

OPIIS PRZEDMIOTU ZAMÓWIENIA

na

Dostawę platformy sprzętowo-programowej dla projektów związanych ze sztuczną inteligencją (Platforma AI).

1. WYMAGANIA OGÓLNE

- 1.1. Sprzęt dostarczony przez Wykonawcę musi być fabrycznie nowy, pochodzić z bieżącej produkcji, nieużywany i niezarejestrowany na innego klienta w bazie klientów producenta sprzętu oraz musi być objęty opieką serwisową producenta.
- 1.2. Zaoferowany sprzęt nie może być starszy niż 6 miesiące od daty produkcji. Wykonawca dostarczy Zamawiającemu najpóźniej w dniu dostawy oferowanych urządzeń oświadczenie producenta lub jego polskiego przedstawicielstwa potwierdzające datę produkcji urządzeń.
- 1.3. Sprzęt dostarczony przez Wykonawcę będzie pochodził z autoryzowanego kanału sprzedaży producenta na rynek Rzeczypospolitej Polskiej. Spełnienie powyższego wymogu zostanie potwierdzone oświadczeniem producenta sprzętu lub jego polskiego przedstawicielstwa, które Wykonawca zobowiązuje się dostarczyć Zamawiającemu najpóźniej w dniu dostawy oferowanych urządzeń.
- 1.4. Korzystanie przez Zamawiającego z dostarczonych urządzeń nie może stanowić naruszenia majątkowych praw autorskich osób trzecich.
- 1.5. Prace prowadzone przez Wykonawcę nie mogą skutkować utratą danych ani uszkodzeniem spójności danych w systemach Zamawiającego.
- 1.6. Wykonawca dostarczy kompletne okablowanie przyłączeniowe do sieci LAN/SAN dla wszystkich dostarczonych urządzeń z interfejsem LAN/SAN oraz kable połączeniowe.
- 1.7. Wymagane jest, aby Wykonawca zapewnił wszystkie niezbędne elementy, konieczne do montażu, uruchamiania urządzeń oraz poprawnego funkcjonowania w szafach rackowych Zamawiającego (tj. śrubki, nakrętki, kable zasilające, konieczne patchcordsy, kable krosowe).
- 1.8. Producent lub wykonawca wykona montaż urządzeń w miejscu wskazanym przez Zamawiającego oraz dokona konfiguracji w uzgodnieniu z Zamawiającym. Instalacji dokona inżynier posiadający certyfikat producenta uprawniający do wykonywanych prac. Na potwierdzenie wykonania konfiguracji zostanie wykonana dokumentacja powykonawcza.
- 1.9. Wykonawca dostarczy dokumentację użytkową oraz administracyjną do oferowanej Platformy AI w języku polskim lub angielskim, w formie papierowej lub elektronicznej (pliki PDF lub EPUB).

- 1.10. Dostawa jest realizowana na koszt i ryzyko Wykonawcy.
- 1.11. Niniejszy dokument zawiera wymagania minimalne, chyba że zapis dla konkretnego parametru mówi inaczej. Każde zaproponowane rozwiązanie, w ramach złożonej oferty zapewniające wyższe standardy techniczne i ekonomiczne, będzie traktowane jako spełniające warunki.
- 1.12. W ramach niniejszego Przedmiotu Zamówienia, wszelkie dostawy i prace instalacyjne muszą zostać wykonane w lokalizacji Zamawiającego - 41-200 Sosnowiec, ul. 3 maja 16a
- 1.13. Termin realizacji przedmiotu Umowy 3 miesiące od podpisania Umowy.

2. WYMAGANIA FUNKCJONALNE ROZWIĄZANIA AI

- 2.1. Wdrażane rozwiązanie (Platforma AI) będzie stanowiło kompletną platformę do wdrażania i działania modeli opartych o sztuczną inteligencję oraz orkiestracji kontenerów.
- 2.2. Rozwiązanie musi być dostarczone w modelu chmury prywatnej stanowiącej platformę dla aplikacji i usług związanych z wytwarzaniem i wykorzystaniem technologii AI / ML. Wykonawca dostarczy gotową, zintegrowaną platformę chmury prywatnej AI, zaprojektowaną specjalnie z myślą o obciążeniach związanych ze sztuczną inteligencją.
- 2.3. Rozwiązanie musi być platformą sztucznej inteligencji przetwarzającą dane w modelu on-premise. System ten musi zapewniać skalowalność, bezpieczeństwo danych oraz zdolność do integracji z innymi narzędziami, umożliwiając wdrażanie i zarządzanie złożonymi modelami i aplikacjami AI.
- 2.4. Zamawiający wymaga dostarczenia gotowej, zintegrowanej przez Producenta lub Wykonawcę platformy chmury prywatnej, składającej się z infrastruktury sprzętowej i dedykowanego oprogramowania, umożliwiającej:
 - 2.4.1. Możliwość zapewnienia wysokiej dostępności platformy AI jako kluczowego systemu dla wsparcia prowadzenia działalności biznesowej Zamawiającego.
 - 2.4.2. Eliminację problemów kompatybilności na każdej warstwie dzięki zapewnieniu zgodności poszczególnych elementów sprzętu, oprogramowania wbudowanego, oprogramowania systemowego i narzędzi do rozwijania aplikacji AI.
 - 2.4.3. Dostosowanie do zmieniających się potrzeb biznesowych poprzez możliwość automatycznego skalowania zasobów.
 - 2.4.4. Zapewnienie wysokiej wydajności platformy dzięki zintegrowanemu stosowi technologicznemu.
 - 2.4.5. Zapewnieni integracji między hardware'em i oprogramowaniem producenta platformy oraz sprzętem i oprogramowaniem firmy NVIDIA .

- 2.5. System musi zapewniać kompleksowe rozwiązania analityczne i AI, dostosowane do wymagań nowoczesnych przedsiębiorstw. W szczególności musi:
- 2.5.1. Umożliwiać przetwarzanie danych w czasie rzeczywistym i wsadowym.
 - 2.5.2. Zapewniać narzędzia do trenowania modeli uczenia maszynowego (ML), modeli sztucznej inteligencji (AI) oraz do ich wdrażania i monitorowania.
 - 2.5.3. Zapewniać narzędzia zdolne wesprzeć każdy etap obróbki i zarządzania danymi – inżynierię danych (ang. data engineering), naukę o danych (ang. data science) oraz analitykę – umożliwiając dostosowanie technologii uczenia maszynowego i sztucznej inteligencji do specyficznych potrzeb Zamawiającego.
 - 2.5.4. Zapewniać narzędzia do zarządzania modelami, automatycznego skalowania procesu inferencji oraz podmiany online wersji modelu zarówno przez GUI, jak i API.
 - 2.5.5. Zapewniać mechanizm automatycznego zarządzania wykorzystaniem GPU.
 - 2.5.6. Oferować łatwe w użyciu interfejsy do generowania raportów, analiz ad-hoc oraz wizualizacji.
 - 2.5.7. System musi być tak zwymiarowany, by móc uruchomić modele AI, które będą dostępne na platformie w formacie zoptymalizowanym do uruchomienia na zaproponowanym akceleratorze graficznym.
- 2.6. Wymagane jest wykorzystanie następujących wzorców architektonicznych, stanowiących dobre praktyki dla współczesnych rozwiązań obliczeniowych:
- 2.6.1. Wykorzystanie architektury przetwarzania rozproszonego ze skalowaniem horyzontalnym (ang. scale-out) w celu maksymalizacji wydajności przetwarzania i usunięcia ograniczeń w skalowaniu mocy obliczeniowej.
 - 2.6.2. Wykorzystanie architektury zdezagregowanych magazynów danych ze skalowaniem horyzontalnym (ang. scale-out) w celu maksymalizacji wydajności obsługi danych i usunięcia ograniczeń w skalowaniu pojemności i wydajności magazynów danych.
 - 2.6.3. Wykorzystanie mechanizmów akceleracji obliczeń w postaci kart GPU zaprojektowanych i dedykowanych do obciążeń AI / ML, jednocześnie zapewniając znaczne zasoby mocy obliczeniowej CPU na potrzeby usług towarzyszących akceleracji obliczeń.
 - 2.6.4. Wykorzystanie mechanizmów akceleracji przepływu danych w systemie obliczeniowym w postaci technologii RDMA, RoCE v2, DCB, GPUDirect i GPUDirect Storage.
 - 2.6.5. Wykorzystanie mechanizmów konteneryzacji aplikacji w celu przyspieszenia wdrażania aplikacji i uproszczenia zarządzania ich cyklem życia.
 - 2.6.6. Wykorzystanie optymalnych modeli realizacji poszczególnych warstw systemu:

