



Progetto "Centro Computazionale di RicErca sui Sistemi COMplessi - CRESCO"
(Prot. 750/6). Programma Operativo Nazionale 2000-2006 "Ricerca Scientifica, Sviluppo Tecnologico, Alta Formazione" Misura II.2 "Società dell'Informazione per il Sistema Scientifico Meridionale", Azione a – Sistemi di calcolo e simulazione ad alte prestazioni. Avviso 1575/2004.

Specifica Tecnica

Specifiche funzionali e caratteristiche tecniche per la fornitura e posa in opera di un sistema di supercalcolo presso il centro ENEA di Portici (NA).



www.cresco.enea.it

INDICE

Premessa	3
1 Specifiche Tecniche	4
Introduzione	4
Oggetto della fornitura	4
Art. 1.1 Sistema di supercalcolo HPC	5
Art. 1.1.1 Sezione 1.....	5
Art. 1.1.2 Sezione 2.....	6
Art. 1.1.3 Sezione 3.....	7
Art. 1.1.4 Caratteristiche di configurazione	7
Art. 1.1.5 Sistema di controllo e gestione	7
Art. 1.2 Rete di interconnessione	8
Art. 1.2.1 Rete ad alta banda e bassa latenza InfiniBand	8
Art. 1.2.2 Rete Gbit Ethernet.....	9
Art. 1.2.3 Rete di interconnessioni alla rete Locale (LAN) 10Gbit Ethernet	9
Art. 1.4 Software	10
Art. 1.4.1 Software di sistema.....	10
Art. 1.4.2 Software di gestione	11
Art. 1.4.3 Middleware ENEA-GRID.....	11
Art. 1.4.4 Software di monitoring	11
Art. 1.4.5 Ambiente software per lo sviluppo di calcolo parallelo	12
Art. 1.4.6 Software applicativo	12
Art. 1.5 Sistema di backup	12
Art. 1.5.1 Libreria nastri	13
Art. 1.5.2 Server di Backup Master	13
Art. 1.5.3 Media Server	13
Art. 1.6 Sistema di sicurezza	13
Art. 1.7 Sistema di alimentazione e di continuità	13
Art. 1.8 Adeguamento sistema condizionamento	14
Art. 1.9 Supporto on site	14
Art. 1.10 Documentazione richiesta	14
2 Specifiche di collaudo	15
Art. 2.1 Verifica formale della fornitura	16
Art. 2.2 Collaudo hardware delle apparecchiature	16
Art. 2.3 Collaudo software di sistema	16
Art. 2.4 Collaudo software applicativo	16
3 Allegati	16



Premessa

Il Progetto CRESCO ha come obiettivo la realizzazione, presso il Centro Ricerche ENEA di Portici (NA), di un importante Polo di calcolo multidisciplinare per lo studio dei sistemi complessi di natura biologica e tecnologica, collegato con facilities di elaborazione, archiviazione e visualizzazione 3D di dati ubicate presso i Centri ENEA di Brindisi e Trisaia (MT).

Sul versante infrastrutturale il Progetto prevede la realizzazione di una piattaforma di calcolo di assoluta eccellenza in grado di erogare una potenza di almeno 20 TeraFlops, corredata di una nuova classe di funzionalità GRID innovative.

Sul piano delle applicazioni e dei contenuti scientifici, il Centro si focalizza sui seguenti ambiti principali:

- lo studio di oggetti biologici dal punto di vista "sistemico" e lo studio di sistemi naturali (comunità animali e sociali) complessi;
- lo studio di sistemi di reti tecnologiche complesse e delle loro mutue interazioni, e la realizzazione di opportuni strumenti per il modelling, la simulazione e il controllo;
- l'implementazione di soluzioni innovative in tema di architetture di sistemi di calcolo e di GRID computing per le attività di R&S di punta dell'ENEA che richiedano l'utilizzo di risorse computazionali estremamente importanti.

Il Progetto prevede una spesa complessiva di € 9.292.000,00 e si dovrà concludere entro e non oltre il 31/12/2008. Il coordinamento generale del Progetto è svolto dal Dipartimento Tecnologie Fisiche e Nuovi Materiali mentre le attività di ricerca prevedono il coinvolgimento di numerose altre Unità organizzative dell'ENEA. È previsto inoltre il coinvolgimento di 15 partner esterni quali Università, Consorzi e PMI la cui attività nell'ambito del Progetto CRESCO è regolata attraverso appositi contratti di ricerca.

La presente specifica tecnica ha lo scopo di fornire tutte le indicazioni necessarie per poter formulare un'offerta tecnico-economica per la fornitura di seguito descritta e sarà parte integrante del Contratto di Fornitura.

Il documento è organizzato nelle seguenti sezioni:

1. **Specifiche Tecniche:** descrive le funzioni che dovranno essere rispettate dal sistema e sul rispetto delle quali sarà valutata l'offerta tecnico-economica.
2. **Specifiche di Collaudo:** riporta le procedure di collaudo.
3. **Allegati:** sono documenti a supporto di quanto descritto nei paragrafi 1 e 2.

