Unit 10: Wonderful World of the Web

The Internet as a System

Learning Objectives

- 27: The student can explain the abstractions in the Internet and how the Internet functions.
- 28: The student can explain characteristics of the Internet and the systems built on it.
- 29: The student can analyze how characteristics of the Internet and systems built on it influence their use.
- 30: The student can connect the concern of cybersecurity with the Internet and systems built on it.
- 31: The student can analyze how computing affects communication, interaction, and cognition.
- 33: The student can connect computing with innovations in other fields.
- 34: The student can analyze the beneficial and harmful effects of computing.
- 35: The student can connect computing within economic, social, and cultural contexts.

Readings/Lectures

- Blown to Bits: Appendix (/bjc-course/curriculum/10-internet/readings/btb-appendix.pdf)
- Reading 10.01: How Google Works (/bjc-course/curriculum/10-internet/readings/01-how-google-works)
- Reading 10.02: Life Before Search (/bjc-course/curriculum/10-internet/readings/02-life-before-search)
- Reading 10.03: Search Takes a Team (/bjc-course/curriculum/10-internet/readings/03-search-takes-a-team)
- Reading 10.04: Blown to Bits Appendix Questions (/bjc-course/curriculum/10-internet/readings/04-bits-questions)

External Resources

- The Rise of Human-Computer Cooperation (http://www.ted.com/talks/shyam_sankar_the_rise_of_human_computer_cooperation.html) - TED Talk
- Packet Switching Demo (http://www.pbs.org/opb/nerds2.0.1/geek_glossary/packet_switching_flash.html) - PBS
- Warriors of the Net Video (http://www.teachertube.com/viewVideo.php?video_id=23140)

Labs/Exercises

- Lab 10.01: How the Internet Works (/bjc-course/curriculum/10-internet/labs/01-how-the-internet-works)
- Lab 10.02: Internet Activity (/bjc-course/curriculum/10-internet/labs/02-internet-activity)
- Portfolio Project 10.03: Internet Portfolio Project (/bjc-course/curriculum/10-internet/labs/03-internet-portfolio-project)

APPENDIX

The Internet as System and Spirit

This Appendix explains how the Internet works and summarizes some larger lessons of its remarkable success.

The Internet as a Communication System

The Internet is not email and web pages and digital photographs, any more than the postal service is magazines and packages and letters from your Aunt Mary. And the Internet is not a bunch of wires and cables, any more than the postal service is a bunch of trucks and airplanes. The Internet is a system, a delivery service for bits, whatever the bits represent and however they get from one place to another. It's important to know how it works, in order to understand why it works so well and why it can be used for so many different purposes.

Packet Switching

Suppose you send an email to Sam, and it goes through a computer in Kalamazoo—an Internet *router*, as the machines connecting the Internet together are known. Your computer and Sam's know it's an email, but the router in Kalamazoo just knows that it's handling bits.

Your message almost certainly goes through some copper wires, but probably also travels as light pulses through fiber optic cables, which carry lots of bits at very high speeds. It may also go through the air by radio—for example, if it is destined for your cell phone. The physical infrastructure for the Internet is owned by many different parties—including telecommunications firms in the

U.S. and governments in some countries. The Internet works not because anyone is in charge of the whole thing, but because these parties agree on what to expect as messages are passed from one to another. As the name suggests, the Internet is really a set of standards for **inter**connecting **net**works. The individual networks can behave as they wish, as long as they follow established conventions when they send bits out or bring bits in.

In the 1970s, the designers of the Internet faced momentous choices. One critical decision had to do with message sizes. The postal service imposes size and weight limits on what it will handle. You can't send your Aunt Mary a two-ton package by taking it to the Post Office. Would there also be a limit on the size of the messages that could be sent through the Internet? The designers anticipated that very large messages might be important some day, and found a way to avoid any size limits.

A second critical decision was about the very nature of the network. The obvious idea, which was rejected, was to create a "circuit-switched" network. Early telephone systems were completely circuit-switched. Each customer was connected by a pair of wires to a central switch. To complete a call from you to your Aunt Mary, the switch would be set to connect the wires from you to the wires from Aunt Mary, establishing a complete electrical loop between you and Mary for as long as the switch was set that way. The size of the switch limited the number of calls such a system could handle. Handling more simultaneous calls required building bigger switches. A circuit-switched network provides reliable, uninterruptible connections—at a high cost per connection. Most of the switching hardware is doing very little most of the time.

So the early Internet engineers needed to allow messages of unlimited size. They also needed to ensure that the capacity of the network would be limited only by the amount of data traffic, rather than by the number of interconnected computers. To meet both objectives, they designed a *packet-switched network*. The unit of information traveling over the Internet is a packet of about 1500 bytes or less—roughly the amount of text you might be able to put on a postcard. Any communications longer than that are broken up into multiple packets, with serial numbers so that the packets can be reassembled upon arrival to put the original message back together.

The packets that constitute a message need not travel through the Internet following the same route, nor arrive in the same order in which they were sent. It is very much as though the postal service would deliver only postcards with a maximum of 1500 characters as a message. You could send *War and Peace*, using thousands of postcards. You could even send a complete description of a photograph on postcards, by splitting the image into thousands of rows and columns and listing on each postcard a row number, a column number, and the color of the little square at that position. The recipient could, in principle, reconstruct the picture after receiving all the postcards. What makes the Internet work in practice is the incredible speed at which the data packets are transmitted, and the processing power of the sending and receiving computers, which can disassemble and reassemble the messages so quickly and flawlessly that users don't even notice.

Core and Edge

We can think of the ordinary postal system as having a *core* and an *edge*—the edge is what we see directly, the mailboxes and letter carriers, and the core is everything behind the edge that makes the system work. The Internet also has a core and an edge. The edge is made up of the machines that interface directly with the end users—for example, your computer and mine. The core of the Internet is all the connectivity that makes the Internet a network. It includes the computers owned by the telecommunications companies that pass the messages along.

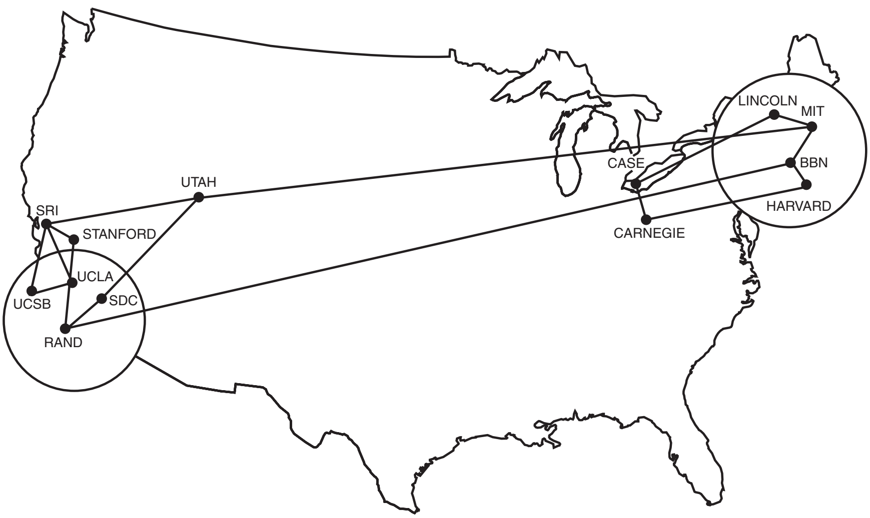
An *Internet Service Provider* or *ISP* is any computer that provides access to the Internet, or provides the functions that enable different parts of the Internet to connect to each other. Sometimes the organizations that run those computers are also called ISPs. Your ISP at home is likely your telephone or cable company, though if you live in a rural area, it might be a company providing Internet services by satellite. Universities and big companies are their own ISPs. The "service" may be to convey messages between computers deep within the core of the Internet, passing messages until they reach their destination. In the United States alone, there are thousands of ISPs, and the system works as a whole because they cooperate with each other.

