



ALLEGATO TECNICO RELATIVO ALLA PROCEDURA DI GARA A LOTTI PER LA FORNITURA DI UN SISTEMA DI STORAGE DISCO E RELATIVA MANUTENZIONE PER IL TIER-1 DELL'INFN

1. Premessa

L'INFN (nel seguito indicato con il termine "committente") invita ditte (nel seguito indicate con il termine "operatore economico") a presentare offerte per la fornitura, installazione e manutenzione di una quantità di storage, suddivisa in due lotti (di tecnologia diversa), di capacità complessiva di almeno 14700 TeraByte , server, switch di management e di un sistema SAN (descritto nel paragrafo 3.3). Detto storage dovrà essere installato presso il data center Tier-1 dell'INFN presso il CNAF.

Sono da intendersi inclusi nella fornitura i seguenti servizi accessori: installazione, messa in funzione dell'intero sistema su rack già presenti nelle sale macchine oggetto dell'installazione (si rimanda al Par. 4 per i dettagli), il cablaggio completo delle alimentazioni e di tutti i collegamenti necessari al funzionamento del sistema e lo smaltimento completo dei materiali di risulta (imballaggi e scatole) da effettuarsi congiuntamente alla fase di installazione.

Lo smaltimento dei materiali di risulta dovrà essere effettuato nel rispetto della normativa vigente: qualora si osservassero modalità palesemente irregolari saranno segnalate alle autorità competenti.

Dovrà inoltre essere incluso un periodo di garanzia e manutenzione on site di almeno 5 anni su tutto il materiale, il cui inizio avverrà a far data dal pagamento dopo l'emissione del certificato di collaudo positivo da parte del personale incaricato del committente nelle modalità descritte successivamente.

Il presente Allegato Tecnico disciplina gli aspetti tecnici della gara.

2. DESCRIZIONE GENERALE

2.1. Glossario delle abbreviazioni

- DWPD - Drive Write Per Day
- FC – Fibre Channel
- GPFS – Spectrum Scale (anche General Parallel File System) di IBM
- IB – InfiniBand
- LAN – Local Area Network
- NVMe - Non-Volatile Memory Express

- NVMe-oF - Non-Volatile Memory Express over Fabrics
- PB – PetaByte (1 PetaByte è pari a 10^{15} Byte)
- PMR – Perpendicular Magnetic Recording
- PDU – Power Distribution Unit o presiere per alimentare gli apparati.
- RHEL – RedHat Enterprise Linux
- RU – Rack Unit (unità rack)
- SAN – Storage Area Network
- SSD - Solid State Drive
- TB – TeraByte (1 TeraByte è pari a 10^{12} Byte)
- TB-N – Per TeraByte netto si intende la capacità utilizzabile al netto della parità e della formattazione contrapposto al TeraByte raw o, più semplicemente, TB.

2.2. Glossario delle definizioni

- Entità logica - sottosistema completamente gestibile in maniera autonoma, costituito, nelle sue parti essenziali, da un'unità di controllo con almeno due controller RAID ridondati ("active - active") e dischi magnetici. All'interno di ogni entità logica ogni controller dovrà essere in grado di gestire tutti i dischi.
- NSD server – disk-server per lotto "GPFS".

2.3. Sede INFN oggetto della fornitura

La fornitura interessa il Tier-1 dell'INFN ospitato nel data center del CNAF. Nella tabella seguente sono riportati l'indirizzo ed il riferimento tecnico.

Sede	Indirizzo	Persona di contatto
CNAF	v.le Bertini Pichat 6/2 - Bologna (per la consegna usare l'accesso carrabile da v. Ranzani 3).	Vladimir Sapunenko (vladimir.sapunenko@cnafe.infn.it), tel. 051 2095450

Per un sopralluogo e/o per la consegna, contattare le persone di riferimento.

Per la modalità di consegna si veda il capitolo 4 del presente Allegato Tecnico.

2.4. Lotti

La fornitura o si articola nei seguenti lotti:

- Lotto "GPFS" – 1 sistema storage da almeno 10700 TB (corrispondenti ad almeno 8732 TB-N)
- Lotto "CEPH" - 1 sistema storage da almeno 4000 TB

2.5. Requisiti tecnici

Gli operatori economici devono possedere la seguente certificazione, **pena l'esclusione**:

- certificazione **ISO 9001:2008** (o 2015);

Inoltre, tutti gli apparati elettrici dovranno essere forniti di certificazione CE. L'operatore economico dovrà fornire prova scritta di rispondenza CE delle apparecchiature proposte in fase di gara.

2.6. Requisiti di sostenibilità

Le apparecchiature fornite dovranno rispettare i seguenti requisiti:

- I requisiti stabiliti nel D. Lgs. n. 81/2008;
- I requisiti di ergonomia stabiliti nella Direttiva CEE 90/270 recepita dalla legislazione italiana con Legge 19 febbraio 1992, n. 142;
- I requisiti di immunità definiti dalla EN55024;
- I requisiti relativi alla restrizione all'uso di sostanze pericolose previsto dalla normativa vigente, ed in particolare dalla direttiva 2011/65/UE, (RoHS II), recepita con D.lgs. 4 marzo 2014, n. 27.
- Per quanto concerne i rifiuti di apparecchiature elettriche ed elettroniche, RAEE, il Fornitore dovrà garantire i requisiti di conformità secondo quanto previsto dal D. Lgs. 14 marzo 2014, n. 49, dal D. Lgs 152/2006 e ss. m.i., dal D.M. 8 marzo 2010, n. 65.

3. DESCRIZIONE TECNICA

Le specifiche tecniche minime riportate nel presente documento, in particolare nel presente paragrafo 3, salvo dove sia espressamente affermato il contrario, devono essere tutte soddisfatte a pena l'esclusione.

Dovranno essere obbligatoriamente indicati la marca ed il modello di tutti gli apparati offerti.

Lotto GPFS

I sistemi disco oggetto del presente lotto saranno composti dai seguenti elementi principali:

- Entità logiche - Le entità logiche dovranno essere omogenee tra loro per caratteristiche e tipologia. Si precisa che dovranno essere fornite almeno due entità logiche.
- NSD Server – Ogni server dovrà essere equipaggiato con due porte Ethernet (2 schede o una bicanale) per il collegamento al core switch e con una scheda bicanale FC o IB per il collegamento alla Storage Area Network (SAN), fornita insieme al sistema disco. Per le caratteristiche complete si veda il par. 3.2.1;

- Un'infrastruttura di Storage Area Network (SAN) in tecnologia IB o FC fra il sistema disco ed i server (par. 3.3.1);
- Le PDU necessarie per l'installazione (par. 3.4);
- Un unico software di configurazione e allarmistica per tutti i sistemi disco;
- Servizi di installazione, configurazione, validazione e manutenzione (assistenza tecnica) on-site per almeno 5 anni.

Lotto CEPH

I sistemi disco oggetto del presente lotto saranno composti dai seguenti elementi principali:

- Almeno 12 disk-server interconnessi tramite switch Ethernet; ogni server dovrà essere equipaggiato con almeno due porte Ethernet, controller RAID e dischi per i dati (per le caratteristiche di tali server si veda il par. 3.2.1);
- 2 Switch Ethernet di collegamento tra i server e con la LAN (si veda par. 3.3.2);
- Le PDU necessarie per l'installazione (par. 3.4);
- Servizi di installazione, configurazione, validazione e manutenzione (assistenza tecnica) on-site per almeno 5 anni.

