

Evaluation de Performance pour différentes approches de tolérance aux pannes

Amri Hafedh
M1 HPC - UVSQ

Plan

- Introduction et contexte
- Objectifs
- Résultats sur Poincare
- Résultats sur Curie
- Conclusion & perspectives

Checkpointing

- Changement d'un composant plusieurs fois par semaine
 - Beaucoup de cœurs utilisé.
 - Beaucoup de temps passé.
- => Pannes matérielles ou logicielles quasi assurées & simulation risque de ne jamais terminée.
- Sauvegarder l'état de simulation sur un stockage stable, un disque indépendant (PFS) pour pouvoir reprendre en cas de panne

Introcution - Gysela5D

- Contexte

Gysela simule les turbulences qui se développent dans les plasma d'un Tokamak.

- Les approches I/O

- **Checkpoint Synchrone.**
- **Checkpoint Asynchrone.**

- Checkpointing en diagnostics : Ce sont des données écrites plus fréquemment qui seront destinés au physiciens.

Introcution - Gysela5D

- **Checkpoint Synchrone** : Calculs puis arrêt pour faire des checkpoints puis continuer les calculs.
- **Checkpoint Asynchrone** : Checkpoint sur le RAM puis reprise du calcul et copie par un thread sur le PFS.

Synchrone		Asynchrone	
Avantages	Inconvénients	Avantages	Inconvénients
- Pas de consommation de RAM	- Moins rapide	- Plus rapide	- Consommation de beaucoup de RAM

Introcution - Gysela5D

- NPROC_R
- NPROC_TH
- NPROC_MU

Paramètres pour
affecter le nombre de
processus

- Nr
- Ntheta
- Nphi
- Nvpar
- Nmu

Paramètres pour la taille
de la grille.

Introcution - Gysela5D

- Paramètres

- asynchrone_wrinting : choix du mode de checkpoint

- Deltat

- dt_diag

Fréquence d'écriture des
diagnostics

- Phi3D_dt_diag

- Moments3D_dt_diag

Introcution - FTI

- FTI est une bibliothèque pour faire du checkpoints
- Utilise des disques SSD locaux aux nœuds de calculs pour accélérer les checkpoints.
- Utilise 1 processus dédié par noeud
- Quatre niveaux dans FTI

Niveaux de FTI

- Level 1 : Chaque processus écrit son checkpoint sur le SSD.

Avantages	Inconvénients
<ul style="list-style-type: none">- Vitesse d'écriture ne diminue pas quand le nombre de nœud augmente.	<ul style="list-style-type: none">- Ne supporte pas les pannes matériels

Niveaux de FTI

- Level 2 : Les processus s'envoient 2 par 2 leurs checkpoints.

Avantages	Inconvénients
- Supporte une panne matérielle	- Temps d'échange des données supplémentaire.

Niveaux de FTI

- Level 3 : Echange des checkpoints entre les processus d'un même groupe (Reed-Solomon).

Avantages	Inconvénients
<ul style="list-style-type: none">- Supporte le nombre de crash jusqu'à la moitié du groupe.	<ul style="list-style-type: none">- Encodage Reed-Solomon prends du temps de calculs et temps de communication.

Niveaux de FTI

- Level 4 : Ecrire un checkpoint asynchrone dans le PFS.

Avantages	Inconvénients
- Supporte tout nombre de crash.	- Plus on a nombre de processus qui écrivent sur le PFS plus ça se passe mal.

Objectifs du stage

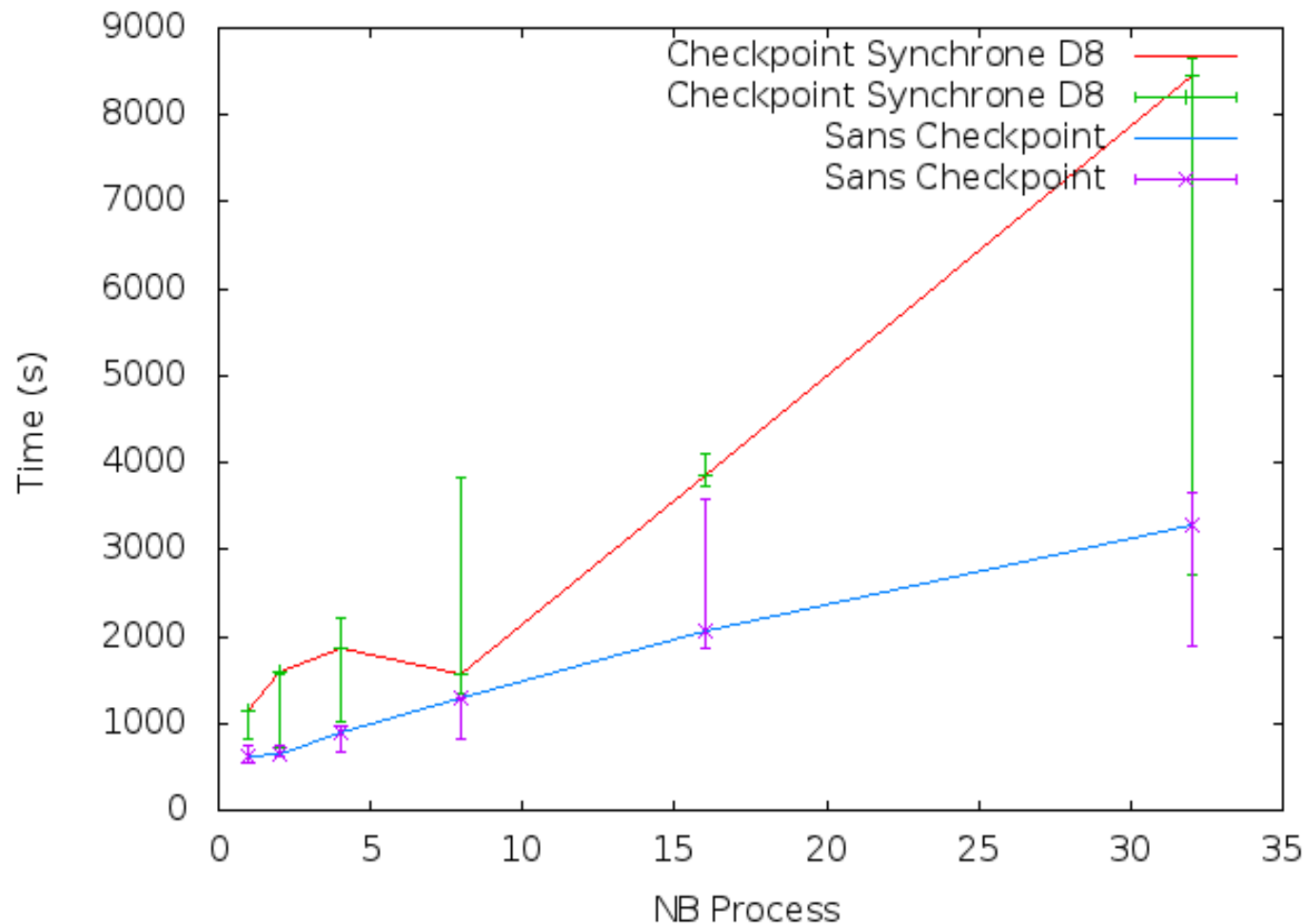
- Evaluation de performance des approches :
 - Checkpoint Synchrone
 - Checkpoint Asynchrone
 - Checkpoint FTI
- Analyse du comportement de chaque approche sur les supercalculateurs
 - Poincare
 - Curie

