

Supercomputing 2007: alla soglia del Petaflop

Claudio Arlandini

CILEA, Segrate

Abstract

Supercomputing è il principale congresso internazionale per quanto riguarda il calcolo, le reti, lo *storage* e l'analisi ad alte prestazioni. I visitatori sono immersi negli ultimi sviluppi di tecnologie, prodotti, ricerche, politiche e iniziative nazionali e internazionali. *Supercomputing* può essere definito l'unico posto al mondo dove si può vedere la tecnologia del domani mentre viene usata per risolvere i grandi problemi dell'oggi. CILEA ha partecipato attivamente all'edizione di quest'anno, presentando il Progetto SEPAC.

Supercomputing is the premier international conference on high performance computing, networking, storage, and analysis. Attendees become immersed in the latest state-of-the-field developments in technology, applications, vendor products, research results, national policy and national/ international initiatives. *Supercomputing* is the one place attendees can see tomorrow's technology being used to solve world-class challenge problems today. CILEA participated actively to the *Supercomputing 2007* conference presenting the SEPAC Project.

Keywords: Supercalcolo; Conferenze; SC07.



Fig. 1 – Il logo del congresso

Nei giorni dal 10 al 16 novembre scorsi si è svolta negli Stati Uniti, a Reno (NV), la ventesima edizione del congresso *Supercomputing* (SC07) [1] (Fig. 1). *Supercomputing* è il principale evento internazionale per quanto riguarda il calcolo, le reti, lo *storage* e l'analisi ad alte prestazioni. I visitatori sono immersi negli ultimi sviluppi di tecnologie, prodotti, ricerche, politiche e iniziative nazionali e internazionali. *Supercomputing* può essere definito l'unico posto al mondo dove si può vedere la tecnologia del domani mentre viene usata per risolvere i grandi problemi dell'oggi.

SC07 comprendeva al suo interno, oltre agli interventi orali divisi su numerose sessioni parallele, una interessante sessione *tutorial*, un ampio spazio espositivo, e una serie di eventi collaterali, il più noto dei quali è sicuramente l'annuncio della nuova classifica TOP500 [2].

Il Progetto SEPAC, di cui CILEA è uno dei membri fondatori, ha avuto uno spazio all'interno dello stand della Hewlett Packard.

In questo articolo verranno presentati alcuni spunti dagli interventi più interessanti, i dettagli della nuova classifica TOP500 e un aggiornamento sul Progetto SEPAC.

Una selezione degli interventi più significativi

“Viviamo in tempi interessanti.” Questo è stato il messaggio di Tim Mattson, Principal Engineer della Intel, durante un tutorial sulla programmazione avanzata OpenMP. Altri presentatori alla conferenza hanno detto sostanzialmente lo stesso: il succo è che lo sviluppo del software per l'HPC è sempre più complesso, e la situazione può solo peggiorare.

Secondo Mattson, “Questa situazione è derivata da un fallimento lato hardware, non da un successo.” Ovvero, dalla incapacità dell'hardware di scalare oltre un certo punto, che ha spinto gli ingegneri a ripensare le tradizionali architetture di von-Neumann. Questo ha condotto dapprima a computer paralleli e multi-core, e più recentemente a hardware speciale come l'uso delle GPU (Graphical Processor

Unit) come acceleratori. Il futuro delle CPU sarà in un insieme eterogeneo di processori con differenti capacità, dice Mattson. "Saranno considerazioni di consumi energetici che ci forzeranno in questa direzione."

Mattson ha comunque predetto che gli acceleratori, specialmente le GPU, saranno alla fine un vicolo cieco, perché le loro funzionalità verranno aggiunte a quelle delle CPU "general purpose". Comunque, questo non vuol dire che siano una cattiva idea sul breve-medio termine. Nel corso del congresso, decine di ricercatori hanno dimostrato speedup di 20-50x o più per specifici algoritmi sulle GPU. "Stiamo imparando moltissimo dagli esperimenti fatti con le GPU," ha detto, indicando che questa esperienza sarà preziosa per il futuro.

Anche qualcosa di semplice come l'I/O non sarà poi così immediato in questo futuro. Considerando le prestazioni di lettura da disco in un classico *benchmark* MPI, utilizzando le nuove funzioni di I/O parallelo di MPI-2 si sono ottenuti miglioramenti di ben 3 ordini di grandezza rispetto al classico I/O posix-style. Ottenere questi miglioramenti però non è stato automatico, il programmatore ha dovuto ripensare a come l'I/O era fatto, pensare cioè in termini di operazioni collettive compiute da un gruppo di processori cooperanti piuttosto che da un insieme di individui ognuno con il proprio compito.

