RFP forGPU Accelerated Server for DC with 2 NVIDIA 96 GB GPU card **&** General-purpose server with 1 NVIDIA 96 GB GPU cardfor 5 years warranty.

**Technical Specification:**

**Item Table:**

|  |  |  |  |
| --- | --- | --- | --- |
| **SL** | **Item** | **Unit** | **Total** |
| 1 | GPU Accelerated Server for DC with 2 NVIDIA 96 GB GPU card | Nos | 1 |
| 2 | General purpose server with 1 NVIDIA 96 GB GPU card | Nos | 1 |

1. **GPU Accelerated Server:**

| **SL** | **Item Name** | **Detail Required Specification** | **Bidder response** |
| --- | --- | --- | --- |
| 1 | Brand | To be mentioned by the bidder. |  |
| 2 | Certification Quality | ISO 9001:2015 for the manufacturer, CE certification for quality assurance.  The server configuration offered must be Energy Star 4.0 Compliant. |  |
| 3 | Model | Offered server must be purpose build OEM NVIDIA GPU Accelerated Server certified to run |  |
| 4 | Country of origin | To be mentioned by the bidder |  |
| 5 | Manufacturing Country | To be mentioned by the bidder |  |
| 6 | Form Factor | Bider must provide Rack Mountable Server with Rail Kit, Cable Management, and Bezel Kits and other accessories for GPU Accelerated Server for DC. |  |
| 7 | Processors | Bidder should be equipped with a minimum of two Intel Xeon processors, with a minimum of 5th generation or higher, each with a minimum of 32 cores per socket and clock speed 2. 5GHz.or higher. |  |
| 8 | Cache L2/L3 | The bidder must offer high cache capacity, capable of delivering maximum performance, with a minimum of 60MB cache or higher |  |
| 9 | Chipset | To be mentioned by the bidder |  |
| 10 | Memory | The system must be configured with a minimum of 2048 GB DDR5 memory, using 64GB/128GB Dual-Rank modules, operating at a minimum speed of 5600 MT/s |  |
| 11 | Storage array controller | Integrated Hardware RAID controller with 8GB cache supporting RAID 0, 1, 10, 5, 6 supporting SAS/SATA/NVMe devices. |  |
| 12 | Internal Storage | Shall be provided with a minimum of 4x 7.68TB NVMe 2.5-inch hot-pluggable SSD. |  |
| 13 | OS Boot Drive | Shall be provided with minimum 2x 480GB NVMe M.2 SSD in Hardware RAID1. |  |
| 14 | GPU | Bidder must be provided with 2x NVIDIA NVL with fully interconnected technology of 96GB Graphics Accelerator kit with NVIDIA AI Enterprise and NVIDIA Ampere NVLink Bridge and NVIDIA Container technology support & subscriptions with 5 years warranty (All SW & HW).  Must support future upgradable up to minimum 4x NVIDIA NVL GPUs slots must be available from day 1. |  |
| 15 | GPU Framework Compatibility | Must support major ML/AI frameworks: TensorFlow, PyTorch, ONNX, RAPIDS, CUDA, and cuDNN libraries with all necessary license with 5 years warranty. |  |
| 16 | Container & Orchestration Support. | Must support Kubernetes, Docker, and NVIDIA GPU Operator out-of-the-box |  |
| 17 | Implementation & Deployment | The deployment must include NVIDIA (NVIDA) container services and any other GPU-related components necessary to fully enable GPU functionality in containerized environments (e.g., Docker, Kubernetes). Deployment must be performed by the OEM or bidder or an authorized partner of the OEM. The deployment method must follow the best practices recommended by NVIDIA, including installation of. |  |