1 Specifiche Tecniche

Introduzione

Le esigenze computazionali espresse dalle attività previste dal progetto CRESCO e, più in generale, dalle attività di Ricerca e Sviluppo dell'ENEA, richiedono l'allestimento di una facility di supercalcolo particolarmente versatile ed in grado di rispondere ad una vasta gamma di tipologie di applicazione che possono essere schematizzate come segue:

- Codici seriali;
- Codici paralleli fortemente accoppiati;
- Codici paralleli debolmente accoppiati;
- Più codici paralleli fortemente accoppiati che interagiscono fra loro (es. accoppiamento fluido struttura);
- Codici con algoritmica particolare in relazione ad applicazioni specifiche (ad es. analisi immagini, genomica)

Queste tipologie di applicazione comportano, inoltre, esigenze diversificate in base ad almeno altre due caratteristiche specifiche:

- l'interazione con i dati di input ed output, in base alle dimensioni e alle frequenze di lettura e scrittura richieste;
- la dimensione della memoria RAM visibile direttamente dal singolo nodo SMP (Symmetric Multi-Processor a memoria condivisa).

Il sistema che si vuole realizzare deve essere in grado di rispondere a tutte queste esigenze contemporaneamente.

Oggetto della fornitura

Oggetto della gara è la progettazione di dettaglio, la fornitura, il trasporto e la messa in opera, presso il centro ENEA di Portici, di un sistema per il calcolo scientifico ad alte prestazioni (**HPC**) completo in tutte le sue parti hardware e software.

Le principali componenti del sistema sono:

1. sistema di supercalcolo HPC;
2. rete di interconnessione HPC;
 - rete ad alta banda e bassa latenza InfiniBand (IB);
 - rete Gbit Ethernet;
 - rete di management dedicata
 - SAN, Storage Area Network;
 - interconnessioni alla rete Locale (LAN) 10Gbit Ethernet;
3. sistema di dischi ad alte prestazioni;
4. software;
 - software di sistema;
 - software di gestione;
 - middleware ENEA-GRID;
 - software di monitoring;

- ambienti per lo sviluppo applicativi ad alto parallelismo;
 - software applicativo;
5. sistema di backup;
 6. sistema di sicurezza;
 7. sistema di alimentazione e di continuità;
 8. adeguamento del sistema di condizionamento.

Tutti i componenti, oggetto della fornitura, dovranno essere collocati in sala calcolo (v. Tav. 1. Allegato 1) con eccezione di quelli esplicitamente posizionati in altro luogo. Tutta la componentistica fornita, ad eccezione dei componenti della sala controllo Art. 1.1.5, dovranno essere installati in rack.

L'ottimizzazione dell'utilizzo dello spazio della sala di calcolo CRESCO, sarà considerato elemento di merito preferenziale. In particolare sarà elemento di merito preferenziale lo spazio che sarà lasciato libero e la sua distribuzione.

Art. 1.1 Sistema di supercalcolo HPC

Il sistema di supercalcolo sarà costituito da 3 Sezioni computazionali e da un sistema di controllo e gestione:

- Sezione 1, dedicata principalmente all'esecuzione di codici in modalità SMP o di codici con elevate esigenze di memoria sul singolo nodo;
- Sezione 2, dedicata principalmente a codici ad elevato parallelismo, che necessitano di una rete di interconnessione ad alte prestazioni;
- Sezione 3, dedicata a sistemi innovativi (multi-core ad architettura asimmetrica).

Per codici ad alto parallelismo ma debolmente accoppiati, le Sezioni 1 e 2 dovranno essere in grado di operare con lo stesso applicativo come un unico sistema e con prestazioni di connettività almeno pari a quelle garantite dalla rete Gbit Ethernet. Il livello di integrazione ed omogeneità della Sezione 1 con la sezione 2 costituirà elemento di merito preferenziale. Altresì costituirà elemento di merito preferenziale la fornitura di un numero di core superiore a quello minimo richiesto per ogni Sezione.

Art. 1.1.1 Sezione 1

La Sezione 1 dovrà essere costituita da **almeno 512** core organizzati ad uno o più nodi SMP ognuno dei quali ad immagine unica, con almeno 16 Core per nodo.

Dovranno essere garantite le seguenti caratteristiche minime:

- Architettura a 64 bit nativa;
- Rpeak minimo per core 8,80 GFlops
($R_{peak} = \text{FrequenzaCPU(GHz)} \times \text{Numero-operazioni-Floating-Point-per-ciclo}$)
- per ogni 16 core almeno doppia interconnessione Ethernet ad almeno 1Gbit/s.
- una memoria RAM di almeno 2 GByte per core (tecnologia almeno DDR2);
- due dischi SAS di almeno 73 GByte per ogni 16 core;
- predisposizione per connessione InfiniBand 4X per ogni nodo SMP;
- possibilità di eseguire contemporaneamente applicazioni sia a 32 che 64 bit;
- controllo della potenza elettrica in funzione del carico;

- alimentazione ridondata per nodo o blade chassis in relazione alla tecnologia proposta.
- la connessione al sistema di storage ad alte prestazioni deve essere garantita o con una scheda FC4 "dual link" per ogni 16 core o tramite l'estensione della rete InfiniBand 4X della sezione 2 prevedendo almeno una connessione del tipo descritto nell' Art. 1.2.1 ogni 16 core.

