

4 Jan, 2008

The Bus (PCI and PCI-Express)

The CPU, memory, disks, and all the other devices in a computer have to be able to communicate and exchange data. The technology that connects them is called the "Bus".

In the first IBM PC, the Bus was just a set of wires that ran through the mainboard. Everything connected to this one Bus and operated at one clock speed. Soon, however, the CPU had to run much faster than anything else, and the memory had to be faster than any I/O device.

In a modern PC, the CPU, memory, and video card are connected to a high speed control chip called the "Northbridge". Each device runs at its own speed:

- The mainboard delivers a clock signal to the CPU at speeds of 200, 266, or 333 Megahertz (ticks per second). The CPU typically transfers data four times a second, so the effective rate of CPU to Memory transfer is typically 800, 1066, or 1333 MHz and is called the Front Side Bus (FSB). Internally the CPU chops each mainboard clock tick into smaller units to generate an internal clock. This is the "multiplier". A multiplier of 10x turns the 200 MHz mainboard clock into an internal 2 GHz clock, which is typically the quoted speed of the CPU.
- The mainboard also gives a clock pulse to the memory. Because the CPU transfers data to the memory four times for every CPU clock tick, the clock speed sent to the memory is synchronized with the clock rate sent to the memory, but they are not the same value. The CPU sends 64 bits at a time, while the memory sends and receives 128 bits at a time. However, while the CPU transfers data four times per clock tick, the memory transfers data only twice per clock tick (Double Data Rate or "DDR"). Not surprisingly, memory vendors describe DDR memory as running at twice the actual clock signal. DDR memory that uses a 400 MHz clock is quoted as 800 MHz memory. However, this can be misleading. Memory has latency (a delay after the request before the data can be transferred). You may double the speed of the clock, but the latency remains the same and performance doesn't improve much.
- Modern video adapter cards plug into a PCI-Express slot with 16 lines of data transfer each of which can transfer data at 250 megabytes per second. That is an aggregate of 4 Gigabytes per second, but no modern video card actually needs that kind of data transfer capability.

The Northbridge is then connected to a second slower control chip called the "Southbridge" that supports all the other devices.

- The old "Parallel" ATA (IDE) transfers data down the flat ribbon cable at 100 Megabytes per second (Mb/s).
- The new Serial ATA transfers data at 150 or 300 Mb/s. This higher speed is misleading, because no desktop hard drive can actually read or write data faster than 40 Mb/s, so even the slower SATA is faster than is necessary.
- The standard PCI bus transfers four bytes of data with every tick of a 33 MHz clock. It therefore also has a rate of 133 Mb/s.
- Each line of a PCI-Express bus transfers data at 250 Mb/s. PCI-e slots on the mainboard can have 1, 2, 4, 8, or 16 lines that combine their data transfer capability. However, just one PCI-e line is faster than the old PCI bus, and only powerful and expensive RAID disk adapters typically need more than one PCI-e line.
- The USB port running in high speed mode transfers 480 megabits per second. That is 60 megabytes per second.

The AMD CPU has a simpler architecture. The memory connects directly to the CPU chip, and a slightly more powerful support chip can add video support to the usual Southbridge functions.

A car drives down the local streets at 25 miles per hour. Then it turns onto a highway ramp and accelerates to 55. Is there one road system, or two? The important thing is that there is a connection that allows a flow of traffic between the two speed zones. So data may flow from the CPU to the Northbridge at 6400 megabytes/second and then queue up to flow down the PCI bus at 133 megabytes/second. The effective data rate will be the slowest bus speed, but data can flow from any device to any other device.

PCI

Even a small modern mainboard can include integrated support for video, 6 channel high definition audio, gigabit Ethernet, lots of USB slots, plus all the usual devices. If you need a different type of audio support, a TV tuner to record programs, a document scanner, or most devices, the 60 megabyte per second transfer rate of USB 2 is more than fast enough.

However, if you need extra gigabit Ethernet, or a RAID disk adapter, or a second video card, then USB is not fast enough. You need to plug these devices into an I/O bus. For the last fifteen years, that bus has been some form of PCI.

Classic PCI

Depending on their size, every desktop computer mainboard has one to five 32 bit PCI adapter slots. This traditional PCI bus transfers 4 bytes of data with every tick of a 33 MHz clock, producing an aggregate bandwidth of 133 megabytes per second. However, this data transfer has to be shared among all of the devices in all of the slots.