- 2.6.6.1. Model on-premise dla warstwy obliczeniowej i warstwy przechowywania i dostarczania danych, zapewniający prywatność, kontrolę i bezpieczeństwo obróbki i przechowywania danych oraz pełnię kontroli nad cyklem życia aplikacji i usług.
- 2.6.6.2. Model hybrydowy dla warstwy zarządzającej, zapewniający lokalne mechanizmy zarządzania zasobami połączone z centralizacją monitorowania i zarządzania instancjami platformy oraz z uproszczeniem procesu aktualizacji platformy.
- 2.6.7. Rozdzielenie warstwy zarządzania od warstwy przetwarzania w celu zagwarantowania dostępności zasobów dla poszczególnych warstw, umożliwienia niezależnego skalowania poszczególnych warstw oraz poprawy bezpieczeństwa rozwiązania.
- 2.7. Rozwiązanie musi obsługiwać szeroką gamę modeli sztucznej inteligencji, w tym uczenie maszynowe (ML), głębokie uczenie (DL) oraz zaawansowane modele językowe, które pozwolą Zamawiającemu na wykorzystywanie AI w różnorodnych zastosowaniach, takich jak analiza predykcyjna, automatyzacja procesów, analiza obrazów czy generatywna sztuczna inteligencja. Platforma musi zawierać narzędzia do:
 - 2.7.1. Trenowania modeli.
 - 2.7.2. Zarządzania wersjami modeli oraz ich parametrami.
 - 2.7.3. Wdrażania modeli na środowiskach produkcyjnych (ang. inference).
 - 2.7.4. Monitorowania jakości i wydajności modelu (ang. model metrics, drift).
 - 2.7.5. Integracji z narzędziami orkiestrowania procesów.
 - 2.7.6. Definiowania wieloetapowych potoków danych ang. (pipelines), zarządzania eksperymentami i automatyzacją procesów.
- 2.8. Rozwiązanie musi zawierać kompleksowe narzędzia do zarządzania całym cyklem życia modeli AI, od etapu koncepcyjnego, przez trenowanie, wdrażanie, aż po monitorowanie i zarządzanie w środowisku produkcyjnym. Dzięki funkcjom takim, jak automatyzacja trenowania modeli, monitorowanie ich wydajności oraz zarządzanie wersjami, musi zapewniać możliwość skutecznej kontroli i optymalizacji.
- 2.9. Platforma AI musi zapewniać dostęp do modeli, komponentów i narzędzi oferowanych w portfolio NVIDIA. Platforma musi wspierać natywnie integrację z repozytorium modeli NVIDIA NGC oraz możliwość publikowania modeli.
- 2.10. Rozwiązanie musi umożliwiać przetwarzanie danych w trybie MPP (Massively Parallel Processing). Platforma musi umożliwiać zapytania rozproszone na dużych zbiorach danych w trybie równoległym przy użyciu języka SQL. System musi

zapewniać zaawansowane narzędzie do wizualizacji. Niezbędny jest komponent, który pozwoli na interaktywną analizę, budowanie pulpitów menedżerskich i raportów wizualizujących dane.

- 2.11. Platforma musi obsługiwać duże zbiory danych i złożone obliczenia. System musi umożliwiać dynamiczne skalowanie zasobów obliczeniowych, dzięki czemu możliwe będzie efektywne przetwarzanie i analizowanie ogromnych ilości danych w czasie rzeczywistym.
- 2.12. System musi posiadać zaawansowane funkcje zabezpieczeń, które spełniają wymagania Polityki Bezpieczeństwa Zamawiającego, a także zapewniać ochronę danych przed nieautoryzowanym dostępem oraz możliwość zarządzania uprawnieniami i kontrolą dostępu.
- 2.13. System musi być zaprojektowany by przechowywać i przetwarzać dane lokalnie, a tym samym zapewniać pełną kontrolę nad danymi oraz elastyczne dostosowanie środowiska obliczeniowego do potrzeb Zamawiającego.
- 2.14. Platforma musi umożliwiać trenowanie własnych modeli AI przy użyciu danych specyficznych dla Zamawiającego i różnych algorytmów uczenia maszynowego – w tym modele głębokiego uczenia (deep learning), klasyczne modele statystyczne oraz inne techniki AI.
- 2.15. System musi umożliwiać użytkownikom dostęp do wstępnie wytrenowanych modeli AI z ekosystemu NVIDIA.
- 2.16. Platforma musi umożliwiać użytkownikom dostęp do wstępnie zbudowanych rozwiązań z ekosystemu NVIDIA Blueprints.
- 2.17. System musi oferować narzędzia do automatyzacji procesu tworzenia i trenowania modeli.
- 2.18. System musi umożliwiać monitorowanie wydajności i efektywności trenowanych modeli. Platforma musi wspierać również zarządzanie wersjami modeli i ich aktualizację.
- 2.19. Platforma musi pozwalać na łatwą integrację z wieloma źródłami danych, takimi jak:
 - 2.19.1. bazy danych,
 - 2.19.2. pliki,
 - 2.19.3. formaty Open Table Format,
 - 2.19.4. systemy przechowywania danych typu S3.
- 2.20. Platforma musi obsługiwać różne rodzaje baz danych, minimalnie:
 - 2.20.1. relacyjne,
 - 2.20.2. NoSQL,
 - 2.20.3. zintegrowane bazy danych,
 - 2.20.4. magazyny danych.

- 2.21. Platforma musi udostępniać narzędzia do czyszczenia, filtrowania i przygotowywania danych. System musi oferować możliwość operacji na danych, przynajmniej:
- 2.21.1. transformacja,
 - 2.21.2. agregacja,
 - 2.21.3. łączenie,
 - 2.21.4. filtrowanie.
- 2.22. System musi wspierać operacje na danych z wykorzystaniem mocy akceleratorów GPU.
- 2.23. System musi posiadać mechanizmy szyfrowania danych, szczególnie w kontekście ochrony prywatności i zgodności z regulacjami prawnymi, takimi jak RODO. Platforma musi oferować mechanizmy kontroli dostępu, umożliwiające nadawanie ról i uprawnień.
- 2.24. System musi wspierać nadzór nad tworzonymi modelami, jego wynikami i decyzjami. Platforma musi zapewnić funkcje monitorowania, które pozwalają na śledzenie działania modeli AI w czasie rzeczywistym.
- 2.25. Częścią rozwiązania musi być nowoczesna platforma do zarządzania operacjami IT (ITOM), która zapewni graficzną i dynamiczną formę stanu systemu, automatyzację i monitorowanie dostarczonej platformy.

3. WYMAGANIA SZCZEGÓŁOWE

- 3.1. Rozwiązanie musi zawierać moduł monitorowania, umożliwiający podgląd szczegółowych informacji o stanie platformy.
- 3.2. Platforma musi dawać możliwość obserwacji danych telemetrycznych za pośrednictwem dostarczonego modułu on-premise lub portalu chmurowego.
- 3.3. Rozwiązanie musi umożliwiać obsługę aktualizacji oprogramowania i oprogramowania układowego za pośrednictwem portalu on-premise lub portalu chmurowego.
- 3.4. Rozwiązanie musi obejmować możliwość bezproblemowego dołączania dodatkowych aplikacji do platformy za pośrednictwem Helm Charts.
- 3.5. Rozwiązanie musi obejmować bezproblemowe logowanie jednokrotne (SSO) do wszystkich dostarczonych elementów.
- 3.6. Platforma musi oferować mechanizm automatycznego zarządzania dostępem do GPU. Platforma musi oferować interfejs graficzny z panelem administracyjnym pokazujący przypisanie ilości akceleratorów graficznych do uruchomionych aplikacji wraz z ich priorytetem oraz maksymalnym czasem bezczynności w zakresie wykorzystania GPU. Musi umożliwiać zmianę czasu bezczynności, po którym

aplikacja zostanie odpisana od GPU, a GPU zostanie przypisane do następnej aplikacji oczekującej na dostęp do GPU.

