



investiamo nel vostro futuro



Dipartimento Interateneo di Fisica
“Michelangelo Merlin”

**FORNITURA DI UNA STRUMENTAZIONE:
“SISTEMA DI STORAGE”**

REFERENTE TECNICO DELLA FORNITURA: PROF. ROBERTO BELLOTTI

IMPORTO COMPLESSIVO DEL LOTTO: € 420.000,00 OLTRE IVA

Firma

e-mail: roberto.bellotti@uniba.it

Data:

30 GEN 2014

Allegato tecnico sistema di storage

(Lotto 2)

1 Specifiche Tecniche del Sistema di Storage

Fornitura di sistemi di storage da rack per uno spazio disco lordo complessivo di almeno 2400 TB oltre che 40 dischi aggiuntivi “*spare*” identici a quelli forniti con i sistemi di storage.

Ogni sistema dovrà essere composto di:

- almeno n. 2 nodi di front-end
- n. 1 unità di storage con spazio disco lordo non inferiore a 480 TB in base al numero di sistemi di storage offerti

I sistemi dovranno essere identici in tutte le loro parti.

1.1 Caratteristiche generali della fornitura

- L'intera fornitura è costituita da almeno 5 Sistemi di Storage ciascuno composto di un insieme indipendente di moduli di storage, l'*unità di storage*, e connesso attraverso canali Fibre Channel (FC) ad almeno due nodi di front-end.
- Per unità di storage si intende un sotto-sistema completamente gestibile in maniera autonoma e costituito da:
 - Un doppio controller Fibre-Channel (FC) ridondato, con uscite attive FC lato host (nodi di front-end) indipendenti dalle altre unità di storage;
 - Un array di dischi di back-end controllati dall'unità stessa;
 - Software di configurazione, gestione ed allarmistica.

1.2 Caratteristiche dei Nodi di Front-End

1.2.1 Caratteristiche fisiche

- L'enclosure dei nodi di front-end deve essere di dimensione 1U o 2U e compatibile per il montaggio nei rack utilizzati nei siti come indicato nel paragrafo 2.2.

1.2.2 Alimentazione elettrica

- Si richiede alimentatore ridondato hot swap.
- Il sistema dovrà essere dotato di cavi di alimentazione con spine compatibili con le presiere (PDU) utilizzate nei rack come indicato nel paragrafo 2.2
- Gli alimentatori devono supportare i requisiti specificati dal costruttore della scheda madre e quelli specificati per tutti i componenti interni all'enclosure.
- I ventilatori, fatta eccezione per quelli delle CPU e degli alimentatori, devono essere ridondanti.

1.2.3 Scheda Madre

- Sulla scheda madre deve essere presente un management controller (BMC) compatibile IPMI versione 2.0 o superiore. Il BMC deve consentire almeno il monitoraggio delle ventole (se presenti), della temperatura di CPU e scheda madre, la gestione remota dell'alimentazione elettrica (possibilità di power-cycle) e l'accesso criptato alla console seriale attraverso la rete (per esempio via RCMP+ oppure SSH).
- La funzionalità IPMI dell'unità di sistema deve essere accessibile sia via web sia attraverso un'applicazione a linea di comando in esecuzione su una macchina Linux remota. Il BMC deve consentire la configurazione dell'utente BMC e dei parametri di rete anche attraverso un'applicazione a linea di comando in esecuzione locale che funzioni sotto Linux
- Il BMC deve mantenere i settaggi, incluse le configurazioni di accesso e di rete, anche qualora l'alimentazione all'unità di sistema venga interrotta; il BMC deve inoltre essere accessibile senza riconfigurazioni una volta che l'alimentazione venga ripristinata.
- La scheda madre deve supportare bootstrap via rete con protocollo PXE 2.0 o superiore. Il BIOS deve consentire la possibilità di eseguire boot via PXE prima del boot locale.
- La scheda madre deve supportare la possibilità di invocare un menu di selezione del device di boot e di selezionare il device di boot all'accensione.
- Il BMC deve supportare la possibilità di cambiare l'ordine dei dispositivi di boot al successivo reboot, tramite un'applicazione a linea di comando sotto Linux o remotamente tramite la sua interfaccia LAN.
- La scheda madre deve supportare la ridirezione della console ad una porta seriale accessibile attraverso IPMI Serial-Over-LAN. Tutti i cambiamenti di settaggi BIOS, l'invocazione del menu di selezione del device di boot e la selezione del device di boot devono essere possibili attraverso questa porta via rete.
- Non è richiesta la ridondanza di BMC o equivalenti.

