

ВИКИПЕДИЯ

# GH200 Grace Hopper

Материал из Википедии — свободной энциклопедии

**NVIDIA GH200 Grace Hopper Superchip** — микропроцессорная сборка (гибридный чип) класса *Server/Datacenter* от корпорации *NVIDIA Corporation* из серии *NVIDIA Grace*, модель середины 2023 года. В состав системы входит 64-битный 72-ядерный ARM-микропроцессор и ускоритель GPGPU-вычислений с тензорными ядрами H100 (Hopper). Разработана Nvidia и производится контрактным производителем TSMC по усовершенствованному 4-нанометровому (N4) техпроцессу<sup>[1]</sup>.

## Содержание

История

Описание

Применение

См. также

Примечания

Ссылки

## История

В марте 2022 года компания Nvidia в ходе своей корпоративной технологической конференции *GTC 2022* анонсировала начало выпуска в первой половине 2023 года давно обещанных ЦПУ Grace и гибридного чипа *Grace Hopper*, — сочетающего процессоры *Grace (ARMv9)* и ускоритель H100 (Hopper)<sup>[2][3]</sup>.

### **NVIDIA GH200 Grace Hopper Superchip**

*Центральный процессор*

<b>Производство</b>	<u>2023 год</u>
<b>Разработчик</b>	<u>NVIDIA Corporation</u>
<b>Производитель</b>	<u>TSMC</u>
<b>Технология производства</b>	4 (N4) <u>нм</u>
<b>Наборы инструкций</b>	<u>Neoverse V2 Cores: Armv9 with 4x128b SVE2</u>
<b>Число ядер</b>	72 ядра
<b>L1-кэш</b>	64KB I-cache + 64KB D-cache per core
<b>L2-кэш</b>	1MB per core
<b>L3-кэш</b>	234MB per superchip
<b>Встроенный графический процессор</b>	Проприетарный Nvidia <u>GPGPU-ускоритель (microarchitecture: Ada Lovelace)</u>
<b>Разъём</b>	
<b>Ядра</b>	

Затем в августе 2022 года компания Nvidia поделилась некоторыми деталями о строении Arm-процессоров Grace и гибридных чипов Grace Hopper<sup>[4]</sup>.

## Описание

---

В мае 2023 года компания Nvidia в ходе выставки Computex 2023 сообщила о начале серийного производства гибридных суперчипов GH200 Grace Hopper, предназначенных для построения НРС-систем и платформ генеративного ИИ<sup>[1]</sup>.

В состав Grace Hopper входят 72-ядерный Arm-процессор Nvidia Grace и ускоритель GPGPU-вычислений H100 (Hopper) с тензорными ядрами для систем искусственного интеллекта с 96 Гбайт НВМ3 памяти, содержащий порядка 80 млрд транзисторов<sup>[5]</sup>. Объём общей для обоих кристаллов памяти составляет 576 Гбайт (480 Гбайт LPDDR5x). Кристаллы соединены между собой шиной NVLink-C2C, обеспечивающей пропускную способность 900 Гбайт/с (в 7 раз быстрее по сравнению с PCIe 5.0). Заявленный уровень производительности GH200 — 4 ПФлопс с использованием Transformer Engine, а реальное потребление GH200 под полной нагрузкой составляет 750—800 Вт<sup>[6][7]</sup>.

Ранее, в июне 2023 года Nvidia уже сообщила, что во всех восьми ИИ-бенчмарках MLPerf Training v3.0 её ускорители H100 установили новые рекорды<sup>[8]</sup>. Теперь, в сентябре 2023 года компания говорит, что её суперчип GH200 Grace Hopper впервые прошёл все тесты MLPerf, и примечательно, что GH200 оказался до 17 % быстрее H100, хотя чип самого ускорителя в обоих продуктах один и тот же. Nvidia объясняет это несколькими факторами: Во-первых, у GH200 больше наборной памяти — 96 Гбайт против 80 Гбайт. Во-вторых, ПСП составляет 4 Тбайт/с, а сам чип является гибридным, так что для передачи данных между LPDDR5x и НВМ3 не используется PCIe<sup>[9]</sup>.

Nvidia также объявила, что в конце 2023 года корпоративным заказчикам станет доступна новая НРС-платформа Nvidia DGX GH200 AI Supercomputer, в которой используются 256 суперчипов GH200 Grace Hopper, объединённых при помощи технологии NVLink Switch System позволяющей всем ускорителям H100 в составе системы функционировать в качестве единого целого. Производительность этой платформы обещана на уровне 1 Эфлопс (~ 9 Пфлопс FP64), а суммарный объём памяти достигает 144 Тбайт — и это почти в 500 раз больше, чем в одной системе Nvidia DGX A100<sup>[10][11]</sup>.

## Применение

---

- **NVIDIA HGX для НРС** — с IV квартала 2023 года<sup>[6]</sup>.

- [NVIDIA DGX GH200 AI Supercomputer](#) ([HPC-платформа](#) объединяющая 256 суперчипов GH200 Grace Hopper) — с IV квартала [2023 года](#)<sup>[10]</sup>.
- [NVIDIA DGX Helios](#) (суперкомпьютер который с помощью [коммутируемой сети](#) Quantum-2 [InfiniBand](#) объединит сразу четыре системы DGX GH200 AI) — с IV квартала [2023 года](#)<sup>[10]</sup>.

