Technical Specifications

SECTION 'C'

Supply, installation and commissioning of High Performance Computing System at IPR-Bhat, Gandhinagar

---

NOTE: Post pre-bid meeting revisions in the document are in emphasised in **BOLD** font

Scope:

Two part tender for Supply, Installation and commissioning of High Performance Computing System at IPR-Bhat, Gandhinagar as per technical specifications, installation of software, benchmarking of acceptance criteria as per terms and condition mentioned in tender document. The system will be installed in the existing racks in data center at IPR. Pre dispatch inspection will be carried out at the bidder's premises prior to dispatch of system to IPR.

Sr.No	Description
1	Sufficient number of compute nodes must be quoted to achieve at least 724 **TF CPU** theoretical peak performance.
2	Sufficient number of compute nodes with **GPU** must be quoted to achieve at least **67 TF CPU and at least 206 TF GPU** theoretical peak Performance.
3	The CPU to be used for Head Nodes, Login Nodes, CPU Nodes, CPU+GPU Nodes, compute nodes with high memory must be identical in technical specification, with Intel Scalable 64 bit Processor, each core capable of performing 32 Double Precision Floating Operations per cycle.
4	OEM refers to Original Equipment Manufacturer for Master/Head Nodes, Login Nodes, Compute Nodes, compute Nodes with GPU, compute nodes with high memory, network fabric and compatible storage solution.
5	Make of Master/Head Nodes, Login Nodes, Compute Nodes, compute Nodes with GPU,compute nodes with high memory will be preferred from same OEM.
6	HPC system will be installed in existing racks in Data Center at IPR. Total Six Racks are available, with cooling capacity of approx 30 KW per rack. Out of six racks, four racks have usable height 41 U and Usable depth 812mm and two racks have usable height 39U and usable depth 812mm. Each rack has two Power Distribution Units(PDU). Each PDU has 24 x IEC 320 socket C13 and 6 x IEC 320 socket C19. Five Racks have three phase power input and Sixth rack has single phase power input. **Solution offered must be setup in available infrastructure only**. All power sockets are industry standard IEC Sockets. Bidder must visit the site to get details about available infrastructure before bidding (For details please refer the Rack Layout).
7	Single infiniband 4xEDR Switch will be quoted to accommodate all the nodes.
8	Bidder should provide all necessary Cables, Connectors, optical drives, Components (Ethernet Cables, Infiniband Cables, KVM cables, necessary sliding rail kits and cable managers etc.) required for quoted system.
9	2 PetaByte Storage configured as single file system will be quoted.
10	**OEM developed or OEM supported cluster Management tools with latest version should be provided.**
11	Bidder should provide Block Diagram of Quoted System clearly showing all configured hardware devices including the Network connectivity.
12	Technical compliance letter, detailed cluster diagram along with datasheets of proposed solution should be inevitably submitted along with the Technical bid/offer.

13 Any component, which is not explicitly mentioned, but is critical for installation, commissioning, required for completing the entire turnkey HPC solution and for utilisation of the entire turnkey HPC solution, must also be offered in the solution.

14 Entire system must have comprehensive on-site support and warranty for 3 years. This is applied to all Head/Master Nodes, Login Nodes, compute nodes, Compute nodes with GPU, compute node with high memory, visualisation server, storage solution, Infiniband Switches, Ethernet Switches, Networking components and other supplied accessories with HPC System.

