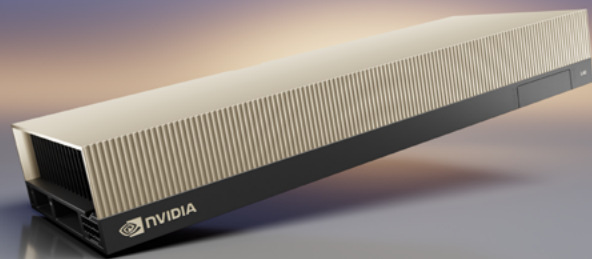




NVIDIA L40

Zapewniając bezprecedensową wydajność obliczeń wizualnych dla centrum danych.



Od aplikacji wirtualnych stacji roboczych po rozbudowane modelowanie i symulację, nowoczesne obliczenia wizualne oraz naukowe przepływy pracy rosną zarówno pod względem złożoności, jak i liczby. Przedsiębiorstwa potrzebują technologii centrum danych, która może zapewnić ekstremalną wydajność i skalować się dzięki wszechstronnym możliwościom, aby sprostać różnorodnym wymaganiom obliczeniowym coraz bardziej skomplikowanych obciążeń.

GPU NVIDIA® L40 dostarcza bezprecedensową wydajność obliczeń wizualnych dla centrum danych, oferując możliwości grafiki, obliczeń i AI nowej generacji. Zbudowana na rewolucyjnej architekturze NVIDIA Ada Lovelace, NVIDIA L40 wykorzystuje moc rdzeni RT, Tensor i CUDA® najnowszej generacji. Razem oferują przełomową wydajność wizualizacji i obliczeń dla najbardziej wymagających obciążeń w centrum danych.

Napędzany architekturą NVIDIA Ada Lovelace

Rdzenie RT trzeciej generacji

Zwiększona przepustowość oraz jednoczesne możliwości śledzenia promieni i cieniowania poprawiają wydajność śledzenia promieni, przyspieszając renderowanie dla projektów produktów, architektury oraz przepływów pracy w inżynierii i budownictwie. Zobacz realistyczne projekty w akcji dzięki sprzętowi przyspieszeniu rozmycia ruchu, które dostarcza oszałamiające animacje w czasie rzeczywistym.

Rdzenie Tensor czwartej generacji

Sprzętowe wsparcie dla strukturalnej rzadkości oraz zoptymalizowany format TF32 zapewniają natychmiastowe zyski wydajności dla szybszego szkolenia modeli AI i nauki o danych. Przyspiesz możliwości grafiki wspieranej AI, w tym DLSS, dostarczając powiększoną rozdzielczość z lepszą wydajnością w wybranych aplikacjach.

Duża pamięć GPU

Radź sobie z aplikacjami i obciążeniami wymagającymi dużej pamięci, takimi jak nauka o danych, symulacje, modelowanie 3D i renderowanie z 48 GB ultraszybkiej pamięci GDDR6. Przydziel pamięć wielu użytkownikom za pomocą oprogramowania vGPU, aby rozdzielić duże obciążenia między zespoły kreatywne, nauki o danych i projektowe.

Gotowe do Centrum Danych

Zaprojektowane do pracy w centrum danych 24/7 z energooszczędnym sprzętem i komponentami, NVIDIA L40 jest zoptymalizowana do wdrożenia na dużą skalę i oferuje maksymalną wydajność dla różnorodnych obciążeń centrum danych. L40 obejmuje bezpieczny rozruch z technologią podstawy zaufania, co zapewnia dodatkową warstwę bezpieczeństwa, oraz jest zgodna z poziomem NEBS Level 3, aby spełniać najnowsze standardy centrum danych. Zapakowana w konstrukcję o podwójnej szerokości, pasywnie chłodzoną i efektywną energetycznie, L40 jest dostępna w szerokiej gamie systemów certyfikowanych przez NVIDIA™ od wiodących dostawców OEM.