Prestazioni richieste

Lotto GPFS

Per il presente lotto (ad eccezione del sistema per i metadati, si veda par. 3.6) verranno richiesti i seguenti valori prestazionali complessivi (verificati in fase di collaudo):

1. **Flusso aggregato complessivo sustained per operazioni sia di Write che di Read sequenziale**, espresso in MBytes/s, misurato secondo le specifiche del protocollo di collaudo descritto successivamente.
2. **Flusso aggregato complessivo su operazioni di tipo random Read**, espresso in MBytes/s ovvero **Numero di IOPS utilizzando un record size di 1024k** misurato secondo le specifiche del protocollo di collaudo descritto successivamente.

Le prestazioni seguenti dovranno essere soddisfatte con la configurazione RAID proposta (si veda paragrafo 3.1).

I valori complessivi verranno ricavati dai seguenti valori unitari (qui indicati per TB netto):

1. **3.5 MBytes/s sustained** per operazioni sia di Write che di Read sequenziale per ogni TB netto di disco;
2. **1.75 MBytes/s sustained** per operazioni contemporanee di Write e di Read sequenziale (50% Read, 50% Write) per ogni TB netto di disco;
3. **1 IOPS/TB-N** per operazioni di tipo random Read utilizzando un record size di 1024k

Lotto CEPH

Verranno richiesti i seguenti valori prestazionali complessivi (verificati in fase di collaudo):

1. **Flusso aggregato complessivo sustained per operazioni sia di Write che di Read sequenziale**, espresso in MBytes/s, misurato secondo le specifiche del protocollo di collaudo descritto successivamente.

Le prestazioni richieste per il singolo server sono le seguenti:

1. **1.1 GBytes/s sustained** per operazioni sia di Write che di Read sequenziale;
2. **1.1 GBytes/s sustained** per operazioni contemporanee di Write e di Read sequenziale (50% Read, 50% Write)

3.1. Specifiche generali del sistema disco

Nel presente paragrafo, il sotto-paragrafo 3.1.1 illustra le caratteristiche relative al "lotto GPFS", mentre il sotto-paragrafo 3.1.2 illustra le caratteristiche relative al "lotto CEPH".

3.1.1. Lotto GPFS

I sistemi disco offerti dovranno essere omogenei per caratteristiche e configurazione.

Per tutte le unità componenti il sistema dovrà essere fornita la ridondanza degli alimentatori, dei controller e delle unità di ventilazione oltre che delle altre parti essenziali come batterie di cache ed eventuali moduli di controllo con sostituzione hot-swap (sostituzione a caldo senza fermo macchina).

Inoltre, almeno 40 TB della capacità offerta, dovrà essere composta da dischi (detti per metadati) in tecnologia a stato solido (SSD) o NVMe con le seguenti caratteristiche:

- Parametro di scrittura su drive per giorno (DWPD) pari almeno a 3
- Capacità minima del singolo disco: 1,5 TB
- Possibilità di installazione e sostituzione in modalità hot-swap senza interruzione di servizio

Entità logiche

Ogni entità logica offerta dovrà avere una capacità minima installata di 2000 TB-N ed essere dotata di almeno due controller active-active ognuno in grado di gestire tutti i dischi presenti nel sistema in caso di non disponibilità dell'altro.

La configurazione di ogni singola entità logica dovrà prevedere controller Raid ridondati, di tipo hot-swap, con cache che potrà essere di una delle seguenti tipologie:

- coerente (ovvero possibilità di mirror della cache su entrambi i controller)
- con flash backend o dischi SSD
- di tipo NVRAM condiviso

La dimensione della cache, per ogni singola entità logica, dovrà essere:

- Pari ad almeno 8 GBytes per entità logiche con capacità fino a 4 PB netti
- Pari ad almeno 16 GBytes per entità logiche con capacità superiore a 4 PB netti

La ridondanza dei controller e la riassegnazione dei volumi logici (LUN), in caso di fallimento di uno dei controller (failover), dovrà essere pienamente compatibile con il sistema operativo RHEL 7.x a 64 bit (e versioni successive attualmente disponibili) ed il relativo tool di multipath.

La configurazione di failover sarà accuratamente testata durante la fase di collaudo descritta successivamente e **costituirà valido motivo per la non accettazione della fornitura qualora la stessa si dimostrasse non rispondente ai requisiti di cui sopra.**

Le entità logiche dovranno essere provviste di un opportuno sistema di autospegnimento (thermal shutdown) in caso di aumento improvviso della temperatura esterna per evitare danni alle parti sensibili di tali apparati e preservare la consistenza dei dati.

Le entità logiche dovranno essere altresì fornite di un sistema altamente configurabile di controllo in background su tutti i volumi dello stato dei Raid e della verifica preventiva e conseguenti azioni di ripristino di settori fisici e/o logici danneggiati.

Il numero minimo di connessioni **attive** in ingresso/uscita **da ogni controller** dovrà essere tale che la banda passante sia almeno pari alla banda richiesta per ogni sistema e comunque non inferiore a 2 per motivi di ridondanza:

$$N_{\text{porte}} = \max(2, \text{Ceil}(V_{\text{TB-N}} * \text{TB-N}_{\text{sistema}} / V_{\text{porta}}))$$

dove:

- $V_{\text{TB-N}}$ indica la velocità "sustained" minima richiesta per TB-N, come definita nel paragrafo precedente, ed espressa in MB/s/TB-N;
- $\text{TB-N}_{\text{sistema}}$ indica la capacità netta di un sistema offerto ottenuta dividendo la capacità netta richiesta per una data sede per il numero di sistemi disco offerti nella stessa sede;
- V_{porta} indica la velocità nominale di una porta FC o IB del sistema offerto espressa in MB/s;
- **Ceil** è la funzione che effettua l'arrotondamento, per eccesso (all'unità superiore), dell'argomento.

Es.: si consideri una fornitura di 10 PB netti di storage: in questo caso la banda totale richiesta sarà di almeno 35 GB/s (ovvero 10,000 TB-N x 3.5 MB/s/TB-N).

*Supponendo che la soluzione proposta preveda 4 sistemi da 2,5 PB netti ciascuno, la banda passante per ogni sistema dovrà essere di almeno 8.75 GB/s. Se ognuno dei sistemi è dotato di 2 controller in tecnologia IB o Ethernet a 100 Gbps, il numero di connessioni di ogni controller dovrà essere almeno pari a 2 (8.75 GB/s * 8 / 100 Gbps arrotondato per eccesso ad 1 e quindi portando a 2 per ridondanza).*

Tipologia dischi

I dischi offerti, ad eccezione fatta per la parte da fornire in tecnologia SSD o NVMe, dovranno essere in tecnologia PMR, "Helium filled" e comunque di classe "Enterprise" in grado di mantenere gli standard di alta disponibilità 24/7.

Le caratteristiche dei dischi offerti dovranno essere omogenee.

La capacità del singolo disco non dovrà essere inferiore a 10 TB.

¹ Si ricordi il fattore 8 di conversione tra Mbit e MB. Ad es. 100 Gbps = 12.5 GB/s

Tipologia RAID

Tutti i dischi offerti per ogni sistema disco dovranno essere raggruppabili in almeno 2 disk-pool (o comunque in un numero pari di disk-pool secondo le raccomandazioni del produttore); ogni disk-pool dovrà essere obbligatoriamente configurato con RAID distribuito (es. tramite tecnologia "Dynamic Disk Pool" o Raid 2.0+ o equivalente) con protezione di tipo 8+2 dischi (comunque con un rapporto fra capacità netta e lorda non superiore a 0.80). Inoltre, lo spazio per la "reserved capacity" dovrà essere almeno equivalente al doppio della capacità di un singolo disco fornito (es. se l'offerta prevedesse dischi da 14 TB, la "reserved capacity" dovrebbe essere almeno pari a 28 TB).