Poincare - Présentation

- Calculateur IBM, principalement composé de serveurs iDataPlex dx360 M4
 - 92 Noeuds de calculs : 16coeurs par nœud, 32Go de mémoire par nœud.
 - 4 Noeuds GPU : 4 processeurs, 32 Go par nœuds.
- GPFS (General parallel file system) : partagé par l'ensemble des nœuds
 - /gpfs1l : 7 To, contient les $\${HOME}$ ainsi que l'environnement logiciel partagé
 - /gpfsdata : 37 To, contient un répertoire de travail pour les grosses données

Poincare: Mesures de performances

Weak Scaling Checkpoint Synchrone et Sans Checkpoint avec D8



Nr : 63

Ntetha : 512

Nphi : 128

Nvpar : 60

Nmu : 1 à 32

-

Deltat : 5

dt_diag : 40

Phi3D_dt_diag : 40

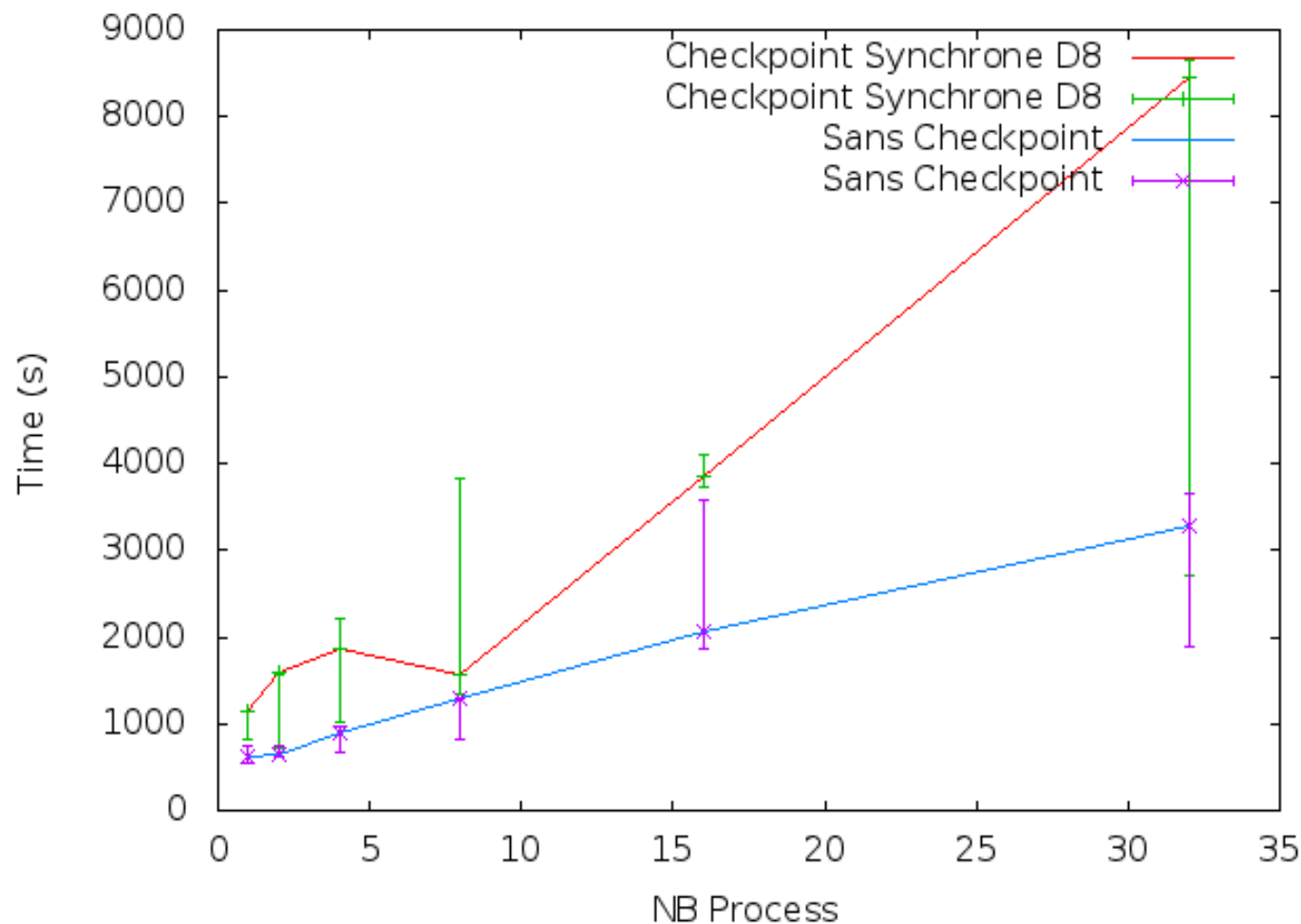
Moments3D_dt_diag : 40

-

asynchrone_writing : false

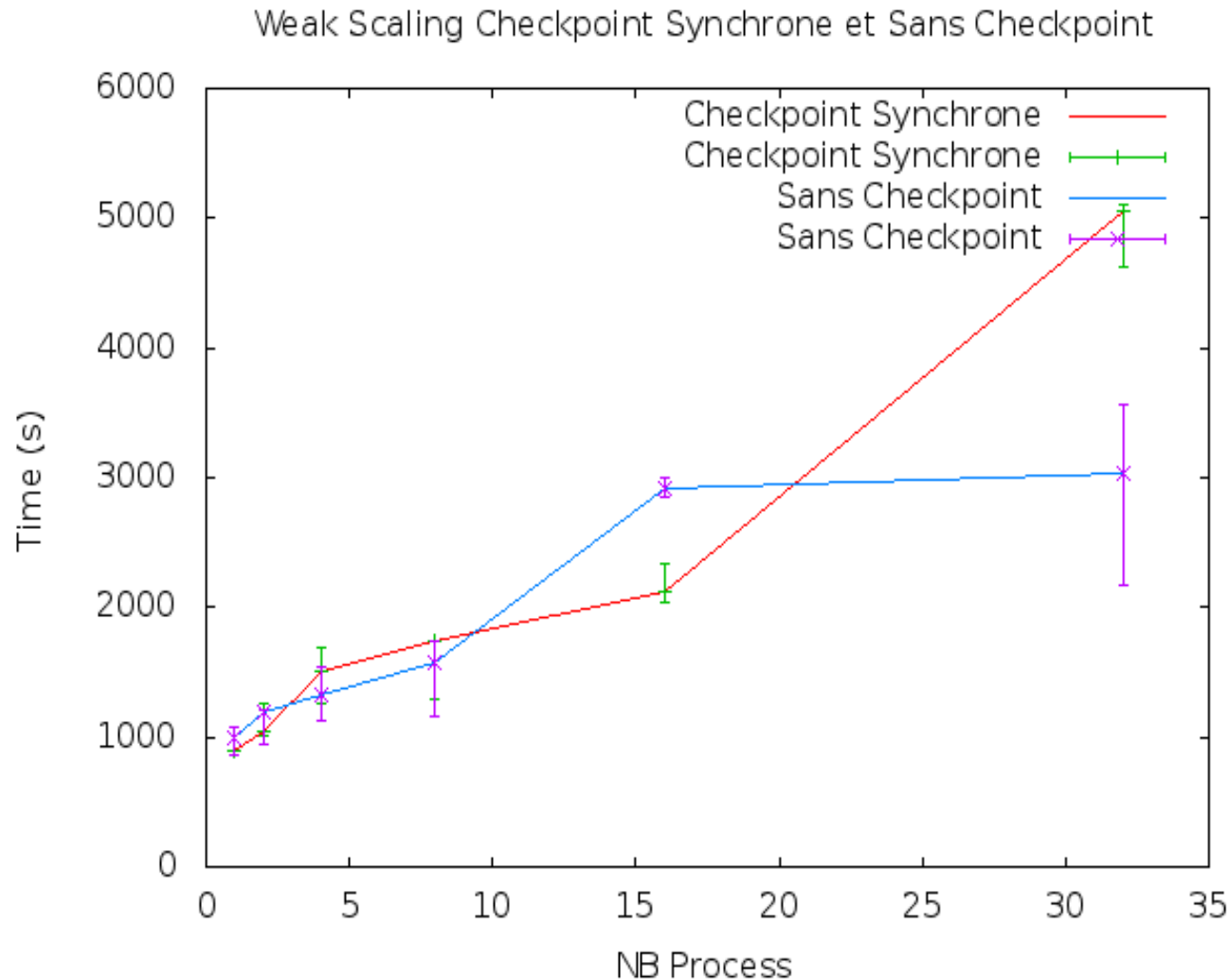
Poincare: Mesures de performances

Weak Scaling Checkpoint Synchrone et Sans Checkpoint avec D8



- Voir la limite du PFS.
- Weak Scaling avec checkpoint Synchrone avec des diagnostics tout les 8 itérations (en rouge) et Weak Scaling sans checkpoint (en bleu).
- les 2 courbes croient en augmentant le nombre de ps.
- Intervalle d'erreur est très long sur 32 ps.
- Pour savoir s'il y a des perturbations, faire des tests pour avoir un intervalle de tolérance.¹⁶