Fundamentally, the Internet consists of computers sending bit packets that request services, and other computers sending packets back in response. Other metaphors can be helpful, but the service metaphor is close to the truth. For example, you don't really "visit" the web page of a store, like a voyeuristic tourist peeking through the store window. Your computer makes a very specific request of the store's web server, and the store's web server responds to it—and may well keep a record of exactly what you asked for, adding the new information about your interests to the record it already has from your other "visits." Your "visits" leave fingerprints!

IP Addresses

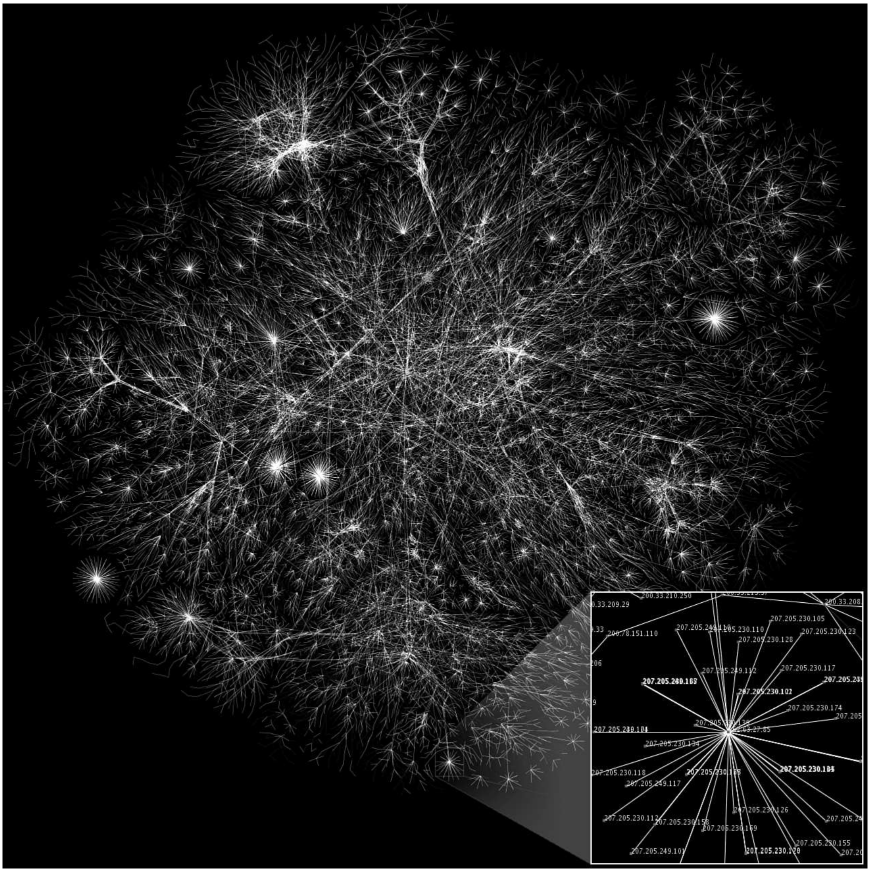
Packets can be directed to their destination because they are labeled with an *IP address*, which is a sequence of four numbers, each between 0 and 255. (The numbers from 0 to 255 correspond to the various sequences of 8 bits, from 00000000 to 11111111, so IP addresses are really 32 bits long. "IP" is an abbreviation for "Internet Protocol," explained next.) A typical IP address is 66.82.9.88. Blocks of IP addresses are assigned to ISPs, which in turn assign them to their customers.

There are 256 × 256 × 256 × 256 possible IP addresses, or about 4 billion. In the pre-miniaturization days when the Internet was designed, that seemed an absurdly large number—enough so every computer could have its own IP address, even if every person on the planet had his or her own computer. Figure A.1 shows the 13 computers that made up the entire network in 1970. As a result of miniaturization and the inclusion of cell phones and other small devices, the number of Internet devices is already in the hundreds of millions (see Figure A.2), and it seems likely that there will not be enough IP addresses for the long run. A project is underway to deploy a new version of IP in which the size of IP addresses increases from 32 bits to 128—and then the number of IP addresses will be a 3 followed by 38 zeroes! That's about ten million for every bacterium on earth.



Source: Heart, F., McKenzie, A., McQuillian, J., and Walden, D., ARPANET Completion Report, Bolt, Beranek and Newman, Burlington, MA, January 4, 1978.

FIGURE A.1 The 13 interconnected computers of the December, 1970 ARPANET (as the Internet was first known). The interconnected machines were located at the University of California campuses at Santa Barbara and at Los Angeles, the Stanford Research Institute, Stanford University, Systems Development Corporation, the RAND Corporation, the University of Utah, Case Western Reserve University, Carnegie Mellon University, Lincoln Labs, MIT, Harvard, and Bolt, Beranek, and Newman, Inc.



Source: Wikipedia, `http://en.wikipedia.org/wiki/Image: Internet_map_1024.jpg`. This work is licensed under the Creative Commons Attribution 2.5 License.

FIGURE A.2 Traffic flows within a small part of the Internet as it exists today. Each line is drawn between two IP addresses of the network. The length of a line indicates the time delay for messages between those two nodes. Thousands of cross-connections are omitted.

An important piece of the Internet infrastructure are the *Domain Name Servers*, which are computers loaded with information about which IP addresses correspond to which "domain names" such as `harvard.edu`, `verizon.com`, `gmail.com`, `yahoo.fr` (the suffix in this case is the country code for France), and `mass.gov`. So when your computer sends an email or requests a web page, the translation of domain names into IP addresses takes place before the message enters the core of the Internet. The routers don't know about domain names; they need only pass the packets along toward their destination IP address numbers.

IP ADDRESSES AND CRIMES

The recording industry identifies unlawful music downloads by the IP addresses to which the bits are sent. But an IP address is rarely the exclusive property of an individual, so it is hard to be sure who is doing the downloading. A provider of residential Internet service allocates an address to a home only temporarily. When the connection becomes inactive, the address is reclaimed so someone else can use it. If NAT is in use or if many people use the same wireless router, it can be impossible to establish reliably who exactly used an IP address. If you don't activate the security on your home wireless router, neighbors who poach your home network signal may get you in serious trouble by their illegal downloads!

An enterprise that manages its own network can connect to the Internet through a single gateway computer, using only a single IP address. Packets are tagged with a few more bits, called a "port" number, so that the gateway can route responses back to the same computer within the private network. This process, called *Network Address Translation* or *NAT*, conserves IP addresses. NAT also makes it impossible for "outside" computers to know which computer actually made the request—only the gateway knows which port corresponds to which computer.

The Key to It All: Passing Packets

At heart, all the core of the Internet does is to transmit packets. Each router has several links connecting it to other routers or to the "edge" of the network. When a packet comes in on a link, the router very quickly looks at the destination IP address, decides which outgoing link to use based on a limited Internet "map" it holds, and sends the packet on its way. The router has some memory, called a *buffer*, which it uses to store packets temporarily if they are arriving faster than they can be processed and dispatched. If the buffer fills up, the router just discards incoming packets that it can't hold, leaving other parts of the system to cope with the data loss if they choose to.

Packets also include some redundant bits to aid error detection. To give a simple analogy, suppose Alice wants to guard against a character being smudged or altered on a post card while it is in transit. Alice could add to the text on the card a sequence of 26 bits—indicating whether the text she has put on the card has an even or odd number of As, Bs, ..., and Zs. Bob can check whether the card seems to be valid by comparing his own reckoning with the 26-bit "fingerprint" already on the card. In the Internet, all the routers do a similar integrity check on data packets. Routers discard packets found to have been damaged in transit.

