

GOLGI: un cluster Opteron per il CILEA

Claudio Arlandini

CILEA, Segrate

Abstract

Il CILEA potenzia il suo parco macchine per il calcolo parallelo a alte prestazioni con un cluster di 55 nodi AMD Opteron, fornito dalla Exadron, Divisione High Performance Computing di Eurotech Spa. Con questa acquisizione il CILEA diversifica la sua offerta nel campo delle macchine a memoria distribuita.

CILEA is strengthening its role in High Performance Computing with a new 55 AMD Opteron nodes cluster, provided by Exadron, High Performance Computing Division of Eurotech Spa. The acquisition is meant to diversify its distributed memory machines offer to customers.

Keywords: Hardware, Supercalcolo, Calcolo Parallelo.

E' stato installato al CILEA il nuovo supercalcolatore che va a potenziare l'offerta a disposizione degli utenti nel settore del calcolo a alte prestazioni. Questo articolo vuole essere una scheda tecnica del nuovo server.

Il Consorzio ha deciso di diversificare maggiormente la propria offerta, che ricordiamo comprendere un cluster di nodi con processori Intel Xeon a 32 bit [1], un cluster di macchine HP con processori a 64 bit [2] e una macchina vettoriale NEC [3], con un cluster di nodi dotati di processori a 64 bit. L'interesse del mondo del calcolo a alte prestazioni per i cluster è sempre più confermato dalla classifica TOP500 [4], dove ormai hanno raggiunto circa il 70% del totale. Circa metà dei sistemi elencati è dotata di processori a 64 bit, equamente divisi tra i due principali costruttori, AMD e Intel.

Sebbene per molte applicazioni calcolatori massicciamente paralleli a memoria condivisa o vettoriali rimangano insostituibili, è pur vero che il rapporto prezzo/prestazioni di un cluster di componenti COTS (Commodities Off-the-Shelf) è molto invitante.

Il CILEA, come è nella sua tradizione, offre ai suoi utenti il meglio che il mercato propone al momento, sia nel campo dei processori, che per quanto riguarda le interconnessioni di rete. Particolare attenzione è stata data alla facilità di uso. Il nuovo cluster, pur con le caratteristiche di un sistema a se stante, non richiede un accesso distinto, ma è utilizzabile,

tramite il sistema di code, dall'usuale punto di accesso, dal cluster "avogadro".

Il server è stato dedicato a un esimio scienziato italiano, come ormai è tradizione CILEA.



Fig. 1 - Bartolomeo Camillo Emilio Golgi (1843-1926)

L'onore è toccato a Camillo Golgi (1843-1926), premio Nobel per la medicina (Fig. 1). Nato a Córteno (allora provincia di Bergamo), si laureò in Medicina con una tesi "Sull'eziologia delle malattie mentali", discussa con Cesare Lombroso, e entrò nel laboratorio istologico fondato da Paolo Mantegazza e diretto da Giulio

Bizzozzero, che fu suo primo maestro di ricerca. Per l'urgenza di trovare un lavoro sicuro, dovette però trasferirsi a Abbiategrasso, dove aveva vinto il concorso per diventare primario del Pio Luogo degli Incurabili. Qui, nella cucina del suo modesto alloggio, allestì un laboratorio di istologia in cui, nel 1873, ideò la rivoluzionaria "Reazione Nera" (Metodo di Golgi). Questo metodo permetteva di colorare, per consentirne finalmente la nitida visione, le cellule nervose e la loro struttura organizzata.

La sua scoperta fu conosciuta e apprezzata nella dovuta misura solo molti anni più tardi, soprattutto per merito del suo principale mentore, il patriarca della biologia ottocentesca Albert Von Koelliker. Nel 1889, i suoi studi sulla cellula lo portarono alla scoperta dell'apparato o Complesso di Golgi, uno dei componenti fondamentali della cellula, cinquanta anni prima dell'invenzione del microscopio elettronico.

