

Część 3 – dostawa serwera do obsługi modeli językowych

Dostawa serwera z co najmniej dwoma kartami GPU o poniższej konfiguracji.

Dostarczone elementy muszą być fabrycznie nowe i objęte obsługą serwisową.

Pozostałe opcje dodatnie dotyczące podzespołów wymieniono na końcu niniejszego dokumentu.

Specyfikacja serwera:

1. Płyta główna

Kategoria	Wymagania
Socket (gniazdo CPU)	<ul style="list-style-type: none">- LGA 4677 lub SP3/TR4- Obsługa wieloprocesorowych układów CPU (dual-socket lub więcej)
Obsługa GPU	<ul style="list-style-type: none">- Minimum 4 złącza PCIe x16 (PCIe Gen 4/5)- Możliwość montażu kart dwuslotowych (np. NVIDIA A100, H100)- Obsługa technologii NVLink między kartami GPU
Obsługa pamięci RAM	<ul style="list-style-type: none">- Minimum 16 slotów DIMM- Typ: DDR5 lub DDR4- Wsparcie dla RDIMM/LRDIMM (ECC)- Maksymalnie: min. 1 TB (rozbudowa do 4 TB)
Złącza dysków i kontrolery	<ul style="list-style-type: none">- Obsługa NVMe M.2 lub U.2 (min. 2-4 złącza)- Obsługa SAS/SATA (hot-swap)- Wsparcie dla RAID (np. RAID 0, 1, 5, 10)
Interfejsy sieciowe (NIC)	<ul style="list-style-type: none">- Wbudowany kontroler 10GbE (lub więcej)- Możliwość rozbudowy o 25GbE lub 100GbE porty SFP+/QSFP- Opcjonalne wsparcie dla RDMA
Zasilanie i chłodzenie	<ul style="list-style-type: none">- Kompatybilność z redundantnymi zasilaczami (min. 1200W, max. 3000W) od 2 do 4 zasilaczy, moc dostosowana do obsługi min. 4 kart GPU- Wsparcie dla zaawansowanego chłodzenia (cieczą lub powietrzem)- Złącza do wentylatorów
Zarządzanie zdalne	<p>IPMI 2.0 (KVM-over-IP) lub iDRAC (Dell) lub ILO (HPE)</p> <p>Umożliwia pełne zdalne zarządzanie serwerem (w tym dostęp do konsoli, BIOS-u, systemów operacyjnych), bez względu na stan systemu operacyjnego.</p>
Rozbudowa i kompatybilność	<ul style="list-style-type: none">- Modułowe zasilanie PCIe- Złącza OCP (do wymiany kart sieciowych)

Certyfikaty i wsparcie	- Wsparcie dla NVIDIA NGC, TensorFlow, PyTorch - Certyfikaty energetyczne (np. 80+ Titanium)
------------------------	---

2. Procesory (CPU):

- Procesor 64 bitowy w architekturze x86
- Liczba procesorów 2
- Taktowanie min. 2.10 GHz
- Taktowanie Turbo min. 4.00 GHz
- liczba rdzeni / wątków min. 8 / 16
- Pamięć Cache min. 25 MB
- Szyna Pamięci min. (MHz) 2666 MHz
- QPI min. 9.60 GT/s min. 2UPI

3. Pamięć RAM:

- Rozmiar min 64GB
- Szyna min. 3200 MT/s
- Typ DDR4 lub DDR5
- Pojemność modułu min. 64 GB
- Korekcja błędów Minimum ECC

4. Karty graficzne (GPU):

- Minimum 2x **GPU** z opcją skalowalności do minimum 4 kart GPU

5. Specyfikacja kart graficznych (GPU):

PARAMETR	WARTOŚĆ
Pamięć:	min. 16GB
Architektura:	Ampere lub nowsza
CUDA Cores	min. 8000
Moc obliczeniowa (FP32)	min. 20 TFLOPS.
Chłodzenie	aktywne albo pasywne
Interfejs	PCIe 4.0
Tensor cores	min 200
Inne	obsługa NVlink

6. Pamięć masowa:

- **Dysk SSD NVMe** o pojemności minimum 960 GB do szybkiego przetwarzania danych.

