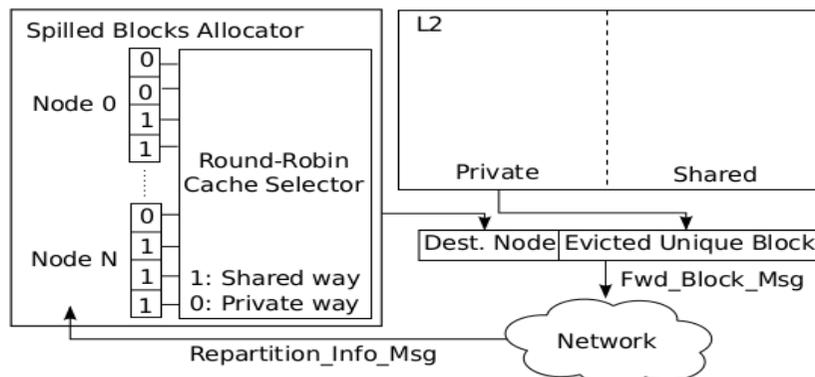


Figure 7 : Mécanisme d'allocation des blocks

```

If Private_LRU_Hit then
    Increase Counter
EndIf
If Shared_LRU_Hit then
    Decrease Counter
EndIf
If Repartition_Cycle then
    If Counter > Upper_Threshold then
        Add_Private_Way
        Send Repartition_Info_Msg
    ElseIf Counter < Lower_Threshold then
        Add_Shared_Way
        Send Repartition_Info_Msg
    EndIf
    Clear Counter
EndIf
    
```

Figure 6 : Protocole de partitionnement

CONTRIBUTION :
MECANISME DE GLISSEMENT DE
DONNEES

Description du principe de glissement de données

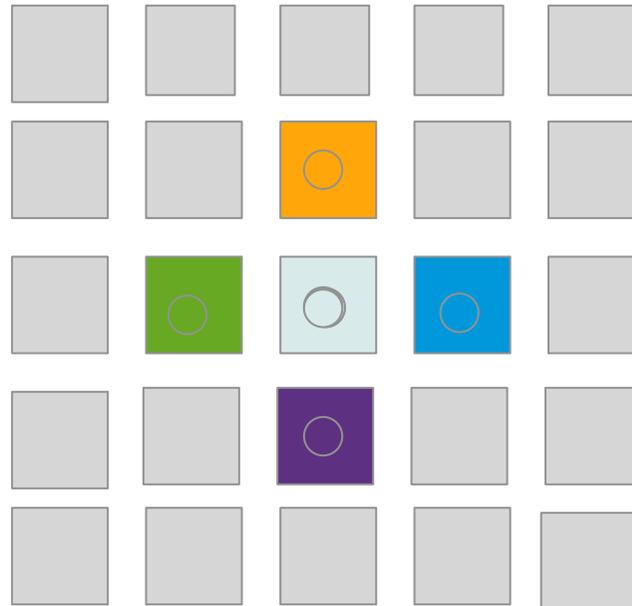


Figure 8: Glissement de données

Technique de remplacement par priorité

■ Technique des compteurs :

- LHC : Local Hit Counter
- NHC : Neighbor Hit Counter

■ Remplacement par priorité :

- LHC > somme(NHC) : LRU partagé remplacé
- LHC < somme(NHC) : LRU privé remplacé
- Distance (LHC , somme(NHC)) < TH_Sliding : activation du glissement des données.