The format for data packets—which bits represent the IP address and other information about the packet, and which bits are the message itself—is part of the *Internet Protocol*, or IP. Everything that flows through the Internet—web pages, emails, movies, VoIP telephone calls—is broken down into data packets. Ordinarily, all packets are handled in exactly the same way by the routers and other devices built around IP. IP is a "best effort" packet delivery protocol. A router implementing IP tries to pass packets along, but makes no guarantees. Yet guaranteed delivery is possible within the network as a whole—because other protocols are layered on top of IP.

Protocols

A "protocol" is a standard for communicating messages between networked computers. The term derives from its meaning in diplomacy. A diplomatic protocol is an agreement aiding in communications between mutually mistrustful parties—parties who do not report to any common authority who can control their behavior. Networked computers are in something of the same situation of both cooperation and mistrust. There is no one controlling the Internet as a whole. Any computer can join the global exchange of information, simply by interconnecting physically and then following the network protocols about how bits are inserted into and extracted from the communication links.

The fact that packets can get discarded, or "dropped" as the phrase goes, might lead you to think that an email put into the network might never arrive. Indeed emails can get lost, but when it happens, it is almost always because of a problem with an ISP or a personal computer, not because of a network failure. The computers at the edge of the network use a higher-level protocol to deliver messages reliably, even though the delivery of individual packets within the network may be unreliable. That higher-level protocol is called "Transport Control Protocol," or TCP, and one often hears about it in conjunction with IP as "TCP/IP."

To get a general idea of how TCP works, imagine that Alice wants to send Bob the entire text of *War and Peace* on postcards, which are serial numbered so Bob can reassemble them in the right order even if they arrive out of order. Postcards sometimes go missing, so Alice keeps a copy of every postcard she puts in the mail. She doesn't discard her copy of a postcard until she has received word back from Bob declaring that he has received Alice's postcard. Bob sends that word back on a postcard of his own, including the serial number of Alice's card so Alice knows which card is being confirmed. Of course, Bob's confirming postcards may get lost too, so Alice keeps track of when she sent her postcards. If she doesn't hear anything back from Bob within a certain amount of time, she sends a duplicate postcard. At this point, it starts getting complicated: Bob has to know enough to ignore duplicates, in case it was his acknowledgment rather than Alice's original message that got lost. But it all can be made to work!

TCP works the same way on the Internet, except that the speed at which packets are zipping through the network is extremely fast. The net result is that email software using TCP is failsafe: If the bits arrive at all, they will be a perfect duplicate of those that were sent.

TCP is not the only high-level protocol that relies on IP for packet delivery. For "live" applications such as streaming video and VoIP telephone calls, there is no point in waiting for retransmissions of dropped packets. So for these applications, the packets are just put in the Internet and sent on their way, with no provision made for data loss. That higher-level protocol is called UDP, and there are others as well, all relying on IP to do the dirty work of routing packets to their destination.

The postal service provides a rough analogy of the difference between higher-level and lower-level protocols. The same trucks and airplanes are used for carrying first-class mail, priority mail, junk mail, and express mail. The loading and unloading of mail bags onto the transport vehicles follow a low-level protocol. The handling between receipt at the post office and loading onto the transport vehicles, and between unloading and delivery, follows a variety of higher-level protocols, according to the kind of service that has been purchased.

In addition to the way it can be used to support a variety of higher-level protocols, IP is general in another way. It is not bound to any particular physical medium. IP can run over copper wire, radio signals, and fiber optic cables—in principle, even carrier pigeons. All that is required is the ability to deliver bit packets, including both the payload and the addressing and other "packaging," to switches that can carry out the essential routing operation.

There is a separate set of "lower-level protocols" that stipulate how bits are to be represented—for example, as radio waves, or light pulses in optic fibers. IP is doubly general, in

IP OVER CARRIER PIGEON

You can look up RFC 1149 and RFC 2549 on the Web, "Standard for the Transmission of IP Datagrams on Avian Carriers" and "IP over Avian Carriers with Quality of Service." They faithfully follow the form of true Internet standards, though the authors wrote them with tongue firmly planted in cheek, demurely stating, "This is an experimental, not recommended standard."

that it can take its bit packets from many different physical substrates, and deliver those packets for use by many different higher-level services.

The Reliability of the Internet

The Internet is remarkably reliable. There are no "single points of failure." If a cable breaks or a computer catches on fire, the protocols automatically reroute the packets around the inoperative links. So when Hurricane Katrina submerged New Orleans in 2005, Internet routers had packets bypass the city. Of course, no messages destined for New Orleans itself could be delivered there.

In spite of the redundancy of interconnections, if enough links are broken, parts of the Internet may become inaccessible to other parts. On December 26, 2006, the Henchung earthquake severed several major communication cables that ran across the floor of the South China Sea. The Asian financial markets were severely affected for a few days, as traffic into and out of Taiwan, China, and Hong Kong was cut off or severely reduced. There were reports that the volume of spam reaching the U.S. also dropped for a few days, until the cables were repaired!

Although the Internet *core* is reliable, the computers on the edge typically have only a single connection to the core, creating single points of failure. For example, you will lose your home Internet service if your phone company provides the service and a passing truck pulls down the wire connecting your house to the telephone pole. Some big companies connect their internal network to the Internet through two different service providers—a costly form of redundancy, but a wise investment if the business could not survive a service disruption.

The Internet Spirit

The extraordinary growth of the Internet, and its passage from a military and academic technology to a massive replacement for both paper mail and telephones, has inspired reverence for some of its fundamental design virtues. Internet principles have gained status as important truths about communication, free expression, and all manner of engineering design.

The Hourglass

The standard electric outlet is a universal interface between power plants and electric appliances. There is no need for people to know whether their power is coming from a waterfall, a solar cell, or a nuclear plant, if all they want to do is to plug in their appliances and run their household. And the same electric outlet can be used for toasters, radios, and vacuum cleaners. Moreover, it will instantly become usable for the next great appliance that gets invented, as long as that device comes with a standard household electric plug. The electric company doesn't even care if you are using its electricity to do bad things, as long as you pay its bills.

The outlet design is at the neck of a conceptual hourglass through which electricity flows, connecting multiple possible power sources on one side of the neck to multiple possible electricity-using devices on the other. New inventions need only accommodate what the neck expects—power plants need to supply 115V AC current to the outlet, and new appliances need plugs so they can use the current coming from the outlet. Imagine how inefficient it would be if your house had to be rewired in order to accommodate new appliances, or if different kinds of power plants required different household wiring. Anyone who has tried to transport an electric appliance between the U.S. and the U.K. knows that electric appliances are less universal than Internet packets.

The Internet architecture is also conceptually organized like an hourglass (see Figure A.3), with the ubiquitous Internet Protocol at the neck, defining the form of the bit packets carried through the network. A variety of higher-level protocols use bit packets to achieve different purposes. In the words of the report that proposed the hourglass metaphor, "the minimal required elements [IP] appear at the narrowest point, and an ever-increasing set of choices fills the wider top and bottom, underscoring how little the Internet itself demands of its service providers and users."

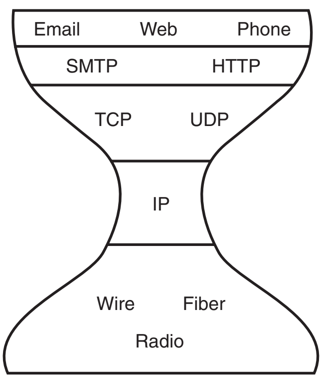


FIGURE A.3 The Internet protocol hourglass (simplified). Each protocol interfaces only to those in the layers immediately above and below it, and all data is turned into IP bit packets in order to pass from an application to one of the physical media that make up the network.

