

***Презентация по дисциплине: «Компьютерная графика»
на тему: «Графические процессоры»***

выполнена:

*студентом 3 курса,
группы 04-ПО-01
Криулиным Д. И.*

Работа

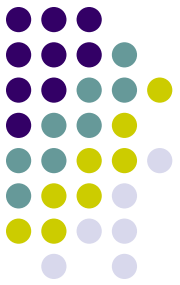
***Краснодар
2006***

Официальные спецификации GeForce 6800

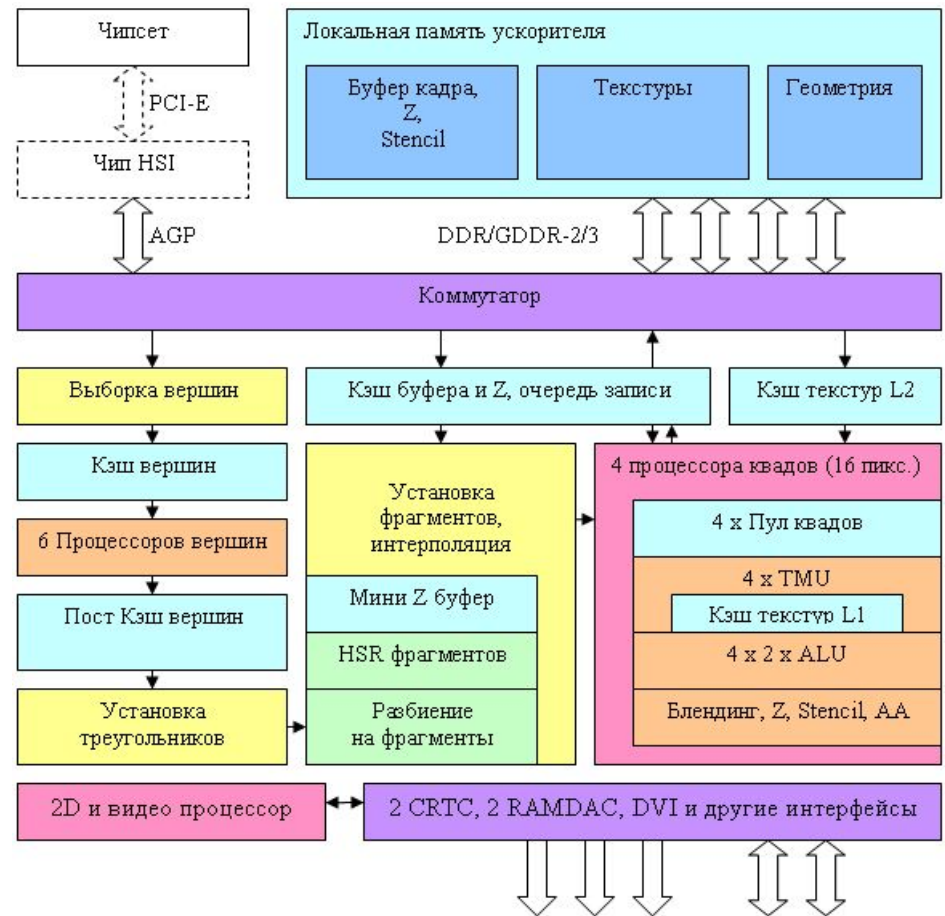


- 16 Пиксельных процессоров, по одному текстурному блоку на каждом с произвольной фильтрацией целочисленных и плавающих текстур (анизотропия степени до 16х включительно).
- 6 Вершинных процессоров, по одному текстурному блоку на каждом, без фильтрации выбираемых значений (дискретная выборка)
- Вычисление, блендинг и запись до 16 полных (цвет, глубина, буфер шаблонов) пикселей за такт.
- Вычисление и запись до 32 значений глубины и буфера шаблонов за такт (если не производятся операции с цветом).
- Поддержка «двустороннего» буфера шаблонов
- Поддержка специальных оптимизаций прорисовки геометрии для ускорения алгоритмов теней на основе буфера шаблонов (т.н. технология Ultra Shadow II).
- Все необходимое для поддержки пиксельных и вершинных шейдеров версии 3.0, включая динамические ветвления в пиксельных и вершинных процессорах, выбор значений текстур из вершинных процессоров и т.д.
- Фильтрация текстур в плавающем формате
- Поддерживается буфер кадра в плавающем формате (включая операции блендинга).
- MRT (Multiple Render Targets - рендеринг в несколько буферов).

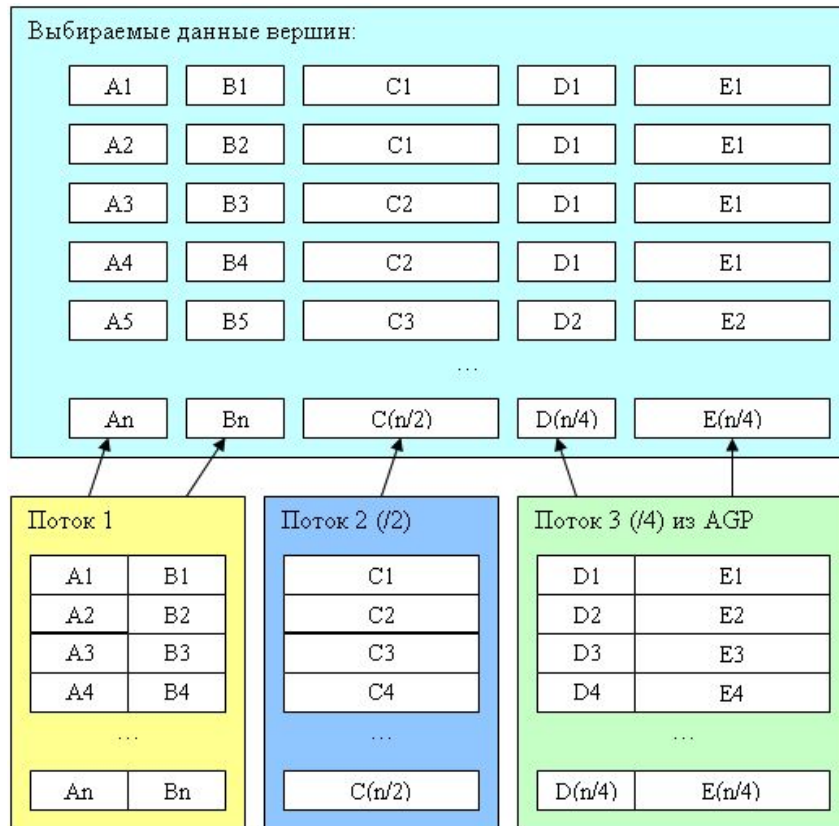
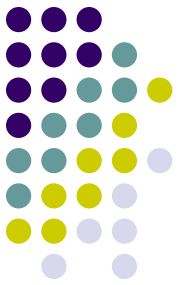
Общая схема чипа NV40