Przyspieszanie Obciążeń Nowej Generacji

- > NVIDIA Omniverse™ Enterprise
- > Renderowanie i Grafika 3D
- > Generatywna AI i Wnioskowanie
- > Wysokowydajne Wirtualne Stacje Robocze z Oprogramowaniem NVIDIA RTX™ Virtual Workstation (RTX vWS)
- > Szkolenie AI i Nauka o Danych
- > Strumieniowanie i Treści Wideo

Specyfikacja

SPECYFIKACJE PRODUKTU

Całkowite zużycie energii	300 W domyślnie 300 W maksymalnie
Rozwiązanie termiczne	Pasywne
Mechaniczny format obudowy	Pełny profil, pełna długość (FHFL) 10,5", dwuslotowy
Taktowanie GPU	Bazowe: TBD MHz (do ustalenia) Boost: TBD MHz (do ustalenia)
Stany wydajności	P0
VBIOS	Rozmiar pamięci EEPROM: 8 Mbit UEFI: Obsługiwane
Interfejs PCI Express	PCI Express Gen4 x16 Obsługiwane odwracanie linii i polaryzacji
Wieloinstancyjny GPU (MIG)	Nie obsługiwane
NVIDIA® NVLink®	Nie obsługiwane
Zero Power	Nie obsługiwane
Złącza	Jedno dodatkowe złącze zasilania PCIe 16-pin Cztery złącza DisplayPort
Waga	Płyta: 1051 g (bez wspornika i przedłużaczy) Wspornik z wkrętami: 20 g Ulepszony prosty przedłużacz: 35 g Długi przedłużacz offsetowy: 48 g Prosty przedłużacz: 32 g

SPECYFIKACJE PAMIĘCI

Taktowanie pamięci	9001 MHz
Typ pamięci	GDDR6
Rozmiar pamięci	48 GB
Szerokość magistrali pamięci	384 bits
Maksymalna przepustowość pamięci	864 GB/s

SPECYFIKACJE OPROGRAMOWANIA

Obsługa SR-IOV	Obsługiwane: 32 VF (funkcji wirtualnych)
Adres BAR (fizyczna funkcja)	BAR0: 16 MiB BAR1: 64 GiB (tryb wyłączonego wyświetlacza; domyślny) BAR1: 8 GiB (tryb włączonego wyświetlacza, 8 GB BAR1) BAR1: 256 MiB (tryb włączonego wyświetlacza, 256 MB BAR1) BAR3: 32 MiB
Adres BAR (funkcja wirtualna)	Tryb wyłączonego wyświetlacza (domyślny): • BAR0: 8 MiB (32 VF × 256 KiB) • BAR1: 64 GiB, 64-bit (32 VF × 2 GiB) • BAR3: 1 GiB, 64-bit (32 VF × 32 MiB) Tryb włączonego wyświetlacza: Rozmiary BAR VF nie mają zastosowania w trybach włączonego wyświetlacza.
Przerwania sygnalizowane poprzez wiadomość (MSI)	MSI-X: Obsługiwane MSI: Nieobsługiwane
Przekazywanie ARI	Obsługiwane

Wsparcie sterownika	Linux: R525 lub nowszy Windows: R525 lub nowszy
Uruchamianie zabezpieczeń	Obsługiwane
Firmware CEC	Nieobsługiwane
Wsparcie dla NVIDIA® CUDA®	CUDA 12.0 lub nowsza
Wsparcie oprogramowania Virtual GPU	Obsługuje vGPU 15.1 lub nowszy: NVIDIA Virtual Compute Server Edition
Tryby pracy	Tryb wyłączonego wyświetlacza (domyślny) Tryb włączonego wyświetlacza, 8 GiB BAR1 Tryb włączonego wyświetlacza, 256 MiB BAR1
Kod klasy PCI	0x03 – Kontroler wyświetlacza
Kod podklasy PCI	0x02 – Kontroler 3D (tryb wyłączonego wyświetlacza) 0x00 – Kontroler VGA (tryb włączonego wyświetlacza)
Zdolność podstawowego urządzenia rozruchowego	Obsługiwane w trybach włączonego wyświetlacza
Wsparcie ECC	Włączone (domyślnie); można wyłączyć za pomocą oprogramowania
SMBus (adres 8-bitowy)	0x9E (zapis), 0x9F (odczyt)
Adres I2C pamięci EEPROM	0x50 (7-bit), 0xA0 (8-bit)
FRU IPMI	
Zarezerwowane adresy I2C	0xAA, 0xAC
Bezpośredni dostęp do SMBus	Obsługiwany
Interfejs SMBPBI (SMBus Post-Box Interface)	Obsługiwany

