

Transcript
Lecture 5: The Internet, Continued
18 October 2006

(00:00:00)

Now that the drum roll is complete, welcome back to computer science E-1. My name is David Malan. This is lecture 5, our second lecture on the internet. Allow me to introduce tonight, by way of the film above you.

video

So, that is a teaser for a short film we will watch toward the end of tonight's lecture, but before we dive into what is TCPIP, and all of that is all about, we are at the point in the semester where it's time to raffle off this iPod T-shirt. Recall that for problem set one, among the students that submitted, if you received higher than a 75%, you are automatically entered in a raffle. The first prize that we will be raffling off tonight, to students, both local and distance alike, is this t-shirt of the first time revealed to the class. This is apparently signifying laptop plus iPod...ok, and on the back we have a little advertisement, it seems. College students buy a Mac and get a free iPod. So, apparently, I'm not sure under what circumstances you get this t-shirt though, but it came to us, none the less, donated by Apple. With that said, I will do a moment of suspense here...and the winner of this t-shirt, this semester, is Mr. Gerald Walden If Gerald here tonight? Gerald, come on down. Around of applause, of course, please.

applause

It is entirely your prerogative if you wish to wear this t-shirt. Congratulations, and well done on problem set one.

STUDENT: Inaudible

So, least ye think there is some funny business going on. Know that Dan Armendariz served as witness of the one line program that I wrote before lecture tonight, that generated a random number between 1-40 that all owed us to select from among those of you who were the lucky winners. I'm going to hold you in a bit more suspense with regard to the iPod Shuffle itself. Let's turn our attention, again, to this topic of the internet.

Frankly, Google Earth is just too cool not to revisit in this class. Tonight will be entirely about Google Earth, if only because it's the coolest thing that I have wasted several hours on over the past couple of weeks. How many of you, out of curiosity, downloaded this program after last week? So, a few of you did. You will be tasked in problem set 4, if you haven't noticed already, with the downloading of this program, in order to satisfy 2 of the problems demands, which is to seek out some fairly popular attractions somewhere in the world, without us giving so much as directions or longitude or latitude, or addresses, and the second part of the problem set will ask that you seek out 3 places of interest to you, personally. This is more than just a fun thing to play with. Know that the

software is entirely free. That means we have the luxury of playing with it, with the expectation that others in the class, and beyond are certainly welcome to do the same. But, this is actually, as we explained in the problem set, a wonderful opportunity, truth be told, to appreciate, if you don't typically do so, that you don't need training to use this new software, or at least you shouldn't need training. You shouldn't need a user's manual to use new software. You don't need to have certifications to use new software, at least in the ideal sense. We try to assure you, in the problem set statement itself, if you struggle the first time you sit down in front of Microsoft Power Point, or Microsoft Excel, or whatever new application you are required to use at your company, (to be honest) it is probably more the fault of the application, its design, its development frankly, than of you-the user. A number of you enter this course with a bit of apprehension as to what you can do with a computer, and daresay you might need training on certain programs and software. Well, the goal of this problem set, in part, besides just having fun and taking a little vacation in your homes, is to introduce you to a program that is fairly well designed. I think it is fairly intuitive. I will give you a cursory tour of some of the features, but to be honest, the second, sort of implicit goal of this problem set is just to give you an opportunity, within the span of just a few minutes, to download a program and then go figure it out. That alone can be sort of empowering.

With that said, we have this spinning globe here. I thought I would take us right away, as we may have done last time, into 1 Oxford Street, Cambridge, Mass 02138. Just a quick tutorial and this is the extent to which you will be trained on this software, for this problem set. That's precisely the intention, to give you very little training. At the top left of this program, if you have not used it before, is a little search box. You can type in an address in there. Google is pretty good about knowing if you type Eiffel Tower, where in the world that is. It's pretty good about guessing where certain popular locations are. But, you can more precisely specify a US address, a foreign address in most countries, a longitude and latitudes. There are other tricks as well. I'm going to go ahead and click enter, after typing that in, and if nothing else, to be honest, it is perhaps the fundamentally unnecessary, there are completely gratuitous animation that makes this program so cool. As we zoom in here on our location, 1 Oxford Street, of course, is where? Right here. So we have now a bird's eye view from a bird that is currently flying at about 3,000 feet. You can see in the bottom right hand of the screen your altitude, at least in feet. You can change that to meters, or kilometers, if you prefer via one of the menu options.

