

## Лекция 5 (4-2 2020 г). Шина PCI-X

История создания

Версии спецификаций

Основные отличия от PCI

Модификация сигналов и протокола PCI (пакетные и одиночные транзакции, последовательность, атрибуты, отложенная и расщепленная транзакция)

Обмен ролями

Уровни протокола, форматы пакетов, кодирование.

Коммутаторы. Режимы PCI-X

Механизм ECC (код исправления ошибок)

Электрический интерфейс PCI-X

Физический интерфейс.



# Версии спецификаций PCI

Стандарт	Разрядность, биты	Напряжение, В	Частота, МГц	Пропускная способность, Мб/с
PCI 2.0	32	5	33	133
PCI 2.1-3.0	32	5 или 3,3	33 или 66	133 (если частота 33 МГц) 266 (если частота 66 МГц)
PCI 64	64	3,3 или 5	33	266
PCI 66	64	3,3	66	533
PCI-X	64	1,5 или 3,3	66-533	1024-4096

Стандарт PCI 66 – это развитие PCI 64. Главное отличие – 66 МГц, слотов только с напряжением 3,3 В. Благодаря повышенной частоте выросла и пропускная способность, которая составляет 533 Мб/с.

**PCI-X – это еще одна модификация стандарта PCI 64.** Основное отличие – зависимость частоты шины от количества устройств:

66 МГц – 4, 100 МГц – 2, выше 133 МГц – 1.

Существуют две версии PCI-X. У первой пропускная способность 1024 МБ/с, у второй – 4096 МБ/с. Напряжение у PCI-X – 1,5 В, но данная шина совместима с платами PCI 64 3,3 В.

**Горячее подключение**

# История создания

Шина PCI-X разработана в 1998 году компаниями IBM, HP и Compaq как расширение шины PCI.

PCI-X 1.0

PCI-X 2.0 – 2003 г.

**Основная цель создания шины – улучшить ключевые характеристики шины PCI, прежде всего пропускную способность и надежность, за счет усложнения протокола обмена данными и увеличения тактовой частоты.**

Шина PCI-X применяется только в рабочих станциях и серверах.

**Совместимость с устройствами PCI** – механическая, электрическая, логическая – сохранена в полном объеме, но при наличии устройства PCI вся шина работает в режиме совместимости.

На сегодня практически полностью вытеснена шиной PCI Express.

# Совместимость с PCI

Первоначально были разработаны три варианта – PCI-X66, PCI-X100 и PCI-X133 с частотами 66, 100 и 133 МГц соответственно, ширина шины 32 или (чаще) **64 бита**.

PCI-X является 64-битной шиной. . В результате пропускная способность автоматически удваивается, равно как число дорожек/контактов и размеры слота.

Но всё остальное, включая протоколы передачи, сигналы и типы разъёмов, обратно совместимо. То есть в слот PCI-X можно установить 32-битную карту PCI (3,3 В). Кроме того, многие 64-битные карты PCI-X могут работать в 32-битных слотах PCI, но, конечно, с заметно сниженной пропускной способностью.

# Шина PCI-X 1.0

## Основные отличия от PCI

- возможность работы на тактовой частоте шины в 100 и 133 МГц;
- возможно использование различных слотов для разных скоростей обмена данными; зависимость частоты шины от количества устройств:
  - 66 МГц – 4,
  - 100 МГц – 2,
  - 133 МГц – 1 (2, если одно или оба устройства не находятся на платах расширения, а уже интегрированы на одну плату вместе с контроллером)
  - 266, 533 МГц и выше – 1
- значительно уменьшено время, выделяемое на операции в PCI-X (все времена в наносекундах);
- разрядность шин данных/адреса 64 бита;
- максимальная пропускная способность возросла до **1064 МБ/с**.
- механизм отдельных транзакций для улучшения производительности при одновременной работе нескольких устройств.

