Annexure -II

PART-2
Technical Specifications of the proposed

Integrated special purpose computing facility for Gravitational Wave research.

IUCAA wants to expand the existing 227 TF (Peak performance) CPU based High Performance Computing facility that uses 10G Ethernet as an Interconnect. The proposed system should be scalable using the same architecture and cooling method available at the site. The current uplink bandwidth is inadequate. Therefore 32 port 100 G ethernet core switch having the switching capacity of 6.4Tb/s with a landmark 4.77 bpps processing capacity in a compact 1RU form factor to be proposed along with required number of 25G switches to add additional 3680 CPU cores to the cluster.

A. Composition of Cluster nodes

The proposed solution should have one compilation node, one graphics node and 92 compute nodes customized for Gravitational wave research. Each note must be individually serviceable. Please note that, the cost of optional items will not be included while identifying a L1 vendor. The bidder is expected to submit the pictorial layout with interconnect and so on of the proposed solution along with relevant product brochures. The bidder is expected to submit the explicit Bill of Material (BOM), and bids with incomplete BOM will not be considered further.

Abbreviations used:

RDIMM	Registered Dual In-line Memory Module
NSPOF	No Single Point of Failure
AMC	Annual Maintenance Contract
IPMI	Intelligent Platform Management Interface
DDR4	Dual Data Rate 4
SLES	Scientific Linux Enterprise Server
HPC	High Performance Computing
OEM	Original Equipment Manufacturer
CPU	Central Processing Unit
GUI	Graphic User Interface
CLI	Command Line Interface
HPL	High-Performance Linpack Benchmark
CMS	Cluster Management Software
ECC	Error Correcting Code
RPM	Rotation Per Minute
RFP	Request for Proposal
PFS	Parallel File System
SFF	Small Form Factor
TF	TeraFlops

1. Specifications of the compute nodes: 92 numbers

- Each node should have 2 nos of Intel Xeon G-6148 (20 cores @ 2.4GHz or better).
- Each node should be configured with 192 GB memory module ECC DDR4 RDIMMs (2666 MHz or better). Each node should be configured with minimum 12 nos of dual rank DDR4 DIMMs. The node should have 16 DIMMs support for future expansion.
- Each node should be configured with 2 x 25G Ethernet ports.
- Each node should be configured with minimum 2 * 1G Ethernet ports.
- Each node should be configured with at least one free slot for an SSD and 1x 1TB @ 7200 RPM SATA / NL-SAS disk.

2. Specifications of the Compilation Node: 1 number

- Each node should have 2 nos of Intel Xeon G-6148 (20 cores @ 2.4GHz or better)
- Each node should be configured with 192 GB memory module ECC DDR4 RDIMMs (2666 MHz or better). Each node should be configured with minimum 12 nos of dual rank DDR4 DIMMs. The node should have 16 DIMMs support for future expansion.
- Each node should be configured with 2 x 25 G Ethernet ports.
- Up to 2U rack-mountable form-factor only. **1U server is the preferred option.**
- It should have USB ports to attach external CD-DVD reader/writer or it could come integrated with an internal CD-DVD reader/writer.
- It should be configured with 4 x 1Gbps RJ45 and dual port 10G HBA
- It should be configured with SFF 2 X 600 GB @10K RPM SAS disks.

3. Specifications of the Graphics Utility Node: 1 number up to 2U form factor.

- The Graphics nodes should have the same hardware and configuration as the compilation node. Additionally, it should have one 1 TB SATA disk and one NVIDIA Quadro P40 GPU PCIE module or better for visualization and the required software for enabling remote user to visualize ie. make use of the features of its graphics card remotely.

4. Features required in the Chassis:

- The Chassis should have hot swappable redundant power supplies with at least N+1 redundancy.
- The Chassis should have redundant fans with at least N+1 redundancy.
- The power consumption for one chassis with all nodes populated with proposed CPU or equivalent wattage CPU, along with the specifications of each power supply, should be reported along with the technical Bid.

5. Additional Features required in Cluster

- IPMI2.0 or equivalent Support with KVM and media over LAN features including licenses, if any.
- The entire compute cluster hardware solution must be based on NSPOF configuration except for the interconnects.

B. Network connectivity:

1. **Primary Interconnect**

   o The primary interconnect should support one 25 G Ethernet for each compute node, compilation and graphics node. The proposed 25G switches should have the support for at least four 100 G uplink ports to the core switch having thirty two 100 G ports.

   o The proposed 32 port core switch should offer switching capacity of at least 6.4Tb/s with a landmark 4.77 bpps processing capacity in a compact 1 RU form factor. Adequate number of 25G switches should be sized to accommodate proposed compute and other nodes.

   o Appropriate number of 25G DAX copper or fibre cables for the proposed compute nodes along with four 10-15 metre 100 G fibre DAX cables per 25G switch for uplinking to 32 port 100 G switches to be quoted. Apart from the above, four additional 100G DAX copper cables would be needed for the existing 10G switch to uplink with 32 port core switch. Please note that the core switch will be housed adjacent to the existing four 10G 48 port switches.

2. **Console / Management Network**

   o IPMI based Console Network for management, should be on different physical switches and should not be clubbed in the Primary Interconnect Enclosure or Switch.

   o All nodes need to be connected by a dedicated Gigabit network for remote management hardware.

   o Managed GigE L2 Switches should be offered with suitable number of cables for the Console Network. Quantity of switches and cables and other required accessories to be provided as per proposed solution requirement. Please use the required patch cords of appropriate length for this network.

   o Advanced license for IPMI based management and monitoring of all nodes should be quoted.

C. Cluster management Software

1. Commercially available licensed Cluster Management Software (CMS) preferably from the same hardware OEM should be offered for provisioning and managing all the compute nodes in the Cluster.