| 18 | Resource Pool | The solution must support the deployment of multiple physical GPU cards grouped under a unified GPU resource pool. From this shared pool, GPU resources (different vGPU profile) must be dynamically assignable to virtual machines or containers based on workload demand. |  |
| 19 | Power supply and Fan kit | The bidder must provide a fully redundant, hot-swappable, and fully provisioned power supply system, including hot-pluggable fans and Titanium-rated hot-plug power supplies (Minimum 2x 2800W or 4x 1000W). |  |
| 20 | Network Card | Each server should be provided with-   * Minimum 1x 2-port 1G NIC. * Minium 3x 2-port 10/25G SFP28 adapter with   25G SFP28 SR transceivers.   * Minimum 2x 1-port 32/64Gbps FC HBA.   Local bidder must provide CISCO 6 Qnty 10G SFP and CISCO 6 Qnty25G SFP transceiver. |  |
| 21 | Expansion Slot (PCIe) | Shall be provided with minimum 6 PCIe slots with version 5.0 or higher. |  |
| 22 | Embedded Remote Management and firmware security | Server should have a dedicated 1Gbps remote management port with the latest Enterprise license for 5 years warranty. |  |
| Server should have a security dashboard: displaying the status of important security features, the Overall Security Status for the system, and the current configuration for the Security State and Server Configuration Lock features. |  |
| 23 | Firmware security | For firmware security, the system should support remote management chips creating a fingerprint in silicon, preventing servers from booting up unless the firmware matches the fingerprint. This feature should be immutable. |  |
| Should maintain repository for firmware and driver’s recipes to aid rollback or patching of compromised firmware. Should also store Factory Recovery recipe preloaded to rollback to factory tested secured firmware |  |
| 24 | Server Management | Software should come with a dashboard view to quickly scan the managed resources to assess the overall health of the data center. It should provide an at-a-glance visual health summary of the resource user is authorized to view. |  |
| The Server Management Software should be of the same brand as the server supplier.  Should support SIEM integration. |  |
| 25 | System Security | * Trusted Platform Module 2.0 must be provided. * Commercial National Security Algorithms (CNSA) * Secure Recovery - recover critical firmware to known good state on detection of compromised firmware |  |
| 26 | Hypervisor & Operating System Support | Should support the following operating systems and hypervisors   * Microsoft Windows Server * Red Hat Enterprise Linux * SUSE Linux Enterprise Server * VMware ESXi * Canonical Ubuntu * Citrix * Proxmox |  |
| 27 | Warranty & services | * 5 Years collaborative warranty for all HW & SW, NVIDIA GPU card with all NVIDIA necessary SW 24x7 mission critical one-point support from OEM. * 30-minutes response for severity 1 incidents (direct connect to product specialist where available) and 4-hour mission critical response back-to-back with OEM. * OEM must maintain an in-country local part depot center; authorization letters must be submitted. * The OEM manufacturer authorization letter must be submitted. |  |