- -многоканальный контроллер памяти. Вместо одной очень широкой 128- или 256-бит шины памяти используются 2 или 4 полностью независимых (разведены отдельные управляющие сигналы) шины памяти шириной 32 или 64 бита соответственно
- -Кэш вершин: ~50..100..200 вершин;
- Пост Кэш вершин: ~16..32 вершины;
- Кэш текстур: ~32..64 килобайт (~512 блоков 4x4);
- Кэш буфера кадра и глубины: 16..32 килобайт (~256 блоков 4x4);
- Минибуфер глубины: ~256 килобайт максимум.

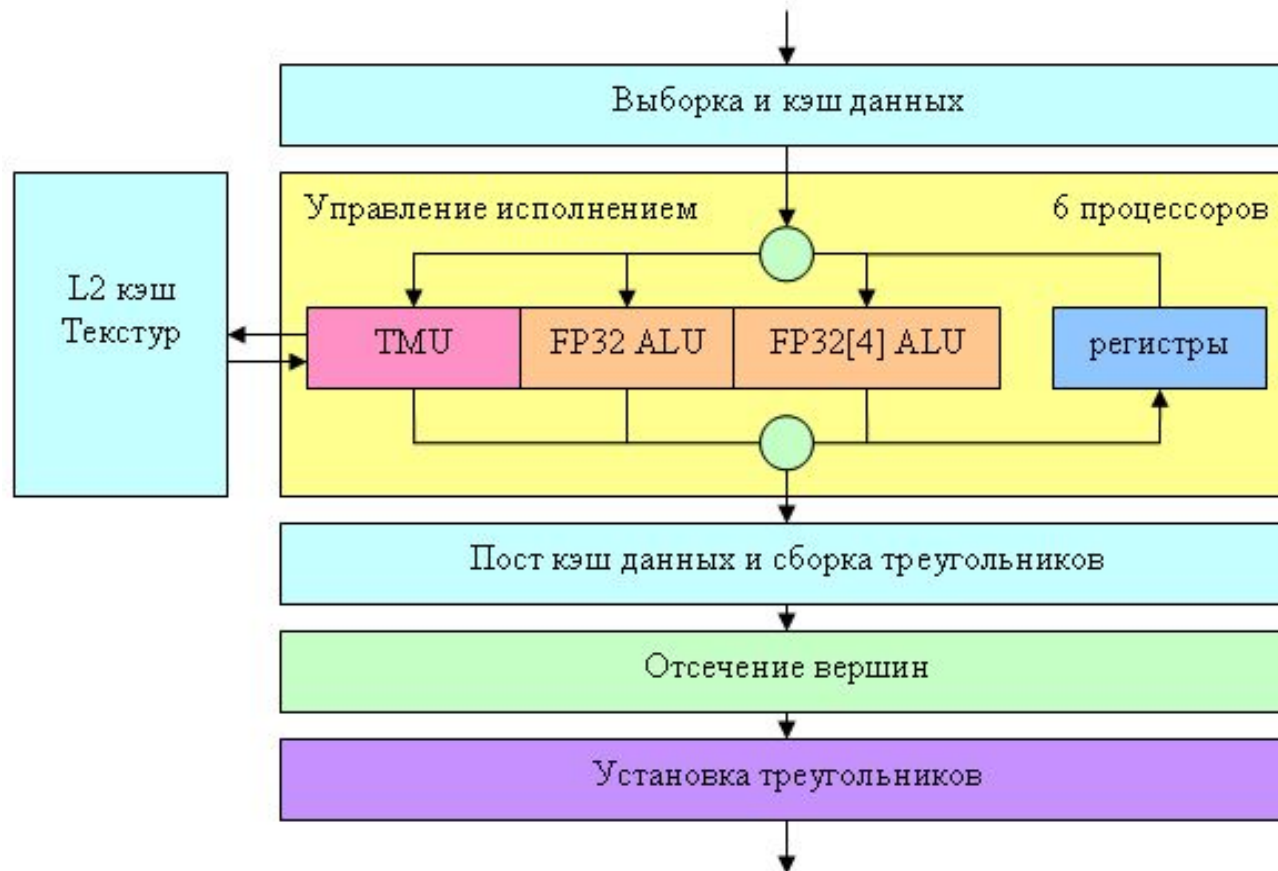
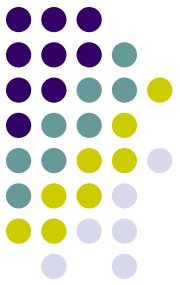


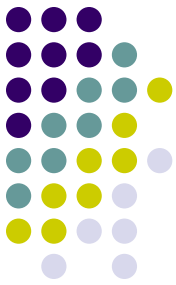
Вершинные процессоры и выборка данных



- Поддержка разных т.н. «делителей частоты» для потоков исходных данных вершин.
- Каждый поток может иметь свой собственный делитель значения индекса (т.н. Frequency Stream Divider). Таким образом, иногда отпадает необходимость дублировать данные, и мы экономим объем и ПСП локальной памяти и системной памяти адресуемой через AGP.