For example, TCP guarantees reliable though possibly delayed message delivery, and UDP provides timely but unreliable message delivery. All the higher-level protocols rely on IP to deliver packets. Once the packets get into the neck of the hourglass, they are handled identically, regardless of the higher-level protocol that produced them. TCP and UDP are in turn utilized by even higher-level protocols, such as HTTP ("HyperText Transport Protocol"), which is used for sending and receiving web pages, and SMTP ("Simple Mail Transport Protocol"), which is used for sending email. Application software, such as web browsers, email clients, and VoIP software, sit at a yet higher level, utilizing the protocols at the layer below and unconcerned with how those protocols do their job.

Below the IP layer are various physical protocol layers. Because IP is a universal protocol at the neck, applications (above the neck) can accommodate various possible physical implementations (below the neck). For example, when the first wireless IP devices became available, long after the general structure of the Internet hourglass was firmly in place, nothing above the neck had to change. Email, which had previously been delivered over copper wires and glass fibers, was immediately delivered over radio waves such as those sent and received by the newly developed household wireless routers.

Governments, media firms, and communication companies sometimes wish that IP worked differently, so they could more easily filter out certain kinds of content and give others priority service. But the universality of IP, and the many unexpected uses to which it has given birth, argue against such proposals to re-engineer the Internet. As information technology consultant Scott Bradner wrote, "We have the Internet that we have today because the Internet of yesterday did not focus on the today of yesterday. Instead, Internet technology developers and ISPs focused on flexibility, thus enabling whatever future was coming."

Indeed, the entire social structure in which Internet protocols evolved prevented special interests from gaining too much power or building their pet features into the Internet infrastructure. Protocols were adopted by a working group called the Internet

The Future of the Internet—and How to Stop It

This excellent book by Jonathan Zittrain (Yale University Press and Penguin UK, 2008) sees the vulnerabilities of the Internet—rapidly spreading viruses, and crippling attacks on major servers—as consequences of its essential openness, its capacity to support new inventions—what Zittrain calls its "generativity." The book reflects on whether society will be driven to use a network of less-flexible "appliances" in the future to avoid the downsides of the Internet's wonderfully creative malleability.

Engineering Task Force (IETF), which made its decisions by rough consensus, not by voting. The members met face to face and hummed to signify their approval, so the aggregate sense of the group would be public and individual opinions could be reasonably private—but no change, enhancement, or feature could be adopted by a narrow majority.

The larger lesson is the importance of minimalist, well-selected, open standards in the design of any system that is to be widely disseminated and is to stimulate creativity and unforeseen uses. Standards, although they are merely conventions, give rise to vast innovation, if they are well chosen, spare, and widely adopted.

Layers, Not Silos

Internet functionality could, in theory, have been provided in many other ways. Suppose, for example, that a company had set out just to deliver electronic mail to homes and offices. It could have brought in special wiring, both economical and perfect for the data rates needed to deliver email. It could have engineered special switches, perfect for routing email. And it could have built the ideal email software, optimized to work perfectly with the special switches and wires.

Another group might have set out to deliver movies. Movies require higher data rates, which might better be served by the use of different, specialized switches. An entirely separate network might have been developed for that. Another group might have conceived something like the Web, and have tried to convince ordinary people to install yet a third set of cables in their homes.

The magic of the hourglass structure is not just the flexibility provided by the neck of the bottle. It's the logical isolation of the upper layers from the lower. Inventive people working in the upper layers can rely on the guarantees provided by the clever people working at the lower layers, without knowing much about *how* those lower layers work. Instead of multiple, parallel vertical structures—self-contained silos—the right way to engineer information is in layers.

And yet we live in an information economy still trapped, legally and politically, in historical silos. There are special rules for telephones, cable services, and radio. The medium determines the rules. Look at the names of the main divisions of the Federal Communications Commission: Wireless, wireline, and so on. Yet the technologies have converged. Telephone calls go over the Internet, with all its variety of physical infrastructure. The bits that make up telephone calls are no different from the bits that make up movies.

Laws and regulations should respect layers, not the increasingly meaningless silos—a principle at the heart of the argument about broadcast regulation presented in Chapter 8.

End to End

"End to End," in the Internet, means that the switches making up the core of the network should be dumb—optimized to carry out their single limited function of passing packets. Any functionality requiring more "thinking" than that should be the responsibility of the more powerful computers at the edge of the network. For example, Internet protocols could have been designed so that routers would try much harder to ensure that packets do not get dropped on any link. There could have been special codes for packets that got special, high-priority handling, like "Priority Mail" in the U.S. Postal Service. There could have been special codes for encrypting and decrypting packets at certain stages to provide secrecy, say when packets crossed national borders. There are a lot of things that routers might have done. But it was better, from an engineering standpoint, to have the core of the network do the minimum that would enable those more complex functions to be carried out at the edge. One main reason is that this makes it more likely that new applications can be added without having to change the core—any operations that are application-specific will be handled at the edges. This approach has been staggeringly successful, as illustrated by today's amazing array of Internet applications that the original network designers never anticipated.

Stupid Networks

Another way to understand the Internet's end-to-end philosophy is to realize that if the computers are powerful at the edge of the network, the network itself can be "stupid," just delivering packets where the packets themselves say they want to go. Contrast this with the old telephone network, in which the devices at the edge of the network were stupid telephones, so to provide good service, the switching equipment in the telephone office had to be intelligent, routing telephone signals to where the network said they should go.

Separate Content and Carrier

The closest thing to the Internet that existed in the nineteenth century was the telegraph. It was an important technology for only a few decades. It put

> **THE VICTORIAN INTERNET**
>
> That is the title of an excellent short book by Tom Standage (Berkley Books, 1999), making the argument that many of the social consequences of the Internet were seen during the growth of the telegraph. The content-carrier conflict is only one. On a less-serious level, the author notes that the telegraph, like the Internet, was used for playing games at a distance almost from the day it came into being.

the Pony Express out of business, and was all but put out of business itself by the telephone. And it didn't get off to a fast start; at first, a service to deliver messages quickly didn't seem all that valuable.

One of the first big users of the telegraph was the Associated Press—one of the original "wire services." News is, of course, more valuable if it arrives quickly, so the telegraph was a valuable tool for the AP. Recognizing that, the AP realized that its competitive position, relative to other press services, would be enhanced to the extent it could keep the telegraph to itself. So it signed an exclusive contract with Western Union, the telegraph monopoly. The contract gave the AP favorable pricing on the use of the wires. Other press services were priced out of the use of the "carrier." And as a result, the AP got a lock on news distribution so strong that it threatened the functioning of the American democracy. It passed the news about politicians it liked and omitted mention of those it did not. Freedom of the press existed in theory, but not in practice, because the content industry controlled the carrier.

> **MORE ON INFORMATION FREEDOM**
>
> The SaveTheInternet.com Coalition is a pluralistic group dedicated to net neutrality and Internet freedom more generally. Its member organizations run the gamut from the Gun Owners of America, to MoveOn.org, to the Christian Coalition, to the Feminist Majority. Its web site includes a blog and a great many links. The blog of law professor Susan Crawford, `scrawford.net/blog`, comments on many aspects of digital information freedom, and also has a long list of links to other blogs.

Today's version of this morality play is the debate over "net neutrality." Providers of Internet backbone services would benefit from providing different pricing and different service guarantees to preferred customers. After all, they might argue, even the Postal Service recognizes the advantages of providing better service to customers who are willing to pay more. But what if a movie studio buys an ISP, and then gets creative with its pricing and service structure? You might discover that your movie downloads are far cheaper to watch, or arrive at your home looking and sounding much better, if they happen to be the product of the parent content company.

Or what if a service provider decides it just doesn't like a particular customer, as Verizon decided about Naral? Or what if an ISP finds that its customer is taking advantage of its service deal in ways that the provider did not anticipate? Are there any protections for the customer?

