

Anlage 5 - Leistungsbeschreibung für Compute Cluster

1. Allgemein

Es soll ein zentrales Compute-Cluster für die Lösung von Aufgaben der numerischen Simulation und Berechnungen überwiegend im Bereich der FEM, CFD und Quantenmechanik beschafft werden. Zudem soll das zu beschaffende Gerät die zunehmenden Bedarfe an Rechenzeit für die Forschung im Bereich künstliche Intelligenz und Machine-Learning decken. Ziel dieser Beschaffung ist ein Cluster mit geringen Latenzen für außerordentlich rechenintensive, parallele Anwendungen in den genannten Bereichen, welches für einen Betriebszeitraum von mindestens 5 Jahren ausgelegt ist. Der zu beschaffende Cluster soll folgendermaßen ausgestattet sein:

Rechencluster:

- Knoten mit einer hohen Anzahl von Rechenkernen (Rechenknoten)
- Knoten mit GPUs (GPU-Knoten)
- Zugangsknoten
- Service-Knoten
- Paralleles Filesystem
- Management Knoten (Main-Knoten)
- Hochleistungskommunikationsnetzwerk zwischen den Rechenknoten
- Integration in die Rechenzentrums-Infrastruktur
- Software- und Programmierpakete

Infrastruktur:

- Vollausgestattete Server-Schränke
- Luft-/Wasser-Wärmetauscher Kühlmodule

Garantie- und Support-Konzept:

- Support in Hard- und Software
- Garantie der Hardwarekomponenten

Eine schlüsselfertige Anlage wird benötigt, die alle geforderten Leistungsparameter erfüllt und ein Maximum an Performance bezogen auf die geforderten Benchmarks bereitstellt. Es werden **drei** Positionen gefordert:

1. Compute Cluster mit einem verfügbaren Budget von maximal **1.500.000€** (inkl. MWST), mit welchem die größtmögliche Rechenleistung zu erzielen ist.
2. Infrastrukturkomponenten. Hier liegt das Budget bei maximal **100.000€** (inkl. MWST)
3. Garantie- und Support-Konzept über 5 Jahre ab der Abnahme. Das verfügbare Budget von **100.000 €** (inkl. MWST) darf nicht überschritten werden.

Der Gesamtpreis über die Positionen fließt mit 30% in die Bewertung ein, die übrigen 70% werden durch die Ausprägung der Bewertungskriterien bestimmt.

2. Leistungsbeschreibung

Die nachstehende Auflistung stellt die geforderten Spezifikationen dar. Gehen Sie in Ihrem Angebot konkret auf die einzelnen Anforderungen ein und geben Sie die Eigenschaften/Parameter aller angebotenen Komponenten an. Es können nur die Angaben bewertet werden, die in Ihrem Angebotstext beschrieben bzw. in der Leistungsbeschreibung ausgefüllt sind (Liefertermin, Gewährleistung, usw.).

Ausschlusskriterien werden im Nachfolgenden mit **A** gekennzeichnet, entsprechend die Bewertungskriterien mit **B**. Bewertungskriterien, die eine Mindestanforderung bzw. maximal Grenze haben, sind mit **B*** gekennzeichnet.

Bewertungskriterien werden mit Punkten bewertet. Zwischenwerte werden grundsätzlich interpoliert und auf zwei Nachkommastellen gerundet. Zum Beispiel (fiktiv):

Für 100 Cores werden 10 Punkte vergeben. 165 Cores würden 16,50 Punkte erhalten.

2.1 Allgemeine Systemanforderungen

Nachfolgend genannte Mindestanforderungen sind zu erfüllen und gelten systemweit für alle Knoten (Main-, Login-, Service-, Rechen- und GPU-Knoten), sofern nicht anders angegeben.

Nr.	Anforderungen	Kriterium
2.1.1	A	homogene x86-64 Systemarchitektur
2.1.2	A	die verfügbaren Speicherkanäle sind maximal zu bestücken, um die größtmögliche Speicherbandbreite zu erzielen

2.1.3	A	Fernwartbarkeit (zum Beispiel über IPMI) muss folgende Funktionalitäten erfüllen: <ul style="list-style-type: none"> • Alle Knoten müssen remote mit direktem Netzwerkzugriff administrierbar und überwachbar sein. • Die Möglichkeit zum Remote-Reset muss gegeben sein. • Die Konsole jedes Rechen- und Managementknotens soll remote zugreifbar sein. Dies beinhaltet auch den Zugriff auf die Firmware (BIOS-Einstellungen o.ä.).
2.1.4	A	Es sollen nur Prozessoren und Grafikprozessoren zum Zeitpunkt der Ausschreibung aktuellsten Generation angeboten werden. CPUs bzw. GPUs die zum Zeitpunkt der Angebotserstellung noch nicht verfügbar sind dürfen ebenfalls angeboten werden. Der genannte Liefertermin muss jedoch zwingend eingehalten werden.

Bereitgestellt werden Speicherbereiche für Home, Anwender- sowie Systemsoftware („Repo“) durch die vorhandene NetApp-Instanz. Hierzu ist eine Anbindung an das Campusnetzwerk notwendig, siehe Abschnitt 2.9 und Abbildungen 1+2.

2.2 Rechenknoten

Unter der Bezeichnung Rechenkern bzw. „Core“ werden die physischen Kerne eines Prozessors verstanden. Die Fähigkeit eines Prozessors mit hardwareseitigem Multithreading zu arbeiten wird nicht bewertet. Allein die Performance der physischen Kerne zählt.

Nr.	Anforderungen	Kriterium
2.2.1	B	Summe aller CPU Cores der Rechenknoten. Je 200 Cores werden 5 Punkte vergeben. Die maximal erreichbare Punktzahl beträgt 300 Punkte.
2.2.2	B	Basis-Taktfrequenz des Prozessors Je ein GHz werden 5 Punkte vergeben, die max. Anzahl an Punkten ist 25.

2.2.3	B*	<p>Mind. 4 GB Arbeitsspeicher mit ECC pro Rechenkern, insgesamt jedoch nicht weniger als 180GB RAM pro Rechenknoten.</p> <p>Für das Minimum von 4GB je Core werden 30 Punkte vergeben. Die maximal erreichbare Punktzahl von 50 wird ab 8GB je Core vergeben. Zwischenwerte werden interpoliert.</p>
2.2.4	B	<p>Transferrate des Arbeitsspeichers in "Megatransfers/Sekunde" (MT/s)</p> <p>Das Minimum von 4000MT/s wird mit 1 Punkt bewertet. Alle 200MT/s wird mit einem weiteren Punkt bewertet. Die maximal erreichbare Anzahl an Punkten ist 25 (bei 8800MT/s).</p>
2.2.5	A	<p>Festplattenfreier Boot und Betrieb („Diskless“)</p>
2.2.6	A	<p>Jeder Rechenknoten muss Zugriff auf die Nutzerdaten haben (Home- und Scratch-Speicher).</p>
2.2.7	A	<p>Jeder Rechenknoten muss Zugriff auf ein zentrales Softwareverzeichnis haben.</p>
2.2.8	A	<p>Jeder Rechenknoten muss die Möglichkeit zur Aufnahme eines lokalen Scratch-Speichers besitzen (zum Beispiel freier NVMe-Anschluss).</p>
2.2.9	A	<p>Drei Rechenknoten sollen mit einem lokalen Scratch-Speicher mit einer Größe von mind. 3TB ausgestattet sein.</p>
2.2.10	A	<p>Einzelne Rechenknoten müssen im laufenden Clusterbetrieb tauschbar/abschaltbar sein, ohne den Weiterbetrieb des Clusters zu beeinträchtigen.</p>

2.3 Mainknoten/Managementknoten

Der Mainknoten stellt zentrale Funktionen für den Rechencluster bereit. Außerdem dient er als Zugangspunkt für administrative Tätigkeiten am Cluster.

Nr.	Anforderungen	Kriterium
2.3.1	A	Prozessor Hersteller und Generation identisch zum Rechenknoten.
2.3.2	A	Mind. 16 Cores.
2.3.3	A	Mind. 48 GB Arbeitsspeicher mit ECC, mind. gleiche Taktfrequenz wie bei den Rechenknoten.
2.3.4	A	Anbindung an das Hochgeschwindigkeits-Netzwerk (siehe Abschnitt 2.9 und Abbildungen 1+2)
2.3.5	A	Bootfestplatte (SSD) ist redundant auszulegen
2.3.6	A	Bootfestplatte verfügt über mind. 1TB Kapazität.
2.3.7	A	Freie Festplattenanschlüsse im Chassis des Mainknotens sind mit weiteren Festplatten (HDD) zu bestücken, Redundanz ist zu gewährleisten, dient als Ablage für administrativ anfallende Daten.