La velocità di ricostruzione di un disco dovrà essere almeno pari a 1 TB per ora (es. quindi per un disco da 12 TB, il tempo totale di ricostruzione non potrà essere superiore a 12 ore).

Footprint/occupazione

Nel computo dell'occupazione dovranno essere considerati tutti gli apparati di storage compresi di controller ed eventuali cassette di espansione. La densità minima, espressa in dischi/RU, non dovrà essere la seguente inferiore a 15 dischi/RU.

Controllo

Il sistema dovrà essere fornito di un software completo di controllo e configurazione che operi sia a livello di singola entità logica che di sistema globale. Il software dovrà includere, **pena l'esclusione**, la gestione dell'allarmistica tramite invio di messaggi di tipo e-mail. Le condizioni di allarme segnalate dovranno almeno comprendere: aumenti della temperatura, fallimento di un controller ridondato, di una unità di alimentazione/ventilazione, fallimento di un disco di hot-spare ed in particolare il fallimento di un singolo disco costituente un array Raid e lo status della relativa ricostruzione automatica dell'array. Tale software dovrà essere installabile su piattaforma Windows o Linux o dovrà essere accessibile via Web da un Browser standard. Il software dovrà essere in grado di gestire e creare array Raid in maniera semplice e ben documentata e dovrà poter permettere il monitoraggio in tempo reale dei parametri suddetti di ogni singola entità logica. Dovrà essere in ogni caso possibile la gestione di tutto il sistema (e quindi di tutte le entità logiche) da un singolo applicativo su un'unica stazione di controllo. Dovrà inoltre essere possibile a livello di singola entità logica disporre di un tool dettagliato di analisi delle performance disco in tempo reale e con la possibilità di memorizzare lo storico.

Parametri da specificare nell'offerta

Dovranno essere obbligatoriamente specificate, oltre alle caratteristiche generali del sistema offerto, le seguenti informazioni (utilizzate in fase di collaudo):

1. **Tipologia dei dischi offerti.** Dovranno essere indicati tecnologia e dimensione dei dischi offerti.
2. **Caratteristiche dei controller**
 - Cache fornita per ogni singolo controller e sistema di protezione della cache stessa.

- Numero, tipologia e velocità (espressa in Gbps) delle connessioni attive verso i dischi da ogni entità logica offerta.
 - Numero, tipologia (IB o FC) e velocità (espressa in Gbps) delle connessioni attive verso la SAN da ogni entità logica offerta.
 - Tipologia di Raid distribuito implementato.
 - Banda passante nominale dei controller espressa sia in numero di operazioni di I/O al secondo (IOPS) che in MByte/s per entità logica.
 - Tempi di ricostruzione di un raid set con la configurazione di RAID usata per ottenere i valori di throughput di cui al punto precedente.
3. **Capacità lorda del sistema**, calcolata moltiplicando la capacità nominale in GByte dei singoli dischi per il numero totale di dischi contenuti nel sistema (inclusi gli eventuali hot-spare).
 4. **Numero consigliato di dischi hot-spare** (si consideri anche lo spazio riservato per ricostruzioni), facenti parte della fornitura, necessari per ogni singola entità logica, supponendo la configurazione con RAID distribuito (con protezione 8+2), descritta precedentemente, sulla totalità dei dischi.
 5. **Capacità netta del sistema disco offerto**. La capacità netta dovrà essere indicata in TB. La capacità netta si riferisce all'effettivo spazio di storage disponibile con il tipo di RAID proposto e nelle condizioni precisate nel protocollo di collaudo ottenute tramite il comando del Sistema Operativo RHEL 7 "df -B TB".
 6. **Caratteristiche fisiche del sistema offerto**:
 - Dimensioni fisiche, peso ed occupazione complessiva in termini di RU (nel computo dovranno essere compresi apparati di storage con cassette di espansione, server, apparati della SAN, qualora previsti, e PDU.
 - Occupazione sistema offerto espressa TB-N per Rack Unit.
 - Consumi e requisiti in termini di condizionamento.

3.1.2. Lotto CEPH

I sistemi disco offerti dovranno essere omogenei per caratteristiche e configurazione.

Per tutte le unità componenti il sistema dovrà essere fornita la ridondanza degli alimentatori e delle unità di ventilazione oltre che delle altre parti essenziali come eventuali moduli di controllo con sostituzione hot-swap (sostituzione a caldo senza fermo macchina).

Server

Ogni server offerto dovrà avere, oltre ai 2 dischi di sistema, 24 dischi di espansione (detti dischi per dati).

Tipologia dischi per dati

I dischi offerti dovranno essere in tecnologia PMR, "Helium filled" e comunque di classe "Enterprise" in grado di mantenere gli standard di alta disponibilità 24/7.

Le caratteristiche dei dischi offerti dovranno essere omogenee all'interno della fornitura.

La capacità del singolo disco non dovrà essere inferiore a 10 TB.

Footprint/occupazione

Nel computo dell'occupazione dovranno essere considerati sia i server che eventuali cassette di espansione. La densità minima, espressa in dischi/RU, non dovrà essere inferiore a 12 dischi/RU.

Parametri da specificare nell'offerta

Dovranno essere obbligatoriamente specificate, oltre alle caratteristiche generali del sistema offerto, le seguenti informazioni (utilizzate in fase di collaudo):

1. Caratteristiche server

- Numero dischi per dati offerti per server (oltre ai 2 dischi di sistema)
- Numero, tipologia e velocità (espressa in Gbps) delle connessioni attive verso i dischi
- Banda passante nominale dei server espressa sia in numero di operazioni di I/O al secondo (IOPS) che in MByte/s.

2. Numero server offerti

3. Tipologia dei dischi offerti. Dovranno essere indicati tecnologia e dimensione dei dischi offerti.

4. Capacità lorda del sistema, calcolata moltiplicando la capacità nominale in GByte dei singoli dischi per il numero totale di dischi contenuti nel sistema.

5. Caratteristiche fisiche del sistema offerto:

- Dimensioni fisiche, peso ed occupazione complessiva in termini di RU (nel computo dovranno essere compresi i server con eventuali cassette di espansione, server, switch di rete e PDU.
- Consumi e requisiti in termini di condizionamento.

3.2. Specifiche SERVER

Nel presente paragrafo, il sotto-paragrafo 3.2.1 illustra le caratteristiche relative al "lotto GPFS", mentre il sotto-paragrafo 3.2.2 illustra le caratteristiche relative al "lotto CEPH".

3.2.1. Lotto GPFS

I server **fanno parte integralmente della presente richiesta di fornitura:** dovrà quindi valere anche per loro la garanzia e manutenzione in maniera analoga all'intero sistema oggetto della fornitura. Sono applicabili inoltre le stesse penali descritte per le condizioni in garanzia della fornitura.