Poincare: Mesures de performances



Nr : 63

Ntetha : 512

Nphi : 128

Nvpar : 60

Nmu : 1 à 32

-

Deltat : 5

dt_diag : 5

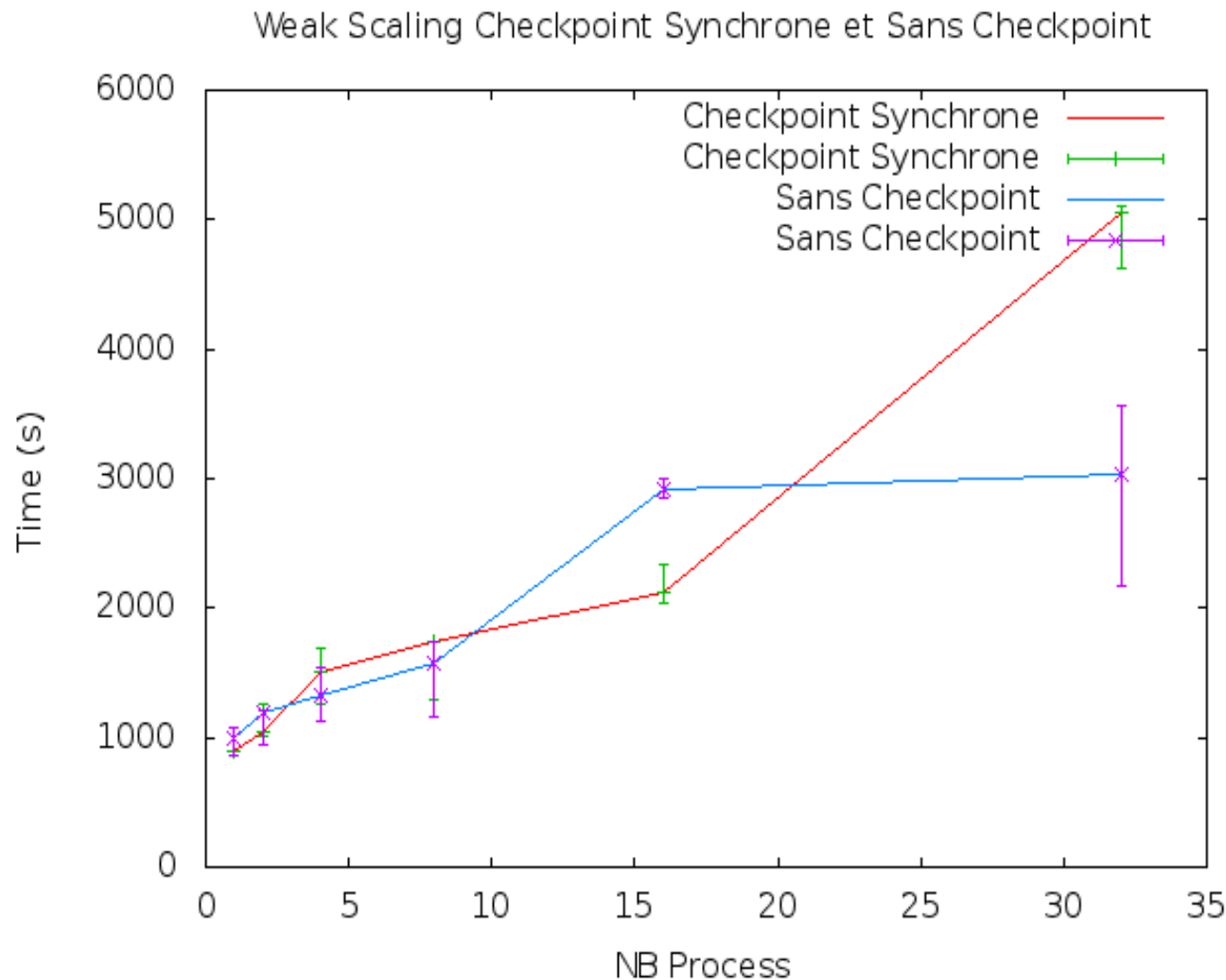
Phi3D_dt_diag : 5

Moments3D_dt_diag : 5

-

asynchrone_writing : false

Poincare: Mesures de performances



- Voir l'impacte des diagnostics sur les Checkpoints

- Courbe du weak scaling avec checkpoint synchrone croient en augmentant le nombre de ps

- Les pts où on utilise 1, 2, 16 ps on un temps de checkpoint négatifs, difficile a expliquer.

- Les résultats ne sont pas significatifs.

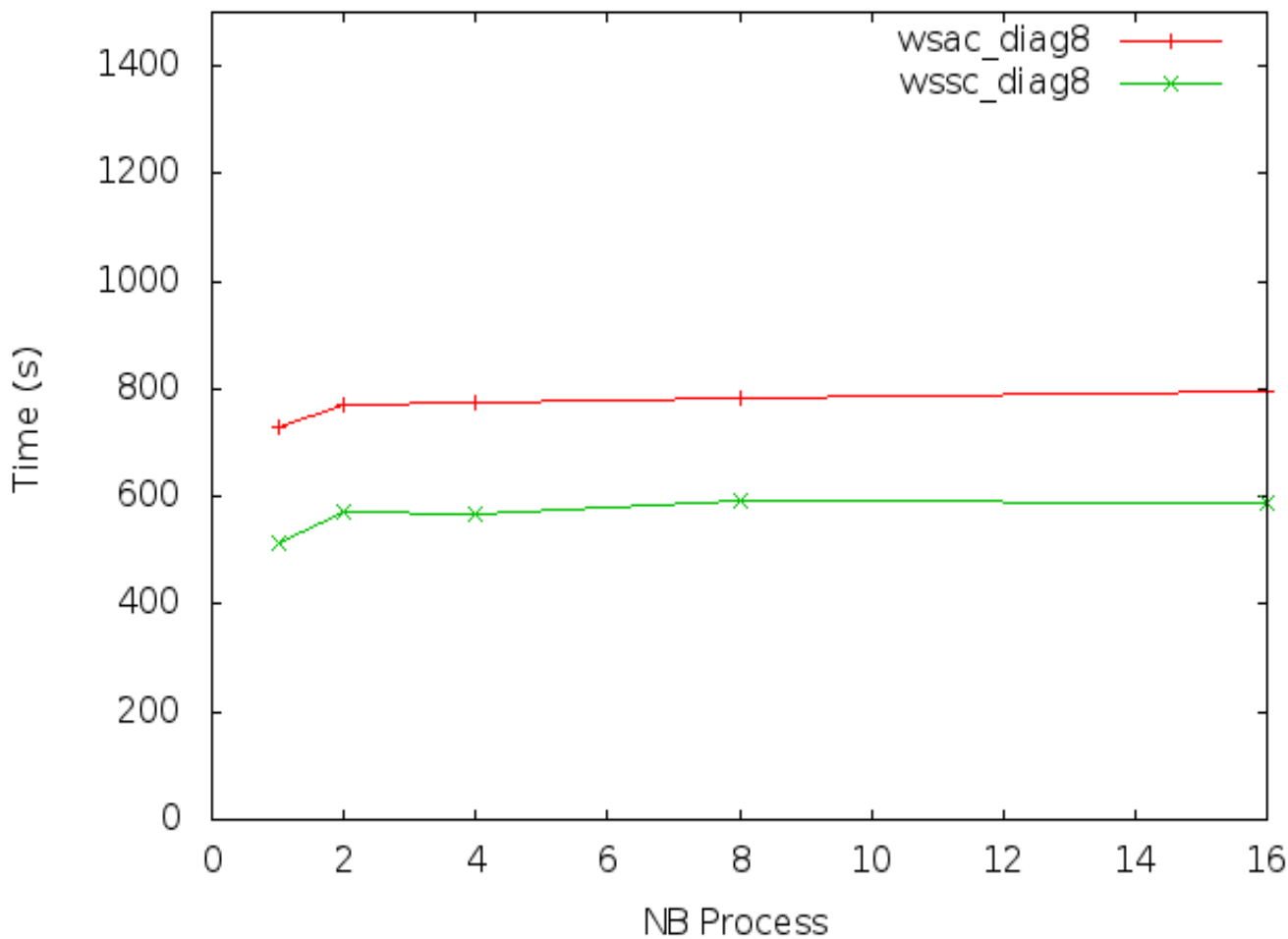
- perturbations dû au nombre des jobs des autres utilisateurs

Curie - Présentation

- Fat Nodes
 - 360 Nodes
 - 4 Processeurs Octo-coeurs
- GPU Nodes
 - 288 Processeurs Intel + 288 Processeurs Nvidia
- Thin Nodes
 - 5040 Nodes
 - 2 Processeurs Octo-coeurs
 - 1 Disque local SSD
- Système de fichiers : 100 Go/s de bande passante théorique

