

Resum

La jerarquia de cache i la xarxa en el xip (NoC) són dos components clau dels xips multiprocessador (CMPs). La majoria del trànsit en la NoC es deu a missatges que les caches envien segons el que estableix el protocol de coherència. La quantitat de trànsit, el percentatge de missatges curts i llargs i el patró de trànsit en general varien depenent de la geometria de les caches i del protocol de coherència. L'arquitectura de la NoC i la jerarquia de cache estan de fet fermament acoblades, i aquests dos components han de ser dissenyats i evaluats conjuntament per a estudiar com al variar un afecta a les prestacions de l'altre. A més, cada component s'ha d'ajustar als requisits i a les oportunitats de l'altre, i a l'inrevés. Normalment diferents classes de missatges s'envien per diferents xarxes virtuals o per NoCs amb diferent ample de banda , separant missatges llargs i curts. No obstant això, una altra classificació dels missatges es pot fer depenent del tipus d'informació que proveeixen: alguns missatges, com les peticions de dades, necessiten camps per aemmagatzemar informació (adreça del bloc, tipus de petició, etc.). Altres, com els missatges de reconeixement (ACK), no proporcionen cap informació excepte per l'ID del node destí. Aquesta segona classe de missatge no necessita molt ample de banda: la latència és molt més important, atès que el node destí està típicament bloquejat esperant la recepció d'ells. En aquest treball de tesi es desenvolupa una xarxa dedicada per a transmetre la segona classe de missatges, la xarxa és molt senzilla i ràpida, i permet el lliurament dels ACKs amb una latència de pocs cicles de rellotge. Reduint la latència i el trànsit en la NoC a causa dels ACKs, és possible:

- accelerar la fase d'invalidació en fase d'escriptura en un sistema que utilitza un protocol de coherència basat en directoris
- millorar les prestacions d'un protocol de coherència basat en broadcast , fins a aplegar a prestacions comparables amb les d'un protocol de directoris però sense el cost d'àrea a causa de la necessitat d'emmagatzemar el directori
- implementar un mapejat dinàmic de blocs a les caches d'últim nivell de forma eficient , amb l'objectiu d'apropar quant al màxim possible els blocs als cores que els utilitzen

L'objectiu final és obtenir un co- disseny de NoC i jerarquia de cache que minimitze els problemes d'escalabilitat dels protocols de coherència. En aquesta tesi s'exploren diferents alternatives per un lliurament ràpid dels ACKs i les oportunitats que això ofereix

al protocol de coherència. Combinant els mecanismes presentats en un sistema final, s'obté una arquitectura adaptable dinàmicament als requisits de múltiples aplicacions en un entorn virtualitzat.