

**PNC - DD 931 del 6 giugno 2022. Avviso per la concessione di finanziamenti destinati ad iniziative di ricerca per tecnologie e percorsi innovativi in ambito sanitario e assistenziale D3 4 Health Digital Driven Diagnostics, prognostics and therapeutics for sustainable Health care – CUP B53C22006140001**

**OSPEDALE PEDIATRICO BAMBINO GESÙ**

**CAPITOLATO TECNICO**

**Server per il calcolo di dati**

**Premessa**

---

Il presente capitolato, nell'ambito della creazione di centri di ricerca nazionale, selezionati con procedure competitive, che siano in grado di raggiungere, attraverso la collaborazione con le Università, centri di ricerca e imprese, una soglia critica di capacità di ricerca e innovazione, ha come oggetto la fornitura degli strumenti necessari per l'incremento del pacchetto strumentale già in dotazione all'Ospedale nell'ambito dell'iniziativa PNRR e la creazione di un Centro Nazionale per Terapia RNA e Terapia Genica.

Nell'implementazione delle attività di ricerca e sviluppo svolte dall'Unità di Ricerca Terapia Genica dei Tumori, afferente all'Area di Ricerca Oncoematologia, si rende necessario l'acquisto di un server dedicato alle attività di analisi e post-analisi di dati genomici e trascrittomici, comprese le immagini di trascrittomica spaziale, ottenute dalle apparecchiature di sequenziamento di nuova generazione in uso per la conduzione del progetto di ricerca.

Il presente Capitolato Tecnico disciplina quindi le condizioni tecniche minime per la fornitura, installazione, configurazione e messa in esercizio di un'infrastruttura di calcolo ad alte prestazioni (HPC) e del relativo storage locale, destinata all'elaborazione di dati genomici e trascrittomici generati dai progetti di ricerca.

L'infrastruttura sarà utilizzata per l'analisi dei seguenti progetti:

- RNA-seq (200 campioni)
- Whole Genome Sequencing Oxford Nanopore (16 campioni)
- Targeted Capture DNA-seq (150 campioni)
- Single-Cell 10x Genomics RNA-seq (50 campioni)

Oltre a questi, ci saranno futuri progetti di analisi dei dati di sequenziamento di nuova generazione (NGS) per la caratterizzazione genomica e trascrittomica delle leucemie pediatriche, di dimensioni simili o superiori a quelli qui elencati.

## Tipologia di Dati da Elaborare

---

I dati da elaborare sono:

- ottenuti dal sequenziamento di acidi nucleici derivati da diversi materiali biologici (umani e animali),
- immagini acquisite da sistemi di microscopia avanzata e da apparati di bioimmagini,
- acquisiti dalle cartelle cliniche elettroniche dei pazienti.

## Risorse di Calcolo Richieste

---

Di seguito sono riportate le risorse richieste per i progetti di immediata conduzione:

Project 1 – RNA-seq (200 campioni):

- CPU: 2000 core (multi-threaded)
- RAM: 1500 GB
- Storage: 8 TB SSD
- GPU: non richieste

Project 2 – Bulk Oxford Nanopore WGS Whole Genome Sequencing (16 campioni):

- CPU: 600 core (multi-threaded)
- RAM: 2400 GB
- Storage: 45 TB SSD
- GPU: 3x GPU (min. 8 GB VRAM ciascuna)

Project 3 – Targeted Capture DNA-sequencing (150 campioni):

- CPU: 700 core (multi-threaded)
- RAM: 1500 GB
- Storage: 8 TB SSD
- GPU: non richieste

Project 4 – Single-Cell 10x Genomics RNA- sequencing (50 campioni):

- CPU: 150 core
- RAM: 200 GB
- Storage: 6 TB SSD
- GPU: 5x GPU (min. 8 GB VRAM ciascuna)

Di seguito tutti i requisiti dei progetti riportati in forma tabellare:

Progetto	Campioni	CPU	RAM (GB)	Storage	GPU
RNA-seq	200	2000	1500	8 TB	N/A
Bulk Oxford Nanopore WGS	16	600	2400	45 TB	3x GPU (8 GB)
Targeted Capture DNA-seq	150	700	1500	8 TB	N/A
Single-Cell 10x Genomics	50	150	200	6 TB	5x GPU (8 GB)

## Altri Requisiti Tecnici

---

Compatibilità con ambienti containerizzati (Docker, Apptainer).  
Scalabilità per progetti futuri.

## Architettura generale della soluzione

---

L'infrastruttura HPC oggetto della fornitura dovrà essere realizzata secondo un'**architettura di riferimento enterprise**, specificamente progettata per ambienti di Calcolo ad Alte Prestazioni (HPC) destinati all'elaborazione di dati scientifici ad alta intensità computazionale, con particolare riferimento ai carichi genomici, trascrittomici.

