Purchasing Services Office
801 Leroy Place
Socorro, NM 87801
(575) 835-5881

Letter of Addendum

TO: All Offerors

FROM: Kimela Miller, CPO

DATE: 12/01/2023

RE: RFP Number: RFP# 2411019C, Amendment No. 3
Commodity: High-Performance Computer System

Please note:

This amendment is issued to provide a list of questions which have been received by NM Tech along with our response.

Q1)   HDD storage File- if there are two devices are they replicating to each other?
A1)   No. the plan was to have open slots for easier incremental expansion.

Q2)   HDD file server : Can they perform this backup function to a NAS protocol device?
A2)   As long as the device(s) is mountable as filesystem and can be managed/backed up efficiently to the cloud and/or other devices (preferably without an additional server to mount and run the backup software).

Q3)   Reference RFP page 14, row 2 of the table: HDD Storage(File) Server - > Is "648Tb" the total capacity of the 2 HDD servers together, or each HDD server is 648Tb capacity? Additionally, please can you specify the usable capacity and the RAID level required here.
A3)   Total Capacity. We would prefer Raid 6 + 2 HS

Q4)   Equipment Location and Existing Resource. What are the considerations for inter rack cabling and cable guidelines between the equipment?

A4)   We do have top-of-rack cable management but it is limited. We would like to limit the cabling to only that necessary. One reason we prefer the fiber InfiniBand is they are smaller. We realize InfiniBand and networking/remote management, etc. will need to interconnect between the racks.

Q5)   Equipment Location and Existing Resource. Can you share the current Rack dimensions?
A5)   Rack specifications are provided in a previous addendum.

Q6)   Equipment location and existing resource. Please share the PDU model used on the racks
Is the dual redundant network required for
·   InfiniBand network
·   1G Cluster management network
·   Uplink network to NMT infrastructure
·   Will the HP iLO, Dell iDRAC, CMC or SuperMicro IPMI network and Cluster in-band management network on 1G be shared?
A6)   PDU model is provided in a previous addendum. The infiniband network will need to have redundancy in power, but not connectivity (dual power supplies not dual connectivity). 1G Management network must have redundant power but not connectivity. Uplink must be dual as they connect to redundant switching.
Concept: Either A or B power may be taken down at any time for maintenance so the remaining power must be able to support the devices. Also either switch for the uplink may be taken down for maintenance as well, same reason and why there needs to be redundancy (bonded).

Q7)   Table 1. HPCS Full Component Summary. Are the exact requirements within the RFP to be met? Can alternate solutions be suggested?
A7)   Yes, we encourage submission of alternate solutions, if, for example, some of the requested components are no longer available (such as GPU model, etc) or a newer hardware would provide a better price to performance ratio.

Q8)   GPU Nodes. Currently A100 has supply constraint. Could you share the workloads details so that we may suggest an alternate GPU?
A8)   We have provided some use cases in the RFP, and you may suggest alternative GPUs that provide similar or better overall performance to the A100 and GPU node configurations in the RFP.

Q9)   GPU Nodes. Can a single GPU node be proposed with dual processors and 4 GPUs on the node as an alternate to node with single processor and 2 GPUs?
A9) Yes.

Q10)   Table 1 – cluster management s/w. Are you existing users of Bright cluster management s/w?
A10)   Please assume we have no existing contract that could be leveraged or added to in any way.

Q11)   Table 1 – on site Installation. Is there a full requirement of site installation and deployment?
A11)   Yes, the vendor must provide on-site installation (or have it arranged). At the end of the installation, the customer should be able to run test/evaluation jobs to evaluate the utility of the system.

Q12)   Table – Warranty. Support of next business day or 4hr mission critical is required?
A12)   NBD (Next Business Day) is sufficient.

Q13)   Table – Warranty. Is the solution expected to be like for like with what you have today? Can you share Vizio of your current set up?
A13)   The warranty is only for the HPC system you supply. Like-for-like for the HPC equipment is the preference.   Better equipment may be acceptable if it doesn't cause any customer costs / significant work to use.