Tutti i server dovranno possedere le caratteristiche minime elencate di seguito:

- Piena compatibilità con il Sistema Operativo RHEL 7.x a 64 bit (e versioni successive attualmente disponibili).
- Piena compatibilità con il software IBM GPFS v. 4.2 (e versioni successive attualmente disponibili); dovrà essere possibile esportare con protocollo GPFS la totalità dello spazio disco costituito dalle entità logiche.
- Almeno 24 core fisici per motherboard con un numero minimo di 3 canali indipendenti PCI Express, di tipologia almeno v. 3.
- Almeno 128 GByte di memoria del tipo DDR4 (o equivalente o comunque superiore), compatibile con il processore offerto, eventualmente con possibilità di espansione. I moduli di memoria presenti dovranno essere tutti uguali (per dimensione e tipologia). Tutti i canali di memoria presenti nel processore dovranno essere utilizzati: dovrà essere quindi prevista una dimensione dei banchi di memoria adeguata allo scopo. Specificare il valore di memory bandwidth sul sistema proposto e tutte le caratteristiche rilevanti della memoria offerta.
- 2 Dischi SSD di capacità almeno 250 Gbyte l'uno con possibilità di sostituzione hot-swap, con un throughput minimo garantito di almeno 80Mbyte/s con sistema Linux.
- Sottosistema Raid integrato per i dischi di sistema con possibilità di almeno Raid 1 (mirroring) hardware sui 2 dischi descritti nel punto precedente.
- 2 Interfacce Ethernet autosensing 100/1000 Mbps on board con uscita in rame RJ45
- Minimo 2 SLOT PCI Express v.3 x16 (o superiore) indipendenti (non sulla stessa raiser card). In ogni caso gli slot PCI dei server offerti dovranno essere pienamente compatibili con la scheda IB o FC di interconnessione al sistema disco e con le interfacce Ethernet a 25 Gb/s o 100 Gb/s descritte successivamente e dovranno permetterne l'utilizzo alla massima velocità nominale.
- Alimentazione ridondata con possibilità di sostituzione hot-swap degli alimentatori (sostituzione a caldo senza interruzione dell'attività del server). Gli alimentatori dovranno supportare i requisiti specificati dal costruttore della scheda madre e quelli specificati per tutti i componenti interni al case.
- Interfacce VGA, Keyboard e Mouse standard o adattatori per connessione standard.
- Montaggio a rack: non saranno in ogni caso considerati server con occupazione maggiore di 2 U.
- Bootstrap via rete con protocollo PXE con sistema Linux e via USB.
- Le ventole, fatta eccezione per quelle delle CPU e degli alimentatori, dovranno essere ridondanti. L'aria calda potrà essere espulsa solamente attraverso la parte posteriore del disk-server.
- Guide ed accessori per montaggio a rack standard 19".
- Cavi di alimentazione con prese standard IEC di lunghezza idonea e cavi di rete standard UTP Cat. 6 di lunghezza idonea per cablaggi all'interno dello stesso rack standard.

- Montaggio, cablatura alimentazione e rete, installazione e validazione (verifica della corretta accensione dei server e delle parti hardware) nei rack già presenti presso il data center del CNAF (vedi Allegato A per maggiori dettagli tecnici a riguardo). La cablatura dell'alimentazione e della rete dovrà essere effettuata secondo lo schema fornito dall'acquirente al momento dell'installazione.
- Remote console e controllo accensione e spegnimento via ipmi.
- Possibilità di lettura stato dei principali componenti hardware tramite comando linux "ipmitool" e protocollo snmp. Minimo richiesto: temperature delle singole CPU, velocità e stato delle ventole di raffreddamento.
- Tool di monitoring e allarmistica a linea di comando in grado di visualizzare i vari parametri fisici del server (temperatura, stato degli alimentatori, velocità e stato delle ventole etc...) e della scheda Raid integrata (stato del Raid 1 sui 2 dischi di sistema) compatibile il sistema operativo specificato in precedenza. Tale tool dovrà essere installato e reso operativo a completo carico della ditta in sede di installazione delle macchine se richiesto al momento del collaudo.
- Configurazione del BIOS delle macchine in modo conforme a quanto specificato dall'acquirente al momento dell'ordine.
- Elenco, su supporto informatico opportuno, contenente, per ogni macchina, le seguenti informazioni: identificativo rack, posizione nel rack, numero seriale macchina, MAC address di tutte le interfacce di rete Ethernet.

Connettività server

I server per i dati dovranno anche avere 2 interfacce Ethernet in fibra complete di transceiver per il collegamento alla LAN ed una scheda bicanale (o due schede) di tecnologia FC o IB per il collegamento verso il sistema disco. Tutte le schede dovranno essere compatibili con le macchine offerte e il sistema operativo RHEL 7. Inoltre, le schede dovranno essere di tipologia PCI-Express v.3 o superiore.

Le porte dovranno essere aggregabili in modo da permettere un throughput complessivo pari al totale. La seguente tabella riassume le possibili combinazioni ammesse per i collegamenti dei server:

	SAN	
<i>Ethernet</i>	<i>Infiniband</i>	<i>Fibre Channel</i>
2x25 Gbps	2x56-2x100 Gbps	2x32 Gbps
2x100 Gbps	2x100 Gbps	NON AMMESSO

Si precisa che le interfacce Ethernet (LAN) verranno collegate in fibra direttamente ai due Switch di CORE Cisco NEXUS 9516 (su moduli N9K-X9732C-EX o N9K-X97160YC-EX). In sala calcolo è disponibile, per il collegamento dei server, un cablaggio duplex MMF 50/125: l'azienda dovrà

verificare la congruità di tale cablaggio con la soluzione proposta e, nel caso in cui non sia utilizzabile, sarà a carico dell'azienda fornire tutto il necessario per il collegamento in rete delle risorse proposte ossia i cavi di collegamento verso il centro stella della rete.

Inoltre, sono comunque a carico dell'azienda i transceiver da inserire negli switch di CORE per collegare i server; le interfacce di rete lato CORE saranno invece a carico del CNAF e saranno dotate di porte QSFP28. Qualunque tecnologia venga scelta per l'interconnessione, i transceiver per gli switch (o i cavi pre-intestati) dovranno essere ufficialmente supportati da Cisco.

Numero server

Dovrà essere fornito un numero minimo di server per i dati secondo le specifiche seguenti:

Interfacce LAN (2 interfacce per server)	# server per PB netto	Throughput minimo server per PB netto
25 Gb/s	2	3 GB/s
100 Gb/s	0.5	12 GB/s

Si intende che il numero ottenuto, se non intero e pari, dovrà essere arrotondato al numero pari superiore e comunque tale da garantire una banda passante pari ad almeno 1.5 il throughput totale richiesto. Verrà considerato come limite massimo di banda per ogni server la banda passante minima fra collegamento Ethernet e collegamento verso il sistema disco; in ogni caso tali valori verranno verificati in fase di collaudo.

3.2.2. Lotto CEPH

I server dovranno possedere le seguenti caratteristiche minime:

- Piena compatibilità con il Sistema Operativo RHEL 7.x a 64 bit
- Macchine biprocessori. I processori dovranno avere architettura x86_64, della famiglia Intel Xeon scalabili o AMD della famiglia EPYC, ognuno con le seguenti caratteristiche minime:
 - almeno 16 core fisici;
 - frequenza di clock di base pari ad almeno 2.1 GHz;
- Almeno 128 GByte di memoria, di tipo DDR4 (o equivalente o comunque superiore), compatibile con il processore offerto, eventualmente con possibilità di espansione. I moduli di memoria presenti dovranno essere tutti uguali (per dimensione e tipologia). Tutti i canali di memoria presenti nel processore dovranno essere utilizzati: dovrà essere quindi prevista una dimensione dei banchi di memoria adeguata allo scopo. Specificare il valore di memory bandwidth sul sistema proposto e tutte le caratteristiche rilevanti della memoria offerta.
- 2 Dischi magnetici di sistema, identici, di capacità di 1 TB l'uno con possibilità di sostituzione hot-swap, con un throughput minimo garantito di almeno 80Mbyte/s con sistema Linux, oppure due dischi SSD (NVMe o SAS),

