

# **Leistungsbeschreibung**

## **Beschaffung einer Multi-GPU-Workstation für den KI-Lehrbetrieb**

**Vergabe Nr.: OeA-253-26**

Auftraggeberin:  
Technische Universität Berlin, Die Präsidentin  
Straße des 17. Juni 135  
10623 Berlin

## Inhaltsverzeichnis

|                                                   |    |
|---------------------------------------------------|----|
| 1 Bezeichnung der Leistung .....                  | 3  |
| 1.1 Dienstvertragliche Bedingungen .....          | 3  |
| 2 Mindestkriterien.....                           | 3  |
| 2.1 Allgemeine Anforderungen .....                | 4  |
| 2.2 Technische Mindestanforderungen .....         | 5  |
| 2.2.1 Grundsystem / Plattform .....               | 5  |
| 2.2.2 Prozessoren.....                            | 6  |
| 2.2.3 Arbeitsspeicher .....                       | 6  |
| 2.2.4 Grafikkarten / GPU .....                    | 7  |
| 2.2.5 Lokaler Speicher.....                       | 8  |
| 2.2.6 Netzteil / Stromversorgung.....             | 10 |
| 2.2.7 Netzwerk .....                              | 10 |
| 2.2.8 Betriebssystem / Linux-Kompatibilität ..... | 10 |
| 2.2.9 Support / Garantie .....                    | 11 |
| 3 Höchstpreis.....                                | 12 |
| 4 Leistungsort.....                               | 13 |
| 4.1 Abnahme.....                                  | 13 |
| 5 Termine / Ausführungszeiten.....                | 13 |
| 6 Einzureichende Unterlagen .....                 | 13 |
| 7 Zuschlagskriterien und Gewichtung.....          | 14 |
| 9.1 Kriterium Gesamtangebotsnettopreis .....      | 15 |
| 9.2 Kriterium Qualität.....                       | 16 |

## **1 BEZEICHNUNG DER LEISTUNG**

Die Technische Universität Berlin (TU Berlin, die Auftraggeberin/AG) setzt die Beschaffung einer zentralen, professionellen und erweiterbaren Multi-GPU-Workstation für den KI-Lehrbetrieb für den Einsatz am Fachgebiet um. Gegenstand der Beschaffung ist die Lieferung der erforderlichen Hardwarekomponenten einschließlich der zugehörigen Peripherie sowie des geforderten Hersteller- bzw. Vor-Ort-Supports.

Das System ist für Einsatz für den Lehrbetrieb im Bereich Künstliche Intelligenz, Machine Learning und kleinere neuronale Netze geplant. Es ist ausdrücklich nicht als große AI-Cloud und nicht für großskaliges LLM-Training vorgesehen.

Folgende Leistungen hat die/der Auftragnehmer\*in zu erfüllen:

- klassische Machine-Learning-Verfahren,
- kleinere neuronale Netze,
- vorbereitete KI-/ML-Übungen,
- parallele studentische Nutzung,
- Nutzung durch ca. 50 Studierende im Kursbetrieb,
- Arbeit überwiegend in Zweiergruppen,
- mehrere kleine bis mittlere Jobs parallel,
- Zugriff über vorhandene Pool-Rechner oder eigene Geräte,
- Betrieb unter Linux mit JupyterHub, SSH und/oder VS Code Server.

Da das vorhandene Standortnetz nur 1 Gbit/s bietet, müssen Datensätze, Arbeitsverzeichnisse und temporäre Arbeitsdaten möglichst lokal auf dem System liegen.

Das System wird in einem klimatisierten Serverraum an der Technischen Universität Berlin in einem Fachgebiet, 5. OG mit ausreichender Stromversorgung betrieben. Lautstärke, Abwärme und Stromaufnahme sind daher nachrangige Kriterien gegenüber GPU-Leistung, Erweiterbarkeit, RAM-Ausstattung, lokalem NVMe-Speicher und Herstellersupport.

### **1.1 DIENSTVERTRAGLICHE BEDINGUNGEN**

Der Vertragsschluss erfolgt ausschließlich auf Grundlage des EVB-IT Kaufvertrags einschließlich der darin einbezogenen Vertragsbedingungen. Abweichende Vertrags- oder Geschäftsbedingungen des Auftragnehmers werden nicht Vertragsbestandteil.

## **2 MINDESTKRITERIEN**

Im Folgenden werden die geforderten technischen und funktionalen Bestandteile und Parameter der zu entwickelnden KI-Lehrplattform angegeben. Alle Bestandteile und Parameter dieses Leistungsverzeichnisses sind vom AN zu erfüllen und verpflichtend, damit der vorgesehene Einsatzzweck erfüllt werden kann.

