

Leistungsbeschreibung zur Vergabe

**Lieferung von drei Rackservern
als Hardwarebasis für einen Proxmox-Cluster
E_3150_PR1284270**

Inhalt

1. Beschreibung des Vorhabens	3
1.1 Allgemeine Vorstellung der Organisation:	3
1.2 Kurzvorstellung des Instituts	3
1.3 Konkrete Vorstellung des Vorhabens	3
1.3.1 Ausgangslage und Zielsetzung	3
1.3.2 Nutzerkreis und Einsatzumgebung	4
1.3.3 Beschaffungsumfang	4
2. Technische Mindestanforderungen	4
2.1 Prozessorleistung	5
2.2 Arbeitsspeicher	5
2.3 Lokaler Speicher	6
2.4 Netzwerk und Schnittstellen	6
2.5 Erweiterbarkeit	6
2.6 Stromversorgung und Energieeffizienz	6
2.7 Verfügbarkeit und Ausfallsicherheit	7
3 Systemumgebung und Kompatibilität	7
4 Lieferung, Einbau und Leistungsort	7
4.1 Lieferung	7
4.2 Leistungsort	7
4.3 Einbau und Inbetriebnahme	8
5 Liefertermin	8
6 Herstellerservice und Support	8
6.1 Supportumfang	8
6.2 Servicezeiten	8
6.3 Vor-Ort-Service am nächsten Arbeitstag	9
6.4 Ersatzteilebereitstellung	9
7 Ausarbeitung des Angebots	9

1. Beschreibung des Vorhabens

1.1 Allgemeine Vorstellung der Organisation:

Die Fraunhofer-Gesellschaft mit Sitz in Deutschland ist eine der führenden Organisationen für anwendungsorientierte Forschung. Im Innovationsprozess spielt sie eine zentrale Rolle – mit Forschungsschwerpunkten in zukunftsrelevanten Schlüsseltechnologien und dem Transfer von Forschungsergebnissen in die Industrie zur Stärkung unseres Wirtschaftsstandorts und zum Wohle unserer Gesellschaft.

Die 1949 gegründete Organisation betreibt in Deutschland derzeit 76 Institute und Forschungseinrichtungen. Die gegenwärtig knapp 32 000 Mitarbeitenden, überwiegend mit natur- oder ingenieurwissenschaftlicher Ausbildung, erarbeiten das jährliche Finanzvolumen von 3,4 Mrd. €. Davon fallen 3,0 Mrd. € auf den Bereich Vertragsforschung, der sich in drei Finanzierungssäulen gliedert: Einen Anteil davon erwirtschaftet Fraunhofer mit Aufträgen aus der Industrie und aus Lizenzerträgen, die sich auf insgesamt 836 Mio. € belaufen. Der hohe Anteil an Wirtschaftserträgen ist das Fraunhofer-Alleinstellungsmerkmal in der deutschen Forschungslandschaft. Ein weiterer Teil aus dem Bereich Vertragsforschung stammt aus öffentlich finanzierten Forschungsprojekten. Bund und Länder komplettieren die Vertragsforschung durch die Grundfinanzierung. Damit ermöglichen die Zuwendungsgeber, dass die Institute schon heute Problemlösungen entwickeln können, die in einigen Jahren für Wirtschaft und Gesellschaft relevant werden.

1.2 Kurzvorstellung des Instituts

Das Institut Fraunhofer FOKUS in Berlin erforscht die Digitale Vernetzung und ihre Auswirkungen auf Gesellschaft, Wirtschaft und Technologie. Seit 1988 unterstützt es Wirtschaftsunternehmen und öffentliche Verwaltung in der Gestaltung und Umsetzung des digitalen Wandels. Dazu bietet Fraunhofer FOKUS Forschungsleistungen von der Anforderungsanalyse über Beratung, Machbarkeitsstudien, Technologieentwicklung bis hin zu Prototypen und Piloten in den Geschäftsbereichen Digital Public Services, Future Applications and Media, Quality Engineering, Smart Mobility, Software-based Networks und Vernetzte Sicherheit an. Fraunhofer FOKUS gehört zu den größten IKT-Instituten der Fraunhofer-Gesellschaft.