15 Appropriate number of commercial licenses with 3 years support/upgrade will be supplied.

Sr.No		Specification
A		**Hardware specifications**
1	**Compute Node Specifications with CPU**	
1.1	CPU	Two Number of 20 core, 64 bit Intel Xeon (2.4 GHz), each core capable of executing 32 (Thirty Two) FLOPS per cycle.
1.2	Memory	384 GB ECC DDR4 2666 MHz.
1.3	Internal Disk	Minimum 2 nos. of 2 TB Enterprise SATA-III disks configured with RAID1
1.4	Form Factor	Rack/blade/dense.
1.5	Infiniband	Single Port 4xEDR with 100Gbps bandwidth with PCI Express x16 3.0 interface.
1.6	Network	2 X 1Gbps ports with PXE boot capability.
1.7	Power supply	The solution should be configured with Hot plug **Redundant Power supplies(N+N)** and fan assemblies.
1.8	Serviceability	All the compute nodes should be individually serviceable without shutting down other compute nodes
1.9	Remote management Port	At least one dedicated port for remote management with KVM over IP, virtual media over LAN, remote power control with required licenses.
2	**Compute Node Specifications with CPU + GPU**	
2.1	CPU	Two Number of 20 core, 64 bit Intel Xeon (2.4 GHz) , each core capable of executing 32 (Thirty Two) FLOPS per cycle.

2.2	Memory	384 GB ECC DDR4 2666 MHz.
2.3	Internal Disk	Minimum 2 nos. of 2 TB each Enterprise SATA-III disks configured with RAID1
2.4	Form Factor	Rack/blade/dense.
2.5	Infiniband	Single Port 4xEDR with 100Gbps bandwidth with PCI Express x16 3.0 interface.
2.6	Network	2 X 1Gbps ports with PXE boot capability.
2.7	Power supply	The solution should be configured with Hot plug **Redundant Power supplies(N+N)** and fan assemblies.
2.8	Serviceability	All the compute nodes should be individually serviceable without shutting down other compute nodes
2.9	Remote management Port	At least one dedicated port for remote management with KVM over IP, virtual media over LAN, remote power control with required licenses.
2.10	GPU enabled	Compute node must have at least 2 GPU slots.
2.11	GPUs	The compute node should be configured with Two numbers of Nvidia P100 GPU PCIe with 16GB RAM.
3	**Compute Node Specifications for Visualisation Server :Qty-01**	
3.1	CPU	Two Number of 20 core, 64 bit Intel Xeon (2.4 GHz) , each core capable of executing 32 (Thirty Two) FLOPS per cycle.
3.2	Memory	At least 384 GB ECC DDR4 2666 MHz.
3.3	Internal Disk	Minimum 2 nos. of 2 TB each Enterprise SATA-III disks configured with RAID1
3.4	Form Factor	Rack/blade/dense.
3.5	Infiniband	Single Port 4xEDR with 100Gbps bandwidth with PCI Express x16 3.0 interface.
3.6	Network	2 X 1Gbps ports with PXE boot capability.
3.7	Power supply	The solution should be configured with Hot plug **Redundant Power supplies(N+N)** and fan assemblies.
3.8	Serviceability	All the compute nodes should be individually serviceable without shutting down other compute nodes
3.9	Remote management Port	At least one dedicated port for remote management with KVM over IP, virtual media over LAN, remote power control with required licenses.

3.10	GPU enabled	Compute node must have at least 2 GPU slots.
3.11	GPUs	The compute node should be configured with Two numbers of Nvidia **P40** GPU with 24 GB cards with required software.
3.12	Software	PBS Works' Display Manager will be installed and configured along with third party software like VisIt, Paraview etc.
4	**Compute Node Specifications with High Memory CPU: Qty-02**	
4.1	CPU	Four Number of 20 core, 64 bit Intel Xeon (2.4 GHz) , each core capable of executing 32 (Thirty Two) FLOPS per cycle.
4.2	Memory	At least 1TB ECC DDR4 2666 MHz.
4.3	Internal Disk	Minimum 2 nos. of 2 TB Enterprise SATA-III disks configured with RAID1
4.4	Form Factor	Maximum **4 U** height with rack/blade/dense form factor.
4.5	Infiniband	Single Port 4xEDR with 100Gbps bandwidth with PCI Express x16 3.0 interface.
4.6	Network	2 X 1Gbps ports with PXE boot capability.
4.7	Power supply	The solution should be configured with Hot plug **Redundant Power supplies(N+N)** and fan assemblies.
4.8	Serviceability	The compute node should be individually serviceable without shutting down other compute nodes
4.9	Remote management Port	At least one dedicated port for remote management with KVM over IP, virtual media over LAN, remote power control with required licenses.
5	**Master /Head Nodes**	
5.1	CPU	Two Number of 20 core, 64 bit Intel Xeon (2.4 GHz) , each core with capable of executing 32 (Thirty Two) FLOPS per cycle.
5.2	Memory	384 GB ECC DDR4 2666 MHz.
5.3	Internal Disk	2 X 1 TB (or higher capacity disk) SAS with 10000 RPM disks configured with hardware RAID1.
5.4	Form Factor	Maximum 2 U height with rack/blade/dense form factor.
5.5	Infiniband	Dual Port 4xEDR with 100Gbps bandwidth with PCI Express x16 3.0 interface.

