How PCI Works (by Jeff Tyson and Ed Grabianowski)
http://computer.howstuffworks.com/pci.htm

Your computer's components work together through a bus. Learn about the PCI bus and PCI card, such as the one above. See more computer hardware pictures.

The power and speed of computer components has increased at a steady rate since desktop computers were first developed decades ago. Software makers create new applications capable of utilizing the latest advances in processor speed and hard drive capacity, while hardware makers rush to improve components and design new technologies to keep up with the demands of high-end software.

There's one element, however, that often escapes notice - the bus. Essentially, a bus is a channel or path between the components in a computer. Having a high-speed bus is as important as having a good transmission in a car. If you have a 700-horsepower engine combined with a cheap transmission, you can't get all that power to the road. There are many different types of buses.

The idea of a bus is simple – it lets you connect components to the computer's processor. Some of the components that you might want to connect include hard disks, memory, sound systems, video systems and so on. For example, to see what your computer is doing, you normally use a CRT or LCD screen. You need special hardware to drive the screen, so the screen is driven by a graphics card. A graphics card is a small printed circuit board designed to plug into the bus. The graphics card talks to the processor using the computer's bus as a communication path.

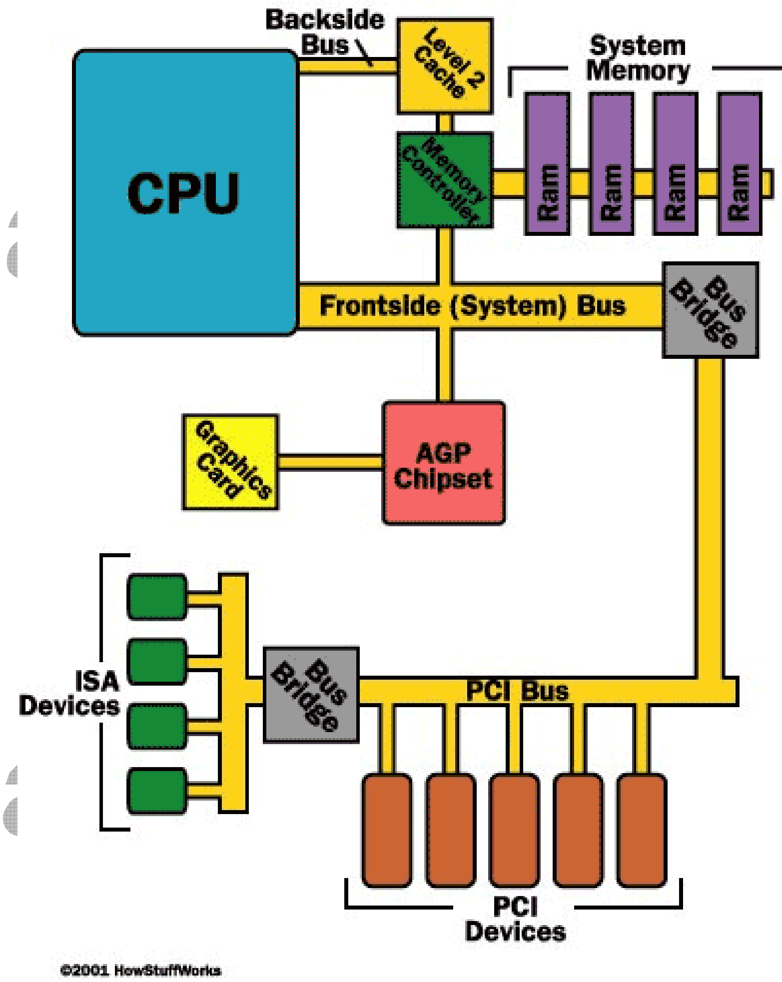
The advantage of a bus is that it makes parts more interchangeable. If you want to get a better graphics card, you simply unplug the old card from the bus and plug in a new one. If you want two monitors on your computer, you plug two graphics cards into the bus. And so on.

