UNIVERSITÀ DEL SALENTO

PROCEDURA APERTA PER L'ACQUISIZIONE DI ATTREZZATURA TECNICO – SCIENTIFICA NELL'AMBITO DEL PROGRAMMA OPERATIVO NAZIONALE "RICERCA E COMPETITIVITA' (PON "R&C") 2007 – 2013 – PROGETTO PONa3_00334 2HE "POTENZIAMENTO DEL CENTRO DI RICERCA PER LA SALUTE DELL'UOMO E DELL'AMBIENTE" – CUP: F81D11000210007 – Codice CIG: 5789278755

CAPITOLATO TECNICO

"Fornitura di un sistema di calcolo parallelo ibrido per le esigenze dell'attività S01 del Progetto PON 2HE"

RIPARTIZIONE LEGALE, ATTI NEGOZIALI ED ISTITUZIONALI
AREA AFFARI NEGOZIALI
UFFICIO APPALTI E CONTRATTI
Piazza Tancredi,7 -73100 Lecce
Tel . 0832 292319 - Fax. 0832 292345

SOMMARIO

1 Introduzione ... 3
2 Caratterizzazione della fornitura ... 3
3 Sistema di calcolo scalare/parallelo ibrido (CPU + GPU) ... 4
3.1 Architettura del sistema ... 4
3.2 Caratteristiche dei nodi di calcolo ... 4
3.2.1 Processori ... 4
3.2.2 GPU ... 4
3.2.3 Memoria ... 5
3.2.4 Storage locale ... 5
3.2.5 Connettività ... 5
3.2.6 Altre richieste ... 5
3.3 Caratteristiche del login node ... 5
3.4 Infrastruttura di rete ... 6
3.4.1 Rete InfiniBand ... 6
3.4.2 Rete in-band ... 6
3.4.3 Rete di management ... 6
3.5 Storage del sistema ... 7
3.6 Rack 42U ... 7
4 Software richiesto ... 8
4.1 Sistema di Management ... 8
4.2 Sistema di code ... 8
4.3 Ambienti di sviluppo ... 8
4.4 Filesystem parallelo ... 8
5 UPS ... 8
6 Installazione, verifica e collaudo della fornitura ... 9
6.1 Installazione ... 9
6.2 Verifica e collaudo ... 9
7 Garanzia e modalita' del servizio di assistenza tecnica e manutenzione ... 9
8 Modalità di fornitura e installazione ... 9

Introduzione

Il presente Capitolato Tecnico disciplina gli aspetti tecnici della fornitura di hardware/software/apparati e quanto altro necessario per la fornitura di un "*Sistema di calcolo parallelo ibrido",* ivi incluso l'installazione, la manutenzione e l'assistenza tecnica specialistica, per il potenziamento del Centro Ricerche per la salute dell'uomo e dell'ambiente nell'ambito del progetto PON Ricerca e Competitività 2007-2013 denominato "2HE Research Center".
Il "*Sistema di calcolo parallelo ibrido",* di seguito denominato **"Sistema",** sarà localizzato presso la sala server del Laboratorio DREAM – Palazzina Uffici, 1° piano, Ospedale Vito Fazzi, ASL Lecce, 73100 Lecce.

Caratterizzazione della fornitura

Oggetto della fornitura è l'insieme delle componenti (apparati, sistemi, software, ecc.) e dei servizi necessari alla realizzazione del **"Sistema"** e al suo corretto ed efficiente funzionamento.

Sinteticamente la fornitura consisterà in:

- **Sistema di calcolo scalare/parallelo ibrido (CPU + GPU);**
- **Sistema di storage;**
- **Gruppo statico di continuità (UPS).**

Eventuali altre componenti e servizi, anche se non esplicitamente menzionati ma comunque necessari per la gestione, l'integrazione e il corretto funzionamento del "**Sistema**" (ad es. cavi di collegamento, strumenti HW/SW per la configurazione, per la gestione e per il monitoraggio del "**Sistema**", firmware, ecc.) dovranno anch'essi essere compresi nella fornitura.
La fornitura, inoltre, dovrà appartenere alla più recente generazione di prodotti rilasciati in commercio ed essere costituita esclusivamente da elementi nuovi di fabbrica.

Sistema di calcolo scalare/parallelo ibrido (CPU + GPU)

L'obiettivo è quello di realizzare un sistema di High Performance Computing (HPC) basato su sistema operativo Linux a supporto delle attività del Centro Ricerche per la salute dell'uomo. Il Sistema sarà formato da N nodi di calcolo ibridi (CPU + GPU) ed un login node.
Il numero dei nodi di calcolo non viene fissato a priori in quanto il sistema che si vuole realizzare dovrà avere una capacità di calcolo di picco, considerando unicamente la "theoretical peak performance" delle CPU, pari ad almeno 2,6TFlops.

