

ВИКИПЕДИЯ

Ampere (микроархитектура)

Материал из Википедии — свободной энциклопедии

Ampere — микроархитектура графических процессоров, разработанная компанией NVIDIA в качестве преемника микроархитектуры Turing. Названа в честь французского физика и математика Андре-Мари Ампера. Была анонсирована 14 мая 2020 года. Ampere используется в графических процессорах NVIDIA Tesla A100 и GeForce 30.

Содержание

Преимущества Ampere

Чипы Ampere

Ускоритель A100 и DGX A100

Модельный ряд

Видеокарты

Чипы

Преемник

Примечания

Ссылки

Преимущества Ampere

- 7 нм техпроцесс GA100 (Tesla A100)
- 8 нм техпроцесс GA10x (GeForce 30)
- Ядра трассировки лучей второго поколения

Nvidia Ampere

Кодовое имя	GA10x
Дата выпуска	Май 2020 (<u>Tesla A100</u>), сентябрь 2020 (<u>GeForce 30</u>)
Производители	<u>TSMC</u> (7 нм), <u>Samsung</u> (8 нм)
Тип памяти	GDDR6X
	<div style="display: flex; justify-content: space-between;"><div style="text-align: center;">← <u>Volta</u> (professional), <u>Turing</u> (consumer)</div><div style="text-align: center;"><u>Hopper</u> (professional), <u>Ada Lovelace</u> (consumer) →</div></div>

- Возможность одновременного использования трассировки лучей, шейдинга и вычислений
- NVLink 3.0
- Новое аппаратное 5-ядерное декодирование JPEG (NVJPG)
- PCI Express 4.0 с поддержкой SR-IOV
- Видеопамять GDDR6X

Чипы Ampere

Чип	GA100 ^[1]	GA102 ^[2]	GA103 ^[3]	GA104 ^[4]	GA106 ^[5]	GA107 ^[6]	GA10B ^[7]	GA10F
Размер чипа	826 мм ²	628 мм ²	496 мм ²	392 мм ²	276 мм ²	200 мм ²	?	?
Транзисторы	54.2 млрд	28.3 млрд	22 млрд	17.4 млрд	12 млрд	8.7 млрд	?	?
Плотность транзисторов	65.6 МТр/мм ²	45.1 МТр/мм ²	44.4 МТр/мм ²	44.4 МТр/мм ²	43.5 МТр/мм ²	43.5 МТр/мм ²	?	?
Кластеров обработки графики (GPC)	8	7	6	6	3	2	2	1
Блоков мультипроцессоров (SM)	128	84	60	48	30	20	16	12
Ядра CUDA	12288	10752	7680	6144	3480	2560	2048	1536
Текстурный блоков (TMU)	512	336	240	192	120	80	64	48
Блоки растеризации (ROP)	192	112	96	96	48	32	32	16
Тензорных ядер	512	336	240	192	120	80	64	48
RT ядер	Н/Д	84	60	48	30	20	8	12
L1 кэш	24 МБ	10.5 МБ	7.5 МБ	6 МБ	3 МБ	2.5 МБ	3 МБ	?
	192 КБ в SM	128 КБ в SM					192 КБ в SM	?
L2 кэш	40 МБ	6 МБ	4 МБ	4 МБ	3 МБ	2 МБ	4 МБ	?

Ускоритель A100 и DGX A100

14 мая 2020 года был анонсирован и выпущен ускоритель A100 на базе Ampere. A100 имеет производительность FP32 19,5 терафлопс, 6912 ядер CUDA, 40 ГБ графической памяти и 1,6 ТБ/с пропускной способности видеопамяти. Ускоритель A100 изначально был доступен только в 3-м поколении серверов DGX. В DGX A100 также входит 15 ТБ хранилища PCIe Gen 4 NVMe, два 64-ядерных процессора AMD Rome 7742 и 1 ТБ ОЗУ. Начальная цена на DGX A100 составляла 199 000 долларов.

- NVIDIA A800 Active PCIe 40GB — для рабочих станций (*альтернатива Nvidia A100 для клиентов в Китае, скорость межсоединения NVLink в GPU снижена до 400 ГБ/с против 600 ГБ/с у A100, представляет собой 2x-слотовую карту PCIe с активным охлаждением*).



Пример 2x-слотовой карты PCIe с активным охлаждением.