Q14) Table – Warranty. Given the complexity of the solution being requested and holiday season would you consider an extension of the existing submission deadline?
A14)   See previous addendum

Q15)   Table – Warranty. Is the cost of liquid cooling included in the RFP? Will that be an additional cost?
A15)   We utilize hot aisle/cold aisle to cool the environment. The amount of cooling available is air cooled.

Q16)   Table – Warranty. Are there future compute or storage growth requirements available that each response should include?
A16)   We do plan on expanding the number of nodes and space (like adding disks to the HDD solution and additional servers to the parallel filesystem / NVMe solution.

Q17)   Table – Warranty. What is the supporting storage and networking needed to complement this solution and are these components open for consideration in this RFP?
A17)   The external to the solution is not up for consideration. The data storage needed for this is specified in the RFP.

Q18)   Table – Warranty. Please share the capacity of UPS and backup power.
A18)   The power limitations specified within the RFP must be conformed to. The back-end UPS / cooling is all based on those requirements.

Q19)   Table – Warranty. Is an OPEX solution or Infrastructure as a Service an option?
A19)   No. We need to own the equipment. This cannot be a cloud or service solution.

Q20)   Table of Requirements   Per the table, are these suggested guidelines or are they written in stone? Are you able to share all HW specs / details that you are looking for the standard compute nodes, large memory workload nodes, AI/ML GPU workload nodes, scientific GPU workload?
A20)   The examples that were provided are exemplars only and not specifications. These are example specifications and will be what we compare the proposal against.

Q21)   Table 1. HPCS Full Component Summary. The example configuration is listed for AMD processor system. Can you share the guidelines/example for an Intel processor based configuration for Compute

nodes, GPU nodes, High Memory nodes and Head nodes as Intel has different core count processors and different balanced memory configuration?

A21)   We will be evaluating the nodes in relation to the examples given.   We have no innate preference between AMD / Intel. At the time we did the examples AMD was less expensive.

Q22)   Table 1. HPCS Full Component Summary. For the internal storage on nodes, can a RI NVME 3.84TB be proposed instead of Mixed Use 3DWPD?

A22)   No, due to expected workloads we need the additional write capabilities for the NVMes.

Q23)   Table 1. HPCS Full Component Summary. Or is 3.2TB Mixed Use 3DWPD acceptable as alternate since 3.84TB 3DWPD NVME is not available?

A23)   Yes, this is fine.

Q24)   Table 1. HPCS Full Component Summary. Can you allow us to use different CPU platforms for CPU and GPU nodes to provide you the best of performance?

A24) Yes, with the condition that the proposal should include some specification for how system users will need to adjust software compilation workflows (i.e., it needs to be done on specific nodes).

Q25)   Table 1. HPCS Full Component Summary. Which workloads/applications will be running on the system/backed up?

A25)   Various. See the "Software" section for some examples, but this is not an exhaustive list.

Q26)   Table 1. HPCS Full Component Summary. Please specify ethernet port requirements? Is it just a management switch that you need for the Idrac? Or what other purpose would the switch have?

A26)   Yes, the Ethernet is XBaseT and is for the management network. Uplink will be dual redundant 1GBaseT. Uplinks to the management server will be dual 10GBaseT.

Q27)   Table 1. HPCS Full Component Summary. How will the new system communicate with the existing clusters?

A27)   This is a standalone cluster that we hope to expand over the years as new funding and capabilities within the hosting environment evolve.

Q28)   Table 1. HPCS Full Component Summary. What is the anticipated timeline for receiving the equipment and completing the deployment?

A28)   As quickly as possible. Delivery time will be considered when evaluating the proposals.

Q29)   Table 1. HPCS Full Component Summary. Are any additional administration node viz. Head nodes, Login nodes, Scheduler nodes etc required for the solution?