Uwaga:
¹Notacja KiB, MiB i GiB podkreśla "potęgę dwóch" tych wartości. Zatem,
• 256 KiB = 256 × 1024
• 16 MiB = 16 × 1024²
• 64 GiB = 64 × 1024³

Temperatura pracy otoczenia	0 °C do 50 °C
Temperatura pracy otoczenia (krótkoterminowa)¹	-5 °C do 55 °C
Temperatura przechowywania	-40 °C do 75 °C
Wilgotność robocza (krótkoterminowa)¹	5% do 93% wilgotności względnej
Wilgotność robocza	5% do 85% wilgotności względnej
Wilgotność przechowywania	5% do 95% wilgotności względnej
Średni czas między awariami (MTBF)	Nieuregulowane środowisko: ² 2 502 369 godzin w temperaturze 35 °C Kontrolowane środowisko: ³ 3 270 359 godzin w temperaturze 35 °C

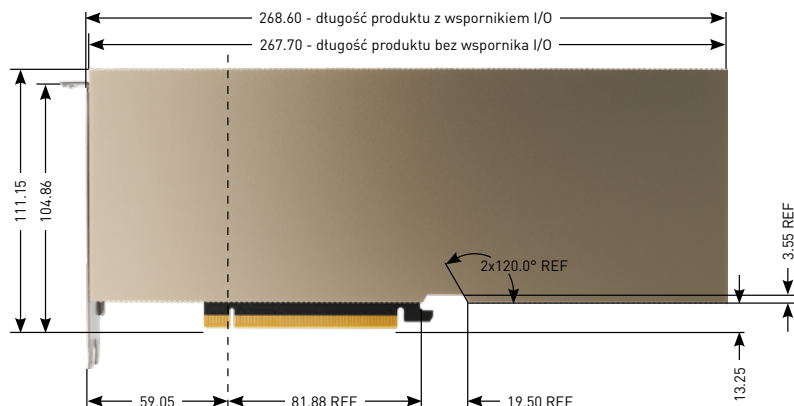
Notatki:

Specyfikacje w tej tabeli dotyczą wysokości do 6000 stóp.

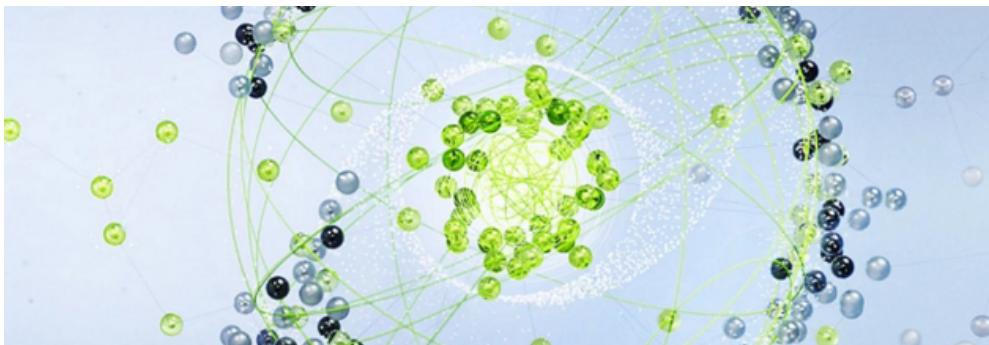
¹ Okres nie dłuższy niż 96 godzin ciągłych, nie więcej niż 15 dni w roku.

² Pewne obciążenie środowiskowe z ograniczoną konserwacją (GF35).

³ Brak obciążenia środowiskowego z optymalną eksploatacją i konserwacją (GB35).