# Шина PCI-X 2.0

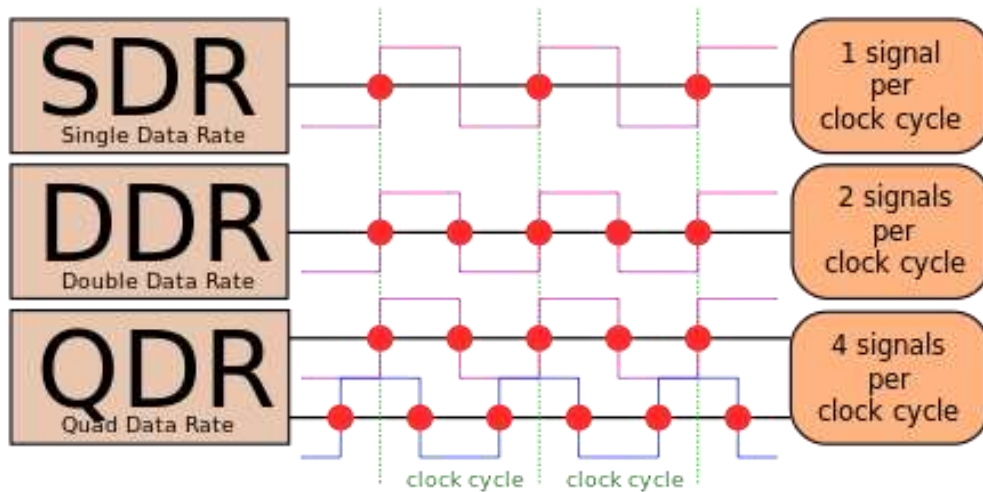
## Основные отличия от PCI

- Режимы PCI-X266 и PCI-X533: частота 133 МГц, обмен данными с частотой 2х и 4х соответственно

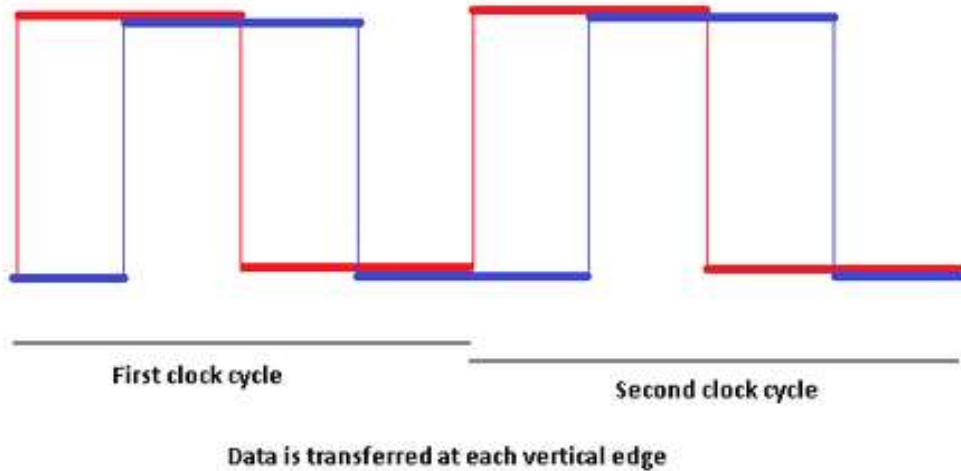
Технология удвоенной передачи данных (DDR - Double Data Rate), когда данные передаются на спаде и возрастании тактового импульса. **2128 МБ/с.**

Технология учетверённой передачи данных (QDR - Quad Data Rate). Intel использует ее для FSB процессоров Pentium 4 и Xeon. **4256 МБ/с.**

- Понижение напряжения питания и уровней сигналов до 1.5 В
- Добавление механизма коррекции ошибок четности при передаче данных (ECC – Error Checking and Correction, Error Correction Code)
- Новый 16-битный интерфейс для микросхем на материнской плате
- Расширение конфигурационного пространства до 4 Кб
- Добавление механизма обмена сообщениями между устройствами (DIM – Device ID Message)
- Допускает расщепление на 4 независимых 16-битных шины, что применяется исключительно во встраиваемых и промышленных системах



A comparison between single data rate, double data rate, and quad data rate



In QDR, the data lines operate at twice the frequency of the clock signal.

# Модификация сигналов и протокола PCI

Направлена на снижение задержек и устранение накладных расходов.

Для определения возможностей карты добавлен новый контакт

PCIXCAP – поддержка протокола PCI-X

(на PCI заземлен, на PCI-X133 соединен с землей через конденсатор (0,01мкФ), на PCI-X66 – параллельной цепочкой RC (10 кОм, 0,01мкФ).