But, notice (and it is a little hard to see if you don't know to look for it, certainly on our screen here, in the top right of the screen, where my cursor is) there are a whole bunch of controls. There is a plus and minus, a vertical scroller, so if you simply click on that plus, as you might expect, things zoom in. If you instead click minus, you are going to zoom out. And there is also a slider there, with which you can move it back and forth without just clicking. It should be fairly intuitive if you click on the thing yourself. Meanwhile, you can certainly move horizontally and vertically and left to right and so forth to get a different vantage point. Just like Google maps, if you are familiar, you can click on the display and just drag it around, which is perhaps even more intuitive. And if you are really adventurous and zoom in, for instance, just a little bit, and you kind of wonder what this building looks like from a different angle, you can use this slider up top to change your tilt. You will get a different perspective on the world. Notice though, a curious thing happens, the more and more we go down. It turns out that the Earth is , indeed, flat. But this is obviously why? This resolve... why are we seeing this?

STUDENT: Inaudible

Right, so you don't have a satellite looking in your window. You have it looking on top of your house from up above, but there are certain parts of the world where you do have multiple angles. You do have software that is designed to interpolate even better what the shapes of those mountains are, or what the canyons are. We will go to one such place in just one moment. This, meanwhile, I'm going to tilt us back up to a more reasonable level. You have this big wheel, which by default points you north, but is you click on the big wheel, and spin yourself around; you can sort of do a nice 180spin in that way as well.

Meanwhile, do take note, if you haven't seen it already, in the bottom left of your screen; you have a whole bunch of layers. This is a fun place to play around because it offers different features of this program. Notice that I have checked terrain. That's a good one to leave on because it is for those areas of the world where Google has, sort of, height information. It will show you more of a 3D perspective, rather than an interpolated flat perspective. There are a whole bunch of other options that you can play around with. If you want to find the nearest pizza place in some area, you can click on such things as dining, or the nearest hotel. It operates in that regard like Google map.

Our purpose tonight is to kind of "wow". So notice in the middle left of the screen, under the thing called places. This is where you can store places that you like. Essentially what problem set 4 asks you to do is to "place mark" a number of specified locations. This is like book marking, but it is a physical location. It is terribly simple. There is a little thumbtack here at the top of the program. You click that thumb tack, and you will notice you get one of these place marks (in the middle of the screen.) You can then click it and drag it around. For instance, I'm putting it above Memorial Hall, which is just across the way, outside this building. Up here you can type a name for it. I could type Memorial hall. It tell you already what the longitude and latitude is, if that's at all of interest. Then you can type a little description here, like where freshmen eat meals, for instance. Ember Hall is inside Memorial hall. Then you click ok, and that's as though you have placed a book mark of sorts into this actual world.

What we are going to ask you to do from problem set 4 is to submit, via the website's drop-boxes, those place marks. The means by which you do this is very simple. Notice that this place mark ended up under this "My Places" section of the page. Moreover, as the problem set does tell you to do, notice that if you, for instance, right click on this with a PC or hold and click on a Mac, notice you can say something like "add to folder" and what we ask you to do in the problem set is to create a folder called "3" I think, and another folder called "4" respective to problems 3 and 4 for the problem set. So, I'm going to type 3 as the name. Notice I get this 3 folder and as you might hope, just simply click and drag "Memorial hall" to your 3 folder. And as with other interfaces the plusses and minuses that appear, that will expand the thing or collapse the thing.