Fece anche importanti scoperte nel campo della malariologia, formulando la "Legge di Golgi". Nel 1885 dimostrò l'associazione tra periodicità delle febbri malariche e ciclo del plasmodio che ne era responsabile e che i due diversi tipi di febbre malarica, la terzana e la quartana, erano provocati da due specie di plasmodio diverse: *P. vivax* responsabile della terzana benigna e *P. malariae* responsabile della quartana. Nel 1889 dimostrò che gli attacchi febbrili si verificavano nel momento in cui gli sporozoi (stadio del ciclo del plasmodio) rompevano i globuli rossi e si liberavano nel circolo sanguigno.

Scoprì anche la terminazione nervosa dei tendini, dette corpuscoli del Golgi.

Trasferitosi a Pavia, città natale del padre, ottenne le cattedre ordinarie di Istologia e Patologia generale, indi fu nominato rettore dell'Università, incarico che ricoprì a più riprese. Si dedicò anche alla politica, o meglio all'amministrazione pubblica, ricoprendo la carica di assessore all'igiene in Comune. Fu anche lungamente membro, e poi presidente, del Consiglio Superiore di Sanità. Propose la costruzione del nuovo Policlinico San Matteo e lottò strenuamente perché l'ateneo pavese mantenesse e accrescesse il suo secolare prestigio.

Fu insignito del Premio Nobel per la Medicina nel 1906 ex aequo con Santiago Ramón y Cajal, per gli studi sulla istologia del sistema nervoso.

Scheda tecnica

Il Comitato Tecnico del CILEA ha scelto come fornitore del sistema, dopo apposita gara, la Exadron, Divisione High Performance Computing di Eurotech SpA, con sede a Amaro (UD) [5].



Fig. 2 - Una foto del cluster golgi.cilea.it durante l'installazione. Un armadio non è rappresentato.

Il cluster (Fig. 2) si compone di 18 nodi quadriprocessori, con processori AMD Opteron 848 2.2 GHz single core e 37 nodi biprocessori AMD Opteron 275 2.2 GHz dual core, di cui uno dedicato a compiti speciali. Tra le caratteristiche tecniche del processore ricordiamo 1 MB di cache per core e tecnologia Hyper-Transport (HT). Ogni processore è dotato di tre canali coerenti HT che permettono una banda di picco di 24 GB/s per le comunicazioni tra processori o con periferiche che supportino il sistema. L'introduzione della tecnologia multicore presenta ovvi vantaggi, pur a scapito di una piccola perdita di prestazioni. Per questo si è deciso di costruire un sistema misto, allo scopo di venire incontro a ogni esigenza.

I nodi sono dotati di disco interno, o in alternativa di Flash Memory, utilizzato solamente per lo swap, in quanto il sistema è installato in configurazione diskless.

L'esperienza ha dimostrato che questa configurazione è più robusta di quella tradizionale. I nodi hanno una memoria RAM di 2 GB per core, aumentati a 4 GB in alcuni nodi per esigenze speciali, e hanno due interfacce di rete: una Gigabit Ethernet per l'amministrazione e l'accesso allo storage condiviso, e una Infiniband 4X [6] per il message passing, con una banda passante di picco di 10 Gbit/s.

Il sistema è contenuto in tre armadi 42U. È stato predisposto un sistema di storage collegato via Fibre Channel con uno dei nodi, capace di 2.6 TB.

Il sistema operativo adottato è Linux CentOS 4.2, con kernel 2.6.14.

Il cluster è amministrato e gestito in maniera diskless tramite un sistema sviluppato in collaborazione con Exadron, basato su UnionFS [7], un tipo di *file system* sviluppato alla Stony Brook University dal 2004. Si tratta di un sistema di unificazione, che può riunire logicamente il contenuto di diverse cartelle, mantenendo separato il loro contenuto fisico.

Permette di unire parti accessibili in lettura e scrittura con altre solo in lettura, così come l'inserimento e la cancellazione di parti al volo.

Si è approfittato dell'installazione del nuovo sistema per sostituire il sistema di gestione delle code con PBS Professional [8], un sistema pressoché identico a livello di comandi con il precedente Torque, permettendo quindi una transizione indolore per gli utenti, ma con innovazioni rilevanti per quanto riguarda l'affidabilità e la gestione di sistemi eterogenei.

