**Learning models and optimization algorithms for static variable precision tuning in the transprecision computing area**

The main purpose is to design solutions for the problem of static Floating-Point (FP) variables precision tuning, with a particular emphasis on reduced energy consumption but tolerable, precise computation errors. This topic is in line with the goal of the H2020 OPRECOMP project, i.e. to develop transprecision hardware and software architectures with improved power efficiency and scalability exploiting reduced precision computations ranging from low-power to high performance computing.

The end goal for this activity is the analysis and study of Artificial Intelligence-inspired methods for the optimal precision assignment for FP variables in scientific applications. Generally speaking, static FP precision tuning is performed by ad-hoc algorithms based on generate-and-test policies, which (albeit capable of obtaining good quality solutions) tend to be extremely slow and not flexible. Furthermore, understanding the impact of different data set fed as input to the scientific benchmarks is an open question still unexplored by state-of-the-art methods.

This work will be focused on the development of Machine Learning models to predict the computation error associated with different configurations of variable precision, and capable of estimating the degree of confidence about the prediction made. Moreover, such ML models will have to take into account the wide range of potentially different data sets taken as input by the target benchmarks. In addition, this work will aim at integrating such ML models within a optimization framework for the optimal tuning of the FP variables, striving to reduce the energy consumption associated to the computation while maintaining sufficient precision. The evaluation of the proposed methods will take place on an architectural simulators, using real application that have been adapted to this context.

**Apprendimento di modelli e algoritmi per la selezione statica della precisione di variabili nell'ambito del paradigma computazionale a transprecisione.**

L'obiettivo principale è lo sviluppo di soluzioni per il problema della selezione statica della precizione di variabili Floating-Point (FP), variabili a virgola mobile, con una particolare enfasi sulla riduzione dei consumi energetici a fronte del rispetto di vincoli sull'errore associato alla computazione. Questo argomento è in linea con gli obiettivi del progetto Oprecomp H2020, ovvero sviluppare piattaforme computazionali sia basso consumo che ad elevate prestazioni che sfruttano una paradigma computazionale a precisione variabile.

Lo scopo finale di questa attività consiste nell'analisi e nello studio di tecniche ispirate all'intelligenza artificiale per l'assegnamento ottimale della precisione di variabili FP all'interno di applicazioni scientifiche. In generale, l'assegnanto statico delle precisioni di variabili FP viene effettuato tramite algoritmi ad-hoc basate su politiche in cui diverse configurazioni vengono iterativamente esplorate, un approccio che (benchè in grado di ottenere soluzioni di buona qualità) tende ad essere molto lento e poco flessibile. Oltretutto, comprendere l'impatto che differenti data set usati come input per le applicazioni scientifiche è ancora una questione aperta che non è stata esplorata in letteratura.

Questo lavoro si concentrerà sullo sviluppo di modelli di apprendimmento automatico (Machine Learning, ML) per predirre l'errore di computazione associato a diverse configurazioni di precisione , modelli anche in grado di stimare il livello di incertezza associato alle predizioni stesse. Inoltre, questi modelli predittivi dovranno anche prendere in considerazione l'ampio spettro di possibili diversi dati presi in ingresso dalle applicazioni scientifiche selezionate come casi di studio. In aggiunta, questo lavoro mira a integrare questi modelli predittivi all'interno di un sistema di ottimizzazione per la selezione delle precisioni ottimali per le variabili FP, con lo scopo di diminuire i consumi energetici associati ad operazioni di computazione, mantenendo al tempo stesso sufficiente accuratezza. La valutazione dei metodi proposti avverà su un simulatore architetturale utilizzando applicazioni reali adattate a tale scopo.