There are a limited number of "interrupt levels" available in a PC. One interrupt level can be shared by two or more devices. Each PCI slot is assigned to an interrupt level, and on most mainboards that interrupt level is already in use by some device built into the mainboard (video, disk controller, audio, Ethernet, etc.). This isn't necessarily a problem, but in practice some adapter cards that you buy don't get along well with specific mainboard devices. If a new PCI card is not working correctly, and there is another free slot in the computer, moving the card may cause the problem to go away. This does not mean there is something wrong with the slot. It is probably just an incompatibility with the device that shares the interrupt level with the slot.

Server PCI [becoming obsolete]

If you need to support two or more gigabit Ethernet cards, or a big RAID adapter, or a FiberChannel interface, then 133 megabytes per second is not enough bandwidth.

Server computers have addressed this problem with expanded versions of the PCI bus. First, the size of the connection between the adapter and the slot can be expanded to support 64 bit data transfer. Then the speed of the bus can be increased to 66, 100, or even 133 MHz. Twice the data at four times the data transfer up to just over a gigabyte per second.

However, you will only find these types of PCI slots on expensive Server computers, and inexpensive mainboards have PCI-e slots that are even faster.

PCI-Express (PCI-e)

[PCI-Express](#) is a new high speed I/O bus.

The old PCI bus starts on the Mainboard at the Southbridge chip. Each of the 32 bits of data is represented by a single wire that runs through each PCI socket. The same 32 wires are used to send data from memory to the devices and from the devices to memory.

It is much easier to change hardware than software, and you don't want to have to wait for a new version of Windows or of the Linux Kernel. So the first requirement of a new bus is that it appear to the OS to be exactly the same as good old PCI. It is not necessary to change a single line of code in the OS or any device driver.

Now if you are going to have a bus that runs at a much higher speed than old PCI, you will have to connect it to the higher speed Northbridge rather than the low speed Southbridge chip on the Mainboard. However, this is not a big problem because the new bus will, among other things, replace the old AGP video interface that the Northbridge chip used to generate.

Now lets build the PCI-Express bus up component by component.

Pair: Each bit of data will be carried on a pair of wires instead of on a single wire. Balanced signals mean that you can run a much higher clock speed with a much lower voltage. The old PCI bus ran at 33 MHz on a desktop and maxed out at 133 MHz on exotic servers. The PCI Express bus runs at 2.5 GHz, but to detect errors and provide timing, it takes 10 clock ticks to transmit an 8 bit byte. Therefore, the pair of wires transmits 250 megabytes per second. PCI-e 2.0 allows transmission at double that rate, 5 GHz or 500 megabytes per second, but will fall back to PCI-e 1.x speed if either the device or mainboard does not support it.

Point-to-Point: Each pair of wires goes from the Northbridge to a single device. The old PCI bus ran the same wire through every device slot. With PCI express, a pair of wires is dedicated to a single slot.

A **Line**: One pair of wires carries data from the Northbridge to the device. A second pair carries data from the device to the Northbridge. A group of two pair of four wires is called a "line".

Uni-Directional: The pairs operate independently, so a device can be both sending and receiving data at 250 megabytes per second on each line. A few vendors claim that the line transfers 500 megabytes per second, but this is misleading because few devices both transmit and receive the same amount of data.

The old desktop PCI bus could transfer a total of 133 megabytes per second of data sent in both directions for all of the five PCI slots on the mainboard. A single line of PCI Express is almost twice as fast in either direction, can transmit data in both directions simultaneously, and is dedicated to a single device. For ordinary desktop use, one line of PCI Express is very fast.

However, the more exotic forms of Server PCI could go faster, as could the old AGP slot. To match or exceed these higher speeds, PCI Express allows one device to use two or more lines at the same time.

Round-Robin: The PCI bus transmitted one bit of data down each wire. The receiver accumulated these bits to form the data. A PCI Express line always sends a complete byte down the wire in 10 ticks of the 2.5 GHz clock (5 GHz on PCI-e 2.0). When a device is connected by more than one line, the bytes are transmitted "round robin" by assigning each consecutive byte to the next line, then wrapping back from the last line to the first. Two lines can carry 500 megabytes per second in each direction, four lines can carry a gigabyte, eight lines can carry 2 gigabytes, and sixteen lines can carry 4 gigabytes per second in each direction (double these numbers for 2.0).

x Notation: The convention is to use an "x" followed by the number of lines in use. This is, unfortunately, often confused with the AGP speed notation. An "x16" PCI Express video card has 16 lines and can transmit data at 4 gigabytes per second in both directions simultaneously (and twice that in PCI-e 2.0). An old 8x AGP card runs at "8 times" the base speed of the interface and can transfer only 2 gigabytes per second totaled over both directions.