- Sottosistema Raid integrato con possibilità di almeno Raid 1 (mirroring) hardware sui 2 dischi di sistema descritti nel punto precedente.
- 2 dischi SSD o NVMe da almeno 1.6 TB ciascuno
- Bus interno SAS da almeno 12 Gbps per collegamento ai dischi per dati
- 24 dischi per dati, da almeno 14 TB, di tipo hot swap (non è ammessa la tecnologia SMR): la configurazione deve essere tale da usare tutti i canali del processore in modo uniforme.
- 1 interfaccia LAN dedicate per IPMI con uscita in rame RJ45 e relativo cavo per collegamento allo switch di management (par. 3.5)
- Connettività LAN: quattro interfacce da 25Gbit e relativi cavi per collegamento agli switch descritti al par. 3.3
- Alimentazione ridondata con possibilità di sostituzione hot-swap degli alimentatori (sostituzione a caldo senza interruzione dell'attività del server). Gli alimentatori dovranno supportare i requisiti specificati dal costruttore della scheda madre e quelli specificati per tutti i componenti interni al case.
- Interfacce VGA, Keyboard e Mouse standard o adattatori per connessione standard.
- Montaggio a rack: non saranno in ogni caso considerati server con densità minore a 12 dischi per Rack Unit.
- Bootstrap via rete con protocollo PXE con sistema Linux e via USB.
- Le ventole, fatta eccezione per quelle delle CPU e degli alimentatori, dovranno essere ridondanti. L'aria calda potrà essere espulsa solamente attraverso la parte posteriore del disk-server.
- Guide ed accessori per montaggio a rack standard 19".
- Cavi di alimentazione con prese standard IEC di lunghezza idonea e cavi di rete standard UTP Cat. 6 di lunghezza idonea per cablaggi all'interno dello stesso rack standard.
- Montaggio, cablatura alimentazione e rete, installazione e validazione (verifica della corretta accensione dei server e delle parti hardware) nei rack già presenti presso il data center del CNAF (vedi Allegato A per maggiori dettagli tecnici a riguardo). La cablatura dell'alimentazione e della rete dovrà essere effettuata secondo lo schema fornito dall'acquirente al momento dell'installazione.
- Remote console e controllo accensione e spegnimento via ipmi.
- Possibilità di lettura stato dei principali componenti hardware tramite comando linux "ipmitool" e protocollo snmp. Minimo richiesto: temperature delle singole CPU, velocità e stato delle ventole di raffreddamento.
- Tool di monitoring e allarmistica a linea di comando in grado di visualizzare i vari parametri fisici del server (temperatura, stato degli alimentatori, velocità e stato delle ventole etc...) e della scheda Raid integrata (stato del Raid 1 sui 2 dischi di sistema) compatibile il sistema operativo specificato in precedenza. Tale tool dovrà essere installato e reso operativo a completo carico della ditta in sede di installazione delle macchine se richiesto al momento del collaudo.

- Configurazione del BIOS delle macchine in modo conforme a quanto specificato dall'acquirente al momento dell'ordine.
- Elenco, su supporto informatico opportuno, contenente, per ogni macchina, le seguenti informazioni: identificativo rack, posizione nel rack, numero seriale macchina, MAC address di tutte le interfacce di rete Ethernet.

3.3. Infrastruttura di collegamento dei server

Nel presente paragrafo, il sotto-paragrafo 3.3.1 illustra le caratteristiche relative al "lotto GPFS", mentre il sotto-paragrafo 3.3.2 illustra le caratteristiche relative al "lotto CEPH".

3.3.1. Lotto GPFS

Dovranno essere forniti 2 switch per la SAN in tecnologia FC (almeno 32 Gbps) o IB (almeno 56 Gbps) con un numero complessivo di porte almeno pari ad almeno alla somma di 2 volte il numero di server ed il numero di porte IB/FC attive sui controller e con tutte le porte "sbloccate" (ovvero con la licenza abilitata).

Inoltre, dovranno essere disponibili e realizzati almeno 2 cablaggi di interlink fra i due switch offerti. Nel caso venga offerta una SAN basata su tecnologia FC, dovranno anche essere realizzati altrettanti cablaggi di interlink tra i 2 switch offerti con altrettanti apparati già in nostro possesso (si tratta di switch di tipo Brocade con Fabric OS v7.2.1d) con i quali quindi dovranno essere pienamente compatibili.

Dovranno essere anche forniti tutti i cavi di lunghezza opportuna (massimo 25 m: in sede di installazione si valuterà eventualmente la sostituzione con cavi di lunghezza minore) per il collegamento tra il Sistema Storage offerto, gli switch e i server oltre che il servizio di cablaggio e attestazione.

Gli switch IB o FC dovranno essere accessibili tramite un tool GUI di configurazione e monitoring remoto.

Le interfacce dei controller dei sistemi storage verso la SAN, al pari di quelle dei server verso la SAN, dovranno essere omogenee.

Gli switch offerti dovranno essere completi di alimentatori e apparati di ventilazione ridondati hot-swap.

Si precisa che gli switch verranno montati a rack in isole ad alta densità (corridoio caldo) e quindi essi dovranno gestire il flusso d'aria rispettando tale vincolo oltre a quelli dati del cablaggio.

Si precisa che anche per il sistema SAN deve valere la garanzia e manutenzione in maniera analoga all'intero sistema oggetto della fornitura. Sono applicabili inoltre le stesse penali descritte per le condizioni in garanzia della fornitura.

3.3.2. Lotto CEPH

Dovranno essere forniti 2 switch in tecnologia Ethernet, ciascuno con le seguenti caratteristiche:

- Almeno 24 porte 25 Gbps (tutte con la licenza abilitata)

- 2 porte di uplink 100 Gbps Ethernet in fibra per collegamento alla LAN (*si precisa che tali interfacce verranno collegate in fibra direttamente ai due switch di CORE Cisco NEXUS 9516 su schede N9K-X9732C-EX o N9K-X97160YC-EX. I transceiver da inserire sui due CORE Cisco NEXUS 9516 dovranno essere forniti dall'azienda. In sala calcolo è disponibile, per il collegamento dei server, un cablaggio duplex MMF 50/125: l'azienda dovrà verificare la congruità di tale cablaggio con la soluzione proposta e, nel caso in cui non sia utilizzabile, sarà a carico dell'azienda fornire tutto il necessario per il collegamento in rete delle risorse proposte ossia i cavi di collegamento verso il centro stella della rete*)
- managed, non-blocking,
- supporto Ethernet bonding
- supporto dei Jumbo frame
- supporto del Vlan Tagging secondo lo standard 802.1Q.
- Inoltre, lo switch dovrà essere configurabile via CLI (Command Line Interface) via IP (telnet, ssh), dovrà supportare il protocollo SNMP per il Monitoring ed il Management e deve essere installabile su rack standard da 19" occupando non più di una unità Rack (1U).
- I due switch dovranno supportare tecnologie di "multi chassis link aggregation" che consentano di gestire i 2 apparati come un unico device. Dovrà essere quindi possibile collegare tutti i server ad entrambi gli switch con interfacce in bonding con tutti i link attivi contemporaneamente.

Gli switch offerti dovranno essere completi di alimentatori e apparati di ventilazione ridondati hot-swap.

Si precisa che anche per gli switch dovrà valere la garanzia e manutenzione in maniera analoga all'intero sistema oggetto della fornitura. Sono applicabili inoltre le stesse penali descritte per le condizioni in garanzia della fornitura.

