

Эволюция архитектуры персонального компьютера

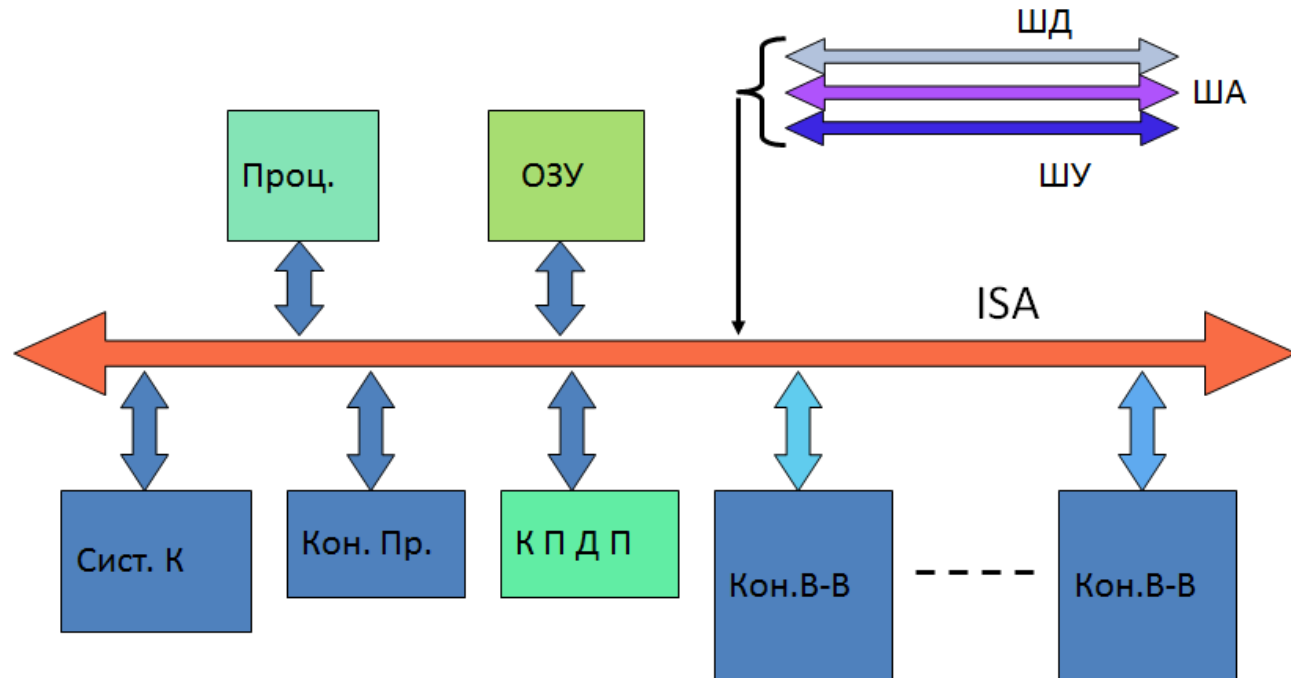
Первый этап

Параллельная шина ISA

Industry Standard Architecture

Одна общая системная шина

- **ISA** - *Industry Standard Architecture*,



- IBM PC/XT/
- Все устройства на **одной общей шине**
- **Два Хозяина шины – процессор и КПДП**
- **Трехшинная архитектура (адрес, данные, управления)**
- **Производительность 8 Мбайт/сек**
- 15 линий прерывания
- 7 линий ПДП

Второй этап

**Появление нескольких
параллельных шин**

Второй этап - появление нескольких шин

- Одна общая шина разделяется на несколько отдельных:
 - Системную шину
 - Шину ввода-вывода (шина PCI - Peripheral Component Interconnect)
 - Шину памяти
 - Шина графического адаптера

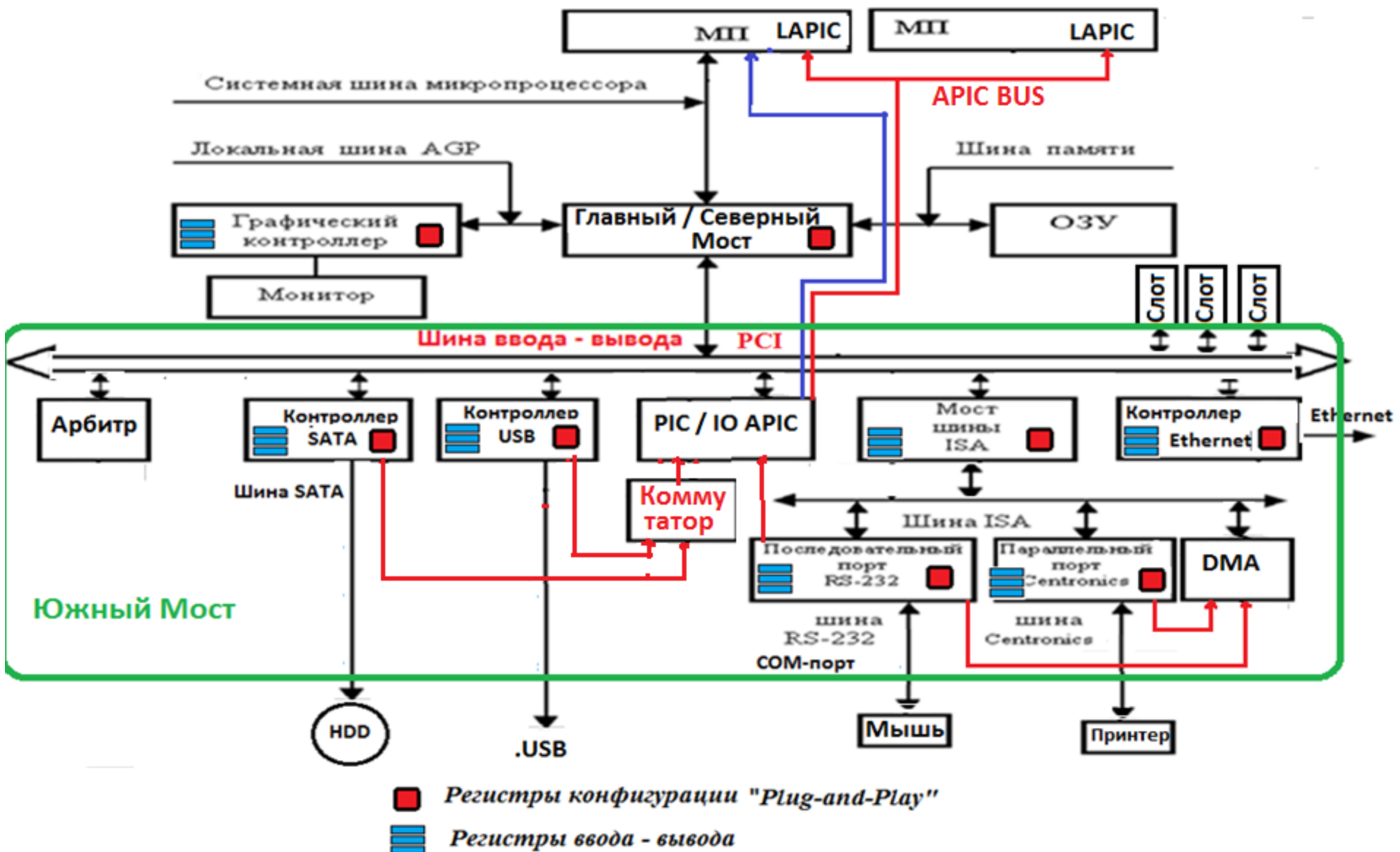
Все шины между собой объединены в одной микросхеме

«Главный мост / Северный мост»

Большинство контроллеров ввода-вывода стали помещать в отдельный чип – *Южный мост*

Мост - устройство объединяющее шины между собой

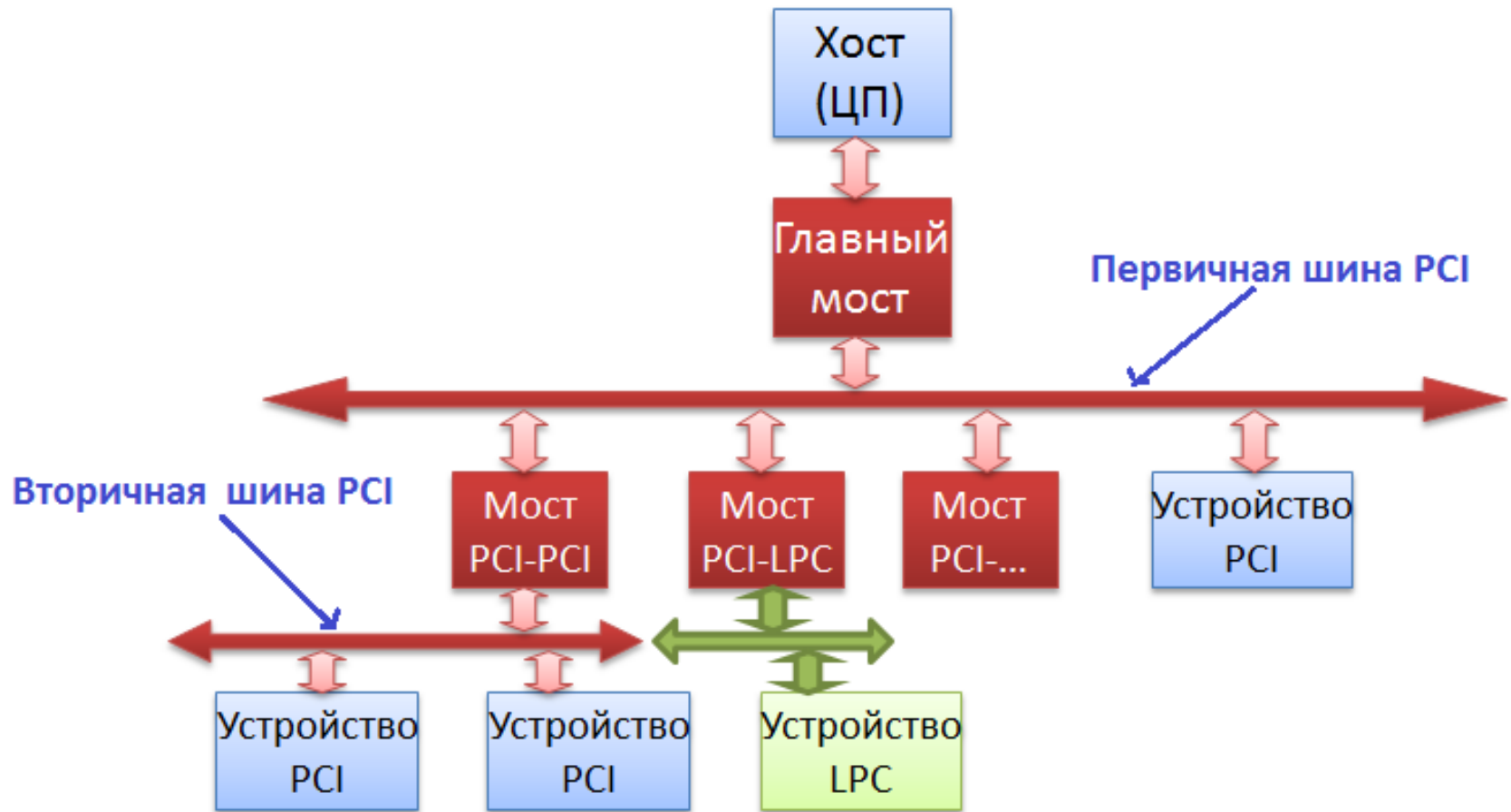
Шина PCI - Peripheral Component Interconnect