In the Internet world, consider the clever but deceptive scheme implemented by Comcast in 2007. This ISP promised customers unlimited bandwidth, but then altered the packets it was handling to slow down certain data transmissions. It peeked at the packets and altered those that had been generated by certain higher-level protocols commonly (but not exclusively) used for downloading and uploading movies. The end-user computer receiving these altered packets did not realize they had been altered in transit, and obeyed the instruction they contained, inserted in transit by Comcast, to restart the transmission from scratch. The result was to make certain data services run very slowly, without informing the customers. In a net neutrality world, this could not happen; Comcast would be a packet delivery service, and not entitled to choose which packets it would deliver promptly or to alter the packets while handing them on.

In early 2008, AT&T announced that it was considering a more direct violation of net neutrality: examining packets flowing through its networks to filter out illegal movie and music downloads. It was as though the electric utility announced it might cut off the power to your DVD player if it sensed that you were playing a bootleg movie. A content provider suggested that AT&T intended to make its content business more profitable by using its carrier service to enforce copyright restrictions. In other words, the idea was perhaps that people would be more likely to buy movies from AT&T the content company if AT&T the carrier refused to deliver illegally obtained movies. Of course, any technology designed to detect bits illegally flowing into private residences could be adapted, by either governments or the carriers, for many other purposes. Once the carriers inspect the bits you are receiving into your home, these private businesses could use that power in other ways: to conduct surveillance, enforce laws, and impose their morality on their customers. Just imagine Federal Express opening your mail in transit and deciding for itself which letters and parcels you should receive!

Clean Interfaces

The electric plug is the interface between an electric device and the power grid. Such standardized interfaces promote invention and efficiency. In the

Internet world, the interfaces are the connections between the protocol layers—for example, what TCP expects of IP when it passes packets into the core, and what IP promises to do for TCP.

In designing information systems, there is always a temptation to make the interface a little more complicated in order to achieve some special functionality—typically, a faster data rate for certain purposes. Experience has shown repeatedly, however, that computer programming is hard, and the gains in speed from more complex interfaces are not worth the cost in longer development and debugging time. And Moore's Law is always on the side of simplicity anyway: Just wait, and a cruder design will become as fast as the more complicated one might have been.

Even more important is that the interfaces be widely accepted standards. Internet standards are adopted through a remarkable process of consensus-building, nonhierarchical in the extreme. The standards themselves are referred to as RFCs, "Requests for Comment." Someone posts a proposal, and a cycle of comment and revision, of buy-in and objection, eventually converges on something useful, if not universally regarded as perfect. All the players know they have more to gain by accepting the standard and engineering their products and services to meet it than by trying to act alone. The Internet is an object lesson in creative compromise producing competitive energy.

RFCs and Standards

The archive of RFCs creates a history of the Internet. It is open for all to see—just use your favorite search engine. All the Internet standards are RFCs, although not all RFCs are standards. Indeed, in a whimsically reflexive explanation, "Not all RFCs are standards" is RFC 1796.

How Google Works

The Problem of Scale

One of the hardest parts in the life of a search engine is managing the extraordinary diversity of content it encounters while doing its job (both in the pages it sees and the questions that it gets asked). Instead of trying to do all of the reasoning and analysis when a search term arrives, it tries instead to organize the information in such a way that it learns a little bit about it, and where to find it, before its even asked a question. If search engines were actually going out to every page to look at its content for the first time when you performed a search, it would take forever to get your results. Imagine going to a big concert and being asked to find the oldest person from Indiana at the show. This would be incredibly difficult if you didn't have any information about the people – but if you were allowed to arrange the people by the origin state as well as their age within each state before the question was asked, then it would become much easier to arrive at a quick and accurate solution. Phone books are alphabetical for a reason!

The Anatomy of a Search Engine

Search engines, including Google, are generally composed of three major parts: the crawler, the indexer, and the searcher. This video was produced by a team at Google describing how search engines work; listen to their description of the process (they are pretty good at it, after all). Don't worry if you don't follow all of it – we'll spend more time on it in a second.

Sounds easy enough, right? Today we're going to use BYOB to create a very simple search engine. Search is a reasonably challenging thing to do at all, and an incredibly difficult thing to master (Google, Microsoft, and some others are spending billions trying!). Breaking up the search engine into multiple parts will make the overall problem much more manageable. Understanding the concepts behind search can be challenging, so we'll spend a little bit more time on those in the following section.

Life Before Search

Now consider a single cluster of technologies: the Internet and web search. These technologies have only been popular for 20 - 30 years, but its hard to imagine life before them.

Before Search Engines and the Internet

Before the Internet became the phenomenon that it is today, it was considerably more difficult to share information with others, especially if they weren't in the same area of the world. In addition to you and your classmates' submissions in the discussion forum, CyberNetNews (http://cybernetnews.com/cybernotes-life-before-the-internet-was-like/) mentions a number of differences between then (15 years ago) and now.

The Internet today provides a tremendous service of making information available to anyone in the world that can connect to it. "Information" can take the form of company websites, email, instant messaging, audio / video, and many, many other possibilities.

Early Search Engines

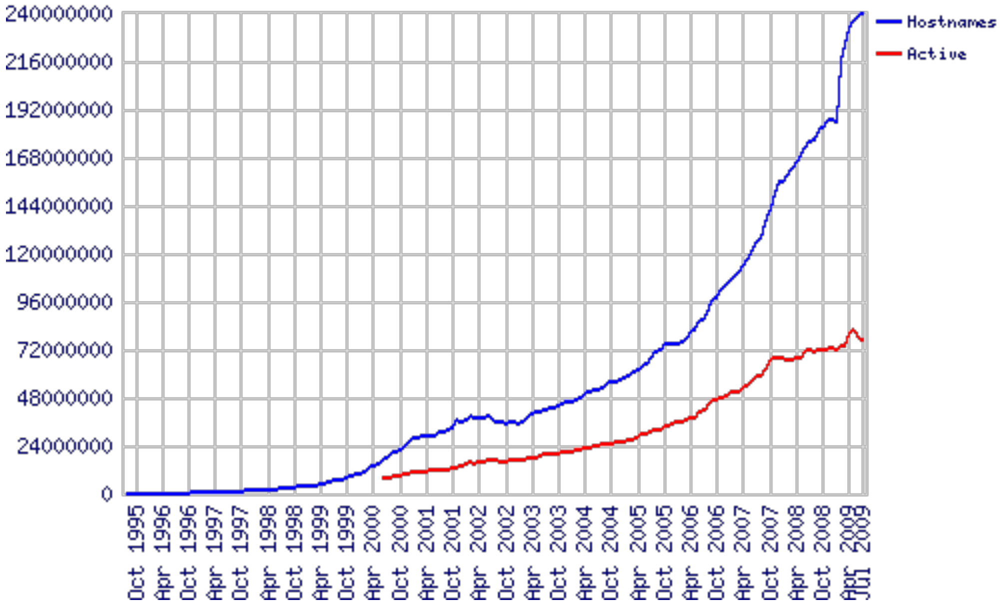
The time of the Internet before search engines was tough as well. The Internet was the same decentralized, wildly diverse collection of information that it is today (although a good bit smaller), but was also terribly difficult to navigate: it was comparable to a telephone network without a phone book. Finding your way to new places on the Internet often involved someone telling you about their site, or stumbling across a site from one you were already familiar with.

Not being ones to shy from problems, computer scientists noticed this shortcoming and began working on technology that would automatically crawl and index the websites that they could find. The directories they created became early search engines. A number of search engines like Lycos, Dogpile, Altavista and Askjeeves used to be big players in the search market. These search engines made huge steps in organizing the Internet by bringing most of it to one place. Nevertheless, the search process was unrefined by modern standards and often required users to search through long lists of sites to find what they needed.