Блок схема вершинного процессора NV40

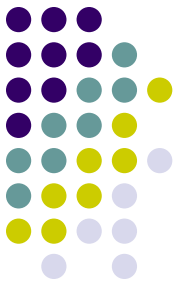




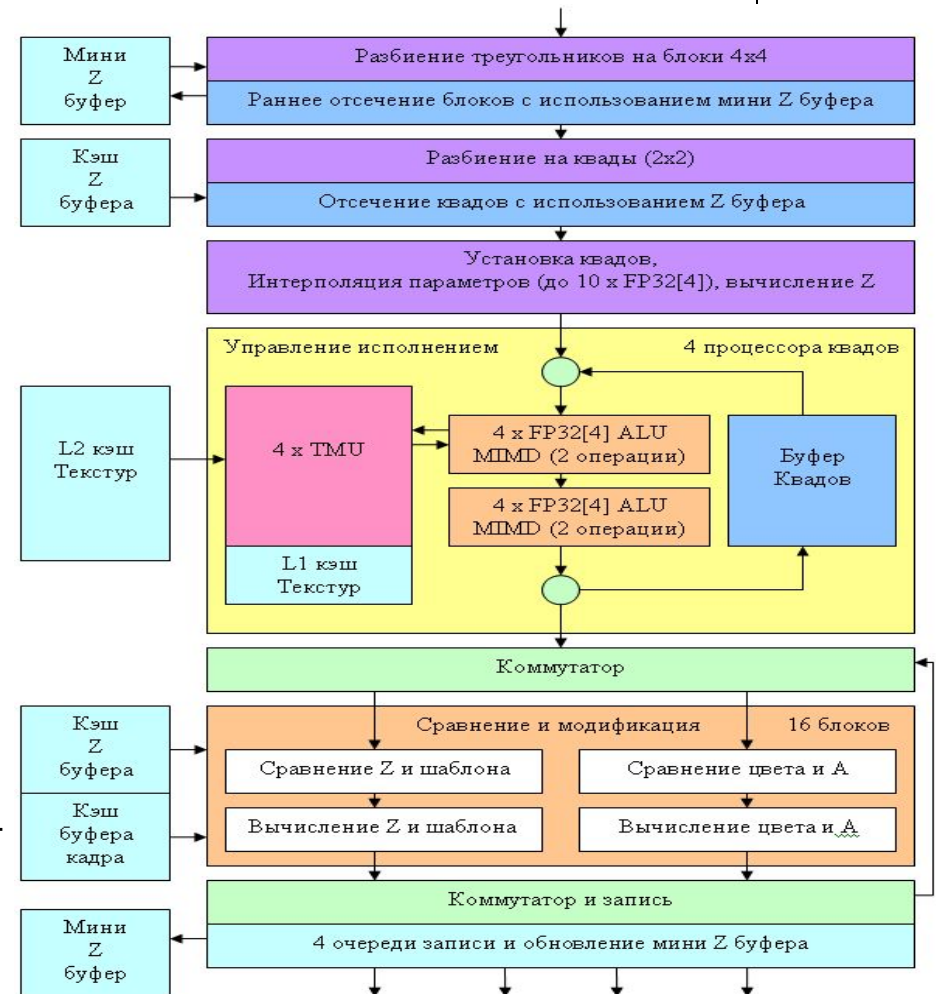
Параметры вершинного процессора NV40

Версия вершинных шейдеров	2.0	2. a	3.0 (NV40)
Число инструкций в коде шейдера	256	256	512 и более
Число исполняемых инструкций	65535	65535	65535 и более
Предикаты	Нет	Нет	Есть
Временных регистров	12	13	32
Константных регистров	256 и более	256 и более	256 и более
Статические переходы	Да	Да	Да
Динамические переходы	Нет	Да	Да
Глубина вложенности динамических переходов	Нет	24	24
Выбор значений текстур	Нет	Нет	Да(4)

Пиксельные процессоры и организация закрашки



- Число TMU – теперь у нас только по одному TMU на каждый пиксель квада. Всего у нас 4 процессора квадов, в каждом по 4 TMU, таким образом, всего их 16.
- Новые TMU поддерживают анизотропную фильтрацию с соотношением сторон до 16:1 включительно (т.н. 16х, у NV3X было до 8х) и, наконец-то, научились выполнять все виды фильтрации с плавающими форматами текстур. Правда, при условии 16 битной точности компонент (FP16).
- Двухуровневая организация кэширования текстур
- На каждый пиксель приходится по два ALU, причем каждое из них может выполнить две различные(!) операции над разным числом произвольно выбираемых компонент вектора (до 4х).
- Поддержка MRT (Multiple Render Targets – рендеринг в несколько буферов).
- Поддержка MSAA.

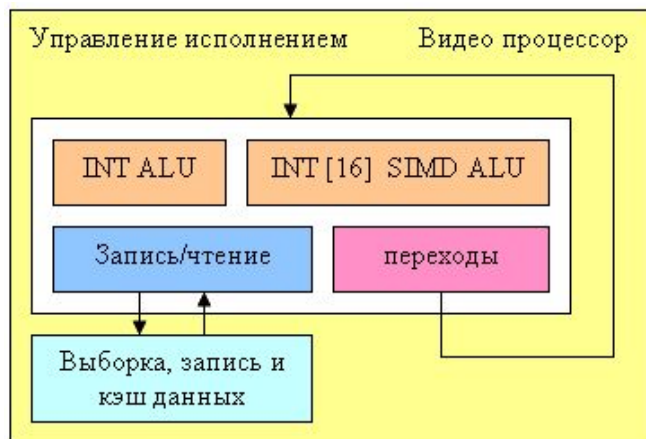


Параметры пиксельного процессора NV40



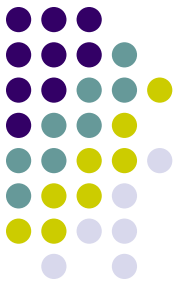
Версия пиксельного шейдера	2.0	2.a	2.b	3.0 (NV40)
Вложенность выборок текстур до	4	Без ограничений	4	Без ограничений
Выборок значений текстур до	32	Без ограничений	Без ограничений	Без ограничений
Длина кода шейдера	32 + 64	512	512	512 и более
Исполняемых инструкций шейдера	32 + 64	512	512	65535 и более
Интерполяторы	2 + 8	2 + 8	2 + 8	10
Предикаты	Нет	Да	Нет	Да
Временных регистров	12	22	32	32
Константных регистров	32	32	32	224
Произвольная перестановка компонент	Нет	Да	Нет	Да
Инструкции градиента (D D X/ D DY)	Нет	Да	Нет	Да
Глубина вложенности динамических переходов	Нет	Нет	Нет	24