Finally what you will do for problem set 4 is when you are all done, and you have found all the locations in the world that we have asked you to find, you will simply (for instance on a PC you would right click. On a Mac you would hold and click on a Mac) and then you are going to say "Save as" and then you get a familiar dialogue window. It can go to your desktop, and say "My place marks" or rather, we will call it "3" in this case. Notice the default file name is going to be .kmz it's a little small, but you will see as much on your own machine, and then just save it. What I just did was

put on my desktop a file called “3.kmz” and that is one of the files that you will upload for submission. What that’s going to let the teaching fellows and me do, is open your place marks, click on those hyperlinks and whisk ourselves away, not only to the places you found, but also the places of personal interest that you drew our attention to.

Any questions at all? To be honest, if that’s not sufficient directions, then frankly, you have the video as the reminder. It’s not meant to be an exercise in remembering what David said about how this program works, it is meant to be exploratory. If you don’t know how to do something, poke around. You are certainly not going to break the software. It’s only so much fun if listening to me talk, why don’t we take ourselves on a very brief tour, perhaps to one of these places where the terrain is of interest.

Notice that Google gives you, by default, this sightseeing folder. We created an empty folder a moment ago, but the program comes with a folder called sightseeing. It has a whole bunch of popular locations. Why don’t we whisk ourselves over to the Grand Canyon? It suffices to just double click on that thing. We are heading out west in this country. We are about to zoom in. The Grand Canyon is a wonderful example. If you have the terrain layer on, you will actually get this 3D map thing. If you really want to have the kind of “Imax” experience, you can use the keyboard to navigate these controls, besides just using the top right, you can “fly” your way through the Grand Canyon. You can give yourself a little tour. It’s really wild. Notice, by contrast, if I turn off terrain, there you go, That’s the grand canyon. So, it does make a difference.

Finally, let me offer you one other demo of my own design. Let’s whisk ourselves away to Boston, Ma. And again, little tricks like Boston, Ma that is sufficient. Google is smart enough to figure out what you mean. We are going to jump back over here, back to Massachusetts. It’s going to zoom in. here we have the main part of downtown. If I drag it over, you can see Boston Common over there (the big green area in the middle.) Let’s focus in on the financial area or downtown and click on one of these layers, namely 3D buildings. First I’m going to zoom in and you can see that these are pretty good photos of these buildings. But if you really want to get a 3D perspective, some cities have some have support of 3D buildings. People have built virtual models of these buildings such that now if we dive into the middle of Boston, and sort of change our perspective, you can actually see what the city looks like from a 3D perspective. You can do this for some of the major cities, New York is another one, and so forth.

Now, this is where you get to partake. Now, before we get back to understanding the internet, let’s use it for just one more minute or so. Where in the world would you like me to take you on this tour?

STUDENT: Inaudible

London... so London, UK... we whisk ourselves over the Atlantic... and in just a moment, we will find ourselves in central London. So, there’s the [Inaudible], looking murky, as always. And if scroll ourselves here, this is the London Eye that it is taking us to. So, the big Ferris wheel of sorts in London.

Oh, one other thing, especially for those of you who use public transportation, let me take up back a moment to Boston, Mass. Let me show you one other layer that's fun. I'm going to turn off 3d buildings by un-checking it, then I'm going to turn on, under transportation, I'm going to turn off transportation. If you have ever wondered, topologically, what the MBTA system looks like, you can see it now with this nice overlay, not only for the subway, but for the commuter rails, if you zoom in. The coolest thing, frankly, I never get tired of this, if you just hold on the keyboard to zoom out, I mean it's like you are making your own movie, 2001 style.

15 minutes 10 seconds

One more place. Where do you want me to take you?

STUDENT: Inaudible

Big Bend national park... let's see. Here we go. Ok, good, another one where terrain is nicely depicted. Here too, I will tilt us back on an axis. You can really see the 3D there. It's pretty wild. So, in any case, have fun We are giving you 20 or so points to sit at home having fun and maybe showing off to your family or friends exactly what your homework is on that part of the problem set.

