



NVIDIA Grace CPU Superchip

Pierwszy na świecie CPU centrum danych bez kompromisów

Najwydajniejszy CPU centrum danych na świecie

Zaprojektowany do wymagających aplikacji, które działają na tradycyjnych CPU, NVIDIA Grace™ to pierwszy CPU bez kompromisów dla chmury, przedsiębiorstw i obliczeń wysokiej wydajności (HPC). Charakteryzuje się dwukrotną wydajnością na wat w porównaniu do najnowszych platform x86, zapewniając najlepszy w swojej klasie przepustowość obliczeń w centrum danych. Grace oferuje również TCO dla wymagających, zoptymalizowanych pod kątem obliczeń chmur, HPC, superkomputerów, infrastruktury korporacyjnej, analityki danych, platform brzegowych i innych zastosowań.

Elastyczny projekt platformy

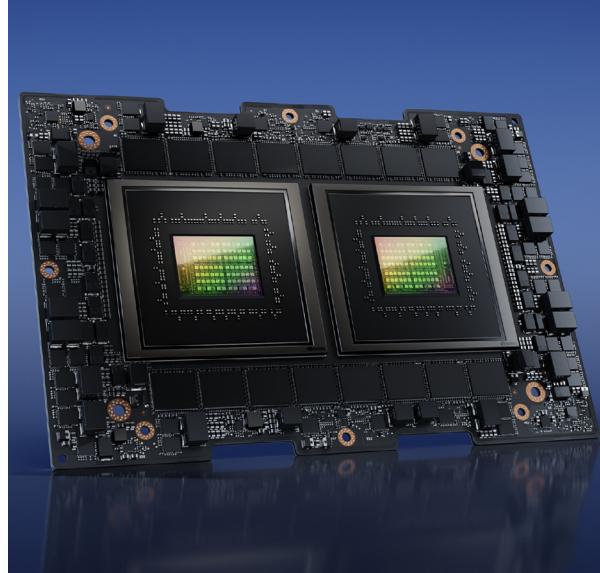
Portfolio NVIDIA Grace obejmuje NVIDIA Grace CPU Superchip, który stanowi serce serwera z dwoma gniazdami, z 144 rdzeniami Arm® Neoverse™ V2 i do 960 GB pamięci LPDDR5X w jednym kompaktowym module, wymagającym jedynie 500 W dla CPU i pamięci — w porównaniu do prawie 900 W w konwencjonalnych serwerach x86.

Dla większej elastyczności, Grace CPU C1 zapewnia wyjątkową wydajność w konfiguracji single-socket, zoptymalizowanej pod kątem chmur, przechowywania, brzegów, telekomunikacji i innych, oferując do 2x większą wydajność na wat w porównaniu do konwencjonalnych serwerów x86.

Innowacyjność i architektura

NVIDIA Grace CPU upraszcza projektowanie energooszczędnych systemów, zapewniając większą wydajność na skalę centrum danych. Łączy 72 rdzenie Arm Neoverse V2 o wysokiej wydajności i efektywności energetycznej, połączone z NVIDIA Scalable Coherency Fabric (SCF), które zapewnia 3,2 TB/s przepustowości bisection — dwukrotnie więcej niż tradycyjne CPU. Ta architektura zapewnia płynny przepływ danych między rdzeniami CPU, pamięcią cache, pamięcią RAM oraz wejściem/wyjściem systemu, maksymalizując wydajność systemu.

Dla poprawy niezawodności, Grace używa pierwszej na rynku pamięci LPDDR5X klasy serwerowej o szerokim podsystemie, zapewniającym do 500 GB/s przepustowości przy zaledwie około 16 W, co stanowi jedną piątą mocy konwencjonalnej pamięci DDR5. Pamięć ta korzysta z nowatorskich technik provisioning i wykrywania błędów, eliminując konieczność serwisowania lub wymiany uszkodzonych modułów w terenie, co znacząco podnosi niezawodność i umożliwia wdrożenie w trudnych lub kosztownych scenariuszach.



