

Новое поколение видеокарт NVIDIA Quadro на чипе Pascal

Одним из главных событий начала 2017 года стал выход на рынок линейки профессиональных видеокарт NVIDIA следующего поколения на графических процессорах семейства GP10x с кодовым именем Pascal. Событие ожидаемое, ведь уже почти год радуют геймеров всей планеты видеоускорители NVIDIA GeForce 1000-й серии на подобных чипах. Задержка в полгода между анонсом карт на SIGGRAPH 2016 и началом продаж обусловлена оптимизацией ПО и сертификацией у разработчиков систем CAD/CAM/CAE.

Перед детальным описанием новых видеокарт стоит вкратце напомнить о ключевых особенностях моделей прошлого поколения, без сравнения с которыми не обойтись.

Итак, основа предыдущих ускорителей – чипы семейства GM20x Maxwell, изготовленные по нормам 28-нм техпроцесса. Число транзисторов на кристалле превысило 8 млрд, что позволило сформировать максимально 3072 шейдерных процессора. Но это лишь на 6% больше, чем в процессорах предыдущего семейства Kepler, (тоже 28-нм).

Таким образом, с 2012 по 2016 год все процессоры NVIDIA – сначала Kepler и Tesla (предназначенные для не визуальных вычислений), а затем Maxwell и Maxwell-2, производились по нормам 28-нм техпроцесса. Этот «застой» связан был, в первую очередь, с необходимостью решения ряда невероятно сложных технических задач, вставших перед производителями микросхем, в связи с переходом к более тонким технологическим процессам.

Другими словами, в «полный рост» дали о себе знать проблемы «второй половины шахматной доски» (см. статью «Закон Гордона Мура и новый GPU NVIDIA Quadro M6000», MediaVision № 7/2015). Поэтому с позиций сегодняшнего дня становится понятным стратегическое решение компании при разработке чипов семейства Maxwell сосредоточить усилия прежде всего на коренном изменении архитектуры, чтобы та наилучшим образом подходила для решения задач визуализации 32-разрядных видеоданных, даже в ущерб производительности при работе с 64-разрядными не визуальными данными.

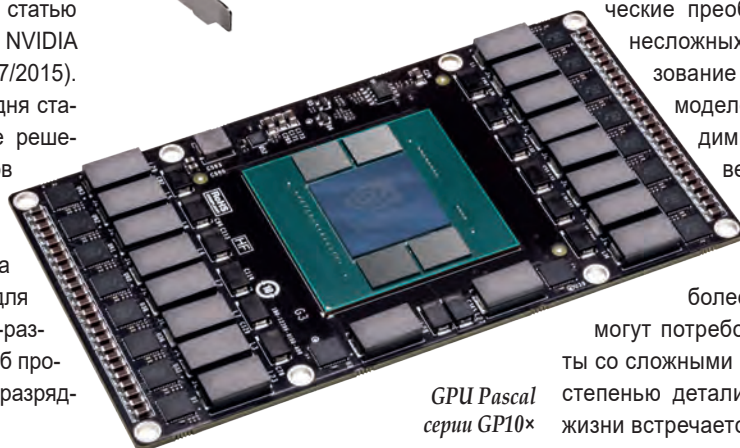
И вот, наконец, преодолев множество подводных камней, основной OEM-партнер NVIDIA – компания TSMC – приступила к серийному выпуску GPU Pascal – флагмана GP102 и GP102/104/106/107, а также GP100 с улучшенной производительностью при работе с 64-разрядными данными. Вот их основные характеристики:

- ◆ производство по новой 16-нм технологии FinFET;
- ◆ число транзисторов на кристалле более 15 млрд;
- ◆ тактовая частота – 1500 и 1800 МГц в базовом и разогнанном (Boost) режимах соответственно;
- ◆ максимальное число процессоров CUDA – 3840;
- ◆ модернизированные чипы памяти GDDR5X, благодаря которым пропускная способность возросла до 480 ГБ/с;
- ◆ максимальный объем памяти – 24 ГБ;
- ◆ максимальная производительность на 32-разрядных данных – более 12 TFLOPS;
- ◆ работа с видео 8K.

Отличие памяти GDDR5X от более привычной GDDR5 состоит в способности «предугадывать» цепочки данных, которые вероятнее



Игровая видеокарта NVIDIA GeForce 1000-й серии



GPU Pascal серии GP10x

Алекс Мастер

всего будут загружены следующими, и заранее загружать их, не дожидаясь соответствующей команды от GPU. Это чем-то сродни конвейеру «спекулятивной обработки», реализованной в центральных процессорах Intel Pentium еще в 90-е годы прошлого века.

По сравнению с видеокартами предыдущего поколения на чипах Maxwell, производительность новых видеокарт поднялась на 50...65%. Это наивысшие возможности GPU на ближайшие 5..6 лет – именно столько может потребоваться для разработки чипов следующего поколения. Может показаться, что срок очень большой, но только не для профессионала. Ведь вкладывая несколько сотен тысяч рублей в видеокарту NVIDIA Quadro, профессиональный пользователь должен быть уверен в защите своих инвестиций, а принудительное подталкивание к частой модернизации систем может привести к отрицательному результату.

