

LINEA DI RICERCA Computational Science e Calcolo ad Alte Prestazioni (High Performance Computing – HPC)

Giuliano Laccetti, Luisa D'Amore, Marco Lapegna

La **Computational Science and Engineering** (CSE) è un campo multidisciplinare di ricerca, che si avvale di esperienze e competenze di matematica numerica, computer science e delle discipline delle scienze e dell'ingegneria, rientrando quindi a pieno titolo nell'ambito del tema **Calcolo Scientifico**. Suo principale scopo è quello di sviluppare e utilizzare metodi computazionali per l'avanzamento dell'innovazione scientifica e tecnologica. La CSE è attualmente uno dei maggiori campi di intervento per la ricerca e l'alta formazione universitaria ed è basata, come pilastro fondamentale, **sull'High Performance Computing** (HPC) rivestendo un ruolo centrale nella "data revolution". In questo ambito si colloca l'attività di un gruppo di ricerca del DMA.

Più precisamente, il gruppo di ricerca del DMA (Prof. Luisa D'Amore, Prof. Giuliano Laccetti, Prof. Marco Lapegna), insieme con un gruppo di assegnisti e studenti di dottorato, svolge attività di ricerca nel contesto scientifico della modellistica e simulazione computazionale di sistemi complessi (ovvero multi-scala e/o multi-physics e/o di grandi dimensioni) e della risoluzione di tali problemi su **sistemi di calcolo ad alte prestazioni**, anche in collaborazione con Università e Centri di Ricerca di rilevanza internazionale; in particolare in questa proposta progettuale l'attenzione sarà rivolta verso la Simulazione della circolazione oceanica su scala locale in mari chiusi e semi-chiusi (Mar Caspio e Mar Mediterraneo). Attività in collaborazione con il Centro Euro Mediterraneo sui Cambiamenti Climatici, (IT), Imperial College London (UK), University of California (USA).

L'innovazione risiede nella metodologia introdotta e/o nel progetto dell'algoritmo e/o nelle scelte implementative adottate oramai indispensabili ai fini della realizzazione di strumenti computazionali affidabili e in grado di sfruttare appieno l'enorme potenza di calcolo che sarà disponibile nei prossimi anni (Exascale Computing, linee guida della Ricerca e Innovazione, Horizon 2020).

Si prevede la sperimentazione e validazione delle metodologie numeriche di Data Assimilation (DA), nella sua formulazione Variazionale tri-dimensionale e quattro dimensionale nei modelli di simulazione della circolazione oceanica. Se da una lato la DA consente di migliorare l'affidabilità dei modelli stessi dall'altro aumenta in maniera significativa la complessità dell'intera simulazione computazionale, pertanto nasce l'esigenza di introdurre paradigmi di parallelismo multilivello nei relativi algoritmi risolutivi.

Ai fini della loro implementazione efficiente su architetture emergenti, gli algoritmi sviluppati dovranno tenere conto dei differenti "livelli architetturali" dei sistemi di calcolo su cui verranno implementati e sperimentati. Si pensi ad esempio ad architetture multiprocessore-multicore, più nodi connessi tra loro, ciascun nodo composto da più cores (computing cores, acceleratori grafici, ecc ..). Ciò porta alla progettazione di algoritmi (e software) di tipo "ibrido", che sappiano sfruttare le diverse potenzialità e caratteristiche dei "diversi" livelli.

Grande importanza inoltre riveste il problema della gestione di grandi quantità di dati (Big Data) in ambito High Performance Computing. Un tale approccio alla implementazione e sperimentazione si avvarrà della collaborazione con l'Argonne National Laboratory del Department of Energy degli USA, e con la University of Chicago.

Infine, nell'ambito della realizzazione di software scientifico del tipo finora descritto, e del suo utilizzo su risorse distribuite, si deve tenere conto di diversi fattori. Qui se ne citano, a titolo di esempio, due. Il primo, quello della cosiddetta "fault-tolerance", cioè la possibilità che, specie in ambienti distribuiti, il software porti comunque a termine il suo compito, anche in presenza di "incidenti" (malfunzionamento di nodi di calcolo, parziali black out elettrici, .. ecc...): di qui la necessità di studiare e realizzare strategie e tools per il cosiddetto checkpointing e/o

migrazione. Il secondo, più proprio di un ambiente di calcolo che caratteristica degli algoritmi e del software, riguarda la gestione efficace ed efficiente delle risorse, rendendole disponibili in maniera "adattativa" alle varie applicazioni. Tali studi e realizzazioni avvengono in collaborazione con il CNR e l'INFN.

Risultati e Prodotti Attesi

Nel campo delle scienze computazionali e dell'HPC il principale risultato atteso – a lungo termine - è la realizzazione di prodotti software effettivamente utilizzabile nelle simulazioni scientifiche, in ambienti di calcolo ad alte prestazioni; un ulteriore contributo sarà alla utilizzazione efficiente di ambienti di calcolo distribuiti (si pensi al caso delle grid e più recentemente, delle cloud). Tuttavia, poiché nessun singolo ricercatore ha le competenze necessarie per far fronte a tutto il dominio di esperienza (dall'analisi dei modelli, alla progettazione di nuovi algoritmi, fino alla loro implementazione in modo efficiente sulle architetture di calcolo, e all'analisi dei risultati, ...), ciò sarà possibile anche con la partecipazione a convegni, l'organizzazione di eventi e seminari, e tenendo anche conto della collaborazione scientifica già in atto con Università ed Enti di Ricerca nazionali e internazionali.