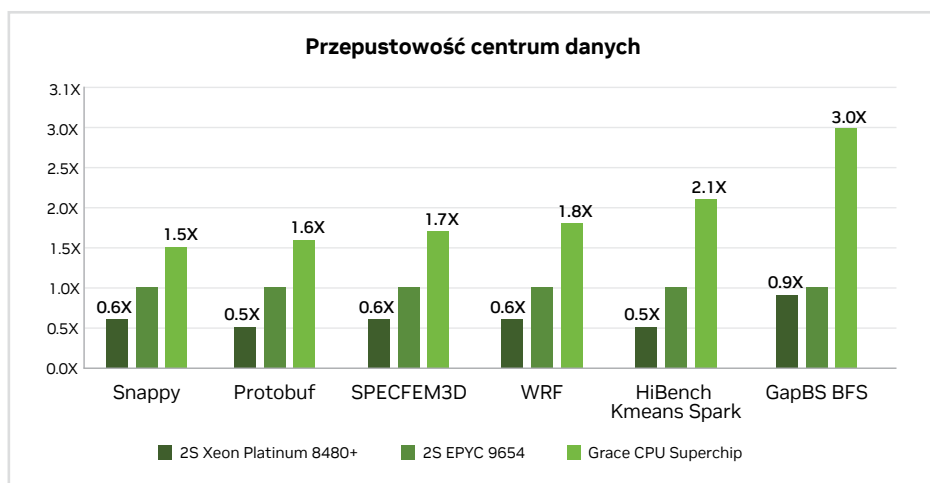
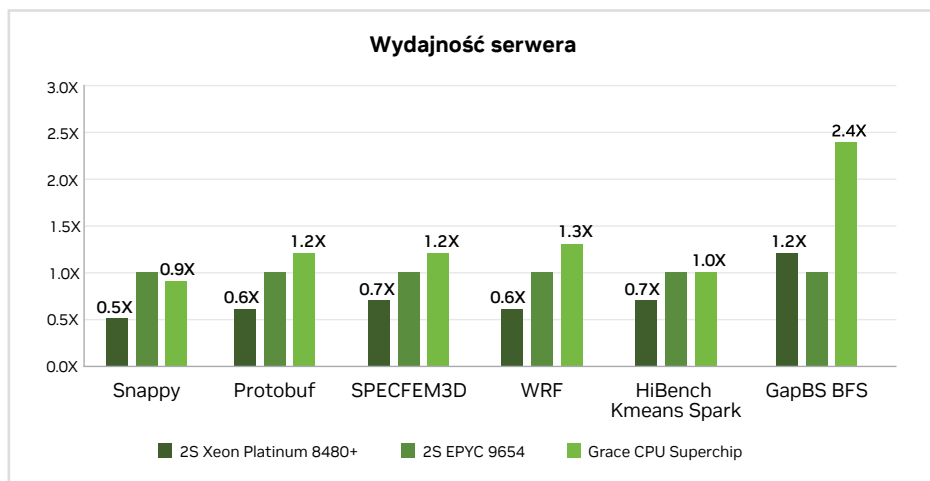
Kluczowe cechy

- > 72 dla NVIDIA Grace CPU C1 lub 144 wysokowydajne dla NVIDIA Grace CPU Superchip
Rdzenie Arm Neoverse V2 z 4x128b SVE2
- > Wysokowydajny NVIDIA Scalable Coherency Fabric z przepustowością 3,2 TB/s (bisection bandwidth)
- > Do 480 GB pamięci LPDDR5X z kodem korekcji błędów (ECC), z maksymalną przepustowością do 512 GB/s na jeden CPU Grace.
- > 900GB/s NVLink®-C2C
- > Do 250W dla NVIDIA Grace CPU C1 lub do 500W dla NVIDIA Grace CPU Superchip (CPU + pamięć).

Wydajność i efektywność energetyczna CPU nowej generacji dla centrum danych

Procesor Grace zapewnia znakomitą wydajność na rdzeń, przepustowość pamięci i możliwości przesyłu danych, osiągając wiodącą wydajność na wątek, co przekłada się na generacyjne zyski w zakresie energooszczędnych obliczeń CPU dla centrów danych. Procesor NVIDIA Grace Superchip oferuje wszechstronność i wydajność dla podstawowych obciążeń centrum danych, takich jak mikroserwisy, analityka danych, analityka grafowa oraz symulacje.

