

Leistungsbeschreibung

GPU-Servercluster mit RTX PRO 6000 SE

Projekt: GPU-Servercluster |

1. Ausgangssituation und Zielsetzung

Das Universitätsklinikum Essen (UK Essen) erweitert seine RCC Rechen-, Speicher- und KI-Infrastruktur für rechenintensive Forschungsanwendungen.

Gegenstand der Beschaffung ist der bezeichnete GPU ServerCluster mit RTX PRO 6000 SE einschließlich Fileserver, GPU-Nodes, Netzwerkkomponenten, NVIDIA DGX Spark Appliances und des ausdrücklich aufgeführten Zubehörs.

Die Lieferung muss den in dieser Leistungsbeschreibung genannten Fabrikaten, Modellbezeichnungen, Stückzahlen und Einzelkonfigurationen **exakt entsprechen**. Sämtliche Geräte sind fabrikneu und in der beschriebenen Konfiguration zu liefern.

1.1 Projektziele

- Bereitstellung der exakt angebotenen GPU-, CPU-, Speicher- und Netzwerkkapazität
- Nahtlose Integration in die bestehende Cluster- und Rechenzentrumsumgebung
- Sicherstellung des Betriebs durch Gewährleistungs- und Supportleistungen
- Vollständige Lieferung aller Systeme, Switches, Rack-Kits und Verbindungskabel

1.2 Beschaffungsumfang

Menge	Position	Exakte Bezeichnung
1	Fileserver	2x16coreCPU-512GB-24x24TB-4x1,9TB-8x15,3TB-2x1,9TB-2x10G-2x100G
2	GPU Nodes	2x128coreCPUs-1536GB-2x960GB-2x15,3TB-2x10G-2x100G-6xRTXPRO6000
1	TOR Switch	SuperMicro SSE-G3748R-SONIC
1	Ethernet Switch	Edge-corE AS5835-54T (DCS 202) SONIC
1	Ethernet Switch	Edge-corE AS7726-32X (DCS204) SONIC
24	DGX Spark-EDU	NVIDIA DGX Spark Founders Edition AI Appliance
12	Rack Kit für 2 Sparc Rechner	UM-NVI-202 - NVIDIA DGX Spark Rack Mount Kit, 2x Slot, 19 inch, 1.33U, Rail System 300mm

Menge	Position	Exakte Bezeichnung
12	Verbindungskabel	200G Ethernet Cable für Spark (Zur Verbindung von jeweils 2)

2. Leistungsumfang

Der Auftragnehmer erbringt folgende Leistungen:

2.1 Lieferung

- Lieferung aller in Abschnitt 1.2 aufgeführten Positionen in den dort genannten Stückzahlen
- Lieferung der Server vollständig bestückt und konfiguriert mit sämtlichen in Abschnitt 3 genannten Hardwarekomponenten
- Lieferung der Switches einschließlich der jeweils ausgewiesenen SONiC-Software- und Supportbestandteile
- Lieferung der NVIDIA DGX Spark Appliances einschließlich der zugehörigen Rack-Kits und 200G-Ethernet-Kabel
- Transport **frei Haus** gemäß Bezugsangebot

2.2 Werksprüfung

- 48h Burn-In/Stresstest für den Fileserver
- 48h Burn-In/Stresstest für jeden der beiden GPU-Nodes
- Dokumentation der gelieferten Seriennummern und Konfigurationsparameter

2.3 Übergabe und Nachweise

- Vollständige Übergabe aller Geräte und Zubehörpositionen
- Nachweis der exakten Modell- und Komponentenbestückung je Server
- Übergabe der Gewährleistungs-, Support- und Subskriptionsnachweise
- Übergabe der für Betrieb und Verwaltung erforderlichen Herstellerdokumentation