- Dodatkowe magazynowanie: minimum 2x **Dysk SSD SAS 2.5"** o pojemności 960 GB, Mixed Use, Hot-Plug, 12 Gbps.

7. Interfejsy sieciowe:

- **2x porty 10GbE** (minimum)
- Dodatkowy moduł sieciowy: **Dual Port 1GbE**.

8. Obudowa i gniazda dysków:

- Obudowa **2U** z możliwością montażu w szafie serwerowej RACK 19cali.
- Minimum 8 zatok 2.5" Hot-Swap SAS/SATA/NVMe.

9. Chłodzenie:

- Zaawansowane, redundantne chłodzenie z obsługą wieloprocessorowych układów GPU.

10. System operacyjny:

- **Bez systemu operacyjnego**, wymagalne wsparcie dla instalacji **Linux** (np. Ubuntu, CentOS) lub **Windows Server 2022**.

Dodatkowe wymagania:

- **Gwarancja i serwis:** Warunki zostały opisane w dokumencie SZCZEGÓŁOWE WARUNKI GWARANCYJNE
- **Wsparcie producenta:** Dostęp do aktualizacji oprogramowania sprzętowego i BIOS przez cały okres gwarancji.

Szkolenie

Wykonawca przeprowadza, krótkie (od 1 do 2 godzin) szkolenie, zapoznające Zamawiającego z dostarczonym sprzętem – serwerami. Szkolenie polega na omówieniu praktycznych i najważniejszych cech użytkowych dostarczonego sprzętu i oprogramowania oraz udzielenie odpowiedzi na pytania Zamawiającego. Szkolenie w języku polskim w dowolnej formie uzgodnionej z Zamawiającym np. on-line albo na miejscu instalacji albo telefonicznie.

Dostawa:

Wykonawca dostarcza w pełni skonfigurowany serwer obliczeniowy, przeznaczony do zadań związanych z AI, z zamontowanymi w nim dyskami SSD i HDD na wskazane przez Zamawiającego miejsce. Dostawa może odbyć się osobiście przez Wykonawcę lub za pośrednictwem firmy kurierskiej.

- **Dostawa osobista:** W przypadku osobistej dostawy, Wykonawca przeprowadza test poprawności działania serwera oraz zamontowanych dysków. Po potwierdzeniu prawidłowego funkcjonowania urządzenia następuje podpisanie protokołu odbioru.
- **Dostawa kurierska:** W przypadku dostawy kurierskiej, Zamawiający wykonuje test poprawności działania serwera wraz z dyskami. Dopiero po pomyślnym przejściu testów następuje podpisanie protokołu odbioru.

Wszelkie usterki lub nieprawidłowości, które pojawią się podczas pierwszego uruchomienia serwera, zostaną odnotowane w protokole odbioru.

Akcesoria i montaż:

Wraz z serwerem, Wykonawca dostarcza wszystkie niezbędne akcesoria potrzebne do instalacji i uruchomienia urządzenia w standardowej szafie RACK. Serwer będzie podłączony do zasilacza UPS, który znajduje się w szafie rackowej Zamawiającego.

OPCJE DODATNIE DLA KOMPONENTÓW

Oferowane części muszą pasować do zaoferowanego sprzętu w zamówieniu podstawowym.

Specyfikacja pojedynczej karty GPU:

PARAMETR	WARTOŚĆ
Pamięć:	min. 16GB
Architektura:	Ampere lub nowsza
CUDA Cores	min. 8000
Moc obliczeniowa (FP32)	min. 20 TFLOPS.
Chłodzenie	aktywne albo pasywne
Interfejs	PCIe 4.0
Tensor cores	min 200
Inne	obsługa NVlink