2D и Видеопроцессор



- Процессор содержит четыре функциональных блока (целочисленное ALU, векторное целочисленное ALU с 16 компонентами, блок загрузки и выгрузки данных и блок управления переходами и условиями) и может таким образом выполнять до четырех различных операций за такт. Формат данных – целые числа, видимо 16 битной или 32 битной точности (точно не известно, но для некоторых алгоритмов 8 бит было бы недостаточно).
- Такой процессор способен существенно разгрузить CPU, особенно в случае больших разрешений видео, таких, как все более набирающие популярность форматы HDTV. Не известно, используются ли возможности этого процессора для ускорения 2D графики, особенно некоторых достаточно сложных функций GDI+ - было бы логично задействовать его на этом поприще, но точной информации об этом аспекте у нас нет.

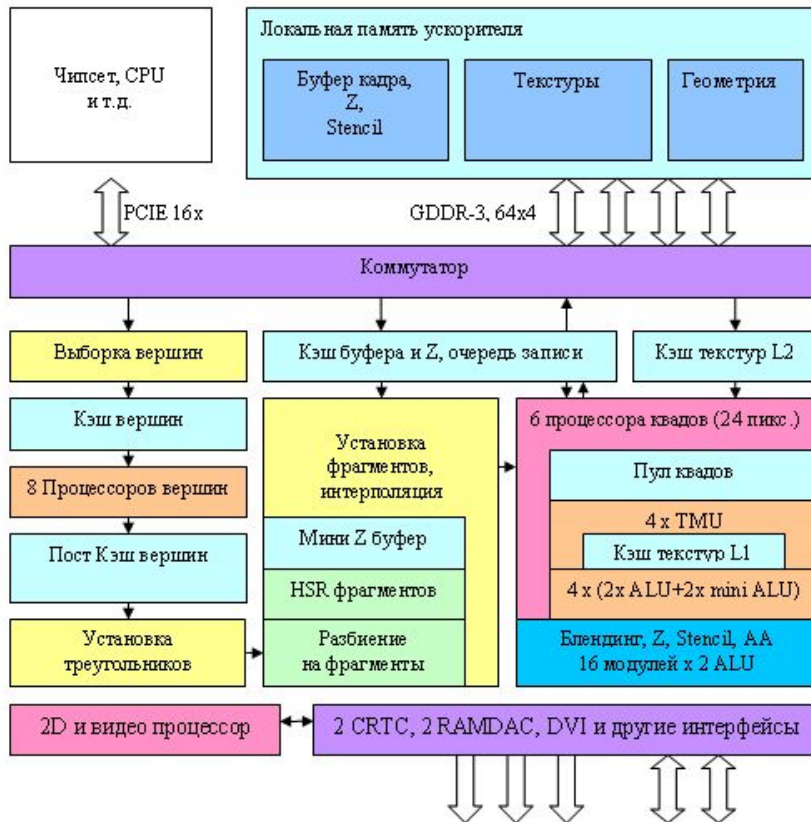
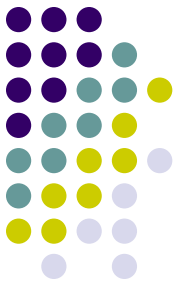
Спецификации GeForce 7800 GTX (кодовое название G70)



- 24 Пиксельных процессора, по одному текстурному блоку на каждом, с произвольной фильтрацией целочисленных и плавающих FP16 текстур (в том числе анизотропия, степени до 16х включительно) и бесплатной нормализацией FP16 векторов. Пиксельные процессоры улучшены по сравнению с NV4X — увеличено число ALU, возможно эффективное выполнения MAD операции.
- 8 Вершинных процессоров, по одному текстурному блоку на каждом, без фильтрации выбираемых значений (дискретная выборка).
- Вычисление, блендинг и запись до 16 полных (цвет, глубина, буфер шаблонов) пикселей за такт
- Вычисление и запись до 32 значений глубины и буфера шаблонов за такт (если не производятся операции с цветом)
- Поддержка «двустороннего» буфера шаблонов
- Поддержка специальных оптимизаций прорисовки геометрии для ускорения алгоритмов теней на основе буфера шаблонов и аппаратные карты теней (т.н. технология Ultra Shadow II)
- Все необходимое для поддержки пиксельных и вершинных шейдеров версии 3.0, включая динамические ветвления в пиксельных и вершинных процессорах, выбор значений текстур из вершинных процессоров и т.д.
- Фильтрация текстур в плавающем формате FP16.
- В вершинных шейдерах аппаратная фильтрация текстур не поддерживается, доступна только выборка значений без фильтрации.
- Поддерживается буфер кадра в плавающем формате (включая операции блендинга в формате компонент FP16 и только запись в формате FP32)
- MRT (Multiple Render Targets — рендеринг в несколько буферов)