Модельный ряд

Видеокарты

Видеокарты, в которых используются чипы с микроархитектурой Ampere:

- Tesla A100
- RTX A6000
- GeForce RTX 3090 Ti
- GeForce RTX 3090
- GeForce RTX 3080 Ti
- GeForce RTX 3080
- GeForce RTX 3070 Ti
- GeForce RTX 3070
- GeForce RTX 3060 Ti
- GeForce RTX 3060
- GeForce RTX 3050 Ti

- [GeForce RTX 3050](#)

Чипы

- GA100
- GA102
- GA104
- GA106
- GA107

Преемник

В декабре 2020 года появилась информация, что NVIDIA планирует показать новую микроархитектуру [Ada Lovelace](#), названную в честь математика [Ады Лавлейс](#), которая придёт на смену Ampere. Возможно, она станет основой для графических ускорителей серии [GeForce RTX 40](#). Количество кластеров обработки графики в Lovelace увеличится вдвое, до 12, что позволит включать в новые графические процессоры до 72 кластеров текстурных процессоров и 144 потоковых мультипроцессоров. Точная дата выхода пока неизвестна^[8].

Примечания

1. [NVIDIA A100 Tensor Core GPU Architecture \(https://images.nvidia.com/aem-dam/en-zz/Solutions/data-center/nvidia-ampere-architecture-whitepaper.pdf\)](https://images.nvidia.com/aem-dam/en-zz/Solutions/data-center/nvidia-ampere-architecture-whitepaper.pdf) (амер. англ.). *NVIDIA Corporation*. Дата обращения: 29 апреля 2024.
2. [NVIDIA GA102 GPU Specs \(https://www.techpowerup.com/gpu-specs/nvidia-ga102.g930\)](https://www.techpowerup.com/gpu-specs/nvidia-ga102.g930) (амер. англ.). *TechPowerUp*. Дата обращения: 29 апреля 2024.
3. [NVIDIA GA103 GPU Specs \(https://www.techpowerup.com/gpu-specs/nvidia-ga103.g989\)](https://www.techpowerup.com/gpu-specs/nvidia-ga103.g989) (амер. англ.). *TechPowerUp*. Дата обращения: 29 апреля 2024.
4. [NVIDIA GA104 GPU Specs \(https://www.techpowerup.com/gpu-specs/nvidia-ga104.g964\)](https://www.techpowerup.com/gpu-specs/nvidia-ga104.g964) (амер. англ.). *TechPowerUp*. Дата обращения: 29 апреля 2024.
5. [NVIDIA GA106 GPU Specs \(https://www.techpowerup.com/gpu-specs/nvidia-ga106.g966\)](https://www.techpowerup.com/gpu-specs/nvidia-ga106.g966) (амер. англ.). *TechPowerUp*. Дата обращения: 29 апреля 2024.
6. [NVIDIA GA107 GPU Specs \(https://www.techpowerup.com/gpu-specs/nvidia-ga107.g988\)](https://www.techpowerup.com/gpu-specs/nvidia-ga107.g988) (амер. англ.). *TechPowerUp*. Дата обращения: 29 апреля 2024.
7. [NVIDIA AGX Orin Series Technical Brief v1.2 \(https://www.nvidia.com/content/dam/en-zz/Solutions/gtcf21/jetson-orin/nvidia-jetson-agx-orin-technical-brief.pdf\)](https://www.nvidia.com/content/dam/en-zz/Solutions/gtcf21/jetson-orin/nvidia-jetson-agx-orin-technical-brief.pdf). *NVIDIA Corporation*. Дата обращения: 29 апреля 2024.

8. Раскрыты спецификации архитектуры NVIDIA Lovelace (<https://4pda.to/2020/12/28/380086/>). 4pda.to (28 декабря 2020). Дата обращения: 20 апреля 2021. Архивировано (<https://web.archive.org/web/20210420183652/https://4pda.ru/2020/12/28/380086/>) 20 апреля 2021 года.

Ссылки

- [Архитектура NVIDIA Ampere \(https://www.nvidia.com/ru-ru/data-center/ampere-architecture/\)](https://www.nvidia.com/ru-ru/data-center/ampere-architecture/) (рус.). Официальный сайт NVIDIA Corporation. Дата обращения: 21 сентября 2023.
 - [NVIDIA назвала запуск Ampere лучшим в истории — несмотря на дефицит и завышенные цены \(https://3dnews.ru/1037531\)](https://3dnews.ru/1037531/) // 3DNews, 16.04.2021
-

Источник — [https://ru.wikipedia.org/w/index.php?title=Ampere_\(микроархитектура\)&oldid=139439148](https://ru.wikipedia.org/w/index.php?title=Ampere_(микроархитектура)&oldid=139439148)

Эта страница в последний раз была отредактирована 5 августа 2024 в 22:26.

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.)