5.6	Network	4 X 1Gbps ports with PXE boot capability.
5.7	Power supply	The solution should be configured with Hot plug **Redundant Power supplies(N+N)**.
5.8	Serviceability	The compute nodes should be individually serviceable without shutting down other compute nodes
5.9	Remote management Port	At least one dedicated port for remote management with KVM over IP, virtual media over LAN, remote power control with required licenses.
5.10	Nodes	Two nodes must be configured in pair to provide 100% redundancy.
6	**Login Nodes**	
6.1	CPU	Two Number of 20 core, 64 bit Intel Xeon (2.4 GHz) , each core with capable of executing 32 (Thirty Two) FLOPS per cycle.
6.2	Memory	384 GB ECC DDR4 2666 MHz.
6.3	Internal Disk	2 X 1 TB (or higher capacity disk) SAS with 10000 RPM disks configured with hardware RAID1.
6.4	Form Factor	Maximum 2 U height with rack/blade/dense form factor.
6.5	Infiniband	Dual Port 4xEDR with 100Gbps bandwidth with PCI Express x16 3.0 interface.
6.6	Network	4 X 1Gbps ports with PXE boot capability.
6.7	Power supply	The solution should be configured with Hot plug **Redundant Power supplies(N+N)** .
6.8	Serviceability	The compute nodes should be individually serviceable without shutting down other compute nodes
6.9	Remote management Port	At least one dedicated port for remote management with KVM over IP, virtual media over LAN, remote power control with required licenses.
6.10	Nodes	Two nodes must be configured in pair to provide 100% redundancy.
6.11	GPU enabled	Nodes must be GPU enabled. Node must have at least 2 GPU slots.
6.12	GPUs	One number of Nvidia p100 GPU PCIe with 16GB RAM in each node.

7	**Networking : Primary Interconnect – Infiniband 4xEDR Switch**	
7.1		Single 4xEDR Infiniband switch with each port having bandwidth of 100Gbps appropriate for HPC system will be supplied.
7.2		Connectivity will be 100% non blocking with FAT Tree architecture between nodes.
7.3		Redundant (N+N) Power Supplies and FANS should be offered in the switch configuration.
7.4		Appropriate numbers of IB spine, leaf modules populated as per requirement will be supplied by bidder.
7.5		Appropriate number and length of IB cables, connectors etc. accessories as per requirement will be supplied by bidder.
7.6		Management module with redundancy and software for Infiniband and all required licenses.
7.7		Optionally quote for one leaf module.
7.8		Bill of material of Infiniband solution must be clearly specified by the bidder in technical bid.
8	**Networking : Admin and Console Network**	
8.1		**Appropriate required number of GigE L2 managed network switches will be used for (a) administration network (Ethernet) (b) IPMI network. Each network / each node should be managed and monitor separately. The number of switches has to be quoted accordingly.**
8.2		All switches with redundant power supply and fan assemblies to be supplied.
8.3		Appropriate number of factory crimped CAT6 cables with suitable length to be supplied.
8.4		These switches will be used for cluster connectivity, management and monitoring.