PCI – устройства : контроллеры и мосты

Производительность **133 Мбайт/сек** (32бит/33МГц) и **533 Мбайт/сек** (64бит/66МГц)

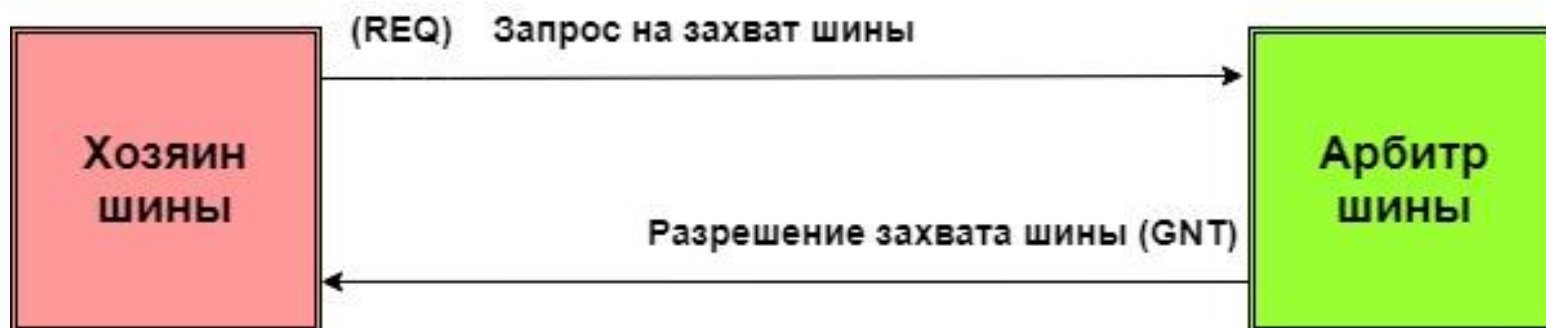
Топология



- Имеет топологию - **многоуровневая звезда**.
- Максимальное количество устройств подключаемых к одному сегменту шины теоретически **32**, реально не более **6**.

Шина PCI

- Все устройства равноправны, нет главного.
- *Bus Master (Хозяин шины)*
 - Любой контроллер или мост) на шине, взявший на себя временное управление шиной.
- *Bus Master* запрашивает арбитр своей шины для захвата шины и приемо/передачи по ней.



Память конфигурации устройства PCI

- У каждого контроллера устройства имеется поле из 256 байт.

31	24	23	16	15	8	7	0	
Device ID				Vendor ID				0x00
Status				Command				0x04
Class code						Revision ID		0x08
BIST		Header Type		Latency Timer		Cache Line Size		0x0C
Base Address Register (BAR) 1								0x10
Base Address Register (BAR) 2								0x14
Base Address Register (BAR) 3								0x18
Base Address Register (BAR) 4								0x1C
Base Address Register (BAR) 5								0x20
Base Address Register (BAR) 6								0x24
<u>Cardbus CIS Pointer</u>								0x28
Subsystem ID				Subsystem Vendor ID				0x2C
Expansion ROM Base Address								0x30
						Capabilities ptr		0x34
								0x38
Max_Lat		Min_Gnt		Interrupt Pin		Interrupt Line		0x3C

- Информация хранится 32 x – битными словами в формате **little endian**
 - Задаёт конфигурацию ресурсов контроллера (адреса портов, тип контроллера, выделенные линии прерывания и.т.д) во время процедуры POST

Организация ПДП

- Любой контроллер ввода-вывода став Хозяином шины может самостоятельно осуществлять прямой доступ к памяти (а также к другому контроллеру).
- **Контроллер ПДП** должен быть реализован в каждом контроллере, который хочет получить доступ по ПДП .
- Классический КПДП контроллер остался для поддержки legacy-устройств шины ISA.

Прерывания шины PCI

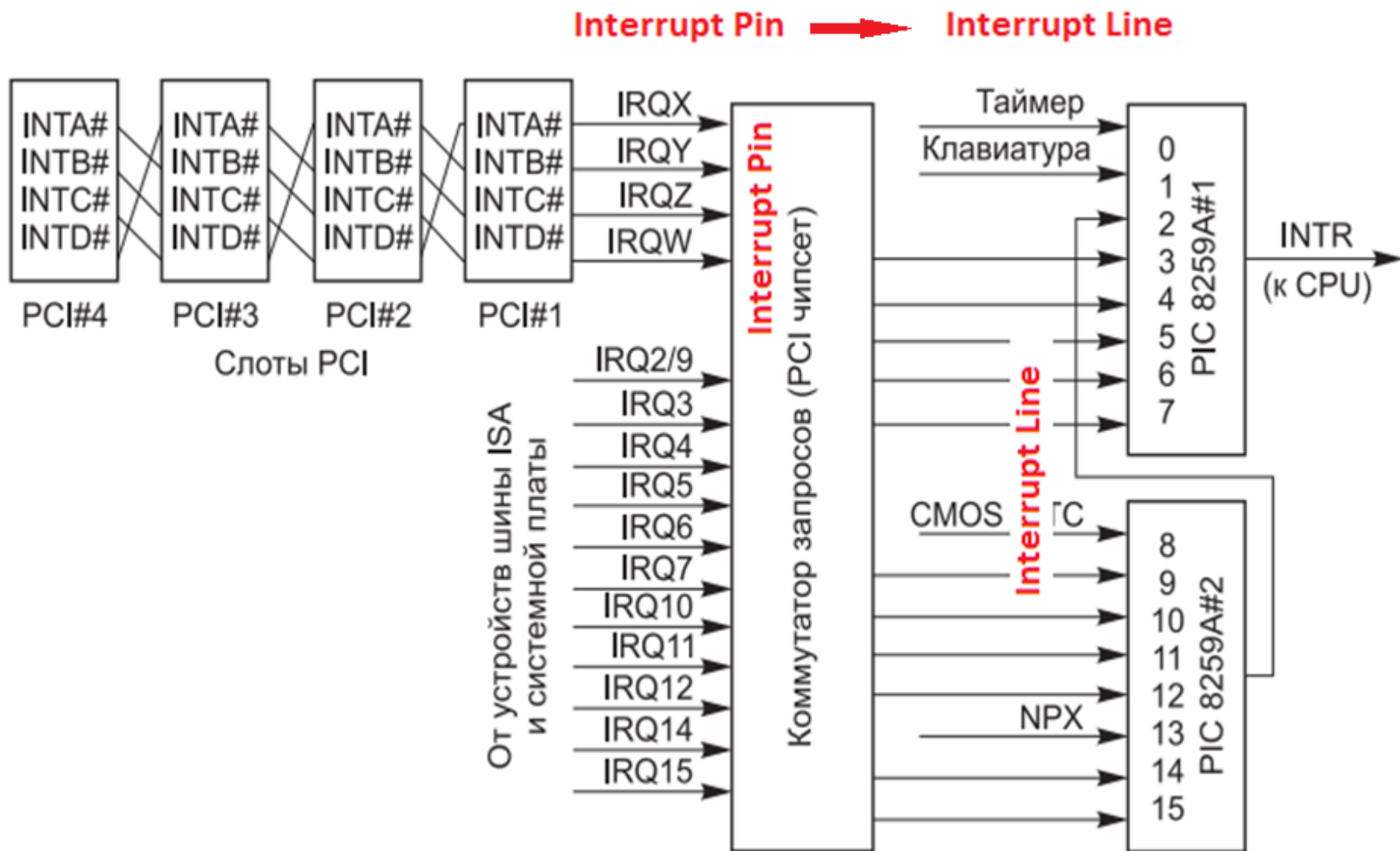
- PCI-устройства могут подавать сигнал запроса на прерывания контроллеру прерываний 2 способами:
 - 1. Проводная сигнализация по проводным линиям INT;
 - 2. Сигнализация с помощью сообщений MSI (Message Signaled Interrupts)

ПРОВОДНАЯ СИГНАЛИЗАЦИЯ ПРЕРЫВАНИЙ

PIC и APIC

- На шине PCI использовались два типа контроллеров прерываний:
 - **Простой PIC (Peripheral Interrupt Controller)**
 - Периферийный контроллер прерываний, программно совместимый с «историческим» контроллером 8259A;
 - **APIC (Advanced Peripheral Interrupt Controller)**
 - Усовершенствованный контроллер прерываний, введенный для поддержки мультипроцессорных (многоядерных) систем..

Простой контроллер PIC (самый первый вариант)



Простой контроллер PIC (самый первый вариант)

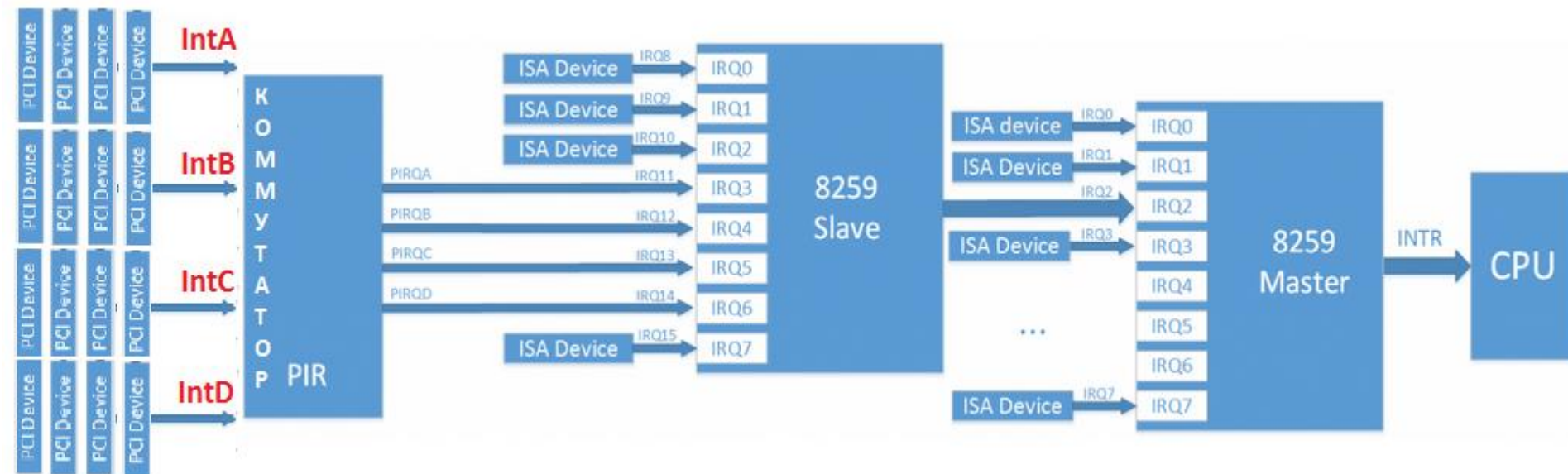
Контакты линий прерываний всех слотов соединили друг с другом с циклическим сдвигом и получили четыре разделяемых линии запроса прерываний Interrupt Pin : INT A, B, C, D.