Negotiate: At startup time, the Northbridge sends a message down each line of PCI Express asking the device at the other end to identify itself. When it gets back the same identity from two or more lines, it configures the device to round-robin the byte transmission across the lines that are connected to that device, and if PCI-e 2.0 is supported by both the mainboard and device, it decides to use this higher transmission speed.

Similarly, when a PCI Express device is plugged into a socket, it does not know how many lines it will actually be able to communicate across. Every PCI Express device must be ready to do everything on just one line if that is all the Mainboard is willing to allocate to it. The extra lines don't add anything except additional transmission capacity.

Power then data: A PCI Express socket has some power pins, a plastic barrier, and then a slot for signal pins. The signal slot can be large enough to accommodate connectors for 1, 4, 8, or 16 lines. A PCI Express card has connectors for the power, a gap that matches the plastic barrier in the slot, and then a tab that plugs into the signal slot.

The card can be shorter than the slot. A PCI Express card with a short tab can always plug into a plastic socket that is longer. Thus a PCI Express card with one line can plug into any Mainboard PCI Express slot even if it is designed for 4, 8, or 16 lines. Alternately, a PCI Express socket large enough to accommodate an x16 card will also accept any other size of card.

The data can be shorter than the slot. A mainboard doesn't have to connect an actual line of transmission capability to every connector on the slot. Several mainboards have "Universal" x16 plastic slots to which only 8, 4, or 2 lines are actually connected. At startup the card will sense which lines are active and will use only the ones it really has.

Thus at startup the Mainboard indicates how many lines it has and the card responds on the number of lines it can accommodate. They end up using the smaller of the two numbers that both can support.

x8: Intel Servers may have plastic slots for x4 and x8 adapter cards. This is useful for x4 adapter cards, typically Serial Attached SCSI (SAS) RAID controller cards.

x1: Desktop boards often have some x1 plastic slot for PCI Express adapter cards. These are often right next to the video card slots. A gamer buys the most powerful video cards available, which typically require two card slots to support a big heat sink and fan. These block the x1 slot which is regarded as expendable. However, if you have unblocked x1 slots, they can be used for SATA controller cards (to get some extra internal or external SATA disk connections), digital or analog TV tuner cards, and even a few sound cards.

Power: The PCI Express standard requires that the mainboard deliver more power to bigger slots than smaller slots. There is a table of required power delivery for x1, x4, x8, and x16 plastic slots. Even when the mainboard doesn't populate all the data connectors with active lines, it must deliver the amount of power indicated for each slot size.

PCI or PCI-e

An ATX board has room for 7 slots. An MATX board has room for 4 slots. Since modern video adapters use 16 lines of PCI-e, there will typically be one full-sized PCI-e slot on the board. The rest of the slots will be divided between PCI and PCI-e based on guesswork. You choose a mainboard based on your own guesswork of how many slots of each type you intend to use.

The very expensive video cards have very hot processing units that require extra cooling. As a result, they are often designed to occupy two card slots instead of one. If you intend to buy one or more of these cards, you must plan to lose access to the card slot next to the video slot. If you have two such cards, they will occupy four slots. While PCI-e is faster and simpler, there are still far more PCI than PCI-e adapter cards in each category. You need to plan your mainboard to allow for any adapters you own or intend to buy.

It is useful to compare the PCI-Express standard from Intel to the HyperTransport standard used by AMD and Apple:

- PCI-Express supports adapter cards plugged into sockets on the mainboard. HyperTransport connects chips soldered to the mainboard and cannot be connected to an I/O socket.
- PCI-Express runs point to point between the Northbridge and one device. HyperTransport is also a point to point connection, but each chip has a bridging capability to the next chip. So one HyperTransport bus can connect to a sequence of chips through bridged point to point links.
- PCI-Express has variable bandwidth represented by the number of lanes dedicated to a device. However, a device has exclusive control of whatever bandwidth is assigned to it. HyperTransport has a fixed bandwidth that is shared by all the bridged chips.

That said, this is mostly a theoretical comparison. If you buy a mainboard with an Nvidia NForce chipset, the board will use HyperTransport between the CPU and the chipset and PCI Express between the chipset and the video adapters. Each bus has its own role and its own devices.

[Copyright 1998, 2008 PCLT](#) -- *Introduction to PC Hardware* -- [H. Gilbert](#)