Также не нужно забывать о том, что сфера производства медиаконтента – не единственная, где нужны профессиональные видеокарты NVIDIA Quadro. Есть еще всевозможные инженерные системы CAD/CAM/CAE для разработки конструкторской документации, рабочие места промышленного дизайна, визуализация результатов медицинских и прочих лабораторных исследований. Рабочая среда приложений CAD/CAM/CAE создает сравнительно небольшую вычислительную нагрузку, основа которой – элементарные геометрические преобразования, наложение несложных 2D-текстур и использование относительно простых моделей освещения. Необходимый и достаточный уровень производительности обеспечивают уже графические ускорители Quadro K4000, и более мощные видеокарты могут потребоваться лишь для работы со сложными 3D-моделями с высокой степенью детализации, что в реальной жизни встречается не часто.

Видеокарты Quadro Pxxxx

Модель	Характеристики						
	Чип	Число ядер CUDA	Память GDDR5X, Гб	Пропускная способность, Гб/с*	Производительность, TFLOPS**	Термопакет, Вт	Выходы видео
P6000	GP102	3840	24	432	12	250	4×DP1.4, DVI-D, стерео
P5000	GP104	2560	16	288	8,9	180	4×DP1.4, стерео
P4000		1792	8	208	5,3	105	
P2000	GP106	1024	5	156	3,0	75	4×DP1.4
P1000	GP107	640	4	121	1,894	47	4×мини-DP1.4
P600		384	2	108	1,195	40	
P400		256	2	89	0,641	30	

* Шины памяти

**При работе с 32-разрядными данными

Иное дело – работа с медиаконтентом, где в последние несколько лет стандартом де-факто стало разрешение 4K, и не за горизонтом переход к производству в 8K. Знают в компании NVIDIA русскую поговорку «Готовь сани летом, а телегу зимой» или нет, но, тем не менее, к повсеместному переходу на разрешение 8K уже подготовили целую линейку новых высокопроизводительных решений, речь о которых ниже.

Сами по себе эти цифры мало что скажут потенциальному пользователю – ему важно знать, для работы с каким контентом хватит этой мощности, или, другими словами, какова область наиболее эффективного применения той или иной модели GPU. Выбор всегда за пользователем, но есть несколько рекомендаций:

- ♦ видеокарты начального уровня Quadro P400 и P600 достаточны для несложного повседневного монтажа HD видео и 2D анимации;
- ♦ видеокарты среднего уровня Quadro P1000 и P2000 подойдут для насыщенного эффектами монтажа HD-видео и цветокоррекции, а также 3D-моделирования;
- ♦ Quadro P4000 – это, можно сказать, золотая середина. Модель дает возможность практически без ограничений работать с 4K- и 3D-видео (данная видеокарта – минимальный вариант в рамках требований VR Ready);

♦ GPU высокого уровня Quadro P5000 позволяет подняться на следующую ступень производительности – полноценно работать с видео 8K и создавать VR-контент;

♦ Quadro P6000 – High End, и этим все сказано. Самый производительный на сегодня одиночный GPU, может работать с любым контентом и не ощутить недостатка мощности на протяжении, по крайней мере, ближайших 10 лет.

Также к этому ряду принадлежит видеокарта Quadro GP100, которая стоит несколько особняком и вот почему. Поиск новых архитектурных решений для достижения более высокой вычислительной мощности привел в начале десятилетия к созданию видеопамати стандарта HBM (High Bandwidth Memory – память с высокой пропускной способностью). Основное ее отличие от ставшей уже классической GDDR в том, что вместо набора очень быстрых чипов памяти (7 ГГц и выше), «общающихся» с GPU через «узкую» шину разрядностью 128...512 бит, используется массив относительно медленных (около 1 ГГц) микросхем, обменивающихся с графическим процессором данными по шине шириной до 4096 бит.

Увеличение сразу в несколько раз разрядности шины данных потребовало кардинально иных технических решений, одним из которых стало размещение массива микросхем памяти в непосредственной близости от графического процессора на единой литографической подложке. Более того, так как чипов памяти теперь намного больше, изменился и способ их монтажа – он стал объемным, «многоэтажным», напоминающим многослойный бутерброд.