Sind die vorgegebenen Bestandteile und Parameter aus dem Angebot nicht eindeutig ersichtlich bzw. werden diese Bestandteile und Parameter nicht erfüllt oder unterschritten, wird das Angebot von der Vergabe ausgeschlossen.

## **2.1 ALLGEMEINE ANFORDERUNGEN**

Anzubieten ist eine professionelle Tower-Workstation oder gleichwertige Workstation-/Serverplattform.

Das System muss als professionelle, für die konkret angebotene Konfiguration freigegebene bzw. validierte Gesamtsystemlösung geliefert werden. Insbesondere die Kombination aus Systemplattform/Mainboard, Prozessoren, Arbeitsspeicher, Stromversorgung und der angebotenen GPU-Bestückung muss für den vorgesehenen Dauerbetrieb geeignet und freigegeben sein.

Die entsprechende Freigabe bzw. Validierung der angebotenen Gesamtkonfiguration ist mit dem Angebot durch geeignete Herstellerunterlagen oder eine entsprechende Bestätigung des Herstellers bzw. eines für das angebotene System autorisierten Systemintegrators nachzuweisen.

Nicht validierte Eigenbau-, Bastel- oder nachträglich zusammengestellte GPU-Konfigurationen ohne entsprechenden Nachweis sind nicht zulässig.

Die angebotene Plattform muss geeignet sein für:

- mindestens vier professionelle GPU-Beschleuniger für KI-/ML-Workloads,
- mindestens sechs nutzbare interne PCIe-Steckplätze,
- spätere Erweiterung bzw. Austausch der GPUs gegen stärkere Workstation-GPUs muss möglich sein,
- mindestens 256 GB ECC-Arbeitsspeicher initial,
- spätere Erweiterung auf mindestens 512 GB RAM muss möglich sein,
- lokalen NVMe-Speicher für System, Daten und Scratch,
- Linux-Betrieb mit GPU-Beschleunigung für gängige KI-/ML-Frameworks,
- 5 Jahre Herstellersupport bzw. Vor-Ort-Service.

Das System kann zunächst als Tower im klimatisierten Serverraum auf einer geeigneten Standfläche betrieben werden.

## 2.2 TECHNISCHE MINDESTANFORDERUNGEN

### 2.2.1 Grundsystem / Plattform

| Merkmal             | Mindestanforderung                                                                                                                                                              |
|---------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Bauform             | Professionelle Tower-Workstation oder gleichwertige Workstation-/Serverplattform                                                                                                |
| Plattform           | Dual-CPU-fähige Workstation-/Serverplattform                                                                                                                                    |
| CPU-Sockel          | 2 CPU-Sockel, beide bestückt                                                                                                                                                    |
| PCIe-Steckplätze    | mindestens 6 nutzbare interne PCIe-Steckplätze                                                                                                                                  |
| GPU-taugliche Slots | mindestens 4 mechanische PCIe-x16-Steckplätze für professionelle GPUs                                                                                                           |
| PCIe-Generation     | mindestens PCIe 4.0, bevorzugt PCIe 5.0 für GPU-Steckplätze                                                                                                                     |
| GPU-Auslegung       | ab Werk für mindestens 4 professionelle GPU-Beschleuniger validiert                                                                                                             |
| RAM-Erweiterung     | mindestens 16 DIMM-Slots oder gleichwertige Erweiterbarkeit                                                                                                                     |
| Speichererweiterung | interne NVMe-/M.2-/PCIe-Erweiterung für System-, Daten- und Scratch-Laufwerke                                                                                                   |
| Betrieb             | Dauerbetrieb im klimatisierten Serverraum                                                                                                                                       |
| Rackeinbau          | Ein Rackeinbau ist nicht Bestandteil der Pflichtkonfiguration. Das System kann als Tower-System im klimatisierten Serverraum auf einer geeigneten Standfläche betrieben werden. |

Die Plattform muss so ausgelegt sein, dass spätere Aufrüstungen von RAM, NVMe-Speicher und GPU-Ausstattung ohne Plattformwechsel möglich sind.

#### 2.2.1.1 OPTIONALE ANGEBOTSPOSITION F:

Über die Mindestkonfiguration mit vier GPU-Beschleunigern hinaus wird eine herstellerseitig unterstützte Erweiterbarkeit auf insgesamt mindestens sechs professionelle GPU-Beschleuniger als qualitativ höherwertige Plattform berücksichtigt. Die Anforderungen hierzu ergeben sich aus der optionalen Angebotsposition F.