Curie – Weak Scaling

Weak Scaling avec et sans checkpoint d8-- Curie



Nr : 63

Ntetha : 512

Nphi : 128

Nvpar : 60

Nmu : 1 à 16

-

Deltat : 5

dt_diag : 40

Phi3D_dt_diag : 40

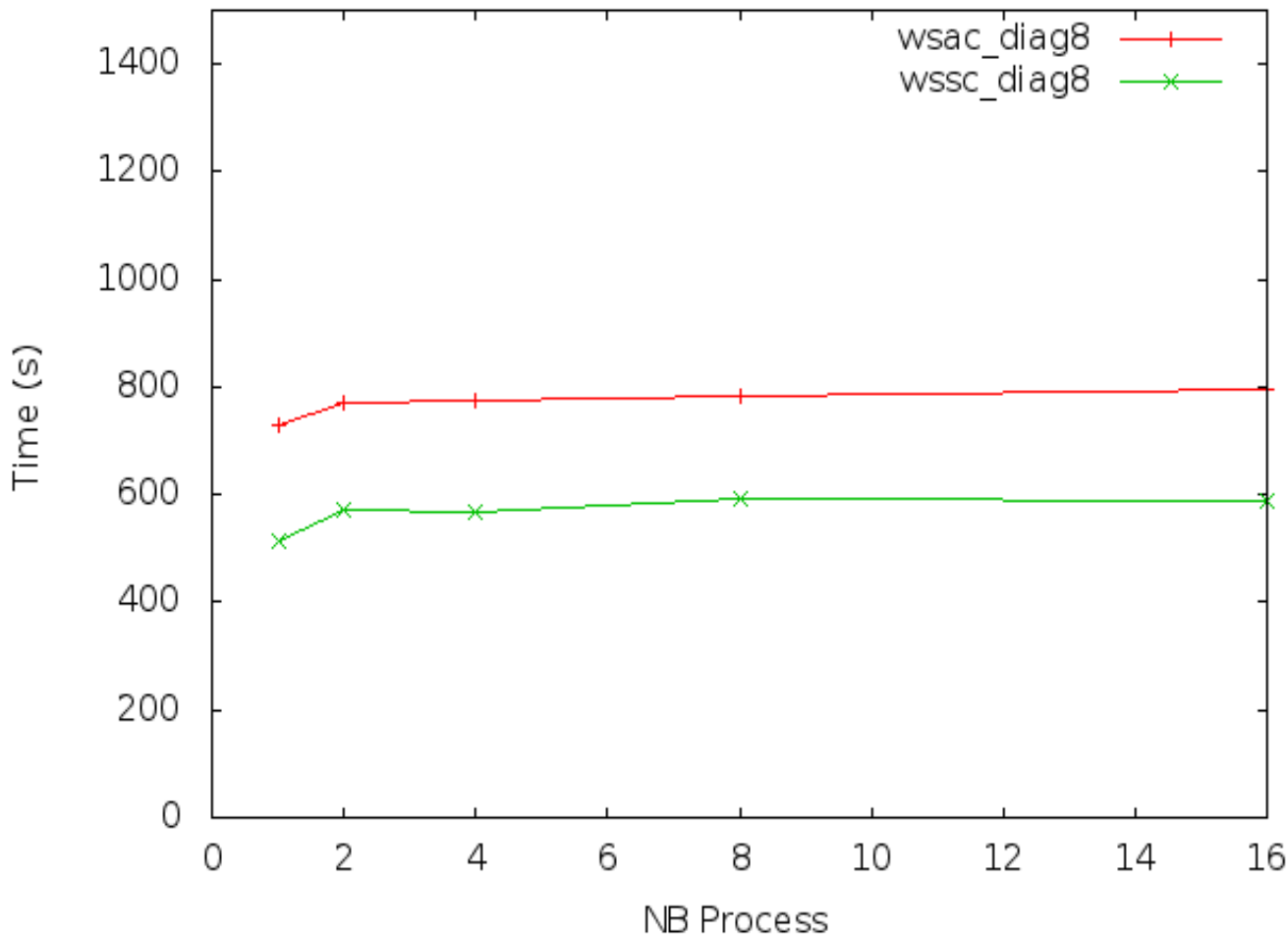
Moments3D_dt_diag : 40

-

asynchrone_writing : false

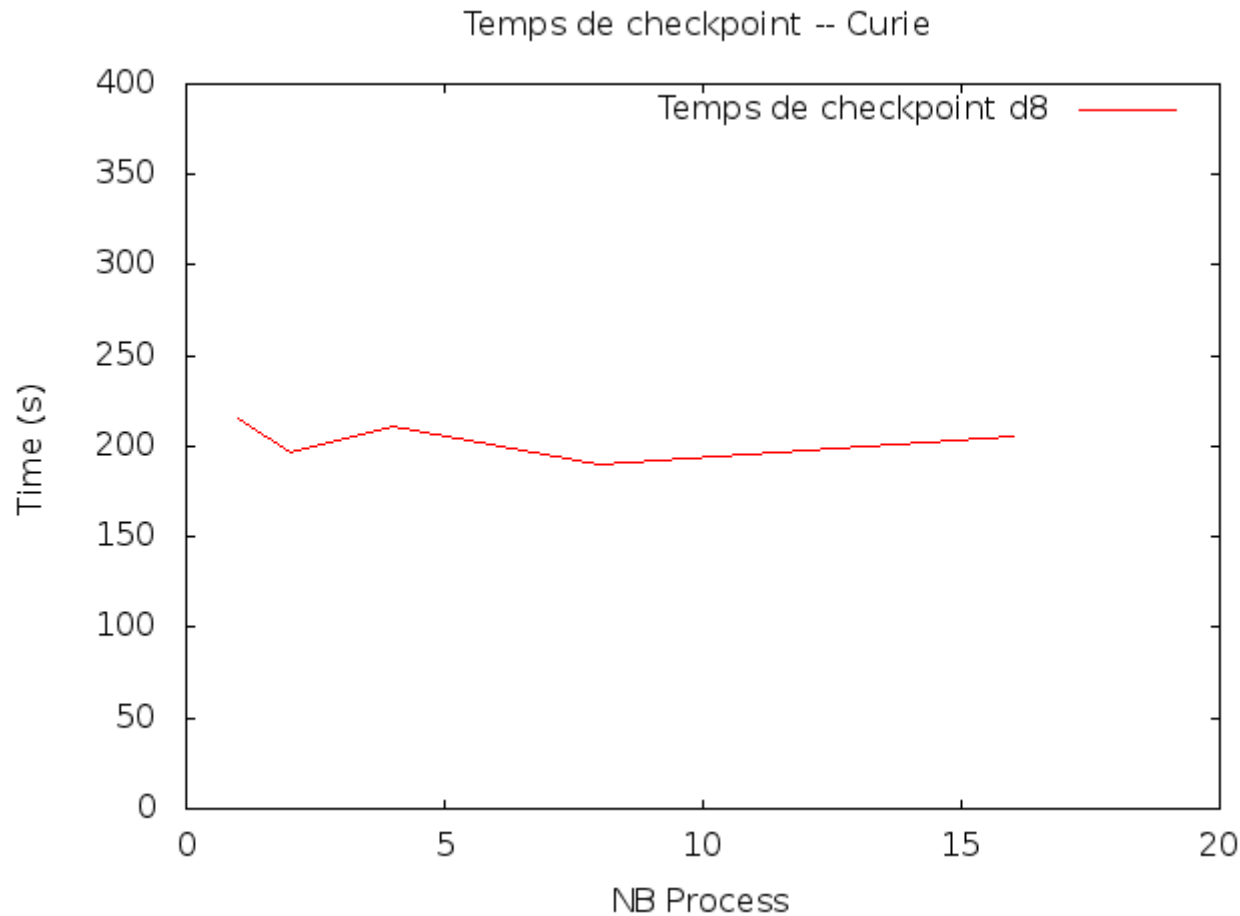
Curie – Weak Scaling

Weak Scaling avec et sans checkpoint d8-- Curie



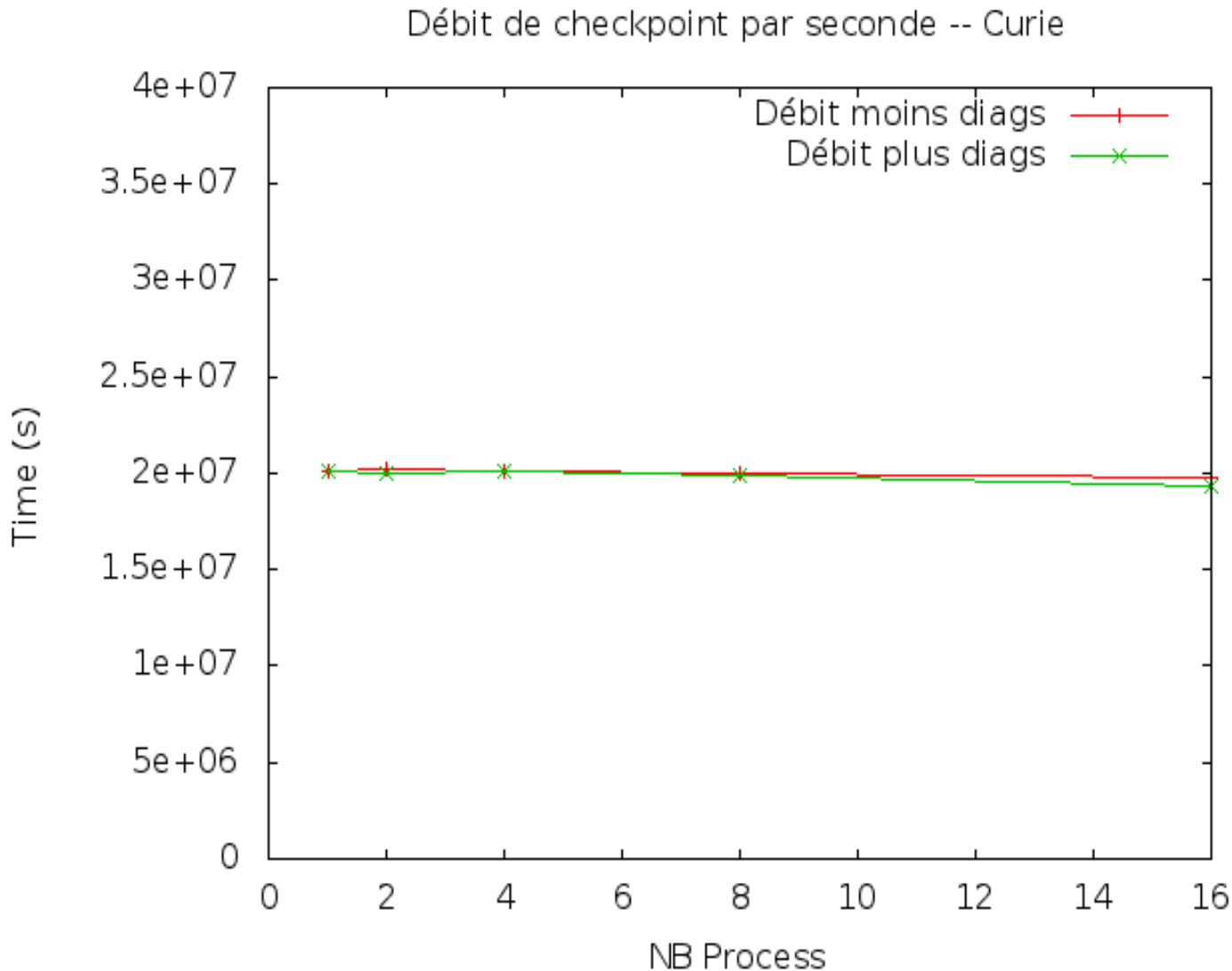
- Pour voir est-ce qu'on atteint la limite du PFS.
- Weak Scaling Avec checkpoint synchrone avec moins de diagnostics et Weak Scaling sans checkpoint.
- Stabilité des 2 courbes à partir de 2 processus.
- Résultats attendu, courbes plates en augmentant le nombre de processus.

Curie – Weak Scaling



- La différence entre le temps de weak scaling avec checkpoint et le temps du weak scaling sans checkpoint.
- Courbe constante en augmentant le nombre de processus.

Curie – Débit de checkpoint



- Vitesse de checkpoint par seconde.