La Sezione 1 comprenderà inoltre le seguenti schede installate e configurate:

- **6 schede FPGA VIRTEX5 LX330**, il relativo ambiente di sviluppo Xilinx ISE 9.1 ed i software di gestione;
- **4 schede di grafica NVIDIA Quadro FX 4500 X2** da installare su alcuni dei nodi.

La modalità di distribuzione delle schede all'interno dei nodi sarà concordata tra ENEA e la ditta vincitrice in fase di implementazione finale.

Costituirà titolo di merito preferenziale:

- il peak transfer-rate della memoria RAM, espresso in GB/s;
- la fornitura di almeno 64 core con memoria maggiore di 2 Gbyte per core;
- l'estensione della rete IB della Sezione 2 alla Sezione 1;
- l'accesso diretto via IB dei nodi SMP al sottosistema dischi ad alte prestazioni di cui all'Art. 1.3.

Art. 1.1.2 Sezione 2

La Sezione 2 dovrà avere almeno **2048** core organizzati ad uno o più nodi SMP ad immagine unica, con almeno 8 Core per nodo.

Dovranno essere garantite le seguenti caratteristiche minime:

- Architettura a 64 bit nativa;
- Rpeak minimo per core 8,80 GFlops
($R_{peak} = \text{FrequenzaCPU(GHz)} * \text{Numero-operazioni-Floating-Point-per-ciclo}$)
- una memoria RAM di almeno 1 GByte per core (tecnologia almeno DDR2);
- un disco SAS da 73 GByte ogni 8 core;
- doppia connessione almeno a 1 Gbit Ethernet per ogni 8 core;
- almeno una connessione InfiniBand 4X (Art. 1.2.1), per ogni 8 core;
- possibilità di eseguire contemporaneamente applicazioni sia a 32 che 64 bit;
- controllo della potenza elettrica in funzione del carico;
- alimentazione ridondata per nodo o blade chassis in relazione alla tecnologia proposta.
- la connessione al sistema di storage ad alte prestazioni deve essere garantita o tramite la rete InfiniBand 4X (Art. 1.2.1) o con una scheda FC4 per ogni 8 core.

Costituirà titolo di merito preferenziale:

- il peak transfer-rate della memoria RAM, espresso in GB/s;
- la fornitura di almeno 128 core con memoria maggiore di 1Gbyte per core;
- l'accesso diretto via IB dei nodi SMP al sottosistema dischi ad alte prestazioni di cui all'Art. 1.3.

Art. 1.1.3 Sezione 3

Nell'ambito del progetto CRESCO è previsto lo sviluppo di specifici tool per l'utilizzo di Cell Processor, prodotto da IBM. Si richiede pertanto la fornitura di una sezione speciale con almeno 4 **Cell Broadband Engine System** IBM aventi le seguenti caratteristiche:

- Cell BE Processor Blade (~500 GFLOPS di picco)
- Dual Cell BE Processor Configuration;
- 3.2 GHz (o l'ultima tecnologia disponibile all'atto dell'offerta);
- 1GB XDRAM (512 MB per processore);
- Blade-mounted 40 GB IDE HDD;
- Dual Gigabit Ethernet (GbE) controllers;
- Predisposizione InfiniBand 4X Host Channel Adapters.
- la connessione al sistema di storage ad alte prestazioni (Sezione 1.3) deve essere garantita o in tecnologia FC4 o in tecnologia IB.

Art. 1.1.4 Caratteristiche di configurazione

Il sistema dovrà essere installato presso il CR ENEA di Portici nei locali descritti nell'**Allegato 1** e configurato come descritto nella sezione relativa ai software.

Le sezioni 1 e 2 dovranno essere configurate in modo da avere ciascuna **almeno 4 nodi di front-end**. La configurazione di tali nodi dovrà essere tale da consentire lo sviluppo di codici nelle varie opzioni seriali e parallele. Pertanto tali nodi dovranno essere dotati di tutti gli ambienti di compilazione come descritto nei successivi paragrafi relativi al software. I nodi di front-end dovranno essere dotati di schede grafiche in grado di supportare la grafica 3D.

I software applicativi richiesti da ENEA (v. Art. 1.4.6) dovranno essere installati nell'ambiente GRID-ENEA, ed utilizzabili da un generico utente del sistema attraverso le code di LSF. Per i software applicativi con licenze d'uso, le licenze saranno messe a disposizione dall'ENEA, ma sarà a carico del fornitore la loro installazione e personalizzazione al fine di consentire il corretto funzionamento dell'intero sistema. Dovranno inoltre essere realizzate le opportune procedure di sottomissione dei software applicativi con LSF in grado di utilizzare tutte le risorse del sistema, sezione 1 e 2.