Alas, this wasn't the case for long. Tremendous progress has been made between then and now, much of which has come from a single company that devoted its entire business to search: Google.

Today's Search

Almost anyone these days is able to find information from across the world in less than half of a second. You can answer almost any question whose solution is known by humanity in under 30 seconds. The amount of information that is available online has taken the world by storm, shifting entire industries and antiquating others. There are tens of millions of websites online at any one time and more are being added every year.



If there is so much information available on the Internet, it seems pretty remarkable that today's search engines can come up with such startlingly accurate results for what we want, when we want it.

Search Takes a Team

The inner workings of search engines are pretty complex systems and can be very intimidating. But what if we were to look at search engines differently? What if we saw them instead as a fruit stand?

There may have been a time when this fruit stand could've been run by one person, but now the demand at the stand is far too high for one person and the job is divided up into three different tasks:

Crawler

The task of the Crawler is straightforward: search all of the fields and forests that you can find for fruit. It doesn't matter what kind it is, what it tastes like, or whether you think that anyone will want to eat it. You find it, you take it. The performance of a Crawler could be measured by how much fruit it could gather over time – the faster the better. Big fruit stands may have multiple crawlers that can easily do their jobs concurrently.

Indexer

The Indexer is the heart of the stand's business. After the Crawler(s) delivers its harvest, the Indexer goes to work organizing the fruit in a way that will be convenient for any Searchers that happen to come by. Depending on the kind of Searchers that the stand typically sees, the Indexer may organize the fruit by type, age, color, size, or otherwise. The cart itself can then be called the index, which is the place where the Indexer places all of the organized fruit. Notice that the word “index” has been used before in this course (in the context of lists as well as loops)…this is the same word but with a different meaning! The Indexer puts the fruit on the stand in an order that it can remember so that it can quickly find a particular type of fruit when a Searcher asks for it.

Searcher

the Searcher is the customer who is looking for the perfect fruit. Different Searchers are usually looking for different fruit, and can often be very picky with their requirements for each individual piece. They search through the fruit stand and find the fruit that best fits what they’re looking for, and want to get in, find their fruit, and get out as quickly as possible.

Reading Assignment

Read the Appendix of Blown to Bits (15 pages). This chapter talks about the Internet as a system.

Reading Questions

Post answers to the following questions on a Portfolio page. We will discuss these in class.

Explain each of the following terms

- router
- ISP
- IP Address
- Domain name
- IP Protocol
- TCP Protocol
- HTTP Protocol
- SMTP Protocol

Answer the following questions:

- What's the difference between a circuit switched and a packet switched network?
- How have open standards helped the Internet develop?
- What is net neutrality and why is it important?

Lab: How the Internet Works

Key Concept A. The Internet is a network of autonomous systems.

Learning Objective 24: The student can explain the abstractions in the Internet and how the Internet functions.

- 24a. Explanation of how the Internet connects devices and networks all over the world.
- 24b. Explanation of how the Internet and the systems built on it facilitate collaboration.
- 24c. Description of evolving standards that the Internet is built on, including those for addresses and names.
- 24d. Identification of abstractions in the Internet and how the Internet functions.

Let’s Watch the Internet

Cnet provides a tool that measures the bandwidth of your Internet connection. Bandwidth (http://en.wikipedia.org/wiki/Bandwidth_%28computing%29) is the amount of data that can be transferred over your connection per unit time. It is usually measured in kilobits per second (Kbps), thousands of bits per second, or Millibits per second (Mbps), millions of bits per second.

Akamai (http://en.wikipedia.org/wiki/Akamai_Technologies) provides some nice visualization tools (http://www.akamai.com/html/technology/visualizing_akamai.html).

There are a couple of command-line tools that we can use to watch packets traveling on the Internet. But first you have to find the command line on your system:

- Find the Terminal or Konsole program on your workstations.
- (Linux Users: Look under System)
- (Mac Users: Open Applications < Utilities < Terminal)
- (Windows User: Open the Command Prompt)

Traceroute (http://en.wikipedia.org/wiki/Traceroute) is a network diagnostic tool that lets you display the path that your packets take across the Internet.

$ traceroute google.com

traceroute: Warning: google.com has multiple addresses; using 74.125.224.163

traceroute to google.com (74.125.224.163), 64 hops max, 52 byte packets

1 192.168.1.1 (192.168.1.1) 1.302 ms 0.919 ms 0.757 ms

2 10.4.144.1 (10.4.144.1) 8.513 ms 9.206 ms 8.140 ms

3 68.9.8.225 (68.9.8.225) 10.512 ms 10.055 ms 9.773 ms

4 ip68-9-7-65.ri.ri.cox.net (68.9.7.65) 11.274 ms 10.933 ms 11.802 ms

5 ip98-190-33-42.ri.ri.cox.net (98.190.33.42) 9.947 ms 9.830 ms 10.313 ms

6 ip98-190-33-21.ri.ri.cox.net (98.190.33.21) 13.648 ms 12.782 ms 11.729 ms

7 provdsrj01-ae3.0.rd.ri.cox.net (98.190.33.20) 11.678 ms 12.195 ms 11.936 ms

8 nyrkbprj01-ae2.0.rd.ny.cox.net (68.1.1.173) 17.775 ms 17.548 ms 17.816 ms

9 209.85.248.178 (209.85.248.178) 21.314 ms 17.731 ms 22.360 ms

10 209.85.251.88 (209.85.251.88) 18.121 ms 21.036 ms 18.488 ms

11 209.85.249.11 (209.85.249.11) 28.026 ms 55.017 ms 28.288 ms

12 66.249.95.149 (66.249.95.149) 39.955 ms 38.041 ms 50.290 ms

13 72.14.233.86 (72.14.233.86) 56.257 ms 65.504 ms 54.386 ms

14 64.233.174.142 (64.233.174.142) 98.997 ms 91.343 ms 89.584 ms

15 64.233.174.189 (64.233.174.189) 90.320 ms 89.137 ms 168.799 ms

16 72.14.236.11 (72.14.236.11) 90.281 ms 92.398 ms 100.019 ms

17 lax02s01-in-f3.1e100.net (74.125.224.163) 91.074 ms 99.050 ms 92.037 ms

$

Ping (http://en.wikipedia.org/wiki/Ping) is a utility to test whether a host on the Internet is reachable: that lets you display the path that your packets take across the Internet and ti reports the round-trip time for messages to that host.

$ ping google.com

PING google.com (74.125.239.14): 56 data bytes

64 bytes from 74.125.239.14: icmp_seq=0 ttl=48 time=91.242 ms

64 bytes from 74.125.239.14: icmp_seq=1 ttl=48 time=89.957 ms

64 bytes from 74.125.239.14: icmp_seq=2 ttl=48 time=91.471 ms

64 bytes from 74.125.239.14: icmp_seq=3 ttl=48 time=90.506 ms

64 bytes from 74.125.239.14: icmp_seq=4 ttl=48 time=91.070 ms

…

Internet Infrastructure

The Internet is a network of independent networks – independent in the sense that the local networks use different protocols (http://sumansinformationtechnology.blogspot.com/2010/01/what-is-internet.html) to transmit data among their computers.

Routers running the Internet Protocol (IP) (http://en.wikipedia.org/wiki/Internet_Protocol) connect the different local networks together, creating the Internet. The IP takes care of routing data through the Internet and translates data from a local protocol (such as Ethernet) to IP and vice versa.

Analogy: You can think of the local networks as different countries where the citizens are connected by different languages. The IP is like a translater that translates from French to English.

Example

Suppose you type the URL of your school's home page into your browser. may seem like your browser (the client) is directly connected to my web page (on the server), as in the top half of the following diagram (http://en.wikipedia.org/wiki/File:IP_stack_connections.svg).