1.2.4 Processori

- I nodi di front-end devono essere dotati di almeno due processori di architettura x86_64 con tecnologia multi-core forniti di estensioni EM64T oppure AMD64.
- I processori devono far parte della famiglia Intel Xeon E5-26xx o della famiglia AMD Opteron 62xx o successive.
- Ogni processore deve avere almeno 8 core fisici. Unità di processing logiche quali quelle fornite da Hyper-Threading oppure Simultaneous Multi-Threading sono considerate come core separati.
- Ogni processore deve avere almeno 16 MB di cache L3.

1.2.5 Memoria

- I nodi di front-end devono essere equipaggiati di almeno 64 GB di memoria RAM ognuno.
- I moduli di memoria devono essere approvati dal costruttore della scheda madre specificamente per l'utilizzo sulla scheda madre fornita.
- La memoria deve essere del tipo DDR-3 Registered ECC ed operante, nel sistema fornito, ad una frequenza di almeno 1333 MHz.
- Le ditte partecipanti alla gara dovranno elencare le specifiche della memoria proposta e la configurazione di installazione nei nodi.

1.2.6 Storage locale

- I nodi devono essere equipaggiati con almeno due dischi con possibilità di RAID 0,1 hardware.
- I dischi devono essere di tecnologia SATA di fascia Enterprise, SCSI o SAS ad almeno 7200 RPM.
- La capacità di ciascun disco deve essere di almeno 500 GB con almeno 16 MB di cache in alloggiamento hot swap.
- Tutti i dischi devono essere certificati dal costruttore per operare continuativamente 24 ore al giorno, 7 giorni la settimana.
- Il controller dei dischi deve verificare lo stato dei dischi attraverso controlli e contatori SMART.

1.2.7 Connettività

- I nodi di front-end devono essere forniti di almeno:
 - n. 2 schede Gigabit Ethernet rame integrate con supporto PXE ciascuna dotata di connettore RJ45 per i cavi in rame (1000baseT);
 - n. 2 schede a 2 porte 10 Gbps con connettore SFP+ con supporto PXE;
 - n. 2 transceiver ottici SFP+ SR multimode corredati delle opportune bretelle ottiche per il collegamento (almeno 1,5 m);
 - n. 2 interfacce Fibre Channel 8 Gigabit/s Qlogic QLE2560 o equivalente corredata delle opportune bretelle ottiche per il collegamento (almeno 1,5 m).
- L'accesso alla console remota IPMI deve essere possibile utilizzando una delle interfacce specificate al paragrafo 1.2.3, ovvero non richiedere necessariamente l'utilizzo di un'interfaccia di rete separata.

1.2.8 Sistema Operativo

- Il sistema operativo da installare sarà Scientific Linux 6.4 o superiore, disponibile gratuitamente dal sito <http://www.scientificlinux.org> compatibile a livello binario con RedHat Enterprise Linux 6 Server x86_64).
- I sistemi non devono contenere componenti o device che richiedano driver non inclusi in tale sistema operativo. Non sono richieste licenze per il sistema operativo e non se ne richiede l'installazione.
- La fornitura e l'installazione saranno a completo carico del fornitore.
- Tutti i computer dovranno essere in grado di fare il boot del sistema operativo citati senza essere connessi a tastiera, video, mouse o console seriali.