Назначения остальных сигналов остались без изменений, кроме C/BE# - они не действуют в пакетных транзакциях (кроме MW - *Memory Write*).

***Ввели дополнения, позволяющие устройствам «предвидеть» грядущие события и выбирать адекватное поведение.***

**В обычной PCI все транзакции начинаются одинаково (с фазы адреса) как пакетные с заранее неизвестной длиной. При этом реально транзакции ввода/вывода всегда имеют лишь одну фазу данных; длинные пакеты эффективны (и используются) только для обращений к памяти.**



# Длина транзакций PCI-X

В PCI-X транзакции по длине разделены на два типа:

*пакетные (Burst)* — все команды, обращенные к памяти, кроме *Memory Read DWORD*;

*одиночные* размером в двойное слово (*DWORD*) — остальные команды.

Введено понятие последовательности (Sequence) – одной или нескольких логически связанных пакетных транзакций (чтение или запись в память), в рамках которых передается единый блок данных. Сколько байт передается в рамках одной транзакции, не определено (она может быть прервана в любой момент), но для последовательности инициатор заявляет суммарный объем в байтах.

Последовательность определяется

- идентификатором (RBN : RDN : RFN) инициатора
  - RBN — *номер шины*,
  - RDN — *номер устройства*,
  - RFN — *номер функции*),
- тэгом (ярлыком) – уникальным 5-битным номером.

# Модификация сигналов и протокола PCI

**Добавлена фаза атрибутов, следующая за фазой адресации перед фазами данных, в которой инициатор сообщает свой идентификатор.**

По шине AD в фазе атрибутов передается:

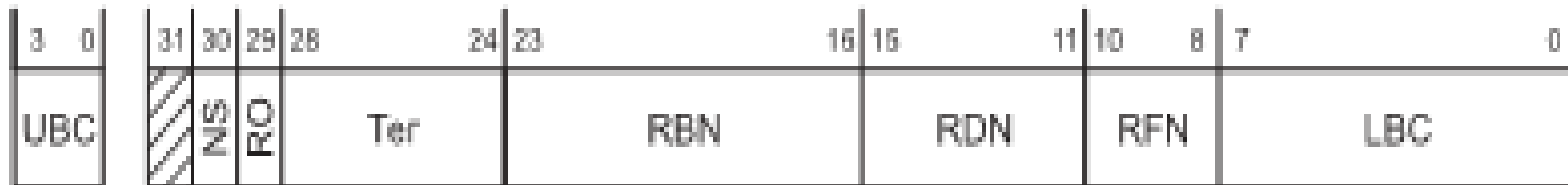
- 8 бит – младшая часть счетчика байт последовательности LBC (сколько байт осталось передать до конца).
- RFN (3 бита), RDN – (5 бит), RBN – (8 бит).
- Тэг (5 бит). Благодаря тегу каждый инициатор может одновременно выполнять до 32 логических передач (повторное назначение тега другой логической передаче возможно только после завершения предшествующей, с тем же тегом).

Дополнительные биты-признаки RO и NS области памяти, к которой относится транзакция. RO (Relaxed Order) – допустим произвольный порядок операций чтения и записи (для возможности переупорядочивания этих операций в целях оптимизации). NS (No Snop) – отсутствие кэширования.

По шине C/BE# передается старшая часть счетчика байтов UBC; итого макс. размер последовательности составляет 4096 байт.

12-битный *счетчик байтов* - только для пакетных транзакций)

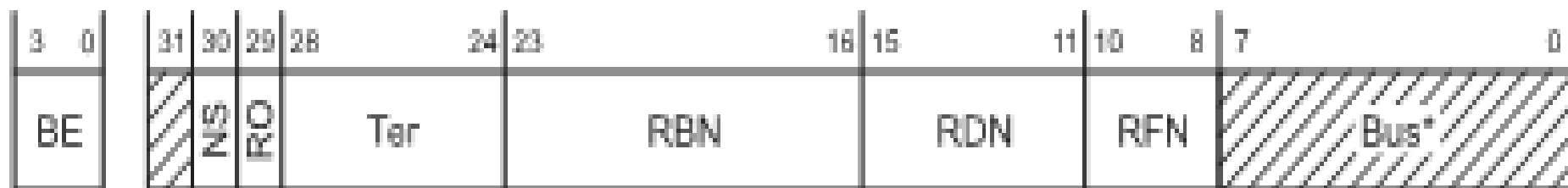
# Форматы атрибутов транзакции PCI-X



C/BE [3:0]#

AD [31:00]