Art. 1.1.5 Sistema di controllo e gestione

Il sistema di controllo e gestione dovrà essere dotato di almeno una consolle per ciascuna Sezione nella sala calcolo e di almeno 3 postazioni di lavoro posizionate nell'area definita come sala controllo (v. Tav. 1. dell'Allegato 1).

Ogni postazione di lavoro dovrà consentire:

- la gestione del sistema;
- le attività di supervisione e monitoring.

Ogni singola postazione deve essere dotata di almeno due monitor 19" non CRT.

Nella sala controllo dovrà essere installato un monitor a parete di almeno 50" che, attraverso un apposito hardware, fornisca la visualizzazione simultanea del contenuto dei 6 monitor di controllo.

Al fine di garantire tutte le funzionalità richieste al sistema è necessario fornire in sala calcolo un rack 42U 19" con i server necessari ai servizi funzionali del sistema e precisamente:

- servizi di File server AFS (almeno 2 server con almeno 2 schede FC4 dual port ognuno);
- servizi di File server GPFS (ove necessari per supportare le prestazioni richieste nell'Art. 1.3);
- servizi di installazione, boot, configurazione e test di sistema;
- servizi di server di licenze ridondati.

Si precisa che ogni servizio principale deve avere almeno un server indipendente e che comunque essi devono essere in numero adeguato a garantire i livelli prestazionali che saranno dichiarati in fase di offerta. Tutti i server dedicati ai servizi devono avere processore Intel/AMD (almeno dual core, 2 GB RAM), doppio disco in RAID1 e l'alimentazione ridondata.

In questo rack dovranno inoltre essere installati con il sistema operativo e connessi in rete ulteriori 15 server 1U, con almeno 1 socket dual-core (Intel/AMD), 2GByte RAM, 2 HD configurati RAID1, per servizi che saranno successivamente installati da ENEA.

Il rack dovrà essere gestito tramite una console dedicata o tramite una delle console relative alla gestione delle sezioni 1, 2 ,3.

Art. 1.2 Rete di interconnessione

Il sistema HPC dovrà essere configurato con almeno le seguenti tipologie e di rete:

- rete InfiniBand per il calcolo ad alte prestazioni ed accesso ai dati ad alta performance (dettagli all'Art. 1.2.1);
- rete Gbit/s per tutte le funzioni di calcolo, di accesso ai dati e di servizio in cui la latenza non è critica (dettagli Art. 1.2.2);
- rete di servizio necessaria a tutte le operazioni di gestione, fisicamente indipendente dalle due reti di cui sopra, collegata direttamente alle console.
- rete 10 Gbit/s per connessione al centro stella della rete LAN dell'ENEA di Portici;
- una Storage Area Network per connessione di tutti i sistemi all'area dati ad alte prestazioni, nella accezione specificata all'Art. 1.2.4;

Nella fase di offerta si dovranno dichiarare le performance delle reti che si implementeranno e predisporre i test per verificarne le rispondenze. Sarà cura del fornitore la configurazione ed l'ottimizzazione di tutte le reti afferenti al sistema offerto, interfacciandosi ove necessario con i responsabili ENEA per il rispetto delle politiche di sicurezza ENEA sulla rete LAN/WAN.

Art. 1.2.1 Rete ad alta banda e bassa latenza InfiniBand



La rete IB 4X deve interconnettere tutti i nodi SMP della Sezione 2 e rispettare le sue caratteristiche nominali. Lo switch IB deve avere un numero minimo del 5% di porte libere attive, rispetto alla configurazione offerta.

La rete IB deve avere al minimo le seguenti caratteristiche fisiche:

- architettura IB 4X DDR (20 Gbps picco), latenza inferiore ai 4 microsec;
- architettura IB full no-blocking.

Nell' offerta devono essere specificate:

- le prestazioni della rete offerta;
- le prestazioni sostenute in termini di scalabilità uno-su-tutti e tutti-su-tutti;
- le prestazioni di I/O complessive;
- l'architettura globale della rete IB e sua integrazione con il sistema dischi ad alte prestazioni (v. Art. 1.3).

Art. 1.2.2 Rete Gbit Ethernet

Come specificato negli Art. 1.1.1, Art. 1.1.2, Art. 1.1.3, Art. 1.1.5, Art. 1.5.1, Art. 1.5.2, Art. 1.5.3, tutti i nodi di calcolo SMP e di servizio dovranno essere interconnessi per mezzo di almeno 2 reti Gbit/s Ethernet. Tutte le connessioni faranno capo a uno o più switch locali, facenti parte della fornitura, che consentiranno la gestione del traffico locale al massimo delle performance. Il dimensionamento degli switch dovrà avere un numero minimo del 10 % di porte libere. Le configurazioni di dettaglio delle reti dovranno essere concordate con il responsabile ENEA.

Art. 1.2.3 Rete di interconnessioni alla rete Locale (LAN) 10Gbit Ethernet

Gli apparati di rete attivi di cui all'Art. 1.2.2 dovranno avere in totale almeno 4 accessi 10 Gbit Ethernet verso il Centro stella dell'ENEA. Per raggiungere il Centro stella l'ENEA mette a disposizione un fascio di fibre ottiche. Sarà cura della ditta vincitrice interfacciare i due sistemi (switch di cui all'Art. 1.2.2 e centro stella) sulla dorsale 10 Gbit Ethernet, in collaborazione con il personale ENEA.