1.2.9 Altre richieste

- Ciascun nodo di front-end deve avere un interruttore on/off di alimentazione raggiungibile quando il nodo è montato su rack.
- Tutti i sistemi oggetto della presente gara devono essere forniti con identiche caratteristiche hardware, versioni di firmware e di BIOS.
- I cavi interni non devono essere punzonati, danneggiati o in tensione meccanica.
- La funzionalità della unità di sistema definibili via setting NVRAM (per esempio ordine di boot, configurazione RAID, etc.) devono essere mantenute anche in caso di accidentale perdita di alimentazione elettrica.

- Tutte le componenti di un server devono rispettare le specifiche approvate dai rispettivi costruttori (per esempio non è consentito overclocking delle CPU).

1.3 Caratteristiche delle Unità di Storage

1.3.1 Caratteristiche delle Unità di Storage

- Ogni unità di storage deve avere almeno un doppio controller con porte verso gli host (nodi di front-end) in tecnologia FC ad almeno 8 Gbps
- Ogni controller deve avere almeno 4 porte FC ad almeno 8 Gbps
- Le porte FC devono essere utilizzabili sia in configurazione DAS o SAN
- Ogni controller deve avere almeno 2 GB di memoria cache con batteria tampone o sistema equivalente.
- E' richiesta la tipologia *active/active* garantita dai due controller.
- Entrambi i controller devono supportare la creazione di raid set con livelli, 0,1,5,6, composti da almeno 4 dischi fino ad un massimo di 15, ovvero sia possibile creare raid set composti da 4,5, .. 15 dischi. Entrambi i controller devono anche supportare configurazioni avanzate basate sui pool di dischi dinamici, o tecnologie equivalenti che consentano l'uso di tutti i dischi per distribuire il carico di lavoro.
- 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 Scientific Linux almeno nella versione 5.4 e 6.4 a 64 bit e successive versioni ed il relativo tool di multipath. E' possibile, in aggiunta al sistema di multipath del sistema operativo Scientific Linux 6.4 a 64 bit e versioni successive, la fornitura di tool di failover e load-balancing specifici delle entità logiche offerte. In tal caso dovrà essere garantito il funzionamento dei tool con il sistema operativo sopra descritto per tutti gli anni di durata contrattuale.
- La configurazione di failover sarà testata durante la fase di verifica di conformità e costituirà valido motivo per la non accettazione della fornitura qualora la stessa si dimostrasse non rispondente ai requisiti di cui sopra.
- Deve essere certificata la possibilità di ospitare dischi SAS e SATA nello stesso enclosure.
- Deve essere possibile definire dischi come hot spare locali e globali.
- Ogni unità di storage deve supportare almeno 170 meccaniche.

1.3.2 Dischi

- I dischi ammessi in fornitura devono avere interfaccia SAS, capacità di 3 o 4 TB, velocità di rotazione di almeno 7200 RPM e almeno 16 MB di cache in cassette hot swap.
- I dischi devono essere certificati per l'uso 24x7 (categoria enterprise).

1.3.3 Caratteristiche fisiche

- Gli enclosure devono essere compatibili per il montaggio nei rack presenti presso l'Università degli Studi di Bari *Aldo Moro*, descritti nel paragrafo 2.2.
- La densità degli enclosure deve essere tale che siano presenti non più di 15 meccaniche per ogni Rack Unit.
- Si richiede alimentatore ridondato hot swap. Il sistema dovrà essere dotato di cavi di alimentazione con spine compatibili con le presiere utilizzate nei rack.

- Le unità di ventilazione degli enclosure contenenti controller e dischi devono essere ridondate con sostituzione a caldo senza fermo macchina.
- Tutte le unità di storage della presente gara devono essere fornite con identiche caratteristiche hardware, versioni di firmware e di BIOS.