9	**Storage Solution with Parallel File System**	
9.1		2 PB(Peta Byte) usable capacity with RAID 6 (with 8D+2P and atleast 1 spare per 30 drives) should be configured with parallel file system as single file system.
9.2		As per requirement parallel file system Metadata will be configured on SSDs **with at least 2% of usable capacity.**
9.3		Storage Solution must be connected to the cluster's Primary Interconnect infiniband switch of 100 Gbps port.
9.4		The storage should be configured with at-least 20 GB/s WRITE , 20 GB/s READ performance individually.
9.5		Read performance should not be less than write performance.
9.6		Storage must be configured with commercial, licensed and OEM supported parallel file system, intel lustre or GPFS.
9.7		Storage will be configured with SSDs or 10K RPM SAS or 7200 RPM NL-SAS Drives.
9.8		Storage should be with minimum two hardware controllers running in active-active mode in high availability pair mode with automatic failover to each other.
9.9		Offered solution must have no single point of failure in the entire storage solution.
9.10		Solution must be fully compatible with quoted system.
9.11		Web based software to monitor health of storage may be provided.
9.12		Solution must have IOR / IOZone individual performance test of atleast 20GB/s write, 20GB/s read and Opensource MDTEST tool performance of minimum 60,000 files create/sec **with 1 MB block size.**
9.13		Total Solution in 16 U Size will be preferred. Maximum available size is 24 U.
10	**Management Console**	
10.1		1U Rack mount LCD Keyboard and mouse with required cables and accessories.

10.2		8port KVM SWITCH with cables compatible with the supplied system.
10.3		All the nodes should be visible through IPMI connectivity in the console.
B		**Software specifications**
1	**Operating system**	Latest version of 64 bit Red Hat enterprise Linux. Licenses must be provided for all the nodes. (Head Node/Master nodes, Login Nodes, Compute nodes, Compute Nodes with GPU, Compute node with high memory, compute node with Visualisation Server, Storage Solution and other, if required.)
2	**JOB Scheduler**	Latest version of Suitable PBS Works' PBS pro for all servers . It must support GPU jobs also. Appropriate license must be provided. **Full support should be provided during the warranty period.**
3	**Cluster Resource Utilisation Software**	Latest version of PBS Works' PBS Analytics with Two concurrent User license.
4	**Cluster Management Software**	**Cluster Management Software**
4.1		Management software with appropriate number of licenses must be from same OEM or OEM supported licensed management software..
4.2		Software for management must support Linux operating system for provisioning, monitoring and management of HPC cluster hardware like CPU, GPU, RAM etc.
4.3		Software should manage all the nodes provided in the solution.
4.4		It should support Web Based access.
4.5		It should provide proactive notifications alerts.
4.6		It should have preferably a Dashboard showing health of HPC.
5	**Compiler**	Latest version of atleast five user floating license of **Intel Parallel Studio Cluster Edition**
6	**Parallel File System Software**	Licensed Intel Sourced Lustre or GPFS parallel file system suitable for proposed solution.

7	**Visualisation Software**	Latest version of PBS works' Display Manager with five concurrent users.
8		Appropriate number of Licenses for all above software along with patches, upgrades etc. with 3 years subscription and support will be supplied and installed. **The Bidder should submit subscription and support certificate for all software mentioned in this tender document from respective OEMs.**
9		Necessary tools for parallel programming on CPU and GPU like Open MP,MPI,MPICH, Open ACC, Open CL, CUDA, GPU libraries should be supplied and installed. **PGI compiler with atleast 2 user floating license should be supplied and installed.**
10		Supplied compliers and libraries must be suitable to offered solution.
C		**Terms and Conditions**
1		**Pre Despatch Inspection (PDI)**
1.1		Pre-dispatch inspection will be carried out at the **Bidder's/OEM** premises prior to dispatch of system to IPR.
1.2		The Bidder has to keep in ready-to-use state all the nodes with required Linux operating system, compilers, GPU drivers and other system management software along with parallel file system storage and Infiniband switch as per tender document.
1.3		The Bidder has to provide head nodes, login nodes and at least 25% of the total compute nodes with CPU and Compute node with CPU+GPU. IPR personnel will select randomly from these ready to use nodes for benchmarking.
1.4		IPR personnel will perform verification of the hardware as per tender document.
1.5		Tests to be conducted by the Bidder in the presence of the personnel from IPR and the log reports of each of the machines should be supplied along with the machines.