Pertanto il numero di nodi sarà determinato dalle caratteristiche della CPU.

Architettura del sistema

L'architettura del sistema di calcolo scalare/parallelo ibrido dovrà essere del tipo multi-nodo con nodi di calcolo e nodi di servizio (I/O nodes, login node, management e monitoring nodes).

L'architettura del sistema che s'intende realizzare è quella standard di un cluster:

- nodi di calcolo,
- nodo/nodi di servizio (I/O nodes, login node, management e monitoring nodes)
- infrastruttura di rete,
- sistema di storage.

il tutto connesso da opportune reti, di seguito specificate. Tutti i dispositivi che compongono il sistema dovranno essere ospitati all'interno di un rack opportunamente predisposto di PDU (Power Distribution Unit).

Caratteristiche dei nodi di calcolo

I nodi di calcolo dovranno avere le caratteristiche minime descritte di seguito.

Processori

- Ciascun nodo dovrà (pena esclusione) essere dotato di 2 processori multi-core x86 a 64bit del tipo INTEL Sandy Bridge o Ivy Bridge o versioni successive;
- ogni processore dovrà avere almeno 8 core fisici;
- frequenza del processore ≥ 2,4 GHz;
- ogni processore deve avere almeno 20 MB di cache L3.

Per il calcolo delle prestazioni di picco del sistema verrà presa in considerazione esclusivamente la frequenza nominale dei processori, escludendo meccanismi di burst, overclocking o similari.

GPU

- Ciascun nodo dovrà (pena esclusione) essere equipaggiato con n. 2 schede GPU di ultima generazione equivalenti alle schede NVIDIA Kepler o modelli successivi;
- Ogni scheda GPU dovrà avere i seguenti requisiti minimi:

- Potenza di picco (doppia precisione in virgola mobile) ≥ 1,17 TFlops;
- Potenza di picco (singola precisione in virgola mobile) ≥ 3,52 TFlops;
- Dimensione della memoria ≥ 5GB.

Le GPU non concorrono al calcolo della potenza di picco del sistema.

Memoria

- Ciascun nodo dovrà (pena esclusione) essere equipaggiato con almeno 64 GB di RAM;
- Ciascun nodo dovrà essere dotato di memorie del tipo DDR-3 registered ECC ed operanti, nel sistema fornito, ad una frequenza effettiva di almeno 1600 MHz;
- I moduli di memoria dovranno essere approvati dal costruttore della scheda madre specificamente per l'utilizzo della scheda madre fornita;
- I canali di memoria dovranno essere popolati in maniera bilanciata ed in base alle indicazioni fornite dal produttore della scheda madre al fine di ottenere le prestazioni ottimali;
- Non sarà permesso combinare moduli di memoria con differente dimensione, tipo, velocità o fabbricante.

Storage locale

I nodi di calcolo dovranno essere dotati di un disco di piccole dimensioni per il solo sistema operativo con tecnologia SATA o near-line SAS. Indicativamente è richiesto un disco di almeno 500 GB di capacità e una velocità di almeno 7.2K RPM.

Connettività

Ciascun nodo dovrà essere fornito di:

- n. 2 schede Gigabit Ethernet rame integrate con supporto PXE ciascuna dotata di connettore RJ45 per i cavi in rame (1000baseT);
- n. 1 scheda Infiniband Dual-port a 40Gb/s minimo.

Altre richieste

- Tutti i sistemi oggetto del presente capitolato dovranno essere forniti con identiche caratteristiche hardware, versioni di firmware e di BIOS;
- I cavi interni non dovranno essere punzonati, danneggiati o in tensione meccanica;
- Le funzionalità dei nodi, definibili via setting NVRAM (per esempio ordine di boot, configurazione scheda madre, etc.), dovranno essere mantenute anche in caso di accidentale perdita di alimentazione elettrica.

Caratteristiche del login node

Il login node del Sistema dovrà avere la stessa identica configurazione dei nodi di calcolo tranne che per la configurazione dello storage locale (Sezione 3.2.4). In tal caso, è richiesta una

configurazione con minimo n.2 hard disk in configurazione RAID 1 ("mirroring") al fine di garantire l'alta disponibilità del servizio ed una capacità raw di almeno 2TB.
Inoltre il login node dovrà essere dotato di alimentazione ridondata con alimentatori di tipo hot swap.