а



C/BE [3:0]#


AD [31:00]

б

а — пакетная транзакция,

б — одиночная транзакция

(\* биты [7:0] используются только в цикле конфигурационной записи),



**Если транслируется транзакция чтения, то мост должен принять ее результаты, чтобы далее переслать их истинному инициатору транзакции.** Этот общий сценарий для различных команд реализуется по-разному, но «при всем богатстве выбора» у моста PCI есть всего два варианта ответа инициатору:

- отложить транзакцию**, ответив условием Retry. Этот вариант называется отложенной транзакцией (delayed transaction); он заставит инициатора через некоторое время повторить попытку данной транзакции. За это время мост должен «проверить» заказанную транзакцию на другой стороне интерфейса;
- сделать вид, что транзакция успешно завершена.** Такой вариант, называемый отправленной записью (posted write), возможен только для операций записи в память. Реальная запись произойдет позже, когда мост сумеет получить **управление шиной на противоположном интерфейсе.**

# Отложенная транзакция

Delayed Transaction (Отложенная транзакция)

Позволяет синхронизировать медленную шину ISA с более быстрой PCI ... 32разрядный буфер для поддержки циклов отложенных транзакций.

Включение функции Delayed Transaction позволяет встроенному 32-битному буферу записи чипсета поддерживать отложенные циклы транзакций. Это означает, что транзакции на шину ISA заносятся в буфер, и шина PCI освобождается для проведения других транзакций, пока шина ISA все еще занята этими транзакциями.

Данная функция должна быть включена (позиция Enabled) для лучшей работы шины PCI и соответствия техническим требованиям PCI 2.1.

# Расщепленные (Split) транзакции

Если целевое устройство не успевает начать обработку транзакции, оно должно не откладывать ее и запрашивать повторение (завершение типа Delay), а выполнять расщепление (Split) – выполнение с обменом ролями.

**В PCI-X отложенные транзакции (Delayed Transaction) заменены на *расщепленные транзакции (Split Transaction)*.**

Любую транзакцию, кроме транзакций записи в память, целевое устройство может завершать либо немедленно, либо с использованием протокола расщепленных транзакций. В последнем случае целевое устройство подает сигнал *Split Response* (расщепление), внутренне исполняет команду, а потом инициирует собственную транзакцию (команда *Split Completion*) для пересылки данных или сообщения о завершении инициатору исходной (расщепленной) транзакции.

# Шина соединения/разъединения

Мост PCI-X для транзакций, транслируемых с шины, работающей в режиме PCI-X, должен вместо *откладывания транзакции* *выполнять ее расщепление*.

**В настоящее время, с точки зрения организации транзакций, используется два типа шин: шины с коммутацией цепей (Circuit– Switched bus) и шины с коммутацией пакетов (Packet – Switched bus).**

Шины с коммутацией пакетов при наличии нескольких задатчиков на шине (master bus) обеспечивают значительно большую пропускную способность по сравнению с шиной с коммутацией цепей за счет разделения транзакции на две логические части: запроса шины и ответа. **Такая методика получила название «расщепления» транзакций (split transaction).** Иногда ее называют шиной соединения/разъединения **(connect/disconnect) или конвейерной шиной (pipelining bus).** Транзакция называется расщепленной, поскольку произвольное количество других пакетов или транзакций могут использовать шину между запросом и ответом.

# Расщепление

Транзакция чтения разбивается на

- транзакцию запроса чтения, которая содержит адрес, и
- транзакцию ответа исполнителя, которая содержит данные.

Каждая транзакция теперь должна быть **помечена (тегирована)** соответствующим образом, чтобы ЦП и исполнители могли сообщить, что есть что.