1.6		HPL benchmark will be run for a period of SIX hours, during which the performance of the system should be same or better as mentioned in Acceptance Criteria of tender document. Systems that show degradation in performance or any other problem will have to be rectified by the Bidder immediately and the procedure repeated until the required performance as per acceptance criteria indicated in the tender document is achieved.
1.7		**High performance Linpack (HPL) efficiency must be at least 60% for pure CPU nodes with Turbo OFF as well as for CPU+GPU Nodes with Turbo OFF.**
1.8		A sample MPI program provided by IPR must run using 20 or more compute nodes with batch queue systems for a period of SIX hours at least.
1.9		For storage solution the performance will be evaluated as specified in the Acceptance criteria as per tender document. Systems that show degradation in performance or any other problem will have to be rectified by the Bidder immediately and the procedure repeated until the required performance as per acceptance criteria indicated in the tender document is achieved.
1.10		Input voltage variation from **180V – 240V** must be conducted on sample set and the machine operation should not get affected. The Bidder has to arrange the required equipments for this testing.
1.11		Bidder should indicate the location of the premises where pre-dispatch inspection can be carried out and the provide infrastructure (electrical power, air-conditioning, floor space etc.) at the same premises.

2	**System Delivery at IPR**	
2.1		HPC system will be delivered at IPR-Bhat within **4 weeks**, after Pre-dispatch inspection is completed and on receipt of despatch clearance letter from IPR
2.2		OEMs have to supply their latest make/model of hardware devices compatible with proposed solution. OEMs should not supply any Hardware, Software and other related devices which are nearing its End-of-Life criteria, even though they may not have been declared at the time of delivery.
3	**Installation and commissioning**	
3.1		Installation, commissioning and acceptance of the HPC system should be completed within 4 weeks from **the date of delivery at IPR-Bhat**.
3.2		Installation and commissioning will be discussed with IPR Team and executed accordingly.
3.3		After installation and commissioning, during performance test, if it is found that, to achieve acceptance criteria as per tender document, additional hardware/software is required, then Bidder/OEM will be responsible for supplying additional hardware/software at no extra cost to IPR.
4	**Implementation Schedule**	
4.1		Bidder will provide complete implementation schedule of the system along with the bid.
4.2		Pre-dispatch inspection, complete scope of supply, installation, commissioning should be completed within 18 weeks from date of contract or Letter of Intent.
5	**Application Installation/Porting / Migration :**	
5.1		Bidder/OEM should provide complete technical help for migration of existing applications to the HPC system.

5.2		All standard libraries / compilers /packages such as ATLAS, LAPACK, SCALAPACK, BLAS, G-FORTRAN, GCC, NetCDF,HDF5, Python,C, C++, openmpi, FFTW, PETSc,PGI, LAMPPS, VisIt, paraview, with compatible version with the HPC system will be installed by the Bidder/OEM.
5.3		Software will be installed as modules as per IPR requirements.
6	**Training**	
6.1		A Three-day HPC –Cluster System Administration and maintenance (Cluster Management, Configuration Monitoring, Node Provisioning, Job Scheduling etc.) training will be scheduled at IPR for HPC Admin Team.
6.2		A Two-day HPC- Users training will be scheduled at IPR for HPC users, for getting familiarity with HPC software stack, job submission and monitoring.
6.3		IPR may request for additional training as per requirement during the warranty period.
7	**Documentation**	
7.1		Documentation on installation and commissioning of the system will be prepared and provided to IPR in soft and hard form.
7.2		Acceptance Test Procedure (ATP) documents will be prepared by Bidder/OEM and submit to IPR as mentioned in Acceptance Criteria as per tender document.
8	**Support**	
8.1		Bidder/OEM will provide for technical support for administration / maintenance as and when required (at both software and hardware level) for HPC during the period of warranty.
8.2		Bidder/OEM will also provide support for installation of third party software as and when required during warranty period.