## 2.2.2 Prozessoren

| <b>Merkmal</b>          | <b>Mindestanforderung</b>                                                                                                                               |
|-------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>CPU-Anzahl</b>       | 2 × CPU                                                                                                                                                 |
| <b>CPU-Klasse</b>       | Aktuelle server- oder workstationfähige Mehrkernprozessoren mit ECC-RAM-Unterstützung und ausreichender PCIe-/Speicherkanal-Anbindung oder gleichwertig |
| <b>Kerne pro CPU</b>    | Mindestens 16 Kerne pro CPU                                                                                                                             |
| <b>Gesamtkerne</b>      | Mindestens 32 Kerne gesamt                                                                                                                              |
| <b>Threads</b>          | Mindestens 64 Threads gesamt                                                                                                                            |
| <b>KI-Unterstützung</b> | Moderne Vektor-, Matrix- oder KI-Instruktionsunterstützung für ML-Inferenz, Datenvorverarbeitung und rechenintensive Workloads oder gleichwertig        |
| <b>Speicherkanäle</b>   | Mindestens 8 Speicherkanäle pro CPU oder gleichwertige Speicherbandbreite                                                                               |
| <b>Aufrüstbarkeit</b>   | CPU-Aufrüstung innerhalb der unterstützten Plattformklasse muss möglich oder herstellerseitig dokumentiert sein                                         |

Die CPU soll ausdrücklich nicht zulasten der GPU-Ausstattung überdimensioniert werden. Ziel ist eine ausgewogene Lehrplattform mit Schwerpunkt auf paralleler GPU-Nutzung.

## 2.2.3 Arbeitsspeicher

| <b>Merkmal</b>            | <b>Mindestanforderung</b>                                   |
|---------------------------|-------------------------------------------------------------|
| <b>RAM initial</b>        | Mindestens 256 GB DDR5 ECC RDIMM                            |
| <b>Speichertyp</b>        | DDR5 ECC RDIMM, Registered ECC                              |
| <b>Bestückung</b>         | Bevorzugt 8 × 32 GB oder 4 × 64 GB                          |
| <b>Erweiterbarkeit</b>    | Mindestens auf 512 GB erweiterbar                           |
| <b>Zielausbau</b>         | 512 GB oder mehr wünschenswert                              |
| <b>DIMM-Slots</b>         | Mindestens 16 DIMM-Slots oder gleichwertige Erweiterbarkeit |
| <b>Speicherverteilung</b> | Möglichst gleichmäßige Bestückung über beide CPUs           |

Der Arbeitsspeicher muss für parallele Jupyter-/Python-/ML-Prozesse ausreichend dimensioniert sein. Eine spätere Erweiterung auf mindestens 512 GB muss ohne Plattformwechsel möglich sein.

### 2.2.3.1 OPTIONALE ANGEBOTSPOSITION A:

Optional ist ein Aufpreis für die Erweiterung des Arbeitsspeichers von 256 GB auf 512 GB DDR5 ECC RDIMM anzugeben.

Sofern der Bieter diese zusätzliche Leistung erfüllen kann, ist hierfür im OeA-253-26 Preisblatt unter Position 3 ein Preis anzugeben. Bei dem anzugebenden Preis handelt es sich ausschließlich um den Aufpreis gegenüber der angebotenen Grundkonfiguration.

### 2.2.4 Grafikkarten / GPU

| Merkmal                        | Mindestanforderung                                                                                                           |
|--------------------------------|------------------------------------------------------------------------------------------------------------------------------|
| <b>GPU-Anzahl</b>              | 4 × professionelle GPU-Beschleuniger für KI-/ML-Workloads                                                                    |
| <b>VRAM pro GPU</b>            | Mindestens 24 GB                                                                                                             |
| <b>VRAM gesamt</b>             | Mindestens 96 GB verteilt                                                                                                    |
| <b>Speicherart</b>             | ECC-geschützter Grafikspeicher oder gleichwertige Fehlerkorrektur, mit ausreichender Speicherbandbreite für KI-/ML-Workloads |
| <b>Architektur</b>             | Aktuelle professionelle GPU-Architektur mit Eignung für KI-/ML-Workloads oder gleichwertig                                   |
| <b>Formfaktor</b>              | Herstellerseitig für das angebotene System validiert                                                                         |
| <b>Treiber</b>                 | Linux-kompatible Treiberunterstützung für die angebotenen GPU-Beschleuniger                                                  |
| <b>Framework-Unterstützung</b> | GPU-beschleunigter Betrieb gängiger KI-/ML-Frameworks unter Linux muss möglich sein                                          |
| <b>Support</b>                 | GPU-Bestückung muss Bestandteil des Gesamtsystem-Supports sein                                                               |

Für den vorgesehenen Lehrbetrieb ist eine Mehr-GPU-Konfiguration einer einzelnen sehr großen GPU vorzuziehen, da mehrere kleine bis mittlere KI-/ML-Jobs parallel ausgeführt werden sollen.