## См. также

---

- [Nvidia Grace](#)
- [Grace Superchip](#)

## Примечания

---

1. [NVIDIA начала массовое производство суперчипов Grace Hopper для генеративного ИИ](#) (<https://servernews.ru/1087470>). ServerNews. (29 мая 2023). Дата обращения: 18 сентября 2023. Архивировано (<https://web.archive.org/web/20231003193827/http://servernews.ru/1087470>) 3 октября 2023 года.
2. [NVIDIA анонсировала 144-ядерные Arm-процессоры Grace и гибриды Grace Hopper](#) (<https://servernews.ru/1062436>). ServerNews. (22 марта 2022). Дата обращения: 18 сентября 2023. Архивировано (<https://web.archive.org/web/20231003133704/https://servernews.ru/1062436>) 3 октября 2023 года.
3. [NVIDIA представила 144-ядерный серверный процессор Grace Superchip](#) (<https://3dnews.ru/1062503/nvidia-predstavila-144yaderniy-serverniy-protssessorgrace-superchip>). 3DNews. (22 марта 2022). Дата обращения: 18 сентября 2023. Архивировано (<https://web.archive.org/web/20231125192207/https://3dnews.ru/1062503/nvidia-predstavila-144yaderniy-serverniy-protssessorgrace-superchip>) 25 ноября 2023 года.
4. [NVIDIA поделилась некоторыми деталями о строении Arm-процессоров Grace и гибридных чипов Grace Hopper](#) (<https://servernews.ru/1072574>). ServerNews. (20 августа 2022). Дата обращения: 18 сентября 2023. Архивировано (<https://web.archive.org/web/20230913075424/https://servernews.ru/1072574>) 13 сентября 2023 года.
5. [Представлен ускоритель вычислений NVIDIA H100 на новейшей архитектуре Hopper и с памятью HBM3](#) (<https://3dnews.ru/1062506/publikatsiya-1062506>). 3DNews. (22 марта 2022). Дата обращения: 18 сентября 2023. Архивировано (<https://web.archive.org/web/20231125192152/https://3dnews.ru/1062506/publikatsiya-1062506>) 25 ноября 2023 года.
6. [Суперчип NVIDIA Grace Hopper](#) (<https://www.nvidia.com/ru-ru/data-center/grace-hopper-superchip/>). Официальный сайт NVIDIA Corporation. Дата обращения: 18 сентября 2023. Архивировано (<https://web.archive.org/web/20230919190132/https://www.nvidia.com/ru-ru/data-center/grace-hopper-superchip/>) 19 сентября 2023 года.
7. [NVIDIA представила суперчип GH200 Grace Hopper с памятью HBM3е и производительностью 4 Пфлопс](#) (<https://3dnews.ru/1091215/new-nvidia-gh200-grace-hopper-hbm3e>). 3DNews. (8 августа 2023). Дата обращения: 18 сентября 2023. Архивировано (<https://web.archive.org/web/20230905093213/https://3dnews.ru/1091215/new-nvidia-gh200-grace-hopper-hbm3e>) 5 сентября 2023 года.
8. [NVIDIA похвасталась рекордами H100 в новом бенчмарке MLPerf для генеративного ИИ](#) (<https://servernews.ru/1089042>). ServerNews. (27 июня 2023). Дата обращения: 18 сентября 2023. Архивировано (<https://web.archive.org/web/20230922085624/https://servernews.ru/1089042>) 22 сентября 2023 года.

9. Много памяти, быстрая шина и правильное питание: гибридный суперчип GH200 Grace Hopper обогнал H100 в ИИ-бенчмарке MLPerf Inference (<https://servernews.ru/1092799>). ServerNews. (11 сентября 2023). Дата обращения: 18 сентября 2023. Архивировано (<https://web.archive.org/web/20230913020959/http://servernews.ru/1092799>) 13 сентября 2023 года.
10. NVIDIA представила 1-Эфлопс ИИ-суперкомпьютер DGX GH200: 256 суперчипов Grace Hopper и 144 Тбайт памяти (<http://servernews.ru/1087469>). ServerNews. (29 мая 2023). Дата обращения: 18 сентября 2023. Архивировано (<https://web.archive.org/web/20230830145504/https://servernews.ru/1087469>) 30 августа 2023 года.
11. DGX GH200: суперкомпьютер Nvidia для искусственного интеллекта (<https://www.osp.ru/articles/2023/0605/13057211>). «Открытые системы». (5 июня 2023). Дата обращения: 18 сентября 2023. Архивировано (<https://web.archive.org/web/20230605072346/https://www.osp.ru/articles/2023/0605/13057211>) 5 июня 2023 года.

## Ссылки

---

- [Суперчип NVIDIA Grace Hopper \(https://www.nvidia.com/ru-ru/data-center/grace-hopper-superchip/\)](https://www.nvidia.com/ru-ru/data-center/grace-hopper-superchip/). Официальный сайт NVIDIA Corporation. Дата обращения: 18 сентября 2023.

---

Источник — [https://ru.wikipedia.org/w/index.php?title=GH200\\_Grace\\_Hopper&oldid=135743299](https://ru.wikipedia.org/w/index.php?title=GH200_Grace_Hopper&oldid=135743299)

---

**Эта страница в последний раз была отредактирована 23 января 2024 в 15:10.**

Текст доступен по лицензии Creative Commons «С указанием авторства — С сохранением условий» (CC BY-SA); в отдельных случаях могут действовать дополнительные условия.

Wikipedia® — зарегистрированный товарный знак некоммерческой организации «Фонд Викимедиа» (Wikimedia Foundation, Inc.)