8.3		If the system requires reinstallation / re-configuration, then the Bidder/OEM will do it without any additional cost to IPR during the warranty period.
8.4		Whenever part replacement is required, then it will be provided latest by next business day onsite at no additional cost to IPR. New part will be brought in to IPR and faulty will be taken back from IPR. **If part is HDD, it may be returned after degaussing.**
8.5		During warranty, if part replacement is required, and same is not available, then it will be replaced with compatible, equivalent or higher rating component.
8.6		The Bidder must have well-established service and support facility at Ahmedabad/Gandhinagar in operation at least for last 3 years or more from date of this tender notice. Company/Office registration document to be submitted. Additionally details like facilities, manpower etc. at their office shall be submitted.
8.7		If the Bidder does not have well-established service and support facility at Ahmedabad / Gandhinagar, then they must have office in India in operation at least for last 3 years from date of publication of this tender notice. Additionally details like facilities, manpower etc. at their office shall be submitted. In this case, **the bidder** will depute at least two number of Engineers at IPR with HPC experience in the area of HPC configuration & operations, at their own cost, during warranty period.
9	**Warranty**	
9.1		Three years comprehensive on site warranty on all the items supplied as per tender document will be provided by the Bidder/OEM.
9.2		During warranty, whenever part replacement is required then, no refurbished parts will be supplied.

9.3		Three years subscription of licenses will be supplied for all the software mentioned in software list as per tender document.
9.4		All software licenses should be perpetual.
9.5		Warranty will start from the date on which IPR accepts the system.
10		**Site Visit**
10.1		Bidder must visit the site at IPR before bidding, to get details about available infrastructure for HPC system installation.
10.2		IPR reserves the right to visit the premises of the Bidder in order to verify the submitted details at any point of time, if required.
11		**Bill of Material**
11.1		Detailed Bill of Material (Un-priced) with part number of entire HPC system should be uploaded along with the Technical Bid, as per tender document.
D		**Acceptance Criteria**
1		Formula to achieve at least **724** TF for CPU Nodes and atleast**67** TF CPU and atleast**206** TF GPU from GPU Nodes must be demonstrated.
2		Master/Head Nodes, Login Nodes, Compute Nodes, compute Nodes with GPU, compute nodes with high memory preferably from same OEM.
3		Processor of Master/Head Nodes, Login Nodes, Compute Nodes, compute Nodes with GPU, compute nodes with high memory must be identical.
4		Head Node, Login Node, Compute Node, Compute Node with GPU, compute node with high memory, visualisation server, Storage Solution should be connected to infiniband 4xEDR 100 Gbps interconnect Switch.
5		In all the servers memory slots must be filled with appropriate size of memory modules, to get maximum performance.

6		**Bidder must demonstrate following performance after installation along with method of arriving at the result as below :**
6.1		High performance Linpack (HPL) must be at least 60% for pure CPU nodes with Turbo OFF. The HPL test will be carried continuously for 72 hours to check hardware consistency prior to acceptance.
6.2		High performance Linpack (HPL) must be at least 60% for CPU +GPU nodes with Turbo OFF. The HPL test will be carried continuously for 72 hours to check hardware consistency prior to acceptance.
6.3		IOR /IO Zone benchmark for storage with at least 20 GB/s Write, 20 GB/s read individually. The test will be carried continuously for 72 hours to check hardware consistency prior to acceptance.
6.4		Open source MDTEST tool performance of minimum 60,000 files create/sec.
6.5		Storage benchmark must show read performance not less than write performance.
7		Total power consumption, heat dissipation and cooling requirements will be tested as per document submitted in mandatory document section.