2.4 Gewährleistung, Support und Lieferzeit

- Server: 3 Jahre Komponenten Schnellaustausch-Service; 3 Jahre Schnellaustausch von Teilen mit Vorabersatz
- NVIDIA DGX Spark: 2 Jahre Gewährleistung
- Edge-corE Switches: 3 Jahre Vorabaustausch gemäß Angebot (kann ausgewählt werden)
- Enterprise SONiC: jeweils 1 Jahr Support beziehungsweise 1-Year Subscription gemäß Position
- Fileserver und GPU-Nodes: ca. 8 bis 10 Wochen ab Bestellung
- Alle Komponenten: ca. 6 bis 10 Wochen oder früher

3. Exakte technische Anforderungen

Die nachfolgenden Spezifikationen geben die zu liefernde Konfiguration **abschließend** wieder. Es handelt sich **nicht um beispielhafte oder lediglich vergleichbare Mindestanforderungen**.

3.1 Fileserver - Menge 1

Modell: 2x16 CPU, **512GB-24x24TB-4x1,9TB-8x15,3TB-2x1,9TB-2x10G-2x100G** - Intel Dual Xeon Hochleistungs-Server

Komponente	Exakte Spezifikation
Chassis	4HE Rack-Chassis mit redundanten 1280W Netzteilen (80Plus Titanium Level / 96%)
Laufwerkseinschübe	12x 2.5" HotSwap Einschub für NVMe 24x 3.5" HotSwap Einschub für SATA/SAS 2x M.2 NGFF Connector (PCI-E 5.0 x4 Slot: 22110) (RAID0,1)
Motherboard	System on Chip 16 DIMM Slots (max. 2048GB DDR5) 2x 10GBase-T 10GbE Ports (Broadcom®BCM57416 Dual-Port 10Gigabit Ethernet Controller) 10x SATA3 (6Gb/s) Ports 4x PCI-E 5.0 x16 Slots (HHFL) 2x PCI-E 5.0 x8 Slots (HHFL) 3x PCI-E 5.0 NVMe x8 Internal Port Integrated IPMI 2.0 with Dedicated LAN Integrated Aspeed AST2600 BMC 4x USB 3.2
CPU	2x Intel Xeon 6515P 16-Core CPU, 150W, 2.30GHz, 72.00MB L3 Cache, DDR5-6400, Turbo Core max. 3.80GHz, -
RAM	512GB (16x 32GB) DDR5-6400 DIMM, REG, ECC, 2R
Controller	1x Broadcom 9580-8i8e MegaRAID SAS3/NVMe Controller, inkl. CacheVault Technology
Storage	24x HDD 24TB SAS3, 24x7, 7,2k RPM, Seagate Exos X24 4x SSD 1.9TB U.2 NVMe, 24x7, 1 DWPD, Micron 7600 PRO 8x SSD 15.3TB U.2 NVMe, 24x7, 1 DWPD, HUAWEI ES3500P 2x SSD 1.9TB U.3 NVMe, 24x7, 1 DWPD, Micron 7500 PRO
Network	1x Broadcom P2100G 100GbE HCA, 2-port, QSFP56 (Broadcom® 57508 Dual-Port 100 Gigabit Ethernet Controller)
Sonstiges	Einbaukit für 19" Rack 48h Burn-In/Stresstest
Gewährleistung	3 Jahre Komponenten Schnellaustausch-Service

Komponente	Exakte Spezifikation
Lieferzeit	ca. 8 bis 10 Wochen ab Bestellung

3.2 GPU Nodes - Menge 2

Modell je System: **2x128CPU-1536GB-2x960GB-2x15,3TB-2x10G-2x100G-6xRTXPRO6000** - AMD Dual EPYC Hochleistungs-Server

