

Resum

Escola Tècnica Superior d'Enginyeria Informàtica
Departament d'Informàtica de Sistemes i Computadors

Doctor en Filosofia
(Enginyeria Informàtica)

Tècniques Dinàmiques amb Control de Consum per a Sistemes Encastats Multinucli de Temps Real

per José Luis March Cabrelles

La contínua reducció de la grandària dels transistors ha permès la implementació de dispositius més complexos i potents en una mateixa àrea, proporcionant una major capacitat i funcionalitat. No obstant, aquest increment en la complexitat comporta un considerable augment del consum energètic. Aquesta situació és crítica en dispositius mòbils, on la capacitat energètica està limitada i, per tant, la vida útil de la bateria defineix la usabilitat del sistema. En conseqüència, el consum energètic s'ha convertit en un aspecte crucial en el disseny de sistemes encastats multinucli de temps real.

Aquesta dissertació proposa diverses tècniques amb l'objectiu d'estalviar energia sense sacrificar la planificabilitat de temps real en aquest tipus de sistemes. Les tècniques proposades actuen sobre diferents components principals del sistema. En particular, aquestes tècniques afecten el particionador de tasques i el planificador, així com el controlador de memòria.

Algunes d'aquestes tècniques estan especialment concebudes per a sistemes multinucli amb dominis DVFS (*Dynamic Voltage and Frequency Scaling*) compartits. Equilibrar la càrrega entre els nuclis en un domini donat té un fort impacte en el consum d'energia, ja que tots els nuclis que comparteixen un domini DVFS han de funcionar a la velocitat requerida pel nucli més carregat.

En aquesta tesi es proposa un nou algorisme de **particionat de càrrega**, anomenat *Load-bounded Resource Balancing (LRB)*. La proposta assigna les tasques als nuclis per equilibrar el consum d'un recurs donat (processador o memòria) entre els nuclis, millorant la planificabilitat de temps real en incrementar el solapament entre processador i memòria. No obstant, distribuir les tasques d'aquesta manera sense tenir en compte la

utilització individual dels nuclis podria donar lloc a distribucions de càrrega desequilibrades. És a dir, un dels nuclis podria estar molt més carregat que la resta. Per tal d'evitar aquest escenari, quan se supera un llindar d'utilització donat, les tasques són assignades al nucli menys carregat.

Desafortunadament, només amb el particionat de càrrega de vegades no és suficient per aconseguir un bon equilibrat de càrrega entre els nuclis. Per tant, aquest treball també explora noves aproximacions basades en **migració de tasques**. Es proposen dues heurístiques amb migració de tasques. La primera heurística, anomenada *Single Option Migration (SOM)*, intenta realitzar només una migració quan la càrrega canvia per millorar l'equilibri de la utilització. S'han ideat tres variants de l'algorisme *SOM*, en funció de l'instant de temps en què es realitza l'intent de migració: quan una tasca arriba al sistema (SOM_{in}), quan una tasca surt del sistema (SOM_{out}), i en ambdós casos (SOM_{in-out}). La segona heurística, anomenada *Multiple Option Migration (MOM)*, explora un particionat de càrrega alternatiu addicional abans de fer l'intent de migració.

Respecte al controlador de memòria, s'han concebut dos **polítics de planificació per al controlador de memòria**. Les polítiques convencionals usades en sistemes que no són de temps real no són apropiades per a sistemes que donen suport a tasques tant de temps real estricte (*Hard Real-Time* o HRT) com de temps real flexible (*Soft Real-Time* o SRT). Aquestes polítiques poden introduir variabilitat en les latències de les peticions de memòria i, per tant, causar una pèrdua de deadlines que podria comportar una fallada crítica del sistema de temps real. Per tractar aquest inconvenient, es proposa una política simple, anomenada *HR-first*, que prioritza peticions de tasques HRT. A més, es presenta una aproximació més elaborada, anomenada *ATR-first*. *ATR-first* prioritza només aquelles tasques HRT que són necessàries per a assegurar la planificabilitat de temps real, millorant la qualitat de servei (*Quality of Service* o QoS) de les tasques SRT.

Finalment, aquesta tesi també aborda la **estimació dinàmica del temps d'execució**. La precisió d'aquesta estimació és important per evitar la pèrdua de deadlines de tasques HRT, però també per incrementar la QoS en sistemes SRT. A més, també pot ajudar a millorar la planificabilitat del sistema i reduir el consum d'energia. Es proposa el model *Processor-Memory (Proc-Mem)*, que dinàmicament prediu el temps d'execució d'aplicacions de temps real per a cada nivell de freqüència. Aquest model mesura en el primer hiperperíode, fent ús de *Performance Monitoring Counters (PMCs)* en temps d'execució, la porció de temps que cada nucli està realitzant còmputos (*CPU*), esperant a memòria (*MEM*), o ambdós (*OVERLAP*). Aquesta informació serà usada per estimar el temps d'execució a qualsevol altra freqüència.