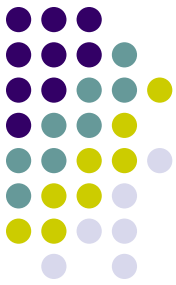


Архитектура ускорителя

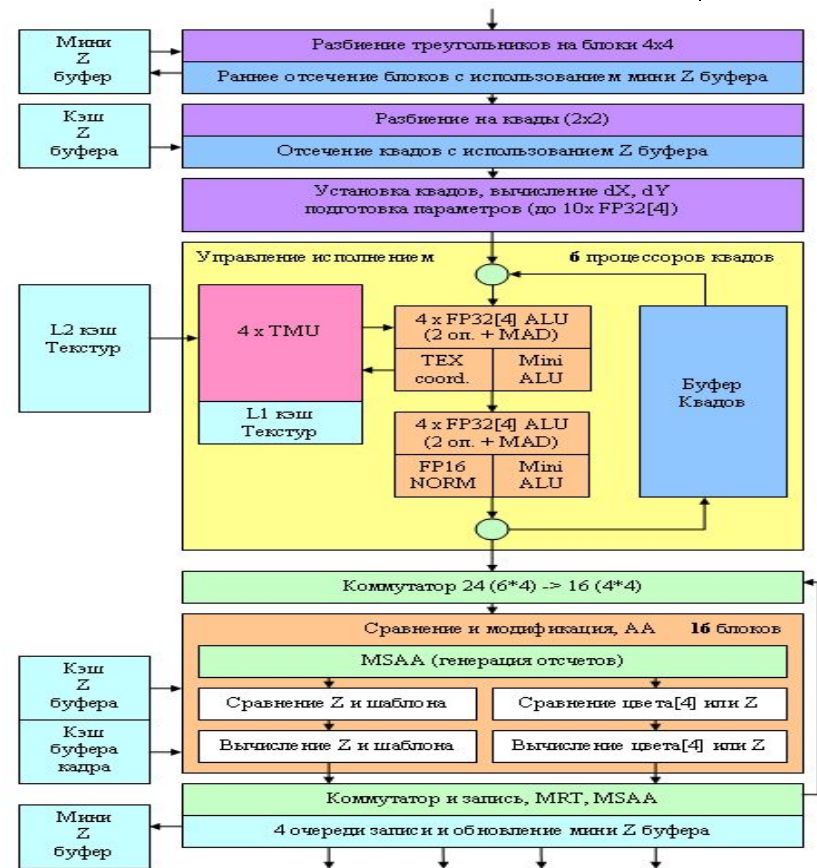


- Наличие 8 вершинных процессоров и 6 процессоров квадов (всего, таким образом, обрабатывается $4 \times 6 = 24$ пикселя) вместо 4 с большим числом ALU для каждого процессора.
- Новый чип может существенно быстрее рассчитывать шейдеры, причем у 24 пикселей параллельно, но по-прежнему записывает не более 16 полноценных пикселей за такт.

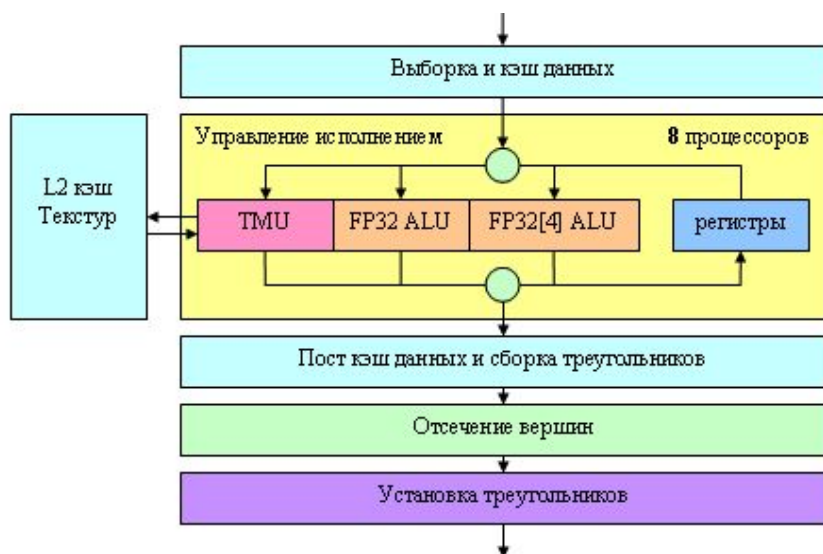
Пиксельный конвейер



- ALU умеют выполнить MAD операцию (одновременное умножение и сложение) без какого либо пенальти
- Выборка значений из текстур с форматом компонент FP32 возможна, но без аппаратной фильтрации
- За массивом из 6 процессоров квадов следует коммутатор, который перераспределяет рассчитанные квады по 16 блокам генерации глубины, AA и блендинга (а точнее по 4 связкам из 4-х блоков, обрабатывающим целый квад, так как геометрическая связанность не должна быть потеряна, так как понадобится при записи и сжатии цвета и буфера глубины).

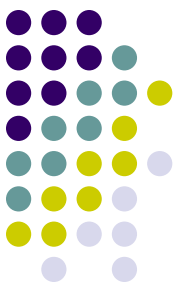


Вершинный конвейер



- Число вершинных процессоров увеличилось с 6 до 8.

Форматы данных, с которыми работает ускоритель

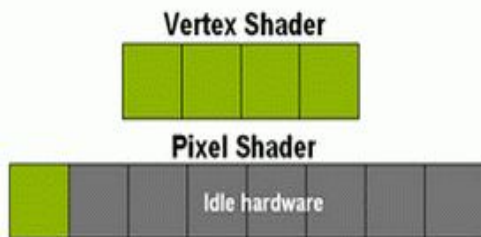


- VS 3.0 — FP32
- PS 3.0 — FP16, FP32
- Текстуры — INT8, FP16, FP32
- Буфер кадра — INT8, FP16, FP32
- А обработка данных (вычисления) происходит в следующих форматах:
- VS 3.0 — FP32
- PS 3.0 — FP32
- Текстуры — INT8, FP16, FP32 (без фильтрации)
- Буфер кадра — INT8, FP16 (без MSAA), FP32 (без блендинга и MSAA)

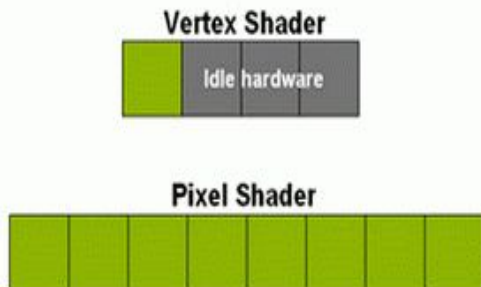
Унифицированная архитектура



Why unify?