Komponente	Exakte Spezifikation
Chassis	5HE Rack-Chassis mit redundanten 6x 2700W Netzteilen (80Plus Titanium / 96%)
Laufwerkseinschübe	4x 2.5" HotSwap Einschub für NVMe 2x 2.5" HotSwap Einschub für SATA 1x M.2 NGFF Connector (PCI-E 3.0 x4 Slot: 22110)
Motherboard	System on Chip 24 DIMM Slots (max. 3072GB DDR5) 2x 10GBase-T 10GbE Ports (Intel® X550 Dual-Port 10Gigabit Ethernet Controller) 2x SATA3 (6Gb/s) Ports 9x PCI-E 5.0 x16 Slots (FHFL) Integrated IPMI 2.0 with Dedicated LAN Integrated Aspeed AST2600 BMC 4x USB 3.2
CPU	2x AMD EPYC 9755 128-Core CPU, 500W, 2.70GHz, 512MB L3 Cache, DDR5-6400, Turbo Core max. 4.10GHz, -
RAM	1536GB (24x 64GB) DDR5-6400 DIMM, REG, ECC, 2R
Storage	2x SSD 960GB SATA3, 24x7, 1.5 DWPD, Kingston DC600M 2x SSD 15.3TB U.2 NVMe, 24x7, 1 DWPD, HUAWEI ES3500P
Network	1x Broadcom P2100G 100GbE HCA, 2-port, QSFP56 (Broadcom® 57508 Dual-Port 100 Gigabit Ethernet Controller)
GPU	6x NVIDIA RTX PRO 6000 Blackwell Server Edition EDU/Inception/Connect, 96GB
Sonstiges	Einbaukit für 19" Rack 48h Burn-In/Stresstest
Gewährleistung	3 Jahre Komponenten Schnellaustausch-Service
Lieferzeit	ca. 8 bis 10 Wochen ab Bestellung

3.3 TOR Switch - Menge 1

Komponente	Exakte Spezifikation
Modell	SuperMicro SSE-G3748R-SONIC
Bauart	TOR Switch, reverse airflow (P2C)
Ports	SuperMicro SSE-G3748R-SONIC 1GbE Switch + 6x 25GbE SFP28

3.4 Ethernet Switch - Menge 1

Komponente	Exakte Spezifikation
Modell	Edge-corE AS5835-54T (DCS 202) SONIC
Hardware	Edge-corE Ethernet Switch AS5835-54T; Intel® Atom® C3558 processor 4-Core; 48-Port 10G RJ-45 with 6x100G QSFP28 uplinks
Airflow	Ethernet Switch P2C airflow
Software/Support	Including Enterprise SONIC mit 1 Jahr Support Broadcom Enterprise SONiC Linux Distribution, Enterprise Base Package, 10G Platform, SmartCare 9x5 CE(S)T Support, 1-Year Subscription
Austausch	3 Jahre Vorabaustausch (kann abgewählt werden)

3.5 Ethernet Switch - Menge 1

Komponente	Exakte Spezifikation
Modell	Edge-corE AS7726-32X (DCS204) SONIC
Switch Ports	32 x QSFP28 40/100G GbE
Logical Ports	Max. 128
Port Modes	1 x 100G (4 lanes 25G NRZ) QSFP28 1 x 40G (4 lanes 10G NRZ) QSFP+ 2 x 50G (2 lanes 25G NRZ) QSFP28 breakout 4 x 25G (1 lane 25G NRZ) QSFP28 breakout 4 x 10G (1 lane 10G NRZ) QSFP+ breakout
Software/Support	Including Enterprise SONIC mit 1 Jahr Support Broadcom Enterprise SONiC Linux Distribution, Enterprise Base Package, 25G/50G/100G Platform, SmartCare 9x5 CE (S)T Support, 1-Year Subscription
Austausch	3 Jahre Vorabaustausch (kann abgewählt werden)

3.6 DGX Spark-EDU - Menge 24

NVIDIA DGX Spark Founders Edition AI Appliance - NVIDIA Grace Blackwell Architecture
(Education Preis)