Разделяемые линии A,D,C,D поступали на вход программно - управляемого коммутатора, где коммутировались в отдельные не разделяемые линии запроса IRQ 0 – IRQ 15 контроллера прерываний PIC

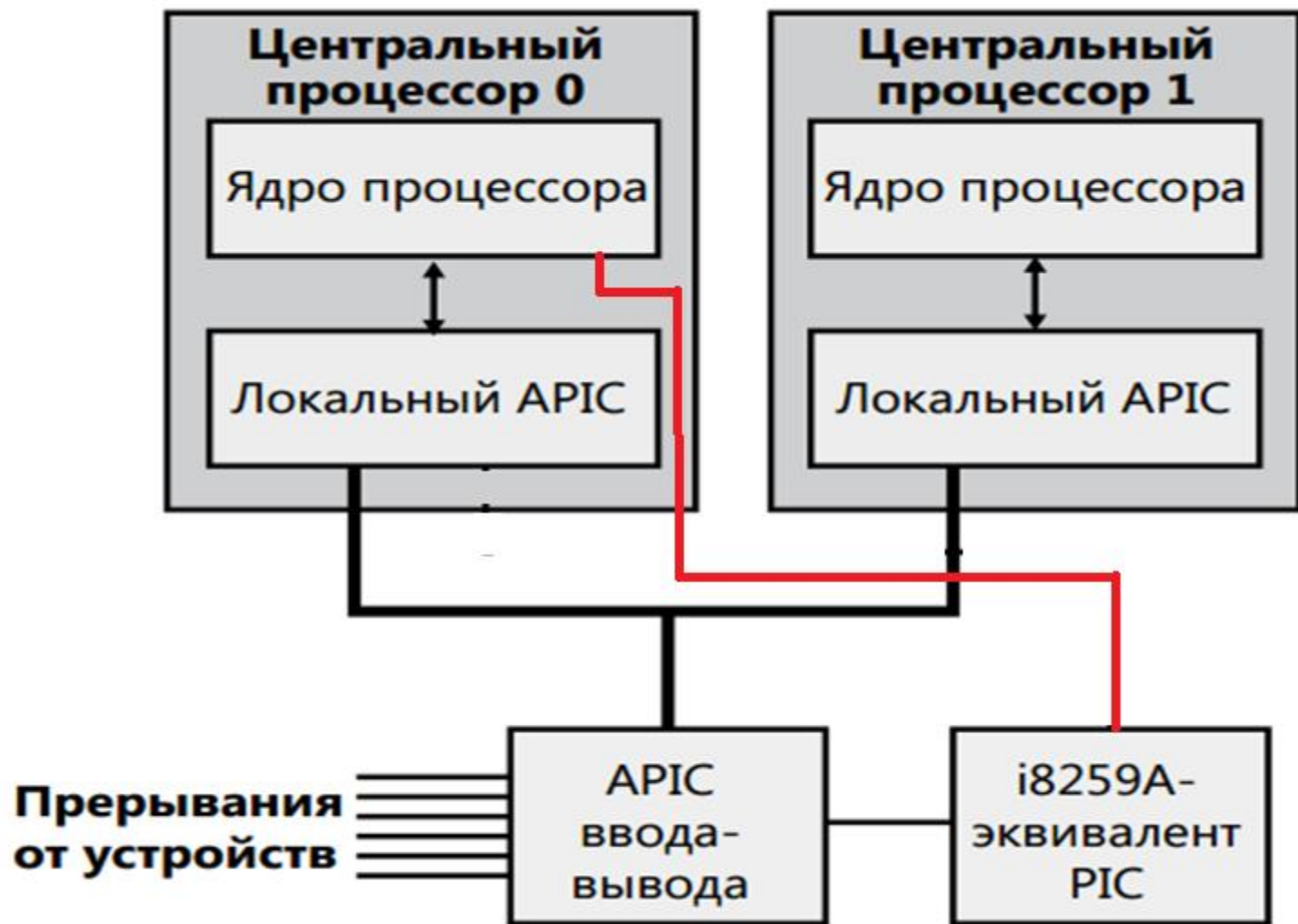
Так как PCI-устройство обычно использует только одно прерывание (один контакт (INTA)), повышаются шансы на получение не разделяемой линии прерывания Interrupt Pin : INT A.

Если PCI-устройство использует несколько прерываний , то вторая линия прерываний Interrupt Pin может быть разделяемая с другим PCI-устройством.

