

Monitoring du noyau Linux

sur une architecture NUMA

Kevin Gallardo
Eric Lombardet
Pierre-Yves Péneau

Université Pierre et Marie Curie

12 *Mai* 2014

Introduction

- Problématique:
architectures NUMA, placement mémoire, performances

Introduction

- Problématique:
architectures NUMA, placement mémoire, performances
- Objectif:
évaluation d'activité, mesures d'évènements, étude
comportementale

Infrastructure de tests

- utilisation mutualisée du Magny Cour → machines virtuelles
- problème: pas d'IBS avec qemu

Infrastructure de tests

- utilisation mutualisée du Magny Cour → machines virtuelles
- problème: pas d'IBS avec qemu

Conséquence

travail en réel sur le noyau pour 50% du projet

Monitoring

Qu'est-ce que c'est ?

- étude bas niveau du comportement matériel et système
- très utile pour le débogage ou l'optimisation poussée
- différentes solutions de monitoring existent

Monitoring

Instruction Based Sampling - Présentation

- technologie AMD
- informations plus précises car spécifique à une famille de processeur
- problème:w
 - ▶ plus difficile à mettre en place

Monitoring

Instruction Based Sampling - Fonctionnement

- tag aléatoirement une instruction
- suivi de l'exécution
- deux types de mesures: fetch/execution sampling

Monitoring

Instruction Based Sampling - Utilisation

- beaucoup d'informations remontées par IBS
- sélection des plus utiles: cache hit/miss

Monitoring

Instruction Based Sampling - Utilisation

- beaucoup d'informations remontées par IBS
- sélection des plus utiles: cache hit/miss

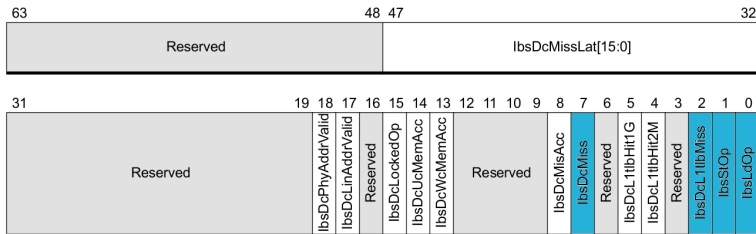


Figure: Schéma du registre MSR IbsOpData3

Monitoring

Instruction Based Sampling - Défauts

- overhead: traitement coûteux des mesures
- pas de vision d'ensemble

Monitoring

Mise en place

- configuration de l'APIC
 - ▶ informer l'APIC de la présence d'interruptions IBS
 - ▶ à faire pour chaque coeur
- enregistrement d'un handler NMI
 - ▶ appelé à chaque interruption IBS
 - ▶ récolte les informations dans les registres MSR

Monitoring

Mise en place

- configuration de l'APIC
 - ▶ informer l'APIC de la présence d'interruptions IBS
 - ▶ à faire pour chaque cœur
- enregistrement d'un handler NMI
 - ▶ appelé à chaque interruption IBS
 - ▶ récolte les informations dans les registres MSR

Attention

le handler doit être enregistré une et une seule fois au niveau du système

Conclusion

Mise en place

20/20 ou chinois tuer et manger ta famille !



"Tu veux un bol de riz ?"