1.3.4 Prestazioni minime dell'unità di storage

- Ogni unità di storage dovrà soddisfare a delle condizioni minime di prestazioni che verranno verificate nella fase di verifica di conformità della fornitura (vedi paragrafo 1.5). La modalità operativa del test sarà così articolata:
 - Almeno due server (OS SL6.x x86_64; si veda par. 1.6 per dettagli sul OS)) connessi all'unità di storage attraverso i due controller; oppure almeno un server munito di scheda FC dual head 8 Gbps connessa ai due controller;
 - test con il tool *iozone* (<http://www.iozone.org/>) 32 bit v3.4.7 con file system XFS (OS SL5.x x86_64) sulle partizioni dell'unità di storage. La riga di comando da utilizzare è: `iozone -r 128k -i 0 -i 1 -i 2 -t <#processi> -s 10G`. (Per la definizione di <#processi> si veda di seguito);
 - partizioni dell'unità di storage con formattazione standard di XFS (`mkfs.xfs /device` senza ottimizzazioni);
 - ogni raidset contenga un solo volume; ogni volume sia mappato in una singola LUN; ogni LUN sia mappata in una singola partizione montata su uno dei disk-server;
 - il numero di processi (<#processi>) per un'unità che contiene circa 480 TB è fissato in almeno 300:
 - i processi devono essere distribuiti in round-robin fra tutte le partizioni disponibili in modo da avere lo stesso numero di processi su ogni partizione o al più con una sola unità di differenza;
- 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*:
 - RANDOM READ (espresso dal valore di: "Children see throughput for <xx> random readers", dove <xx> è il valore <#processi>):
 - il valore aggregato per una unità di storage equipaggiata con circa 480 TB, nel caso di random read, deve essere almeno 760.000 KB/s;
 - SEQUENTIAL READ (espresso dal valore di: "Children see throughput for <xx> readers", dove <xx> è il valore <#processi>):
 - il valore aggregato per unità di storage equipaggiata con circa 480 TB, nel caso di sequential read, deve essere almeno 1.700.000 KB/s;
 - SEQUENTIAL WRITE (espresso dal valore di: "Children see throughput for <xx> initial writers", dove <xx> è il valore <#processi>):
 - il valore aggregato per unità di storage equipaggiata con 480 TB, nel caso di sequential write, deve essere almeno 1.300.000 KB/s
 - Tutti i valori, sia per il numero di processi sia per i valori di performance, vanno scalati con una proporzione lineare in base al numero di TB con cui le unità di storage vengono equipaggiate.

1.3.5 Gestione e Monitoraggio dell'unità di Storage

- Deve essere garantita la gestione, la configurazione e il monitoraggio remoto via seriale, ssh web, snmp e il supporto per lo shutdown e reset remoto dei controller RAID, in base a soglie e politiche configurabili.
- I sistemi dovranno essere dotati di un sistema configurabile di controllo in background su tutti i volumi dello stato dei Raid.
- Il sistema dovrà essere fornito di un software completo di controllo e configurazione. Tale 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, fallimento di un'unità di alimentazione/ventilazione, fallimento di un disco di hot-spare e in particolare il fallimento di un singolo disco costituente un array Raid. Tale software dovrà essere installabile su piattaforma Windows e/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 ben documentata e dovrà permettere il monitoraggio in tempo reale dei parametri suddetti.

1.4 Descrizione delle caratteristiche dei sistemi di storage

- Per ogni Unità di Storage offerta dovranno essere obbligatoriamente specificate le seguenti informazioni che verranno anche utilizzate in fase di verifica di conformità:
 1. Marca e modello dell'unità di storage;
 2. 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 hot-spare);
 3. Caratteristiche dei dischi offerti;
 4. Numero totale di enclosure componenti l'unità di storage;
 5. Numero di Rack Unit e profondità del singolo enclosure completo di alimentatori, dischi e quant'altro;
 6. Numero di alimentatori e ventole di raffreddamento;
 7. Configurazione usata per la verifica delle performance richieste
 8. Tutti i dischi offerti dovranno essere configurabili nel Raid Set specificato;
 9. Capacità del più esteso volume logico gestibile dal controller che potrà essere costituito da partizioni di uno stesso Raid Set o da partizioni risultanti dalla combinazione di più Raid Set;
 10. Capacità massima dell'eventuale futura espansione dei sistemi possibile con i controller presenti nella configurazione offerta, in maniera non intrusiva, quindi che richieda al massimo un fermo macchina ma che non comporti il rischio di perdita di dati o di informazioni.
 11. Risultati dei test di performance eseguiti sui sistemi descritti nel paragrafo 1.3.4 e la configurazione usata per ottenerli. Tali valori saranno verificati in fase di verifica di conformità;
 12. Apparecchiature e/o licenze necessarie per l'utilizzo del sistema in configurazione DAS e per la sua conversione in modalità SAN;
 13. Potenza assorbita del sistema di storage: dovranno essere indicate la potenza AC apparente (in VA), la potenza AC attiva (in W) e il fattore di potenza del sistema pienamente popolato di tutti i dischi offerti, sotto pieno carico e in modalità idle;
 14. Lista indicativa degli switch compatibili con le schede FC.