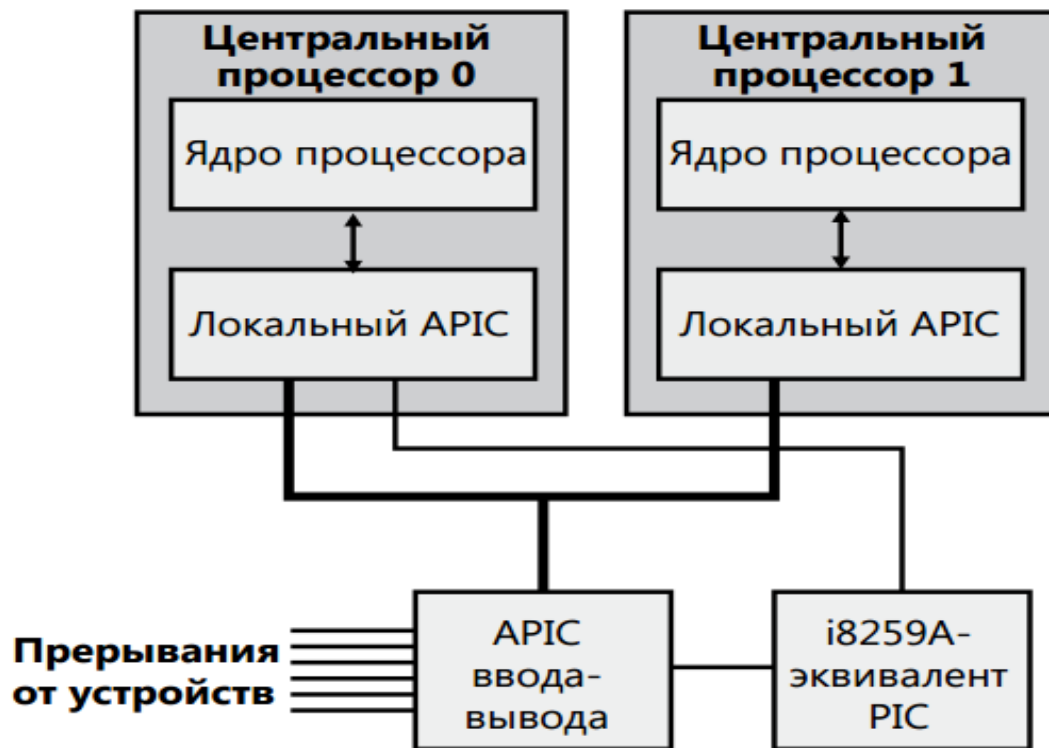
Простой спаренный PIC



Расширенный контроллер прерываний APIC



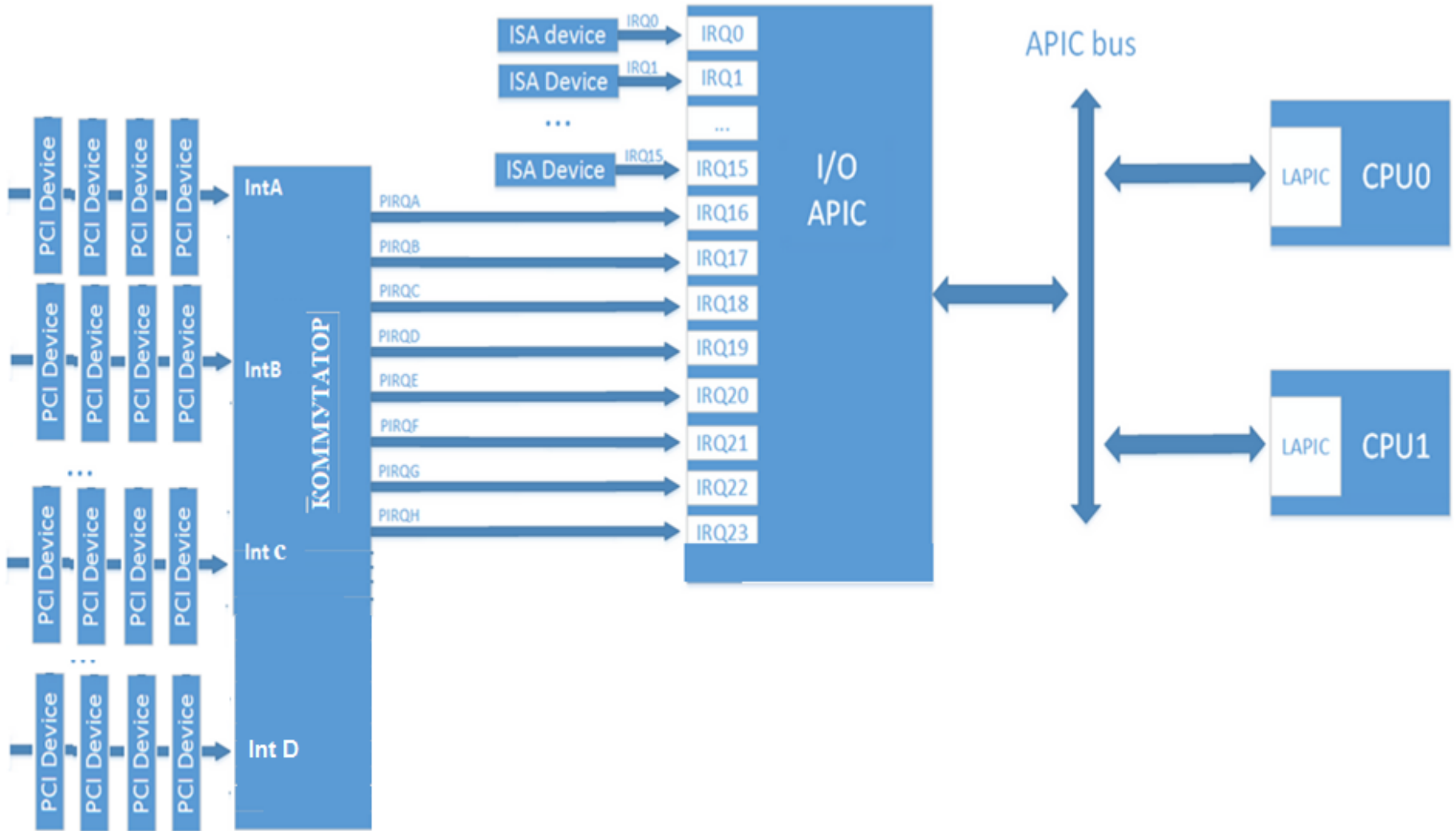
Расширенный контроллер прерываний APIC



Состоит из двух частей **I/O APIC** и **LAPIC**

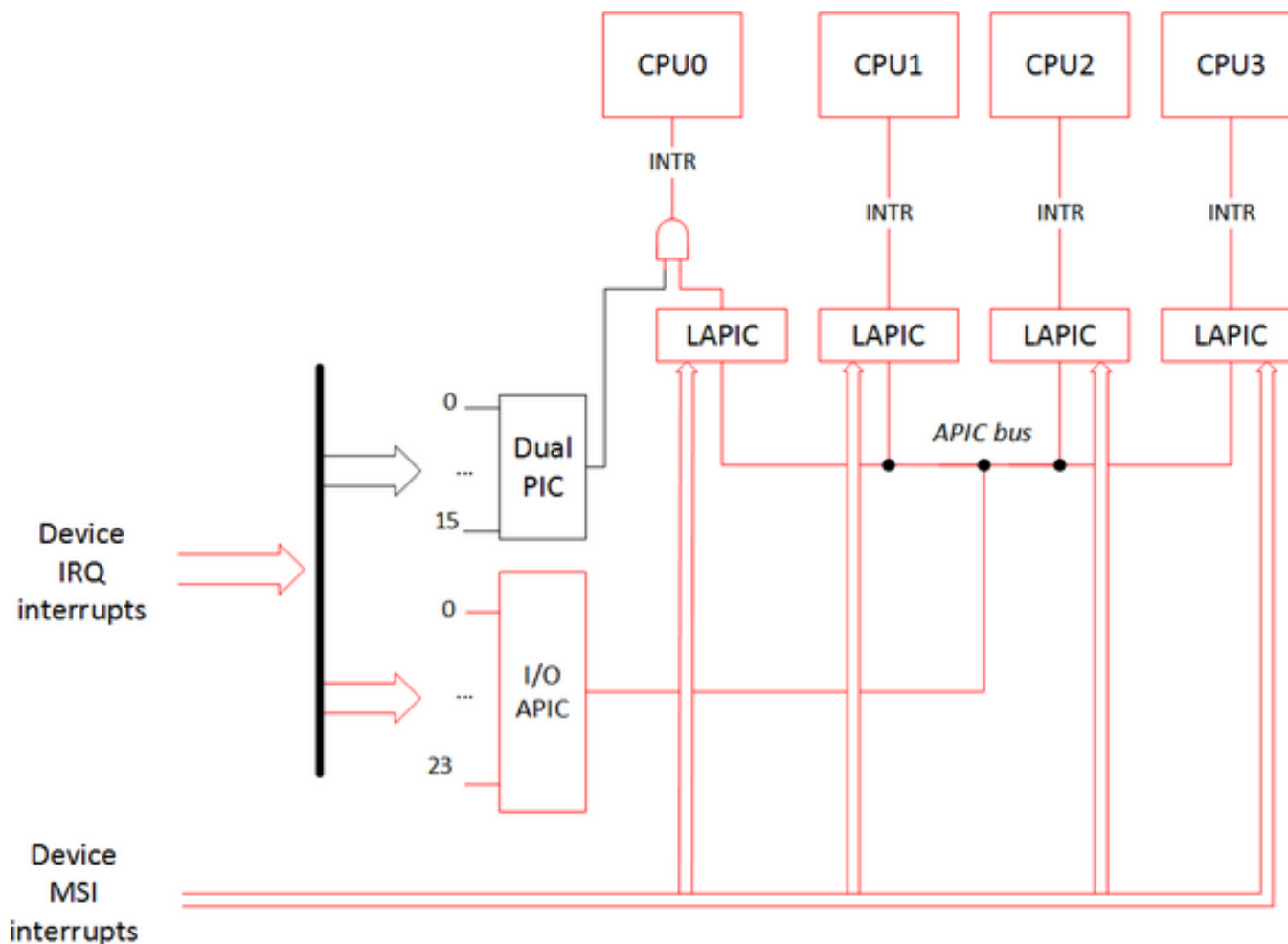
- **I/O APIC** - контроллер прерываний ввода-вывода
- **LAPIC** – локальный контроллер прерываний для каждого процессора (ядра).
- **I/O APIC** принимает запросы на прерывание и передает их соответствующему LAPIC по специальной шине

APIC



■ 23 линии

Различные типы прерываний



- Операционной системой могут поддерживаться различные типы доставки прерываний

СИГНАЛИЗАЦИЯ ПРЕРЫВАНИЙ С ПОМОЩЬЮ СООБЩЕНИЙ

Необходимость MSI

- Из-за малого количества линий запроса прерываний Interrupt PIN, при большом количестве контроллеров несколько из них могут оказаться **на одной разделяемой линии запроса** .
- Для определения источника разделяемых прерываний необходимо вызвать обработчик прерывания (*ISR*, Interrupt Service Routine), входящий в драйвер данного устройства.
- Для диспетчеризации разделяемых прерываний, ОС выстраивает их обработчики (*ISR*-ы) в цепочку и последовательно запускает их. Каждый *ISR* должен опрашивать флаг наличия прерывания в регистре состояния своего контроллера. В результате увеличилось время обработки прерываний.
- Для решения проблемы разделяемых прерываний решили перейти от запроса прерываний по физическим линиям на запросы прерываний с помощью сообщений.

Прерывания сообщением (MSI) - идея

- MSI – (*Message Signaled Interrupts*)
- Решат проблему разделяемых прерываний
- Устройство, захватившее шину (Хозяин шины) **посылает сообщение о прерывании по той же шине, что и другие данные, в специальный регистр контролера (/OAPIC)**
- **Адрес регистра (FEC0 0000) и формат сообщения** задается в пространстве конфигурации PCI-устройства при инициализации в режиме **POST**
- **I/OAPIC передает запрос на прерывание LAPIC конкретного процессора, после чего начинается процедура обработки прерываний.**
- Устройство, использующему MSI запрещается использовать проводные линии запроса INTx.

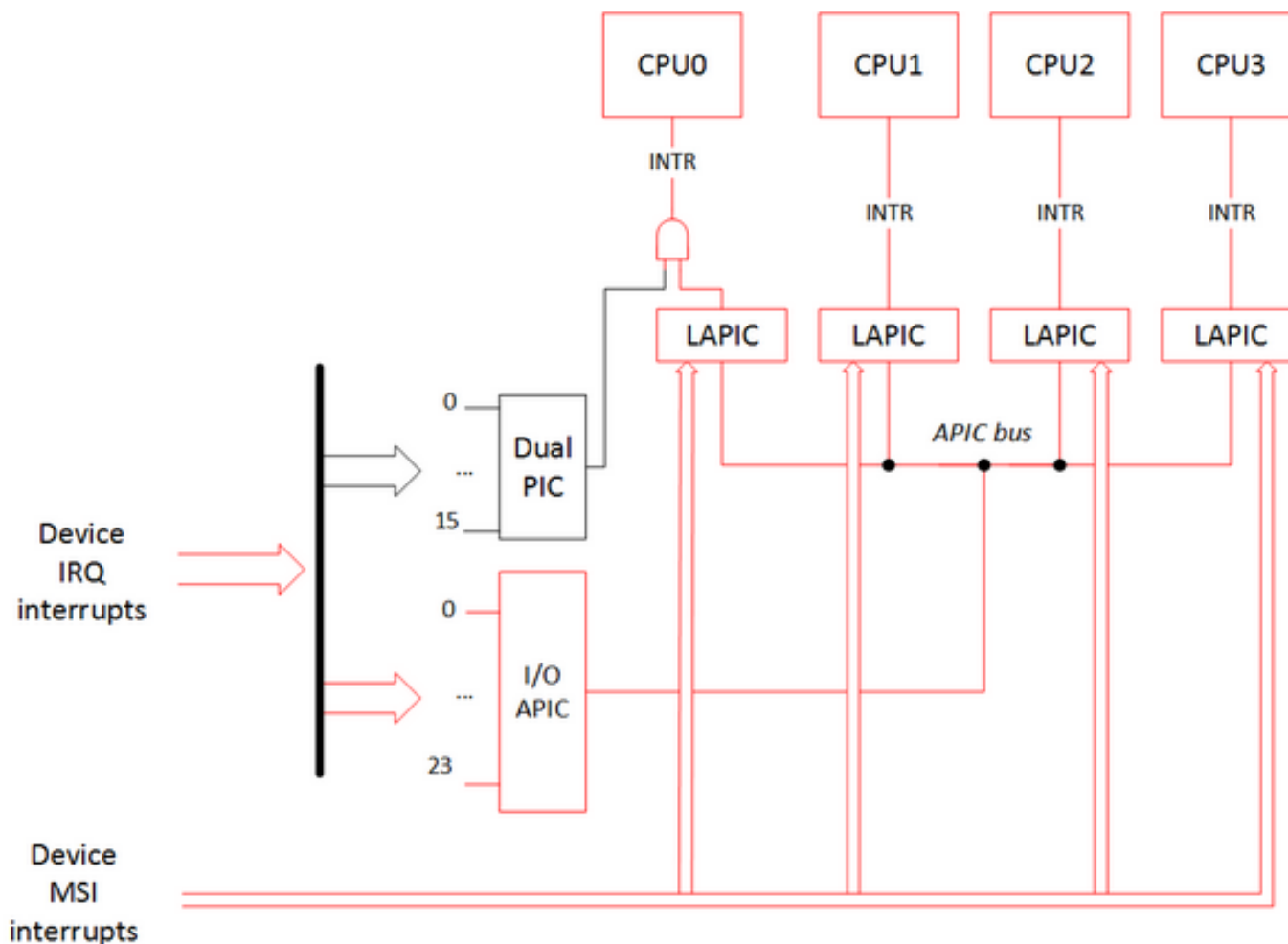
Регистр MSI APIC

Непосредственно адресуемые регистры APIC

Адрес	Размер, бит	Тип	Назначение
FECO_0000h	8	R/W	Index Register, индекс для доступа к косвенно адресуемым регистрам
FECO_0010h	32	R/W	Data Register, данные для обращений к косвенно адресуемым регистрам
FECO_0020h	8	WO	IRQ Pin Assertion Register, регистр программной установки запросов прерываний (запись числа 0–23 эквивалентна подаче сигнала на соответствующий вход INTINn)
FECO_0040h	8	WO	EOI Register — регистр завершения прерываний для входов, чувствительных к уровню. Запись байта — вектора прерывания — вызывает сброс бита Remote_IRR для всех входов, которым назначен данный вектор (аналогичное действие IOAPIC выполняет по сообщению EOI, полученному по локальной шине)