Heavy Geometry
Workload Perf = 4



Heavy Pixel
Workload Perf = 8

Why unify?



Heavy Geometry
Workload Perf = 12



Heavy Pixel
Workload Perf = 12

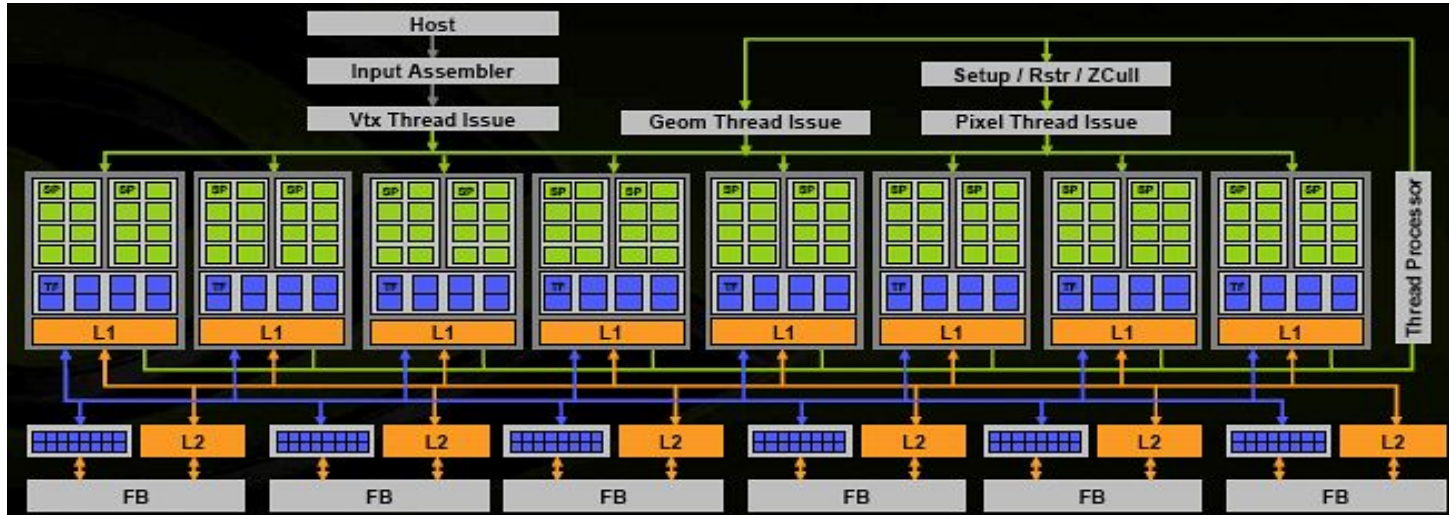
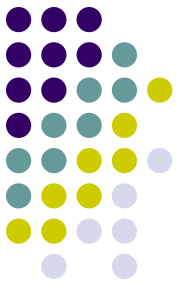
Графический ускоритель GeForce 8800



- **Унифицированная архитектура** (массив общих процессоров для потоковой обработки вершин и пикселей, а также других возможных видов данных)
- Аппаратная поддержка последних новшеств DirectX 10, в том числе и новая шейдерная модель - SM4, генерация геометрии и запись промежуточных данных из шейдеров.
- 384 бит шина памяти, 6 независимых контроллеров шириной 64 бита, поддержка GDDR4 (1.8ГГц)
- Частота ядра 575 МГц (GeForce 8800 GTX)
- 128 **скалярных** (не векторных, внимание!) ALU с плавающей точкой (целочисленные и плавающие форматы, поддержка FP 32бит точности в рамках стандарта IEE 754, MAD+MUL без потери тактов)
- ALU работают на удвоенной частоте (1.35ГГц для 8800GTX)
- 32 текстурных блока, поддержка FP16 и FP32 компонент в текстурах
- 64 блока билинейной фильтрации (то есть, возможна честная бесплатная трилинейная фильтрация, а также вдвое более эффективная по скорости анизотропная фильтрация)
- Возможность динамических ветвлений в пиксельных и вершинных шейдерах - размер блока планирования - 8x4 (32) пикселя.
- 6 широких блоков ROP (24 пикселя) с поддержкой AA до 16 семплов на пиксель, в том числе при FP16 или FP32 формате буфера кадра (то есть, возможны HDR+AA). Каждый блок состоит из массива гибко конфигурируемых ALU и отвечает за генерацию и сравнение Z, MSAA, блендинг. Пиковая производительность всей подсистемы до 96 MSAA отсчетов (+ 96 Z) за такт, в режиме без цвета (Z only) - 192 отсчета за такт.
- Запись результатов до 8 буферов кадра одновременно (MRT)
- Все интерфейсы вынесены на внешний дополнительный чип NVIO (2 RAMDAC, 2 Dual DVI, HDMI, HDTV)
- Очень хорошая масштабируемость архитектуры, можно по одному блокировать или убирать контроллеры памяти и ROP (Всего 6), шейдерные блоки (Всего 8 блоков TMU+ALU)
- Очень большой размер кристалла.