3.4. Alimentazione

Tutte le apparecchiature fornite (server, sistemi disco, switch etc..) dovranno essere di tipo Titanium.

Dovranno essere inoltre fornite le PDU idonee agli attacchi presenti nel Rack Schneider, comprensive di spine di collegamento IEC 309 interbloccate, da 400 V (5P) o 230 V (3P).

3.5. Specifiche switch di management

Dovrà essere fornito uno switch per le connessioni di management (storage, server e PDU) con le seguenti caratteristiche:

- Switch Gigabit ethernet **con** 48 porte Gigabit ethernet su rame (10/100/1000) con connettori RJ45
- almeno 2 Uplink in Gigabit Ethernet dotati di transceiver (SFP) su fibra ottica multimodale SR (per brevi distanze)

- alimentazione ridondata con utilizzo complessivo di 1 Unità Rack senza dovere ricorrere ad un dispositivo esterno allo switch stesso e con la possibilità di collegare gli alimentatori a due sorgenti di energia differenti
- supporto dei Jumbo frame
- supporto del Vlan Tagging secondo lo standard 802.1Q.
- Inoltre, lo switch deve essere configurabile via CLI (Command Line Interface) via IP (telnet, ssh), deve supportare il protocollo SNMP per il Monitoring ed il Management e deve essere installabile su rack standard da 19" occupando non più di una unità Rack (1U).

4. FORNITURA DEL SISTEMA

Tutto il materiale dovrà essere montato nei Rack già presenti: Rack da 19" di tipologia APC Netshelter VX integrati in "isole ad alta densità".

Nel caso si ritenga non adatta questa tipologia di rack, sarà a carico del fornitore, previo sopralluogo e nulla osta del committente per l'inserimento nelle isole presenti nel data center, fornire il rack (o i rack) di installazione.

Il numero, la tipologia della presa di alimentazione e la lunghezza di tutti i cavi, dipendono dal tipo di soluzione tecnica presentata e non può pertanto essere specificato rigidamente prima dell'aggiudicazione della gara: sarà dovere del vincitore richiedere prima della consegna la tipologia di presa necessaria.

Le presiere dovranno essere montate a rack e collegate alle blindo-sbarre.

Dovranno essere forniti cavi di alimentazione con spine/prese il cui standard e lunghezza verrà specificato da ogni sede. A tale scopo contattare le persone indicate nel Par. 2.2.

Sarà possibile prima della presentazione dell'offerta effettuare un sopralluogo nella sala macchine oggetto dell'installazione.

L'installazione e la messa in funzione dell'intero Sistema sui rack già presenti nella sala macchine ed il cablaggio completo delle alimentazioni e di tutti i collegamenti necessari al funzionamento del sistema dovrà essere incluso nella fornitura, oltre allo smaltimento completo dei materiali di risulta (imballaggi e scatole) da effettuarsi congiuntamente alla fase di installazione.

Per tutte le unità componenti la fornitura dovrà essere fornita la ridondanza degli alimentatori, dei sistemi di controllo RAID e delle unità di ventilazione oltre che delle altre parti essenziali come eventuali batterie di cache ed eventuali moduli di controllo con sostituzione hot-swap (sostituzione a caldo senza fermo macchina).

In particolare, dovranno essere fornite tutte le indicazioni sui consumi, i requisiti in termini di condizionamento, le dimensioni fisiche e i pesi degli oggetti offerti oltre che, obbligatoriamente il calcolo complessivo di occupazione a rack della soluzione proposta.

4.1. Rischi

Nella sede della fornitura non esistono rischi da interferenze per le attività di installazione. In ogni caso, l'installazione dovrà essere concordata con un anticipo di almeno 10 giorni, in modo da consentire l'eliminazione dei rischi da interferenza eventualmente sopravvenuti. Per la sicurezza la persona di contatto è:

- Ing. Carlo Crescentini (carlo.crescenti@bo.infn.it), tel. 051 20 95266

5. GARANZIA E ASSISTENZA

Per i beni oggetto del contratto, in base agli artt. 1490 e 1495 del c.c., l'appaltatore dovrà fornire idonea garanzia non inferiore a 12 mesi.

L'assistenza tecnica dovrà essere fornita direttamente dalla azienda produttrice (casa madre) degli apparati offerti. Qualora questo non sia tecnicamente possibile (ad es. nel caso la casa madre non possieda centri di assistenza in Italia) dovranno essere fornite opportune giustificazioni tecniche in merito. Nel caso comunque il concorrente non intenda effettuare il servizio di manutenzione, dovrà richiedere, in sede di offerta, l'autorizzazione al subappalto. In questi casi l'assistenza tecnica dovrà essere comunque fornita da azienda autorizzata e certificata dalla casa madre. Per l'assistenza tecnica si richiede inoltre il possesso della certificazione ISO 9001 con indicazione della data di conseguimento e dell'ente certificatore.

5.1. Costo del servizio

Dovrà essere incluso nell'offerta economica ed indicato **separatamente** da quello della fornitura con la dicitura "costo totale di manutenzione ed assistenza" il costo annuale di un contratto di manutenzione e assistenza on-site (modalità dettagliate nel Par. 5.2), **della durata di almeno 5 anni**, per tutto il materiale della fornitura, con scadenze di pagamento annuali posticipate alla fine dell'anno di manutenzione corrispondente, fatto salvo l'applicazione delle eventuali penali.

5.2. Tipologia di servizio richiesto

Nel periodo indicato di 5 o più anni, l'operatore economico si impegna a sostituire ed installare a sua cura e spese quelle parti della fornitura hardware che, per qualsiasi motivo, dovessero risultare in un qualsiasi momento difettose o difformi dalle specifiche, nonché ad effettuare tutte le prestazioni conseguenti per tutto il periodo di copertura contrattuale.

La ditta che effettuerà la manutenzione dovrà mettere a disposizione del CNAF un centro per la ricezione e gestione delle chiamate riguardanti le richieste di manutenzione in garanzia. Tale centro dovrà essere operativo, con operatori addetti, tutti i giorni dell'anno, con esclusione di sabato, domenica e festivi, dalle ore 09:00 alle ore 18:00. Al di fuori di tale fascia oraria potrà essere attivata una segreteria telefonica o un fax per la registrazione delle chiamate, le quali dovranno intendersi come ricevute alle ore 9:00 del giorno lavorativo successivo.

Si richiede inoltre che la ditta dichiari **un tempo massimo di ripristino garantito dall'apertura della chiamata e che dovrà essere comunque chiaramente specificato nella offerta.** Tale tempo di ripristino garantito va inteso come il "tempo necessario per riportare esattamente nelle condizioni ottimali immediatamente antecedenti il guasto l'oggetto ove il guasto stesso si sia

verificato senza causare nel contempo una perdita di dati o di informazioni” e sarà utilizzato come riferimento nel contratto di assistenza tecnica e suscettibile delle penali descritte nel “Disciplinare di Gara”.

Dovrà essere obbligatoriamente previsto un sistema che preveda supervisione remota degli apparati da parte di tecnici specializzati in caso di problemi critici ed eventualmente sistemi automatici di apertura delle chiamate di assistenza.

La ditta deve in ogni caso mettere a disposizione dell’INFN, al momento della firma del contratto, le funzionalità di un Call Center che funzioni da centro di ricezione e gestione delle chiamate relative alle richieste di manutenzione in garanzia per il malfunzionamento delle apparecchiature; dovranno quindi essere comunicati:

- un numero telefonico dedicato,
- un numero di fax dedicato
- un indirizzo di e-mail dedicato.