Die Plattform muss einen späteren Austausch der vorhandenen GPU-Beschleuniger gegen leistungstärkere professionelle GPU-Beschleuniger mit höherem

Grafikspeicher ermöglichen, ohne dass hierfür ein Austausch der Systemplattform erforderlich wird. Darüber hinaus ist eine Erweiterungsmöglichkeit um zusätzliche professionelle GPU-Beschleuniger wünschenswert. Die konkrete Erweiterung muss im Rahmen der herstellerseitig unterstützten PCIe-, Stromversorgungs-, Kühlungs- und Gehäusekonfiguration möglich sein. Leistungsstärkere bzw. zusätzliche GPU-Beschleuniger müssen für das konkrete System freigegeben und Bestandteil des Gesamtsystem-Supports sein.

#### 2.2.4.1 OPTIONALE ANGEBOTSPPOSITION B:

Optional ist ein Aufpreis für eine leistungsstärkere GPU-Ausstattung anzugeben.

Das Qualitätskriterium ist erfüllt, wenn anstelle der Mindestanforderung von 4 × professionellen GPU-Beschleunigern mit jeweils mindestens 24 GB Grafikspeicher eine Ausstattung mit 4 × professionellen GPU-Beschleunigern mit jeweils mindestens 32 GB Grafikspeicher angeboten wird. Der insgesamt verfügbare verteilte GPU-Grafikspeicher muss damit mindestens 128 GB betragen.

Die angebotenen GPU-Beschleuniger müssen für das angebotene System herstellerseitig freigegeben sein, unter Linux mit gängigen KI-/ML-Frameworks nutzbar sein und Bestandteil des Gesamtsystem-Supports sein.

Sofern der Bieter diese zusätzliche Leistung erfüllen kann, ist hierfür im OeA-253-26 Preisblatt unter Position 4 ein Preis anzugeben. Bei dem anzugebenden Preis handelt es sich ausschließlich um den Aufpreis gegenüber der angebotenen Grundkonfiguration.

#### 2.2.5 Lokaler Speicher

| Bereich                             | Mindestanforderung                                                              |
|-------------------------------------|---------------------------------------------------------------------------------|
| Systemlaufwerke                     | 2 × NVMe SSD, mindestens je 960 GB                                              |
| Systemschutz                        | Spiegelung per RAID 1, ZFS Mirror, mdraid oder gleichwertig                     |
| Daten-/Scratch-Speicher             | Mindestens 2 × NVMe SSD, mindestens je 1,92 TB                                  |
| Physisch verbauter Speicher initial | Mindestens 5,76 TB NVMe brutto                                                  |
| Erweiterbarkeit                     | Erweiterung auf mindestens 12 bis 16 TB lokalen NVMe-Speicher muss möglich sein |
| SSD-Typ                             | Enterprise- oder Workstation-taugliche NVMe SSD                                 |
| Schnittstelle                       | PCIe 4.0 oder besser                                                            |
| Speicherzellen                      | TLC oder höherwertig                                                            |

| Bereich         | Mindestanforderung                                             |
|-----------------|----------------------------------------------------------------|
| Verschlüsselung | OPAL oder hardwareseitige Verschlüsselungsoption wünschenswert |

Empfohlene Startbestückung:

| Zweck                        | Konfiguration                          |
|------------------------------|----------------------------------------|
| Betriebssystem / System      | mindestens 2 × 960 GB NVMe, gespiegelt |
| Daten / Workspaces / Scratch | mindestens 2 × 1,92 TB NVMe            |
| Zielausbau später            | 8 bis 16 TB Daten-/Scratch-Speicher    |

Aufgrund der vorhandenen 1-Gbit/s-Netzanbindung ist lokaler NVMe-Speicher für Datensätze, Arbeitsverzeichnisse und temporäre Jobs erforderlich.

#### 2.2.5.1 OPTIONALE ANGEBOTSPOSITION C:

Optional ist ein Aufpreis für die Erweiterung des Daten-/Scratch-Speichers von mindestens 2 × 1,92 TB auf mindestens 2 × 3,84 TB NVMe SSD anzugeben.

Das Qualitätskriterium ist erfüllt, wenn anstelle der Mindestanforderung von 2 × NVMe SSD mit jeweils mindestens 1,92 TB eine Ausstattung mit 2 × NVMe SSD mit jeweils mindestens 3,84 TB angeboten wird.

Sofern der Bieter diese zusätzliche Leistung erfüllen kann, ist hierfür im OeA-253-26 Preisblatt unter Position 5 ein Preis anzugeben. Bei dem anzugebenden Preis handelt es sich ausschließlich um den Aufpreis gegenüber der angebotenen Grundkonfiguration.