2.3.8	A	Netto-Kapazität dieser zusätzlichen Festplatten mind. 5TB.
2.3.9	A	Auf dem Mainknoten laufen die für den Betrieb des Gesamtsystems erforderlichen Dienste: <ul style="list-style-type: none"> • Clustermanagementsoftware • Batchsystem (siehe Software) • Monitoringsysteme • Lizenzdienste • Logserver • weitere notwendige Dienste

2.4 Loginknoten

Der Loginknoten dient dem Zugriff auf den Rechencluster und die Speicherbereiche für die Nutzer. Zudem soll er zum interaktiven Pre- und Postprocessing, Kompilieren und Debuggen von Software dienen.

Nr.	Anforderungen	Kriterium
2.4.1	A	Prozessor Hersteller und Generation identisch zum Rechenknoten
2.4.2	A	L3 Cache identisch zum Prozessormodell der Rechenknoten
2.4.3	A	Mind. 32 Cores und mind. 750GB Arbeitsspeicher (Spezifikationen identisch zu den Rechenknoten)
2.4.4	A	Anbindung an das Hochgeschwindigkeits-Netzwerk (siehe Abschnitt 2.9 und Abbildungen 1+2)

2.4.5	A	Festplattenfreier Boot und Betrieb („Diskless“)
2.4.6	A	Lokaler Scratch-Speicher von mind. 512 GB, redundant ausgelegt
2.4.7	A	Rechenjobabgabe an das Batchsystem ist möglich (siehe Software)
2.4.8	A	Einbindung der Verzeichnisstruktur für Softwarekomponenten, identisch zu den Rechenknoten.
2.4.9	A	Auf dem Loginknoten laufen keine für den Betrieb des Gesamtsystems erforderlichen Dienste (z.B. Batchsystem-Server, Clustermanagementssysteme und evtl. weitere).

2.5 Service-Headnode

Der Service-Headnode dient als Hosting-Plattform für zusätzliche Dienste. Hier sollen per Virtualisierung mittels KVM/QEMU und mit einer Container-basierten Lösung (z.B. Docker) Anwendungen bzw. Plattformen wie JupyterHub und OpenOnDemand für die Nutzer bereitgestellt werden können.

Nr.	Anforderungen	Kriterium
2.5.1	A	Prozessor identisch zum Mainknoten
2.5.2	A	Mind. 250 GB Arbeitsspeicher
2.5.3	A	Bootfestplatte (SSD) ist redundant auszulegen
2.5.4	A	Bootfestplatte verfügt über mind. 4TB Kapazität
2.5.5	A	Anbindung an das Hochgeschwindigkeits-Netzwerk (siehe Abschnitt 2.9 und Abbildungen 1+2)
2.5.6	A	Der Server muss für potentielle Erweiterungen im Betriebszeitraum mindestens zwei freie PCI-Express Steckplätze besitzen. Die Anzahl der auf diesen freien Ports zur Verfügung stehenden PCIe-Lanes und Slot-Größen bzw. -Formfaktor sind anzugeben.
2.5.7	A	Eine zusätzliche PCIe-Steckkarte für das Hochgeschwindigkeitsnetzwerk sowie ein passendes Kabel zum Anschluss wird mitgeliefert, jedoch nicht verbaut. Diese Karte wird bei Bedarf im Betriebszeitraum eingebaut.

2.6 GPU-Knoten

Die GPU-Knoten sind vorgesehen für Berechnungen, die durch einen Grafikprozessor beschleunigt werden. Die Knoten sollen für die Anwendungen in den Feldern *Machine Learning*, *Deep Learning* oder *Artificial Intelligence* geeignet sein. Um anspruchsvolle Aufgaben im Bereich Künstliche Intelligenz (KI) und wissenschaftliches Rechnen (SciML) zu bewältigen – insbesondere solche, die physikbasierte neuronale Netze und große Sprachmodelle (LLMs) mit einem Kontext von 128 Gigabyte oder mehr nutzen – ist eine große Menge an High Bandwidth Memory (HBM) notwendig. Zudem finden sie Anwendung im Rechencluster zum Pre- und Postprocessing, sowie zur Vorbereitung der Visualisierung von Rechenergebnissen.

Nr.	Anforderungen	Kriterium
2.6.1	A	Prozessor Hersteller und Generation identisch zum Rechenknoten
2.6.2	A	L3 Cache identisch zum Prozessormodell der Rechenknoten
2.6.3	A	Arbeitsspeicher von mind. 512GB
2.6.4	A	Es sollen ausschließlich GPUs für Rechenzentren (Data-Center GPUs), die für den Hochleistungsrechenbetrieb geeignet sind, angeboten werden.
2.6.5	A	Grafikprozessor mit nativer Kompatibilität zu folgenden Schnittstellen: CUDA (mind. 12.0), OpenCL (mind. 3.0)
2.6.6	A	Insgesamt mind. 16 GPUs Die Bewertung der angebotenen GPUs erfolgt in dem Punkt 2.6.7.

2.6.7	B*	<p>Mind. 4 GPUs müssen verbaut sein, die die Fähigkeit zur Verarbeitung von 64-Bit-Gleitkommazahlen (FP64) besitzen.</p> <p>Die GPUs werden nach Typen getrennt bewertet.</p> <p>Es werden je 5 Punkte für eine GPU der Klasse FP64 vergeben, hier ist die maximale Punktzahl 30 Punkte.</p> <p>Für GPUs ohne Verarbeitung von FP64 werden 5 Punkte vergeben, hier ist die maximale Punktzahl 70 Punkte.</p>
2.6.8	A	<p>GPUs der Klasse FP64 müssen über mindestens 140 GB HBM-Speicher verfügen. GPUs der non-FP64-Klasse müssen mindestens 48 GB GPU-Speicher (HBM oder GDDR) besitzen.</p>
2.6.9	A	<p>Es dürfen max. zwei verschiedene GPU-Modelle angeboten werden. Der Hersteller muss identisch sein.</p>
2.6.10	A	<p>Die GPUs müssen für den globalen Markt entwickelt sein und dürfen keine regionalspezifischen Anpassungen aufweisen.</p>

2.7 SCRATCH

Das parallele Scratch-Filesystem stellt den zentralen Arbeitsbereich des Clusters dar. Rechenjobs nutzen diesen Speicher u.a. als temporären Speicherplatz und für Ein- und Ausgabedateien, die während der Jobabarbeitung entstehen. Im Pre- und Postprocessing wird ebenfalls mit angefallenen Daten in diesem Speicherbereich gearbeitet. Die Performanceanforderungen an dieses System sind besonders hoch, da sonst der eigentliche Rechenbetrieb ausgebremst werden kann.

Nr.	Anforderungen	Kriterium
2.7.1	B*	Mind. 1PB Brutto-Kapazität des Gesamtsystems. Es wird 1 Punkt je 13TB vergeben. Die maximale Anzahl an Punkten ist 120.
2.7.2	B*	Mind. 150TB Brutto-Kapazität des SSD-Speichers. Es wird 1 Punkt je 8TB an SSD-Speicher vergeben. Die maximale Punktzahl ist 25.
2.7.3	A	Management via Webinterface
2.7.4	A	Management via Kommandozeilenschnittstelle, die auch mit Skripten automatisierbar ist.
2.7.5	A	Monitoring via SNMP Protokoll
2.7.6	A	Netzwerkanbindung an das Hochgeschwindigkeitsnetzwerk; siehe Abschnitt 2.9.

2.7.7	A	Bei Stromunterbrechung/-ausfall dürfen keine Nutzerdaten verloren gehen, entsprechende Sicherheitsmechanismen müssen ab Werk integriert sein.
2.7.8	A	Das Speichersystem darf keinen „single point of failure“ besitzen. Der Ausfall einzelner Komponenten des Scratch-Speichers muss durch werksseitige, automatisierte Mechanismen abgefangen werden und darf Nutzerdaten nicht beschädigen.
2.7.9	A	Alle Funktionsmerkmale müssen ohne zusätzliche Lizenzkosten nutzbar sein.
2.7.10	A	Der Scratch-Speicher folgt dem „Appliance“-Ansatz und wird in vollem Umfang (Hard- und Software) von einem Hersteller unterstützt.
2.7.11	A	Eine unterbrechungsfreie Erweiterung im Betriebszeitraum muss realisierbar sein.
2.7.12	B	Die Erstellung von Speicher-Volumes, die nur den SSD-Speicher des Gerätes nutzen, ist möglich. Ist die Möglichkeit gegeben werden 5 Punkte erzielt.
2.7.13	A	Das angebotene Scratch-Speichersystem muss eine Erasure Coding-Implementierung einsetzen, die mindestens die Schutzklasse eines RAID6-Äquivalents bietet.
2.7.14	A	Integrierter, für Nutzer transparenter Mechanismus zur Verteilung der Daten auf verschiedene Performanceklassen der Datenspeicher: <ol style="list-style-type: none"> 1. NVMe SSD 2. SATA/SAS SSD 3. SATA/SAS HDD Mind. 2 der genannten Geräteklassen müssen vorhanden sein.

