

Dall'High Performance Computing all'Exascale:

le sfide della ricerca e le nuove applicazioni

Milano, 11 ottobre 2018 • DEIB, Aula Alpha • Edificio 24 • via Golgi 40 • ore 10:00 - 16:30

PRESENTAZIONE

L'HPC (High-Performance Computing) è una risorsa strategica per il nostro futuro in quanto consente ai ricercatori di studiare e simulare fenomeni complessi consentendo ai policy maker di prendere decisioni migliori e all'industria di cogliere nuove opportunità per l'innovazione di prodotti e servizi.

L'HPC oggi utilizza reti di milioni di processori ad alte prestazioni per affrontare alcuni dei problemi più complessi al mondo. Tutte le discipline scientifiche oggi stanno diventando computational intensive. La ricerca scientifica richiede una potenza di calcolo e una capacità di gestire enormi volumi di dati. L'industria e le PMI fanno sempre più affidamento sulla potenza dei supercalcolatori per ridurre i costi ed il time to market di prodotti e servizi.

I problemi computazionali che stiamo affrontando oggi guidano scoperte e innovazioni fondamentali per la qualità della nostra vita; la necessità di una potenza computazionale sempre maggiore cresce con la complessità delle domande che la scienza si pone. Le esigenze sociali, scientifiche ed economiche sono i driver per la prossima generazione di sistemi HPC con prestazioni Exascale (computer in grado di eseguire 10¹⁸ cioè miliardi di miliardi di operazioni in virgola mobile al secondo). L'Exascale computing è considerato un risultato fondamentale per le architetture dei calcolatori, poiché si ritiene che l'Exascale sia l'ordine di grandezza della potenza di elaborazione del cervello umano a livello neurale.

L'Europa e molti paesi (Stati Uniti, Giappone, Europa, Cina e Brasile) hanno annunciato piani ambiziosi per la costruzione della prossima generazione di sistemi HPC con prestazioni Exascale e l'implementazione di supercomputer all'avanguardia.

Tuttavia, molte sono le sfide ancora aperte associate all'Exascale computing. Si suppone che un supercomputer Exascale comprenderà oltre cinque milioni di processori interconnessi e richiederà circa 20MW di potenza (pari alla potenza richiesta da una città di medie dimensioni), con sufficiente larghezza di banda di memoria per alimentare il flusso di calcolo ed elevata capacità di storage. Questi requisiti a livello di sistema implicano che ciascun nodo possa offrire oltre 10 Teraflop con meno di 200W. Oggi esiste un gap pari a 7 tra i flop per watt del supercomputer più efficiente dal punto di vista energetico al mondo e l'obiettivo dell'Exascale. Una macchina Exascale sfrutterà i progressi di tecnologie chiave, tra cui sistemi eterogenei basati su CPU a bassa potenza e alte prestazioni e GPU sempre più efficienti abbinati a memorie 3D a banda larga, uso di tecnologie chiplet, reti e sistemi di storage avanzati e nuovi framework di programmazione parallela.

Obiettivo dei seminari di oggi è quello di presentare una panoramica delle principali sfide e delle soluzioni attuali per costruire i sistemi di calcolo Exascale di prossima generazione coniugando gli aspetti di ricerca scientifica con aspetti industriali e illustrare alcune nuove applicazioni che sfrutteranno tali capacità di calcolo.



MODALITÀ DI PARTECIPAZIONE

La partecipazione al Seminario è gratuita. L'iscrizione può essere effettuata accreditandosi sul sito www.aeit.it alla Sezione Manifestazioni oppure via mail all'indirizzo aict@aeit.it

SEDE

DEIB Polimi • Aula ALPHA • Edificio 24
Via Golgi 42 - Milano

SEGRETARIA ORGANIZZATIVA

E-mail: fontana@aeit.it

ORGANIZZATO DA



PROGRAMMA

9.30-10.00 Registrazione

10.00-10.15 Welcome

Ing. Andrea Penza • *Presidente AICT*

Prof. Danilo Ardagna • *Politecnico di Milano*

10.15-10.45 • Il progetto ANTAREX: un approccio adattativo per sistemi HPC applicato alla scoperta di nuovi farmaci e alla navigazione nelle smart cities