1.3 Konkrete Vorstellung des Vorhabens

1.3.1 Ausgangslage und Zielsetzung

Anlass der Beschaffung ist die Erweiterung der vorhandenen Rechen- und Virtualisierungskapazitäten. Beschafft werden drei einheitlich konfigurierte, hochperformante und ausbaufähige Rackserver. Sie sollen als Knoten einer On-Premise-Virtualisierungsumgebung auf Basis von Proxmox VE eingesetzt werden, neue virtualisierte Workloads aufnehmen und geeignete, bislang auf VMware betriebene Systeme schrittweise übernehmen.

Ziel ist eine zukunftssichere Serverplattform mit hoher Rechen- und Arbeitsspeicherkapazität, redundanter Stromversorgung, umfangreicher Netzwerkanbindung, Remote-Management und Herstellerservice. Die Systeme sind für die gemeinsame Nutzung von Rechen-, Speicher- und Netzwerkressourcen durch mehrere Forschungsprojekte vorgesehen.

1.3.2 Nutzerkreis und Einsatzumgebung

Die Hardware ist für die IT-Abteilung und Administratoren bestimmt. Der Einsatz erfolgt in einem Rechenzentrum beziehungsweise Serverraum. Als Rahmenbedingungen sind IT-Sicherheitsanforderungen, räumliche Vorgaben und eine begrenzte Stromversorgung zu berücksichtigen.

1.3.3 Beschaffungsumfang

- Drei Rackserver mit identischer Konfiguration
- Rack-Schienen für den Einbau in ein Standard-19-Zoll-Rack
- Redundante, hot-plug-fähige Netzteile
- Vorbereitung für den Einbau von lokalem Speicher
- Vorbereitung für den Einbau von zwei NVidia H100-Karten
- Erforderliche Netzwerk- und Managementschnittstellen
- Technische Unterlagen sowie Kostenaufschlüsselung
- ProSupport oder gleichwertiger Herstellerservice mit Vor-Ort-Service am nächsten Arbeitstag für 60 Monate

2. Technische Mindestanforderungen

Sämtliche Anforderungen dieses Kapitels sind verbindliche Mindestanforderungen. Angebote, die eine Mindestanforderung nicht erfüllen, sind nicht wertbar. Gleichwertige Lösungen sind zugelassen, sofern die jeweils geforderten Funktionen, Leistungswerte, Kompatibilitäten und Qualitätsmerkmale nachweislich erfüllt werden. Hersteller-, Marken-, Typen- oder Produktbezeichnungen werden nur verwendet, wenn der Auftragsgegenstand andernfalls nicht hinreichend genau und allgemein verständlich beschrieben werden kann oder eine sachlich gerechtfertigte technische Notwendigkeit besteht. Soweit eine solche Bezeichnung verwendet wird, gilt sie grundsätzlich mit dem Zusatz „oder gleichwertig“; die Gleichwertigkeit ist anhand der in den Vergabeunterlagen festgelegten Kriterien nachzuweisen.

Nr.	Bereich	Mindestanforderung
1	Bauform	19-Zoll-Rackserver, maximal 2 HE, einschließlich geeigneter Rack-Schienen
2	Konfiguration	Drei Server mit identischer Hardwarekonfiguration
3	Prozessor	2 Serverprozessoren je Server, mit mindestens 48 physischen Prozessorkernen je Prozessor
4	Arbeitsspeicher	Mindestens 768 GB ECC-RAM DDR5-6400 je Server, Das System muss nach Lieferung mindestens 50 % freie RAM-Steckplätze für einen späteren Ausbau aufweisen.
5	Speicher	Boot-Controller mit M.2 und mindestens 960 GB im RAID-1-Verbund

6	Netzwerk	Mindestens vier 25-GbE-SFP28-, zwei 100 GbE QSFP28 und zwei 10 GbE – SFP+ oder gleichwertige Ethernetanschlüsse je Server
7	Netzwerkredundanz	Verteilung der produktiven Netzwerkanschlüsse auf mindestens zwei getrennte Netzwerkpfade; VLAN-Unterstützung
8	Erweiterbarkeit	Erweiterbarkeit mit 2 NVIDIA H100 Karten muss möglich und vorbereitet sein
9	Remote-Management	Dedizierter Out-of-Band-Managementanschluss mit Funktionen für Fernwartung, Diagnose und Neustart, z. B. iDRAC Enterprise oder gleichwertig
10	Stromversorgung	Zwei hot-plug-fähige, redundant ausgelegte Netzteile je Server; Anschluss an getrennte A/B-Strompfade, Versorgung von Erweiterungen wie 2x H100
11	Virtualisierung	Eignung für Proxmox VE und parallelen Betrieb virtueller Linux- und Windows-Gastsysteme
12	Kompatibilität	Netzwerkkarten, Boot-Controller und Firmware müssen von der eingesetzten Proxmox-VE-Version beziehungsweise dem zugrunde liegenden Linux-Kernel unterstützt werden
13	Konformität	CE-Kennzeichnung, RoHS-Konformität sowie Einhaltung geltender Sicherheits- und EMV-Anforderungen
14	Betrieb	Eignung für produktiven 24/7-Dauerbetrieb in einer Rechenzentrumsumgebung