# Прерывание транзакции

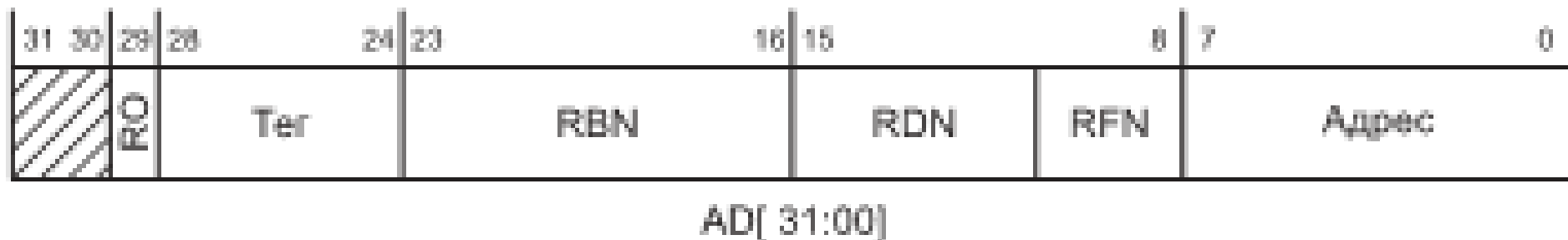


- **BCM** (Byte Count Modified) — признак того, что будет передано меньше байтов данных, чем просил запросчик (передается с данными завершения);
- **SCE** (Split Completion Error) — признак ошибки завершения, устанавливается при передаче сообщения завершения как ранний признак ошибки (до декодирования самого сообщения);
- **SCM** (Split Completion Message) — признак сообщения (отличает сообщение от данных).

# Транзакция завершения

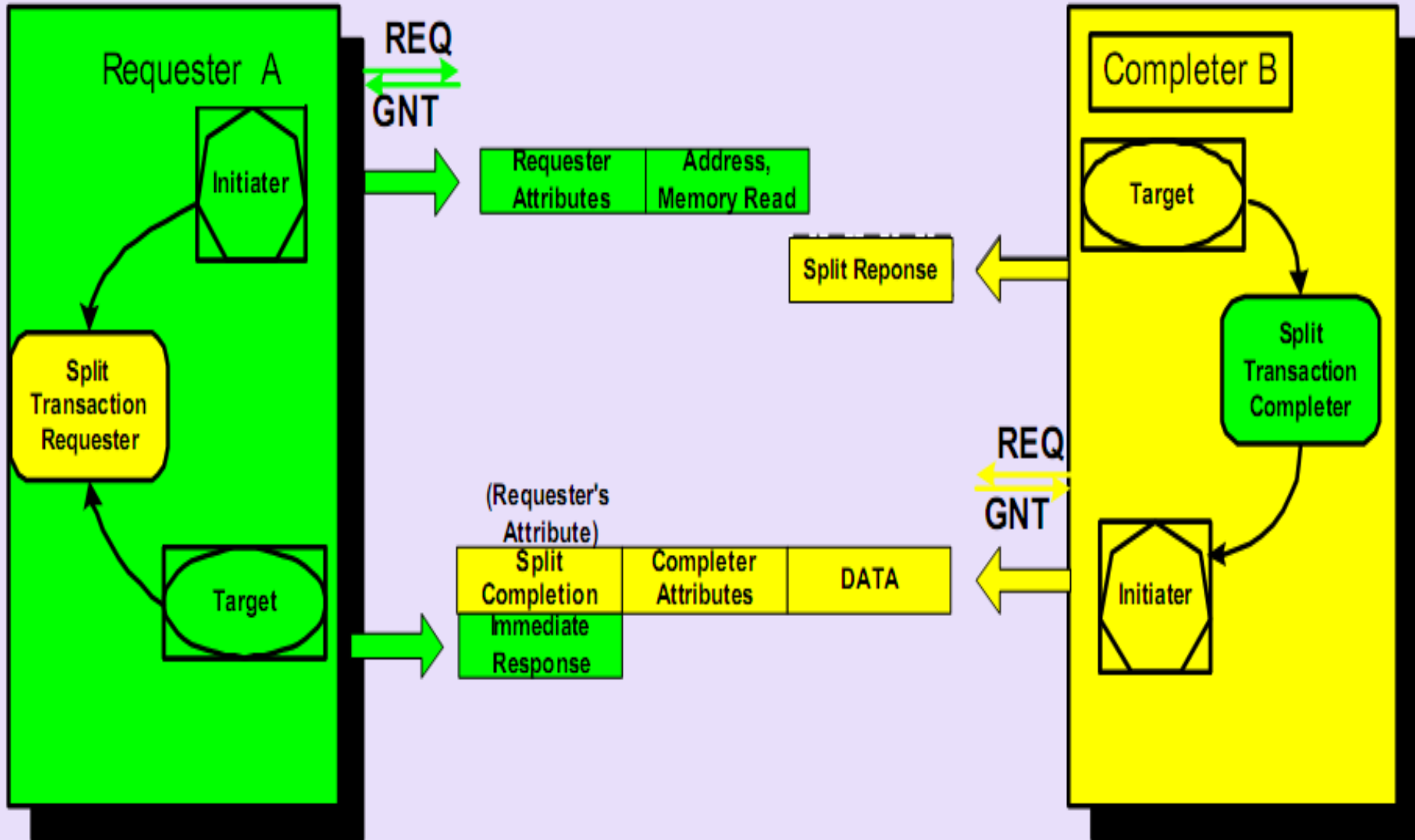
Прервав транзакцию, целевое устройство должно выполнить запрос, а потом вызвать транзакцию завершения (Split Completion) – код команды 1100 (вместо Memory Read Multiple, которая у PCI-X отменена).