La soluzione proposta dovrà basarsi su **uno stack tecnologico omogeneo e integrato**, comprendente nodi di calcolo, sistema di storage parallelo e infrastruttura di rete ad alte prestazioni, **progettato, certificato e supportato dal medesimo produttore**, al fine di garantire elevata affidabilità ed elevate prestazioni.

## Oggetto della Fornitura

---

La fornitura dovrà comprendere:

- Cluster HPC con nodi CPU e nodi GPU
- Storage locale ad alte prestazioni per dati grezzi e intermedi
- Sistema operativo e software scientifico e container runtime
- Installazione completa, testing e collaudo
- Formazione personale
- Garanzia e assistenza tecnica

Il fornitore dovrà prevedere e includere nell'offerta tutte le componenti hardware, software, licenze, cablaggi, moduli, transceiver, configurazioni e attività necessarie a garantire la piena connettività dell'intero ambiente HPC, assicurando il corretto e completo funzionamento del cluster nel suo complesso, senza necessità di ulteriori integrazioni o forniture aggiuntive da parte dell'Amministrazione.

La fornitura, la messa in esercizio ed il collaudo finale **DEVE concludersi entro e non oltre il 30 ottobre 2026**.

Si richiede di suddividere l'offerta in:

### Nodi CPU e GPU

#### CPU:

- 2× CPU Intel Xeon / AMD EPYC
- ≥32 core fisici per CPU

#### Memoria RAM:

- ≥512 GB DDR5 ECC o superiore

### Storage Sistema Operativo:

2× SSD M.2  $\geq 900$  GB in RAID1  
hot-swap

### Storage locale dati:

$\geq 45$  TB  
RAID 0/1/5/6/10  
dischi hot-swap  
ottimizzato per HPC

### Connettività:

1× scheda 2× SFP+ (10/25 Gbps) per connettività cluster  
1× scheda 2× SFP+ (10/25 Gbps)  
1× porta 1 GbE per management (BMC/ILO/IDRAC)

### Supporto GPU:

chassis e alimentazione compatibile per schede GPU (es. NVIDIA H100)

### Nodi di calcolo e integrazione con lo storage

I nodi di calcolo (CPU e GPU) dovranno essere progettati per l'accesso concorrente e massivo a un file system condiviso ad alte prestazioni, attraverso client nativi sviluppati e mantenuti dal produttore dello storage e ottimizzati per carichi HPC.

L'accesso allo storage dovrà avvenire **senza** ricorso a layer di astrazione generici o protocolli di rete tradizionali non progettati per HPC

### Storage locale (Requisiti minimi)

- **Capacità minima complessiva:**  $\geq 45$  TB raw  
(adeguata ai requisiti del progetto WGS e multi-progetto)
- **Tipologia:** SSD 6 Gbps, 1 DWPD oppure tecnologia equivalente o superiore
- **Supporto a RAID 0/1/5/6/10**
- **Hot-swap** per tutti i dischi
- **Interfaccia:** SAS/SATA con controller dedicato oppure tecnologia equivalente o superiore
- **Affidabilità:** MTBF certificato dal produttore

### Requisiti architetturali dello storage

Lo storage locale a servizio del cluster HPC dovrà essere basato su un **file system parallelo enterprise**, conforme allo standard POSIX idoneo a consentire l'accesso concorrente ai medesimi file da parte di più nodi di calcolo senza degradazione delle prestazioni e specificamente progettato per ambienti HPC per l'accesso concorrente da migliaia di core di calcolo **HPC e workload bioinformatici (Whole Genome Sequencing – WGS)**, elevato throughput e gestione simultanea di file di grandi e piccole dimensioni.

### Funzionalità minime del file system parallelo

Il file system HPC dovrà supportare nativamente le seguenti funzionalità avanzate:

- distribuzione dei metadati su più nodi dedicati, con replica e meccanismi di alta affidabilità;
- accesso concorrente massivo in lettura e scrittura;

- snapshot consistenti a livello di file system;
- compressione e tecnologie avanzate di ottimizzazione dello spazio;
- replica e caching remoto opzionale per scenari di espansione futura.