In caso di chiamata per malfunzionamento la ditta dovrà assegnare, e quindi comunicare alla sede INFN interessata, un numero univoco di chiamata (identificativo della richiesta di intervento) contestualmente alla ricezione della chiamata con l’indicazione della data ed ora di registrazione; i termini di erogazione del servizio di manutenzione in garanzia decorreranno dall’ora di registrazione della richiesta di intervento. In caso di contestazione sulla data di richiesta intervento farà fede la ricevuta della corretta trasmissione della richiesta di intervento via fax o via e-mail.

Per ogni intervento di manutenzione e assistenza on site dovrà essere redatto dalla ditta un apposito “Verbale di Manutenzione”, sottoscritto da un incaricato dell’INFN e da un incaricato della ditta, nel quale dovrà essere dato atto della tipologia di intervento, delle attività svolte e dei livelli di servizio ottenuti; dovranno essere riportate, fra le altre, le seguenti informazioni:

- la data e il luogo dell’avvenuto intervento,
- l’identificativo unico di installazione (assegnato dalla ditta all’atto dell’installazione),
- un identificativo unico dell’intervento,
- un identificativo unico della chiamata (corrispondente a quello assegnato dal Call Center all’atto dell’apertura della richiesta di intervento) ed il corrispondente orario e data di apertura,
- il numero delle apparecchiature oggetto del servizio,
- il quantitativo (numero) e la tipologia delle apparecchiature e della componentistica opzionale consegnata ed installata, nonché l’elenco delle caratteristiche tecniche,
- una descrizione delle attività svolte durante l’intervento,
- in caso di sostituzione di componenti, gli identificativi (part number) delle componenti sostituite e di quelle di rimpiazzo,
- l’orario e la data di inizio dell’intervento,
- l’orario e la data di termine dell’intervento,
- l’orario e la data di ripristino dell’operatività delle apparecchiature

Alla fine dell’intervento, l’incaricato della ditta dovrà compilare e firmare il “Verbale di Manutenzione”; tale rapporto è controfirmato e timbrato dall’INFN che ne tiene una copia,

attestando così formalmente il lavoro eseguito e concludendo l'attività di manutenzione in garanzia per quel specifico evento di guasto.

Per il mancato rispetto dei tempi di risposta o della tempestività di risoluzione dei guasti entro i tempi di ripristino garantiti saranno applicate delle penali come previsto dal "Disciplinare di Gara".

5.3. Caratteristiche specifiche della manutenzione

Nel presente paragrafo, il sotto-paragrafo 5.3.1 illustra le caratteristiche relative al "lotto GPFS", mentre il sotto-paragrafo 5.3.2 illustra le caratteristiche relative al "lotto CEPH".

5.3.1. Lotto GPFS

Le modalità specifiche di erogazione del servizio di garanzia e manutenzione on-site sono le seguenti.

È richiesto l'intervento in loco presso la sala macchine Tier1 dell'INFN CNAF entro 4 ore dall'apertura del guasto, durante tutti i giorni lavorativi in orario 09:00-18:00 con fornitura ed installazione delle parti di ricambio e il ristabilimento delle precedenti condizioni di funzionamento del sistema.

Il tempo di ripristino non potrà essere superiore a 3 giorni lavorativi dall'apertura della chiamata.

5.3.2. Lotto CEPH

È richiesto l'intervento in loco presso la sala macchine Tier1 dell'INFN CNAF entro il giorno lavorativo successivo (NBD) dall'apertura del guasto, durante tutti i giorni lavorativi in orario 09:00-18:00 con fornitura ed installazione delle parti di ricambio e il ristabilimento delle precedenti condizioni di funzionamento del sistema.

Il tempo di ripristino non potrà essere superiore a 3 giorni lavorativi dall'apertura della chiamata.

5.4. Decorrenza del periodo di manutenzione

Per tutta la fornitura, il periodo di garanzia e manutenzione on-site di 5 anni avrà inizio dalla data di formale completa accettazione della fornitura con relativo certificato di collaudo come nulla osta alla fatturazione dello stesso. La mancanza o il ritardo nella consegna di componenti richiesti; la mancata sostituzione di componenti errati o la mancata risoluzione di problematiche relative al collaudo che precluda il raggiungimento delle prestazioni richieste, comporterà il ritardo dell'accettazione della fornitura e di conseguenza anche la partenza della garanzia. Tale principio è da considerarsi un requisito essenziale, pertanto si richiede che l'azienda aggiudicataria fornisca adeguati elementi a prova del fatto che la garanzia e la manutenzione siano riconosciute ed effettuate dalla casa produttrice degli apparati nei suddetti termini

temporali (ad esempio una certificazione scritta del produttore, oppure del partner/rivenditore che si impegna a far valere il periodo pattuito con il produttore).

Si sottolinea che in qualunque caso farà fede la data di pagamento della fornitura da parte dell'INFN per ciò che riguarda l'inizio del periodo di garanzia e manutenzione.

5.5. Varie

Su esplicita richiesta del cliente dovrà essere incluso un aggiornamento di firmware non intrusivo da effettuarsi a caldo (quindi che non richieda un fermo macchina delle entità logiche e che non comporti in nessun caso il rischio di perdita di dati o di informazioni) all'ultima release disponibile per tutti gli elementi costituenti la fornitura nella misura di massimo un upgrade per anno. Restano esclusi da tale conteggio di minimo tutti gli aggiornamenti di firmware considerati necessari al buon funzionamento del sistema da parte del supporto tecnico e che sono quindi da considerarsi attività del contratto di assistenza.

6. INSTALLAZIONE VALIDAZIONE E TRAINING

L'intero sistema dovrà essere assemblato, reso operativo e validato (per validazione si intende la messa in condizione pienamente operativa dell'apparato secondo il giudizio della ditta) a completo carico della ditta, sotto la coordinazione del personale tecnico dell'INFN, in modo da soddisfare tutte le specifiche del presente capitolato tecnico.

L'installazione dovrà comprendere in particolare:

1. alimentazione elettrica del sistema;
2. configurazione dei server;
3. configurazione degli switch;
4. Cablaggio della rete.

Inoltre, per il "lotto GPFS", dovranno essere effettuati:

5. cablaggi tra la SAN (ove presente), i server e le entità logiche offerte (ove presenti);
6. configurazione del sistema storage;
7. installazione e configurazione del software di Management e Allarmistica dell'intero sistema fornito (in particolare per le temperature, le ventole, lo stato degli alimentatori e lo stato dei Raid);
8. configurazione del sistema di failover per le connessioni tra i server e le entità logiche offerte.

Inoltre, sempre per il solo "lotto GPFS":

Dovrà essere inoltre previsto un periodo di training (separato dal servizio di installazione e collaudo) per addestrare opportunamente il personale tecnico della sede INFN di installazione alla gestione e manutenzione dell'intero sistema. In particolare, l'addestramento dovrà focalizzarsi sulla gestione completa delle entità logiche (creazione volumi logici, tool di

management, monitoring e allarmistica ed ottimizzazione dei parametri del sistema). Il periodo di training non dovrà essere inferiore a 3 giorni e dovrà essere svolto on-site dalla ditta incaricata dell'assistenza tecnica con opportuni strumenti didattici su richiesta dell'INFN entro tutto il periodo di copertura contrattuale.