В этой транзакции в фазе адреса передается идентификатор устройства, запросившего транзакцию, тэг, а также младшие 6 бит адреса (вместо счетчика) – для непоследовательного завершения. В фазе атрибутов передается счетчик байт, подлежащих завершению, признак ошибки или успешного выполнения (в старших битах AD). Фазы данных может не быть, если сообщается только факт завершения транзакции (бит 29 AD).



# Выполнение с обменом ролями

PCI Bus



# Выполнение с обменом ролями

ЦУ подает сигнал Split Response (расщепление), внутренне исполняет команду, а потом инициирует собственную транзакцию (команда Split Completion) для пересылки данных или сообщения о завершении инициатору исходной (расщепленной) транзакции.

Устройство, вызвавшее расщепленную транзакцию, называется **запросчиком** (Requester).

Устройство, завершающее расщепленную транзакцию, — **исполнителем** (Completer).

Для завершения транзакции исполнитель должен запросить управление шиной у арбитра; запросчик на этапе завершения выступает в роли ЦУ. Транзакция завершения (команда Split Completion) во многом напоминает пакетную транзакцию записи, но вместо полного адреса пространства памяти или ввода-вывода по шине AD передаются идентификатор последовательности (с номерами шины, устройства и функции запросчика), к которой относится это завершение, и только младшие 6 бит адреса. По этому идентификатору (номеру шины запросчика) мосты доводят транзакцию завершения до устройства-запросчика. Последовательность может отрабатываться и не одной транзакцией завершения, а серией транзакций.

Запросчик должен быть всегда готов к получению данных начатых им последовательностей, причем данные разных последовательностей могут приходить в произвольном порядке. Исполнитель может выдавать транзакции завершения на несколько последовательностей также в произвольном порядке.

# Режимы PCI-X

Каждый сегмент PCI-X (физическая шина) должен работать в самом прогрессивном режиме, доступном всем его абонентам, включая и главный для этой шины мост. В стандартной шине PCI «прогрессивность» определяется только допустимой тактовой частотой (33 или 66 МГц), и свои способности карта сообщает по контакту B49 (M66EN).

**В шине PCI-X появляются новые возможности: поддержка собственно протокола PCI-X (Mode 1 в терминах PCI-X 2.0) и ускоренных передач (Mode 2).**

Эти возможности карта сообщает через контакт B38 (PCIXCAP).

Мост, которому подчиняется данная шина, проверяет состояние линий M66EN и PCIXCAP по началу сигнала сброса. В соответствии с увиденными возможностями (они будут соответствовать самому слабому абоненту) мост выбирает режим работы шины. Этот режим доводится до всех абонентов с помощью шаблона инициализации (PCI-X Initialization Pattern) — уровней сигналов PERR#, DEVSEL#, STOP# и TRDY# в момент окончания сигнала RST# (по его нарастающему фронту). К этому моменту на слоты уже подается соответствующее напряжение +V<sub>I/O</sub>.