- Адреса регистров APIC **отражены на память**
- Регистр с адресом **FECO_0020h** APIC служит для приема запросов прерываний в виде сообщений

Различные типы прерываний



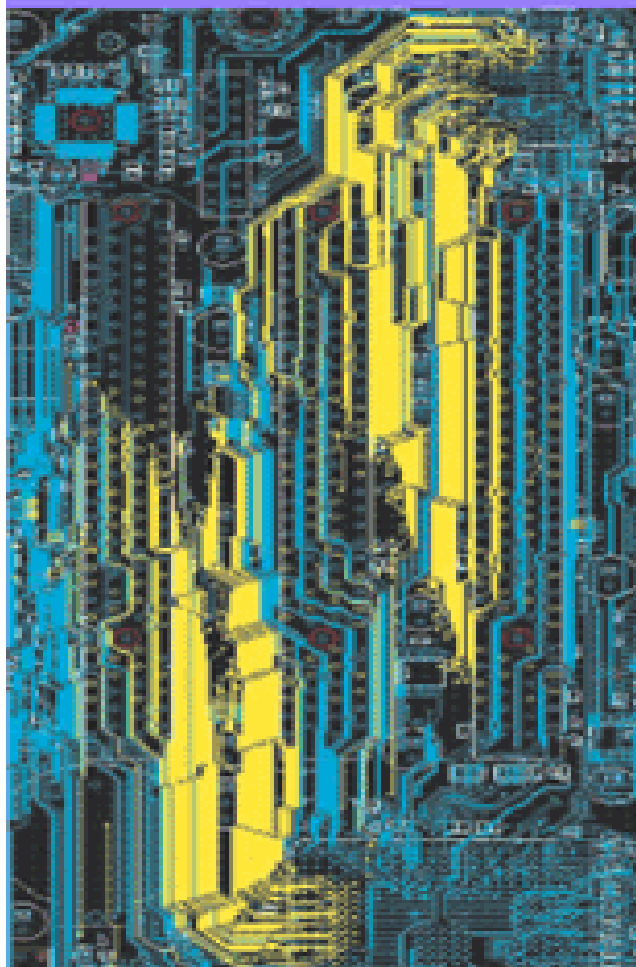
- Операционной системой могут поддерживаться различные типы доставки прерываний

Ограничения параллельных шин

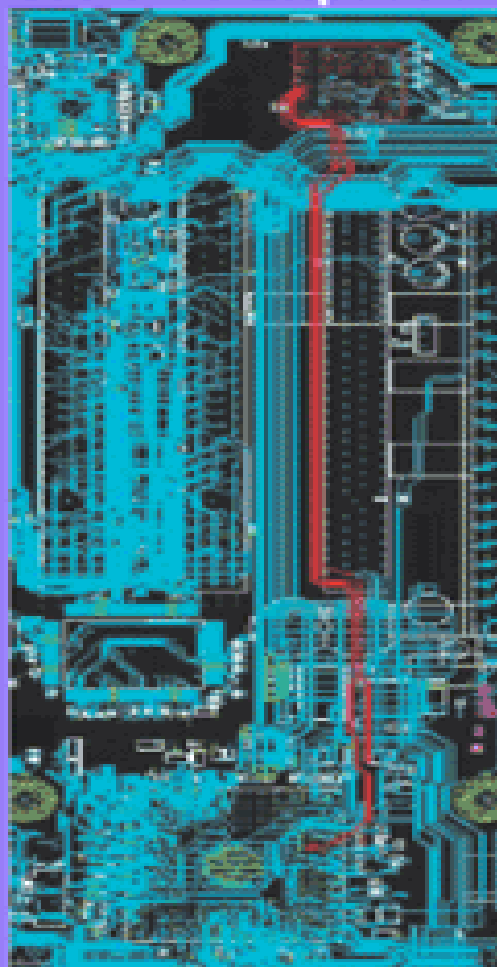


Сравнительный анализ шин

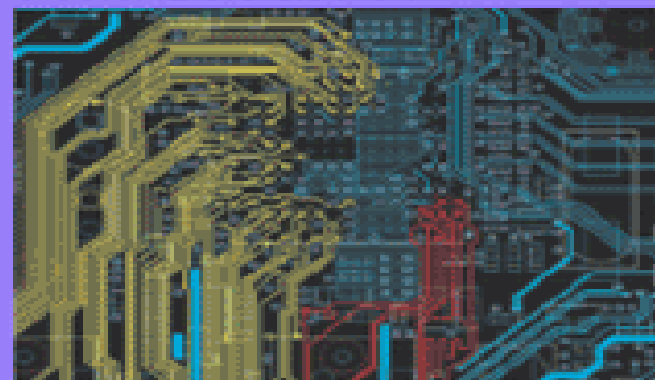
32-bit PCI



X1 PCI Express



Южный мост

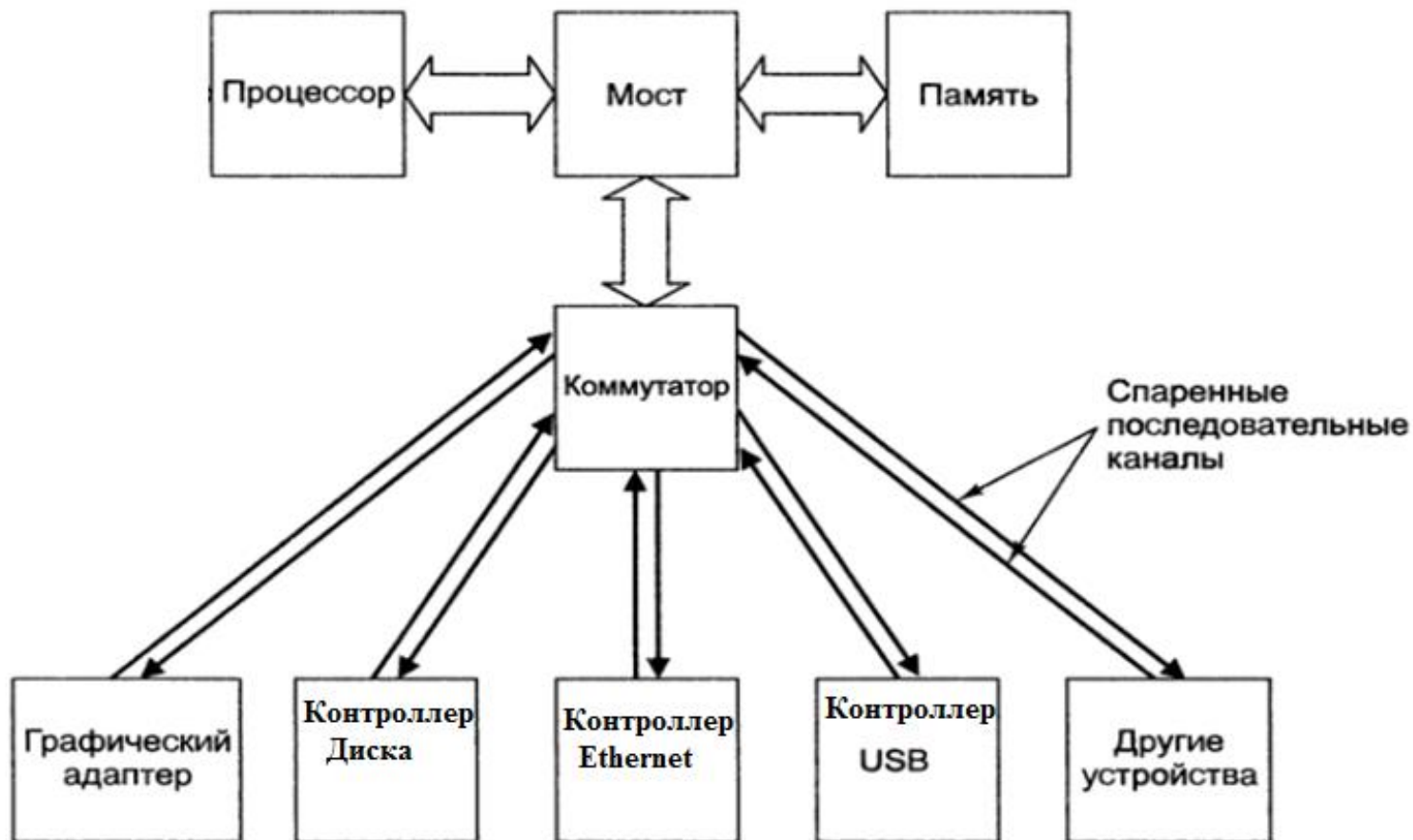


- трассировка PCI
- трассировка PCI Express

Третий этап

Последовательно-параллельная шина
PCI-Express

Архитектура шины (общий вариант)