2.1 Prozessorleistung

Je Server sind zwei für den Serverbetrieb geeignete Prozessoren mit jeweils mindestens 48 physischen Prozessorkernen erforderlich. Die TDP beider Prozessoren darf zusammen höchstens 600 W betragen. Die Prozessoren müssen Arbeitsspeicher nach dem Standard DDR5-6400 oder neuer unterstützen. Die erforderliche Rechenleistung ist anhand eines einheitlichen, herstellernutralen und nachprüfbaren Leistungsmaßstabs festzulegen. Sie wird benötigt, um mehrere virtuelle Maschinen gleichzeitig stabil zu betreiben, Lastspitzen und künftige Erweiterungen abzudecken und beim Ausfall eines Clusterknotens virtuelle Maschinen übernehmen zu können. Eine geringere Prozessorleistung kann zu CPU-Engpässen, verlängerten Antwortzeiten und Einschränkungen der virtualisierten Fachanwendungen führen. Der AMD EPYC 9455 dient ausschließlich als unverbindliches Beispiel für die geforderte Leistungsklasse; gleichwertige Prozessoren anderer Hersteller sind zugelassen.

2.2 Arbeitsspeicher

Jeder Server muss mit mindestens 768 GB ECC-Arbeitsspeicher ausgestattet sein. Die Speicher müssen mindestens dem Standard DDR5-6400 entsprechen. Das System muss nach Lieferung

mindestens 50 % der maximal unterstützten Systemkapazität für einen späteren Ausbau aufweisen. Dieser Speicher wird für den dauerhaften Betrieb mehrerer virtueller Maschinen, deren Betriebssysteme, Fachanwendungen und Datenbankdienste sowie den Proxmox-Hypervisor benötigt. Sie muss Lastspitzen, Wachstum und die Übernahme virtueller Maschinen beim Ausfall eines Clusterknotens abdecken. Eine geringere Ausstattung kann zu Speicherdruck, Auslagerungsvorgängen und nicht vertretbaren Leistungseinbußen führen.

2.3 Lokaler Speicher

Jeder Server muss mit einem RAID-Controller ausgestattet sein, der den Betrieb des Servers mit lokalen Festplatten in den Kombinationen RAID 0,1,5,6,10,50 und 60 ermöglicht. Lokale Platten sollen noch nicht verbaut sein, die Möglichkeit muss aber vorhanden sein. Zusätzlich muss jeder Server ein Boot-Controller beinhalten, der zwei M.2-SSD Laufwerke mit mindestens 960 GB Speicherkapazität im RAID-1-Verbund bereitstellt, und als Boot-Laufwerk für das System verwendet werden kann. Dieser Boot-Controller muss kompatibel zu Proxmox

Installationen sein.

2.4 Netzwerk und Schnittstellen

Je Server sind mehrere redundant nutzbare Ethernetanschlüsse erforderlich, darunter mindestens vier 25 GbE-SFP28-, zwei 100 GbE-QSFP28 und zwei 10 GbE-SFP+ – oder gleichwertige Schnittstellen für Cluster-Kommunikation, VM-Datenverkehr, Live-Migrationen und gegebenenfalls Storage-Verkehr. Die Anschlüsse sind auf mindestens zwei getrennte Switch- beziehungsweise Netzwerkpfade zu verteilen. VLANs müssen unterstützt werden. Zusätzlich ist ein dediziertes Out-of-Band-Management für Fernwartung, Diagnose und Neustart erforderlich.