- 3.7. Platforma musi posiadać mechanizmy zarządzania i separacji dostępu do poszczególnych aplikacji i danych za pomocą mechanizmu projektów. Platforma musi umożliwiać zarządzanie cyklem życia projektów (tworzenie, usuwanie, modyfikacja) oraz konfigurację dostępu użytkowników do projektów (tworzenie projektów prywatnych, współdzielonych, zmiana dostępu dla poszczególnych użytkowników). Platforma musi ponadto umożliwiać ograniczenie zużycia zasobów GPU za pomocą mechanizmu kwot – możliwość ograniczenia zużycia ilości GPU przez poszczególne projekty.
- 3.8. Platforma musi umożliwiać przypisywanie osobnych poziomów dostępu do poszczególnych źródeł danych i udostępnianych modeli na poziomie projektu – każdy projekt musi mieć możliwość indywidualnej konfiguracji dostępu do danych i modeli platformy.
- 3.9. Platforma musi zawierać zestaw oprogramowania orkiestrującego zgodnego z technologią i standardami Kubernetes (CNCF).
- 3.10. Platforma musi zawierać oprogramowanie, które umożliwia dostarczanie i zarządzanie maszynami wirtualnymi.
- 3.11. Platforma musi mieć charakter otwarty – musi umożliwiać instalację dodatkowego oprogramowania na platformie za pomocą technologii konteneryzacji aplikacji.

4. WYMAGANIA TECHNICZNE

- 4.1. Wymaganiem jest aby oferowana platforma było podzielona na dwa rodzaje serwerów:
 - 4.1.1. serwery obliczeniowe do przetwarzania danych (ang. worker node)
 - 4.1.2. serwery zarządzające pełniące rolę zarządzania systemem (ang. control node)
- 4.2. Żaden z serwerów nie może pełnić obu tych ról jednocześnie.
- 4.3. Parametry sprzętowe platformy:

Warstwa	Zasoby
Serwery zarządzające	Minimum 3 serwery do zarządzania platformą i koordynacji zasobów. Każdy serwer musi spełniać wymagania z punktu 4.4 „Wymagania dotyczące serwera zarządzającego”.
Serwery obliczeniowe	Minimum 1 serwer zoptymalizowany pod kątem obciążeń AI. Każdy serwer musi spełniać wymagania z punktu 4.5 „Wymagania dotyczące serwera obliczeniowego” oraz być wyposażony w co najmniej 1 kartę z 2 portami 200Gb/s każdy oraz 4 karty z przynajmniej 1 portem 400Gb/s.

	Możliwość co najmniej dwukrotnego zwiększenia mocy obliczeniowej platformy.
Pamięć masowa	Minimum 100 TB współdzielonej użytkowej przestrzeni dyskowej oraz minimum 2 redundantne kontrolery dyskowe. Pamięć masowa musi spełniać wymagania z punktu 4.6 „Wymagania dotyczące macierzy plikowej”.
Sieć obliczeniowa	2 przełączniki 400Gb/s Ethernet ze wsparciem dla RoCE v2 oraz DCB na potrzeby aktualnych zasobów obliczeniowych oraz na potrzeby przyszłej rozbudowy mocy obliczeniowej. Każdy przełącznik musi spełniać wymagania z punktu 4.7 „Przełącznik obliczeniowy”.
Sieć zarządzająca	2 przełączniki 1Gb Ethernet sieci zarządzającej. Przełączniki spełniające szczegółowe wymagania z punktu 4.8 „Przełącznik zarządzający”.

4.4. Wymagania dotyczące serwera zarządzającego:

Element konfiguracji	Wymagania minimalne
Obudowa	Maksymalnie 1U RACK 19 cali wraz z szynami montażowymi.
Płyta główna	Płyta główna z fabrycznym oznaczeniem logo producenta (dopuszcza się logo producenta na module zarządzania trwale zintegrowanym na płycie głównej).
Procesor	Zainstalowany 1 procesor x86 - 64 bity, posiadający 32 rdzenie i pracujący z nominalną częstotliwością 2.6 GHz. Wymagany cache L3 Cache min. 128MB.
Pamięć operacyjna	384GB RDIMM DDR5 4800 MT/s w modułach min. 64GB.
Gniazda rozszerzeń	Serwer musi być wyposażony w 1 aktywne gniazdo PCI-Express generacji 5 lub nowsze gotowe do obsadzenia kartami sieciowymi, każde gniazdo x16, pełnej wysokości i pełnej długości.
Dysk twardy	Zainstalowane: <ul style="list-style-type: none"> 8 dysków każdy 1.6TB NVMe Gen4 Mixed Use, DWPD 3, IOPS dla zapisów co najmniej 270 000 . Dyski podłączone bezpośrednio do CPU. 2 dyski każdy 1.6TB SAS 24G Mixed Use, DWPD 3, IOPS dla zapisów co najmniej 160 000. Dyski podłączone do kontrolera dyskowego, skonfigurowane w RAID 1.
Kontroler	Serwer wyposażony w sprzętowy kontroler RAID zapewniający obsługę RAID 0/1/10.
Interfejsy sieciowe i FC	Serwer musi być wyposażony w: <ul style="list-style-type: none"> 2 karty dwuportowe 10/25Gb Ethernet SFP28
Karta graficzna	Zintegrowana karta graficzna z obsługą rozdzielczości minimum 1920 x 1200 60Hz (32 bpp)
Bezpieczeństwo	Wbudowany moduł TPM 2.0.
Porty	Wbudowane 4 porty USB 3.2 w tym 2 wewnętrzne.

	1x VGA z tyłu obudowy.
Zasilacz	2 szt., typu Hot-plug typu Titanium, redundantne, każdy o mocy minimum 1000W.
Chłodzenie	Zestaw wentylatorów redundantnych typu hot-plug.
Karta/moduł zarządzający	<p>Niezależna od systemu operacyjnego, zintegrowana z płytą główną serwera lub jako dodatkowa karta w slotcie PCI Express, jednak nie może ona powodować zmniejszenia minimalnej liczby gniazd PCIe w serwerze, posiadająca minimalną funkcjonalność:</p> <ul style="list-style-type: none"> • monitorowanie podzespołów serwera: temperatura, zasilacze, wentylatory, procesory, pamięć RAM, kontrolery macierzowe i dyski (fizyczne i logiczne), karty sieciowe • możliwość pracy w trybie bezagentowym – bez konieczności instalacji agentów zarządzania w systemie operacyjnym z generowaniem alertów SNMP • dostęp do karty zarządzającej poprzez <ul style="list-style-type: none"> – dedykowany port RJ45 z tyłu serwera – lub przez współdzielony port zintegrowanej karty sieciowej serwera • dostęp do karty możliwy <ul style="list-style-type: none"> – z poziomu przeglądarki webowej (GUI) – z poziomu wiersza poleceń; – poprzez IPMI 2.0 (Intelligent Platform Management Interface) • wirtualna zdalna konsola, tekstowa i graficzna, z dostępem do myszy, klawiatury i możliwością podłączenia wirtualnych napędów CD/DVD, USB oraz wirtualnych folderów • monitorowanie zasilania oraz zużycia energii przez serwer w czasie rzeczywistym z możliwością graficznej prezentacji • konfiguracja maksymalnego poziomu pobieranej mocy przez serwer (capping) • zdalna aktualizacja oprogramowania (firmware) • wsparcie dla Microsoft Active Directory • wsparcie dla IPv4 oraz IPv6, obsługa SNMP v3 oraz RESTful API • możliwość autokonfiguracji sieci karty zarządzającej (DNS/DHCP)
Wsparcie minimum dla systemów operacyjnych i systemów wirtualizacyjnych	<p>Microsoft Windows Server 2022, 2025 Red Hat Enterprise Linux (RHEL) 9.4, 10.0 SUSE Linux Enterprise Server (SLES) 15 SP4 lub nowszy VMware ESXi 8.0 U3 i 9.0 i nowsze</p>