2. CMS is required to manage the complete cluster. It should have the ability to verify and ensure consistency in hardware and system settings across the Cluster's resources from a single console. CMS must support the following:
   i. It must be a GUI/Web-based tool accessible from any client.
   ii. It should provide a single interface for management and control of the complete cluster.
   iii. It should provide the facility to dynamically add, remove or configure any individual node.
   iv. It should provide for remote booting/resetting of individual or group of nodes.
   v. It should provide for monitoring of vital parameters and provide predictive failure analysis and alarms.

vi.   The supplied Cluster Management Software Suite must be a commercially licensed product with latest version. All necessary software device drivers are to be provided.

3. The tool should have support for GUI/CLI interface.

4. Should support provisioning of stable latest version of SLES.

5. Following features should be supported by the CMS,
   i.   Managing as many different images as needed for different software stacks, different operating systems, or different hardware.
   ii.   Cloning from one to many nodes at a time with a scalable algorithm which is reliable and does not stop the entire cloning process if any nodes are broken.
   iii.   Replicating available images on any number of compute nodes in the cluster.
   iv.   Customizing reconfiguration scripts associated with each image to execute specific tasks on compute nodes after cloning.

D. Software Stack: All the software quoted must be licensed and perpetual.

E. Operating System:
The entire solution should support the latest version of SLES 7 Operating System.

F. Optional Items to be quoted with 5 year comprehensive on-site warranty:

1. Cost of single compilation node
2. Cost of single compilation node with 32GB, 64GB and 96GB RAM respectively instead of 192 GB RAM.
3. Cost of single compilation node with 32GB, 64GB and 96GB RAM instead of 192 GB RAM and processor 2 x Intel Xeon G-5118 instead of 2 x Intel Xeon G-6148.
4. Cost of single graphics utility node.
5. Cost of proposed fully populated compute nodes in a single chassis.
6. Cost of 384 GB dual rank RAM in place of 192 GB dual rank RAM.
7. Cost of 2 TB SATA disk in place of 1TB SATA disk.
8. Cost of minimum 200 GB SSD for compute nodes
9. Cost of minimum 400 GB SSD for Compute nodes
10. Cost of Intel compiler for a single node license.
11. Cost of proposed 25G switch
12. Cost of proposed 32 port core switch
13. Cost of 100G Copper DAX cable (3metres)
14. Cost of 100G fibre DAX cable (10 metres)
15. Cost of 100G fibre DAX cable (15 metres)
16. Cost of 25G DAX copper cable (3 metres)
17. Cost of 25G DAX copper cable (5 metres)
18. Cost of 40G DAX copper cable (3 metres)
19. Cost of two GPU nodes up to 2U form factor having following specifications:
    o   Each node should have 2 nos of Intel Xeon G-6148 (20 cores @ 2.4GHz or better) and 1 no of Nvidia Tesla V100 GPU-32GB PCIE module.
    o   The rest of the configuration are identical to the compilation node.
20. Cost of any other multiple GPU in a single enclosure configuration (upto 6U Chassis) with

sufficient memory and CPUs that can be integrated in this facility with appropriate network uplink connectivity proportional to the number of nodes present. The corresponding peak performance of such configuration must be included.

PART 3

A. Warranty and support:

All the equipments and components supplied must have 5 (Five) years onsite comprehensive warranty with 24x7 support along with 4 hours response time and resolution time of Next Business Day (NBD), covering all parts & labor starting from the date after the successful installation, demonstration of performances and acceptance by IUCAA. The defects, if any, during the warranty period are to be rectified free of charge by arranging free replacement wherever necessary. It should be completed within Next Business Day resolution for entire solutions after the intimation of fault.

B. Installation and Support:

1. The OEM should do the integration/installation of the cluster at site.

2. The entire system should be integrated on site in consultation with IUCAA team as per the agreed architecture.

3. During the warranty period, OEM will have to undertake comprehensive maintenance of the entire hardware and its components.

4. Quarterly review of HPC cluster health and its report should be submitted within a week.

5. Half yearly review of various firmware related to hardware components should be done.

5. Once a year upgrading of required firmware to maintain stability of various components should be done.

6. Upgrading of firmware if necessary on the basis of criticality apart from the above schedule.

7. HPC cluster implementation should be monitored by a dedicated project manager of the OEM for smooth implementation.

8. OEM should provide onsite support during maintenance window such as DC shutdown, power outage and firmware upgrade.

9. Installation/configuration and upgradation of HPC cluster activities should be carried out by direct OEM engineers only.

C. <u>**Acceptance Test Procedure:**</u>

1. **Inventory check:** All the hardware and software will be checked against the Purchase Order.

2. **Functional Test:** All the functionalities of the proposed cluster will be tested including the connectivity, with existing storage system, entire suite of condor manager, cluster management tool etc.

3. **Performance Test:** HPL ratings (peak & sustained) for entire cluster configuration should be demonstrated after installation at site. Sustained HPL efficiency of the installed solution for 16 nodes should be more than 60% of the offered theoretical peak performance.

4. **Training:** The OEM should give 2 days System Administration training to a group of IUCAA personnel on installed hardware (Compute/storage/interconnect), operating system, installed system software and development tools including API. The training must be arranged at IUCAA.

5. **Documentation:** Documentations should be submitted for the following:

    i. Procedure for bringing up and shutting down the fully integrated cluster.
    ii. Procedure to get user accounting for Compute nodes.
    iii. Procedure for basic troubleshooting of Compute nodes, Compilation and Graphic nodes
    iv. Step by step installation guide for the entire HPC implementation/configuration from scratch.
    v. Project documentation listing hardware/software with serial numbers, configuration and connectivity.
    vi. Any other document/manual useful for daily administration.