#### 2.2.5.2 OPTIONALE ANGEBOTSPOSITION D:

Optional ist ein Aufpreis für die Erweiterung des Daten-/Scratch-Speichers auf mindestens 4 × 3,84 TB NVMe SSD anzugeben.

Das Qualitätskriterium ist erfüllt, wenn eine Ausstattung des Daten-/Scratch-Speichers mit 4 × NVMe SSD mit jeweils mindestens 3,84 TB angeboten wird.

Sofern der Bieter diese zusätzliche Leistung erfüllen kann, ist hierfür im OeA-253-26 Preisblatt unter Position 6 ein Preis anzugeben. Bei dem anzugebenden Preis handelt es sich ausschließlich um den Aufpreis gegenüber der angebotenen Grundkonfiguration.

## 2.2.6 Netzteil / Stromversorgung

| <b>Merkmal</b>        | <b>Mindestanforderung</b>                                                            |
|-----------------------|--------------------------------------------------------------------------------------|
| <b>Netzteilklasse</b> | Hochleistungsnetzteil für Multi-GPU-Workstation                                      |
| <b>Leistung</b>       | Mindestens 1800 W                                                                    |
| <b>Bevorzugt</b>      | 2 × 1850 W Hot-Swap oder gleichwertige PSU-Konfiguration                             |
| <b>Effizienz</b>      | Mindestens 80 PLUS Platinum oder gleichwertig                                        |
| <b>Betriebsmodus</b>  | Redundanz- oder Team-/Performance-Modus, sofern herstellerseitig vorgesehen          |
| <b>GPU-Reserve</b>    | Ausreichend dimensioniert für spätere stärkere Workstation-GPUs                      |
| <b>Validierung</b>    | Stromversorgung muss für die angebotene GPU-Bestückung vom Hersteller validiert sein |

Die angebotene Stromversorgung muss herstellerseitig für die geforderte Multi-GPU-Konfiguration validiert und ausreichend dimensioniert sein.

Die Stromversorgung muss so dimensioniert sein, dass die Startkonfiguration mit vier GPUs sicher betrieben werden kann und spätere Aufrüstungen auf stärkere Workstation-GPUs möglich bleiben, sofern diese vom Hersteller freigegeben sind.

Sofern die optionale Angebotsposition F angeboten wird, muss die Plattform einschließlich Stromversorgung, Kühlung, Gehäuseauslegung und PCIe-Anbindung herstellerseitig für den Betrieb von insgesamt mindestens sechs professionellen GPU-Beschleunigern geeignet bzw. freigegeben sein.

## 2.2.7 Netzwerk

| <b>Merkmal</b>          | <b>Mindestanforderung</b>                         |
|-------------------------|---------------------------------------------------|
| <b>Netzwerk onboard</b> | Mindestens 1 GbE                                  |
| <b>Zusätzlich</b>       | Mindestens 10 GbE onboard oder als Adapter        |
| <b>Nutzung</b>          | JupyterHub, SSH, VS Code Server, Datenübertragung |

Obwohl am Standort aktuell nur 1 GbE zur Verfügung steht, soll das System bereits für spätere schnellere Netzanbindungen vorbereitet sein.

## 2.2.8 Betriebssystem / Linux-Kompatibilität

| <b>Merkmal</b>                 | <b>Mindestanforderung</b>                                                                |
|--------------------------------|------------------------------------------------------------------------------------------|
| <b>Betriebssystem</b>          | Linux-kompatibel                                                                         |
| <b>Zielbetrieb</b>             | Ubuntu Server LTS oder Rocky Linux                                                       |
| <b>Herstellerfreigabe</b>      | Linux-Support oder Linux-Zertifizierung für das angebotene Modell                        |
| <b>GPU-Treiber</b>             | Linux-kompatible Treiber für die angebotenen GPU-Beschleuniger müssen unterstützt werden |
| <b>Framework-Unterstützung</b> | GPU-beschleunigter Betrieb gängiger KI-/ML-Frameworks unter Linux muss möglich sein      |
| <b>Container</b>               | Docker, Podman oder vergleichbare Containerumgebungen müssen nutzbar sein                |
| <b>Zugriff</b>                 | JupyterHub, SSH und VS Code Server müssen betreibbar sein                                |

Das System muss im späteren Betrieb ohne zwingenden Windows-Einsatz als Linux-basierte KI-Lehrplattform genutzt werden können.

