

SUPERCOMPUTER, PIÙ OTTANI NEL MOTORE DELLA RICERCA



Massimo Nannini*

Come una benzina con più ottani riesce ad evitare perdite di energia dovute a detonazioni involontarie durante la fase di compressione della miscela aria-benzina, anche i moderni supercomputer, grazie alle loro sempre crescenti capacità di calcolo, permettono agli scienziati di concentrare tutti i loro sforzi sull'ideazione di nuove teorie.

Come definisce in modo semplice e diretto il suo nome, un supercomputer non altro che un computer in grado di svolgere miliardi di operazioni in virgola mobile al secondo. Per indicare la velocità con cui operano si usano come unità di misura i "FLOPS" (Floating Point Operations Per Second) o suoi multipli come il TeraFLOPS (1012 FLOPS) e già questo rende l'idea degli ordini di grandezza con cui ci confrontiamo.

I supercomputer non sono un'invenzione moderna, ma esistono dagli anni '70, come superelaboratori dedicati al calcolo scientifico. Inizialmente realizzati con tecnologie ed architetture proprietarie e dedicate, si sono via via evoluti fino ad arrivare ad oggi, dove i più potenti si basano su centinaia di migliaia di unità di calcolo non troppo differenti, se prese singolarmente, da quelle di un computer di fascia alta ma, connesse fra di loro con reti ad alta velocità, risultando così milioni di volte più veloci di un singolo PC commerciale.

Per ottenere queste prestazioni si potrebbe pensare vengano utilizzati processori dedicati, ma niente di più sbagliato. Si impiegano i migliori processori sul mercato ma, si tratta pur sempre

di processori commerciali come per esempio CPU Power9 o Xenon e GPU Nvidia V100.

Un supercomputer non è però solamente un superelaboratore più veloce e più grande, ma una macchina in grado di sfruttare al massimo l'elaborazione parallela dei processi. È stata infatti l'introduzione del concetto di esecuzione parallela la novità introdotta dai primi supercomputer, poi successivamente adottata da tutti i computer commerciali.

Alla base dell'esecuzione parallela vi è l'idea di potere suddividere l'elaborazione da eseguire in più task che possono essere elaborati singolarmente e contemporaneamente su microprocessori diversi o su più core dello stesso processore dando così luogo al cosiddetto massive parallel processing.

Riepilogando un supercomputer moderno non è altro che una rete di elaboratori ad alte prestazioni in grado di scambiarsi informazioni ad altissima velocità e coordinati da un SO concorrente generalmente Linux.

Cluster e grid computer

In questo mondo fatto di migliaia di processori che lavorano insieme esistono due approcci architetture differenti. Il cluster, che definisce una architettura con un grande numero di processori "vicini" che comunicano attraverso connessioni superveloci per esempio Infiniband e il grid che vede il collegamento tra vari mainframe situati generalmente in università e/o centri di ricerca,

che non lavorano necessariamente tutti insieme, ma che all'occorrenza possono operare come se fossero uno stesso supercomputer tra loro interconnessi attraverso reti private ed Internet.

Sistemi operativi e software

Anche in questo ambito l'evoluzione ha portato a sacrificare parte delle prestazioni a favore della compatibilità, non si utilizzano più SO proprietari ma SO standard, basti pensare che la quasi totalità dei supercomputer più veloci al mondo utilizzano Linux. Come abbiamo già accennato in precedenza, sfruttare l'architettura parallela significa utilizzare paradigmi di programmazione concorrente che permettano di sfruttare al massimo le potenzialità di tale architettura e linguaggi che anche se "datati" generano codice oggetto più efficiente, quali il Fortran e C/C++.

La Top 500

Questa classifica include i più veloci supercomputer mondiali. Attualmente il primo in classifica è supercomputer Giapponese Fugaku che dispone di 7,630,848 Cores ed è in grado di svolgere 442,010 TFlop/s. La classifica è sempre in movimento e vede attualmente una forte presenza di supercomputer Statunitensi e Cinesi, ma anche Europei se pur con notevole distacco come è facilmente verificabile visitando il sito <https://www.top500.org>.

Risulta evidente che siamo di fronte ad una importantissima sfida mondiale a cui anche l'Europa vuole partecipare. Per questo motivo ha lanciato il progetto EuroHPC (European High Performance Computer) per coordinare gli sforzi tra i paesi membri e dotarsi di una infrastruttura di supercomputer in grado di competere con USA e Cina. EuroHPC ha selezionato 8 centri in 8 diversi Paesi, fra cui l'Italia che ospiterà a Bologna il primo supercomputer (Leonardo). Gli altri centri sono: Sofia (Bulgaria), Ostrava (Repubblica Ceca), Kajaani (Finlandia), Bissen (Lussemburgo), Minho (Portogallo), Maribor (Slovenia), Barcellona (Spagna).

Italia al quarto posto per capacità di calcolo

Uno studio condotto da Atos e Ambrosetti (Gli impatti dell'high performance computing e la centralità per il posizionamento dell'Emilia-Romagna come hub di riferimento in Europa per la

ricerca e lo sviluppo sui big data) ha valutato che oggi l'Italia è terza in Europa per potenza di supercalcolo installata e sesta nel mondo, con una quota del 3,2% della potenza globale. Quando saranno accesi i circuiti del supercomputer europeo Leonardo che verrà installato all'interno del Tecnopolo di Bologna e gestito dal consorzio interuniversitario CINECA alla fine di quest'anno, i rapporti di forza cambieranno: la potenza di supercalcolo italiana aumenterà di 4,2 volte rispetto a quella attuale, posizionando il nostro paese al quarto posto mondiale nella classifica delle nazioni per capacità computazionale.

Supercomputer acceleratori per la ricerca

I supercomputer sono dunque un "acceleratore" per i progetti di ricerca in quanto aiutano gli scienziati nello sviluppo della medicina, nella progettazione di nuovi farmaci e nuovi materiali ma anche nelle modellazioni di sistemi complessi per esempio lo studio delle evoluzioni climatiche o le previsioni metereologiche. Recentemente poi hanno svolto un ruolo essenziale nella mappatura del genoma del Covid-19 contribuendo così, come mai prima d'ora, i tempi per arrivare ad un vaccino efficace contro la pandemia. In tutti questi casi sono le enormi quantità di dati generati dai sistemi complessi a farla da padrone ed i supercomputer, grazie alle loro capacità di elaborazione, sono in grado di fare emergere le informazioni nascoste sotto una montagna di dati altrimenti impenetrabile.



***Massimo Nannini**, ingegnere elettronico e libero professionista: si occupa di consulenza informatica, project management e formazione di impresa.

Contatti

Email: info@gemaxconsulting.it
www.gemaxconsulting.it