PCI Express

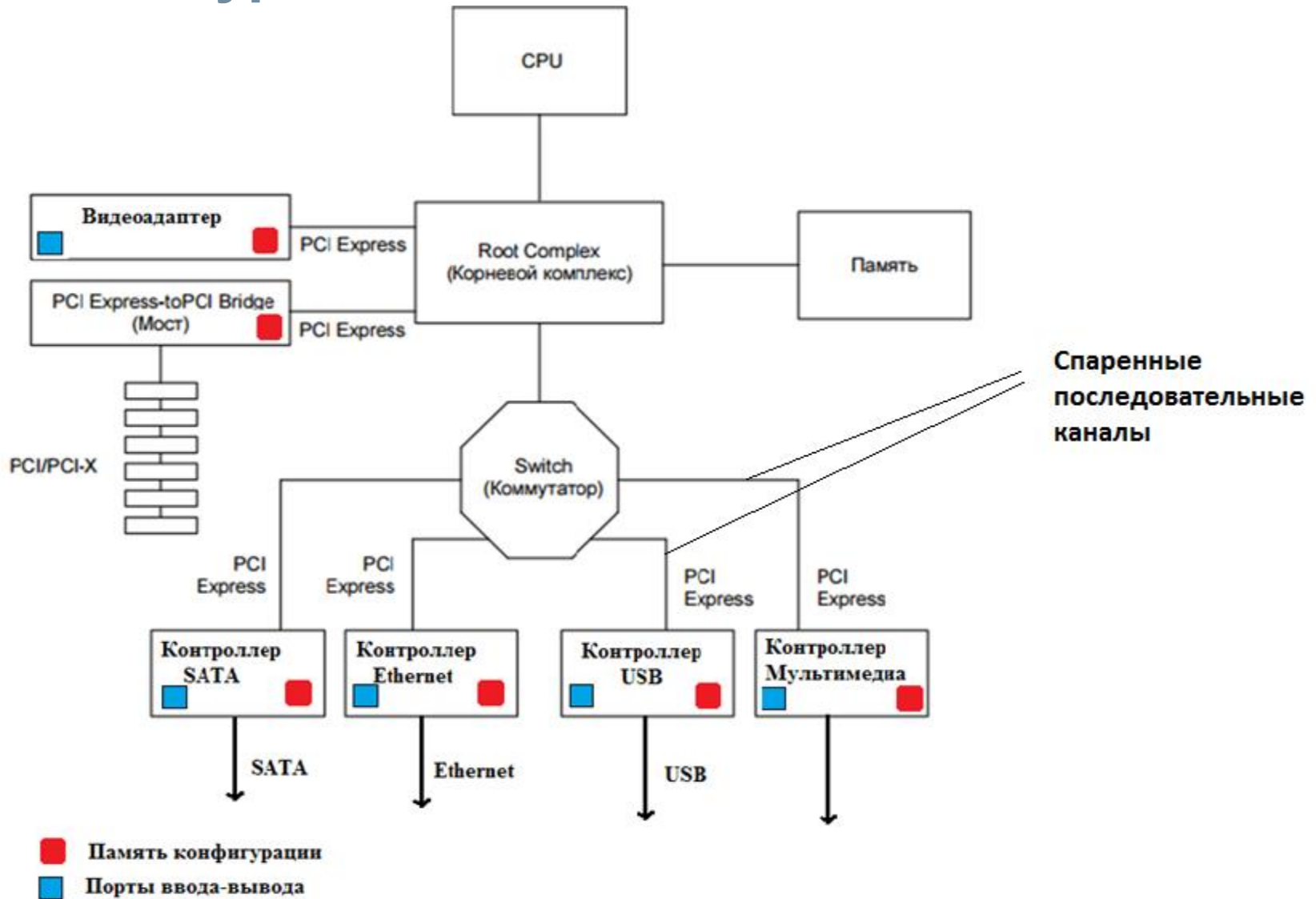
- Последовательно - параллельная шина - совокупность независимых последовательных каналов работающих параллельно.
- Данные и вся управляющая информация передается пакетами
- Стандартизированы 1, 2, 4, 8, 12,16 и 32 канальные варианты.
- Каждый канал является дуплексным и состоит из двух дифференциальных пар проводников (необходимо только 4 контакта) работающих в режиме дуплекса.
- Основана на идеологии многоуровневой сетевой модели OSI.
- Шина PCI Express - локальная сеть в пределах компьютера,напоминающая Gigabit Ethernet.

PCI Express

Пропускная способность PCI Express, Гбайт/с

Год выпуска	Версия PCI Express	Кодирование	Скорость передачи	Пропускная способность на x линий				
				x1	x2	x4	x8	x16
2002	1.0	8b/10b	2,5 ГТ/с	250 Мбайт/с	0.50 Гбайт/с	1.0 Гбайт/с	2.0 Гбайт/с	4.0 Гбайт/с
2007	2.0	8b/10b	5 ГТ/с	500 Мбайт/с	1.0 Гбайт/с	2.0 Гбайт/с	4.0 Гбайт/с	8.0 Гбайт/с
2010	3.0	128b/130b	8 ГТ/с	984.6 Мбайт/с	1.97 Гбайт/с	3.94 Гбайт/с	7.88 Гбайт/с	15.8 Гбайт/с
2017	4.0	128b/130b	16 ГТ/с	1.969 Гбайт/с	3.94 Гбайт/с	7.88 Гбайт/с	15.75 Гбайт/с	31.5 Гбайт/с
2019	5.0	128b/130b	32 ГТ/с	3.938 Гбайт/с	7.88 Гбайт/с	15.75 Гбайт/с	31.51 Гбайт/с	63.0 Гбайт/с
2021	6.0	128b/130b	64 ГТ/с	8 Гбайт/с	16 Гбайт/с	32 Гбайт/с	64 Гбайт/с	128.0 Гбайт/с

Архитектура шины



Корневой комплекс

(**Root Complex** или **корневой комплекс**) отвечает за связь с процессором и системной памятью, а также за управление и конфигурирование всей шины PCI Express .

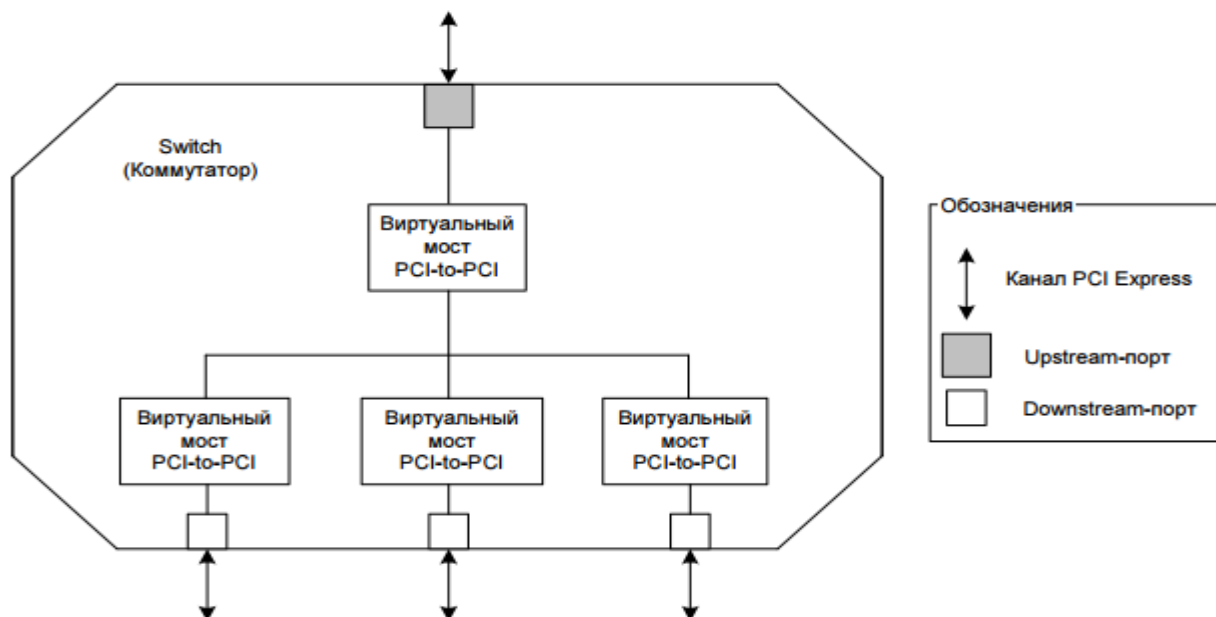
Root Complex содержит несколько портов PCI Express

К каждому порту может подключаться коммутатор (switch), или PCIExpress - устройство (Endpoint).

Root Complex отвечает за доступ к пространству конфигурации, портам и пространству памяти.

В настоящее время входит в состав северного моста

Коммутатор PCI Express

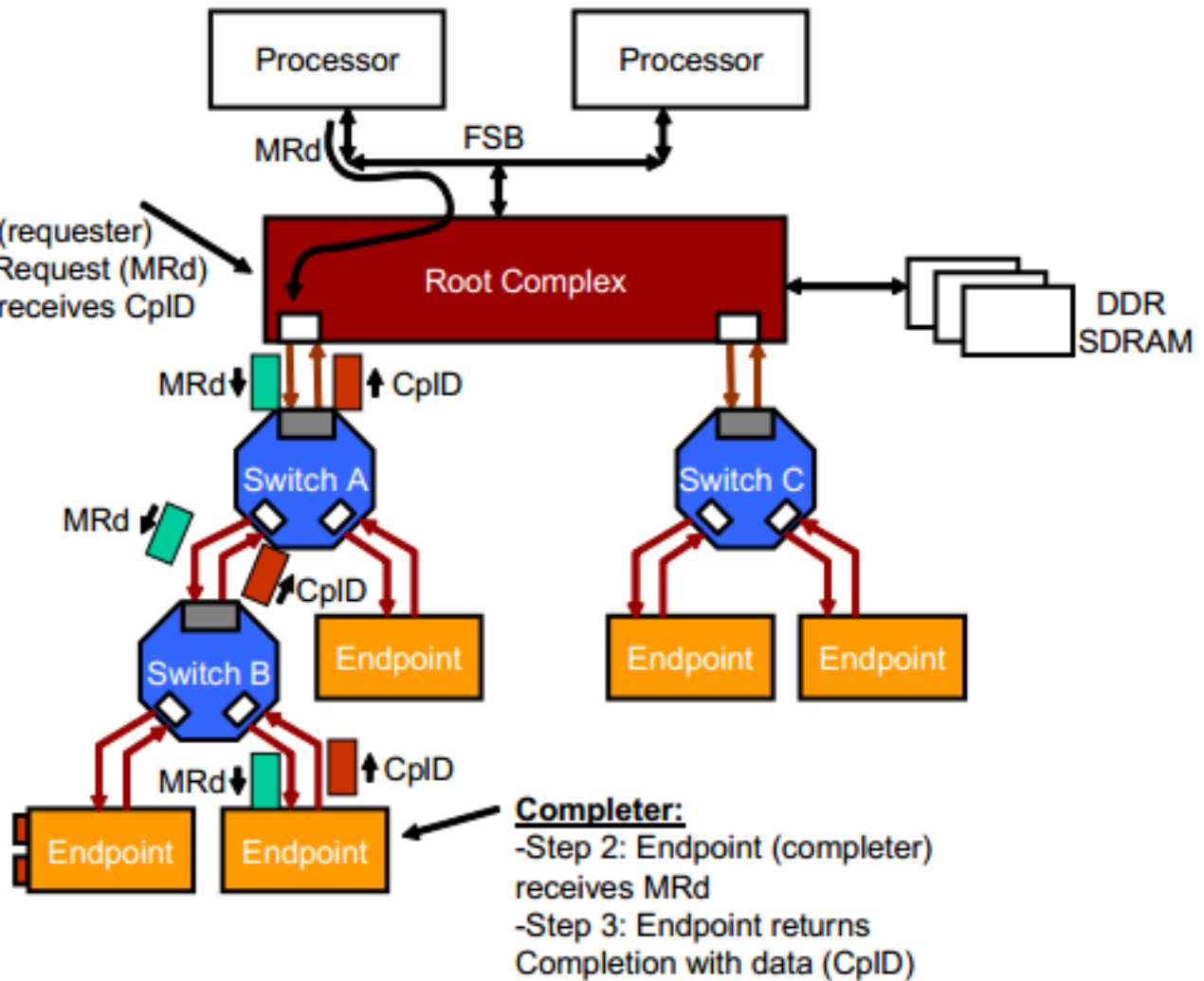


- Коммутатор передает пакеты с одного порта на другой и служит для расширения количества подключаемых устройств
- Программно коммутатор представляет собой **набор PCI - мостов, т.е. каждая шина PCI-Express представляется как виртуальный мост PCI.**