### 2.2.9 Support / Garantie

| <b>Merkmal</b>       | <b>Mindestanforderung</b>                                                                                                                                                                                                                                                                                         |
|----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Laufzeit</b>      | 5 Jahre                                                                                                                                                                                                                                                                                                           |
| <b>Art</b>           | Hersteller-Vor-Ort-Service, Premier Support, ProSupport oder gleichwertig<br><br>Der Auftragnehmer stellt den Herstellersupport beziehungsweise Vor-Ort-Service für 5 Jahre ab Abnahme des Systems sicher. Der Support umfasst die Behebung hardwarebezogener Störungen im Rahmen des angebotenen Supportmodells. |
| <b>Reaktionszeit</b> | Next Business Day oder besser<br>Der 5-jährige Vor-Ort-Service mit einer Reaktionszeit von Next Business Day oder besser ist Bestandteil der Pflichtkonfiguration. Eine längere Reaktionszeit erfüllt die Mindestanforderung nicht.                                                                               |
| <b>Abdeckung</b>     | Gesamtes System inklusive Mainboard, Netzteil, CPUs, RAM, GPUs und SSDs                                                                                                                                                                                                                                           |
| <b>Ersatzteile</b>   | Ersatzteilversorgung über Hersteller oder autorisierten Servicepartner                                                                                                                                                                                                                                            |

| Merkmal       | Mindestanforderung                                   |
|---------------|------------------------------------------------------|
| Supportmodell | Gesamtsystem-Support, keine reine Einzelteilgarantie |

Die angebotene GPU-Bestückung muss Bestandteil des Herstellersupports sein. Nicht herstellerseitig validierte Eigenbau- oder Nachrüstkonfigurationen sind nicht zulässig.

Der angebotene Support muss die konkret angebotene Gesamtkonfiguration einschließlich der vollständigen GPU-Bestückung abdecken. Die entsprechende Abdeckung ist mit dem Angebot nachzuweisen.

### 2.2.9.1 OPTIONALE ANGEBOTSPOSITION E:

Optional ist ein Aufpreis für einen erweiterten Support mit kürzerer Reaktionszeit anzugeben.

Die Mindestanforderung beträgt 5 Jahre Herstellersupport beziehungsweise Vor-Ort-Service mit Reaktionszeit Next Business Day oder besser. Das Qualitätskriterium ist erfüllt, wenn für das angebotene Gesamtsystem ein 5-jähriger Herstellersupport beziehungsweise Vor-Ort-Service mit einer Reaktionszeit von spätestens 4 Stunden innerhalb der vereinbarten Servicezeiten angeboten wird.

Die Supportleistung muss das Gesamtsystem einschließlich Mainboard, Netzteil, Prozessoren, Arbeitsspeicher, GPU-Beschleunigern und Massenspeicher abdecken.

Sofern der Bieter diese zusätzliche Leistung erfüllen kann, ist hierfür im OeA-253-26 Preisblatt unter Position 7 ein Preis anzugeben. Bei dem anzugebenden Preis handelt es sich ausschließlich um den Aufpreis gegenüber der angebotenen Grundkonfiguration.

## 3 HÖCHSTPREIS

Der **Höchstpreis** für den Auftrag (exkl. optionale Leistungen) beträgt **26.050,42 Euro netto / 31.000,00 Euro brutto** und ist bindend für alle Bieter\*innen/ Bietergemeinschaften. **Das heißt, Bieter\*innen/ Bietergemeinschaften, die einen höheren Angebotspreis einreichen, werden vom weiteren Verfahren ausgeschlossen.**

Im Höchstpreis sind sämtliche gegebenenfalls anfallenden kostenpflichtigen Nebenleistungen bereits vollständig zu berücksichtigen und einzukalkulieren. Die optionalen Leistungen bzw. qualitativen Kriterien gemäß Punkt 9.2 der Leistungsbeschreibung sind nicht Bestandteil des Höchstpreises. Über das Angebot hinausgehende Forderungen können gegenüber der TU Berlin nicht geltend gemacht werden.

## 4 LEISTUNGsort

Technische Universität Berlin, Pascalstr. 8-9, 10587 Berlin

### 4.1 ABNAHME

Die Abnahme erfolgt nach Lieferung und Prüfung des Systems durch die Auftraggeber/in. Geprüft wird insbesondere, ob die gelieferte Hardware der angebotenen und beauftragten Konfiguration entspricht und die geforderten technischen Mindestanforderungen erfüllt. Mängel oder Abweichungen sind durch den Auftragnehmer innerhalb einer angemessenen Frist zu beheben.

## 5 TERMINE / AUSFÜHRUNGSZEITEN

Ausführungsbeginn: Die Ausführung kann ab dem Tag der Auftragsvergabe (voraussichtlich 15.10.2026) erfolgen.