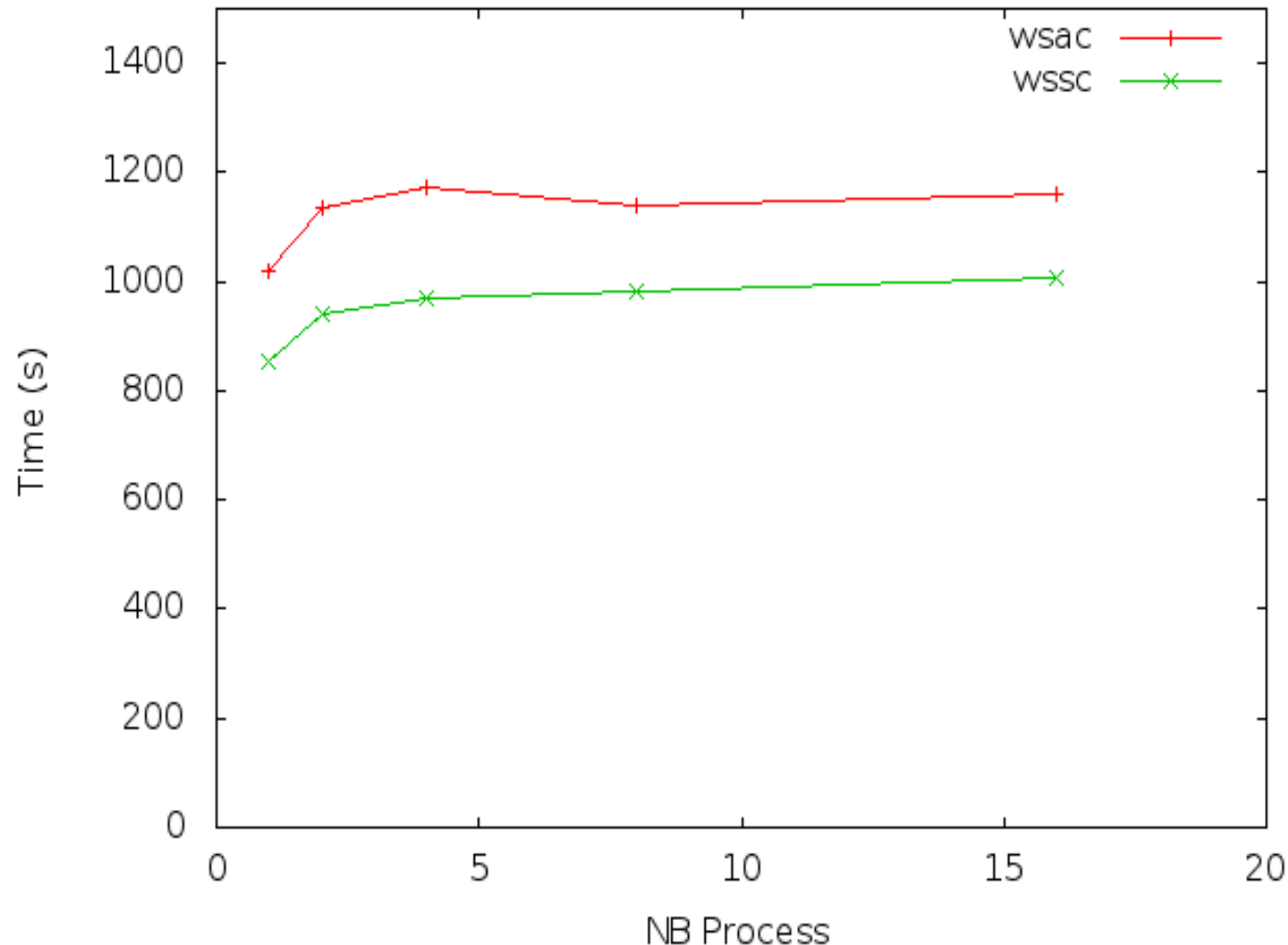
- 20Mo/s : c'est le débit par nœud.

- Stabilité des 2 courbes sur 16 premiers processus.

- $16 \text{ ps} * 20 \text{ Mo/s} = 360 \text{ Mo/s}$ (théorique 100Go/s donc normal)