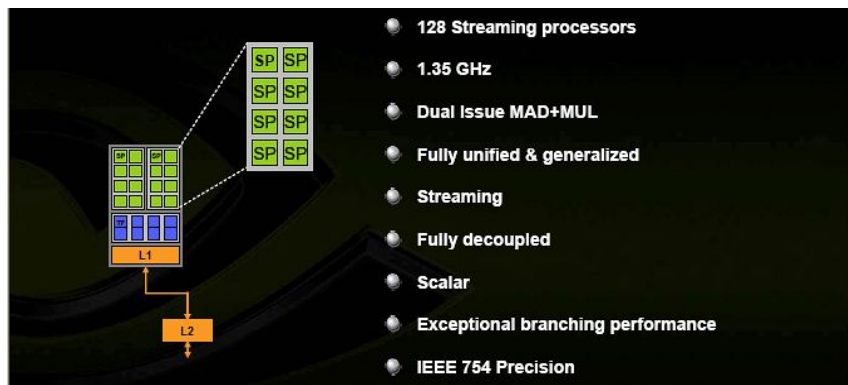


Архитектура



- Чип состоит из 8 универсальных вычислительных блоков (шейдерных процессоров), и хотя NVIDIA говорит о 128 процессорах, заявляя, что каждое ALU является таковым, это несколько неверно - единица исполнения команд - такой процессорный блок, в котором сгруппированы 4 TMU и 16 ALU. Всего, таким образом, мы имеем 128 ALU и 32 TMU, но гранулярность исполнения составляет 8 блоков, каждый из которых в один момент может заниматься своим делом, например, исполнять часть вершинного, или пиксельного, или геометрического шейдера над блоком из 32 пикселей.
- Каждый такой процессор снабжен собственным КЭШем первого уровня, в котором теперь хранятся не только текстуры, как ранее, но и другие данные, которые могут быть запрошены шейдерным процессором.
- Кроме управляющего блока и 8 вычислительных шейдерных процессоров в наличии 6 блоков ROP, исполняющих определение видимости, запись в буфер кадра и MSAA (синие, рядом с блоками КЭШа L2), сгруппированные с контроллерами памяти, очередями записи и КЭШем второго уровня.

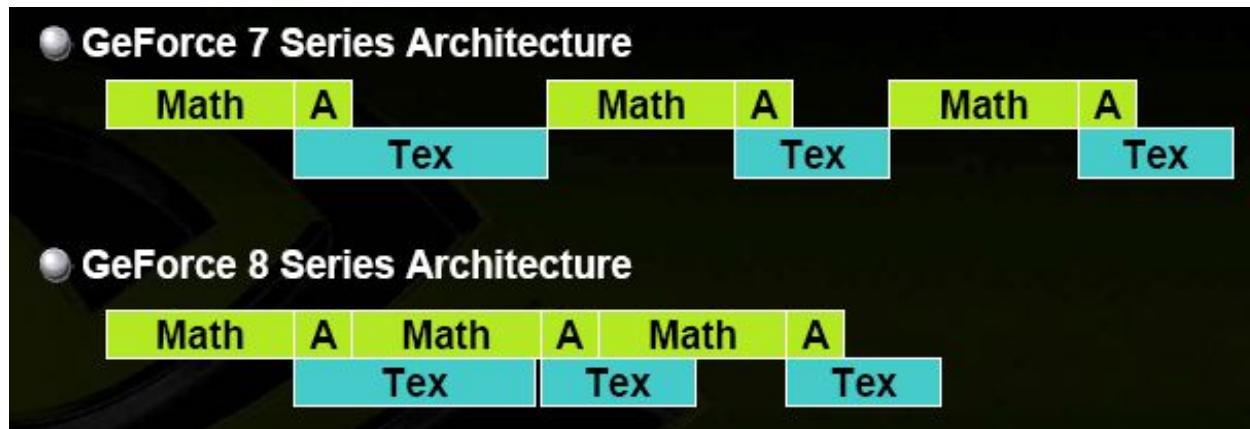
Шейдерный процессор и его TMU/ALU



	G71	R580	GeForce 8800
Structure	Vec3+scalar MAD	Vec3+scalar MAD Vec3+scalar ADD	Scalar MAD + Scalar MUL
# Units	48	48	128
Clock Rate (GHz)	.65	.65	1.35
GMuls	125	125	345

- В каждом из 8 шейдерных блоков в наличии 16 скалярных ALU. Что, опять таки, дает нам потенциальную возможность увеличить КПД их нагрузки вплоть до 100% вне зависимости от кода шейдера. ALU работают на удвоенной частоте и таким образом, соответствуют или превосходят (в зависимости от операций в шейдере) 8 четырехкомпонентных векторных ALU старого «образца» (G70) на равной базовой частоте ядра.

ALU



- Операция выборки и фильтрации текстур не требует ресурсов ALU и может теперь производиться полностью параллельно математическим вычислениям. Генерация текстурных координат по-прежнему отнимает часть времени ALU. Это снова логично, если мы хотим использовать транзисторы чипа на все 100%, ведь генерация текстурных координат требует стандартных плавающих операций, и заводить для нее отдельные ALU было бы непредусмотрительно.

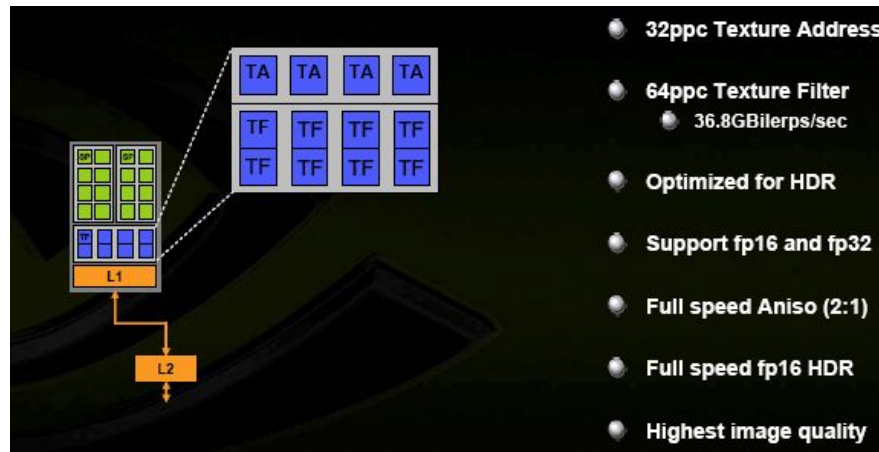
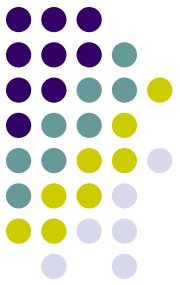
Параметры пиксельного процессора G80