NVIDIA Omniverse i generatywna AI



Procesor graficzny NVIDIA L40 to potężne rozwiązanie zaprojektowane z myślą o doskonaleniu obsługi NVIDIA Omniverse i generatywnych aplikacji AI. Zbudowany na architekturze Ada Lovelace, zawiera rdzenie RT trzeciej generacji i rdzenie Tensor czwartej generacji, dzięki czemu idealnie nadaje się do śledzenia promieni w czasie rzeczywistym i przepływów pracy przyspieszanych przez sztuczną inteligencję. Jako silnik NVIDIA Omniverse w centrum danych, L40 zapewnia solidne możliwości RTX i AI do obsługi rozszerzonej rzeczywistości (XR) i wirtualnych obciążeń produkcyjnych. W przypadku generatywnych zadań AI, L40 zapewnia wyjątkową wydajność, oferując do 5 razy wyższe prędkości wnioskowania w porównaniu do swojego poprzednika, NVIDIA A40. Dzięki 48 GB pamięci GDDR6 zapewnia wystarczającą pojemność do obsługi złożonych generatywnych modeli AI i wielkoskalowych symulacji 3D. Wszechstronność L40 obejmuje obsługę procesów 3D opartych na uniwersalnym opisie scen (OpenUSD) i zadań symulacyjnych w ekosystemie Omniverse. To połączenie wysokowydajnej grafiki i akceleracji AI sprawia, że NVIDIA L40 jest doskonałym wyborem dla organizacji, które chcą wykorzystać najnowocześniejsze technologie w tworzeniu wirtualnych światów, generowaniu treści opartych na sztucznej inteligencji i wciągających doświadczeniach wizualnych.

Grafika 3D i wizualizacja



Procesor graficzny NVIDIA L40 to potężny procesor graficzny 3D i wizualizacji, oferujący niespotykaną wydajność w środowiskach centrów danych. Zbudowany w oparciu o zaawansowaną architekturę Ada Lovelace, zawiera rdzenie RT trzeciej generacji, które znacznie zwiększają możliwości śledzenia promieni w czasie rzeczywistym, zapewniając nawet dwukrotnie większą wydajność w porównaniu do swojego poprzednika. Dzięki 48 GB ultraszybkiej pamięci GDDR6, L40 z łatwością radzi sobie ze złożonymi modelami 3D, teksturami o wysokiej rozdzielczości i symulacjami na dużą skalę. Tak duża pojemność pamięci pozwala kreatywnym profesjonalistom pracować nad skomplikowanymi projektami i wydajnie renderować fotorealistyczne sceny. Rdzenie Tensor czwartej generacji zastosowane w L40 przyspieszają funkcje graficzne wzmocnione sztuczną inteligencją, takie jak DLSS (Deep Learning Super Sampling), umożliwiając uzyskanie wyższej rozdzielczości wyjściowej i lepszą wydajność w obsługiwanych aplikacjach. W przypadku profesjonalnych zastosowań wizualizacyjnych, w tym projektowania wspomaganego komputerowo (CAD) i produkcji wirtualnej, L40 umożliwia interaktywne renderowanie z pełną wiernością w czasie rzeczywistym, umożliwiając artystom i projektantom szybsze iterowanie i szybsze ukończenie projektów. W połączeniu z oprogramowaniem NVIDIA RTX Virtual Workstation (vWS), L40 może dostarczać potężne wirtualne stacje robocze z centrum danych lub chmury, zapewniając elastyczność i wysoką wydajność dostępu do wymagających aplikacji z dowolnego urządzenia.

Wirtualne stacje robocze i zdalna współpraca



Procesor graficzny NVIDIA L40 to potężne rozwiązanie dla wirtualnych stacji roboczych i zdalnej współpracy, oferujące wyjątkową wydajność i wszechstronność w środowiskach centrów danych. W połączeniu z oprogramowaniem NVIDIA RTX Virtual Workstation (vWS), L40 zapewnia wysoką wydajność wirtualnych stacji roboczych z centrum danych lub chmury na dowolne urządzenie, umożliwiając kreatywnym i technicznym profesjonalistom zdalny dostęp do wymagających aplikacji za pomocą wydajności dorównującej fizycznym stacjom roboczym. 48 GB pamięci GDDR6 pozwala na obsługę złożonych modeli 3D, tekstur o wysokiej rozdzielczości i dużych zbiorów danych, dzięki czemu idealnie nadaje się do zadań takich jak renderowanie 3D, edycja wideo i projektowanie CAD. Zaawansowane możliwości graficzne laptopa L40, obsługiwane przez rdzenie RT trzeciej generacji i rdzenie Tensor czwartej generacji, umożliwiają śledzenie promieni w czasie rzeczywistym i funkcje wzmocnione sztuczną inteligencją, znacznie poprawiając jakość i szybkość pracy w zakresie zdalnej wizualizacji. W przypadku projektów opartych na współpracy, L40 doskonale obsługuje aplikacje NVIDIA Omniverse Enterprise, ułatwiając korzystanie z rzeczywistości rozszerzonej (XR) i rzeczywistości wirtualnej (VR), współpracę projektową i tworzenie cyfrowych bliźniaków. To sprawia, że NVIDIA L40 jest doskonałym wyborem dla organizacji, które chcą zwiększyć swoje możliwości pracy zdalnej, usprawnić współpracę i utrzymać wysoką produktywność rozproszonych zespołów w takich dziedzinach, jak architektura, produkcja oraz media i rozrywka.