1. **General purpose Server with NVIDIA GPU card:**

| **SL** | **Item Name** | **Detail Required Specification** | **Bidder response** |
| --- | --- | --- | --- |
| 1 | Brand | To be mentioned by the bidder. |  |
| 2 | Quality | ISO 9001:2015 for the manufacturer, CE certification for quality assurance.  The server configuration offered with Energy Star 4.0 Compliant. |  |
| 3 | Model | Offered server must be purpose build OEM GPU Accelerated Server certified to run |  |
| 4 | Country of origin | To be mentioned by the bidder |  |
| 5 | Manufacturing Country | To be mentioned by the bidder |  |
| 6 | Form Factor | Bider should mention Rack Mountable Server with Rail Kit, Cable Management, and Bezel Kit. |  |
| 7 | Processors | Bidder should be equipped with a minimum of two Intel Xeon processors, with a minimum of 5th generation or higher, each with a minimum of 32 cores per socket and clock speed 2.5GHz or higher. |  |
| 8 | Cache L2/L3 | The bidder must offer high cache capacity, capable of delivering maximum performance, with a minimum of 60MB cache or higher |  |
| 9 | Chipset | To be mentioned by the bidder |  |
| 10 | Memory | The system must be configured with a minimum of 2048 GB DDR5 memory, using 64GB/128GB Dual-Rank modules, operating at a minimum speed of 5600 MT/s |  |
| 11 | Storage array controller | Integrated Hardware RAID controller with 8GB cache supporting RAID 0, 1, 10, 5, 6 supporting SAS/SATA/NVMe devices. |  |
| 12 | Internal Storage | Shall be provided with a minimum of 4x 7.68TB NVMe 2.5-inch hot-pluggable SSD. |  |
| 13 | OS Boot Drive | Shall be provided with minimum 2x 480GB NVMe M.2 SSD in Hardware RAID1. |  |
| 14 | GPU | Bidder must be provided with 1x NVIDIA NVL with fully interconnected technology of 96GB Graphics Accelerator kit with NVIDIA AI Enterprise and NVIDIA Ampere NVLink Bridge and NVIDA Container technology support & subscriptions with 5 years warranty (All SW & HW).  Must support future upgradable up to minimum 1 x NVIDIA NVL GPUs, slots must be available from day1 on the compatible general-purpose server. |  |
| 15 | GPU Framework Compatibility | Should support major ML/AI frameworks: TensorFlow, PyTorch, ONNX, RAPIDS, CUDA, and cuDNN libraries. |  |
| 16 | Container & Orchestration Support | Must support Kubernetes, Docker, and NVIDIA GPU Operator out-of-the-box. |  |
| 17 | Implementation & Deployment | The deployment must include NVIDIA (NVDA) container services and any other GPU-related components necessary to fully enable GPU functionality in containerized environments (e.g., Docker, Kubernetes). Deployment must be performed by the OEM or bidder or an authorized partner of the OEM. The deployment method must follow the best practices recommended by NVIDIA, including installation of. |  |
| 18 | Resource Pool | The solution must support the deployment of multiple physical GPU cards grouped under a unified GPU resource pool. From this shared pool, GPU resources (different vGPU profile) must be dynamically assignable to virtual machines or containers based on workload demand. |  |
| 19 | Power supply and Fan kit | The bidder must provide a fully redundant, hot-swappable, and fully provisioned power supply system, including hot-pluggable fans and Titanium-rated hot-plug power supplies (Minimum 2x 2800W or 4x 1000W). |  |
| 20 | Network Card | Each server should be provided with-   * Minimum 1x 2-port 1G NIC. * Minimum 2x 2-port 10/25G SFP28 adapter with 25G SFP28 SR transceivers. * Minimum 2x 1-port 32/64Gbps FC HBA.   Local bidder must provide CISCO 4 Qnty 10G SFP and CISCO 4 Qnty25G SFP transceiver. |  |
| 21 | Expansion Slot (PCIe) | Shall be provided with minimum 6 PCIe slots with version 5.0 or higher. |  |
| 22 | Embedded Remote Management and firmware security | Server should have a dedicated 1Gbps remote management port with advanced/enterprise license/subscription. |  |
| Server should have a security dashboard: displaying the status of important security features, the Overall Security Status for the system, and the current configuration for the Security State and Server Configuration Lock features. |  |
| 23 | Firmware security | For firmware security, the system should support remote management chip creating a fingerprint in the silicon, preventing servers from booting up unless the firmware matches the fingerprint. This feature should be immutable. |  |
| Should maintain repository for firmware and driver’s recipes to aid rollback or patching of compromised firmware. Should also store Factory Recovery recipe preloaded to rollback to factory tested secured firmware |  |
| 24 | Server Management | Software should come with a dashboard view to quickly scan the managed resources to assess the overall health of the data center. It should provide an at-a-glance visual health summary of the resource’s user is authorized to view. |  |
| The Server Management Software should be of the same brand as of the server supplier. |  |
| 25 | System Security | * Trusted Platform Module 2.0 must be provided. * Commercial National Security Algorithms (CNSA) * Secure Recovery - recover critical firmware to known good state on detection of compromised firmware |  |
| 26 | Hypervisor & Operating System Support | Should support the following operating systems and hypervisors   * Microsoft Windows Server * Red Hat Enterprise Linux * SUSE Linux Enterprise Server * VMware ESXi * Canonical Ubuntu * Citrix * Proxmox |  |
| 27 | Server compatible with NVDIA GPU card | The proposed server must be compatible with the provided NVIDIA GPU card. |  |
| 28 | Warranty & services | 5 Years warranty for all HW (Server and others), SW and NVIDIA GPU card with all NVIDIA necessary mention SW 24x7 mission critical one-point support from OEM.   * 30-minutes response for severity 1 incidents (direct connect to product specialist where available) and 4-hour mission critical response back-to-back with OEM. * The OEM manufacturer authorization letter must be submitted. |  |