There is more to the problem set, and realize that the problem set is not due until the 8th of November. It may seem like a long way off, but this problem set is fairly involved, not in a difficult way, but in a fun, as we say, way.

Among the other challenges in the problem set, is to actually go to a local computer store and take your new-found vocabulary, jargon, and comfort (presumably) with hardware and software, and the internet to a local computer store and "buy" with some virtual dollars of ours... \$2,000 virtual, some compute equipment. This process is all about empowering you with a bit more comfort in 1 reading the tags on the shelves; and 2, giving you, finally, that sense, if you have never experience t before, that you do, in fact, know more than the sales man or sales woman who is trying to sell you that computer; and 3 as we qualify in the problem set, we do suggest (because we do this every year and alienate all local computer stores about once a year) that you do not advise the staff until you are done with your assignment that your dollars are, in fact, virtual. You won't get nearly as much assistance.

With that said, are there any questions about problem set 4... before it, or anything beyond this point in the course? All right, so enough about Google.

So, I connected my mom's computer, remotely, in Connecticut last night and installed this on her computer Google Earth, and that killed the rest of her evening; taking a look at her favorite places. It's addictive, so beware.

All right, so... The internet. Last week, and certainly part of tonight, we focus on the things you can do with the internet; the types of services that it offers. Let's take one step back to last week and answer the following questions since it is one that recurs every year... are the internet and the World

Wide Web the same thing? So, no. let's qualify that now. Why, what's the distinction between these two?

STUDENT: Inaudible

So, the internet is the operating system within the World Wide Web. I tweaked that slightly, and before I tweak, let me see if there is another suggestion.

STUDENT: Inaudible

Let's go with this one because it is more precise. The way the internet is a network of networks. It is an infrastructure. I would beware using the term operating system only because it is too easily conflated with desktop operating systems. It is an infrastructure, it's a framework. It's a backbone on which applications and services run. So when we talk about the World Wide Web we are actually talking about something you can do on the internet. And this use of preposition is not coincidence. It is really the fact that you are using a service on top of the internet; whereas the internet, itself, though you can not necessarily point to it and say that's the internet, because it is this network of networks. It is more the physical infrastructure underlying all the service you run.

So tonight is ultimately about is about "lifting up the hood" of the internet; looking at how it works. When it breaks, what does that mean? When your own computer doesn't connect to the internet properly, what does that mean? When you call up Comcast or Verizon and they are trying to diagnose why your internet connect is down, what kind of question can you, yourself, ask or answer for them? How, in short, might you go about picking up that phone and rattling off so many technical words in the right order, and with the right usage, that the person on the other end is convince that you don't need to spend the next hour on the phone with them. You can get bumped right up to, lets say level 2, or level 3 tech support. It's a wonderful skill, if you can rattle off more information than the person on the other line understands, usually, though not always, they will realize that it will be easier for them to just pass things up to someone else. Frankly, certainly early on, you can probably get away with using all these words and not necessarily in the right order, but enough of them that you get your way past the initial tech support call anyway.

So, let's start, as always, from the bottom up and build up to the more juicy and interesting things. You've got a couple of computers. Let's say you have 3 computers. How could you go about topologically, connecting these things together? What does this mean? Well, last week we talked about the smallest possible network, that's at least interesting. I had 2 computers and I drew our very dated picture of a computer like this. We said we can connect it by some wire or some connection of some sort. And there you have what kind of network?

STUDENT: Inaudible

Peer-to-peer network. And again, to be sure that term has taken on a broader meaning these days in terms of the Napster-like software of the world and so forth. Peer-to-peer can quite literally mean that. 2 peers, 2 equal computers effectively communicating back and forth. This is just to make the distinction between a similar pairing, but in a relationship in which I intentionally drew one computer bigger and when we drew one small computer connected to one big computer via some

connection, we described that relationship in terms of 2 other buzz words- client server. None of that was meant to say that servers are always bigger than clients, but is just meant to suggest that just like a restaurant can handle multiple clients, multiple customers at once, so can a server handle multiple requests, for instance if it is a web server, from multiple clients at once.

