



Central Electronics Engineering Research Institute Pilani (Rajasthan)-333 031,INDIA

Fax : 01596 – 242135,242294

Phone : 01596 – 244710

Gram : “ELECTRONIC” Pilani

No. SPO/EOI/10-Pur/2018/2

Dated: 09.10.2018

EXPRESSION OF INTEREST FOR PROCUREMENT OF “ AI Workload Storage Solution and Interconnect”

Director, CSIR- CEERI, Pilani invites interested manufacturer / authorized distributors only to participate in per-indent conference for procurement of the “**AI Workload Storage Solution and Interconnect**” which is required for working in the area of Artificial Intelligence, deep learning, big data analysis. The high speed performance is required to run deep learning models of the DGX machines. These deep learning models will require large datasets to work on and they require AI workload SAN storage solution. The interconnect will provide the high speed connectivity between DGX server and the storage solution. To cater the needs of the storage for the deep leaning models the high speed storage compliable with DGX cluster is required. Broad specifications of proposed storage solution are as follows:

SN	Features	Specifications
1.	General Features and protocols	<ul style="list-style-type: none">• Hosting Parallel File System.• File sharing via NFS and CIFS.• Shared storage between DGX-1 (training cluster) and Inferencing Servers over IP Switch (RoCE Protocol) / Infiniband• Shared storage for workstation over Ethernet.
2.	Capacity	<ul style="list-style-type: none">• Performance Disks: The solution must be populated with 25TB usable all-flash NVMe/ SSDs Disks with 10GB/s Write & Read performance.• Archival Disks: The solution must be populated with 75 TB usable capacity must provide with higher capacity 7.2K RPM NL-SAS/SATA disks.

3.	Connectivity	<ul style="list-style-type: none"> Storage solution must offer 100 Gbps or Higher Ethernet/ EDR Infiniband connectivity to GPU based computing NVIDIA DGX Cluster and Inferencing engine servers.
4.	Complete Solution	<ul style="list-style-type: none"> Storage solution manufacturer must have at least one reference of similar solution serving Artificial intelligence (A.I) workloads connected for NVIDIA DGX solution.
5.	Performance requirement	<ul style="list-style-type: none"> The bandwidth for 100% Read workload in the above system in Raid 6 should be minimum 10GB/s. The design/ performance sizing document should be shared.
6.	Integration with Clusters [Training cluster (DGX) and Inferencing cluster (GPU-servers)]	<ul style="list-style-type: none"> The storage should be compatible with DGX Cluster and Inferencing Servers Cluster using IB/RoCE Switch
7.	Interconnect	<ul style="list-style-type: none"> High speed 100 Gbps Ethernet of IB switch.

The meeting will be held on 29.10.2018 at 11.00 A.M. in CSIR-CEERI Conference Room. The interested participants should visit CSIR-CEERI, Pilani at their own cost and should bring supporting documents i.e. leaflets/technical literature and also be able to deliver technical presentation of the AI Workload Storage Solution and Interconnect.

Stores & Purchase Officer