Передача пакета

Requester:

- Step 1: Root Complex (requester) initiates Memory Read Request (MRd)
- Step 4: Root Complex receives CplD

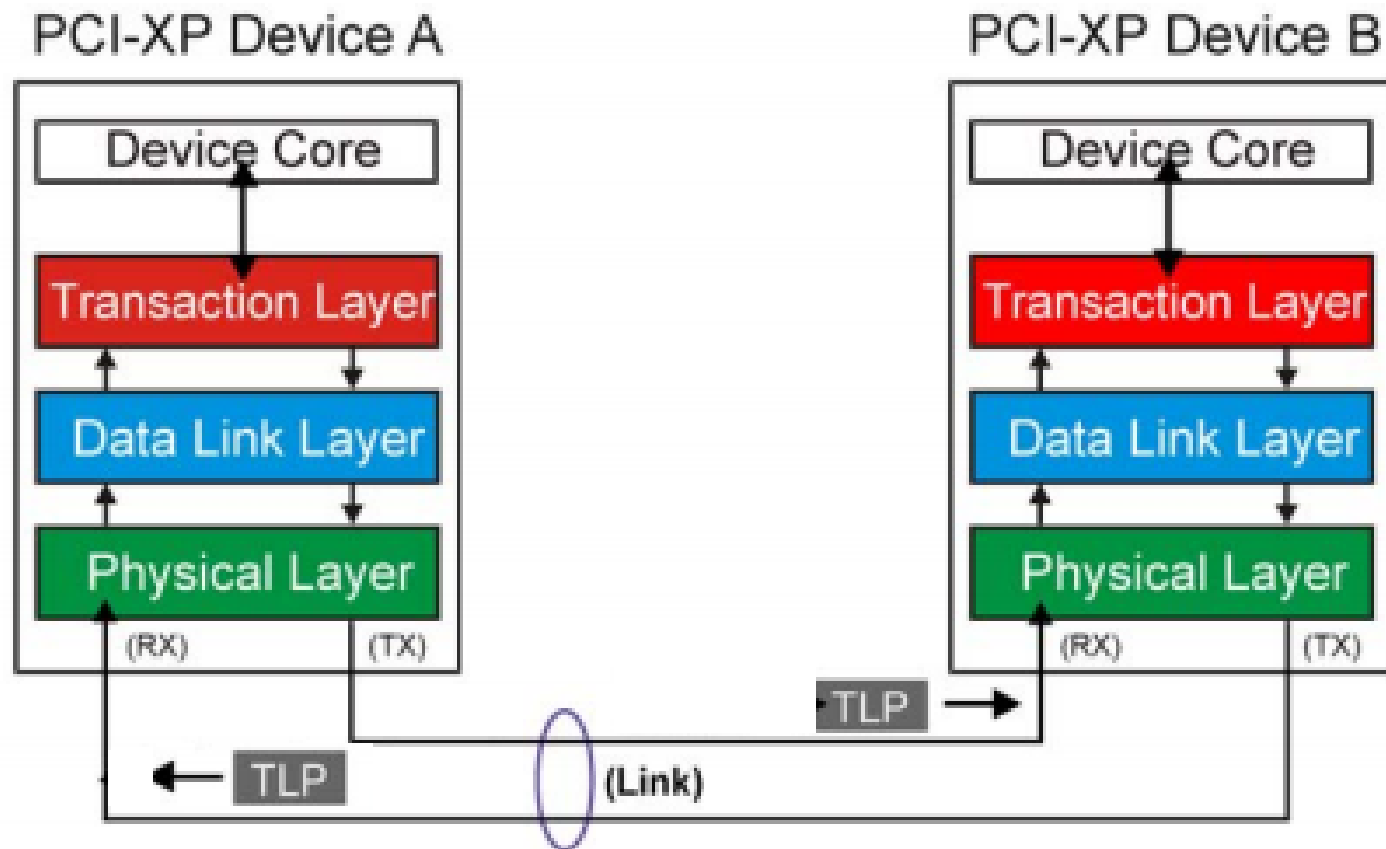


Многоуровневая модель шины PCI Express



Каждый уровень добавляет к данным свой заголовок

Передача /прием пакетов



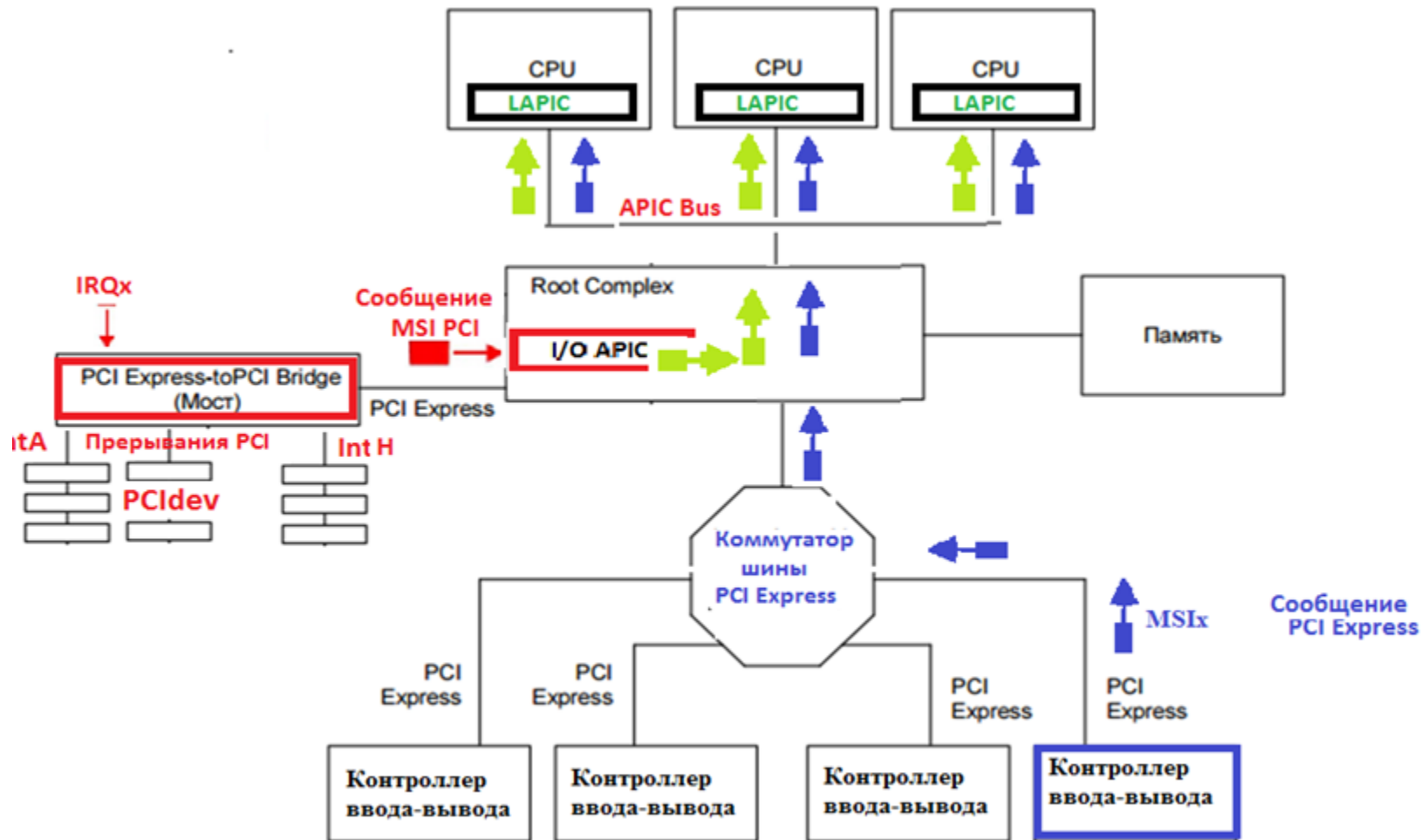
Формат пакетов шины PCI Express



- **Кадр** — начальная и конечная последовательность бит необходимых для распознавания начала и конца каждого пакета.
- **Порядковый номер** — добавляется на канальном уровне чтобы пакеты направленные одному получателю можно было собрать в нужной последовательности;
- **Заголовок** — описывает тип пакета, адрес получателя, приоритет и другие свойства
- **Данные** — данные пакеты;
- **CRC** — контрольная сумма пакета.