Ausführungsende: Die Lieferung muss spätestens innerhalb von 10 Wochen nach Zuschlagserteilung erfolgen. Der 5-jährige Support beginnt ab Abnahme des Systems.

## 6 EINZUREICHENDE UNTERLAGEN

Neben den geforderten Eigenerklärungen (siehe Vergabeunterlagen) sind folgende Punkte im Angebot schriftlich darzustellen. Werden die folgenden Punkte nicht schriftlich dargestellt, so wird das Angebot ausgeschlossen.

- **TUB - Angebotsschreiben des Bieters:** Bitte reichen Sie dieses Dokument ausgefüllt mit dem Angebot ein.
- **TUB – Wirt-Erklärungen (EK<sup>1</sup>):** Bitte reichen Sie diese Dokumente ausgefüllt mit dem Angebot ein:

---

<sup>1</sup> EK = Eignungskriterium

- o Wirt-124

Sofern zutreffend:

- o Wirt-235
- o Wirt-236
- o Wirt-238

- **Spezifikationsdokumente (EK):** Das Angebot hat Spezifikationsdokumente (technische Beschreibung, Datenblätter o.ä.) zu enthalten, aus denen die Erfüllung sämtlicher geforderter Spezifikationen einschließlich der erforderlichen technischen Anschlussparameter eindeutig hervorgeht.

Bei Angebot der optionalen Angebotsposition F ist zusätzlich ein geeigneter Nachweis einzureichen, aus dem die herstellerseitige Eignung bzw. Freigabe der angebotenen Plattform für den Betrieb von insgesamt mindestens sechs professionellen GPU-Beschleunigern hervorgeht.

Zusätzlich ist mit dem Angebot ein geeigneter Nachweis einzureichen, aus dem hervorgeht, dass die angebotene Gesamtkonfiguration einschließlich Systemplattform/Mainboard, Prozessoren, Arbeitsspeicher, Stromversorgung und vollständiger GPU-Bestückung für den vorgesehenen Betrieb freigegeben bzw. validiert ist. Als Nachweis können geeignete Herstellerunterlagen oder eine entsprechende Bestätigung des Herstellers bzw. eines für das angebotene System autorisierten Systemintegrators vorgelegt werden.

- **TUB – Preisblatt (ZK<sup>2</sup>):** Bitte reichen Sie dieses Dokument ausgefüllt mit dem Angebot ein.
- **TUB - Qualitätskriterien (ZK):** Sofern Sie die unter den Positionen 3 bis 8 des OeA-253-26 Preisblatts aufgeführten zusätzlichen Leistungen anbieten können, geben Sie hierfür den jeweiligen Preis im Preisblatt an. Die angebotenen zusätzlichen Leistungen müssen die in der Leistungsbeschreibung genannten Anforderungen erfüllen. Die entsprechenden Angaben sind mit dem Angebot einzureichen und durch Spezifikationsdokumente zu bestätigen.

## 7 ZUSCHLAGSKRITERIEN UND GEWICHTUNG

Unter Berücksichtigung aller Umstände wird der Zuschlag jeweils auf das wirtschaftlichste Angebot erteilt. Die Bewertung erfolgt auf Grundlage des schriftlichen

---

<sup>2</sup> ZK = Zuschlagskriterium

Angebots. Daher liegt es im Interesse der/ des Bieters\*in, alle angeforderten Informationen so detailliert und korrekt wie möglich zur Verfügung zu stellen.

Zuschlagskriterium ist zu 30,00 % der Gesamtangebotsnettopreis und zu 70,00 % die Qualität. Die sich aus dem jeweiligen Kriterium ergebenden Punkte werden addiert. Der/ die Bieter\*in mit der Höchstpunktzahl erhält den Zuschlag. Alle Zahlenwerte werden auf zwei Nachkommastellen gerundet.

| Kriterium                | Gewichtung      | Punkte               |
|--------------------------|-----------------|----------------------|
| Gesamtangebotsnettopreis | 30,00 %         | 30,00                |
| Qualität                 | 70,00 %         | 70,00                |
| <b>Gesamt</b>            | <b>100,00 %</b> | <b>100,00 Punkte</b> |

### 9.1 KRITERIUM GESAMTANGEBOTSNETTOPREIS

Aus der Menge der wertungsfähigen Angebote, d.h. alle Angebote, die nicht aus formalen Gründen, nicht aus Gründen der fehlenden Unternehmereignung und nicht aus Gründen eines im Verhältnis zur Leistung unangemessenen Preises ausgeschlossen wurden, wird der niedrigste Angebotspreis bestimmt.