7. COLLAUDO DEL SISTEMA

La ditta è tenuta a fornire garanzia, e assistenza on-site su tutto il materiale costituente la fornitura, con le condizioni presenti nel presente contratto, durante la validazione e per i 2 mesi successivi durante i quali verrà effettuato il collaudo della stessa. Tale periodo si concluderà con l'emissione di un certificato attestante il buon esito del collaudo. Quindi, salvo problemi, entro 2 mesi dalla data di completamento di tutte le forniture, lavori d'installazione e validazione di tutte le apparecchiature oggetto della presente fornitura, il committente provvederà alle verifiche tecniche, prove e constatazioni necessarie per accertare la possibilità di emissione del certificato di collaudo positivo della fornitura. Nel caso il collaudo abbia esito negativo il periodo di collaudo sarà esteso per 1 mese ulteriore. Nel caso i problemi non siano risolti il Committente si riserva il diritto di risolvere il contratto e di rivalersi interamente sulla cauzione descritta nel "Capitolato d'Oneri".

Il sistema sarà collaudato dal personale tecnico dell'INFN opportunamente coadiuvato dai tecnici della ditta che dovranno fornire tutta l'assistenza tecnica necessaria. La fase di collaudo inizierà una volta terminata l'installazione completa e la validazione come descritto nel paragrafo precedente.

Inoltre, sempre per il solo "lotto GPFS":

In accordo con il personale tecnico della sede INFN, il sistema disco sarà opportunamente configurato in array Raid e suddiviso in volumi logici come indicato dalla ditta nel precedente punto 2) del paragrafo 3.1 Specifiche generali. Successivamente, verranno effettuati i test come descritto nel paragrafo successivo. La ditta dovrà prendersi carico di tutto il lavoro sistemistico e di tuning degli apparati coinvolti nel sistema, comprese le macchine server fino a raggiungere le performance previste dai requisiti di gara.

7.1. Protocollo di collaudo

Nel presente paragrafo, il sotto-paragrafo 7.1.1 illustra le caratteristiche relative al "lotto GPFS", mentre il sotto-paragrafo 7.1.2 illustra le caratteristiche relative al "lotto CEPH".

Durante la fase di collaudo verrà misurata la banda passante "sustained" (accesso sequenziale) e "random" in lettura e scrittura del sistema disco offerto.

I server offerti verranno usati in parallelo per effettuare i test sui diversi volumi logici, equamente distribuiti fra tutte le entità logiche, utilizzando la configurazione degli array Raid consigliata per ottenere il miglior compromesso fra prestazioni ed affidabilità (come specificato dalla ditta nel precedente punto 2) del paragrafo 3.1 "Specifiche generali").

Preliminarmente verrà verificata l'effettiva possibilità dei server di usare tutta la banda di rete offerta dalle schede installate tramite l'utility iperf.

7.1.1. Lotto GPFS

Sui server verrà installato congiuntamente con i tecnici della ditta il software GPFS (versione minima 4.2) e l'intero spazio disco verrà formattato in un unico file-system GPFS con caratteristiche di Block allocation type=scatter e Block size = minimo 4MByte verificandone la piena funzionalità operativa e la compatibilità con quanto richiesto dal presente documento. Successivamente, tramite accesso con protocollo GPFS via rete LAN verranno testate le performance in accesso al file-system sopra descritto.

A tale scopo si utilizzeranno i seguenti comandi:

1. comando Linux "dd" (con un numero crescente di sessioni in parallelo), da /dev/zero al file-system (per scritture) e dal file-system a /dev/null (per le letture) effettuando la scrittura e la lettura (sia in modo sia esclusivo che concorrente ovvero operazioni miste di scrittura/lettura) di file di dimensioni pari ad almeno 3 volte la grandezza fisica della memoria dei server e con block size pari a 1024 e 4096 Kbyte.
2. comando Linux "iozone" (lanciati simultaneamente su tutti i server offerti) con opzioni di test esclusivi write/rewrite, read/reread e random access (opzioni -i 0 -i 1 -i 2), numero di processi concorrenti crescente fino a 256 (quindi opzione al massimo -t 256), size dei file scritti da ogni server di dimensione pari ad almeno 3 volte la grandezza fisica della memoria del server e record size di 1024k (-r 1024k) e di 4096k (-r 4096k). I campi utilizzati come valori di risultato dei test saranno i RANDOM READ, SEQUENTIAL READ, SEQUENTIAL WRITE.
3. Il test di cui al punto 2) verrà effettuato anche impegnando contemporaneamente metà dei server in write e l'altra metà in read.

I risultati di entrambe le fasi di test dovranno essere pari almeno a quanto richiesto nel Paragrafo 2 del presente Capitolato.

Successivamente verrà effettuata la fase di collaudo dell'intera fornitura con la configurazione di produzione ovvero installando un cluster GPFS v. 4.2 o successiva e verificandone la piena funzionalità operativa e la compatibilità con quanto richiesto dal presente capitolato in particolare configurando il sistema con un file-system unico (includendo i dischi cosiddetti "dedicati ai metadati") e ripetendo i test di performance sopra descritti.

Verrà effettuata inoltre la misura del tempo di ricostruzione di un raid-set al fine di verificare quanto richiesto riguardo alle caratteristiche tecniche dei controller (vedere paragrafi 2 e 3.1).

7.1.2. Lotto CEPH

Nel prosieguo con NDD indichiamo il numero di dischi dati per server e definiamo $NP=2 \times NDD$.

Su ogni server verrà lanciato un test sulle partizioni dell'unità di test con il tool *iozone*. La riga di comando che verrà utilizzato è:

```
iozone -r 128k -i 0 -i 1 -i 2 -i 8 -t <NP> -s 10G -F /dir1/file /dir2/file /dir3/file ...
```

il numero di processi è da suddividere equamente su tutte le partizioni (potranno esserci delle differenze di massimo uno tra il numero di processi che insistono sulle varie partizioni).

I requisiti minimi risultanti dal test dovranno essere come di seguito indicato, tenendo presente che i valori indicati in KB si intendono omogenei nell'unità di misura con quanto riportato nei risultati da *iozone*. Per ogni parametro di seguito valutato, ogni server dovrà superare la soglia indicata.

1. **SEQUENTIAL READ** (espresso dal valore di: "Children see throughput for <NP> readers"): il valore complessivo sul singolo server dovrà essere uguale o maggiore di **1.1 GB/sec**;
2. **SEQUENTIAL WRITE** (espresso dal valore di: "Children see throughput for <NP> initial writers"): il valore complessivo sul singolo server dovrà essere uguale o maggiore di **1.1 GB/sec**;
3. **MIXED WORKLOAD** (espresso dal valore di: "Children see throughput for <NP> mixed workload"): il valore complessivo sul singolo server dovrà essere uguale o maggiore di **1.1 GB/sec**;

7.2. Varie

Quanto prescritto dal presente paragrafo si applica al solo "lotto GPFS".

Dovrà inoltre essere collaudato in particolare il sistema di failover e di monitoring e allarmistica con notifica via e-mail, effettuando simulazioni di fallimenti hardware su tutti gli apparati oggetto della fornitura in particolare riguardo alimentatori, ventole, dischi e controller. Verrà verificato il corretto comportamento sia dal lato sistema in cui si è simulato il fallimento (notifica via e-mail del problema, e opportune azioni automatiche correttive dello stesso) sia dal lato server, in particolare con la corretta esecuzione del failover via multipath di RHEL. Sarà facoltà del personale incaricato del collaudo di ampliare i test richiesti al fine di approfondire maggiormente alcuni aspetti tecnici e si ricorda che, in ogni caso, **il mancato collaudo ovvero la mancata verifica effettiva che il sistema risponda in ogni sua parte a quanto richiesto dal presente capitolato tecnico costituirà valido motivo per la non accettazione della fornitura.**