	DX8 SM1.x	DX9 SM2.0	DX9c SM3.0	DX10 SM4.0
Vertex Instructions	128	256	512	64K
Pixel Instructions	4+8	32+64	512	64K
Vertex Constants	96	256	256	16x4096
Pixel Constants	8	32	224	16x4096
Vertex Temps	16	16	16	4096
Pixel Temps	2	12	32	4096
Vertex Inputs	16	16	16	16
Pixel Inputs	4+2	8+2	10	32
Render Targets	1	4	4	8
Vertex Textures	n/a	n/a	4	128
Pixel Textures	8	16	16	128
2D Texture Size			2K x 2K	8K x 8K
Int Ops	-	-	-	+
Load Op	-	-	-	+
Derivatives	-	-	+	+
Vertex Flow Control	n/a	Static	Static/Dynamic	Dynamic
Pixel Flow Control	n/a	n/a	Static/Dynamic	Dynamic

Налицо значительные количественные и качественные перемены - все меньше и меньше ограничений для шейдеров, все больше и больше общего с CPU. Пока без особого произвольного доступа (такая операция появилась в SM4, - пункт Load Op на диаграмме, но ее эффективность для общих целей пока сомнительна, особенно в первых реализациях), но нет сомнений, что в скором времени и этот аспект будет развит, как была развита за эти 5 лет поддержка FP форматов - от первых проб в NV30 до тотального, сквозного FP32 конвейера во всех режимах сейчас - в G80.

Текстурные модули(TMU)

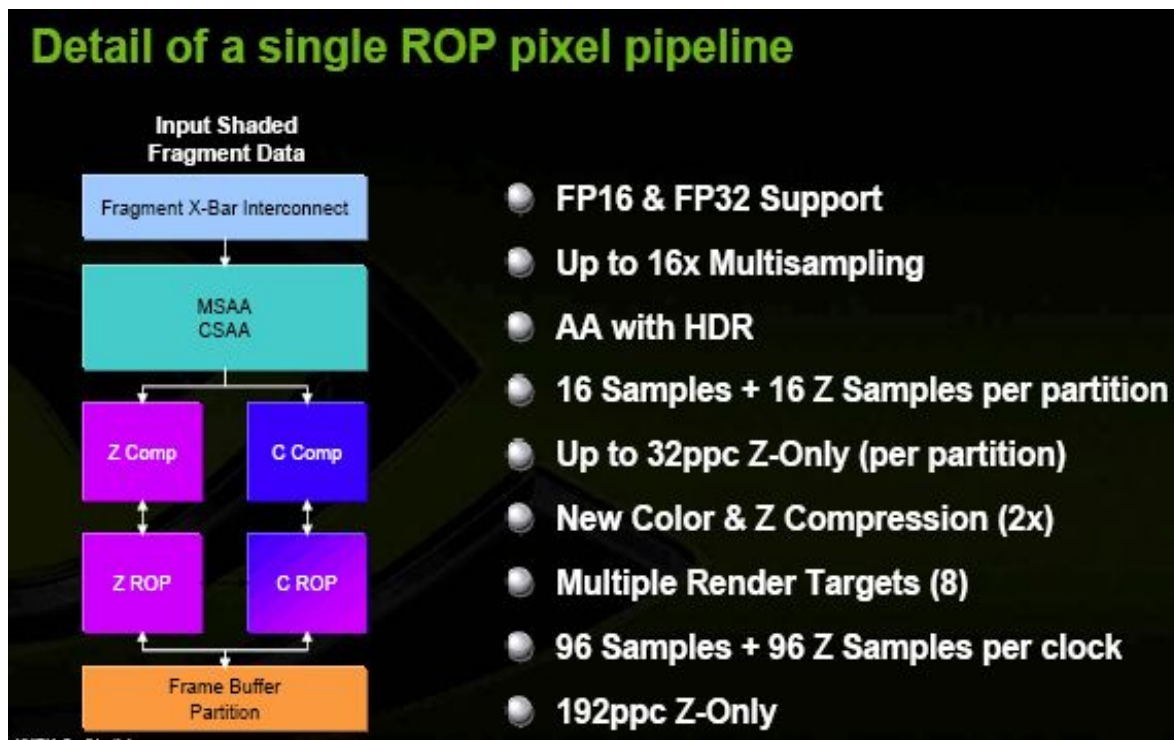


- В наличии 4 модуля для адресации текстур TA (определения по координатам точного адреса для выборки) и вдвое больше модулей для билинейной фильтрации TF.

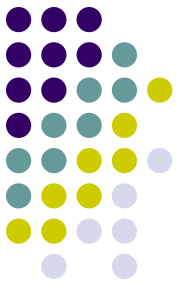
Блоки ROP, запись в буфер кадра, сглаживание



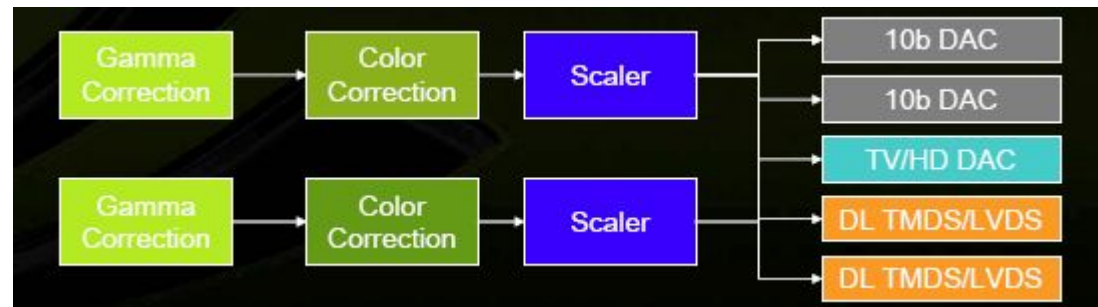
- полноценная поддержка FP32 и FP16 форматов буфера кадров ВМЕСТЕ с AA.



NVIO



- В нем интегрированы:
- 2 * 400 МГц RAMDAC
- 2 * Dual Link DVI (или LVDS)
- HDTV-Out
- Подсистема вывода, таким образом, выглядит так:



- Точность при этом составляет 10 бит на компоненту на всем ее протяжении.

Конвейер с точки зрения API DirectX 10