Infrastruttura di rete

Il Sistema sarà interconnesso tramite reti diverse per scopi diversi. In particolare, dovranno essere previste le seguenti tre reti distinte: rete InfiniBand, rete in-band, rete di management.
La fornitura dovrà comprendere, per tutte le reti sotto descritte, i relativi apparati di interconnessione e cablaggio, come meglio specificato nei paragrafi successivi.

Rete InfiniBand

Al fine di connettere tra di loro tutti i nodi di calcolo, il nodo di login/management e il sistema di storage, tramite una rete ad alta velocità e bassa latenza si richiede la fornitura di una rete InfiniBand.
Pertanto si richiede la fornitura di:

- n. 1 Switch InfiniBand che soddisfi i requisiti minimi di seguito indicati;
- tutti i cavi necessari, in rame o fibra, per collegare i singoli nodi allo switch.

La rete di interconnessione InfiniBand dovrà avere i seguenti requisiti minimi:

- Switch di connessione QDR o versioni successive almeno a livello di backplane e schede sui nodi QDR o versioni successive;
- 1:1, full bandwidth (non-blocking);
- banda passante punto punto non inferiore a 40Gbs;
- latenza punto punto internodo non superiore a 1.5 microsecondi.

Rete in-band

Tutti i nodi del Sistema dovranno essere connessi con rete interna (in-band) di tipo Ethernet con velocità di 1Gbps.
La fornitura dovrà prevedere uno switch Layer 2 stackable con un opportuno numero di porte. La fornitura dovrà essere compresa di cablaggio in Cat. 6.

Rete di management

Al fine di gestire dal nodo di management tutti i nodi (per mezzo delle relative Schede di Remote Management), si richiede la fornitura di una rete di management esterna (out-of-band). La tipologia richiesta è Ethernet con velocità di almeno di 1Gbps.
La fornitura dovrà prevedere uno switch Layer 2 stackable con un opportuno numero di porte. La fornitura dovrà essere compresa di cablaggio in Cat. 6.

Sarà accettata anche una soluzione in cui la rete di management esterna (out-of-band) e la rete di connessione interna (in-band) condividono lo stesso switch purchè l'offerta preveda un numero

sufficiente di porte per entrambi le reti. In questo caso (infrastruttura unica) la soluzione dovrà prevedere la possibilità di gestione separata delle reti tramite VLAN.

Storage del sistema

Il cluster dovrà essere dotato di un sistema di storage con le seguenti caratteristiche minime:

- capacità iniziale di almeno 60TB raw;
- performance in scrittura/lettura del sistema di almeno 1GBytes/sec;
- compatibilità con soluzioni di filesystem paralleli (Lustre 2.x e/o GPFS) che comunicheranno via rete IB con i clients (nodi di calcolo);
- doppio controller per garantire la massima affidabilità e disponibilità, capaci di gestire configurazioni con dischi hot spare, dischi hot swap;
- entrambi i controller dovranno supportare la creazione di RAID set con livelli 0,1,5,6.

Il sistema di storage sarà partizionato in un'area scratch (maggioritaria) e in una area utenti (/home max 15 TB) entrambi gestiti da un file system parallelo.

Rack 42U

Tutti i dispositivi precedentemente descritti dovranno essere ospitati in un armadio Rack da 19" che dovrà esser consegnato con tutti gli accessori (rotaie, guide telescopiche, viti specifiche, dadi, etc.) atti al montaggio degli stessi; tutte le apparecchiature dovranno poter scorrere su rotaie telescopiche ed essere estraibili dal rack per ispezione/manutenzione.

L'armadio Rack dovrà essere dotato di un numero adeguato di PDU per consentire l'alimentazione di tutte le apparecchiature fornite con il massimo carico previsto.

Software richiesto

Il Sistema oltre ad avvalersi di un Sistema Operativo Unix/Linux dovrà essere corredato di software rispondente ai seguenti requisiti:

Sistema di Management

Software in grado di amministrare e gestire l'intero Sistema (installazioni, upgrade, inventari HW/SW, system monitoring, ecc..) da un unico punto di controllo, in command line e da interfaccia grafica.

Sistema di code

Software (tipo LSF) per l'implementazione di un sistema di code (batch, pipe, ecc.) che permetta la sottomissione di job seriali/paralleli e la gestione dell'allocazione delle risorse.