Il nuovo cluster permetterà inoltre ai nostri utenti l'utilizzo di applicazioni grafiche a alte prestazioni. Un nodo infatti è equipaggiato con una scheda video di alta potenza NVIDIA Quadro FX 4400, con 512MB di memoria GDDR3. Il software AMIRA [9] verrà utilizzato per la compressione del flusso di immagini a utenti remoti. Dettagli su questo nuovo servizio sono forniti in un apposito articolo [10].

Nella seguente tabella riportiamo le caratteristiche salienti del calcolatore.

	Cluster Exadron 64 bit
CPUs (nodi)	220 (55)
Processore	AMD Opteron 2.2 GHz
RAM per core	2-4 GB
Spazio disco	2.6 TB
OS	Linux CentOS 4.2
Kernel	2.6.14
Interfacce di rete	2 (Infiniband 4X, gigabit ethernet)

Modalità di utilizzo

La configurazione attuale prevede che gli utenti accedano direttamente al cluster collegandosi al nodo di front-end, di indirizzo avogadro.cilea.it, in modalità SSH v.2.

L'utilizzo degli altri nodi è consentito unicamente attraverso la modalità batch, usando il sistema di gestione di code PBS Professional. In particolare al momento sono definite una coda *opteron1*, a accesso libero, e una coda *reserved* a accesso ristretto.

Per i principali pacchetti applicativi sono stati preparati opportuni script ottimizzati per il corretto lancio in coda.

L'esecuzione di un job può avvenire sia a linea di comando, con *qsub*, sia con interfaccia grafica *xpbs*, che consente anche in maniera intuitiva il controllo e l'interruzione dei lavori. Seguirà un articolo dettagliato sull'utilizzo del sistema da parte degli utenti. Per ora rimandiamo gli interessati all'assistenza dell'autore (arlandini@cilea.it) oppure del Dott. Cremonesi (cremonesi@cilea.it).

Attualmente sono disponibili i seguenti pacchetti:

Ambienti di sviluppo:

(referente Maurizio Cremonesi)

Package	Release
GCC	3.4.4-2
Intel Fortran Compiler	9.0
Intel C++ Compiler	9.0
G95	4.0.1
Intel MKL	7.2
MPICH	1.2.5
Portland Fortran	6.0

Analisi strutturale:

(referente Maurizio Cremonesi)

Package		Release
ABAQUS		6.5
ANSYS		9.0
LS-DYNA		960
MSC suite	NASTRAN	2005
	PATRAN	2005
	MARC	2005
RADIOSS		4.4

Termo-fluidodinamica:

(referente Paolo Ramieri)

Package		Release
FLUENT suite	FLUENT	6.2.16
	FIDAP	8.7.2
	Gambit	2.2.30
	Tgrid	3.5
STARCD suite	Starcd/hpc	3.26
COMET		2.100.B
CFX		5.7.1

Elettromagnetismo:

(referente Maurizio Cremonesi)

Package	Release
ANSYS	9.0

Chimica e scienza dei materiali:

(referente Maurizio Cremonesi)

Package	Release
GROMACS	3.3
TURBOMOLE	5.6
CPMD	3.9.1
GAMESS	December 12, 2003
AMBER	7
CP2K	
OCTOPUS	2.0

Bibliografia

- [1] A. Cantore, "Un nuovo supercalcolatore al CILEA: un cluster da oltre un TeraFlops", Bollettino del CILEA, n. 90, Dicembre 2003.
- [2] C. Arlandini, "GALILEO: un nuovo server HP SuperDome per il calcolo parallelo al CILEA", Bollettino CILEA, n. 81, Febbraio 2002.
- [3] F. Bonini, A. Mattasoglio, "Nuovo calcolatore vettoriale del CILEA: Nec SX-5/4s", Bollettino del CILEA, n. 76, Febbraio 2001.
- [4] TOP500, URL: <http://www.top500.org>
- [5] EXADRON, URL: <http://www.exadron.com>
- [6] INFINIBAND specifications, URL: <http://www.infinibandta.org/home>
- [7] UnionFS, URL: <http://www.unionfs.org>
- [8] PBS Professional, URL: <http://www.pbs.com>
- [9] AMIRA, URL: <http://www.amiravis.com/>
- [10] P. Dagna, "Grafica alte prestazioni a distanza", Bollettino CILEA, n. 101, Aprile 2006.