But your request and the server's response travel through several Internet abstraction layers as shown in the bottom half of this diagram.

Tracing the Data Flow

At the **application layer**:

- Your browser uses the HyperText Transfer Protocol (HTTP) (http://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol) protocol, the primary World Wide Web (WWW) protocol.
- The HTTP protocol requests a translation of the host name portion of the URL www.cs.trincoll.edu into an IP Address (http://turing.cs.trincoll.edu/~ram/cpsc110/inclass/internet/) of the server that stores my home page from a Domain Name System (DNS) (http://en.wikipedia.org/wiki/Domain_Name_System) server.

There are many other application layer protocols, including:

- File Transfer Protocol (FTP) (http://en.wikipedia.org/wiki/File_Transfer_Protocol)
- Simple Mail Transfer Protocol (SMTP) (http://en.wikipedia.org/wiki/Simple_Mail_Transfer_Protocol)
- Post Office Protocol (http://en.wikipedia.org/wiki/Post_Office_Protocol)
- Secure Shell (http://en.wikipedia.org/wiki/Secure_Shell) – remote terminal sessions

At the **transport layer**:

- The Transmission Control Protocol (TCP) (http://en.wikipedia.org/wiki/Transmission_Control_Protocol) to is used to insure a reliable, ordered delivery of a stream of data from your browser to the trincoll.edu server.

At the **internet layer**:

- The Internet Protocol (IP) (http://turing.cs.trincoll.edu/~ram/cpsc110/inclass/internet/) is used to transmit 1500 byte packets of data through the interet. Each packet is sent separately from the client to the server.

At the link layer, the **hardware layer**.

Various protocols are used to transmit the data across different local area networks. Some of the protocols include:

- Ethernet (http://en.wikipedia.org/wiki/Transmission_Control_Protocol)
- Digital Subscriber Line (DSL) (http://en.wikipedia.org/wiki/Digital_subscriber_line)
- Point-to-Point (PPP) (http://en.wikipedia.org/wiki/Point-to-Point_Protocol)

Packet Switching

The Internet (the IP) is based on packet switching (http://en.wikipedia.org/wiki/Packet_Switching). Data are broken into 1500 byte blocks (8 bits per byte), which are transmitted from router to router through the Internet.

This contrasts with **circuit switching**, the technology that land-line telephones used to use, in which a continuous circuit was set up through switches from one end of the call to the other.

IP Addresses

An IPv4 address (http://en.wikipedia.org/wiki/IP_address) uses a 32-bit IP address, broken into 4 8-bit segments each represented by a decimal number. An IPv6 address uses a 128-bit address, broken into 8 16-bit segments represented as Hexadecimal numbers:

Question: Why do you think Hex is used in IPv6 instead of decimal?

Domain Names

An Internet domain name is organized into a number of levels as shown in this diagram (http://en.wikipedia.org/wiki/Domain_name). A domain name takes the following form:

```
fourth-level-domain  .  third-level-domain  .  second-level-domain   .   top-level domain

      turing         .          cs          .       trincoll         .         edu
```

A hostname is a domain name that is associated with an IP address. For example, the following are examples of hostnames:

- trincoll.edu
- www.trincoll.edu
- turing.cs.trincoll.edu

But top-level domain names, such as com and edu are not hostnames.

Domain Name Service

Domain names are mapped into IP address by the Domain Name System (http://en.wikipedia.org/wiki/Domain_Name_System), a network of servers that keep track of the mappings.

In this example, three look-ups are need before the IP address is resolved.

The World Wide Web

The World Wide Web (WWW) is not a network – technically speaking. It is a system of interlinked hypertext documents (http://turing.cs.trincoll.edu/~ram/cpsc110/inclass/internet/http://turing.cs.trincoll.edu/~ram/cpsc110/inclass/internet/) that can be accessed via the Internet.

The Web is a service governed by the HyperText Transfer Protocol (HTTP) (http://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol) protocol.

It was invented by Sir Tim Berners-Lee (http://en.wikipedia.org/wiki/Tim_Berners-Lee), who, to his credit, turned his invention into an open standard.

“I just had to take the hypertext idea and connect it to the Transmission Control Protocol and domain name system ideas and ta-da! the World Wide Web.”

Discussion Questions

- Would the Web be the success that it is if Berners-Lee had made it a proprietary system?
- Would the Internet be the success that it is if it was not based on open standards?

Internet Activity

This activity is a followup of How the Internet works. Create a single written document (such as Microsoft Word) that answers the questions below. Work by yourself or with a partner to answer these questions.

The purpose is to try to understand something about what the Internet is and explore a few examples of how you can using computation and computational thinking in ways related to the Internet.

There are four general parts to this: a map of the internet, understanding that the Internet is a network of autonomous systems (regional networks), thinking about IP addresses and DNS — the Domain Name System, and tracking how packets travel on the Internet.

Each of the following four sections have questions that you (and your partner, if you have one) will answer.

Map of the Internet (10-15 minutes)

First let's look at a few examples to get the conversation started. Here's a website that says it's a map of the Internet: http://internet-map.net/ (http://internet-map.net/)

Click around, zoom, and in general play with this for a couple of minutes.

Task: Add a question or observation you have about the map. There can be more than one.

There are other maps of the Internet.

http://www.peer1.com/map-of-the-internet (http://www.peer1.com/map-of-the-internet)

http://www.telegeography.com/telecom-maps/global-internet-map/ (http://www.telegeography.com/telecom-maps/global-internet-map/)

http://xkcd.com/195/ (http://xkcd.com/195/)

Autonomous Systems (10-15 minutes)

Conventionally the Internet is often referred to as a "network of networks". A slightly different view is that the Internet is a network of autonomous systems that communicate with each other using the Border Gateway Protocol (BGP). Let's find out something about autonomous systems in this activity.

Task: Answer the three questions about Autonomous Systems

- How many autonomous systems (AS) are there today on the Internet?
- What is the AS number of Duke University?
- What entity/company is associated with AS 16631?

IP Addresses and Domain Names (10-15 minutes)

An IP address identifies a machine/device connected to the Internet; it's a 32- or 128-bit number, e.g., 184.84.228.110 for an IPv4 dotted-quad. People like names better than machines, names like whitehouse.gov. We'll explore names and addresses, specifically how they tie together and how they work (briefly).

You can use http://www.kloth.net/services/nslookup.php (http://www.kloth.net/services/nslookup.php) or http://www.kloth.net/services/dig.php (http://www.kloth.net/services/dig.php) to answer these questions, there are other services as well that wrap dig and nslookup in a website. For geolocation of an IP address you can try http://geomaplookup.net/ (http://geomaplookup.net/) or http://www.infosniper.net/ (http://www.infosniper.net/)

Task: Enter answers to the following four questions

- What is the domain name registered to 65.39.205.54?
- What is the IP address associated with girlswhocode.com?
- What is the IP address associated with csprinciples.org?
- In what city is the IP address of csprinciples.org located?