4.5. Wymagania dotyczące serwera obliczeniowego:

Element konfiguracji	Wymagania minimalne
Obudowa	Maksymalnie 4U RACK 19 cali wraz z szynami i ramieniem na kable.
Płyta główna	Płyta główna z fabrycznym oznaczeniem logo producenta (dopuszcza się logo producenta na module zarządzania trwale zintegrowanym na płycie głównej).
Procesory	Zainstalowane 2 procesory x86 - 64 bity, posiadające po min.84 rdzenie każdy i pracujące z nominalną częstotliwością min.1.9 GHz. Wymagany cache L3 Cache min. 336 MB.
Pamięć operacyjna	1TB RAM RDIMM DDR5 6400 MT/s w modułach o pojemności 32GB.
Dedykowane karty GPU	Serwer wyposażony w 4 kart GPU RTX PRO 6000 Blackwell Server Edition 96GB PCIe lub równoważne. Warunki równoważności dla pojedynczej karty przedstawiono poniżej: <ul style="list-style-type: none"> • Pamięć karty graficznej: 96GB GDDR7 • Typ interfejsu: PCI Express 5.0 x16 • Przepustowość pamięci GPU: 1.6 TB/s • Wydajność teoretyczna FP32 117 TFLOPS
Gniazda rozszerzeń	Po zainstalowaniu wymaganych kart GPU serwer musi być wyposażony w 5 aktywnych gniazd PCI-Express generacji 5 pełnej wysokości (full height) x 16.
Dyski	Zainstalowane 4 dyski każdy dysk o parametrach: 3.2TB NVMe Gen5 typu Mixed Use, DWPD 3, IOPS dla zapisów co najmniej 550 000. Dyski podłączone bezpośrednio do CPU. Zainstalowany moduł z dwoma dyskami Hot-Plug NVMe M.2 zapewniających minimalną pojemność 480 GB i redundancję danych RAID-1.
Interfejsy sieciowe	Serwer musi być wyposażony w: <ul style="list-style-type: none"> • Minimum cztery karty jednoportowe 400Gb Ethernet ze wsparciem dla RoCE v2, NVIDIA GPUDirect, NVIDIA GPUDirect Storage. • Minimum jedna karta dwuportowa 200Gb Ethernet ze wsparciem dla RoCE v2, NVIDIA GPUDirect, NVIDIA GPUDirect Storage.
Karta graficzna	Zintegrowana karta graficzna, umożliwiająca wyświetlanie obrazu w rozdzielczości minimum 1920 x 1200 pikseli.
Porty	Wbudowane 4 x USB 3.0 (1 z przodu, 2 z tyłu, 1 wewnętrzny) Dodatkowy port USB/USB-C z przodu obudowy serwera umożliwiający dostęp do interfejsu zdalnego zarządzania. 1x VGA
Zasilacz	8 szt., typu Hot-plug, redundantne, każdy o mocy minimum 2400W, efektywność Titanium.
Bezpieczeństwo	Wbudowany moduł TPM 2.0.
Chłodzenie	Zestaw wentylatorów redundantnych typu hot-plug.

Karta/moduł zarządzający	<p>Niezależna od system operacyjnego, zintegrowana z płytą główną serwera lub jako dodatkowa karta w slotcie PCI Express, jednak nie może ona powodować zmniejszenia minimalnej liczby gniazd PCIe w serwerze, posiadająca minimalną funkcjonalność:</p> <ul style="list-style-type: none"> • monitorowanie podzespołów serwera: temperatura, zasilacze, wentylatory, procesory, pamięć RAM, kontrolery macierzowe i dyski (fizyczne i logiczne), karty sieciowe • możliwość pracy w trybie bezagentowym – bez agentów zarządzania instalowanych w systemie operacyjnym z generowaniem alertów SNMP • dostęp do karty zarządzającej poprzez <ul style="list-style-type: none"> – dedykowany port RJ45 z tyłu serwera lub – przez współdzielony port zintegrowanej karty sieciowej serwera dostęp do karty możliwy <ul style="list-style-type: none"> – z poziomu przeglądarki webowej (GUI) – z poziomu linii komend; – poprzez interfejs IPMI 2.0 (Intelligent Platform Management Interface) • wirtualna zadalna konsola, tekstowa i graficzna, z dostępem do myszy i klawiatury i możliwością podłączenia wirtualnych napędów CD/DVD i USB i i wirtualnych folderów • monitorowanie zasilania oraz zużycia energii przez server w czasie rzeczywistym z możliwością graficznej prezentacji • konfiguracja maksymalnego poziomu pobieranej mocy przez server (capping) • zdalna aktualizacja oprogramowania (firmware) • wsparcie dla Microsoft Active Directory • wsparcie dla IPv4 oraz IPv6, obsługa SNMP v3 oraz RESTful API • możliwość autokonfiguracji sieci karty zarządzającej (DNS/DHCP)
Wsparcie minimum dla systemów operacyjnych i systemów wirtualizacyjnych	<p>Microsoft Windows Server 2022, 2025 Red Hat Enterprise Linux (RHEL) 9.4 i 10 lub nowsze SUSE Linux Enterprise Server (SLES) 15SP6 i 16 lub nowsze VMware ESXi 8.0U3 i 9.0 lub nowsze</p>

4.6. Wymagania dotyczące macierzy plikowej:

- 4.6.1. Elementem oferowanej platformy musi być macierz plikowa zaprojektowana w oparciu o architekturę mikrousługową, która w natywny sposób realizuje dostęp do danych po protokołach NFS, SMB i S3.

- 4.6.2. Oferowana macierz musi umożliwiać niezależne skalowanie wydajności i pojemności poprzez możliwość rozbudowy niezależnie o kontrolery i półki dyskowe. Architektura macierzy musi zapewniać dostęp każdego kontrolera dyskowego do każdego zainstalowanego dysku NVMe.
- 4.6.3. Oferowana macierz musi zapewniać globalny (w ramach macierzy) dostęp zarówno do danych, jak i metadanych, bez konieczności tworzenia kopii danych w pamięci cache kontrolerów dyskowych. Awaria jednego lub większej liczby kontrolerów dyskowych nie może wymagać przebudowy lub równoważenia danych pomiędzy dyskami.
- 4.6.4. Oferowana macierz musi być w całości oparta o architekturę NVMe, co oznacza, że komunikacja pomiędzy kontrolerami dyskowymi a dyskami musi wykorzystywać protokół NVMe-oF lub linie PCIe, a wszystkie dyski muszą być w technologii SSD NVMe.
- 4.6.5. Oferowana macierz musi być wyposażona w redundantne kontrolery dyskowe oraz przestrzeń dyskową zbudowaną z wykorzystaniem dysków SSD NVMe. Ilość kontrolerów oraz pojemność użytkowa (bez uwzględniania mechanizmów deduplikacji i kompresji) nie może być mniejsza niż wymagana w szczegółowych parametrach dla platformy sprzętowej (punkt 4.3).
- 4.6.6. W celu minimalizacji opóźnień zapisu macierz musi oferować dodatkowy bufor dyskowy zrealizowany z wykorzystaniem szybkiej pamięci klasy SCM (Storage Class Memory). W przypadku braku obsługi pamięci klasy SCM, kontrolery dyskowe muszą być wyposażone w 1TB pamięci RAM każdy.
- 4.6.7. Kontrolery dyskowe oferowanej macierzy muszą być wyposażone w wielordzeniowe procesory x86 (minimum 32 rdzenie per procesor). Zamawiający zaakceptuje jedynie rozwiązania wyposażone w procesory firm Intel lub AMD. Każdy kontroler musi posiadać co najmniej 256 GB pamięci RAM. Zastosowane procesory muszą wspierać minimum standard PCIe Gen4.
- 4.6.8. Oferowana macierz nie może wykorzystywać pamięci DRAM lub innej pamięci ulotnej jako pamięci podręcznej do zapisu, niezależnie od sposobu zabezpieczenia takiej pamięci (dotyczy to także modułów NVDIMM).
- 4.6.9. Każdy z kontrolerów dyskowych musi być wyposażony w co najmniej 2 porty Ethernet 100 Gb/s przeznaczone do dostępu do danych oraz co najmniej 2 porty Ethernet 100 Gb/s przeznaczone do obsługi półek dyskowych.
- 4.6.10. Oferowana macierz musi natywnie obsługiwać protokoły NFS v3, NFS v4.1, SMB 2.x, SMB 3.x Multi-channel i S3.
- 4.6.11. Oferowana macierz musi obsługiwać uwierzytelnianie Kerberos dla protokołów NFS i SMB oraz uwierzytelnianie NTLMv2 dla klientów SMB.

