

I. Leistungsbeschreibung

KI-Server

Leistungsparameter erfüllt:		ja	nein
1. Systemtopologie und Chassis • Bereitstellung eines Rackmount-Gehäuses im standardisierten 19-Zoll-Format mit einer maximalen Bauhöhe von vier Höheneinheiten (4U) sowie Rackmount-Schienen. • Die Mainboard- und Gehäusearchitektur muss auf eine maximale Skalierbarkeit für GPGPU-Computing ausgelegt sein und physikalisch wie thermisch die Aufnahme von 8 Erweiterungskarten im Dual-Slot-Format (PCIe 5.0) unterstützen. • Implementierung von insgesamt 4 im laufenden Betrieb austauschbaren (Hot-Swap) Netzteilmodulen in einem redundanten 3+1 Setup. Jedes Einzelnetzteil muss eine Nennleistung von mindestens 3200 Watt aufweisen, um die Leistungsaufnahmen der Hochleistungs-GPUs stabil abzusichern. • Integration eines dedizierten kryptographischen Koprozessors gemäß TPM 2.0 Spezifikation.			
2. CPU und Arbeitsspeicher • Integration von zwei x86-64-basierten Serverprozessoren der aktuellen Architektur-Generation im Dual-Socket-Verbund. • Jeder der beiden Prozessoren muss über eine physikalische Kernanzahl von mindestens 96 Cores verfügen. • Eine garantierte Basistaktfrequenz von mindestens 2,6 GHz ist gefordert. • Die Gesamtkapazität des Arbeitsspeichers muss mindestens 768 GB betragen. • Zur Sicherstellung einer maximalen Speicherbandbreite durch Memory-Interleaving ist die Kapazität zwingend auf 12 identische Speichermodule mit einer Kapazität von jeweils 64 GB zu verteilen. • Der Arbeitsspeicher muss auf der DDR5-Technologie basieren und über einen hardwarebasierten Fehlerkorrekturcode (ECC Registered) verfügen, mit einer Speichertaktrate von mindestens 6400 MHz.			
3. Hardwarebeschleuniger für neuronale Netze • Ausstattung mit zwei dedizierten, für den Dauerbetrieb im Rechenzentrum zertifizierten GPGPU-Beschleunigereinheiten, konzipiert für hochparallele Tensor-Operationen. • Jede Beschleunigereinheit muss über einen direkt angebundenen High-Bandwidth-Memory (HBM3e) von mindestens 140 GB verfügen, um die In-Memory-Vorhaltung großparametrischer KI-Modelle zu ermöglichen. • Die logische Anbindung an die Host-Systemarchitektur hat zwingend über den PCIe 5.0 Standard zu erfolgen.			

• Zur Umgehung des PCIe-Flaschenhalses und zur Latenzminimierung beim Datenaustausch zwischen den Beschleunigern ist eine zusätzliche, proprietäre High-Speed-Interconnect-Brücke (2-Way) zur direkten physischen Kopplung der beiden GPUs zu verbauen.		
4. Netzwerk-, Kommunikations- und Speicherschnittstellen • Netzwerkkonnektivität (NIC): Bereitstellung einer dedizierten Netzwerkkarte mit zwei Ports und einer Durchsatzrate von 100 Gbit/s pro Port. Zudem muss eine weitere dedizierte Netzwerkkarte zur Verfügung gestellt werden, welche über zwei RJ45-Ports mit je 10 Gbit/s verfügt. • Basiskonnektivität (Onboard): Das System muss über mindestens einen RJ-45-Netzwerkanschluss für standardisierte Ethernet-Verbindungen sowie über gängige USB-Schnittstellen für Peripherie- oder Diagnosezwecke verfügen. • Speicher-Controller: Integration eines Tri-Mode fähigen Host Bus Adapters (HBA) zur potenziellen Ansteuerung von SAS-, SATA- oder NVMe-Medien. • Hinweis zur Datenspeicherung: Die Bereitstellung von weiteren Speichermedien (SSDs/HDDs) ist explizit nicht Gegenstand dieser Beschaffung. Die intern notwendigen Kabelstränge zur späteren Laufwerksanbindung sind jedoch vollständig vorzuinstallieren.		
5. Qualitätskontrolle und Service-SLA • Das System ist im Vorfeld der Auslieferung durch den Auftragnehmer vollständig zu assemblieren, zu verkabeln und einer strengen Funktionsevaluierung gemäß Herstellerspezifikation zu unterziehen. • Als integrativer Bestandteil des Service-Level-Agreements wird ein Vorabtausch-Service für als defekt identifizierte Systemkomponenten gefordert, um operativen Ausfallzeiten präventiv zu begegnen.		

Kaufmännische Bestandteile

	Leistungsparameter erfüllt:		ja	nein
Festpreis lt. Anschreiben				
Lieferdatum: schnellstmöglich bis spätestens 18.12.2026				
Allgemeine Gewährleistungsdauer von 36 Monaten				
Lieferung frei Haus				
Lieferung eines hardwareseitig assemblierten KI-Servers (ohne Speichermedien)				
CE-Zertifizierung: bestätigen				
Alle netzbetriebenen Geräte sind für die in Deutschland üblichen 230V / 50 Hz Wechselstrom vorgesehen. Alle Geräte und Ausstattungen müssen sämtliche erforderliche Normen und Sicherheitsbestimmungen für den Laborbetrieb in Deutschland erfüllen. Insbesondere sind die Vorschriften und Empfehlungen der DIN, VDE, IEC, CE - Norm, der Unfallverhütungsvorschrift usw. in ihrer derzeit gültigen Form einzuhalten. Die Geräte müssen CE - konform und entsprechend gekennzeichnet sein oder es muss eine schriftliche Erklärung der Übereinstimmung beigelegt werden.				
Nachhaltigkeit: Um einen ressourcenschonenden und langfristigen Lebenszyklus zu gewährleisten, muss das System auf einer vollständig modularen Architektur basieren, die eine bedarfsgerechte Skalierung (insbesondere die nachträgliche Bestückung mit bis zu sechs weiteren GPUs sowie zusätzlichem Arbeitsspeicher) ermöglicht. Die Netzteile sollen über hocheffiziente 80Plus Titanium Zertifizierungen verfügen. Bei der Verpackung und Lieferung ist auf Nachhaltigkeit in Form von recyclebaren Verpackungsmaterial und CO2-Neutraler Lieferung zu achten.				

Unterschrift der Vertragsunterlagen als Zeichen der Akzeptanz

.....