Traceroute (10-15 minutes)

The traceroute command shows the “hops” that a packet will take in getting from one place on the Internet to another. This command/tool can be run via several websites, some that offer visualizations. We’ll use http://traceroute.org (http://traceroute.org) which maps several sites in other countries to your IP address or one you specify. For visualizing you can try either http://traceroute.monitis.com (http://traceroute.monitis.com) or http://www.yougetsignal.com/tools/visual-tracert/ (http://www.yougetsignal.com/tools/visual-tracert/)

For example, using the pages at traceroute.org, if you start from the site given in the country of Kyrgyzstan 21 hops are shown to reach Duke (output from website)

traceroute to 152.3.250.1 (152.3.250.1), 30 hops max, 72 byte packets

1 r2d2 (212.42.96.42) 0.255 ms 0.262 ms 0.246 ms (0% loss)

2 R76 (212.42.96.80) 0.340 ms 0.310 ms 0.314 ms (0% loss)

3 195.218.197.85 (195.218.197.85) 57.587 ms 57.609 ms 74.506 ms (0% loss)

4 p4.Moscow.gldn.net (195.239.11.166) 81.353 ms 85.248 ms 83.839 ms (0% loss)

5 mx01.Stockholm.gldn.net (194.186.156.41) 75.101 ms 75.684 ms 75.102 ms (0% loss)

6 xe-11-0-0-xcr1.skt.cw.net (166.63.220.65) 75.182 ms 90.087 ms 75.216 ms (0% loss)

7 xe-0-3-1-xcr2.amd.cw.net (195.2.9.217) 201.529 ms 176.642 ms 177.156 ms (0% loss)

8 ae0-xcr2.amd.cw.net (195.2.30.105) 175.810 ms 175.820 ms 175.656 ms (0% loss)

9 xe-7-0-0-xcr2.lsw.cw.net (195.2.25.38) 176.593 ms 176.578 ms 176.607 ms (0% loss)

10 xe-2-2-0-xcr2.lnd.cw.net (195.2.28.185) 178.416 ms 175.892 ms 175.479 ms (0% loss)

11 xe-3-1-0-xcr1.bkl.cw.net (195.2.21.214) 103.025 ms 102.920 ms 103.093 ms (0% loss)

12 xe-8-1-0-xcr1.nyk.cw.net (195.2.25.26) 176.424 ms 176.293 ms 200.133 ms (0% loss)

13 nyc-brdr-02.inet.qwest.net (63.146.27.61) 174.859 ms 174.929 ms 174.798 ms (0% loss)

14 * * * (100% loss)

15 65.121.156.210 (65.121.156.210) 200.693 ms 216.372 ms 200.503 ms (0% loss)

16 rtpcrs-gw-to-wscrs-gw.ncren.net (128.109.212.9) 223.205 ms 203.419 ms 203.567 ms (0% loss)

17 rlgh7600-gw-to-rtp-crs-gw.ncren.net (128.109.9.6) 206.194 ms 206.184 ms 206.192 ms (0% loss)

18 roti-gw-to-rlgh7600-gw.ncren.net (128.109.70.18) 215.218 ms 215.790 ms 215.372 ms (0% loss)

19 tel1u-datacenter-vrf.netcom.duke.edu (152.3.234.25) 204.385 ms 204.414 ms 205.012 ms (0% loss)

20 tel2u_3_12_4.netcom.duke.edu (152.3.250.254) 215.001 ms 215.041 ms 214.813 ms (0% loss)

21 dukedns1.netcom.duke.edu (152.3.250.1) 204.437 ms 204.365 ms 204.392 ms (0% loss)

Your task is to pick 3 cities/locations from http://traceroute.org (http://traceroute.org) and use tools to find the number of hops to either your machine (sometimes that's the default) or to Duke's nameserver which is 152.3.250.1 — if you have a choice of how to connect, use traceroute. Many of the links at traceroute.org don't work, click around as needed.

Task: Enter the three cities and the number of hops

Reflection on these Internet Tools

- What impact has the Internet had on solving problems and gathering information?
- What problems has the Internet created? What are we doing to address these problems?

This is a test/project grade.

Portfolio Project: Internet

Learning Objectives

The Internet Portfolio Task addresses the following CS Principles Learning Objectives:

- 27: The student can explain the abstractions in the Internet and how the Internet functions.
- 28: The student can explain characteristics of the Internet and the systems built on it.
- 29: The student can analyze how characteristics of the Internet and systems built on it influence their use.
- 30: The student can connect the concern of cybersecurity with the Internet and systems built on it.
- 31: The student can analyze how computing affects communication, interaction, and cognition.
- 33: The student can connect computing with innovations in other fields.
- 34: The student can analyze the beneficial and harmful effects of computing.
- 35: The student can connect computing within economic, social, and cultural contexts.

Task

- Identify and describe a significant, contemporary problem and potential solution that are connected by the Internet to a societal, economic, or cultural context. The problem and/or potential solution must have a strong connection to computing and the Internet.
- You will create a document describing the problem, the potential solution, the societal, economic, or cultural contexts, cyber security concerns, and the relevant characteristics of the Internet. The societal, economic, or cultural context should affect a significant population more than a few hundred people).
- Note: In the task description and the text below the phrase “problem and/or solution” means that either the problem, or the solution, or both should have the characteristics referred to. For example, the task description requires that the problem, or the solution, or both, must have a strong connection to computing and the Internet.

Prepare and submit the following

A single written document that addresses the task and meets all the requirements listed below.

- The document can include illustrations and non&textual examples.
- The document does not need to be an essay. For example, it can be a sequence of written paragraphs each of which addresses one of the requirements below, but which does not include words to connect the paragraphs together.
- You can also write an essay in which the paragraphs are connected.

Requirements

- Maximum length of essay is 1000 words.
- MLA Format, Double-spaced, Times New Roman 11 pt
- Works Cited, MLA format
- You will also do a 30-second Talking Point in class
- Identify at least one cyber security concern related to the problem and/or the potential solution depending on which of these are connected to the Internet).

- Describe the cyber security concern and explain how it relates to the problem or its potential solution or both.
- Provide a brief but clear description of the problem.
- Identify and describe the problem clearly.
- What population is affected and how?
- If the problem is connected the Internet describe how. Be sure to address the specific characteristics of the Internet (see Description of Internet Characteristics) relevant to this problem and how the Internet and the problem are connected.
- Provide a brief but clear description of the potential solution.
- Describe the potential solution clearly and how it addresses the problem.
- If the solution is connected to the Internet describe how. Be sure to address the specific characteristics of the Internet (see Description of Internet Characteristics) relevant to this solution and how they enable the solution.
- Connect your references to the descriptions of the problem and the solution. That is, your descriptions must make explicit references to your sources.
- Describe one potentially beneficial and one potentially harmful effect the Internet has on society, culture or economics in the context of the problem or the potential solution.
- Identify and describe one field other than computing that contributed to or is impacted by the problem or the proposed solution, citing at least one specific example. Describe how the field contributes to or is impacted by the problem or its solution.

Your document must make references to two external sources that provide context for the problem and solution.

- The references should be high quality sources.
- Each reference should include a URL when appropriate, but also a citation for the reference, e.g., the author and title of the reference.
- You must include a date on which you accessed the source if the source is online.
- Each reference should be to something written between September 2012 and today.
- The references should be to sources of information that anyone can use to learn about the problem or its solution.

Description of Internet Characteristics

When you describe the characteristics of the Internet relevant to the problem or its solution or both) you must identify, describe, and explain at least the characteristics below that are relevant and related to the problem or the solution. You only need to identify, describe, and explain those characteristics below that are relevant and related to the problem and/or solution. You may identify characteristics of the Internet in addition to those below.

- Evolving Internet standards and abstractions e.g., addresses and names)
- Hierarchy and redundancy in the Internet
- Interfaces and protocols of the Internet that enable widespread use
- The trust model of the Internet and its role in cybersecurity
- How cryptography affects cybersecurity

Rubric

Essay (85 pts)

- Identify and describe a significant, contemporary problem
- Identify and describe a potential solution to the problem
- Identify societal, economic, or cultural context of the problem/solution
- Describe how population is impacted
- Identify and describe the problem's/solution's connection to the Internet
- Identify & describe at least one cyber security concern related to the problem and/or the potential solution

- Identify and describe one field other than computing that contributed to or is impacted by the problem or the proposed solution, citing at least one specific example. Describe how the field contributes to or is impacted by the problem or its solution.
- At least two appropriate references used as assigned

Works Cited (10 pts)

30 sec Talking Points (5 pts)