A30)   There is a head node specified and it is expected that it would be the login/scheduler. If the solution you specify requires an additional node you may quote it but you must live within the space and power limitations.

Q31)   User equipment (Networking infrastructure and protocols, plugs, etc.) must be compatible with corresponding data center equipment – What is the corresponding data center equipment?

A31)   The PDU was addressed in a previous addendum. The Networking types were specified within this document (and also in a previous addendum).

Q32)   Table 1. HPCS Full Component Summary. Are Cable mgmt arms a hard requirement? We generally recommend against the use of cable management arms on any node with an InfiniBand fabric that uses copper cables. When the arm is folded the cable bend radius is tighter than recommended by Nvidia/Mellanox and can cause stability issues on the fabric. Fiber IB cables are still questionable to use with the management arms but are less likely to have problems if it is an absolute requirement.
A32)   We would prefer Fiber InfiniBand and understand about the bend radii. We really want cable management arms for maintenance and manageability.

Q33)   Table 1. HPCS Full Component Summary. What is the liquid cooled solution, and do you expect vendor to configure this?
A33)   This should be hot/cold aisle cooling.

Q34)   Table 1. HPCS Full Component Summary. Can you please detail the connectivity?
A34)   See previous.

Q35)   Table 1. HPCS Full Component Summary. What are you using TOR for switches?
A35)   We prefer to not give brands but we have redundant 10/40G switches.

Q36)   Table 1. HPCS Full Component Summary. Are you looking for the solution to be scalable in the rack or scalable outside the rack?
A36)   The proposed equipment needs to live within the space provided in the RFP and previous addendums. The system will hopefully be expanded in the future (see A27)

Q37)   Table 1. HPCS Full Component Summary. Mission critical support 4 hrs or NBD do you require onshore resources for installation and configuration?
A37)   NBD is adequate.

Q38)   Table 1. HPCS Full Component Summary. Regarding HDD Storage (File) Server: Can you clarify what "Enterprise class" means? For example: does it mean you want enterprise class components, but NMT staff would be willing to load, admin a filesystem on it to provide 648TB?
A38)   It means we want new enterprise-grade components.   The system should be turnkey so the filesystems should be installed and working with the systems after installation.

Q39)   Table 1. HPCS Full Component Summary. Can you share more requirements on the "backups" use of this storage? For example, will it reside in another datacenter? Is combining both tiers with the NVME storage acceptable?
A39) At this stage, we have not decided on a location for the backup storage, and it may reside in the same or different data center. We are fine with both tiers being combined, as long as it does not hinder the performance of the NVMe-based file server. The HDD file server would also need to have a separate head node, but could be managed under the same system.

Q40)   Table 1. HPCS Full Component Summary. How will users outside of the HPC access the NVME tier of storage?
A40)   They will not. The data will pass through the head-node to be loaded onto the NVMe.

Q41)   Table 1. HPCS Full Component Summary. Does dual network apply for all networks or just uplink?
A41)   The uplink. As long as there are redundant power supplies to the components, that is good enough.

Q42)   Infiniband Switch and Cables. Which version of IB switch is required – HDR/NDR?
A42)   See previous addendum.

Q43)   Infiniband Switch and Cables. Which IB Card is required on the nodes?
A43)   An IB card that is well supported and operates the the full speed of the infiniband switch that will be included in the product offering for this RFP.

Q44)   Infiniband Switch and Cables. Should transceivers be included for the 25Gb/s network ports?
A44)   No. The transceivers should operate with the Infiniband switch contained in the product offering for this RFP.

Q45)   Infiniband Switch and Cables. Can you share the BOM for the switches and cables to be included in installation services? This would help scope and price the network SOW.
A45)   The switches/cables may vary depending on configuration and we do not want to constrain any respondents to a specific solution.

Q46)   Infiniband Switch and Cables. Can vendor optimized linpak binary be used for HPL benchmarking or is it mandatory to compile from netlib?
A46)   Netlib is necessary to determine performance from regular off the shelf options but numbers from vendor optimized binaries will be considered to see what is possible.