- 4.6.12. Macierz musi obsługiwać wieloprotokołowy dostęp do obiektów za pośrednictwem interfejsu API Amazon S3 oraz plików za pośrednictwem NFS i SMB zlokalizowanych w pojedynczej, rozproszonej przestrzeni nazw.
- 4.6.13. Oferowana macierz musi obsługiwać mechanizm NFS nconnect w celu zwiększenia wydajności NFS. Musi istnieć możliwość zdefiniowania co najmniej 16 połączeń TCP pomiędzy każdym klientem a systemem pamięci masowej.
- 4.6.14. Oferowana macierz musi obsługiwać mechanizm NFS over RDMA i multi-pathing w celu zwiększenia wydajności NFS podczas łączenia klienta z systemem pamięci masowej. Multi-pathing musi działać w połączeniu z NFS nconnect i NFS over RDMA.
- 4.6.15. Oferowana macierz musi obsługiwać mechanizm alternatywnych strumieni danych (ADS) dla protokołu SMB, umożliwiający przypisanie dodatkowych metadanych do plików.
- 4.6.16. Oferowana macierz musi obsługiwać mechanizmy ciągłej dostępności dla udziałów SMB oraz dostęp wielokanałowy w celu zwiększenia przepustowości i zapewnienia redundancji połączeń.
- 4.6.17. Macierz musi obsługiwać blokowanie i wersjonowanie obiektów.
- 4.6.18. Oferowana macierz musi oferować wieloudział (ang. multitenancy) rozumiany jako możliwość utworzenia wirtualnych serwerów plikowych, które funkcjonują w izolowanych domenach bezpieczeństwa.
- 4.6.19. Oferowana macierz musi oferować funkcje wirtualnych adresów IP w celu dystrybucji i równoważenia obciążenia klienta na wszystkich oferowanych kontrolerach.
- 4.6.20. Oferowana macierz musi umożliwiać tworzenie wielu wirtualnych pul adresów IP, tak aby do danej puli można było przypisać dedykowany zestaw kontrolera i wybranych portów dostępowych w ramach tego kontrolera w celu uzyskania środowiska wielodostępowego. Musi istnieć możliwość zdefiniowania wirtualnych pul adresów IP z możliwością ich przypisania do różnych wirtualnych serwerów plikowych.
- 4.6.21. Wirtualna pula adresów IP powinna obsługiwać zarówno protokół IPv4, jak i IPv6. Musi istnieć możliwość utworzenia oddzielnej wirtualnej puli adresów IP dla dostępu do danych oraz replikacji.
- 4.6.22. Oferowana macierz musi umożliwiać definiowanie limitów pojemności dla poszczególnych katalogów, użytkowników i grup użytkowników. Wykorzystana pojemność musi być mierzona na podstawie wykorzystania bajtów w pliku, a nie na logicznym rozmiarze pliku.

- 4.6.23. Zarządzanie limitami pojemności musi obejmować miękkie i twarde limity. Po osiągnięciu twardego limitu pojemności macierz musi blokować możliwość dalszego zapisu danych.
- 4.6.24. Oferowana macierz musi umożliwiać natywne zarządzania wydajnością poprzez definiowanie limitów wydajności w IOPS i MB/s na poziomie poszczególnych udziałów sieciowych / katalogów, na poziomie wirtualnych serwerów plikowych oraz na poziomie poszczególnych użytkowników.
- 4.6.25. Musi istnieć możliwość zdefiniowania statycznych limitów wydajnościowych dla udostępnianego zasobu w zakresie parametrów przepustowości i operacji we/wy (IOPS) oraz dynamicznych limitów w zakresie ww. parametrów dla wykorzystanej i udostępnionej pojemności tak, aby wraz ze wzrostem pojemności rosła wydajność.
- 4.6.26. Oferowana macierz musi posiadać wbudowane mechanizmy kompresji i deduplikacji danych wykonywane w locie, przed zapisem danych na dyski SSD NVMe.
- 4.6.27. Mechanizm deduplikacji musi identyfikować i redukować nie tylko identyczne bloki danych, ale także bazować na podobieństwach, to znaczy zapisywać jedynie różnicę między podobnymi blokami pod warunkiem, że jest to korzystne z punktu widzenia uzyskania lepszego współczynnika redukcji danych.
- 4.6.28. Oferowana macierz musi obsługiwać szyfrowanie danych. Szyfrowanie danych musi wykorzystywać algorytm AES-256.
- 4.6.29. Oferowana macierz musi obsługiwać zarówno wewnętrzne, jak i zewnętrzne zarządzanie kluczami.
- 4.6.30. Kontrolery dyskowe oferowanej macierzy muszą pracować w trybie bezstanowym tak, aby awaria wielu kontrolerów nie miała wpływu na pojemność i dostępność do danych. Macierz musi zapewniać odporność na awarię N-1 (wszystkich bez jednego) kontrolerów dyskowych obsługujących dostęp do danych.
- 4.6.31. Oferowana macierz musi zapewniać zabezpieczenie przed utratą spójności danych w przypadku jednoczesnej awarii czterech dowolnych dysków. Dane i parzystość muszą być rozproszone po wszystkich dyskach macierzy bez konieczności dedykowania dysków parzystości i dysków zapasowych. W przypadku awarii dysku macierz musi zapewniać automatyczną odbudowę danych na pozostałych dyskach.
- 4.6.32. Oferowana macierz nie może posiadać pojedynczego punktu awarii, który powodowałby brak dostępu do danych. Musi być zapewniona pełna redundancja komponentów, w szczególności zdublowanie kontrolerów, zasilaczy i chłodzenia.

- 4.6.33. Oferowana macierz musi umożliwiać wykonywanie aktualizacji mikrokodu w trybie online bez przerywania dostępu do zasobów.
- 4.6.34. Każda para kontrolerów dyskowych musi zapewniać wydajność odczytu sekwencyjnego na poziomie nie mniejszym niż 25 GB/s.
- 4.6.35. Zarządzanie zaoferowaną macierzą musi być możliwe z poziomu lokalnego interfejsu graficznego, linii poleceń oraz poprzez REST API.
- 4.6.36. Oprogramowanie do zarządzania musi pozwalać na stałe monitorowanie stanu macierzy oraz umożliwiać konfigurowanie jej zasobów dyskowych. Narzędzie musi pozwalać na obserwację danych wydajnościowych oraz prezentację ich w postaci wykresów oraz czytelnych raportów. Wymagane jest monitorowanie bieżących parametrów pracy macierzy.
- 4.6.37. Wbudowany graficzny interfejs zarządzania musi umożliwiać monitorowanie wykorzystania pojemności, w tym także zużycia pojemności przez poszczególne udziały sieciowe / katalogi wraz z rozkładem zajętej pojemności przez podkatalogi oraz szczegółowymi informacjami na temat redukcji rozmiaru danych.
- 4.6.38. Oferowana macierz musi zapewniać obsługę uwierzytelniania AD, LDAP i NIS.
- 4.6.39. Oferowana macierz musi posiadać konfigurowalny dziennik audytu dla protokołów NFS i SMB pozwalający na logowanie wszystkich krytycznych działań, takich jak tworzenie plików, ścieżek i katalogów. Musi istnieć możliwość zdefiniowania okresu retencji dla plików dziennika audytu od kilku godzin do nieskończoności.
- 4.6.40. Oferowana macierz musi zapewniać integrację z Kubernetes przy użyciu sterownika CSI (Container Storage Interface). Mechanizm musi obsługiwać zarówno dynamiczne przypisywanie zasobów, jak i integrację z kopiami migawkowymi.
- 4.6.41. Wszystkie wymagane funkcje muszą być zawarte w pełnej wersji wraz z odpowiednimi licencjami (jeśli takie licencje wymagane) dla pełnej pojemności oferowanej macierzy plikowej.

