HOW THE INTERNET WORKS: UNDER THE HOOD

What Happens When You Click?

“Internet” stands for “interconnected networks" because it's really a network of networks. The computer in your home is connected in a local network. That network is connected to another network operated by your Internet Service Provider (ISP). The ISPs network is connected to other ISPs’ networks. Those networks may be comprised of many different types of computers. That’s the hardware or physical view of the Internet.

A variety of physical media can be used to make the connections: Ethernet cable, telephone or power transmission lines, radio signals (satellite or wifi), and beams of visible light (fiber optics) are all the same to the Internet. The key thing is that a medium be capable of transmitting information according to the protocols of the Internet.

A protocol, on or off the Internet, is an agreed language for communicating, and a set of rules for doing something. There are fire drill protocols; CPR protocols; and Internet Protocol. The last is the “IP” in the acronym, “TCP/IP.”

IP determines where data goes and how it travels; TCP makes sure it gets there quickly and intact. The Internet Protocol is the set of rules followed to deliver data from point A to point B on the Internet based on the destination machine’s IP address. TCP stands for Transmission Control Protocol: the set of rules followed to ensure fast, error-checked transmission of data between two points on the Internet.

IP Addresses and the Domain Name System

A numeric IP address is similar to the address written on a postcard. Applying the rules of the Internet Protocol to an IP address should get data from the sending (host) machine to the one with that IP address. These addressing and routing rules are found in the Domain Name System (DNS).

The core of the DNS is a huge, two-column table of domain names and IP addresses. When you type “scsccbkk.org” into your browser’s address bar, this is what happens:

The browser sends “scsccbkk.org” to a DNS server along with a request: “What’s the IP address that corresponds to ‘scsccbkk.org?’” The DNS server consults its table and sends the answer, if it has one. If the DNS server can’t find the answer, it sends the request to a higher-level DNS server that has more names and addresses. The request keeps getting kicked up to a higher level DNS server until the answer is found, if it exists. In the whole wide world, there are only 13 “root” DNS servers that know every name and address pair; most DNS requests are resolved (successfully answered) at much lower levels.

When your browser receives the correct IP address, it sends a request for Web content to that address using the HTTP or HTTPS protocol. (The latter specifies that certain security measures be taken to protect the privacy of communications; see below). When the Web server at that IP address gets the request, it collects the requested data and sends it back to the requesting browser’s IP address.

You Don’t Go to a Webpage; It Comes to You (Think Amazon order!)

You don’t really “go to” a Web page. Web pages come to you in response to your browser's requests, just as packages come to you from Amazon in response to your purchase orders.

A Web page may consist of thousands or millions of bytes of data. They don’t all arrive at once in one huge package. The data your browser requests is broken up into blocks of 1,000 to 1,500 bytes. Each block is packaged with header and footer information that specify where it’s going, what larger body of data it comes from, and where it fits in the jigsaw puzzle of blocks that will have to be re-assembled at the destination address.

Data blocks rarely follow each other in single file over the same path from a server to the machine that requested them. Instead, each packet of data is sent along the path of least resistance (fastest speed) by each router that handles it on its way back to you.

The illusion that you are visiting a website at Amazon is created by software. Or if you prefer, magic. Clarke's Third Law states: "Any sufficiently advanced technology is indistinguishable from magic."

What About Security?

In theory, any data travelling across the Internet can be seen by persons who have access to the computers or routers in the local network or Internet backbone. On a public wifi connection, you are even more exposed, because everything you can see in your web browser or email program is also visible to others on the same wifi network. In practical terms, that means everyone in the same coffee shop, airport lounge, library, computer classroom, or hotel.

The answer is encryption. When the web address shown in your browser says HTTPS instead of HTTP, that means your data is encrypted before hitting the Internet. To anyone who might be "sniffing" it will appear as a random jumble of numbers and letters.

The HTTPS protocol combines HTTP with a security protocol called TLS/SSL. Actually, TLS (Transport Layer Security) is a modern, more secure replacement for SSL (Secure Sockets Layer), but both are commonly used and so appear together. Using digital certificates and public key encryption technology, TLS/SSL first authenticates the destination server, verifying that it is indeed “scsccbkk.org” and not a malware-spewing imposter. Then an encrypted “tunnel” is created between the destination server and the requesting host machine, through which data is exchanged safe from eavesdropping. The math involved is mind-bogglingly complex, but that need not concern mere mortals.

All the extra activity of authentication, encryption and decryption of data adds some overhead to an Internet communication stream and the machines on each end. The Web may seem a bit slower but the added security and privacy are more than worth the sacrifice. Using a secure HTTPS connection is pretty much standard for most websites these days.