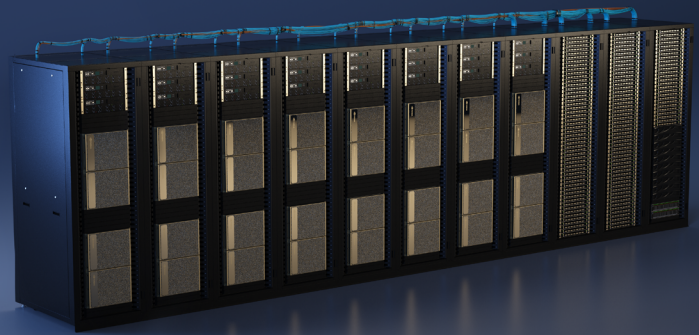




NVIDIA DGX SuperPOD

Infrastruktura fabryki AI dla współczesnych przedsiębiorstw

NVIDIA DGX SuperPOD™ to rozwiązanie infrastruktury fabryki AI, które zapewnia niezrównaną wydajność dla każdego użytkownika i obciążenia. DGX SuperPOD łączy w sobie bezkonkurencyjne doświadczenie NVIDIA w projektowaniu i używaniu superkomputerów AI oraz stanowi kompleksowe rozwiązanie infrastruktury AI, które pozwala przedsiębiorstwom na szybkie uruchomienie systemu w ciągu tygodni, a nie miesięcy. DGX SuperPOD oferuje przyspieszoną infrastrukturę klasy lidera z skalowalną wydajnością dla najbardziej wymagających zadań treningowych i wnioskowania AI, z potwierdzonymi wynikami w branży, co pozwala działom IT na dostarczanie wydajności bez kompromisów.

Kompletna infrastruktura fabryki AI

NVIDIA DGX SuperPOD charakteryzuje się zoptymalizowanym połączeniem obliczeń AI, sieci, pamięci masowej i oprogramowania. Jego fundament obliczeniowy oferuje wysoką gęstość, wydajność i elastyczność, połączony z ultra-wysoką przepustowością i niską latencją w sieci NVIDIA. Certyfikowane rozwiązania pamięci masowej integrują wydajną strukturę danych do zasilania obciążeń AI. W pełni zoptymalizowany stos oprogramowania, który działa w całym systemie, oferuje narzędzia klasy światowej dla workflow deweloperów oraz kompletną zarządzanie klastrem, aby zmniejszyć obciążenie operacyjne i maksymalizować wykorzystanie.



Architektura NVIDIA DGX SuperPOD

DGX SuperPOD

Sprzęt

- > Systemy NVIDIA DGX™ GB300, DGX GB200, DGX B300 lub DGX B200
- > Sieć NVIDIA
- > Wydajna pamięć masowa

Oprogramowanie

- > NVIDIA AI Enterprise
- > NVIDIA Mission Control
- > System operacyjny NVIDIA DGX

Usługi cyklu życia*

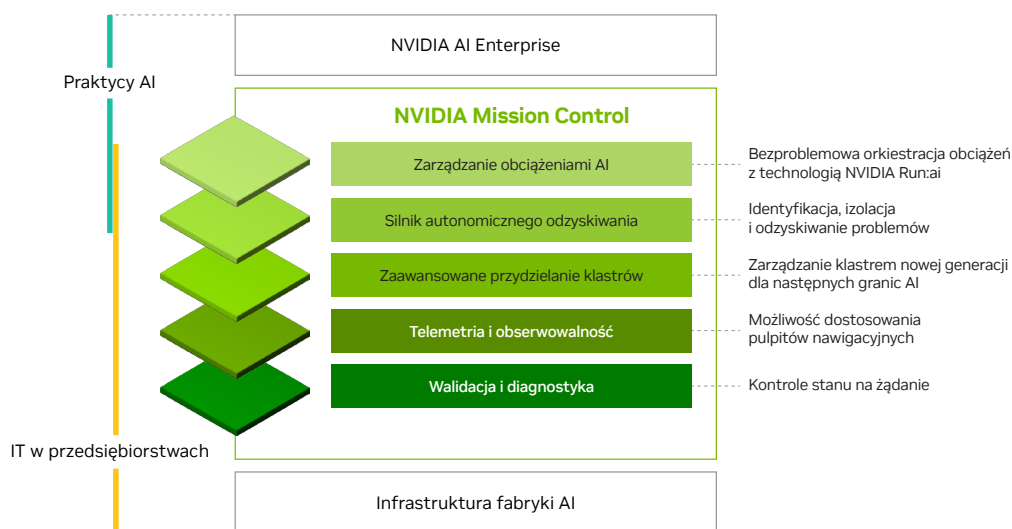
- > Planowanie/Wdrożenie **
 - Planowanie i projektowanie centrów danych/systemów
 - Przygotowanie poza siedzibą
 - Wdrożenie fizyczne i logiczne na miejscu
 - Testowanie i walidacja
- > Szkolenie/Optymalizacja
 - Testowanie wydajności aplikacji
 - Pakiet dokumentacji lokalizacji
 - Szkolenie dla użytkowników/DevOps
 - Szkolenie oparte na obciążeniach w NVIDIA Deep Learning Institute
 - Wsparcie ekspertów 24/7