Con l'avvento di calcolatori massicciamente paralleli, GPU, FPGA, NUMA e altre tecnologie, i programmatori si ritrovano ad affrontare una sfida crescente, e non è chiaro come questa sfida verrà affrontata.

Il Premio Nobel George Smoot ha entusiasmato i presenti alla sessione plenaria del mercoledì, parlando della nuova visione dell'universo resa possibile dai supercalcolatori negli ultimi anni. Esaminando infinitesime variazioni della radiazione cosmica di fondo, Smoot e collaboratori sono stati in grado di spiegare perché l'universo è granulare, con galassie raggruppate in ammassi e strutture di scala superiore. Queste fluttuazioni sono state scoperte per la prima volta dal satellite COBE nei primi anni '90, e osservazioni successive hanno aggiunto misurazioni di sempre maggior dettaglio. Il satellite Planck, a cui l'Italia fornisce un importante apporto alla realizzazione, sarà lanciato l'anno prossimo. Con questo tipo di dati gli scienziati possono lanciare enormi simulazioni per esplorare come queste superstrutture si sono formate e predire cosa succederà all'universo nel futuro.

Smoot ha condotto il pubblico in un tour virtuale dell'universo, soffermandosi in un'intersezione di diversi "filamenti" dove giacciono migliaia di galassie. Ci sono solamente centomila circa di queste intersezioni nel nostro universo, secondo Smoot. "Dovrebbe essere un posto piuttosto interessante in cui stare." (Fig. 2).



Fig. 2 – Un'immagine del tour cosmico in cui il Premio Nobel George Smoot ha condotto il pubblico.

Il Big Bang ha creato un effetto molto simile a quello di un sasso nello stagno. Un'onda d'urto continua a propagarsi da quel punto di origine. La superficie sferica dell'onda, la "superficie di ultimo scattering", è il limite di quello che noi possiamo osservare. Data a circa quattrocentomila anni dopo il Big Bang, quando i primi atomi sono stati in grado di formarsi. Per ragioni non ancora comprese, questa superficie non sta rallentando sotto l'influenza della gravità, come ci aspetteremmo, ma sta in effetti accelerando. Capire perché questo stia succedendo è uno dei grandi misteri del nostro tempo, e i supercalcolatori stanno aiutando gli scienziati come Smoot a svelarlo.

Le parole "cambio di paradigma" non danno giustizia al potenziale del *quantum computing*. Tutti sono d'accordo che stia arrivando e che modificherà alla radice l'industria attuale, se non il mondo intero. Il fondatore di D-Wave Geordie Rose afferma che il quantum computing è già tra noi, con una dimostrazione dell'ultimo prototipo della sua compagnia.

Il prototipo in realtà non era fisicamente alla conferenza, perché avrebbe avuto le dimensioni dell'intero stand della D-Wave, quindi le dimostrazioni avvenivano remotamente con il laboratorio, anche se si sta lavorando per rendere il sistema di refrigerazione più piccolo e

autocontenuto in vista della possibile commercializzazione del sistema.

L'ultima versione del chip della DWave ha 28 qubit (quantum bit), secondo Rose. Pensa di poter mostrare una macchina da 512 qubit l'anno prossimo e uno da 1024 nel 2009. Lo zoccolo può sostenere fino a un milione di qubit, ma la strada è ancora lunga. "Se non riusciamo ad arrivare a 512 qubit l'anno prossimo, siamo nei guai," ha ammesso.

Il calcolatore quantico della DWave lavora cercando lo stato a energia minore una volta fornito un set di input e limitazioni. È un po', dice Rose, come cercare la stazione in una radio FM. Quando si trova un segnale forte, si va avanti e indietro con la manopola fino a trovare il punto migliore di ascolto. Solo che qui ci sono migliaia di manopole da girare. Questo dà un'idea di quello che succede a un calcolatore quantico quando esegue un programma, solo che invece di diventare gradualmente più debole o più forte il segnale salta tra stati quantici discreti. Occasionalmente succede che il sistema fornisca una risposta errata, spiega Rose, così ogni programma è eseguito in realtà 8 volte e si prende come risposta finale quella fornita la maggioranza delle volte. È strano, ma sembra funzionare.