2.7.15	A	Das Speichersystem muss mind. einem nativen, für parallelen Zugriff optimierten Protokoll arbeiten.
2.7.16	A	Die Clientsoftware für das Zugriffsprotokoll muss ein Installationspaket passend zur Paketverwaltung des Betriebssystems besitzen oder als Quellcode bereitgestellt werden, sodass eine manuelle Erstellung eines Pakets möglich ist.
2.7.17	B	Die Möglichkeit zur Verwendung der alternativen Storageprotokolle NFS und S3 ist ohne zusätzliche Kosten gegeben. Ist diese Möglichkeit gegeben werden für NFS 5 Punkte vergeben, für S3 werden ebenfalls 5 Punkte vergeben.
2.7.18	B	Das Speichersystem muss die Erstellung (manuell und per Zeitplan) von Snapshots (oder ähnliche Mechanismen) unterstützen. Der Zugriff auf die Daten in den Snapshots ist ohne Eingriff eines Administrators auf Nutzerebene möglich. Für die Möglichkeit Snapshots zu erstellen werden 10 Punkte vergeben.
2.7.19	A	Die Speicherlösung unterstützt Online-Updates für einzelne Unterkomponenten.
2.7.20	A	Einzelne Komponenten müssen vor Ausfall (bei steigender Fehlerzahl) aus dem System entfernt werden können, um einen transparenten Austausch zu ermöglichen.
2.7.21	A	Die Möglichkeit zum Setzen von Prioritäten im Falle einer Rekonstruktion von Daten muss gegeben sein.

2.7.22	A	Um die Verfügbarkeit der Komponenten bis zu 5 Jahren sicherzustellen, dürfen nur Enterprise-Komponenten, insbesondere HDD und SSD, im System verbaut werden.
2.7.23	A	Ein automatisches Ausbalancieren der Daten nach dem Austausch einzelner, defekter Komponenten bzw. nach der Erweiterung des Systems durch neue Komponenten muss gegeben sein.

2.8 Integration in das Rechenzentrum

Für die Integration in die Infrastruktur des Rechenzentrums sind folgende Punkte von Bedeutung: Der Einbau der Hardware erfolgt in drei zu liefernde Serverschränke mit je einem Luft-/Wasser-Wärmetauscher. Die Serverschränke und Wärmetauscher sind an das bestehende Rittal-Racksystem anzureihen (vorhandenes LCP SK 3312.239 (30kW)) und in die vorhandene Management- und Monitoring-Lösung Rittal CMC zu integrieren. Um eine reibungslose Integration in die bestehende Infrastruktur des Rechenzentrums zu gewährleisten ist ein verpflichtender Vororttermin vorgesehen. Aufbau und Betrieb muss ohne Beeinträchtigung der vorhandenen Technik erfolgen. Erstellen Sie eine eigene Position in Ihrem Angebot für die Infrastrukturkomponenten. Das Budget von **100.000€** (inkl. MWST) darf nicht überschritten werden.

Nr.	Anforderungen	Kriterium
2.8.1	A	Integration aller Komponenten in drei zu liefernde Server-Racks (Höhe 2200mm, Tiefe 1200mm, Breite 800mm). Die Server-Racks sind mit 200mm Sockeln auszustatten. Die Schränke werden in die vorhandene Rauchfrüherkennung eingebunden, ein automatisches Öffnen der Schranktüren im Erkennungsfall zur Einleitung des Löschgases ist zu gewährleisten. Türen frontseitig sind als Sichttüren (transparent) auszuführen. Die Schränke sind verschließbar. Bei der Bestückung der Schränke ist wenigstens 15% freie Reserve für Erweiterungen je Schrank zu planen.
2.8.2	A	Die Nutzkühlleistung eines einzelnen Schrankes beträgt 30 kW. Die Kühlung erfolgt über Luft-/Wasser-Wärmetauscher (Höhe 2200mm, Tiefe 1200mm, Breite 300mm). Die Wärmetauscher sind zwischen den Server-Racks zu platzieren und mit 200mm Sockeln zu versehen. Es ist insgesamt 15% Reserve an Kühlleistung für Erweiterungen einzuplanen. Die Wärmetauscher bieten Funktionen zur Leckageerkennung sowie zur Überwachung technischer Parameter (Durchfluss, Ein- und Ausgangstemperaturen). Der Wasseranschluss der Luft-/Wasser-Wärmetauscher erfolgt über im Doppelboden vorhandene Anschlüsse.

2.8.3	A	Für die Stromversorgung der Server sind entsprechende PDUs zu liefern. Pro Schrank stehen je zwei 3-Phasen-Einspeisungen (32A) für Allgemein- und USV-Strom zur Verfügung (CEE-Anschlüsse). Die PDUs sind je Ausgangsdose mit Funktionalitäten zur Messung des Energiebedarfs, zur Schaltung sowie zur Überwachung ausgestattet. Die PDUs werden in das Monitoring integriert und sind per SNMPv3 ausles- und steuerbar. Bei der Anzahl der Steckplätze ist insgesamt 15% Reserve für Erweiterungen einzuplanen.
2.8.4	A	Zum sicheren Schutz der Hardware vor Überhitzung bzw. Ausfall der Klimatisierung muss der Auftragnehmer ein entsprechendes Eskalationsverfahren bis hin zur Komplettabschaltung des HPC-Clusters in die gelieferten Schränke implementieren. Das Prinzip ist im Angebot darzustellen.
2.8.5	A	Ein Aufstellungs- und Verkabelungsplan ist beizulegen.
2.8.6	A	Die drei zu liefernden Racks und Luft/Wasser-Wärmetauscher sowie das zugehörige Management- und Monitoringsystem sind in das Bestandssystem basierend auf Rittal CMC III zu integrieren. Die Racks sind an ein vorhandenes LCP Rittal SK 3312.239 (30kW) anzureihen.
2.8.7	A	Es ist ein Wartungsvertrag über 5 Jahre für die Kühlmodule anzubieten. Dieser soll als optionaler Posten aufgeführt sein.

2.9 Netzwerke

Die Anforderungen an die Kommunikationshardware sind im folgenden Abschnitt dargestellt.

Das zuliefernde System beinhaltet drei Arten von Netzwerken:

1. Clusternetzwerk
Dieses Netzwerk dient zur Kommunikation clusterinterner Dienste, z.B. des Stapelsystems, Zugriff auf Home- und Software-Speicherbereiche, u.a. auch zum festplattenfreien Boot der Knoten.
2. Managementnetzwerk
Dieses Netzwerk dient zum Zugriff auf alle Management-Zugänge der verbauten Komponenten, z.B. IMPI-Funktionalität der Knoten.
3. Hochgeschwindigkeitsnetzwerk
Dieses Netzwerk dient zur parallelen, latenzarmen Kommunikation während der Ausführung von Jobs.

Das Clusternetzwerk soll über einen zentralen, vollständig redundanten Core-Switch verfügen. Der Login-, Main- sowie Serviceknoten sind jeweils redundant mit 25Gbit daran anzubinden. Dieser Core-Switch wird per LACP-Trunk an zwei vorhandene 25G-SR-Interfaces an Cisco Nexus 9000-Switches angeschlossen. Über diese ist der vorhandene Storage für die weiteren Speicherbereiche (siehe „2.1 Allgemeine Systemanforderung“), sowie das Campus-Netz und das Internet erreichbar.

Die Integration der einzelnen HPC-Cluster-Nodes in das clusterinterne Netzwerk ist abhängig von der konkreten Ausgestaltung des Knoten-Setups, ist aber mit mindestens 10Gbit/s pro Knoten zu realisieren. Auf eine redundante Anbindung der Rechenknoten ist zu verzichten. Der in der Abbildung 1 angedeutete "Computenode Switch" ist exemplarisch, eine nicht-redundante Anbindung über verfügbare Ports am Core-Switch ist zulässig.

Alle Geräte, die ein dediziertes Out-of-Band Management Interface besitzen (z.B. IPMI oder ähnlich), sollen an das IPMI-Management-Netz (siehe Abschnitt 2.9 Managementnetzwerk) angeschlossen werden, die Anbindung geschieht per 1Gbit-Links.