Ohne geeignete und redundante Netzwerkanbindung sind Cluster-Kommunikation, Erreichbarkeit virtueller Maschinen, Live-Migration und Hochverfügbarkeitsfunktionen beeinträchtigt. Der Ausfall einer einzelnen Netzwerkkarte, eines Kabels oder Switch-Ports darf nicht zur vollständigen Nichterreichbarkeit des Serverknotens führen.

2.5 Erweiterbarkeit

Die Systeme müssen für den späteren Einbau von zwei GPU-Beschleunigern je Server vorbereitet sein. Erforderlich sind geeignete Einbauplätze und PCIe-Anbindungen, eine für den vorgesehenen Vollausbau ausreichende Stromversorgung und Kühlung sowie die Unterstützung der erforderlichen Kommunikation zwischen CPU und GPU. Maßgeblich sind die in den Vergabeunterlagen festgelegten technischen Mindestparameter, insbesondere Bauform, Schnittstelle, Leistungsaufnahme und Kühlbedarf. Die GPUs selbst sind nicht Bestandteil des Lieferumfangs. Die NVIDIA H100 dient ausschließlich als unverbindliches Beispiel für die vorgesehene Leistungsklasse; technisch gleichwertige Lösungen sind zugelassen.

Diese Vorbereitung ist notwendig, wenn im Betriebszeitraum des Systems die Notwendigkeit für KI-spezifische Workloads entsteht.

2.6 Stromversorgung und Energieeffizienz

Jeder Server muss über zwei hot-plug-fähige, redundant ausgelegte Netzteile verfügen und über getrennte A/B-Strompfade versorgt werden können. Bei Ausfall eines Netzteils, einer PDU, eines Stromkreises oder einer USV muss der Clusterknoten weiterbetrieben werden können. Energieeffiziente Komponenten und Einstellungen sind erwünscht. Wegen der erforderlichen Rechenleistung, Speicherausstattung und Ausfallsicherheitsreserven ist ein besonders niedriger Energieverbrauch jedoch kein vorrangiges Mindestkriterium.

2.7 Verfügbarkeit und Ausfallsicherheit

Das Gesamtsystem ist als Cluster mit mindestens drei Knoten zur Sicherstellung des Quorums und mit N+1-Kapazitätsreserve auszulegen. Bei Ausfall eines Knotens müssen die verbleibenden Knoten priorisierte virtuelle Maschinen übernehmen können. Erforderlich sind redundante Netzwerkanbindungen und Stromversorgungen.

3 Systemumgebung und Kompatibilität

Die Server werden als Knoten eines On-Premise-Proxmox-Clusters eingesetzt. Die Hardware muss Proxmox VE ab Version 9.2 oder höher als Linux-basierten Virtualisierungshypervisor unterstützen und den parallelen Betrieb virtueller Linux- und Windows-Gastsysteme ermöglichen. Netzwerkkarten, Boot-Controller und Firmware müssen mit Proxmox VE ab Version 9.2 d. h. dem jeweils zugrunde liegenden Linux-Kernel kompatibel sein. Der Bieter hat die Kompatibilität der angebotenen Netzwerkkarten, Boot-Controller und Managementkomponenten mit Proxmox VE Version 9.2 oder höher bzw. dem zugrunde liegenden Linux-Kernel durch Herstellerunterlagen, technische Freigaben oder andere nachvollziehbare technische Nachweise zu belegen.

Die Netzteile müssen für den Betrieb an Wechselspannung 200–240 V AC geeignet sein und den Vollausbau des angebotenen Systems einschließlich aller beschleunigenden Recheneinheiten (GPU-Beschleuniger) unterstützen.

Die Rack-Einbausysteme müssen mit 19" Racks mit 740mm Montageabstand kompatibel sein.

4 Lieferung, Einbau und Leistungsort

4.1 Lieferung

Der Auftragnehmer liefert drei vollständig konfigurierte Rackserver einschließlich aller zum bestimmungsgemäßen Betrieb erforderlichen Komponenten, Rack-Schienen, Dokumentationen und bestellten Serviceleistungen. Transport, Verpackung, Versicherung und Entladung sind in den Angebotspreisen enthalten.

4.2 Leistungsort

Die Lieferung erfolgt zur Warenannahme des Instituts.

Anschrift:

Fraunhofer Fokus, Kaiserin-Augusta-Allee 31, 10589 Berlin.