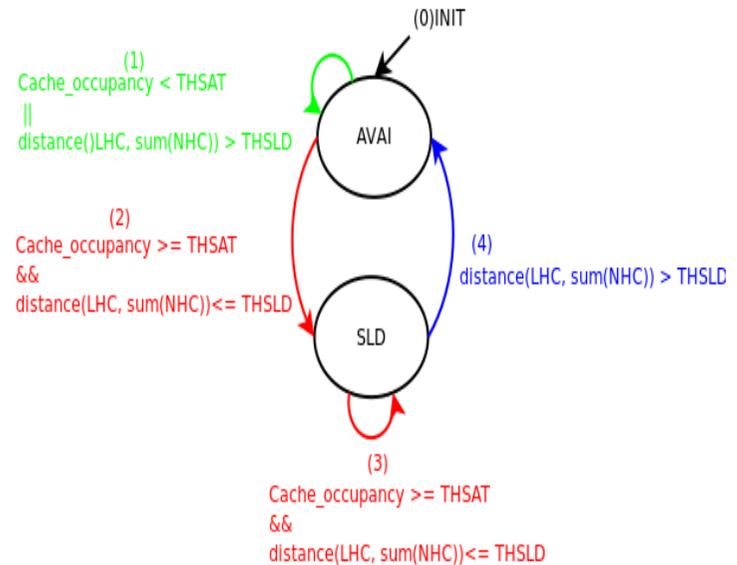


Figure 9 : Machine d'état du protocole élastique par glissement

=> Le choix du bloc qui va migrer vers le voisin proche prend en compte la fréquence d'accès aux données.

Technique du choix du meilleur voisin (Best Neighbor)

■ Principe:

Le **meilleur voisin** est celui qui sollicite le moins la zone partagée $\min(\text{NHC})$.

■ Implémentation :

Technique des compteurs

=> Le choix du nœud destinataire prend en compte le besoin en mémoire de celui-ci.