Zwischen allen Knoten sowie dem parallelen Filesystem besteht ein Hochgeschwindigkeitsnetzwerk mit mindestens 200Gbit-Links (siehe Abschnitt 2.9 Hochgeschwindigkeitsnetz).

Der IPMI-Zugang zum Mainknoten ist per 1Gbit-Kupfer-Link in das URZ-Management-Netz zu integrieren. Eine Skizzierung der Netzwerkstruktur ist in Abbildung 1 visualisiert.

Die für die Nutzer erreichbaren Dienste des Login- und Serviceknoten werden über eine vorhandene DMZ bereitgestellt. Für beide Nodes werden öffentliche IP-Adressen

aus dem TUBAF-Netzbereich bereitgestellt, der Datenverkehr ist mit einem VLAN zu taggen.

Das clusterinterne Netzwerk und das IPMI-Netzwerk verwenden private IP-Adressen. Das clusterinterne Netzwerk ist per VLAN zu taggen.

Das Hochgeschwindigkeitsnetz hat keine Berührungspunkte mit dem Campus- und clusterinternen Netz und wird ungetaggt über eigene Switches realisiert.

Eine schematische Darstellung der Netzbereiche finden Sie in Abbildung 2.

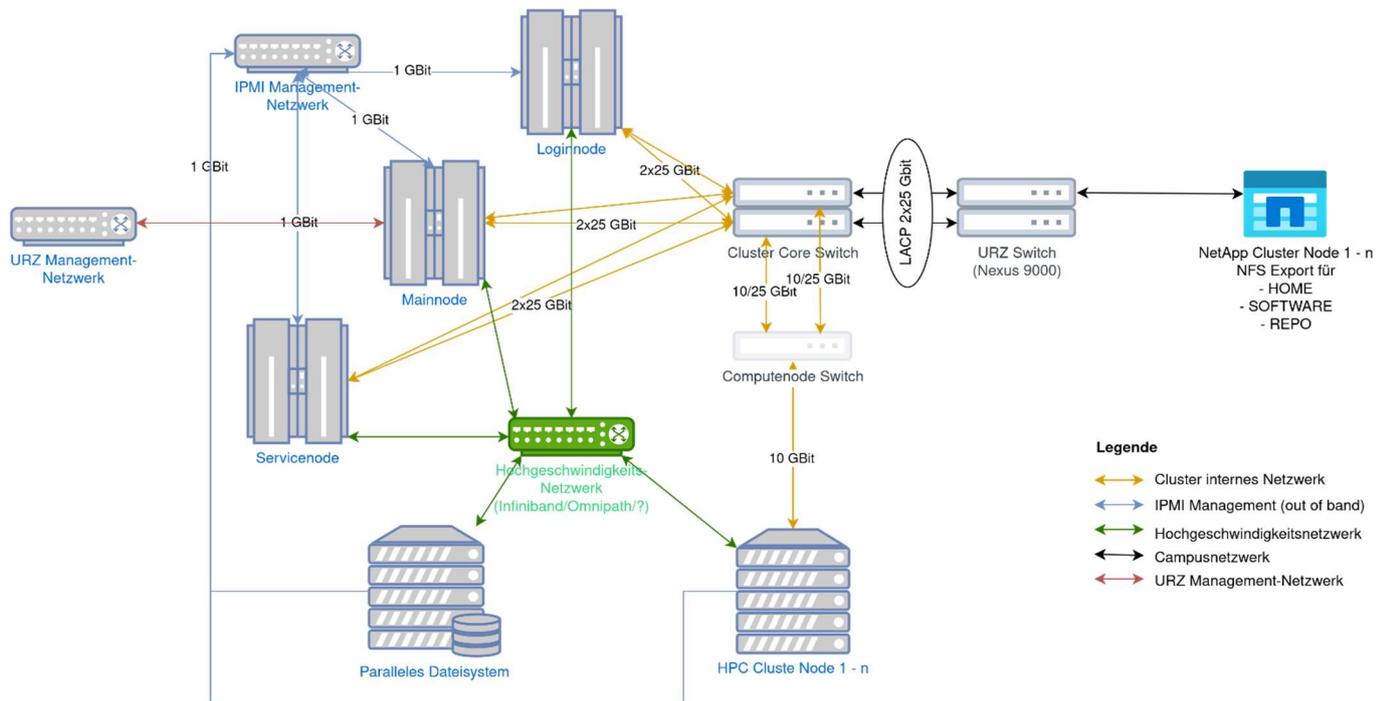


Abbildung 1; Skizzierung der vorgesehenen Netzwerkstruktur und Anbindung an das Campusnetz.

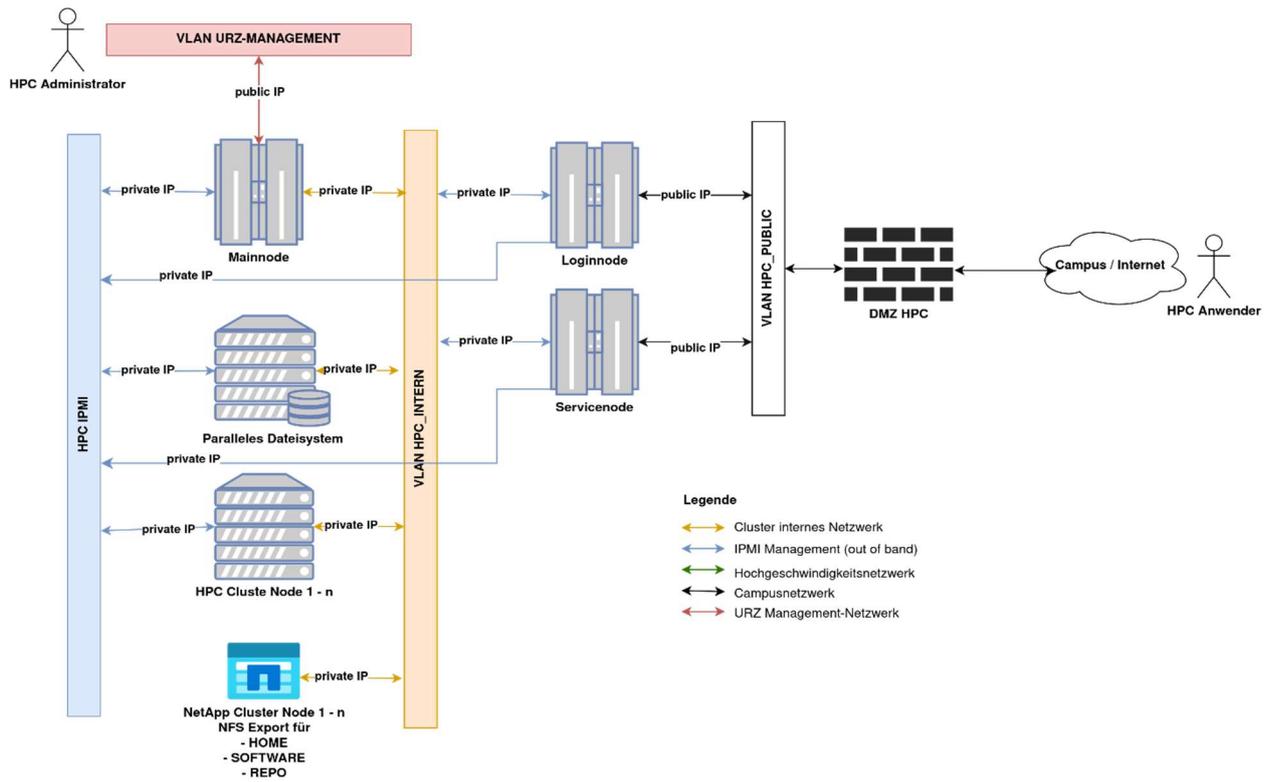


Abbildung 2: Skizzierung der VLAN- und Netzstruktur, sowie Vergabe der IP-Adressen.

Hochgeschwindigkeitsnetzwerk

Nr.	Anforderungen	Kriterium
2.9.1	A	Das Cluster ist mit einem Hochgeschwindigkeitsnetzwerk auszustatten.
2.9.2	A	Bandbreite mind. 200 GBit/s
2.9.3	A	Die Topologie des Hochgeschwindigkeitsnetzwerk ist so zu gestalten, dass der Blockingfaktor maximal 1:2 beträgt. Die Anzahl und Größe der Switches ist anzugeben und die Topologie des Netzwerkes ist zu dokumentieren.