Art.1.2.4 Storage Area Network (SAN)

Il fornitore dovrà fornire il progetto preliminare del sistema SAN. Data la criticità del sistema esso dovrà essere ridondato. Il progetto del sistema SAN sarà fortemente condizionato dalla soluzione tecnologica globale che sarà adottata per l'accesso da parte della Sezione 1, Sezione 2 e Sezione 3 al Sistema di storage ad alte prestazioni Art. 1.3. In ogni caso deve essere prevista una SAN che connette almeno :

- tutte le sezioni del sistema ove necessario;
- i server dotati di schede FC4 Art. 1.1.5;
- il sistema di Backup e relativi server, Art. 1.5;
- il sistema di storage ad alte prestazioni con almeno 8 FC4, Art. 1.3;

Inoltre il sistema SAN dovrà avere un numero minimo di 12 porte libere FC4 attive.

L'implementazione del sistema SAN puo' anche essere realizzato sfruttando l'integrazione con il sistema InfiniBand tramite opportuni gateway.

Art. 1.3 Sistema di storage ad alte prestazioni

Tutte e tre le sezioni del sistema dovranno accedere al sistema di storage ad alte prestazioni. Per la Sezione 1 e la Sezione 2 dovranno essere garantite le prestazioni globali di seguito riportate.

Sarà cura della ditta fornitrice definire l'architettura globale del sistema (SAN, Server GPFS, Server AFS, connessione IB)

Il sistema dischi dovrà avere al minimo le seguenti caratteristiche:

- capacità grezza minima 120 TByte;
- capacità in configurazione RAID6 96 TByte;
- prestazioni di picco minime 3 GByte/s;
- prestazioni sostenute di scrittura aggregata con File system GPFS di almeno 2GByte/s (con file system di almeno 50 TB)
- interfacce minime:
 - n. 8 FC4;
 - supporto nativo InfiniBand (IB);
 - IB e FC4 devono poter essere interconnesse allo stesso controller.

Sono caratteristiche di merito preferenziale della fornitura:

- l'attivazione delle connessioni IB con la Sezione 2;
- una uguale velocità in scrittura e lettura;
- la verifica della parità in scrittura e lettura;
- il livello di degrado delle prestazioni nella fase di ricostruzione del dato espresso in percentuale di degrado rispetto al valore delle prestazioni di picco;
- la potenza elettrica necessaria al funzionamento e l'eventuale presenza di funzioni "sleep mode";
- la possibilità di avere dischi con diverse capacità (500GByte e 1000GByte) sotto lo stesso controller.

Art. 1.4 Software

Il sistema dovrà essere installato e configurato con i software di seguito descritti.

Il fornitore dovrà installare i software secondo le indicazioni che saranno fornite da ENEA. Il fornitore dovrà garantire:

- il supporto necessario alla realizzazione delle procedure per il corretto utilizzo degli applicativi (Art. 1.4.6);
- il corretto funzionamento dei componenti software e dei servizi principali di sistema (Artt. 1.4.2, 1.4.3, 1.4.4, 1.4.5).

Art. 1.4.1 Software di sistema

Il sistema dovrà essere configurato con tutti i driver necessari al corretto funzionamento delle componenti hardware.

Il sistema operativo dovrà essere **Linux 64 bit** supportato a livello industriale.

Costituiranno titolo di merito preferenziale:

- l'adozione di una distribuzione Redhat o Suse/Novel
- il livello e la qualità di supporto garantito;
- la certificazione dell'hardware della Sezione 1 e della Sezione 2 per il sistema operativo prescelto;

- la certificazione del software applicativo elencato all'Art. 1.4.6. che può essere sostituita da dichiarazione equivalente del fornitore;

Il sistema operativo prescelto dovrà essere integrato, a cura del fornitore, con i componenti del middleware di ENEA-GRID (v. Art. 1.4.3).

Art. 1.4.2 Software di gestione

Dovrà essere fornito un sistema di gestione efficiente ed unificato per le Sezioni 1 e 2. Il sistema di gestione dovrà essere descritto in tutte le sue funzionalità. Nella valutazione del sistema di gestione sarà posta particolare attenzione alle funzionalità di installazione e configurazione da rete, analisi ed individuazione automatica dei guasti, monitoring di tutte le componenti di rete del sistema. Costituirà elemento di merito preferenziale per la valutazione, e sarà oggetto di collaudo, il tempo massimo per l'esecuzione della procedura automatica di installazione completa di tutti i nodi della Sezione 1 e della Sezione 2.