### Requisiti minimi

- **Capacità complessiva:** secondo quanto esposto le risorse necessarie per i progetti di immediata conduzione richiedono almeno **45 TB raw**, adeguata ai requisiti del progetto WGS e a scenari multi-progetto. Oltre a questi, ci saranno futuri progetti di analisi dei dati di sequenziamento di nuova generazione (NGS) per la caratterizzazione genomica e trascrittomica delle leucemie pediatriche, di dimensioni simili o superiori a quelli qui elencati. Pertanto, la capacità minima richiesta sin dalla prima configurazione è di **100 TB**, con possibilità di espansione futura senza interruzione del servizio.
- **Supporti di memorizzazione:** dischi di tipo **SSD enterprise** con resistenza minima pari ad **1 DWPD** oppure tecnologia equivalente o superiore, idonei a carichi I/O intensivi.
- **Protezione dei dati:** supporto a meccanismi di protezione quali **RAID 0/1/5/6/10** o soluzioni equivalenti a livello di file system/storage parallelo, inclusi schemi di ridondanza distribuita.
- **Accesso parallelo:** il file system dovrà supportare operazioni di **I/O parallelo** e accesso concorrente ai file da parte di più client, garantendo elevate prestazioni aggregate e consistenza dei dati.
- **Connettività:** interfacciamento tramite rete dedicata ad alte prestazioni (es. **Ethernet  $\geq 25/40/100$  GbE e/o InfiniBand**), idonea a non rappresentare un collo di bottiglia rispetto alle prestazioni dei dischi.
- **Interfacce disco:** SAS/SATA con controller dedicato oppure tecnologia equivalente o superiore, coerente con l'architettura di storage parallelo proposta.
- **Hot-swap:** tutti i dischi dovranno essere sostituibili a caldo senza interrompere il servizio.
- **Affidabilità:** componenti con **MTBF certificato dal produttore** e progettazione orientata all'alta affidabilità e alla continuità operativa.
- **Scalabilità:** possibilità di incremento sia della capacità sia delle prestazioni mediante l'aggiunta di ulteriori nodi storage, senza riconfigurazioni invasive dell'infrastruttura esistente.
- **Compatibilità HPC:** piena integrazione con ambienti Linux e cluster di calcolo, incluso il supporto a workload WGS, analisi genomiche e pipeline bioinformatiche.

La soluzione di storage parallelo dovrà essere fornita completa di tutte le licenze software necessarie, della documentazione tecnica e delle certificazioni richieste, ed essere idonea all'utilizzo in contesti di ricerca scientifica e sanitaria ad alta criticità operativa.

### Connettività infrastruttura

L'offerta dovrà includere tutto quanto necessario a garantire la completa connettività dell'intero ambiente HPC (nodi di calcolo, rete, storage e servizi), mediante rete ad alte prestazioni (es. InfiniBand o Ethernet  $\geq 100$  Gb/s), comprensiva di switch, moduli di espansione, cavi, transceiver e configurazioni, dimensionata per supportare carichi HPC.

L'infrastruttura dovrà prevedere una **rete ad altissime prestazioni**, parte integrante dell'architettura HPC certificata, dimensionata per supportare:

- comunicazione ad alta velocità fra nodi di calcolo;

- accesso parallelo allo storage;

La rete dovrà essere pienamente certificata dal produttore dello stack HPC per l'utilizzo con il file system parallelo adottato.

### Sistema operativo (Compatibile con stack HPC)

- Red Hat Enterprise Linux Server 64 bit, (con licenza per ciascun nodo).
- Oppure sistemi operativi linux compatibili con la tecnologia richiesta (Rocky Linux / AlmaLinux )

### Dimensioni ed assorbimento:

- Unità rack per dimensionamento spazi rack
- Assorbimento elettrico
- Btu per condizionamento

### Installazione e collaudo

---

Il fornitore deve:

- consegnare e installare l'infrastruttura presso i locali indicati da Opgb
- cablare e configurare nodi, rete e storage
- le configurazioni di rete saranno in collaborazione col personale OPBG
- installare OS, software base, container runtime
- verificare performance e corretto funzionamento
- effettuare test di I/O, RAM, CPU, GPU (dove installate)

Il collaudo si considera superato dopo l'esecuzione dei benchmark minimi di sistema e la validazione da parte del personale OPBG (IT è Lab. Di Ricerca).

### Documentazione e Supporto

---

- Il fornitore dovrà rilasciare la documentazione relativa all'installazione e alla configurazione di quanto installato.
- La fornitura dovrà essere corredata da manuale tecnico e manuale d'uso, redatti in lingua italiana ove disponibili, o in alternativa in lingua inglese, in formato elettronico o resi disponibili tramite accesso online.
- La documentazione dovrà includere tutte le informazioni necessarie al corretto utilizzo, alla gestione operativa e alla diagnostica del sistema, comprese le procedure di avviamento, gestione dei fermi, interventi in caso di guasto e operazioni consentite durante le attività di elaborazione.
- Il manuale d'uso dovrà fornire adeguate indicazioni per la manutenzione del prodotto, incluse le informazioni relative alle parti di ricambio.
- Il Fornitore si impegna ad aggiornare e, ove necessario, sostituire la documentazione tecnica per l'intero ciclo di vita dell'attrezzatura.