In this article, you will learn about some of those buses. We will concentrate on the bus known as the Peripheral Component Interconnect (PCI). We'll talk about what PCI is, how it operates and how it is used, and we'll look into the future of bus technology.

System Bus vs. PCI Bus
Twenty or 30 years ago, the processors were so slow that the processor and the bus were synchronized – the bus ran at the same speed as the processor, and there was one bus in the machine. Today, the processors run so fast that most computers have two or more buses. Each bus specializes in a certain type of traffic.

A typical desktop PC today has two main buses:

- The first one, known as the system bus or local bus, connects the microprocessor (central processing unit) and the system memory. This is the fastest bus in the system.
- The second one is a slower bus for communicating with things like hard disks and sound cards. One very common bus of this type is known as the PCI bus. These slower buses connect to the system bus through a bridge, which is a part of the computer's chipset and acts as a traffic cop, integrating the data from the other buses to the system bus.



The illustration above shows how the various buses connect to the CPU.

Technically there are other buses as well. For example, the Universal Serial Bus (USB) is a way of connecting things like cameras, scanners and printers to your computer. It uses a thin wire to connect to the devices, and many devices can share that wire simultaneously. Firewire is another bus, used today mostly for video cameras and external hard drives.

PCI History

The original PC bus in the original IBM PC (circa 1982) was 16 bits wide and operated at 4.77 MHz. It officially became known as the ISA bus. This bus design is capable of passing along data at a rate of up to 9 MBps (megabytes per second) or so, fast enough even for many of today's applications.

Several years ago, the ISA bus was still used on many computers. That bus accepted computer cards developed for the original IBM PC in the early 1980s. The ISA bus remained in use even after more advanced technologies were available to replace it.

There were a couple of key reasons for its longevity:

- Long-term compatibility with a large number of hardware manufacturers.
- Before the rise of multimedia, few hardware peripherals fully utilized the speed of the newer bus.

As technology advanced and the ISA bus failed to keep up, other buses were developed. Key among these were Extended Industry Standard Architecture (EISA) – which was 32 bits at 8 MHz– and Vesa Local Bus (VL-Bus). The cool thing about VL-Bus (named after VESA, the Video Electronics Standards Association, which created the standard) is that it was 32 bits wide and operated at the speed of the local bus, which was normally the speed of the processor itself. The VL-Bus essentially tied directly into the CPU. This worked okay for a single device, or maybe even two. But connecting more than two devices to the VL-Bus introduced the possibility of interference with the performance of the CPU. Because of this, the VL-Bus was typically used only for connecting a graphics card, a component that really benefits from high-speed access to the CPU.

During the early 1990s, Intel introduced a new bus standard for consideration, the Peripheral Component Interconnect (PCI) bus. PCI presents a hybrid of sorts between ISA and VL-Bus. It provides direct access to system memory for connected devices, but uses a bridge to connect to the frontside bus and therefore to the CPU. Basically, this means that it is capable of even higher performance than VL-Bus while eliminating the potential for interference with the CPU.

Frontside Bus, Backside Bus and PCI Cards

The frontside bus is a physical connection that actually connects the processor to most of the other components in the computer, including main memory (RAM), hard drives and the PCI slots. These days, the frontside bus usually operates at 400-MHz, with newer systems running at 800-MHz.

The backside bus is a separate connection between the processor and the Level 2 cache. This bus operates at a faster speed than the frontside bus, usually at the same speed as the processor, so all that caching works as efficiently as possible. Backside buses have evolved over the years. In the 1990s, the backside bus was a wire that connected the main processor to an off-chip cache. This cache was actually a separate chip that required expensive memory. Since then, the Level 2 cache has been integrated into the main processor, making processors smaller and cheaper. Since the cache is now on the processor itself, in some ways the backside bus isn't really a bus anymore.