Art. 1.4.3 Middleware ENEA-GRID

Poiché l'infrastruttura CRESCO deve essere inserita all'interno della esistente infrastruttura ENEA-GRID, è necessario che i componenti essenziali di ENEA-GRID siano installati. Il middleware di ENEA-GRID consiste di:

OpenAFS: l'implementazione Open Source del file system OpenAFS è disponibile in rete e tutte le macchine della infrastruttura dovranno essere installati come client OpenAFS della cella AFS enea.it. Saranno anche installate le componenti File server di OpenAFS sulle macchine dedicate a tale servizio (v. Art. 1.1.5).

LSF: ENEA-GRID utilizza LSF/LSF Multicluster, il gestore di risorse di Platform nella sua versione Multicluster. Tale componente dovrà essere reso disponibile in numero di licenze pari al numero totale di core dei nodi relativi alla Sezione 1 e Sezione 2 aumentato del 15%. Le licenze dovranno essere quelle di LSF base + Multicluster, di tipo "perpetuo" e dotate di supporto fino alla scadenza della garanzia dell'intero sistema.

GPFS: il file system parallelo GPFS dovrà essere reso disponibile nelle componenti client sui nodi di calcolo delle Sezioni 1 e 2 e nelle componenti server sulle macchine dedicate a tale servizio (v. Art. 1.1.5). L'installazione deve garantire almeno le performance richieste nella sezione dedicata allo storage (v. Art. 1.3).

L'accesso utente ai nodi di front-end (Art. 1.1.4) sarà garantito dalla disponibilità del servizio ssh, che deve essere integrato nel sistema di autenticazione Kerberos 5 di ENEA-GRID, inclusa la compatibilità con OpenAFS.

L'accesso a tali nodi di front-end dovrà anche essere possibile attraverso il portale ENEA-GRID, basato su Citrix/Java. Sarà compito del fornitore supportare l'ENEA nella risoluzione di eventuali problemi legati all'implementazione del servizio.

Art. 1.4.4 Software di monitoring

Il sistema dovrà essere dotato di un software di monitoring in grado di controllare anche da remoto le funzionalità di tutte le sue componenti. La ditta dovrà specificare le caratteristiche salienti del sistema offerto ed in particolare la sua integrazione con le componenti di gestione.

Art. 1.4.5 Ambiente software per lo sviluppo di calcolo parallelo

Il sistema dovrà essere dotato dei compilatori Intel, Portland e GNU per le piattaforme delle Sezioni 1 e 2 e di compilatori ed ambienti di sviluppo specifici per i nodi della sezione 3 e per le schede FPGA.

Dovranno essere fornite almeno le librerie parallele matematiche associate ai suddetti compilatori. La quantità e qualità di librerie aggiuntive fornite costituirà elemento di merito preferenziale per la valutazione dell'offerta.

Dovranno essere fornite le librerie di supporto del servizio MPI nelle sue diverse opzioni, per tutti i compilatori e per le reti di interconnessione disponibili (IB, GigaEthernet).

Dovrà essere fornito almeno un ambiente di sviluppo, debugging ed ottimizzazione di codici paralleli.

Art. 1.4.6 Software applicativo

Sul sistema dovrà essere installato, in modalità ENEA-GRID, almeno il seguente software applicativo di terze parti:

- Fluent;
- Abaqus;
- ANSYS;
- MCNP;
- MCNPX;
- CPMD;
- GROMACS
- IDL;
- AVSexpress.

A tale proposito, si precisa che la modalità di installazione ENEA-GRID si basa su:

- installazione del software sul file system distribuito OpenAFS condiviso da tutti i nodi, in un apposito spazio AFS predisposto da ENEA;
- configurazione di una appropriata procedura autoconsistente di esecuzione del software, con particolare attenzione alla gestione delle licenze e alle modalità di esecuzione in ambiente parallelo;
- procedura di sottomissione in ambiente LSF.

Per tutti i software saranno provate le procedure di sottomissione e per i software CPMD, FLUENT e MCNP saranno effettuate le prove di funzionalità e scalabilità riportate nell'Allegato 4 che costituisce parte integrante del collaudo finale.

Art. 1.5 Sistema di backup



Il sistema HPC dovrà essere dotato di un sistema di backup basato sul software Veritas. Un sistema basato su tale software è già operante presso il centro ENEA di Portici. Il nuovo sistema dovrà integrarsi con il sistema già esistente e sarà cura del fornitore acquisire, nel corso del sopralluogo, tutte le informazioni necessarie per progettare e realizzare l'integrazione. Il fornitore dovrà inoltre integrare le licenze disponibili in ENEA con quelle necessarie per il sistema fornito. In particolare si sottolinea la necessità di effettuare il backup di AFS e GPFS. Il sistema dovrà avere le componenti minime come descritto nei paragrafi successivi.

Art. 1.5.1 Libreria nastri

La libreria dovrà al minimo:

- comprendere una fornitura iniziale di nastri per almeno 100 TByte;
- essere espandibile con la sola aggiunta di nastri fino a 300 TByte;
- essere realizzato in tecnologia LTO4;
- avere almeno 4 driver con tecnologia Fibra Ottica (FC) a 4Gbit/s;
- essere predisposto per un alloggiare almeno 12 driver;
- se necessario, includere eventuali licenze aggiuntive Veritas oltre a quelle in dotazione in ENEA;
- essere contenuto all'interno di un singolo rack 42U 19".