* Połączenie usług NVIDIA i partnerów

** Wdrożone na miejscu lub w centrum danych DGX-Ready

Uruchamiaj modele, automatyzuj podstawowe operacje z NVIDIA Mission Control

NVIDIA Mission Control napędza każdy aspekt operacji fabryki AI, od obciążeń deweloperów po infrastrukturę i obiekty, wykorzystując umiejętności zespołu operacyjnego klasy światowej, które są teraz dostarczane jako oprogramowanie. Zapewnia natychmiastową elastyczność dla wnioskowania i treningu, jednocześnie oferując inteligencję pełnostackową dla odporności infrastruktury. Mission Control pozwala każdemu przedsiębiorstwu uruchomić AI z wydajnością na poziomie hiperskalowym, przyspieszając eksperymentowanie z AI. Dodatkowo, NVIDIA AI Enterprise, oferująca zestaw oprogramowania upraszczającego rozwój i wdrożenie AI, jest zoptymalizowana do pracy na systemach NVIDIA DGX. Użyj mikroserwisów NVIDIA NIMTM dla optymalnego rozwoju modeli, oferując szybkość, łatwość użycia, zarządzanie i bezpieczeństwo.



Oprogramowanie NVIDIA Mission Control napędza NVIDIA DGX SuperPOD

Ogromne superkomputery dla nowych granic AI

Skalując się do dziesiątek tysięcy procesorów graficznych NVIDIA, DGX SuperPOD zajmuje się treningiem i wnioskowaniem dla nowoczesnych modeli generatywnej AI. Ta platforma klasy lidera zapewnia najwyższą wydajność i skalowalność dla największych obciążeń AI, przy jednoczesnym zmniejszeniu kosztów operacyjnych i złożoności infrastruktury. W przypadku wnioskowania AI, DGX SuperPOD oferuje wiodącą na świecie platformę do wnioskowania na poziomie przedsiębiorstw, dostarczając największą domenę NVIDIA NVLink, niespotykaną pamięć całkowitą GPU oraz przepustowość pamięci między GPU, a także wydajność obliczeniową klasy lidera. W przypadku treningu AI klastry superkomputerowe oparte na architekturze DGX SuperPOD przyspieszają czas uruchamiania dla przedsiębiorstw, regularnie zajmując miejsca w czołówce list TOP500 i Green500 oraz ustanawiając rekordy benchmarków MLPerf.²



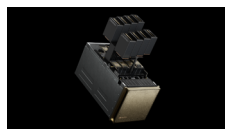
NVIDIA DGX GB300



NVIDIA DGX GB200



NVIDIA DGX B300



NVIDIA DGX B200

Kompletny cykl życia ekspertyzy, wspierany przez NVIDIA

Firmy potrzebują szybszej drogi do uczynienia infrastruktury obliczeń przyspieszonych użyteczną operacyjnie dla ich biznesów, niż tylko projekt architektury. Potrzebują doświadczenia wdrożeniowego, które jest gotowe do użycia, szybkie i zoptymalizowane pod kątem ich środowiska IT—tak aby naukowcy danych mogli zacząć działać od pierwszego dnia—i aby mogli nieustannie doskonalić się w czasie.

Dzięki NVIDIA DGX SuperPOD, firmy korzystają z pełnych usług cyklu życia i wsparcia od początku do końca, zapewnianych przez ekspertów NVIDIA. Wyznaczone wskazówki i wsparcie obejmują planowanie i instalację, walidację, zarządzanie infrastrukturą oraz cały cykl produkcyjny. NVIDIA zapewnia firmom dostęp do ekspertów pokrywających pełny stos DGX SuperPOD, aby utrzymać ich obciążenia AI w szczytowej wydajności.

Doświadczenie, które napędza sukces AI

DGX SuperPOD wykorzystuje wiedzę tysięcy badaczy i inżynierów z NVIDIA, którzy codziennie korzystają z platformy, aby wprowadzać nowe innowacje na rynek. DGX SuperPOD oferuje gotowe rozwiązanie dla fabryki AI i może być wdrożone przez klientów w ich własnym centrum danych lub z różnych dostawców usług zarządzanych, opcji kolokacyjnych lub rozwiązań chmurowych.



Superkomputer NVIDIA DGX Eos jest wykorzystywany do badań i rozwoju w NVIDIA.

Gotowy, aby zacząć?

Aby dowiedzieć się więcej o NVIDIA DGX SuperPOD, odwiedź:

nvidia.com/dgx-superpod

1. Zobacz top500.org, aby uzyskać więcej informacji.
2. Zobacz mlperf.org, aby przeczytać więcej.

© 2025 NVIDIA Corporation. Wszelkie prawa zastrzeżone. NVIDIA, logo NVIDIA, DGX, DGX SuperPOD i NIM są znakami towarowymi i/lub zarejestrowanymi znakami towarowymi NVIDIA Corporation w Stanach Zjednoczonych i innych krajach. Inne nazwy firm i produktów mogą być znakami towarowymi odpowiednich firm, z którymi są związane. Wszystkie pozostałe znaki towarowe są własnością ich odpowiednich właścicieli. 3529054. Październik 2024.