Curie – Weak Scaling

Weak Scaling avec et sans checkpoint -- Curie



Nr : 63

Ntetha : 512

Nphi : 128

Nvpar : 60

Nmu : 1 à 16

-

Deltat : 5

dt_diag : 5

Phi3D_dt_diag : 5

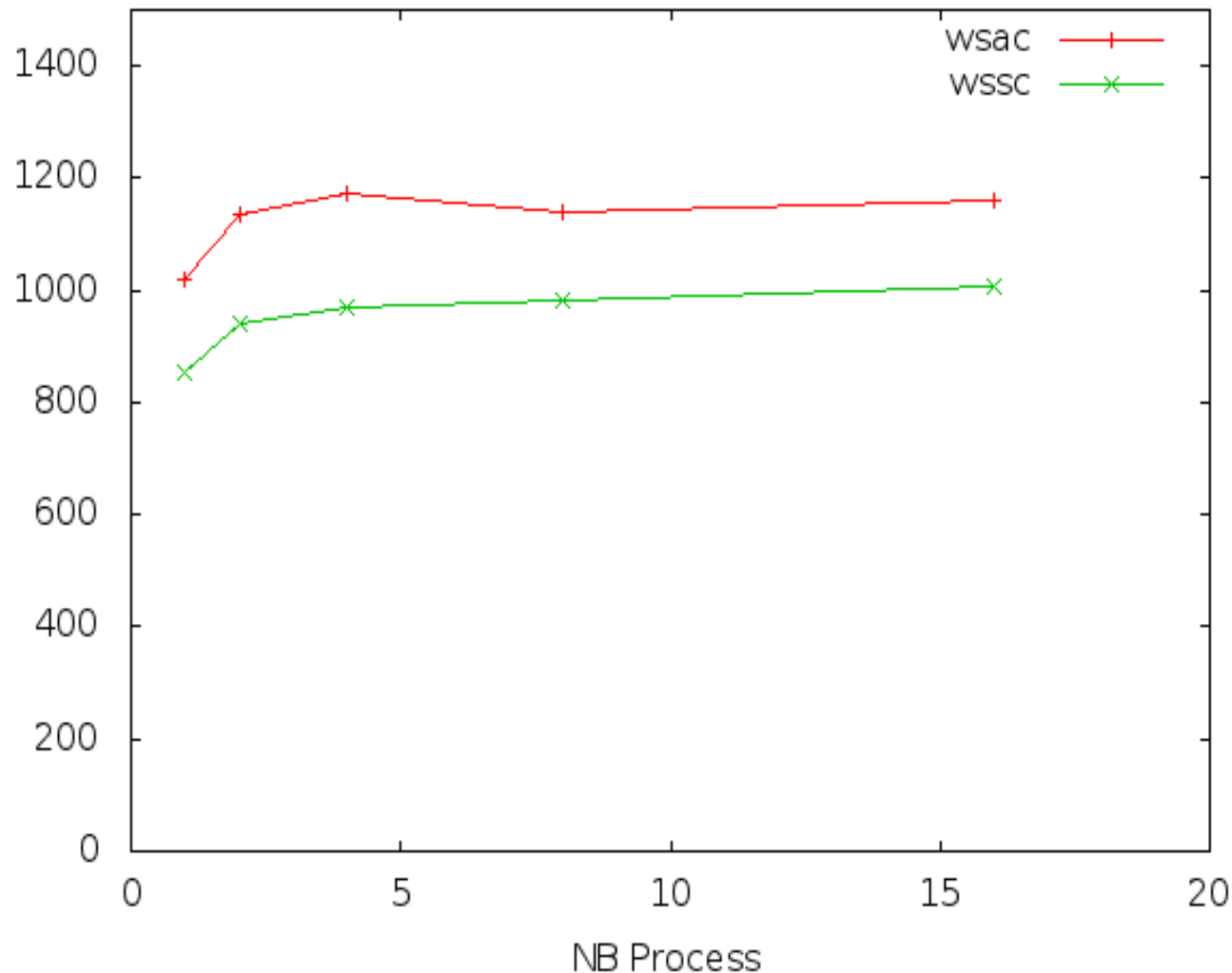
Moments3D_dt_diag : 5

-

asynchrone_writing : false

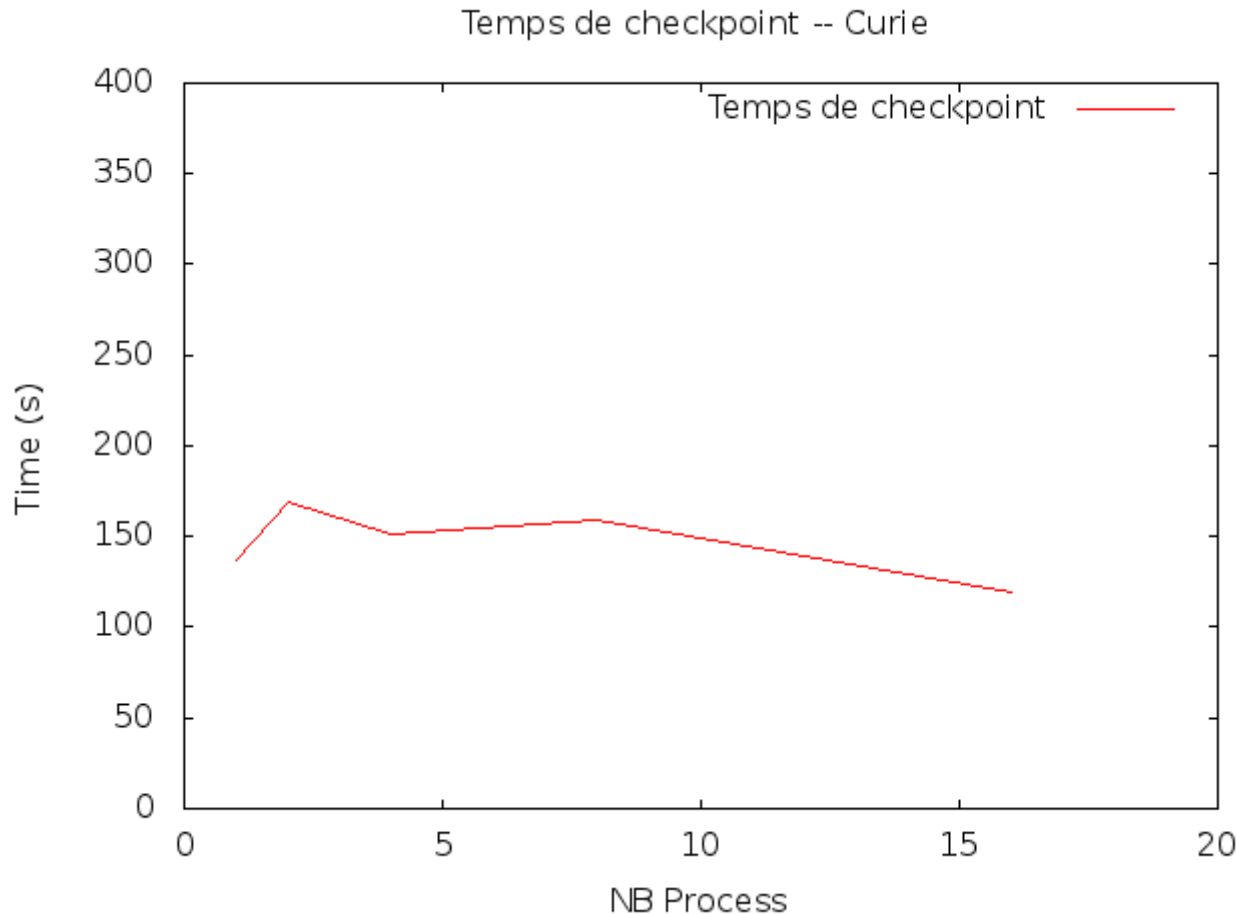
Curie – Weak Scaling

Weak Scaling avec et sans checkpoint -- Curie



- Voir l'impacte des diagnostics sur les Checkpoints.
- Weak Scaling avec un checkpoint synchrone et un Weak Scaling sans checkpoint.
- Les 2 courbes sont stable à partir de 2 processus.
- Par rapport a poincare on a moins de perturbations

Curie – Weak Scaling



- Différence entre le temps de Weak Scaling avec checkpoint synchrone avec beaucoup de diagnostics et le temps de Weak Scaling sans checkpoint.
- Courbe constante en augmentant le nombre de processus.

Conclusion & Perspective

- Mesures sur poincare ne donne pas des résultats pertinents.
- Résultats sur Curie c'est très stable, et que jusqu'à 16 nœuds ce n'est pas suffisant pour atteindre la limite du PFS.
- Implémentation de FTI n'est pas encore débogué pour l'évaluer.
- Evaluer le mode asynchrone.