4.7. Przełącznik obliczeniowy:

Element konfiguracji	Wymagania minimalne
Obudowa	Maksymalnie 1U RACK 19 cali wraz z zestawem do montażu w szafie rack.
Porty sieciowe	32 porty 40/100/200/400Gb Ethernet
Wspierane technologie i protokoły	<ul style="list-style-type: none"> • IPv4 • IPv6 • Bridging and Spanning Tree (802.1D) • QoS (802.1p) • VLAN Tagging (802.1Q)

	<ul style="list-style-type: none"> • Rapid Spanning Tree (802.1w) • Multiple Spanning Tree Protocol (802.1s) • Link Layer Discovery Protocol (802.1AB) • ETS (802.1Qaz) • PFC (802.1Qbb) • Link Aggregation with LACP (802.3ad) • Flow Control (802.3x) • Border Gateway Protocol (BGP) • Open Shortest Path First (OSPF) • RDMA over Converged Ethernet (RoCEv2) • Virtual routing and forwarding (VRF) • Bidirectional forwarding detection (BFD) • VXLAN • MLAG
--	--

4.8. Przełącznik zarządzający:

Element konfiguracji	Wymagania minimalne
Obudowa	Maksymalnie 1U RACK 19 cali wraz z zestawem do montażu w szafie rack.
Porty sieciowe	<ul style="list-style-type: none"> • 48 portów 1Gb Ethernet Base-T • 4 porty 1/10/25Gb Ethernet
Wspierane technologie i protokoły	<ul style="list-style-type: none"> • IPv4 • IPv6 • LACP (802.3ad) • Multiple Spanning Tree (802.1s) • Ethernet Ring Protection Switching (ERPS) • Virtual Router Redundancy Protocol (VRRP) • Bidirectional Forward Detection (BFD) • Traffic prioritization (IEEE 802.1p) • Deficit Weighted Round Robin (DWRR) • Egress Queue Shaping (EQS) • Mechanizm łączenia (ang. stacking) przełączników, pozwalający zarządzać grupą przełączników jako pojedynczym logicznym przełącznikiem

5. WYMAGANIA DLA OPROGRAMOWANIA

5.1. Platforma musi dostarczać wspierane przez producenta oprogramowanie przechowujące i udostępniające modele. Oprogramowanie musi zapewnić interfejs GUI i API, umożliwiać automatyczne skalowanie wykorzystywanych zasobów na potrzeby procesu inferencji na podstawie polityki definiowanej przez użytkownika, a także umożliwiać podmianę modelu na inny w sposób bezprzerwowy

dla aplikacji, która komunikuje się z modelem oraz bez potrzeby modyfikacji parametrów aplikacji (np. adresu sieciowego modelu).

- 5.2. Dołączone modele muszą być dystrybuowane jako obrazy OCI w lokalnym rejestrze kontenerów.
- 5.3. Modele muszą być przechowywane na wielu warstwach obrazu OCI, aby zapewnić efektywne pobieranie modeli.
- 5.4. Wszystkie spakowane modele muszą być składowane na wysokowydajnej pamięci masowej NVMe.
- 5.5. Wszystkie modele muszą być udostępnione użytkownikom za pośrednictwem mechanizmu katalogu, pozwalając na uproszczone i zautomatyzowane wdrażanie modeli dla aplikacji AI.
- 5.6. Platforma musi dostarczać mechanizm kontroli dostępu do modeli, tak aby tylko autoryzowani użytkownicy mogli wdrażać i korzystać z odpowiednich modeli.
- 5.7. Platforma musi zawierać zoptymalizowane pod kątem użytych akceleratorów graficznych duże modele językowe w celu poprawy wydajności uczenia maszynowego i aplikacji AI.
- 5.8. Spakowane modele językowe muszą być dostarczane jako łatwe w użyciu mikrousługi.
- 5.9. Platforma musi zawierać wyselekcjonowane i obsługiwane przez producenta zestawy narzędzi typu open source, aby wspierać inżynierów danych i analityków danych z poziomu interfejsu użytkownika. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.
- 5.10. Proponowana platforma musi zawierać oprogramowanie do automatyzacji potoków inżynierii danych. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.
- 5.11. Proponowana platforma musi zawierać oprogramowanie do obsługi pulpitów nawigacyjnych BI i wizualizacji danych. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.
- 5.12. Platforma musi zawierać rozproszony silnik SQL na potrzeby inżynierii danych. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.
- 5.13. Silnik SQL musi obsługiwać zapytania federacyjne dystrybuowane między różnymi zewnętrznymi źródłami danych, takimi jak relacyjne bazy danych, baza danych noSQL i hurtownie danych. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.

- 5.14. Silnik SQL musi obsługiwać rozproszone buforowanie często używanych zestawów danych z czasem ich wygaśnięcia. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.
- 5.15. Silnik SQL musi obsługiwać optymalizacje zapytań, takie jak wypychanie predykatów, wypychanie projekcji, filtrowanie dynamiczne i optymalizacja oparta na kosztach. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.
- 5.16. Platforma musi zawierać wielojęzyczny rozproszony aparat analityczny typu open source do obsługi obciążeń analizy danych. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.
- 5.17. Platforma musi posiadać zintegrowane oprogramowanie będące silnikiem przetwarzania danych. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.
- 5.18. Interfejs użytkownika musi posiadać funkcje uruchamiania zadań silnika przetwarzania danych w jednym z dwóch trybów: w trybie wsadowym oraz w trybie interaktywnym. W trybie interaktywnym interfejs użytkownika musi zapewnić możliwość uruchomienia interaktywnej sesji do silnika przetwarzania danych z poniższymi parametrami:
 - 5.18.1. ilość pamięci i liczba rdzeni przypisana do procesu zarządcy,
 - 5.18.2. ilość pamięci, liczba procesów i liczba rdzeni przypisana do procesu wykonawcy.
- 5.19. W trybie wsadowym interfejs użytkownika musi posiadać funkcje uruchamiania i harmonogramowania zadań wsadowych oraz przydzielania im:
 - 5.19.1. ilość pamięci i liczba rdzeni przypisana do procesu zarządcy,
 - 5.19.2. ilość pamięci, liczba procesów i liczba rdzeni przypisana do procesu wykonawcy.
- 5.20. Platforma sprzętowo-programowa musi zawierać oprogramowanie do tworzenia i wdrażania przenośnych i skalowalnych przepływów pracy AI/ML. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.
- 5.21. Platforma sprzętowo-programowa musi zawierać rejestr modeli, aby umożliwić zarządzanie cyklem życia uczenia maszynowego. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.
- 5.22. Platforma sprzętowo-programowa musi zawierać rozwiązanie typu open source, umożliwiające skalowanie aplikacji AI/ML pisanych w języku Python, obsługę

obciążeń rozproszonych i równoległych. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.

- 5.23. Platforma sprzętowo-programowa musi obejmować środowisko IDE zoptymalizowane pod kątem nauki o danych dla poszczególnych użytkowników dla interaktywnych środowisk obliczeniowych, które umożliwiają analitykom danych opracowywanie i uruchamianie aplikacji do nauki o danych (ang. data science). Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.
- 5.24. Platforma sprzętowo-programowa musi umożliwiać użytkownikom łączenie do zewnętrznych aplikacji i narzędzi BI.
- 5.25. Platforma sprzętowo-programowa musi zawierać interfejs do udostępniania dzienników i metryk aplikacjom innych firm. Oprogramowanie to musi być zintegrowane z platformą i posiadać wsparcie producenta platformy sprzętowo-programowej.
- 5.26. Platforma sprzętowo-programowa musi umożliwiać użytkownikom korzystanie z własnych aplikacji kontenerowych obok dołączonych narzędzi open source.
- 5.27. Musi istnieć możliwość monitorowania statystyk pomiarów, w tym śledzenia zużytych zasobów dla poszczególnych aplikacji.
- 5.28. Aktualizacje elementów platformy muszą być zautomatyzowane. Musi istnieć mechanizm powiadamiający o uaktualnieniach przez portal producenta rozwiązania, tak aby użytkownik mógł je widzieć na portalu dostępowym do systemu. Administrator platformy musi mieć możliwość wyboru momentu rozpoczęcia aktualizacji, rozpoczęta aktualizacja musi przebiegać automatycznie, bez potrzeby dalszej ingerencji administratora (o ile nie wystąpią problemy w czasie aktualizacji). Aktualizacja platformy w postaci jednego elementu (paczki) zawierającego niezbędne komponenty do aktualizacji platformy, bez potrzeby instalacji poszczególnych komponentów pojedynczo. Aktualizacja za pomocą zintegrowanej paczki co najmniej w zakresie: narzędzi zawartych w platformie (w tym: open source), systemów konteneryzacji, systemów wirtualizacji, oprogramowania układowego serwerów, oprogramowania układowego macierzy dyskowych, aktualizacji o nowe funkcjonalności platformy.
- 5.29. Platforma sprzętowo-programowa musi być w stanie łączyć się z zewnętrznymi źródłami danych strukturalnych, w tym z relacyjnymi bazami danych SQL.
- 5.30. Platforma sprzętowo-programowa musi być w stanie podłączyć się do zewnętrznych źródeł danych nieustrukturyzowanych, takich jak magazyny NFS i S3.
- 5.31. Platforma sprzętowo-programowa musi zawierać silnik inferencji, który można zintegrować z wewnętrznymi i zewnętrznymi rejestrami modeli.