Szkolenie i wnioskowanie AI (sztucznej inteligencji)



Procesor graficzny NVIDIA L40 to potężne rozwiązanie do obciążeń związanych ze szkoleniem AI i wnioskowaniem, oferujące wyjątkową wydajność i wszechstronność. Zbudowany na architekturze Ada Lovelace, wyposażony w rdzenie Tensor czwartej generacji, które znacznie przyspieszają obliczenia AI. L40 zapewnia do 5 razy wyższe prędkości wnioskowania w porównaniu do swojego poprzednika, NVIDIA A40, dzięki czemu szczególnie dobrze nadaje się do wnioskowania z dużym modelem językowym (LLM) i generatywnych zadań AI. Dzięki 48 GB pamięci GDDR6 zapewnia wystarczającą pojemność do obsługi złożonych modeli sztucznej inteligencji i dużych zbiorów danych, umożliwiając wydajne szkolenie i wnioskowanie w zaawansowanych sieciach neuronowych. Zaawansowane możliwości sztucznej inteligencji L40 sprawiają, że idealnie nadaje się do szerokiego zakresu zastosowań, w tym do przetwarzania języka naturalnego, wizji komputerowej i systemów rekomendacji. Jego wydajność zarówno w scenariuszach szkoleniowych, jak i programistycznych dotyczących sztucznej inteligencji z jednym procesorem graficznym sprawia, że jest to wszechstronna opcja dla organizacji chcących przyspieszyć przepływ pracy w zakresie sztucznej inteligencji. W połączeniu z kompleksowym pakietem oprogramowania NVIDIA, w tym NVIDIA AI Enterprise, L40 oferuje solidną platformę do wdrażania i skalowania aplikacji AI w środowiskach centrów danych, zapewniając moc obliczeniową potrzebną do napędzania innowacji w badaniach i rozwoju AI.

Kodowanie wideo i strumieniowanie



Procesor graficzny NVIDIA L40 oferuje wyjątkowe możliwości w zakresie kodowania wideo i przesyłania strumieniowego, co czyni go potężnym rozwiązaniem do strumieniowego przesyłania transmisji, produkcji wideo i transkrypcji. Zbudowany na architekturze Ada Lovelace, L40 posiada trzy silniki kodowania i dekodowania wideo, znacznie zwiększające jego zdolność do jednoczesnej obsługi wielu strumieni wideo wysokiej jakości. Kluczowym postępem jest dodanie obsługi kodowania i dekodowania AV1, która zapewnia przełomową wydajność i niższy całkowity koszt posiadania dla twórców treści i platform przesyłania strumieniowego. Ta funkcja pozwala na uzyskanie wyższej jakości wideo przy niższych przepływnościach, z korzyścią zarówno dla dostawców treści, jak i użytkowników końcowych. Potężne przyspieszenie sprzętowe L40 w połączeniu z 48 GB pamięci GDDR6 umożliwia mu wydajne przetwarzanie złożonych obciążeń wideo, w tym transkodowanie w czasie rzeczywistym i tworzenie treści w wysokiej rozdzielczości. Organizacjom zajmującym się transmisją strumieniową na żywo, usługami wideo na żądanie lub produkcją wirtualną, NVIDIA L40 zapewnia wydajność i wszechstronność niezbędną do spełnienia wymagań współczesnego tworzenia i dystrybucji treści wideo.

Gotowy, aby zacząć?

Aby dowiedzieć się więcej o GPU NVIDIA L40, odwiedź:

www.nvidia.com/L40

© 2023 NVIDIA Corporation and affiliates. All rights reserved. NVIDIA, the NVIDIA logo, NVIDIA-Certified Systems, NVIDIA Omniverse and NVIDIA RTX are trademarks and/or registered trademarks of NVIDIA Corporation and affiliates in the U.S. and other countries. Other company and product names may be trademarks of the respective owners with which they are associated. 2436245. JUL24