PROCEDURE DE TEST

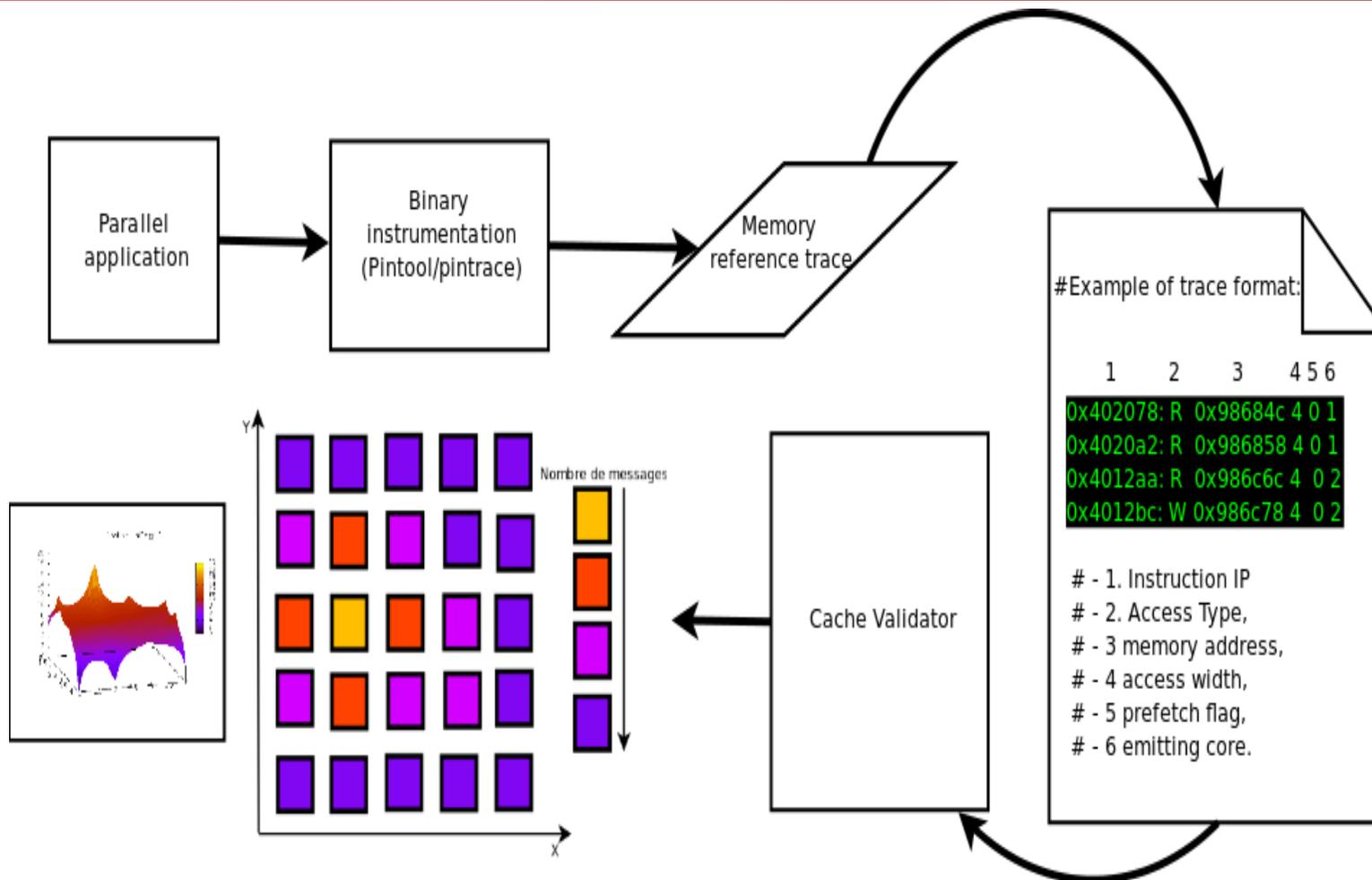


Figure 10 : Processus de validation

■ Évaluation du trafic réseau:

Message vers → Home Node (Défaut de cache)

Message vers → Neighbor (Récupération de la donnée de chez le voisin)

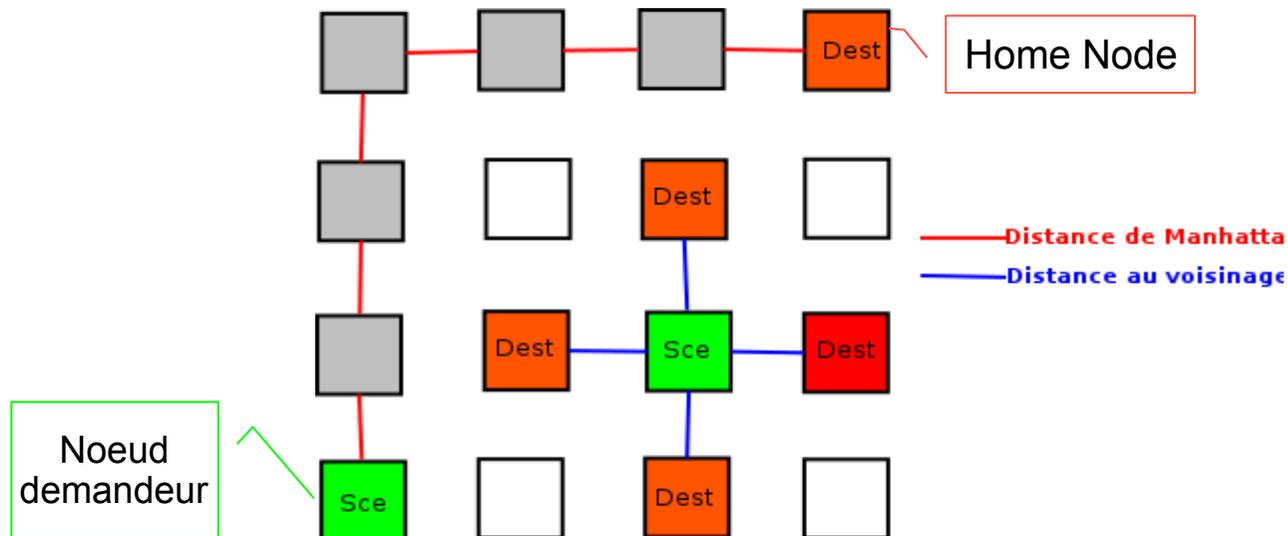
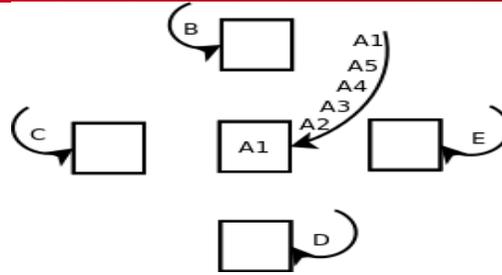
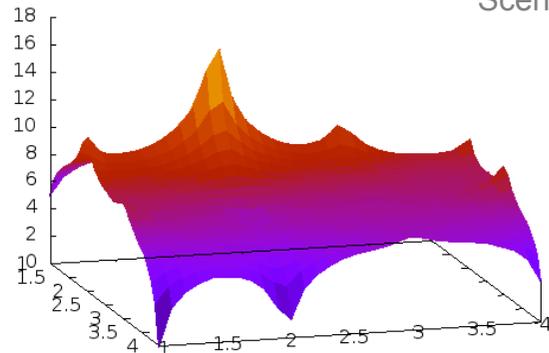