2.9.4	A	Das Hochgeschwindigkeitsnetzwerk bietet eine native RDMA-Unterstützung auf Hardwareebene.
2.9.5	A	Die Möglichkeit zur IP-Kommunikation über das Hochgeschwindigkeitsnetz muss gewährleistet sein.
2.9.6	A	Um die zukünftige Erweiterbarkeit des Clusters zu gewährleisten, sollen 15% der Netzwerkanschlüsse frei zur Verfügung stehen.
2.9.7	B	Die Möglichkeit zum modularen Austausch von Lüftern und/oder Netzteilen ist gegeben. Lüfter tauschbar – 2 Punkte Netzteil tauschbar – 2 Punkte Beides tauschbar – 5 Punkte

Clusternetzwerk

Nr.	Anforderungen	Kriterium
2.9.8	A	Basierend auf Ethernet
2.9.9	A	Mind. 10 Gbit/s Anbindung der Rechen- und GPU-Knoten
2.9.10	A	Mind. 25 Gbit/s redundante Anbindung des Main-, Login- und Serviceknoten an den Clustercoreswitch.

2.9.11	A	Um die zukünftige Erweiterbarkeit des Clusters zu gewährleisten, sollen mind. 15% der verfügbaren Netzwerkanschlüsse frei bleiben.
2.9.12	A	Die Netzwerkkomponenten müssen VLAN-fähig sein.
2.9.13	B	Die Möglichkeit zum modularen Austausch von Lüftern und/oder Netzteilen ist gegeben. Lüfter tauschbar – 2 Punkte Netzteil tauschbar – 2 Punkte Beides tauschbar – 5 Punkte

Managementnetzwerk

Nr.	Anforderungen	Kriterium
2.9.14	A	Basierend auf Ethernet
2.9.15	A	Mind. 1 Gbit/s Anbindung der Management-Schnittstellen an allen Knoten (Main-, Login-, Service-, Rechen- und GPU-Knoten) sowie weiterer managebarer Komponenten
2.9.16	A	Interconnect der Management-Switches (falls vorhanden) mind. 10 Gbit/s
2.9.17	A	Um die zukünftige Erweiterbarkeit des Clusters zu gewährleisten, sollen mind. 15% der verfügbaren Netzwerkanschlüsse frei bleiben.

2.9.18	B	Die Möglichkeit zum modularen Austausch von Lüftern und/oder Netzteilen ist gegeben. Lüfter tauschbar – 2 Punkte Netzteil tauschbar – 2 Punkte Beides tauschbar – 5 Punkte
--------	---	---

Anbindung an das Campusnetz

Alle vom Cluster ausgehenden Netzwerkverbindungen sollen auf einem redundanten Stapel aus Cisco Nexus 9000 Switches zusammenlaufen. Dieser muss um entsprechende Ports erweitert werden. Daher muss das Angebot folgende Leistungsmerkmale umfassen:

Nr.	Anforderungen	Kriterium
2.9.19	A	Entsprechende Anzahl an 25GBit SFP+ Module (multimode)
2.9.20	A	Die passende Anzahl an LWL-Kabeln
2.9.21	A	Es ist eine redundante Anbindung mit 25 GBits/s an den Switch des URZ (Cisco Nexus 9000) bereitzustellen. Diese Verbindung soll für die Anbindung des zentralen Software- und Homeverzeichnis genutzt werden. (siehe Abbildung des Netzwerkstruktur)

2.10 Software

Folgende Softwareausstattung soll dem System beigefügt und installiert werden:

Nr.	Anforderungen	Kriterium
2.10.1	A	Einheitliches Betriebssystem für alle Knoten (Main-, Login-, Service-, Rechen und GPU-Knoten), basierend auf einer quelloffenen Linuxdistribution.
2.10.2	A	Das Betriebssystem soll für mindestens 5 Jahre upgradefähig sein. Die Upgradefähigkeit zu neueren Versionen im Laufe des Betriebszeitraums ist zu gewährleisten und darf durch keine Softwarekomponente, wie z.B. durch das Clustermanagementsystem, verhindert werden.
2.10.3	A	Cluster-Management-Systemkomponenten: 1) Monitoring a) Überwachung von Systemstatusinformationen (logs) aller Knoten. b) Meldung von Ausfällen c) Protokollierung der Auslastung d) Abrufbar per Web-Schnittstelle e) Online-Updatefähig 2) Automatisierte Installation der Rechenknoten 3) Imageverwaltung des Betriebssystems
2.10.4	A	Nutzerverwaltung, Autorisierung und Authentifizierung basierend auf der vorhandenen Microsoft Active Directory-Struktur. Die Möglichkeit zur Verwendung von Active Directory Sicherheitsgruppen in den Dateisystemen und den Queues des Stapelsystems des Clusters ist zu gewährleisten.
2.10.5	A	Batch-System inklusive Betriebslizenz (falls benötigt) für unbestimmte Zeit. Folgende Leistungsmerkmale müssen abgedeckt sein: a) Gruppenbasierende Jobverteilung b) Priorisierung der Nutzergruppen c) Anpassbare Jobsortierung d) Unterstützung für Linux Control Groups (cgroups) e) Skriptbasierende Erweiterbarkeit von Funktionalitäten

2.10.6	A	<p>Compiler und Bibliotheken:</p> <ul style="list-style-type: none"> • Intel, Fortran, C, C++ Compiler-Suite inklusive aller notwendigen Lizenzen. Floating-Lizenz für einen Nutzer, ausführbar auf allen Knoten. • Intel MPI Bibliothek • Hochoptimierte mathematische Bibliotheken, die mind. die volle BLAS- und LAPACK Funktionalität enthalten • OpenMP-Bibliotheken für C/C++ und Fortran • OpenMPI mind. MPI-Standard 4.0 • MPICH MPI-Libraries • Die Bereitstellung der Bibliotheken sollte über die Open-Source Software Spack erfolgen
--------	---	---

2.11. Umwelt und Energie

Nr.	Anforderungen	Kriterium
2.11.1	B	<p>Wirkungsgrad der Netzteile der Server (80 PLUS Rating).</p> <p>Die Höchstpunktzahl von 10 wird mit 80 PLUS Titanium erreicht. Mit jeder Stufe weniger werden 2 Punkte abgezogen (demnach 80 PLUS Bronze = 2 Punkte, rein 80 PLUS oder keines = 0 Punkte).</p>
2.11.2	B*	<p>Kühlbedarf des Gesamtsystems unter Maximallast (max. 75 kW um den Kühlbedarf potentieller Erweiterungen zu decken). Folgende Parameter sind verbindlich anzugeben:</p> <ol style="list-style-type: none"> 1) Energiebedarf in kW <ol style="list-style-type: none"> a) Bei einer zu erwartenden Produktivlast von 70% b) Unter Maximallast 2) Kühlbedarf in kW <ol style="list-style-type: none"> a) Bei einer zu erwartenden Produktivlast von 70% b) Unter Maximallast <p>Es können maximal 40 Punkte erreicht werden. 75 kW – 10 Punkte 70 kW – 20 Punkte 65 kW – 30 Punkte <=60 kW – 40 Punkte</p>

2.12 Erweiterbarkeit

Nr.	Anforderungen	Kriterium
2.12.1	A	Eine Erweiterung an Knoten im Betriebszeitraum ist vorgesehen. Diese müssen sich störungsfrei integrieren lassen
2.12.2	A	Eine Erweiterung des SCRATCH-Speichers im Betriebszeitraum ist vorgesehen. Diese sollte sich störungsfrei integrieren lassen.

3. Benchmarks

3.1 Vorbemerkungen

Im Rahmen der vorliegenden Ausschreibung sind Benchmarks verpflichtend durchzuführen.

Nachdem die Installation des Clustersystems erfolgt ist, werden alle hier aufgeführten Benchmarks auf diesem System durchgeführt und müssen mind. die gleiche Benchmarkleistung erreichen, die der Anbieter in seinen Tests für das Angebot erzielt hat. Eine Abweichung von bis zu 5% ist zulässig. Sollten diese nicht erreicht werden ist der Bieter verpflichtet Nachbesserungen zu leisten, bis die zugesicherte Leistung erzielt wurde.

Die von uns bereitgestellte Software sowie alle Skripte und weitere Dateien (außer Open Source Pakete) sind nicht an Dritte weiterzugeben und nach Abschluss des Benchmarks wieder zu löschen.

Folgende Benchmark-Pakete sollen durchgeführt werden:

- **HPCG**
- **STREAM**
- **OpenFOAM**

Die Verwendung von vorkompilierten aus dem Internet geladenen Containern ist zulässig und zum Teil vorgesehen, die Quelle des Containers ist anzugeben. Ebenso zulässig ist die Verwendung eines Paketmanagers (zum Beispiel Spack) zum Installieren der Software.

Sollten Sie die Performance durch Compiler-Flags oder ähnliches optimieren, so ist Ihr Vorgehen genau zu dokumentieren.