Komponente	Exakte Spezifikation
GPU	NVIDIA GB10 Grace Blackwell
CPU	20 core Arm, 10 Cortex-X925 + 10 Cortex-A725 Arm CPU
Tensor Performance*	1 PFLOP
System Memory	128 GB LPDDR5x, unified system memory
Memory Interface	256-bit
Memory Bandwidth	273 GB/s
Storage	4 TB NVME.M2 with self-encryption
USB	4x USB TypeC
Ethernet	1x RJ-45 connector, 10 GbE
NIC	ConnectX-7 Smart NIC
Wi-Fi	WiFi 7
Bluetooth	BT 5.3
Audio-output	HDMI multichannel audio output
Display Connectors	1x HDMI 2.1a
NVENC NVDEC	1x 1x
OS	NVIDIA DGX™ OS
System Dimensions	150 mm L x 150 mm W x 50.5 mm H
System Weight	1.2 kg
Gewährleistung	2 Jahre

3.7 Zubehör

Komponente	Exakte Spezifikation
12x Rack Kit für 2 Sparc Rechner	UM-NVI-202 NVIDIA DGX Spark Rack Mount Kit 2x Slot 19 inch 1.33U Rail System 300mm Rack Kit to mount two NVIDIA DGX Spark units with their associated power supplies on 1.33U of space in a 19 inch rack
12x Kabel	200G Ethernet Cable für Spark (Zur Verbindung von jeweils 2)

4. Standort und Mitwirkung des Auftraggebers

Der Auftraggeber stellt für Anlieferung, Einbau und Betrieb bereit:

Komponente	Exakte Spezifikation
Rack-/Stellfläche	Ausreichender Platz für die in Abschnitt 3 genannten Bauformen sowie die 12 Rack-Kits
Stromversorgung	Anschlüsse und Leistungsreserven entsprechend den gelieferten Systemen und deren Netzteilen
Kühlung	Ausreichende Kühlleistung für den Betrieb der gelieferten Systeme
Netzwerk	Anschlusspunkte für Management-, 10GbE-, 25GbE-, 100GbE- und 200G-Verbindungen gemäß Lieferumfang
Zutritt	Zugang zum Rechenzentrum und geeignete Anlieferungsroute
Ansprechpartner	IT-Betrieb vor Ort bei Übergabe und Abnahme

5. Abnahmekriterien

Die Abnahme erfolgt nach erfolgreicher Erbringung aller Leistungen gemäß folgender Checkliste:

- ☐ 1 Fileserver des exakt bezeichneten Modells vollständig geliefert
- ☐ 2 GPU-Nodes des exakt bezeichneten Modells vollständig geliefert
- ☐ 1 SuperMicro SSE-G3748R-SONIC vollständig geliefert
- ☐ 1 Edge-corE AS5835-54T (DCS 202) SONIC vollständig geliefert
- ☐ 1 Edge-corE AS7726-32X (DCS204) SONIC vollständig geliefert
- ☐ 24 NVIDIA DGX Spark Founders Edition AI Appliances vollständig geliefert
- ☐ 12 Rack-Kits UM-NVI-202 vollständig geliefert
- ☐ 12 200G-Ethernet-Kabel für Spark vollständig geliefert
- ☐ Fileserver und GPU-Nodes entsprechen vollständig der Komponentenbestückung in Abschnitt 3
- ☐ Switches entsprechen vollständig den Port-, Software- und Supportangaben in Abschnitt 3
- ☐ DGX Spark Appliances entsprechen vollständig den technischen Angaben in Abschnitt 3
- ☐ 48h Burn-In/Stresstest für Fileserver und beide GPU-Nodes dokumentiert
- ☐ Alle Seriennummern und Konfigurationsparameter dokumentiert
- ☐ Gewährleistungs-, Support- und Subskriptionsnachweise vollständig übergeben
- ☐ Lieferung frei Haus erfolgt
- ☐ Übergabeprotokoll unterzeichnet