- Il Fornitore si impegna a supportare audit, tracciabilità, continuità operativa

Supporto tecnico per almeno 12 mesi.

Formazione per il personale.

### Responsabilità e supporto

Il fornitore dovrà assumere la **piena responsabilità funzionale dell'intero stack HPC**, comprendente:

- nodi di calcolo;
- file system parallelo;
- infrastruttura di rete;
- integrazione complessiva della soluzione.

Il supporto di terzo livello dovrà essere erogato **direttamente dal produttore della tecnologia principale di storage HPC**, garantendo un unico punto di responsabilità per la diagnosi e risoluzione delle problematiche.

### Certificazioni

---

Il produttore della soluzione HPC e del file system parallelo dovrà essere in possesso delle certificazioni ISO 9001, ISO/IEC 27001 e garantire supporto diretto di terzo livello.

Il file system parallelo dovrà essere certificato per ambienti HPC, Red Hat Enterprise Linux e RDMA, ed essere adottato in contesti di supercalcolo e ricerca scientifica. Soluzioni prive di tali certificazioni non saranno considerate equivalenti.”

Saranno valutate positivamente soluzioni basate su tecnologie utilizzate in sistemi inclusi nella TOP500.

### Referenze/Esperienze

---

Il fornitore deve dimostrare di avere i seguenti requisiti:

- quale requisito di capacità economico e finanziaria: devono aver conseguito un nei migliori tre anni degli ultimi cinque anni precedenti a quello di indizione della procedura, un fatturato globale pari a: € 1.000.000,00 (unmilione/00)
- quale requisito di capacità tecnica e professione: devono eseguito negli ultimi dieci anni dalla data di indizione della presente procedura contratti analoghi a quello oggetto di affidamento, anche a favore di soggetti privati, per un importo di Euro 5.000.000,00 (cinquemilioni/00).
- devono dimostrare di aver realizzato, negli ultimi 5 (cinque) anni, almeno n.3 (tre) infrastrutture HPC per ambito genomico/biomedico, sanitario o altri contesti, per un valore totale pari o superiore ad Euro 1.000.000,00 (unmilione/00), ciascuna con:
  - file system parallelo enterprise
  - workload WGS / RNA-seq / Single Cell

## Criteri di Valutazione

Cod.	Criterio	Descrizione	Punteggio
T1	Rispondenza ai requisiti tecnici	Grado di conformità della configurazione offerta ai requisiti minimi e migliorativi del capitolato tecnico.	<b>14</b>
T2	Scalabilità e affidabilità	Capacità di espansione futura (CPU, memoria, dischi, slot di espansione) e soluzioni di ridondanza/alta affidabilità dei componenti critici.	<b>8</b>
T3	Life Cycle Management	Strumenti e funzionalità di gestione del ciclo di vita dell'hardware: provisioning, monitoraggio, aggiornamento firmware e gestione remota.	<b>6</b>
T4	Dimensioni fisiche	Ingombro in unità rack e compatibilità con gli spazi e i vincoli del data center.	<b>7</b>
T5	Alimentazione	Efficienza energetica, consumo elettrico e ridondanza degli alimentatori.	<b>4</b>
T6	Prestazioni e benchmark	Risultati prestazionali documentati su benchmark standard di riferimento per i carichi di lavoro previsti.	<b>5</b>
T7	Container & software	Supporto a tecnologie di containerizzazione/virtualizzazione e compatibilità con lo stack software richiesto.	<b>5</b>
T8	Installazione & collaudo	Modalità, tempi, completezza e qualità delle attività di installazione, configurazione e collaudo presso la sede.	<b>4</b>
T9	Documentazione	Completezza e qualità della documentazione tecnica e operativa fornita a corredo della fornitura.	<b>2</b>
T10	Tempi di consegna	Tempistiche di consegna offerte rispetto ai termini richiesti dal capitolato.	<b>7</b>
T11	Tempi di garanzia (on-site)	Durata della garanzia comprensiva di assistenza on-site inclusa nella fornitura.	<b>3</b>
T12	Supporto on-demand	Disponibilità e modalità del supporto tecnico su richiesta: tempi di risposta, canali e copertura oraria.	<b>3</b>
T13	Esperienza documentata del fornitore in ambito HPC genomico/biomedico	Referenze e progetti pregressi del fornitore in contesti HPC applicati alla genomica e alla biomedicina.	<b>2</b>
		<b>Totale</b>	<b>70 pt</b>

Roma,

Il Responsabile Unico del Progetto  
 Ing. Andrea Proietti Pannunzi