4.3 Einbau und Inbetriebnahme

Einbau und Inbetriebnahme erfolgt durch den Auftraggeber, nicht durch den Auftragnehmer.

5 Liefertermin

Die Lieferung erfolgt zum nächstmöglichen Zeitpunkt. Die vollständige Lieferung muss spätestens am letzten Werktag der Kalenderwoche 50 am vereinbarten Bestimmungsort eingegangen sein. Der Termin ist verbindlich. Die Rechte des Auftraggebers bei nicht fristgerechter Leistung richten sich nach den Vertragsbedingungen und den gesetzlichen Vorschriften.

6 Herstellerservice und Support

Für alle drei Server ist ein Herstellerservice auf dem Niveau "ProSupport und Vor-Ort-Service oder gleichwertig am nächsten Werktag" für eine Laufzeit von 60 Monaten ab Wareneingang vorzusehen. Der Service muss sämtliche gelieferte Hardwarekonfiguration umfassen.

6.1 Supportumfang

Der Auftragnehmer stellt für die angebotenen Serversysteme einen Herstellerservice für einen Zeitraum von 60 Monaten bereit.

Der Service umfasst mindestens:

- Hardwarefehlerdiagnose durch qualifiziertes Fachpersonal
- Remote-Support zur Fehleranalyse
- Bereitstellung von Firmware- und Treiberupdates
- Lieferung von Ersatzteilen
- Vor-Ort-Service durch qualifiziertes Servicepersonal
- Koordination und Eskalation von Störungsfällen bis zur Wiederherstellung der Hardwarefunktion

6.2 Servicezeiten

Der Support muss rund um die Uhr erreichbar sein.

Reaktionszeit

Die Reaktionszeit ist Vorfallabhängig, beträgt bei kritischen Problemen höchstens vier Stunden.

Als Reaktionszeit gilt der Zeitraum zwischen Eingang der Störungsmeldung und Beginn der qualifizierten Bearbeitung durch den Support.

6.3 Vor-Ort-Service am nächsten Arbeitstag

Kann eine Störung nicht per Remote-Diagnose behoben werden und wird ein Hardwaredefekt festgestellt, muss spätestens am nächsten Arbeitstag ein Vor-Ort-Einsatz am Leistungsort erfolgen.

Der Vor-Ort-Einsatz umfasst:

- Anreise des Servicetechnikers
- Austausch defekter Hardwarekomponenten
- Funktionsprüfung der ersetzten Komponenten

Als „nächster Arbeitstag“ gilt der erste Werktag nach abgeschlossener Fehlerdiagnose.

6.4 Ersatzteillbereitstellung

Erforderliche Ersatzteile sind vom Auftragnehmer bereitzustellen.

Der Auftragnehmer trägt sämtliche Kosten für:

- Ersatzteile
- Versand
- Vor-Ort-Einsatz
- Arbeitszeit

Defekte Datenträger

Für alle Datenträger mit möglicher Speicherung personenbezogener, vertraulicher oder schutzbedürftiger Daten gilt:

- Defekte Datenträger verbleiben beim Auftraggeber und müssen nicht an den Hersteller zurückgegeben werden.

Der Auftragnehmer hat eine entsprechende Leistung (vergleichbar „Keep Your Hard Drive“) bereitzustellen.

7 Ausarbeitung des Angebots

Die Fraunhofer-Gesellschaft bittet um ein – dem oben genannten Bedarf entsprechendem – Angebot:

- Die Angabe einer konkreten Ansprechperson inkl. Kontaktdaten, welche für die Ausschreibung verantwortlich ist
- Gesamtpreis und nachvollziehbare Kostenaufschlüsselung je Server und Serviceposition
- Dem Angebot sind folgende technische Dokumentationen hinzuzufügen
 - Vollständige technische Datenblätter und Konfigurationsübersicht
 - Hersteller- und Typbezeichnungen
 - Übersicht aller verbauten Komponenten
 - Firmware- und Treiberstände

- Dokumentation der Netzwerk- und Managementschnittstellen
- Garantie- und Servicenachweise
- CE- und RoHS-Nachweise beziehungsweise entsprechende Herstellererklärungen
- Nachweis der Proxmox-VE- beziehungsweise Linux-Kompatibilität der wesentlichen Komponenten
- Beschreibung des angebotenen Herstellerservices und seiner Laufzeit
- Verbindlicher Liefertermin