Prof. Cristina Silvano • *Politecnico di Milano*

Il progetto ANTAREX colloca il Politecnico di Milano nell'avanguardia mondiale della ricerca nel settore dell'efficienza energetica per sistemi di calcolo ad elevate prestazioni (HPC). In particolare ANTAREX propone un approccio fortemente innovativo per supportare l'adattività dinamica delle applicazioni per sistemi di supercalcolo eterogenei ad elevata efficienza energetica in vista di raggiungere il livello di prestazioni previsto dall'era Exascale, ovvero un miliardo di miliardi di operazioni al secondo. È in corso la fase di sperimentazione della ricerca su due casi di studio industriali di alta rilevanza scientifica e sociale. Il primo caso consiste nell'ottimizzazione energetica di un'applicazione biofarmaceutica per accelerare la ricerca di nuovi farmaci sfruttando le potenzialità di calcolo offerte dal CINECA. Il secondo caso di studio consiste nello sviluppo di un sistema di navigazione adattativo da usare nel contesto di "smart cities" sfruttando le interazioni tra il lato client e il lato server ad alte prestazioni offerto dal sistema di supercalcolo installato presso IT4Innovations in Repubblica Ceca. Oltre a dieci ricercatori del Politecnico di Milano, il progetto ANTAREX coinvolge due centri di supercalcolo: il CINECA, Consorzio interuniversitario italiano di supercalcolo e IT4I, Centro nazionale di supercalcolo della Repubblica Ceca. Il progetto coinvolge inoltre l'ETH di Zurigo, l'Universidade do Porto e l'Institut National de Recherche en Informatique et en Automatique di Rennes. I partner industriali comprendono una delle principali aziende biofarmaceutiche italiane, Dompé, focalizzata sullo sviluppo di soluzioni terapeutiche innovative per malattie rare spesso orfane di cura, e Sygic, la principale azienda europea di software di navigazione.

10.45-11.15 • HPC and more in Microsoft Azure

Dr.ssa Anna Debora Maria Schampers • *Microsoft Italia*

Il bisogno di sistemi HPC scalabili sta crescendo sensibilmente sotto la spinta dell'aumentata quantità di dati generati e delle conseguenti informazioni da trattare. I cluster HPC tradizionali, progettati per gestire l'elaborazione di big data, presentano carenze intrinseche, che possono essere superate migrando verso un ambiente più flessibile di cloud computing. Il Cloud Computing è una realtà sempre più diffusa e discussa, anche se il suo vero potenziale non è ancora stato pienamente esplorato e utilizzato. La maggior parte delle applicazioni HPC richiede risorse ed ambienti dedicati, sempre disponibili ed altamente personalizzati per il tipo di calcolo attuato, poiché fanno un uso intensivo delle risorse a disposizione. Queste esigenze erano tradizionalmente trattate da cluster e supercomputer, che sono spesso difficili da configurare, gestire o operare. Il seminario illustrerà come la flessibilità "on demand" del cloud si presti molto bene per carichi di lavoro ad alte prestazioni e può portare anche ad implementazioni in modalità Ibrida e discuterà di come con Microsoft Azure sia possibile fruire di un accesso semplice e rapido a molti altri servizi di tipo Platform-as-a-Service, come Machine learning, Cognitive Services, Intelligenza Artificiale, Deep learning, HPC Batch Services etc., che possono portare anche all'interazione di due aree di interesse diverse come HPC e AI.

11.15-11.30 • BREAK

11.30-12.00 • COMPSs: un framework di alto livello per lo sviluppo di applicazioni parallele

Dr. Daniele Lezzi • *Barcelona Supercomputing Center, Spagna*

L'uso di workflow in ambito scientifico ha rappresentato un metodo classico per descrivere ed implementare esperimenti scientifici, tipicamente studi parametrici ed esplorazioni di soluzioni ottimali. L'applicazione di tali tecniche richiede dei processi molto diversi tra loro a seconda dell'area di ricerca stessa in quanto l'applicazione finale deve rispecchiare esattamente il metodo sperimentale del ricercatore. I dati sono il fulcro dello sviluppo di workflow scientifici, dati che possono essere generati dalle applicazioni o processati da esse ma che comunque, considerata la sempre crescente dimensione dei dati stessi, richiedono soluzioni avanzate per l'analisi in-situ. Tali soluzioni comportano la necessità di sviluppare applicazioni di tipo workflow che includano esecuzioni di calcolo puro affiancate a calcoli di tipo data-analytics. La imminente introduzione di architetture Exascale, comporta per la comunità scientifica nuove opportunità per l'implementazione di workflow "intelligenti" che includano simulazioni HPC tradizionali (che richiedono l'uso di nodi di calcolo di un supercomputer), big data analytics (intervallati o sovrapposti al calcolo) e operazioni di visualizzazione. L'applicazione risultante non sarà statica, ma le diverse parti saranno istanziate dinamicamente a seconda dei requisiti dello studio al centro della simulazione. Il gruppo di Wor-

kflows and Distributed Computing del Barcelona Supercomputing Center, ha realizzato negli ultimi anni COMPSs, un framework di programmazione che permette l'ottimizzazione di applicazioni su piattaforme distribuite (HPC, Cloud, etc). Gli ultimi sviluppi permettono il coordinamento di operazioni complesse con l'obiettivo finale di realizzare la convergenza tra HPC, High Throughput Computing e HPC Data Analytics.