Now enter in computer number 3. Well, it would seem, just intuitively, that we already have those 2 computers together, how in the world can we introduce a third? Right? We could disconnect the middle guy and plug the left guy into the right guy, but now we still have a peer-to-peer network of 2 computers. Ok, so using a server, I heard.

STUDENT: Inaudible

Turn one into a server. “What else might we do? What else might we do? So, let’s go there. This is probably a device that you are all perhaps owners of, if you have multiple computers connected to your internet connection. Well, rather than trying to connect them all together, we could certainly connect them to one central computer as was suggested, a server. And that’s server can sort of figure out how information can go from (and this might be our server here) how information might go from A to the server, and from the server could send it to B, or the server could send it to C. Well, it turns out that that is a very reasonable solution, but it also turns out that this notion of routing data among computers from A to B, from B to C, from C to A, it doesn’t even require a computer or a server in the conventional sense.

This problem can be implemented in various pieces of hardware. So, yes you can view these pieces of hardware, routers as they are called, as computers and they are, but they are not computers in the monitor/keyboard/ mouse sense, at least.

So, one way we might connect all these computers together then, is not so much directly from one to the other. And we don’t need a whole other desktop and monitor to connect these guys to; all we need is some central point. Connect this computer here, this computer here, and this computer here. This solves the problem of these computers, presumably only having one port on the back on the back into which you can plug in a network cable. Recall in our first or second lecture when we looked briefly at some ports on the back of a computer. Maybe in section you did this as well. Well, you take these telephone-jack-like plugs. They are a little bit fatter than the typical telephone jack, well, typical computers these days only has one of these jacks. If you only have one of them, you’ve got to put the other end of the cable into something and it clearly doesn’t work just to connect two together, because then the buck stops here. You can’t go any further. So, you connect them to a central device like this that clearly must have multiple Ethernet jacks, as they are called. So, the technology via which many computers, daresay, most computers today, at least so far as we are concerned, the technology they tend to use, is something called Ethernet.

So, if you ever hear someone talking about an Ethernet cable, or an Ethernet jack, for all intents and purposes, they are talking about this standard telephone-like cord and the jack it plugs in to. The jack, specifically just to throw out something you can impress others with, is called an RJ45 connector. This is in contrast with a telephone (random trivia) that is called an RJ 11 connector. Just think of, it doesn’t necessarily grow out of this, but 45 is bigger than 11, so 45 is the fatter cord.

That might be a simple way of remembering that. Frankly, if you never remember that, you won't be terribly worse off than you are right now.

But, seeing it in print is useful, for instance, if you go to the store and you want to buy a cable so you can connect your computer to one of these central devices. What do you get? Well, you ask for an Ethernet cable, a network cable, a cable with RJ45 connectors on the other end. You might see things labeled differently depending on the circumstances.

But let's slap a label on this thing. We called it a router, but it turns out that you don't need a router, per say, to have data transmitted among 3 different computers, called A, B, and C. This device is simply something known as a switch. So, a switch is quite simply, a relatively cheep, these days, and relatively dumb device whose sole purpose in life upon receiving data- that is bits from computer A- it looks at those bits and realizes that , you know what, these bits are intended for computer B, so it sends it from this jack (this RJ45 jack) out this jack. If it instead receives some bits from A and somewhere imbedded in those bits is the label C, well, a switch is designed to send those bits, instead to computer C. And switches are... they aren't completely dumb, they don't just spit out all the information that comes into them, rather they do intelligently select, almost always, which jack out of which it should send the received data. So, the reason we qualify switches as being distinct from routers, is that those little blue and black devices that you buy, or those silver devices that you buy called home routers or you call firewalls, they are so many things on one these days. If you have one of these devices that gives you the ability to run multiple computers off of one internet connection, you would be correct, most likely, in calling that a switch, a router, an access point, a firewall, a proxy server, and daresay, a whole bunch of other things as well. That's just because, at the end of the day, it is a computer. It has no monitor. It has no direct keyboard or mouse, but it is a little computer that has imbedded in it software that does all of those things. And it has the hardware that is the jacks that physically allow those kinds of features happen. And tonight we will tease apart some of what those features are, for instance, what a router is, what a switch is, but for now it suffices to realize that, you know what, we can in fact, connect multiple computers together with a device like this.