- 5.32. Platforma sprzętowo-programowa musi zawierać wstępnie zainstalowane i łatwo dostępne duże modele językowe zoptymalizowane pod kątem producenta akceleratora graficznego.
- 5.33. Wstępnie zainstalowane duże modele językowe zoptymalizowane pod kątem producenta akceleratora graficznego muszą być zoptymalizowane pod kątem wyszukiwania, dostrajania i zarządzania osadzaniem.
- 5.34. Platforma sprzętowo-programowa musi zawierać technologię serwowania modeli typu open source.
- 5.35. Zamawiający oczekuje, że integralnym elementem oferowanej platformy będzie system monitorowania. Musi on obejmować zestaw narzędzi pozwalających monitorować, kontrolować oraz optymalizować działanie dostarczonej platformy. W szczególności musi zapewniać:
- 5.35.1. monitorowanie urządzeń pamięci masowej będącej elementem platformy,
 - 5.35.2. monitorowanie serwerów, będących elementem platformy,
 - 5.35.3. monitorowanie systemów operacyjnych i aplikacji będących elementem platformy,
 - 5.35.4. monitorowanie będącej elementem platformy infrastruktury sieciowej,
 - 5.35.5. automatyczne tworzenie map infrastruktury,
 - 5.35.6. zarządzanie alertami i incydentami
 - 5.35.7. generowanie raportów
- 5.36. Proponowane rozwiązanie musi zapewniać monitorowanie wszystkich komponentów platformy AI.
- 5.37. Proponowane rozwiązanie musi zapewniać tworzenie raportów w formatach przynajmniej XLSX, CSV i PDF.
- 5.38. Oferowane oprogramowanie wirtualizacyjne musi zapewnić wysoką dostępność maszyn wirtualnych oraz obsługiwać procesory od co najmniej 2 producentów.
- 5.39. Oferowane oprogramowanie wirtualizacyjne musi wspierać następujące systemy operacyjne jako systemy gościa w maszynach wirtualnych: RHEL, CentOS, SUSE Linux Enterprise Server, Windows Server.
- 5.40. Oferowane oprogramowanie wirtualizacyjne musi obsługiwać migrację uruchomionej maszyny wirtualnej z jednego hosta na drugi w ramach tego samego klastra bez przestojów maszyny wirtualnej.
- 5.41. Oferowane oprogramowanie wirtualizacyjne musi automatycznie uruchamiać ponownie maszyny wirtualne na innym hoście w tym samym klastrze w przypadku nieoczekiwanej awarii hosta w klastrze.

- 5.42. Oferowane oprogramowanie wirtualizacyjne musi dynamicznie planować rozmieszczenie maszyn wirtualnych w klastrze w oparciu o optymalny rozkład obciążenia w klastrze.
- 5.43. Oferowane oprogramowanie wirtualizacyjne musi obsługiwać migrację dysków wirtualnych uruchomionej maszyny wirtualnej z jednego magazynu danych pamięci masowej do innego bez przestojów.
- 5.44. Oferowane oprogramowanie wirtualizacyjne musi mieć wbudowane rozwiązanie do tworzenia kopii zapasowych danych, chroniących maszyny wirtualne.
- 5.45. Natywny mechanizm tworzenia kopii zapasowych powinien być w stanie używać co najmniej protokołów CIFS, NFS, jako protokołów dostępowych do miejsca składowania kopii zapasowych.
- 5.46. Oferowane zintegrowane oprogramowanie do tworzenia kopii zapasowych musi być tak zintegrowane z mechanizmem tworzenia i edycji instancji / maszyn wirtualnych, aby wszystkie nowo utworzone instancje / maszyny wirtualne były automatycznie chronione i archiwizowane.
- 5.47. Oferowane zintegrowane oprogramowanie do tworzenia kopii zapasowych powinno obsługiwać funkcje, takie jak planowanie tworzenia kopii zapasowych, okres przechowywania kopii zapasowych, tworzenie kopii zapasowych na żądanie itp.
- 5.48. Oferowane oprogramowanie do wirtualizacji musi umożliwiać tworzenie udziałów sieciowych z wykorzystaniem protokołów NFS i CIFS, które mogą być używane do przechowywania kopii zapasowych, archiwów, wdrożeń i obrazów wirtualnych.
- 5.49. Możliwe musi być przeglądanie, przesyłanie, pobieranie lub usuwanie plików z udziałów sieciowych. Obsługiwane muszą być różne protokoły udostępniania plików, takie jak CIFS, NFS, pamięć lokalna i wiodące w branży macierze do przechowywania plików.
- 5.50. Oferowane oprogramowanie do wirtualizacji musi obsługiwać uruchamianie maszyn wirtualnych z wykorzystaniem pamięci zewnętrznej podłączonych z wykorzystaniem protokołów iSCSI, NFS i Fibre Channel.
- 5.51. Oferowany silnik zarządzania mechanizmami wirtualizacji musi posiadać i wykorzystywać koncepcję grupowania zasobów we wspólną tożsamość, składającą się z zasobów takich jak chmury, hosty, maszyny wirtualne, sieć, pule zasobów, magazyny danych itp. w taki sposób, aby móc do tych zasobów przypisać wymaganych użytkowników.
- 5.52. Oferowany silnik zarządzania mechanizmami wirtualizacji musi pozwalać administratorowi na stworzenie planu usług lub rozmiaru maszyny wirtualnej (rozmiar według szablonu) w oparciu o procesor, pamięć i pamięć masową, które

będą następnie dostępne dla użytkowników podczas tworzenia / udostępniania instancji / maszyn wirtualnych.

- 5.53. Oferowane oprogramowanie do wirtualizacji musi posiadać wewnętrzny silnik zarządzania użytkownikami, a także integrację z zewnętrznymi dostawcami usług katalogowych.
- 5.54. Oprogramowanie musi umożliwiać odwzorowanie użytkowników zewnętrznego dostawcy integracji na role oferowane w wirtualizatorze.
- 5.55. Oferowane oprogramowanie do wirtualizacji musi mieć integrację z zewnętrznymi dostawcami DNS, aby zautomatyzować tworzenie rekordów DNS dla maszyny wirtualnej podczas procesu udostępniania
- 5.56. Oferowane oprogramowanie do wirtualizacji musi obsługiwać wykonywanie skryptów Bash i PowerShell na powołanych i wykrywanych maszynach wirtualnych w celu realizacji zadań operacyjnych.
- 5.57. Dostęp do grupowania zasobów musi być kontrolowany poprzez odpowiednie role podczas przypisywania do użytkowników. Role muszą zapewniać dostęp do zasobów przy użyciu odpowiednich uprawnień. Możliwe musi być przynajmniej skonfigurowanie następujących kluczowych uprawnień:
 - 5.57.1. Dostęp do natywnej konfiguracji ochrony danych,
 - 5.57.2. Odczyt lub pełny dostęp do tworzenia planów serwisowych,
 - 5.57.3. Dostęp do API do wykonywania skryptów na instancjach / maszynach wirtualnych,
 - 5.57.4. Dostęp umożliwiający użytkownikom korzystanie z dynamicznego harmonogramu obciążeń do umieszczania maszyn wirtualnych i przypinania maszyn wirtualnych do określonego hosta,
 - 5.57.5. Dostęp do tworzenia skryptów automatyzacji,
 - 5.57.6. Uprawnienie do zmiany rozmiaru instancji,
 - 5.57.7. Dostęp do instancji / VM – Konsola, Dodawanie lub usuwanie instancji / VM.
- 5.58. Oferowane oprogramowanie do wirtualizacji musi obsługiwać przynajmniej następujące funkcje zarządzania maszynami wirtualnymi:
 - 5.58.1. Tworzenie / Usuwanie / Ponowne uruchamianie / Uruchamianie / Zatrzymywanie / Wstrzymywanie i wykrywanie maszyn wirtualnych,
 - 5.58.2. Operacje migawkowe – tworzenie/usuwanie/przywracanie maszyn wirtualnych,
 - 5.58.3. Etykietowanie – dodawanie / usuwanie / edytowanie etykiet dla maszyn wirtualnych,
 - 5.58.4. Migracja na żywo maszyny wirtualnej, maszyny wirtualnej VM HA i przypinania do określonego hosta,
 - 5.58.5. Klonowanie maszyny wirtualnej, klonowanie do szablonu maszyny wirtualnej,

5.58.6. Zarządzanie sprzętem wirtualnym –dodawanie i usuwanie sprzętu wirtualnego, takiego jak dyski twarde, interfejsy sieciowe, procesor i pamięć z zarządzanej maszyny wirtualnej.