PCI	33	66			МГц
PCI-X	66	133	266	533	МГц

# Возможные режимы шины

Возможные режимы шины и их шаблоны инициализации приведены во второй таблице.

PERR#	DEVSEL #	STOP#	TRDY#	Режим	Частота, МГц	Контроль
H	H	H	H	PCI	0-33	PAR
H	H	H	H	PCI	33-66	PAR
H	H	H	L	PCI-X M1	50-66	PAR
H	H	L	H	PCI-X M1	66-100	PAR
H	H	L	L	PCI-X M1	100-133	PAR
H	L	L	L	PCI-X M1	100-133	ECC
L	L	L	H	PCI-X M2	66-100	ECC

# Протокол режима Mode 2

В PCI-X 2.0 введен новый режим работы шины – Mode 2:

- Напряжение питания понижено до 1.5 В
- Частота составляет 133 МГц
- Добавлен механизм ECC
- Увеличена задержка декодирования адреса (от FRAME# до DEVSEL#) с 1 до 2 тактов
- В транзакции Memory Write Block (код команды 1111) линии C/BE# используются для удвоенной или учетверенной синхронизации данных; пары BE#[1:0] и BE#[3:2] – дифференциальные строб-сигналы для AD[15:0] и AD[31:16] соответственно
- Поддержка 16-битной шины: используются линии AD[31:16] и C/BE[3:2], все фазы занимают по два такта (первыми идут младшие биты)

# Механизм обмена сообщениями (DIM)

**DIM – Device ID Message, транзакция не по адресу памяти или портов в-в, а по идентификатору устройства. Поддержка DIM введена в PCI-X 2.0, она необязательна для устройств, только для мостов.**

В фазе адреса передается:

- Код сообщения, 8 бит – зависит от класса сообщения
- CBN:CDN:CFN – ID устройства назначения
- Класс сообщения – 4 бита

В фазе атрибутов старший бит AD – признак первой транзакции (начала сообщения).

Сообщение – это последовательность, его длина может достигать 4096 байт.



# Ужесточение правил обмена

- Инициатор не может вводить холостые такты.
- Первая порция данных выставляется на шину через 2 такта после фазы атрибутов, вторая – через 2 такта после DEVSEL#.
- Если сигнала TRDY# нет, инициатор начинает повторять первые две порции до его появления.
- Целевое устройство может задерживать (холостым ходом) только первую фазу данных.
- Прерывание транзакции возможно только на границе 128 байт.
- В пакетных транзакциях, кроме MW, все байты разрешены и валидны.

# Механизм ЕСС (код исправления ошибок)

Этот метод включает определение ошибки не только в одиночном разряде, но и двух, трех и четырех разрядах. Кроме того ЕСС может также исправлять ошибку в одиночном разряде.

Во время считывания результат ЕСС слова сравнивается с рассчитанным, подобно тому, как происходит в методе проверки четности. Основное различие состоит в том, что в проверке четности каждый бит связан с одним байтом, в то время как **ЕСС слово связано со всеми 8 байтами**. Это означает, что разрядные значения для ЕСС не будут теми же, что и индивидуальные биты для проверки четности для тех же восьми байтов, поэтому модули ЕСС не могут использоваться в режиме четности (однако, паритетные модули, могут использоваться в режиме ЕСС, как описано ниже).

# Механизм ЕСС

Обязателен для Mode 2, может быть включен для Mode 1.

Устройство может не исправлять ошибки (отключено в конфиг. регистрах), но обязано проверять ЕСС.

Сигналы ЕСС передаются по отдельным линиям шины. В 32-битном режиме используются ЕСС[6:0], в 64-битном – ЕСС [7:0].

Биты ЕСС относятся к данным AD предыдущей фазы данных, и к данным C/BE# за две фазы от текущей.

Ошибка в одном бите исправляется, в двух и более – считается неисправимой, с сигнализацией по PERR# (фаза данных) или SERR# (фаза адреса или атрибутов).