Ambienti di sviluppo

Suite per lo sviluppo e il debugging di codici seriali e paralleli costituita da:

- compilatori Fortran, C e C++ con l'opzione di parallelizzazione automatica e/o OpenMP nelle ultime versioni disponibile per l'architettura proposta;
- ambiente software per lo sviluppo e l'esecuzione di applicazioni parallele basate su message-passing MPI con supporto del network di interconnessione del cluster;
- debugger e profiler, anche per programmi paralleli, dotati di GUI;
- CUDA package: CUDA driver, CUDA libraries, CUDA SDK, GPU burn-in (per sistemi GPU).

Filesystem parallelo

Software per implementare un filesystem parallelo come, ad esempio, le soluzioni Lustre o GPFS. Il software offerto, nella versione più aggiornata disponibile, dovrà essere inclusivo di tutte le licenze necessarie per la completa funzionalità dello stesso anche per quelle caratteristiche non elencate sopra.

UPS

Al fine di garantire la continuità dell'alimentazione elettrica, si richiede la fornitura di un UPS dotato di sistema di protezione da sovratensione.
La configurazione dell'UPS dovrà soddisfare i seguenti requisiti minimi:

- essere opportunamente dimensionato in termini di potenza per proteggere tutti i dispositivi che fanno parte del sistema di calcolo e storage offerto;
- garantire un'autonomia di funzionamento dell'intero Sistema per almeno 15 minuti;
- essere di tipo rack-mountable.

Installazione, verifica e collaudo della fornitura

Installazione

- Tutto l'hardware fornito dovrà essere installato nei rack, reso operativo e validato ad opera del fornitore;
- L'installazione dovrà comprendere per tutte le apparecchiature la cablatura all'alimentazione elettrica e alla rete;
- Per quanto riguarda il software, l'installazione dovrà essere condotta in collaborazione con il committente per il necessario "training on the job" dello strumento di installazione scelto dal fornitore, qualora questo non sia conosciuto dal committente. Lo strumento di installazione scelto dovrà comunque essere totalmente compatibile con le caratteristiche specificate nelle procedure di collaudo.

Verifica e collaudo

La verifica deve accertare che la fornitura, per quanto riguarda il numero e la tipologia dei componenti, tecniche e metodologie impiegate, l'esecuzione e le funzionalità, siano in tutto corrispondenti a quanto previsto dai documenti della procedura in questione.

Il collaudo consiste in una serie di prove atte a verificare la perfetta esecuzione del contratto quindi di quanto eseguito e consegnato e il rispetto di tutte le funzionalità richieste.

Garanzia e modalita' del servizio di assistenza tecnica e manutenzione

Si richiede per tutti i dispositivi hardware un servizio di manutenzione on site di tipo NBD (Next Business Day) della durata di tre anni.

Modalità di fornitura e installazione

Il servizio di consegna ed installazione dovrà essere erogato dal Fornitore, attraverso proprio personale specializzato. Tali attività si intendono comprensive di ogni onere relativo al trasporto, facchinaggio, consegna "al piano", posa in opera, asporto dell'imballaggio e di qualsiasi altra attività ad esse strumentale. Il Fornitore, inoltre, dovrà dotarsi di mezzi opportuni e/o di quanto altro necessario a trasportare, scaricare e a collocare la fornitura nei siti suddetti.

Il Fornitore garantirà il rispetto delle seguenti normative:

- D.LGS. 81/08 "Testo unico in materia di salute e sicurezza nei luoghi di lavoro"
- DM 37/08 "Disposizioni in materia di attività di installazione degli impianti all'interno degli edifici."

E' altresì a cura del Fornitore il rispetto delle normative di impiantistica, secondo la legislazione in vigore.

Inoltre, contestualmente alla consegna della fornitura e durante il periodo di validità del contratto, per ogni apparecchiatura HW dovrà essere consegnata al referente tecnico la seguente documentazione:

- manualistica (installation guide, hardware technical reference, operation's guide, ecc.), in lingua italiana o inglese, su supporto cartaceo e/o su CD;
- i certificati di Conformità CE, recanti le norme tecniche e legislative cui il produttore/importatore si è attenuto in tutte le fasi di progettazione, costruzione e collaudo;

per i software:

- la manualistica (installation guide, user's guide, ecc.) in lingua italiana o inglese, su supporto cartaceo e/o su CD;
- i CDROM/DVD, i certificati di licenza intestati all'Università del Salento, le key di attivazione e i relativi contratti di manutenzione validi per l'intero periodo contrattuale.

A tal riguardo, tutti i software forniti dovranno avere licenza d'uso permanente.