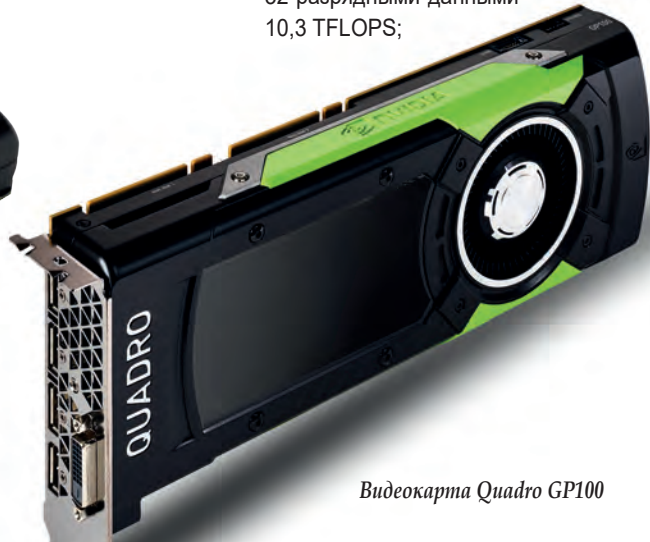
В видеокарте Quadro GP100 используется память уже второго поколения – стандарта HBM2, чипы которой изготовлены по нормам 20-нм техпроцесса, а пропускная способность может достигать 256 Гб/с (для «бутерброда» из 4 чипов емкостью по 1 Гб).

Основные характеристики видеокарты Quadro GP100:

- ♦ чип GP100 (15,3 млрд транзисторов на кристалле 610 мм²);
- ♦ 3584 CUDA-ядра;
- ♦ 16 Гб памяти HBM2;
- ♦ пропускная способность шины памяти – 720 Гб/с;
- ♦ производительность при работе с 32-разрядными данными – 10,3 TFLOPS;



Quadro P6000 –
мощнейший на сегодня GPU



Видеокарта Quadro GP100

- ◆ термопакет – 235 Вт;
- ◆ видеовыходы – 4× DP1.4, DVI-D, стерео.

На первый взгляд может показаться, что Quadro GP100 – лишняя в вышеописанной линейке графических ускорителей, тем более что она даже немного дороже, чем флагман NVIDIA Quadro P6000, и ей вряд ли найдется применение. Однако есть одно но, позволяющее по-новому взглянуть на роль этого GPU, которую он может сыграть в составе перспективной рабочей станции. Это – производительность GPU при обработке 64-разрядных данных. Для GP100 она составляет 5,2 TFLOPS, то есть половину от 32-разрядной производительности, тогда как у остальных видеокарт линейки производительность при работе с 64-разрядными данными падает в 32 (!) раза.

Иными словами, Quadro GP100 хороша там, где важна точность вычислений, а это моделирование реальных процессов – физических, химических, биологических, медицинских и пр. А благодаря наличию собственных видеовыходов видеокарта GP100, в отличие от ускорителей Tesla P100, может обойтись без поддержки дополнительных GPU, самостоятельно выводя картинку на экраны мониторов.

Отдельного внимания заслуживают расширенные возможности вывода видео, имеющиеся у новых видеокарт. Все GPU Pascal получили обновленный дисплейный контроллер DisplayPort 1.4. Если ранее 4 интерфейса DP видеокарт Maxwell позволяли вывести 4 независимые картинки 4K/60 Гц, то теперь появилась возможность подключить 4 монитора 5K/60 Гц, а изображения 4K выводить с кадровой частотой до 120 Гц. Объединенные в общий пул, 4 видеовыхода DP позволяют работать с 8K-монитором (7680×4320) – например, с недавно появившейся моделью Dell UP3218K.

Для тех, кому подключения одновременно четырех мониторов недостаточно, могут воспользоваться новым решением NVIDIA для создания многомониторных конфигураций – картой синхронизации Quadro Sync II. Благодаря ей число синхронизируемых видеокарт может достигать до 8, а, соответственно, мониторов – до 32. Их, в свою очередь, можно объединить в общее рабочее пространство с помо-



Плата синхронизации Quadro Sync II

щью фирменного приложения NVIDIA Mosaic. Не забыто и более бюджетное решение NVIDIA для объединения видеокарт – с помощью фирменного интерфейса NVLink (вместо устаревшего SLI) пара GPU будет стабильно работать в одной «упряжке».

Секреты операторского мастерства – из первых рук!

«Отдам в хорошие руки»

В книге заслуженного деятеля искусств России А.М. Кириллова рассказывается о различных операторских приемах, о том, как прямо на съемочной площадке добиться тех или иных эффектов, часто реализуемых лишь на стадии монтажа и обработки материала.

Книга богато иллюстрирована фотографиями, кадрами из кинофильмов и схемами, поясняющими конфигурацию съемочного пространства, расположение камеры, объекта съемки, осветительных приборов и применяемых приспособлений.

Книга будет полезна как начинающим, так и опытным кинооператорам.



**Стоимость книги с учетом доставки:
495 руб. 60 коп, в т.ч. НДС 18% – 75 руб 60 коп.**

Кириллов А.М.

Отдам в хорошие руки. – М.: «Издательство Медиавижн», 2013. – 96 с.

Чтобы приобрести книгу, нужно отправить заявку на адрес электронной почты: book@mediavision-mag.ru

Необходимая для приобретения информация:

Для юридического лица: название организации, юридический адрес, ИНН, КПП, почтовый адрес, по которому следует выслать заказ, адрес электронной почты для отправки электронных версий счета на оплату и других документов.

Для физического лица: ФИО, почтовый адрес, по которому будет выслана книга, адрес электронной почты для отправки электронной версии счета на оплату и информации для отслеживания почтового отправления.