Q47)   Infiniband Switch and Cables. Utilizing the previous generation processors might allow for higher density of nodes per rack due to low power consumption. Is this acceptable?
A47)   Yes, as long as the overall number of CPUs, performance, and price does not significantly differ from using the current generation of processors.

Q48)   Infiniband Switch and Cables. Is 68kVA when derated to each system and PSU?
A48)   We don't understand the question. 68KVA is far more than the power budget of 17KW for one rack and 7 KW for the second.

Q49)   Infiniband Switch and Cables. Is redundancy required for PSUs of hardware?
A49) Yes, see previous answers, but equipment must be able to run off either A or B power solely.

Q50)   Infiniband Switch and Cables. The Processor will have primary channel and secondary channels for memory. Do we interpret this as all primary memory channels must be populated with DIMMS no less than 32GB?

A50)   Memory should be configured to provide the best performance in balance with cost for the proposal.

Q51)   HDD Storage (File) Server. What's the RAID level required?

A51)   We prefer Raid 6.

Q52)   HDD Storage (File) Server. The requested number is 2. Does that mean two separate 648TB storages? How was this number deduced?

A52)   See previous questions, but a total of 657 TB. We are hoping for room to expand in the trays.

Q53)   Can copper cables be used for cables?

A53)   Fiber InfiniBand are strongly preferred for cable management both within the rack and between the racks.

Q54)   What is the room ambient temperature?

A54)   The current maximum inlet temp we've seen is 63.3F.

Q55)   Can you provide/describe the type and length of uplink from management switch?

A55)   1GBastT and to the top of rack switches

Q56)   Is the cabling run in overhead trays?

A56)   Cabling will be in the top of rack cable management preferably. Limited space is available, which is why we prefer fiber InfiniBand cables.

Q57)   What file system will be used on backup storage servers?

A57)   Unspecified at this time. The HDD backup will most likely be scp/sftp copies of data sets pre/post processing. If the backup server is backed up it will likely be to the cloud, but this is not the out of the box configuration.

Q58)   Is the 648TB capacity referenced for the HDD Storage (File) Server, the total capacity of the 2 servers combined? Or is that 648TB per server?

A58)   Combined. We hope to have spare slots (see previous answers).

Q59)   I had a question about the rack diagram. Are the spaces marked as reserved for MRI going to be used for the equipment in this RFP?

A59)   Yes.

Q60) For Section 1.5, Table 1, the RFP requests 3.84TB NVMe SSDs (Compute and GPU Nodes), and 15.36TB NVMe SSDs with 3 DWPD. The higher endurance 3DWPD SSDs come as 3.2TB or 6.4TB (for the Compute/GPU Nodes). Would the University like 3.2TB or 6.4TB SSDs? Similarly, the 15.36TB SSDs would

be 12.8TB. Anything larger than this is difficult to find, and we'd need to do a very high capacity SSD and overprovision it to get to 3DWPD. Can the High Memory Nodes use 12.8TB SSDs?
A60) The 3.2TB are fine for the 3.84TB devices. The larger NVMes can be build with multiple drives and the nearest size the respondant can reasonably find.

Q61) In the same table, there is a requirement for Cable Management Arms. Most nodes do not come with nor do they need cable management arms. Can this requirement please be relaxed?
A61) There is a strong preference for Cable Management Arms.

Q62) For storage, do we need to add licenses for filesystems such as BeeGFS? Or will the University procure this separately?
A62) If you are specifying BeeGFS in your design, then all licensing, maintenance and warranties needed should be included in the response.

 Q63) For software (including parallel filesystem and cluster management), would the University like to see a 3 year or 5 year option?
A63) We would prefer to see both options.

All other terms and conditions of the RFP remain unchanged. All Offerors are required to confirm the receipt of this amendment in their RFP response. The proposal due date is December 12, 2023 at 2:00 p.m. local time.

Xc: File