- Per i nodi di front-end offerti dovranno essere obbligatoriamente specificate le seguenti informazioni:
 - 15. Caratteristiche del processore offerto;
 - 16. Numero e caratteristiche dei dischi offerti;
 - 17. Caratteristiche della memoria e GB di RAM offerti;
 - 18. Caratteristiche delle schede di connettività offerte;
 - 19. Potenza assorbita dai nodi di front-end: dovranno essere indicate la potenza AC apparente (in VA), la potenza AC attiva (in W) e il fattore di potenza sotto pieno carico e in modalità idle;

- In tutti i casi, nell'offerta andranno indicati chiaramente per i diversi componenti i sistemi di gestione e monitoraggio *non presidiato (tramite automatismi software)*:
 - requisito preferibile è la presentazione delle metriche di monitoring e l'implementazione di tecnologie di controllo remoto tramite protocolli standard come SNMP, IPMI, ecc.
 - In particolare il monitoraggio non presidiato sarà effettuato da un nostro server con sistema operativo Linux sul quale gireranno script, programmi e/o daemon che dovranno poter:
 - monitorare ventole, alimentatori, tensioni, temperature, eventuali altri componenti attivi elettrici ed elettronici;
 - eseguire power on, power off e reset dei nodi di front-end e shutdown o reset dei controller RAID dello storage, in base a soglie e politiche configurabili.

2 Consegna e Installazione delle apparecchiature

2.1 Consegna delle apparecchiature

- Il Fornitore dovrà provvedere alla consegna delle apparecchiature entro 30 giorni solari dal momento in cui verrà sottoscritto il contratto e comunque non prima del 1 maggio 2014. L'intero sistema dovrà essere installato nei rack, reso operativo e validato entro il 31 luglio 2014.
- L'indirizzo di consegna è il seguente:
 - **Dipartimento Interateneo di Fisica "Michelangelo Merlin"**
Campus Universitario

Via E. Orabona n. 4
70125 Bari

2.2 Installazione delle apparecchiature

- Tutto l'hardware fornito dovrà essere installato nei rack, reso operativo e validato ad opera del fornitore.
- L'installazione dovrà comprendere per tutte le apparecchiature la cablatura all'alimentazione elettrica e alla rete, inoltre, per i Sistemi di Storage del Lotto 2, dovrà comprendere il cablaggio tra le unità di storage e i nodi di front-end, il montaggio e la configurazione delle schede Fibre Channel in modo da avere 2 canali ridondati per nodo.
- Il fornitore dovrà inoltre provvedere al settaggio di rete delle console remote dei sistemi.
- Le guide e gli accessori (bulloni e viti varie) necessari per il montaggio e l'installazione a rack, così come i cavi di alimentazione, di rete e le fibre, fanno parte della fornitura.

- I rack utilizzati presso l'Università degli Studi di Bari *Aldo Moro* sono del seguente tipo:
 - Rack della ditta APC, modello SX AR3100
- Le PDU, disponibili presso l'Università degli Studi di Bari *Aldo Moro*, sono del tipo IEC C13.
- L'installazione dovrà essere concordata con un anticipo di almeno 10 giorni, in modo da impedire l'accesso ai locali per altre attività. A tale scopo contattare il Direttore dell'esecuzione del contratto, Prof. Roberto Bellotti (080-5443171, <roberto.bellotti@uniba.it>). Il Prof. Roberto Bellotti sarà garante dell'eliminazione dei rischi da interferenza.