Art. 1.5.2 Server di Backup Master

Il Server deve essere provvisto di almeno 2 processori Intel/AMD dual core, con minimo: 2 GBytes di RAM, 4 connessioni di rete Gbit, 3 slot PCI, un HBA dual port FC4, N° 2 HDD da 146GB (RAID1), OS Linux compatibile con Veritas Netbackup 6.X

Art. 1.5.3 Media Server

I Media Server, che sovrintendono alle funzioni di trasferimento dei dati dallo storage ai tape drive, **sono 2**. Tali media server devono essere corredati di 4 porte FC4, dei collegamenti necessari ad accedere al sistema a dischi indicato nell'Art. 1.3 e sostenere un throughput complessivo verso i tape drive di almeno 300MB/s.

Art. 1.6 Sistema di sicurezza

Il fornitore dovrà configurare la rete privata del sistema di supercalcolo in modo da consentire un livello di sicurezza minimo tipico di un Centro di calcolo per attività di ricerca e compatibile con le regole di sicurezza della rete ENEA. Il fornitore collaborerà con ENEA per la messa a punto dei sistemi di sicurezza. Non è compito del fornitore la perimetrazione della LAN e la gestione del firewall del Campus di Portici.

Art. 1.7 Sistema di alimentazione e di continuità

La sala calcolo è servita con un armadio elettrico a doppia alimentazione trifase da 400V, una da 300 KVA per il sistema di calcolo ed una da 100 KVA per il sistema di condizionamento. Il posizionamento degli armadi è illustrato in Allegato 1, Tavola 7.

Partendo dalla situazione installata ed eventualmente verificata con sopralluoghi presso il centro di Portici, si dovrà fornire:

- un sistema di continuità in grado di garantire almeno 20 minuti di funzionamento a pieno carico dell'intero sistema.
- un sistema di spegnimento automatico, al mancare dell'alimentazione, a protezione delle risorse critiche nel tempo di funzionamento del sistema di continuità

Le caratteristiche minime del sistema di continuità sono riportate in Allegato 2.

Art. 1.8 Adeguamento sistema condizionamento

La sala calcolo del Centro di Portici è già dotata di un sistema di condizionamento predisposto per soddisfare le esigenze previste per il supercalcolatore. Essendo il dimensionamento di tale sistema stato realizzato in assenza delle caratteristiche di dettaglio del supercalcolatore, sarà cura della ditta fornitrice verificarne l'efficacia ed eventualmente integrarlo.

Il sistema attuale è costituito da quattro gruppi refrigeranti, le cui caratteristiche e posizionamento sono riportate in Allegato 1, Tavola 5. L'impianto è predisposto per ricevere l'installazione di un quinto gruppo (linee liquidi ed elettriche). Costituirà elemento di merito preferenziale la presentazione di un progetto riguardante la fluidodinamica termica della sala. Nel corso del sopralluogo presso il Centro di Portici la ditta fornitrice avrà modo di raccogliere tutte le informazioni che riterrà necessarie per un corretto dimensionamento del sistema.

Art. 1.9 Supporto on site

Il fornitore dovrà garantire, dopo il collaudo con esito positivo, un'attività on site, di almeno 80 giorni, per la formazione del personale ENEA. Tale attività avrà una cadenza tipica di 2 giorni/settimana. Dovrà inoltre essere fornito un supporto remoto per la gestione straordinaria finalizzata all'utenza.

Questi supporti si aggiungono a quelli di garanzia, legati ad anomalie o guasti del sistema.

Art. 1.10 Documentazione richiesta

In risposta al presente Bando, il fornitore dovrà presentare un progetto esecutivo di dettaglio del sistema offerto sulla base delle specifiche più sopra riportate.

Il progetto dovrà contenere almeno:

- lo schema a blocchi della realizzazione dell'intero sistema;
- il programma temporale per la realizzazione del progetto;
- il layout fisico dei componenti e la loro collocazione nella sala di calcolo di Portici e/o relative aree di servizio;
- l'elenco di tutti i componenti principali hardware e le loro caratteristiche tecniche;
- l'elenco di tutti i componenti principali software e le loro caratteristiche tecniche;
- il progetto preliminare delle Sezioni 1, Sezione 2, Sezione 3 più sopra citate;
- il progetto preliminare dell'infrastruttura di rete complessiva che comprende la cablatura InfiniBand, la SAN, e la connessione alla LAN;