Sollten Sie bei den vorgeschriebenen Versionen Konflikte feststellen, so können Sie die Versionen ändern. Änderungen sind anzugeben und zu dokumentieren.

3.2 Hardware

Die Hardware, auf denen die Benchmarks ausgeführt werden, soll den angebotenen Komponenten möglichst nahekommen. Zudem sollte die jeweils aktuellste Firmware installiert sein. Sollte die angebotene Hardware ein noch nicht veröffentlichtes Release-Datum aufweisen, sind Benchmarks auf Basis vergleichbarer, aktuell verfügbarer Hardware durchzuführen. Die zugesicherte Leistung muss bei Abnahme des Systems mit der tatsächlichen Leistung der verbauten Hardware übereinstimmen oder diese übertreffen.

3.3. Zusammenstellung der Ergebnisse und Punktevergabe

Ein Benchmark erhält 0 Punkte im Falle einer unvollständigen Dokumentation bzw. sollten nicht alle log und output-Dateien eingereicht werden. Wir werden zwischen den Werten generell linear interpolieren. Die Anzahl der erreichbaren Punkte ist im jeweiligen Benchmark genannt.

Die Dokumentation muss mindestens folgende Werte/Dateien enthalten:

- Art der Knoten (Anzahl der CPUs und Cores, Hauptspeicher, Speicherbandbreite)
- Prozessormodell
- Alle relevanten lokalen build-Prozess Outputs in Logfiles
- Bei Verwendung eines Batchsystems: der „Submit“-Befehl und/oder das Jobskript
- Quellen bei der Verwendung von vorkompilierten Containern.

Nach dem Test:

- Outputdateien
- Laufzeit (Wallclocktime)

Im Folgenden bezeichnen wir die GPUs der FP64-Klasse mit **Typ(FP64)** und die GPUs der non-FP64-Klasse mit **Typ(FP32)**.

3.3.1 HPCG (maximal 30 Punkte)

Führen Sie den HPCG Benchmark auf den Rechenknoten (CPU) und den GPU Knoten aus. Zur Beurteilung des Benchmarks werden wir den GFlops/s RAW Wert verwenden. Alle anderen Werte werden nicht berücksichtigt.

Rechenknoten

Für die Rechenknoten installieren Sie den HPCG Benchmark in Version 3.1 (Downloadlink <https://www.hpcg-benchmark.org/downloads/hpcg-3.1.tar.gz>).

Verwenden Sie bitte zum Kompilieren den GCC Compiler Version 11.4.0 und für das Message Passing Interface (MPI) nutzen Sie OpenMPI Version 4.1.8.

Der Inhalt der hpcg.dat Datei sollte dem folgenden entsprechen:

```
HPCG benchmark input file
Sandia National Laboratories; University of Tennessee, Knoxville
104 104 104
60
```

Führen Sie den Benchmark auf folgenden CPU-Ressourcenkonfigurationen aus:

- 1 Knoten, 1 Sockel mit 20 MPI-Prozessen

<15 GFlop/s	0 Punkte
15 GFlop/s	1 Punkt
>45 GFlop/s	5 Punkte
- 1 Knoten, alle Sockel, mit Verwendung aller möglichen CPU Cores (kein Hyper-Threading). Der GFlop/s Wert wird hier durch die Anzahl der vorhandenen Cores geteilt.

<0,6 GFlop/s	0 Punkte
0,6 GFlop/s	1 Punkt
>3,0 GFlop/s	5 Punkte
- 2 Knoten, je ein Sockel mit je 20 MPI-Prozessen.

<25 GFlop/s	0 Punkte
25 GFlop/s	1 Punkte
>75 GFlop/s	5 Punkte
- 2 Knoten, alle Sockel, mit Verwendung aller möglichen CPU Cores (kein Hyper-Threading). Der GFlop/s Wert wird hier durch die Anzahl der vorhandenen Cores geteilt.

<1,2 GFlop/s	0 Punkte
1,2 GFlop/s	1 Punkt
>6,0 GFlop/s	5 Punkte

GPU-Knoten

Nutzen Sie zur Durchführung des HPCG-Benchmarks auf den GPUs den Nvidia HPC Benchmark Container (<https://catalog.ngc.nvidia.com/orgs/nvidia/containers/hpc-benchmarks>). Für den Download ist ein kostenfreier Account bei NGC (<https://ngc.nvidia.com/signin>) und ein API-Key notwendig. Eine Anleitung zum Erhalt eines API Keys finden hier:

<https://docs.nvidia.com/ai-enterprise/deployment/spark-rapids-accelerator/latest/appendix-ngc.html>

Nutzen Sie bitte Singularity als Containerlösung (die Verwendung von Apptainer ist ebenfalls zulässig). Singularity benötigt den golang-compiler.

Der Nvidia HPC Benchmark Container kann wie folgt geladen werden:

```
export SINGULARITY_DOCKER_USERNAME='$oauthtoken'
```

```
singularity pull --docker-login hpc-benchmarks:25.04.sif  
docker://nvcr.io/nvidia/hpc-benchmarks:25.04
```

Enter Docker / OCI registry password:

Bei dieser Abfrage muss der Nvidia API Key eingefügt werden.

Führen Sie den Benchmark auf jedem angebotenen GPU-Typ aus. Sollten Sie nur einen GPU-Typ (FP64) anbieten (siehe Leistungsverzeichnis 2.6.6 – 2.6.8), werden die Ergebnisse für beide Punktevergaben gewertet.

Typ 1 (FP64)

Führen Sie den Container wie folgt aus

```
singularity run --nv hpc-benchmarks\\:25.04.sif /workspace/hpcg.sh --  
nx 256 --ny 256 --nz 256 --rt 2
```

<350 GFlop/s	0 Punkte
350 GFlop/s	1 Punkt
>550 GFlop/s	5 Punkte

Typ 2 (FP32)

Führen Sie den Container wie folgt aus

```
singularity run --nv hpc-benchmarks\\:25.04.sif /workspace/hpcg.sh --  
nx 128 --ny 128 --nz 128 --rt 2
```

<150 GFlop/s	0 Punkte
150 GFlop/s	1 Punkt
>350 GFlop/s	5 Punkte

3.3.2 STREAM (maximal 15 Punkte)

Führen Sie den STREAM Benchmark auf den Rechenknoten (CPU) und den GPU Knoten aus. Zur Beurteilung des Benchmarks werden wir den ausgegebenen „Triad“-Wert verwenden. Alle anderen Werte werden nicht berücksichtigt.

Rechenknoten

Nutzen Sie bitte zur Installation den GCC-Compiler Version 11.4.0 und führen Sie die folgenden Schritte durch:

```
git clone https://github.com/jeffhammond/STREAM.git
cd STREAM/
make
```

Bitte beachten Sie, dass Sie je nach Hardware die `STREAM_ARRAY_SIZE` in `stream.c` anpassen und neu kompilieren müssen, bevor Sie den Benchmark ausführen.

Die `STREAM_ARRAY_SIZE` sollte $10^6 \cdot \text{L3 Cache-Size in MB}$ entsprechen.

Dokumentieren Sie Ihre Kompilierung. Nutzen Sie für den Benchmark den C-Code.

Führen Sie den STREAM Benchmark auf einem Sockel mit allen verfügbaren Cores aus.

<150 GB/s	0 Punkte
150 GB/s	1 Punkt
>500 GB/s	5 Punkte

GPU Knoten

Nutzen Sie für den Stream Benchmark auf den GPUs ebenfalls den Nvidia HPC Benchmark Container. Instruktion hierzu finden Sie im Abschnitt des HPCG Benchmarks.

Führen Sie den Benchmark auf jedem angebotenen GPU-Typ aus. Sollten Sie nur einen GPU-Typ (FP64) anbieten (siehe Leistungsverzeichnis 2.6.6 – 2.6.8), werden die Ergebnisse für beide Punktevergaben gewertet.

Führen Sie den STREAM, Benchmark auf der GPU vom Typ FP64 aus wie folgt aus:

```
singularity run --nv hpc-benchmarks\\:25.04.sif /workspace/stream-gpu-test.sh
```

<2000 GB/s	0 Punkte
2000 GB/s	1 Punkt
>4000 GB/s	5 Punkte

Führen Sie den STREAM Benchmark auf der GPU vom Typ FP32 aus wie folgt aus:

```
singularity run --nv hpc-benchmarks\\:25.04.sif /workspace/stream-gpu-test.sh --dt fp32
```

<500 GB/s	0 Punkte
500 GB/s	1 Punkt
>900 GB/s	5 Punkte

3.3.3 OpenFOAM (maximal 45 Punkte)

Dieser Benchmark dient zur Bestimmung der Performance von unterschiedlichen HPC-Systemen. In diesem wird die dreidimensionale, laminare, transiente Strömung in einer Kavität berechnet. Allerdings werden nur die ersten 100 Zeitschritte berechnet, was statistisch belastbare Rechenzeiten ergeben sollte. Als Software basiert der Case auf OpenFOAM v2412. In die Bewertung fließt die *Execution Time* ein, welche in der durch den Benchmark erzeugten `log.icoFoam-` Datei zu finden ist. Andere Werte werden nicht berücksichtigt.