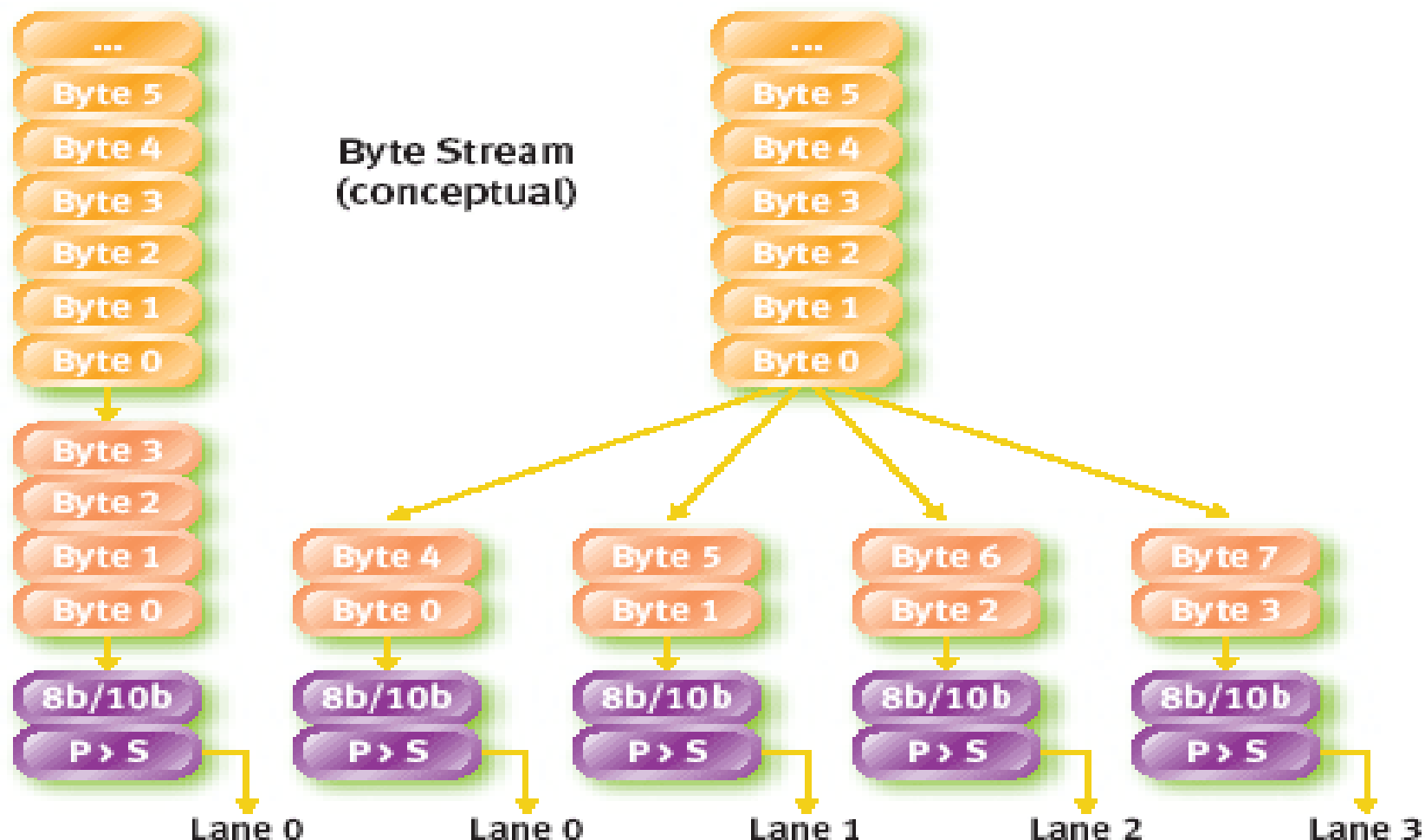
Прерывания PCI Express

MSI - (*Message Signaled Interrupts*)

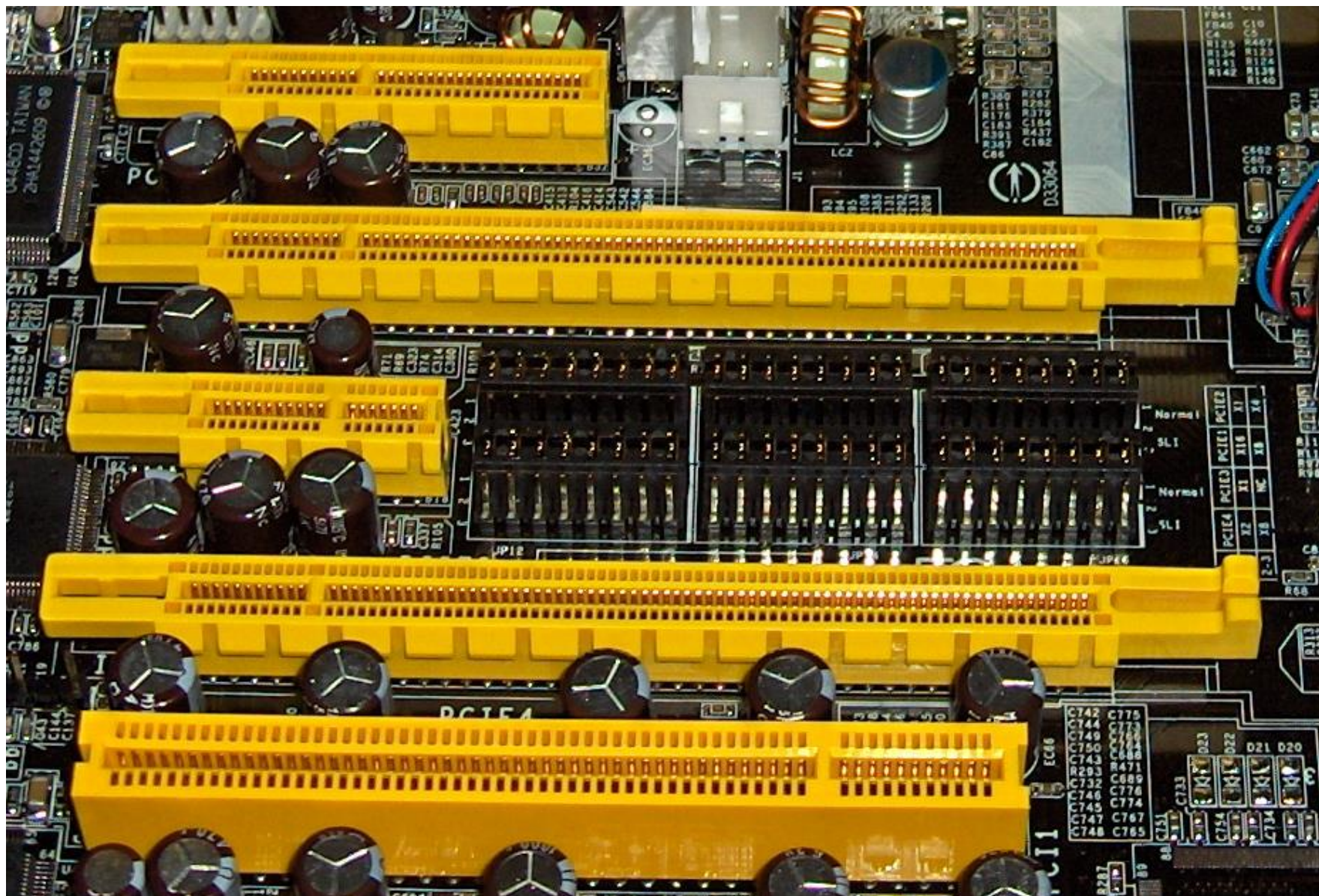


Для запроса прерывания контроллеру надо послать специальный пакет на адрес регистра в LAPIC (MSIx) или I/O APIC (MSI)

Масштабируемость шины PCI Express



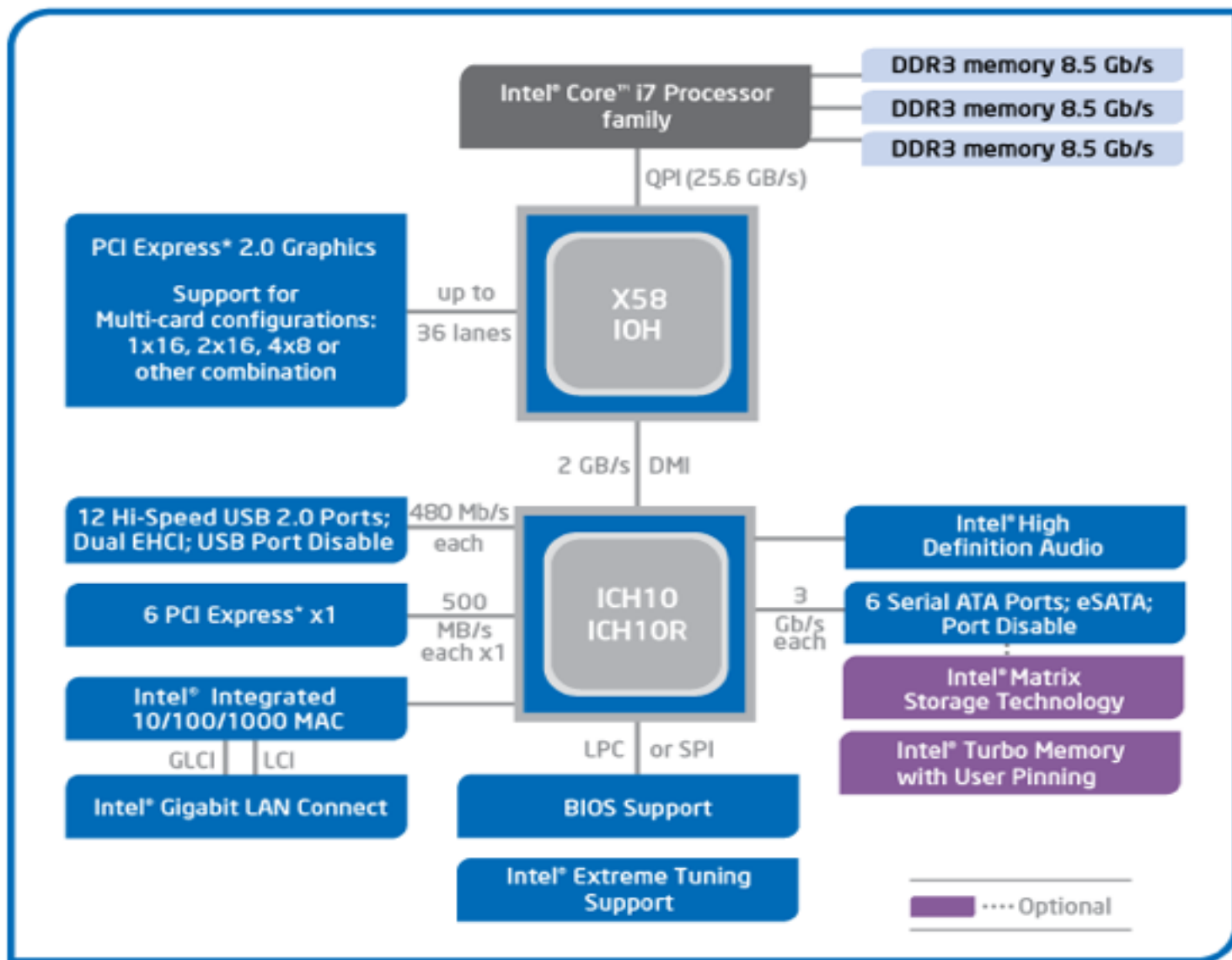
Последовательный поток байт может передаваться параллельно по нескольким каналам (до 32 каналов)



Чипсет

- **Чипсет** - набор микросхем которые обеспечивают взаимодействие между процессором, памятью, и устройства ввода-вывода.
- Состав - два чипа:
 - **северным мост**(Northbridge) или MCH (Memory Controller Hub)
 - **южный мост** (Southbridge) или ICH (*I/O Controller Hub,)*

АРХИТЕКТУРА НА ОСНОВЕ ШИНЫ PCI Express



Чипсет для процессора Intel Core i9