Das Angebot mit dem niedrigsten Gesamtpreis (exkl. MwSt.) erhält die Höchstpunktzahl von 30 Punkten. Ist der Gesamtpreis eines Angebotes z.B. um 100 % höher im Vergleich zum Angebot mit dem niedrigsten Gesamtpreis, dann erhält dieses Angebot 0 Punkte. Null Punkte stellen die Untergrenze dar, d.h. es werden keine Negativpunkte vergeben. Die Punkte dazwischen werden nach der folgenden Gleichung interpolär berechnet:

Formel:

$$Punkte_{preis} = \left( 2 - \frac{Preis}{Preis_{min}} \right) * 30$$

Beispielrechnung zur Punktevergabe:

|            |                      |                |
|------------|----------------------|----------------|
| Angebot 1: | Preis von 15.000 EUR | = 15,00 Punkte |
| Angebot 2: | Preis von 12.000 EUR | = 24,00 Punkte |
| Angebot 3: | Preis von 10.000 EUR | = 30,00 Punkte |

Punkte für Angebot 2:

$$Punkte_{preis} = \left( 2 - \frac{12.000}{10.000} \right) * 30 = 24,00$$

## 9.2 KRITERIUM QUALITÄT

Nachfolgend werden qualitativ höherwertige Aspekte der Leistung angegeben. Ob, diese zusätzlichen Aspekte oder nur bestimmte zusätzliche Aspekte tatsächlich mitbeauftragt werden, werden mit Zuschlagserteilung gegenüber dem zuschlagsbegünstigten Bieter bekannt gegeben.

Auf die Erfüllung der zusätzlichen Aspekte werden insgesamt maximal **70 Punkte** vergeben.

Jedes Angebot, das die angegebenen zusätzlichen Aspekte erfüllt, erhält die angegebenen Punkte. Im OeA-253-26 Preisblatt ist das jeweilige Kriterium Qualität mit dessen Kosten unter den Positionen 3 - 8 entsprechend anzugeben. Wird dort keine Eintragung gemacht, so erhält das Angebot 0 Punkte.

### Punktevergabe: Optionale Angebotspositionen

Folgende Positionen dürfen jedoch nicht zulasten der Pflichtkonfiguration gehen:

| Kriterium                 | Beschreibung                                                                                                                                | Punkte |
|---------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|--------|
| <b>Kriterium Option A</b> | – Erweiterung RAM von 256 GB auf 512 GB DDR5 ECC RDIMM                                                                                      | 10     |
| <b>Kriterium Option B</b> | - Leistungsstärkerer professioneller GPU-Beschleuniger mit höherem Grafikspeicher und herstellerseitiger Freigabe für das angebotene System | 20     |
| <b>Kriterium Option C</b> | - Erweiterung Daten-/Scratch-Speicher von mindestens 2 × 1,92 TB auf mindestens 2 × 3,84 TB NVMe                                            | 10     |
| <b>Kriterium Option D</b> | – Erweiterung Daten-/Scratch-Speicher auf mindestens 4 × 3,84 TB NVMe                                                                       | 5      |
| <b>Kriterium Option E</b> | - Erweiterter Support mit kürzerer Reaktionszeit                                                                                            | 10     |
| <b>Kriterium Option F</b> | – Erweiterte GPU-Plattform / Erweiterbarkeit auf mindestens sechs professionelle GPU-Beschleuniger                                          | 15     |

### **Optionale Angebotsposition F:**

Das Qualitätskriterium ist erfüllt, wenn die angebotene Plattform herstellerseitig dafür ausgelegt und freigegeben ist, zusätzlich zu den vier GPU-Beschleunigern der Pflichtkonfiguration mindestens zwei weitere professionelle GPU-Beschleuniger aufnehmen und betreiben zu können. Hierfür müssen mindestens zwei zusätzliche für professionelle GPU-Beschleuniger geeignete PCIe-x16-Steckplätze bzw. eine

technisch gleichwertige GPU-Anbindung sowie die erforderliche Gehäuse-, Kühlungs- und Stromversorgungsreserve vorhanden sein.

Die Erweiterbarkeit auf insgesamt mindestens sechs professionelle GPU-Beschleuniger ist durch geeignete Herstellerunterlagen oder eine entsprechende Bestätigung des Herstellers bzw. eines für das angebotene System autorisierten Systemintegrators nachzuweisen.

Sofern der Bieter dieses Qualitätskriterium erfüllen kann, ist hierfür im OeA-253-26 Preisblatt unter Position 8 eine entsprechende Angabe vorzunehmen. Sofern für die erweiterte GPU-Plattform gegenüber der angebotenen Grundkonfiguration kein zusätzlicher Aufpreis anfällt, ist dies im Preisblatt entsprechend mit 0,00 Euro auszuweisen.