- il progetto preliminare e le caratteristiche tecniche del sistema di alimentazione e dei gruppi di continuità;
- il progetto preliminare del sistema di storage ad alte prestazioni;
- il progetto preliminare del sistema di backup;
- il progetto preliminare di adeguamento del sistema di condizionamento;
- la descrizione hardware e software del sistema di installazione, configurazione monitoraggio e controllo;
- i risultati dei benchmark: HPC Challenge Benchmarks, Versione 1.0.0 come disponibile presso <http://icl.cs.utk.edu/hpcc>; in particolare si richiedono i risultati dei test riportati nel seguito (con le unità di misura) ognuno di essi dovrà essere rispettato in fase di collaudo del sistema, Art.2.4; l'insieme dei dati forniti costituirà elemento di merito preferenziale nella valutazione:
 - Benchmark G-HPL (TFlop/s);
 - Benchmark G-PTRANS (GB/s)
 - Benchmark G-FFTE (GFlop/s)
 - Benchmark G-Random Access (Gup/s)
 - Benchmark G-STREAM Triad (GB/s);
 - Benchmark EP-STREAM Triad (GB/s);
 - Benchmark EP-DGEMM Triad (Gflop/s);
 - Benchmark Random Ring Bandwidth (GB/s)
 - Benchmark Random Ring Latency (usec);

Il prefisso G indica il test globale che coinvolge la comunicazione tra processi su una intera sezione mentre il prefisso EP indica il caso "Embarassing Parallel".

I test devono riferirsi a 3 differenti configurazioni: test individuali delle sezioni 1 e 2 e test del sistema complessivo formato dall'insieme delle sezioni 1 e 2, per un totale di 9 test per 3 configurazioni. I test devono essere effettuati con almeno un processo per ogni core.

- l'elenco dei software applicativi certificati per la Sezione 1 in Gbit Ethernet e per la Sezione 2 in InfiniBand;
- la potenza elettrica necessaria al funzionamento a pieno carico per l'intero sistema escluso il sistema di condizionamento;
- il profilo della potenza elettrica necessaria in funzione del carico sui nodi di calcolo;
- tutta la documentazione necessaria a valutare gli elementi di merito preferenziale espressi nella Specifica Tecnica e/o degli eventuali elementi migliorativi proposti.

Ogni caratteristica dichiarata dall'Impresa per cui sia prevista l'attribuzione di un punteggio costituisce obbligazione contrattuale.

2 Specifiche di collaudo

Nel presente capitolo si riportano i contenuti tecnici del collaudo, mentre le regole contrattuali e penali sono riportate nell'apposito articolo del modello di contratto di fornitura.

Il collaudo deve consentire al collaudatore ENEA e suoi assistenti di poter verificare tutte le dichiarazioni effettuate dal fornitore in fase di offerta. Sarà cura del fornitore, ove necessario, implementare delle procedure che consentano di verificare le dichiarazioni di performance dichiarate in fase di offerta.

In particolare, il collaudo si articolerà come specificato nel seguito.



Art. 2.1 Verifica formale della fornitura

Si verificherà che tutto il materiale fornito sia conforme all'ordine di fornitura e sue eventuali variazioni.

Art. 2.2 Collaudo hardware delle apparecchiature

In questa fase saranno verificate tutte le funzionalità hardware e di sicurezza:

- sistema di alimentazione;
- sistema di condizionamento;
- sistema di protezione e spegnimento in caso di caduta dell'energia elettrica e conseguente entrata in servizio del gruppo di continuità.
- eventuali sostituzioni a caldo di componenti, ove dichiarato;
- continuità di funzionamento nelle condizioni di massimo carico per 72 h.

Art. 2.3 Collaudo software di sistema

In questa fase sarà verificata la corretta configurazione come richiesto nelle specifiche del software di sistema. Verranno verificate tutte le funzionalità dichiarate per i prodotti di configurazione, monitoraggio e gestione, con particolare riguardo al sistema di installazione e configurazione automatica da rete.

Art. 2.4 Collaudo software applicativo

Il collaudo dei software applicativi sarà organizzato in due parti, una dedicata ai benchmark HPC Challenge Benchmarks standard, ed una dedicata alla verifica del corretto funzionamento di alcuni software commerciali installati.

Per quanto riguarda i test HPC, Art. 1.10, si richiede il rispetto di tutti i dati di prestazione specificati in sede di offerta. L'esito positivo del collaudo è condizionato dal rispetto dei dati di prestazione dichiarati in fase di offerta. In caso di mancato rispetto di qualunque dei dati dichiarati il fornitore dovrà effettuare tutte le modifiche hardware e software necessarie al rispetto di tutti i parametri.

Per quanto concerne i software applicativi che il fornitore deve installare, di cui all'Art. 1.4.6, ai fini del collaudo dovranno essere eseguiti i benchmark riportati in Allegato 4 ai fini di verificarne il corretto funzionamento, sul sistema fornito. I benchmark si riferiscono ai codici FLUENT, CPMD, MCNP, GROMACS.

3 Allegati

Gli allegati costituiscono parte integrante della specifica tecnica. Il loro scopo è solo quello di dare indicazioni puntuali su alcuni aspetti specifici che si ritengono importanti per la formulazione dell'offerta. Al fine di fornire un quadro riassuntivo si riporta l'elenco degli allegati con una sintesi del contenuto relativo.

Allegati	Descrizione
Allegato 1	Piantine Locali
Allegato 2	Caratteristiche minime del Gruppo di continuità
Allegato 3	Modello di attestato di avvenuto sopralluogo
Allegato 4	Test Funzionali e di scalabilità