5.59. Oferowane oprogramowanie do wirtualizacji musi obsługiwać zarówno rozbudowę, jak i zmniejszanie klastra obsługującego maszyny wirtualne.

6. USŁUGI WDROŻENIOWE

6.1. Wymagany jest montaż w szafach rack Zamawiającego.

6.2. Konfiguracja sprzętowo-systemową dostarczonego sprzętu oraz podzespołów.

6.3. Uruchomienia dostarczonego sprzętu i weryfikację poprawności pracy wszystkich dostarczonych podzespołów.

6.4. Uruchomienie dostarczonego oprogramowania i weryfikację poprawności pracy wszystkich wymaganych w OPZ funkcjonalności.

6.5. Wykonawca wykona aktualizację oprogramowania układowego oraz mikrokodów (firmware) dostarczonego sprzętu do najnowszej, stabilnej wersji oprogramowania opublikowanego przez producenta co najmniej na dzień dostawy sprzętu.

6.6. Opracowanie i przekazanie dokumentacji powykonawczej obejmującej:

6.6.1. szczegółowy opis zrealizowanej instalacji fizycznej i logicznej dostarczonych komponentów;

6.6.2. opis konfiguracji infrastrukturalnej i przyłączy;

6.6.3. opis konfiguracji dostarczonego oprogramowania systemowego;

6.6.4. dokumentację konfiguracji środowisk;

6.6.5. procedury ciągłości działania na wypadek awarii;

6.6.6. procedury udostępniania nowych środowisk obliczeniowych;

6.6.7. dokumentację postępowania na wypadek awarii,

7. SZKOLENIA

7.1. Wykonawca zobowiązany jest do dostarczenia puli voucherów szkoleniowych na subskrypcje szkoleniowe producenta rozwiązania. Termin realizacji vouchera (rozpoczęcie subskrypcji szkoleniowej) minimum 12 miesięcy od daty podpisania Umowy.

7.2. Wymagane jest zapewnienie voucherów dla 3 administratorów Zamawiającego po 5 dni szkoleniowych dla każdego.

8. GWARANCJE I WSPARCIE

8.1. Wykonawca udziela gwarancji jakości na przedmiot zamówienia na okres 3 lat od daty podpisania Protokołu Odbioru Końcowego.

8.2. Odpowiedzialność z tytułu gwarancji jakości obejmuje wady powstałe z przyczyn tkwiących w przedmiocie zamówienia.

- 8.3. Transport, części zamienne, dostawa części zamiennych i urządzeń lub elementu nie posiadającego wad w miejsce urządzenia lub elementu wadliwego będą realizowane na koszt i ryzyko Wykonawcy.
- 8.4. Zamawiający wymaga powołania opiekuna po stronie Wykonawcy, którego zadaniem będzie koordynacja prac świadczonych na rzecz Zamawiającego. Powinien on posiadać certyfikat ITIL na poziomie minimum podstawowym lub równoważny certyfikat potwierdzający zakresem znajomość najlepszych praktyk w zakresie implementacji usług informatycznych.
- 8.5. Wymagane jest dostarczenie autoryzowanego wsparcia producenta dla wymienionego dostarczonego sprzętu i oprogramowania o następujących cechach:
- 8.5.1. Możliwość zgłaszania incydentów dotyczących sprzętu i oprogramowania - 24 godziny na dobę, przez 7 dni w tygodniu.
- 8.5.2. Gwarantowany czas naprawy jest liczony od momentu zgłoszenia problemu technicznego do serwisu i nie może przekraczać 24 godziny
- 8.5.3. Naprawa i utrzymanie Platformy AI, w szczególności zabezpieczenie poprawnego działania mechanizmów tworzenia i udostępniania zasobów obliczeniowych przez cały okres trwania Umowy.
- 8.5.4. Obsługa zgłoszeń w języku polskim.
- 8.5.5. Wykonawca zapewni Zamawiającemu możliwość aktualizacji do najnowszych wersji oprogramowania oraz mikrokodów (firmware) dla dostarczonej Platformy AI wraz ze wszystkimi jej komponentami, podzespołami, elementami i modułami w okresie gwarancji.
- 8.5.6. W ramach świadczonego serwisu, od dnia podpisania umowy i na cały okres trwania gwarancji, Wykonawca zapewni bezpośredni dostęp dla pracowników Zamawiającego do portalu internetowego producenta sprzętu i oprogramowania zawierającego narzędzia wsparcia elektronicznego. W szczególności narzędzia te muszą umożliwiać:
- 8.5.6.1. Przeszukiwanie bazy wiedzy producenta dotyczącej sprzętu i oprogramowania objętego wsparciem zawierającej wykaz znanych symptomów nieprawidłowego działania systemów oraz sposobów naprawy, jak również opisy i specyfikacje produktów oraz dokumentację techniczną.
- 8.5.6.2. Pobieranie poprawek i nowych wersji oprogramowania oraz mikrokodów (firmware) dla dostarczonego rozwiązania wraz ze wszystkimi ich komponentami, podzespołami, elementami w okresie gwarancji.

- 8.5.6.3. Uzyskiwanie informacji o statusie umowy serwisowej oraz o elementach nią objętych.
- 8.5.7. Uruchomienie i utrzymanie oprogramowania do monitorowania stanu Platformy AI umożliwiającego:
 - 8.5.7.1. Automatyczne otwieranie zgłoszeń serwisowych w autoryzowanym serwisie producenta w razie wystąpienia problemu technicznego.
 - 8.5.7.2. Powiadomienia e-mail administratorów Zamawiającego w przypadku wystąpienia problemu technicznego.
 - 8.5.7.3. Monitoring i diagnostykę oraz logowanie zdarzeń z pracy systemu.
 - 8.5.7.4. Inwentaryzację konfiguracji.
- 8.6. Wykonawca wykona minimum cztery razy w roku okresowy przegląd serwisowy Platformy AI pod kątem:
 - 8.6.1. Analizy zbiorczej incydentów i awarii za okres od ostatniego przeglądu serwisowego do obecnego przeglądu serwisowego. Pierwszy przegląd serwisowy wykonany zostanie w terminie do 3 miesięcy od daty podpisania Protokołu Odbioru Końcowego.
 - 8.6.2. Analizy dostępnych oraz zainstalowanych wersji oprogramowania systemowego i narzędziowego oraz mikrokodów (firmware) dla zainstalowanego sprzętu.
 - 8.6.3. Przegląd musi zostać przedstawiony w formie raportu zawierającego powyższe dane jak i rekomendację działań następnych.
 - 8.6.4. Po zaakceptowaniu przez Zamawiającego rekomendowanych działań następnych oraz w ustalonym z Zamawiającym terminie, Wykonawca wykona okresową aktualizację wersji oprogramowania systemowego, Platformy AI, narzędziowego oraz mikrokodów (firmware) dla zainstalowanego Rozwiązania.
- 8.7. Wykonawca wykona instalację bieżących poprawek systemowych na dostarczonym sprzęcie minimum raz na trzy miesiące. Zgodnie z harmonogramem przerw serwisowych Zamawiającego – cykl kwartalny (możliwa realizacja prac w godzinach nocnych oraz soboty i niedziele; możliwa praca zdalna z wykorzystaniem łącza VPN).
- 8.8. Wykonawca obejmie asystą techniczną prace związane z uruchomieniem i utrzymaniem produkcyjnym dostarczonej Platformy AI na okres 6 miesięcy od podpisania Protokołu Odbioru Końcowego.
- 8.9. Zamawiający wymaga od Wykonawcy utrzymania dokumentu Planu Obsługi Serwisowej zawierającego minimum:
 - 8.9.1. Inwentaryzację systemu objętego serwisem.

- 8.9.2. Specyfikację zobowiązań serwisowych.
- 8.9.3. Harmonogram planowanych prac serwisowych.
- 8.9.4. Historię zmian wersji oprogramowania systemowego i narzędziowego oraz mikrokodów (firmware) dla zainstalowanej Platformy AI.
- 8.9.5. Historię zgłoszeń serwisowych.