2x więcej przepustowości dla centrum danych w chmurze i aplikacjach HPC¹



Nowy standard dla infrastruktury oprogramowania

Procesor NVIDIA Grace podąża za głównymi zasadami projektowania CPU, jest programowany tak samo jak inne serwerowe CPU i jest wspierany przez pełny ekosystem NVIDIA. Architektura NVIDIA Grace jest w pełni kompatybilna z ekosystemem ARM, co gwarantuje, że każde oprogramowanie zaprojektowane dla 64-bitowego ARM w centrach danych będzie działać bezproblemowo na Grace,

¹ NVIDIA Grace Superchip 480GB LPDDR5X, AMD EPYC 9654 768GB DDR5, i Intel Xeon Platinum 8480+ z 1TB DDR5. OS: Ubuntu 22.04. Kompilatory: GCC 12.3, chyba że zaznaczono inaczej. Zużycie energii, dla efektywności, obejmuje CPU + pamięć mierzoną mocą.

Kompresja: Snappy (Commit af720f9a3b2c831f173b6074961737516f2d3a46 | N instancji równocześnie); Mikrosługi: Google Protobufs (Commit 7cd0b6fbf1643943560d8a9fe553fd206190b27f | N instancji równocześnie); Przetwarzanie danych sejsmicznych: SPECFEM3D four_material_simple_model; HPC SDK 24.3; CFD: OpenFOAM; Motocykl | Large v2212 MD: CP2K RPA 2023.1; Pogoda: WRF CONUS12km | x86: ICC 2024.01; Analityka danych: HiBench+K-means Spark (HiBench 7.1.1, Hadoop 3.3.3, Spark 3.3.0; Grace: NVHPC 24.5, x86: Intel 2021.4); Analityka grafów: The Gap Benchmarks Suite BFS arXiv:1508.03619 [cs.DC], 2015

i odwrotnie. Wszystkie główne dystrybucje Linux oraz ogromne kolekcje dostępnego oprogramowania działają w pełni i bez modyfikacji na procesorze NVIDIA Grace CPU.

Aby deweloperzy mogli szybko rozpocząć pracę, NVIDIA Grace CPU Superchip jest wspierany przez pełny stos oprogramowania NVIDIA, w tym NVIDIA HPC, NVIDIA AI i NVIDIA Omniverse™.

Specyfikacje produktu

	NVIDIA Grace CPU C1	NVIDIA Grace CPU Superchip
Konfiguracja	1x CPU Grace	2x CPU Grace
Liczba rdzeni	72 rdzenie Arm Neoverse V2 z 4x 128b SVE2	144 rdzenie Arm Neoverse V2 z 4x 128b SVE2
Pamięć L1	64 KB i-cache + 64 KB d-cache na rdzeń	64 KB i-cache + 64 KB d-cache na rdzeń
Pamięć L2	1 MB na rdzeń	1 MB na rdzeń
Pamięć L3	114 MB	228 MB
Wielkość pamięci LPDDR5X	Opcje pamięci na module: 120 GB, 240 GB i 480 GB	Opcje pamięci: 240 GB, 480 GB i 960 GB
Przepustowość pamięci	Do 384 GB/s dla 480 GB	Do 512 GB/s dla 120 GB, 240 GB
Przepustowość NVLink-C2C	Nie dotyczy	Do 900 GB/s
Łącza PCIe	Do 4x PCIe Gen 5x16 z opcją rozdzielenia (bifurcation)	Do 8x PCIe Gen 5x16 z opcją rozdzielenia (bifurcation)

Gotowy, aby zacząć?

Aby dowiedzieć się więcej o NVIDIA Grace CPU Superchip, odwiedź nvidia.com/grace-cpu

Aby uzyskać informacje o pakietach wsparcia NVIDIA, prosimy o kontakt z naszym przedstawicielem handlowym NVIDIA [FORMAT.COM.PL](https://format.com.pl)

© 2024 NVIDIA Corporation i jej spółki zależne. Wszelkie prawa zastrzeżone. NVIDIA, logo NVIDIA, Grace, NVLink oraz Omniverse są znakami towarowymi i/lub zarejestrowanymi znakami towarowymi NVIDIA Corporation i jej spółek zależnych w Stanach Zjednoczonych i innych krajach. Inne nazwy firm i produktów mogą być znakami towarowymi odpowiednich właścicieli, z którymi są związane. Wszystkie inne znaki towarowe są własnością ich odpowiednich właścicieli. 3462416. PAŹDZIERNIK 2024.