12.00-12.30 • Mirar le stelle: Deep Learning per l'astrofisica delle alte energie

Dr. Andrea Bulgarelli • *Istituto Nazionale di Astrofisica*

Oggi la comunità scientifica sta affrontando un enorme aumento di volume, velocità e dimensionalità dei dati. La ragione di questa esplosione di dati è nel progresso della scienza stessa. Nuovi esperimenti richiedono livelli senza precedenti di accuratezza e precisione. Sono inoltre necessarie nuove simulazioni su vasta scala per esplorare il problema scientifico. In questo contesto INAF sta sperimentando le tecniche di Deep Learning nei progetti di astrofisica. In questo intervento verranno descritti alcuni problemi connessi all'astrofisica dei raggi gamma e verranno mostrate alcune soluzioni utilizzando tecniche di HPC e Deep Learning.

12.30-14.00 • BREAK

14.00-14.30 • Un sistema di monitoraggio e di controllo dell'efficienza energetica per sistemi HPC applicato al centro di supercalcolo del CINECA ed al supercalcolatore D.A.V.I.D.E.

Prof. Andrea Bartolini • *Università di Bologna, Italia*

Nella corsa verso l'exascale, i sistemi di calcolo ad alte prestazioni si trovano a dover affrontare importanti sfide tecnologiche che ne limitano l'efficienza. Tra tutte, con la fine della legge di Dennard l'avanzare del processo elettronico causa un aumento della densità di potenza, che negli anni ha reso i sistemi di calcolo limitati dal consumo di potenza. Inoltre, l'affidabilità e la sicurezza dei sistemi di calcolo, dei componenti HW e dei componenti SW pone nuove sfide nella gestione olistica del datacenter. D.A.V.I.D.E. [Development for an Added Value Infrastructure Designed in Europe], è un cluster di calcolo ad alte prestazioni innovativo progettato da E4 Computer Engineering per PRACE [Partnership for Advanced Computing in Europe]. D.A.V.I.D.E. oltre all'uso di componentistica efficiente e performante, introduce un innovativo strato software progettato per il monitoraggio e controllo della potenza e grana fine. In questa presentazione lo speaker presenterà queste estensioni sul D.A.V.I.D.E. dando spunti sul design dei sistemi di calcolo ad alte prestazioni efficienti energeticamente.

14.30-15.00 • IBM: Advanced Deep Learning su piattaforme HPC

Ing. Giovanni Mariani • *HPC Zurich division, IBM Svizzera*

Il deep learning è sempre computationally intensive e si rendono molto spesso sistemi e software HPC specializzati per fornire servizi di deep learning su larga scala. Per realizzare appieno il potenziale delle tecnologie di deep learning, IBM vuole rendere il deep learning più facilmente accessibile agli sviluppatori e ai data scientist in modo che si possano concentrare sulle problematiche dei dati automatizzando l'addestramento dei modelli di rete neurali su grandi data set in modo da creare modelli all'avanguardia. In questo intervento presenteremo nuove infrastrutture e tecnologie per l'implementazione di servizi di deep learning, tra cui tecniche avanzate di ottimizzazione automatica per il l'addestramento e il tuning dei modelli di deep-learning.

15.00-15.30 • Quantum Computing: dalla ricerca al mainstream in cinque anni

Dr. Federico Mattei • *IBM Quantum Computing Leader, IBM Italia*

Tra cinque anni, gli effetti del Quantum Computing andranno oltre i laboratori di ricerca. Sarà usato in modo estensivo da nuove categorie di professionisti e sviluppatori che guardano a questo metodo di calcolo emergente per risolvere problemi un tempo considerati irrisolvibili. Durante il seminario verrà illustrata la strategia di IBM per il Quantum Computing e le principali aree di ricerca in quest'area.

15.30-16.00 • Software-Defined WAN e il progetto Atmosphere

Prof. Giacomo Verticale • *Politecnico di Milano, Italia*

L'accesso affidabile, veloce, sicuro e ubiquo ai servizi in cloud è indispensabile per garantire che i vantaggi dell'Exascale si traducano in applicazioni e servizi per gli utenti finali. Il paradigma SD-WAN fornisce gli strumenti per la virtualizzazione e l'automazione di rete, massimizzando il traffico cloud-to-cloud ed eliminando i colli di bottiglia che compromettono le prestazioni. Le funzioni di rete sviluppate nel progetto Atmosphere consentiranno di gestire le reti verso i sistemi cloud e tra sistemi cloud con l'affidabilità necessaria per ogni tipo di servizio.

16.00-16.30 Conclusioni

Ing. Andrea Penza • *Presidente AICT;*

Prof. Danilo Ardagna • *Politecnico di Milano*