Устройство PCI-X может не подавать PERR#, а попытаться исправить ошибку повтором транзакции, если это возможно.

# Контроль четности и ЕСС-контроль

Для контроля достоверности передачи информации на шине PCI применяется проверка четности адреса и данных;

в PCI-X используется и ЕСС-контроль с исправлением однобитных ошибок.

ЕСС-контроль обязателен при работе в Mode 2, он может использоваться и при работе в Mode 1. Метод контроля достоверности сообщается мостом в шаблоне инициализации по окончании аппаратного сброса шины. **Мост выбирает тот метод контроля, который поддерживают все абоненты его вторичной шины (и он сам).** Для сообщения об ошибках служат сигналы PERR# (протокольная сигнализация между устройствами) и SERR# (сигнал фатальной ошибки, вызывающий, как правило, немаскируемое прерывание системы).

При контроле четности используются сигналы PAR и PAR64, обеспечивающие четность числа «единиц» на наборах линий AD[31:0], C/BE[3:0]#, PAR и AD[63:32], C/BE[7:4]#, PAR64. Сигналы четности PAR и PAR64 вырабатываются тем устройством, которое в данный момент управляет шиной AD (выводит команду и адрес, атрибуты или данные). Сигналы PERR# и SERR# вырабатываются приемником информации в такте, следующем за тактом, в котором нарушено условие четности.

# Исправление ошибок

При ECC-контроле в 32-разрядном режиме для контроля линий AD[31:0] и C/BE[3:0]# применяется 7-битный код ECC с сигналами ECC[6:0]; в 64-разрядном режиме применяется 8-битный код с сигналами ECC[7:0]; в 16-разрядном режиме применяется несколько измененная схема ECC7 + 1.

**В любом из режимов ECC-контроль позволяет исправлять только одиночные ошибки и обнаруживать большинство ошибок с большей кратностью.**

Исправление ошибок может быть запрещено программно (через регистр управления ECC-контролем), при этом обнаруживаются все ошибки кратности 1, 2 и 3. В любом случае в регистрах ECC-контроля сохраняется диагностическая информация. Биты ECC выводятся на шину по тем же правилам и с теми же задержками, как и биты четности. Однако сигналы PERR# и SERR# вырабатываются приемником информации через такт после действительных бит ECC — «лишний» такт отдается на анализ синдрома ECC и попытку исправления ошибки.

# Исправление ошибок

Обнаруженная ошибка четности, как и ошибка более чем в одном бите, обнаруженная при ECC-контроле, считается неисправимой (unrecoverable).

Достоверность информации в фазе адреса, а в PCI-X и в фазе атрибутов, проверяется целевым устройством. В случае обнаружения неисправимой ошибки в этих фазах целевое устройство подает сигнал SERR# (в течение одного такта) и устанавливает в своем регистре состояния бит 14 — Signaled System Error.

В фазе данных достоверность проверяет устройство-приемник данных; в случае обнаружения неисправимой ошибки оно подает сигнал PERR# и устанавливает в своем регистре состояния бит 15 — Detected Parity Error.

# Электрический интерфейс PCI-X

Напряжение питания и уровни сигналов – 3.3 В, с возможностью работы на 1.5 В при поддержке режима Mode 2.

Щелевой разъем имеет ту же конфигурацию, но иное назначение некоторых контактов. В частности, добавлены сигналы ECC.

Режим работы шины определяется мостом по началу сигнала сброса (RST#).

Помимо линии M66EN, проверяется контакт PCIXCAP, по сопротивлению резистора определяется максимальная частота каждого устройства.

Проверив устройства, мост выбирает режим работы шины (частота, режим PCI-X или PCI, контроль достоверности данных) и сообщает его всем устройствам двоичным кодом по линиям PERR#, DEVSEL#, STOP# и TRDY# по спаду сигнала RST#.

# Заключение

PCI-X является существенно доработанной версией параллельной шины Peripheral Components Interconnect (PCI). Она построена на классической шинной топологии и требует для подключения большое число дорожек/контактов. Как мы уже упоминали выше, доступная пропускная способность разделяется между всеми устройствами.

Обратная совместимость является большим плюсом PCI-X.

Администраторы желают быть абсолютно уверенными, что новое оборудование будет работать правильно.