Figure 11 : Accès au Home Node versus accès au voisin



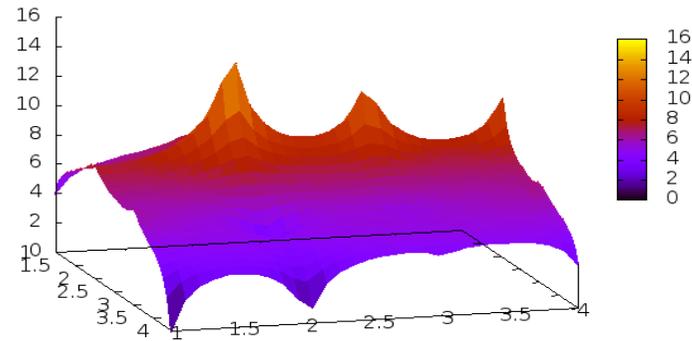
Scénario de test

Nb de Messages



Cache coopératif élastique

Nb de Messages

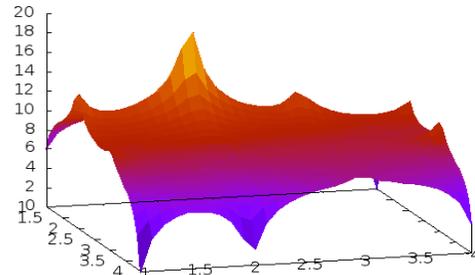


Glissement de données

- Réduit la zone de trafic sur puce : limitation des accès Home Node (défauts de cache).

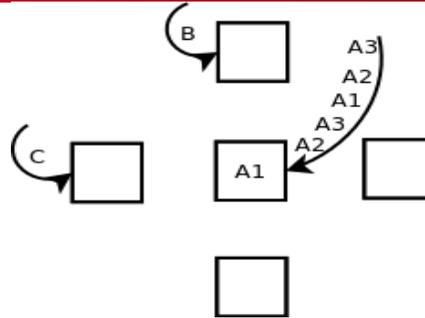
- Favorise la communication sur le voisinage proche.