Installation (Basis-Instruktionen)

Zur Installation sind folgende Abhängigkeiten zu installieren:

- gcc Version 11.4.0
- cmake
- boost Version 1.74
- fftw Version 3.3.10
- openmpi Version 4.1.6
- gmp Version 6.2.1

OpenFOAM liefert eine Liste an Abhängigkeiten

<https://develop.openfoam.com/Development/openfoam/blob/develop/doc/Requirements.md>

Der Benchmark ist für Version 2412 ausgelegt, diese kann hier geladen werden:

<https://dl.openfoam.com/source/v2412/OpenFOAM-v2412.tgz>

Laden Sie zusätzlich noch die Abhängigkeiten:

<https://dl.openfoam.com/source/v2412/ThirdParty-v2412.tgz>

Eine ausführliche Anleitung zur Installation finden Sie hier:

<https://develop.openfoam.com/Development/openfoam/-/blob/master/doc/Build.md>

Das Benchmark ist in folgenden Konfigurationen auf den Rechenknoten durchzuführen. Die Input Dateien und alle zusätzlich notwendigen Skripte können Sie von <https://cloud.tu-freiberg.de/s/4akXDrBbbs8x5dR> mit Passwort: „benchTUBAF“ herunterladen.

Führen Sie für jeden Test zunächst das „Allpre“ Skript im jeweiligen Testordner aus, was das jeweilige Netz und die Prozessor-Ordner erstellt. Dieser Vorgang kann 2-3 Minuten dauern.

Anschließend erfolgt der eigentliche Benchmark mit dem „Allrun“ Skript:

- 1 Sockel mit 32 Cores (1_32Cores Ordner)

>470 s	0 Punkte
470 s	1 Punkt
<300 s	15 Punkte

- 2 Sockel mit je 32 Cores (2_64Cores Ordner)

>200 s	0 Punkte
200 s	1 Punkt
<150 s	15 Punkte

- 4 Sockel (aufgeteilt auf 2 Knoten) mit je 32 Cores (128 insgesamt) (3_128Cores Ordner)

>100 s	0 Punkte
100 s	1 Punkt
< 75 s	15 Punkte

Sollte es aufgrund der Hardware-Ausstattung nicht möglich sein, die geforderte Anzahl an Sockeln zu verwenden, so dürfen hier Anpassungen vorgenommen werden. Die Gesamtzahl der Cores darf jedoch nicht verändert werden.

Das „Allclean“-Skript ermöglicht ein zurücksetzen des Case-Ordners.

4. Verpflichtender Besuch Vor-Ort

Für ein qualifiziertes Angebot wird es als notwendig angesehen, dass der Bieter sich vorab den vorgesehenen Raum und darin vorhandene Infrastruktur anschaut.

Der Bewerber hat dafür in dem Zeitraum vom 12.08.-05.09.2025 einmalig die Möglichkeit. Der Besuch ist in diesem Zeitraum werktags zwischen 9.00Uhr -12.00Uhr und 14:00Uhr-16:00Uhr möglich und vorab mit der TU Bergakademie Freiberg

abzustimmen. Die Anfragen zur Besichtigung sind bis zum 29.08.2025 einzureichen. Melden Sie Ihren Besuch unter folgende E-Mail-Adresse an:

ausschreibung@zuv.tu-freiberg.de

Die Besichtigung erfolgt dann gemeinsam mit einem Mitarbeiter der TU Bergakademie Freiberg. Der Bewerber erhält nach dem Besuch eine Bestätigung über den Vor-Ort-Termin. Diese Bestätigung ist den Angebotsunterlagen beizulegen. Sollte dieses Formular nicht beiliegen oder nicht bestätigt beiliegen, führt dies zum zwingenden Ausschluss vom Bewerbungsverfahren.

5. Garantie und Support

Um einen reibungslosen Betrieb über einen Zeitraum von fünf Jahren sicherzustellen, ist ein umfassendes Garantie- und Support-Konzept für diese Laufzeit vorzulegen. Dazu ist eine eigene, detaillierte Position zu erstellen, die die nachfolgenden Anforderungen erfüllt. Die Kosten dieser Position dürfen das Budget von **100.000€** (inkl. MWST) nicht überschreiten. Alle unten genannten Anforderungen sollen ohne zusätzliche Kosten erfüllt werden.

Die Schlüssigkeit des Konzeptes, sowie Umfang und Vertragsdauer fließen mit 175 Punkten in die Bewertung mit ein.

Die Bewertung wird wie folgt vorgenommen:

- ❖ Mangelhaft: 0 Punkte
 - Das erbrachte Angebot genügt nicht den Anforderungen.
- ❖ Ausreichend: 40 Punkte
 - Das erbrachte Angebot erfüllt im Mindestmaß die Anforderungen. Gesamtkonzept zeigt jedoch Unstimmigkeiten auf.
- ❖ Befriedigend: 85 Punkte
 - Das erbrachte Angebot erfüllt die Anforderungen.
- ❖ Gut: 130 Punkte
 - Das erbrachte Angebot erfüllt die Anforderungen im vollen Umfang als schlüssiges Gesamtkonzept.
- ❖ Sehr gut: 175 Punkte
 - Das erbrachte Angebot erfüllt die Anforderungen im vollen Umfang und darüber hinaus.

Anforderungen

Die Garantie umfasst alle Komponenten des Clusters.
Die Koordination von Garantieansprüchen wird grundsätzlich vom Anbieter übernommen, sofern keine Garantie/Supportanfragen direkt beim Hersteller möglich sind.

Für die gesamte Software sind Support und Updates zur Verfügung zu stellen. Falls ein Softwarehersteller innerhalb des Betriebszeitraums von fünf Jahren ein Produkt abkündigt („End of Life“), ist sicher zu stellen, dass ein Upgrade auf eine unterstützte Version verfügbar ist.

Bietet ein Hersteller keine 5-jährigen Supportzeiträume an, so ist in Ausnahmefällen ein kürzerer Zeitraum möglich. Eine schriftliche Bestätigung des Softwareherstellers ist für diesen Ausnahmefall beizulegen.

Im Störungsfall erwarten wir „Next Business Day Support“ für alle Komponenten, dies beinhaltet Hard- und Softwarekomponenten.

Für Anfragen zu administrativen Tätigkeiten bezüglich Hard- und Software ist während der Betriebszeiten (9:00 bis 17:00 Montag bis Freitag) Support anzubieten. Die Reaktionszeit muss unter 72 Stunden liegen.

Für den Fall einer Störung der kritischen Komponenten soll ein direkter Kontakt (Durchwahl) zu einem Mitarbeiter/Mitarbeiterin des Bieters möglich sein. Die Kontaktperson muss Deutsch oder Englisch (Level C1) sprechen. Ein Kontakt über die Standard-Servicehotline ist nicht vorzuschlagen.

Bei kritischen Komponenten sollte die Antwortzeit unter 4 Stunden während der regulären Arbeitszeit (9:00 bis 17:00 Montag bis Freitag) liegen.

Bei Bedarf sind Fehler in Komponenten durch Service Techniker vor Ort zu beheben, andernfalls im Rahmen eines Austauschservice („Collect and Return“).

Teil des Garantie- und Wartungskonzept soll der Support des Parallelen Filesystem sein. Hier wird erwartet, dass der Bieter den Support initial direkt mit dem Hersteller vermittelt. Dies beinhaltet die Einrichtung eines Zugangs zu einem Supportportal des Herstellers. Servicefälle im Laufe des Betriebszeitraumes, können dann direkt zwischen Käufer und Hersteller erfolgen. Der Support soll alle Software-Upgrades für diese Komponente enthalten.

Im Falle eines nicht-kommerziellen Batch-Systems ist Support zu leisten (initiale Einrichtung, Störfälle, Sonderanforderungen von Nutzern, Konfigurationsänderungen). Störfälle fallen unter die Kategorie „kritische Komponenten“ für welchen wir die oben genannte Reaktionszeit erwarten. Alles weitere fällt unter die 72-stündige Reaktionszeit.

Im Haverie-Fall (Ausfall 25% der Cluster Komponenten und/oder Potentieller Datenverlust) wird ein Austausch der Hardware an Werktagen innerhalb von 48 Stunden gefordert.