Sembra esserci un certo scetticismo nei confronti del lavoro di Rose, in quanto ha pubblicato pochissimo dei suoi risultati su riviste scientifiche di primo piano. L'ufficio brevetti degli Stati Uniti ha comunque garantito alla compagnia decine di brevetti sull'argomento, e altre decine sono in corso di emissione. La D-Wave ha sicuramente il maggior numero di brevetti al mondo nel settore. Questo è forse il vero piano per il futuro successo della compagnia. Anche se il sistema che hanno costruito alla fine non si dimostrerà essere pratico, la DWave potrebbe garantire un profitto ai suoi investitori dal proprio portfolio brevetti quando qualcuno costruirà veramente un calcolatore quantico commercialmente praticabile.

Le prospettive del quantum computing sembrano essere reali. Un dirigente HP mi ha confermato che la multinazionale ha cominciato a investire moltissimo nel settore.

La nuova classifica TOP500

Il nome di Big Blue continua a sveltare nella classifica ufficiale dei supercomputer più potenti del mondo (Fig. 3).



Fig. 3 – Il logo della classifica TOP500

Blue Gene/L, situato al Lawrence Livermore National Laboratory (Fig. 4), nella sua ultima espansione è in grado di raggiungere 478,2 TFlop/s: il passaggio da 130.072 a 212.992 processori ha permesso di raggiungere prestazioni decisamente superiori a quelle registrate sei mesi fa, quando Blue Gene/L aveva confermato la leadership con i suoi 280,6 TFlop/s. Analogamente la nuova installazione di Blue Gene/P (la cui nascita era stata annunciata lo scorso giugno), è operativa presso il consorzio di ricerca di Julich in Germania: con 167 TFlop/s è il numero due al mondo e il più veloce in Europa.



Fig. 4 – Blue Gene/L, il calcolatore più potente al mondo secondo la più recente classifica TOP500

La novità della classifica consiste nel fatto che attualmente solo una delle prime cinque posizioni è occupata da uno dei sistemi del Dipartimento dell'Energia USA: Blue Gene/L a parte, gli altri sistemi fanno capo ad altre organizzazioni, mentre sei mesi fa i primi tre supercomputer erano legati al Dipartimento. Infatti, in terza posizione troviamo l'Altix ICE 8200 di SGI (126,9 TFlop/s), impiegato presso il Computing Applications Center nel Nuovo Messico; al quarto posto abbiamo poi la *new entry* dell'India, che per la prima volta compare addirittura nella top-ten, grazie al sistema Cluster Platform 3000 BL460c (117,9 TFlop/s)

fornito da Hewlett-Packard ai Tata Computational Research Laboratories, nella città di Pune; la quinta posizione è ancora merito del BL460c di HP (102,8 TFlop/s), questa volta installato in Europa, presso un'agenzia non meglio precisata del governo svedese.

Il Dipartimento dell'Energia USA ha promesso rivincita per il 2008, quando entreranno in operazione i primi due calcolatori a superare il traguardo del Petaflop/s di potenza sostenuta, ovvero il cluster fornito da Cray all'Oak Ridge National Laboratory e "Roadrunner", il sistema ibrido di processori AMD e Cell che IBM installerà presso i Los Alamos National Laboratory.

Il tasso di cambiamento nella lista è stato ben superiore al solito, evidenziando l'ottimo stato di salute del settore. Il livello di ingresso è balzato in sei mesi da 4.0 TFlop/s a 5.9 TFlop/s, tanto che l'ultimo sistema della lista attuale sarebbe stato alla posizione 255 solo sei mesi prima. Il numero di sistemi etichettati come cluster ha superato quota 80%. Intel rafforza il suo ruolo dominante come produttore di processori per i sistemi ad alte prestazioni, con il 71%. AMD sconta i ritardi del suo processore quad-core, ma rimane al secondo posto, grazie al crollo del numero di sistemi con processori IBM Power. HP e IBM dominano oltre il 75% del mercato, con una leggera predominanza del secondo.

Mentre l'Europa si scalda, con due siti nei top-ten e ben cinque siti che si sono dichiarati disposti a ospitare sistemi di classe Petaflop nei prossimi anni, nell'ambito dell'iniziativa PRACE [3], rimane non entusiasmante la situazione italiana, con soli sei sistemi, di cui cinque di proprietà del CINECA, o comunque ospitati presso il medesimo centro, e uno del nuovo centro di calcolo dell'ENEA.

