

Applications and Programming Models for Scalable Many-Cores Architectures

The context of this work is the increasing irrelevance of Moore's Law, which observed that the number of transistors that could be put on a chip at the same price doubled every 18 to 24 months. For more to fit they had to get smaller, which let them run faster, albeit hotter, so performance rose over the years — but so did expectations. Today, those expectations remain, but processor performance has plateaued. This leads to the search of new architectures with higher energy efficient that just faster single processors.

The goal is to develop software concepts related to scalable many core heterogeneous systems exploiting in-package wireless communication and in-memory computing accelerators in both high-performance and deeply embedded domain. This is in line with the goal of the WiPlash project i.e. develop heterogeneous many-core systems and applications exploiting the emerging topic of wireless plasticity to improve performance, energy efficiency and scalability of heterogeneous accelerators based on in-memory computing processors.

The main topics that will be studied in this work are:

1. Heterogeneous many-core architectures
2. Hardware-software optimization
3. Programming Models for scalable architectures
4. Optimization and benchmarking of applications

Applicazioni e Modelli di Programmazione per Architetture di Calcolo Scalabili

Il contesto di questo lavoro è la crescente irrilevanza della Legge di Moore, che era basata sull'osservazione empirica che il numero di transistor che potrebbero essere messi su un chip allo stesso prezzo veniva raddoppiato ogni 18-24 mesi. Cio' implica che i transistori CMOS devono essere sempre più piccoli, il che permette loro di funzionare più velocemente, anche se a temperature operative più alte. In questo modo così le prestazioni aumentarono nel corso degli anni, ma anche le aspettative. Oggi, queste aspettative rimangono, ma le prestazioni dei processori si sono stabilizzate in termini di frequenza operativa, a causa dell'intollerabile aumento della densità di potenza dissipata. Questo porta alla ricerca di architetture con una maggiore efficienza energetica piuttosto che focalizzarsi esclusivamente su processori singoli più veloci.

L'obiettivo è lo sviluppo di concetti software legati a sistemi scalabili a bassa potenza sia nel dominio ad alte prestazioni che in quello profondamente integrato, sfruttando concetti emergenti come in-memory computing e canali di comunicazione wireless. Questo è in linea con l'obiettivo del progetto WiPlash, ovvero sviluppare acceleratori paralleli con efficienza energetica e scalabilità migliorate sfruttando anche concetti emergenti di plasticità wireless per la comunicazione on-chip ed off-chip e accelerazione in-memory computing.

Gli argomenti principali che saranno studiati in questo lavoro sono:

1. Architetture multi-core eterogenee
2. Ottimizzazione hardware-software
3. Modelli di programmazione per architetture many core
3. Ottimizzazione e valutazione di applicazioni