Bus Type	Bus Width	Bus Speed	MB/sec
ISA	16 bits	8 MHz	16 MBps
EISA	32 bits	8 MHz	32 MBps
VL-bus	32 bits	25 MHz	100 MBps
VL-bus	32 bits	33 MHz	132 MBps
PCI	32 bits	33 MHz	132 MBps
PCI	64 bits	33 MHz	264 MBps
PCI	64 bits	66 MHz	512 MBps
PCI	64 bits	133 MHz	1 GBps

PCI can connect more devices than VL-Bus, up to five external components. Each of the five connectors for an external component can be replaced with two fixed devices on the motherboard. Also, you can have more than one PCI bus on the same computer, although this is rarely done. The PCI bridge chip regulates the speed of the PCI bus independently of the CPU's speed. This provides a higher degree of reliability and ensures that PCI-hardware manufacturers know exactly what to design for.

PCI originally operated at 33 MHz using a 32-bit-wide path. Revisions to the standard include increasing the speed from 33 MHz to 66 MHz and doubling the bit count to 64. Currently, PCI-X provides for 64-bit transfers at a speed of 133 MHz for an amazing 1-GBps (gigabyte per second) transfer rate!

PCI cards use 47 pins to connect (49 pins for a mastering card, which can control the PCI bus without CPU intervention). The PCI bus is able to work with so few pins because of hardware multiplexing, which means that the device sends more than one signal over a single pin. Also, PCI supports devices that use either 5 volts or 3.3 volts.

Although Intel proposed the PCI standard in 1991, it did not achieve popularity until the arrival of Windows 95 (in 1995). This sudden interest in PCI was due to the fact that Windows 95 supported a feature called Plug and Play (PnP), which we'll talk about in the next section.

Plug and Play

Plug and Play (PnP) means that you can connect a device or insert a card into your computer and it is automatically recognized and configured to work in your system. PnP is a simple concept, but it took a concerted effort on the part of the computer industry to make it happen. Intel created the PnP standard and incorporated it into the design for PCI. But it wasn't until several years later that a mainstream operating system, Windows 95, provided system-level support for PnP. The introduction of PnP accelerated the demand for computers with PCI, very quickly supplanting ISA as the bus of choice.

To be fully implemented, PnP requires three things:

- PnP BIOS - The core utility that enables PnP and detects PnP devices. The BIOS also reads the ESCD for configuration information on existing PnP devices.
- Extended System Configuration Data (ESCD) - A file that contains information about installed PnP devices.
- PnP operating system - Any operating system, such as Windows XP, that supports PnP. PnP handlers in the operating system complete the configuration process started by the BIOS for each PnP device. PnP automates several key tasks that were typically done either manually or with an installation utility provided by the hardware manufacturer.

These tasks include the setting of:

- Interrupt requests (IRQ) - An IRQ, also known as a hardware interrupt, is used by the various parts of a computer to get the attention of the CPU. For example, the mouse sends an IRQ every time it is moved to let the CPU know that it's doing something. Before PCI, every hardware component needed a separate IRQ setting. But PCI manages hardware interrupts at the bus bridge, allowing it to use a single system IRQ for multiple PCI devices.
- Direct memory access (DMA) - This simply means that the device is configured to access system memory without consulting the CPU first.
- Memory addresses - Many devices are assigned a section of system memory for exclusive use by that device. This ensures that the hardware will have the needed resources to operate properly.
- Input/Output (I/O) configuration - This setting defines the ports used by the device for receiving and sending information.

While PnP makes it much easier to add devices to your computer, it is not infallible.

Variations in the software routines used by PnP BIOS developers, PCI device manufacturers and Microsoft have led many to refer to PnP as "Plug and Pray." But the overall effect of PnP has been to greatly simplify the process of upgrading your computer to add new devices or replace existing ones.

PCI Standards and PCI Express

As processor speeds steadily climb in the GHz range, many companies are working feverishly to develop a next-generation bus standard. Many feel that PCI, like ISA before it, is fast approaching the upper limit of what it can do. All of the proposed new standards have something in common. They propose doing away with the shared-bus technology used in PCI and moving to a point-to-point switching connection. This means that a direct connection between two devices (nodes) on the bus is established while they are communicating with each other. Basically, while these two nodes are talking, no other device can access that path. By providing multiple direct links, such a bus can allow several devices to communicate with no chance of slowing each other down.