Il Progetto SEPAC



Fig. 5 – Il logo del Progetto SEPAC.

All'interno dello spazio espositivo di Hewlett Packard è stato generosamente concesso un angolo per la presentazione dell'iniziativa SEPAC (Fig. 5) (Southern European Partnership for Advanced Computing [4]). Si tratta dell'evoluzione del Progetto TAPAC [5], una cooperazione sovranazionale tra centri di calcolo del Sud Europa basata su tecnologie di griglia. La missione dell'iniziativa è quella di fornire alla comunità di ricercatori un ambiente che sia stabile, affidabile e facile da usare per il lancio di simulazioni in un ambiente di calcolo distribuito.

Oltre al CILEA, socio fondatore, ne fanno attualmente parte il Consorzio SPACI, il Politecnico (ETH) di Zurigo, il Centro di supercalcolo svizzero (CSCS) di Lugano, e l'Università di Zurigo, sotto l'egida di Hewlett Packard. Nel corso del 2007 due nuovi partner sono entrati nel progetto, allargandone gli scopi e la base geografica, il Centro di supercalcolo di Catalogna (CESCA) di Barcellona e il Centro di Ricerca Nucleare del Negev (NRCN) di Be'er Sheva, Israele.

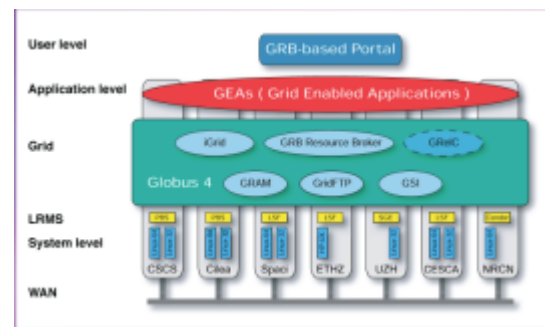


Fig. 6 – L'architettura del Progetto SEPAC

L'infrastruttura (Fig. 6) che lega i centri di calcolo in questione è il middleware Globus v. 4, su cui poggiano una serie di servizi, sia appartenenti al Globus Toolkit, sia costruiti ad-hoc, come il sistema informativo iGrid, sviluppato dal gruppo del consorzio SPACI, che è un'alternativa più potente e veloce del comune pacchetto MDS. Gli utenti trovano a loro disposizione una serie di pacchetti applicativi *grid-enabled*, il cui numero è in costante aumento, che vengono acceduti tramite il portale GRB (Grid Resource Broker), anch'esso sviluppato dal gruppo SPACI, che si occupa anche dello *scheduling* delle simulazioni interagendo con i sistemi di code e di *load balancing* locali. Il punto di forza di questa architettura sta proprio nella eterogeneità delle infrastrutture sottostanti, che vengono com-

pletamente mascherate all'utente lasciandolo libero di fare scienza.

Conclusioni

Il Petaflop è ormai una certezza, ma il futuro è costellato di punti interrogativi, avendo l'impressione di essere giunti al termine di un lungo percorso tecnologico e non avere un chiaro riferimento di fronte a sé. Già si parla di Exaflop in termini realistici, anche se nessuno può dire come sarà il processore del primo calcolatore di questa classe. Quello che è chiaro è che con le tecnologie dell'oggi avrebbe bisogno di una centrale nucleare dedicata. E il calcolatore quantistico ha ancora una lunga strada da percorrere, se pur se ne potranno mai ricavare calcolatori "general purpose", come ora li intendiamo. Nuove tecnologie si affacciano alla ribalta, come gli acceleratori, ma tutte richiedono un completo ripensamento del modo di costruire il software per il calcolo ad alte prestazioni, e questa è la parte più difficile, per di più con la costante preoccupazione di aver imboccato un vicolo cieco.

Davvero, "Viviamo in tempi interessanti".

Bibliografia

- [1] SC07, URL:<http://sc07.supercomputing.org>
- [2] TOP500 Computing Sites, URL:<http://www.top500.org>
- [2] Iniziativa PRACE, URL:<http://www.prace-project.eu>
- [4] Progetto SEPAC, URL: <http://www.sepac-grid.org>
- [5] C. Arlandini, "Il progetto Trans-Alpine Partnership for Advanced Computing", *Bollettino del CILEA*, n. 89, ottobre 2003.