Nb de Messages



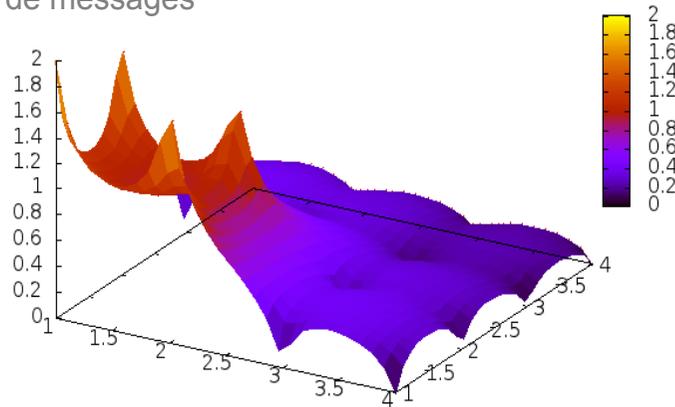
Protocole Baseline

TEST 2 : BEST NEIGHBOR



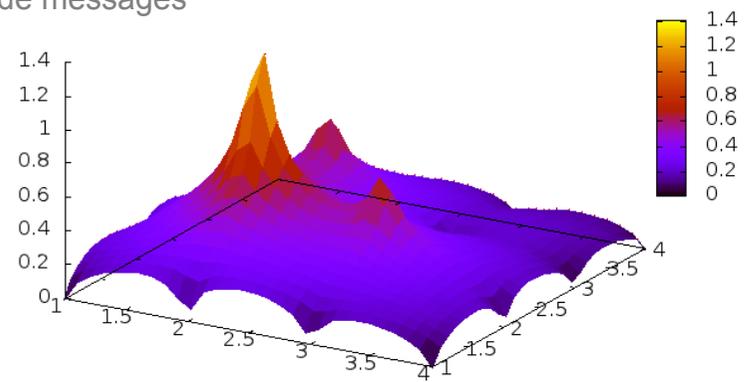
Scénario de test

Nb de messages



Round Robin

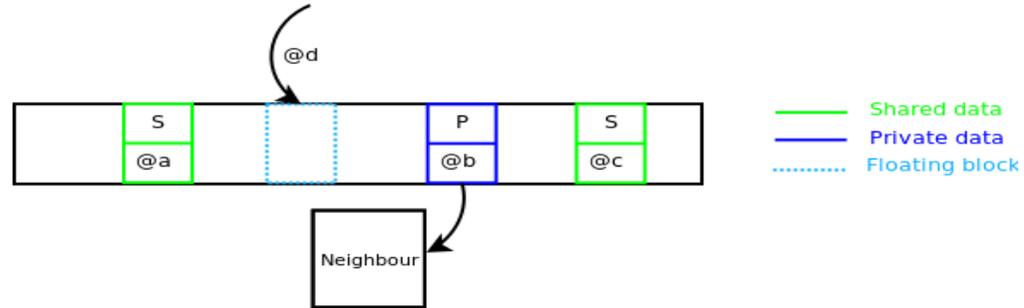
Nb de messages



Best Neighbor

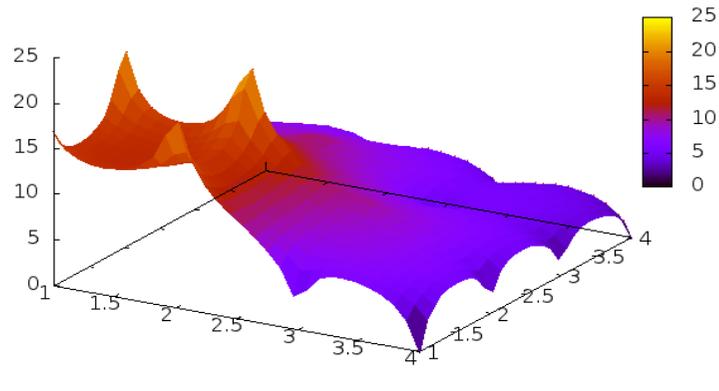
- Réduit la zone de trafic sur puce : limitation des accès Home Node (défauts de cache).

- Favorise la coopération avec les voisins libres plutôt que ceux chargés (d'où la pertinence du choix du voisin destinataire).