1. **Annexure 1 : SLA**

**Annexure 1**

**High level Service Level Agreement (SLA):**

|  |  |  |  |
| --- | --- | --- | --- |
| **Severity Level** | **General Evaluation Criteria** | **Response Time (24/7/365)** | **Resolution Time (24/7/365)** |
| High | A problem that critically effects on ability to do regular business. System functionality is unavailable or down due to hardware or software failure. | By Email/Phone: 01 Hour  Onsite: 02 Hour | 04 Hours |
| Medium | Any Hardware or software component failure that leads to degradation of system performance. | 08 Hours |
| Low | System is having minor issues. Insignificant error without impacting usual business or system performance. | 24 Hours |

**Penalty Clause on Service Failure**

0.5% of the work order value will be penalized for each failure to meet service level.

If resolution time is greater than agreed service level, then 0.025 % (of the work order value) will be penalized for each hour delay for severity level ‘High’ and 0.007 % (of the work order value) will be penalized for each hour delay for severity level ‘Medium’ and ‘Low’.

**Scope of Warranty Service:**

The scope of service is as below but not limited to:

* The product will be delivered as per RFP (All HW, GPU card, GPU enterprise necessary SW and other HW & SW accessories.
* Complete implementation & deployment all HW, SW ,GPU Card (NVIDIA with necessary SW) & Container & Orchestration as per RFP in the Brac Bank premises will be carried out by the bidder along with OEM & local partner.
* Warranty for Server, GPU Card & all necessary software of NVIDIA mention in the RFP for 5 years
* Install, configure in the Server systems as per the requirement of the Bank
* Maintain provided server, GPU card and software 24/7/365 days without any interruption.
* Provide patch, firmware, software upgradation, Security Enhancement, root cause analysis, VA mitigation of all HW & SW as per RFP.
* Pro-active support and notification from OEM or Local Partners for emergency Security fixes and bugs.
* Complete parts replacement of server, GPU card replacement for faulty will be carried out by the bidder along with OEM & local partner.
* Bridge between OEM and BRAC Bank for knowledge transfer, problem troubleshoot and best use of the product. OEM back-to-back support for bidding products will be applicable.
* Ensure new integration of hosts/servers with GPU card as per the requirement of the Bank
* To ensure a single point of contact, backup contact and support contact escalation matrix
* To provide a quarterly report on the health status of the Server system along with performance forecast and performance improvement plan
* To ensure engineer support for any kind of server movement.

1. **Annexure 2: Local partner information for technical evaluation.**

**Annexure 2**

|  |  |
| --- | --- |
| Item | Bidders’ response |
| No of Proposed products sold in the Bank |  |
| No of Similar product sold in the Bank |  |
| No of affiliation with Bank (Warranty/AMC) |  |
| No of engineer for technical support |  |
| No of certified engineer for proposed product OEM |  |

**For Local partner:** Bidder (Local partner) provides the information below.

\*\* All information provided should be validated with supported document

1. Annexure 3: OEM information for technical evaluation.

Annexure 3

|  |  |
| --- | --- |
| Item | OEM response |
| Confirm the local depo in BD |  |
| Confirm Direct support form OEM |  |
| Year of production announcement of proposed product |  |

For OEM: OEM will provide the information

\*\* All information provided should be validated with supported document