6. Schulungen und Dokumentation

Für das Personal an der TU Bergakademie Freiberg ist ein umfassendes Schulungs- und Dokumentationskonzept vorzulegen.

Anforderungen

Es muss eine mind. 1- tägige Vor-Ort-Schulung für die Administratoren angeboten werden. Es müssen alle wichtigen Schnittstellen für die Administratoren vorgestellt werden. Die Schulung ist in deutscher Sprache durchzuführen. Bei Bedarf ist der Schulungszeitraum zu erweitern.

Ein „Hands-On“ Training für jene Komponenten muss angeboten werden, die auch vom Auftraggeber selbst gewartet werden könnten. Das Hands-on-Training ist in deutscher Sprache durchzuführen. Dieses Training kann Teil der oben genannten Schulung sein.

Eine ausführliche Dokumentation in druckbarer und elektronisch durchsuchbarer Form (z.B. PDF) muss mitgeliefert werden. Diese sollte durch Tabellendokumente oder Ähnliches ergänzt werden. Diese Dokumente sind in deutscher Sprache abzugeben.

7. Bewertungsmatrix

Die Bewertung der Angebote erfolgt nachfolgenden Bewertungskriterien: Allen Angeboten werden Bewertungspunkte für ausgewählte Leistungskriterien zugeordnet. Die unten aufgeführte Tabelle fasst die maximal möglichen Bewertungspunkte aus dem Leistungsverzeichnis zusammen:

Kriterium	Kurzbeschreibung	Bewertungspunkte
2.2.1	Summe aller CPU Cores der Rechenknoten	300
2.2.2	Basistaktfrequenz des Prozessors (in GHz)	25

2.2.3*	RAM in GB (mind. 4GB pro Core und nicht weniger als 180GB pro Knoten)	50
2.2.4	Transferrate des Arbeitsspeichers	25
2.6.7-1	Anzahl GPU (Typ 1 (FP64), mind. 4)	30
2.6.7-2	Anzahl GPU (Typ 2 (FP32))	70
2.7.1*	Brutto-Kapazität Gesamtspeicher des SCRATCH (mind. 1PB)	120
2.7.2*	Brutto-Kapazität des SSD-Speichers des SCRATCH (mind. 150 TB)	25
2.7.12	Möglichkeit zur Erstellung reiner SSD Volumes	5
2.7.17	Alternative Storage Protokolle	10
2.7.18	Erstellung von Snapshots	10
2.9.7/13/18	Modular tauschbare Lüfter und Netzteile	15
2.11.1	Wirkungsgrad der Netzteile (80 PLUS Rating)	10
2.11.2	Kühlbedarf in (kW)	40
Bench1	HPCG	30
Bench2	STREAM	15
Bench3	OpenFOAM	45
Support	Support	175
Gesamt	Summe aller Punkte	1000

8. Liefermodalitäten und Abnahmebedingungen

Es sind ausschließlich fabrikneue, originale Produkte des Herstellers zu liefern, die in der EU zugelassen sind. Die Lieferbarkeit der angebotenen Geräte muss ab Vertragsbeginn in vollem Umfang gesichert sein.

Die Lieferung und Inbetriebnahme sollte bis 6 Monate nach Zuschlagserteilung erfolgen. Geben Sie im Angebot eine Terminkette für die Aufbau- und Installationsarbeiten mit an.

Die Lieferung erfolgt „frei Verwendungsstelle“, umfasst weiterhin das Abladen, den Einbau im Serverraum und beinhaltet alle Kosten für Verpackung, Transport, Zölle usw.). Die Anlage muss „schlüsselfertig“ vor Ort aufgebaut werden. Vor Abnahme werden die Benchmarks durchgeführt und überprüft ob die zugesicherte Leistung erbracht wird. Ebenfalls sind Test zum Erreichen der Maximallast durchzuführen um auch hier die zugesicherte Leistung und maximale Kühlbelastung zu bestätigen. Sollte die zugesicherte Leistung bzw. Kühllast nicht erreicht werden erfolgt solange keine Abnahme bis zur Nachverbesserung der Leistung.

Ort der Lieferung:

TU Bergakademie Freiberg
Universitätsrechenzentrum - IT-Sicherheitszelle
Bernhard-von-Cotta-Straße 1

09599 Freiberg

9. Zahlungsbedingungen

Das Angebot (Cluster, Infrastrukturkomponenten, Garantie-Support) darf in Summe einen Gesamtauftragswert von 1.700.000,00 € brutto (inkl. MwSt.) nicht überschreiten.

Die eingesetzten **Preise sind Festpreise** und beziehen sich auf die angebotenen Ausführungen der jeweiligen Positionen einschließlich Lieferung sowie Verpackung frei Leistungs- und Erfüllungsort, Abladen sowie etwaiger Zölle und der Installation.

Vorauszahlungen (Zahlungen vor Endabnahme vor Ort in Freiberg, also auch „bei Lieferung“ oder Werksabnahme) sind nur möglich, insoweit diese branchenüblich sind (§ 56 Bundeshaushaltsordnung/Sächsische Haushaltsordnung) und der Auftragnehmer eine **gültige Bürgschaft** eines in der Europäischen Union zugelassenen und vom Auftraggeber akzeptierten Kreditinstitutes vorlegt. Die **Bürgschaftserklärung muss unbefristet**, schriftlich und unter Verzicht auf die

Einrede der Vorausklage nach § 771 BGB, der Anfechtbarkeit nach § 770 Abs. 1 BGB sowie der Aufrechenbarkeit nach § 771 Abs. 2 BGB abgegeben werden.

Die Bürgschaftsurkunde wird nach erfolgreicher Endabnahme und damit nach Erfüllung aller Leistungen an das Kreditinstitut zurückgegeben.

Vorauszahlungen (Zahlungen vor Endabnahme vor Ort in Freiberg, also auch „bei Lieferung“ oder Werksabnahme) werden zudem nur bis zu einer Höhe von insgesamt 30% des Auftragswertes geleistet. Die übrigen 70% des Auftragswertes werden erst nach erfolgreicher Abnahme gezahlt. Für die Zahlung gilt § 3 Abs. 3 Satz 1 und 2 der AVB der TU Bergakademie Freiberg entsprechend, mit der Maßgabe, dass die Vorauszahlung nach Eingang der Auftragsbestätigung und Rechnungslegung erfolgt.

Der Zahlungsplan ist mit der Angebotsabgabe bindend und kann nach der Zuschlagserteilung nicht verändert werden. Achten Sie auf die maximal möglichen Prozentsätze. Sollte die Anzahlung nicht in Anspruch genommen werden, tragen Sie bitte bei Anzahlung eine „0“ ein und bei Schlusszahlung eine „100“.

- Prozentsatz (max. 30% des Gesamtauftragswertes)
für die Anzahlung gegen eine unbefristete Bankbürgschaft: % (bitte eintragen)
- Prozentsatz für die Schlusszahlung
(nach Lieferung und Abnahme): % (bitte eintragen)

Ein abweichender Zahlungsplan ist nicht zulässig.

Zahlungsbedingungen:

30 Tage Netto

oder

..... % Skonto bei Zahlung innerhalb von Tagen (mind. 14 Tage gemäß Anlage 4 AVB),
30 Tage Netto

10 Nebenangebote

Nebenangebote sind nicht zugelassen.

11. Bewertung des Angebotes

Die Bewertung der Angebote erfolgt mit der folgenden Gewichtung der Bewertungskriterien:

Kriterium 1:	Preis	Gewichtung 30 %
Kriterium 2:	Leistung (siehe Leistungsverzeichnis/7. Bewertungsmatrix)	Gewichtung 70%

Im Kriterium Leistung werden ausschließlich die Angaben des Bieters zu den Bewertungskriterien in der Leistungsbeschreibung bewertet.

Die Wertung des Preises bemisst sich am Bewertungspreis, der sich aus dem Angebotsnettopreis und anfallenden Steuern (Einfuhr-/Umsatzsteuer), die von Auftraggeber zu tragen sind und zwar ohne Rücksicht auf die Steuerschuldnerschaft, zusammensetzt.

Der Zuschlag wird auf das wirtschaftlichste Angebot erteilt. Der punktstärkste Bieter erhält den Zuschlag.

Zur Ermittlung des wirtschaftlichsten Angebots wird die Preisquotientenmethode angewandt:

$$\text{Gesamtpunktzahl} = \text{Gewichtung Leistung} * \frac{\text{erreichte Leistungspunktzahl}}{\text{Maximal mögliche Leistungspunktzahl}} + \text{Gewichtung Preis} * \frac{\text{Niedrigster Angebotspreis}}{\text{Angebotspreis}}$$