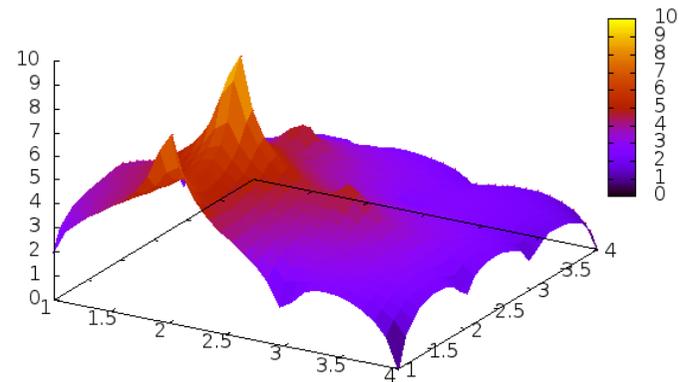
Scénario de remplacement des données

Nb de messages



Partitionnement cyclique

Nb de messages



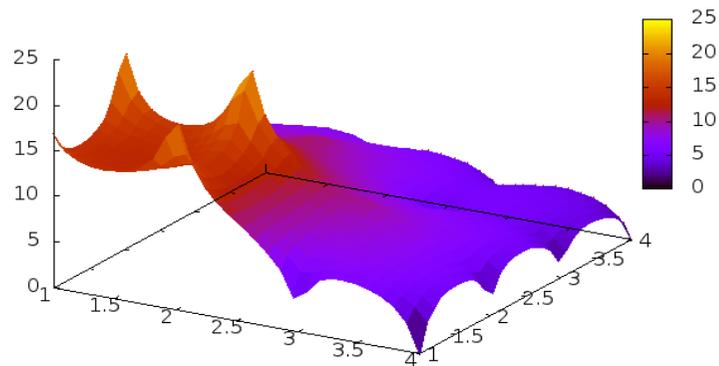
Remplacement par priorité

- Réduit le nombre max de messages : de 25 à 10 (-60 %).
=> Taux d'éjection réduit (moins de défauts de cache)



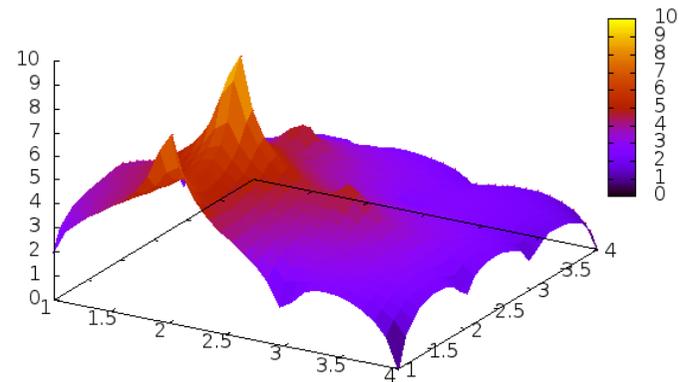
Scénario de remplacement des données

Nb de messages



Partitionnement cyclique

Nb de messages



Remplacement par priorité

- Réduit le nombre max de messages : de 25 à 10 (-60 %).
=> Taux d'éjection réduit (moins de défauts de cache)

■ Améliorations:

- Abandonne la notion de partitionnement cyclique.
- La prise de décision sur le bloc à migrer est effectuée à chaque requête de stockage

=> En cas de voisinage fortement stressé, le glissement de données de proche en proche réduit l'éjection hors puce.

CONCLUSION

- **Contribution:** Amélioration du modèle de cache coopératifs élastique pour architecture massivement parallèle par **glissement de données** :
- ** **Technique de partitionnement** : réduction du nombre de défauts de cache (60 %)
- ** **Technique du choix du meilleur voisin** : meilleure répartition de données sur le voisinage (limitation de l'éjection hors puce des données)
- **Perspectives:**
 - Evaluation des techniques de glissement sur plate-forme de simulation (ex : SocLib).
 - Utilisation d'application à forte charge de données (Compression de données, Traitement d'image, Streaming, Recuit simulé sur NUMA).
 - Glissement étendu (N-chance forwarding) basé sur une évaluation globale de la répartition de la charge au niveau de la puce.



MERCI !

Commissariat à l'énergie atomique et aux énergies alternatives
Centre de Saclay | 91191 Gif-sur-Yvette Cedex
T. +33 (0)1 69 08 60 00 | F. +33 (0)1 XX XX XX XX
Etablissement public à caractère industriel et commercial | RCS Paris B 775 685 019

Direction
Département
Service