How much is it going to run you if you have a few computers at home and you want to connect them all together. Well, how much are you going to pay for this kind of switch? And all a switch is, and you can see this perhaps in section, is, or you can see it tin your own home, it has a power adapter which you plug into a wall. And it is a little plastic for metal box that has a whole bunch of theseRJ45 jacks, that's pretty much it; no moving parts or switches. No pun intended.

How much?

STUDENT: Inaudible

\$20? Almost always in this class, frankly, these days, everything is about \$20, \$10, \$30. It is relatively cheap these days. And if you watch for rebates and you are really careful with your money, you can sometimes get these things for free after rebates, plus shipping.

So, what topology are computers usually connected in these days. This sort of mish-mash here is meant to beg the question... When you have multiple computers, how do you connect them

together? Well, what you have seen an example of is what's known as a (I'm going to skip ahead) a star topology.

This is probably the most common way of connecting multiple computers together in a network, these days. You wire all of them to some central point. In fact, if you are really a geek, or live with a geek, who has him or her self multiple computers; or you work in an office, were you have a server room; or you have jacks into which all of your co-workers computers plug in... well, ask your IT guys sometime this week, where is your server room or where is your wiring closet, or where do all these wires go? And if the place is laid out nicely, there is somewhere in your office or building, some kind of closet or room that you can walk into and see a whole bunch of wires coming out of the wall or coming through the ceiling or from the floor. All of them get plugged in to... all of these little jacks... on a rack. So, computers and server rooms are usually laid out in big metal racks. And what you see is switches that don't just have 3 jacks, but rather might have 24 or 48 or 100, so that you can plug a whole bunch of computers into.

Well, many of you probably have those internet jacks on your walls. Let's see if there is one in the room... most everything is wireless these days. So, back in the day there would have been, on the wall somewhere, one of these jacks into which you, at work, plugged your computer. Well, all that is... is a cheep little box that has a jack and behind that jack is one of these little wires. That wire is most likely leads to that closet or room in your building or company and on the other end is one of these RJ45 jacks, and it plugs into a really big switch. And by big, I just mean that physically it is large so that you can fit a bunch of these things, so that all of the computers are through that switch. They are interconnected. And so if you have a file server at work, or a printer that is shared at work, it all boils down to them being connected, usually these days, via Ethernet, via cables and switches like these.

30 minutes 27 seconds

In yesteryear, and the reason for the other pictures, is that you could connect computers in other topologies, which we won't dwell on. And since it is not as common these days, but you could go the ring approach, which several years ago was common. If only because that's what some of the original protocols, languages that computers spoke, when networked were designed to do. And it is also cheaper when cabling was more expensive. You don't necessarily run wires from every computer in your office all the way to the basement, or all the way to the attic. Think about the waste. It is much more sensible, it would seem, to run those wires to the next closest computer and from there to the next closest and save on costs.

These days you pay pennies for this kind of cable. In fact, 2 or 3 years ago the teaching fellows and I bought a box, a spindle, of this kind of cable. The cable itself is called, usually, CAT 5, category 5. it just means it is capable of internet speeds today. We spent \$50 maybe \$80, and I thought every year we would have to get a new box because in section we use this cable. But we are still working off the same box for 3 years now, which means we spend pennies per foot of this thing. You may wonder why this teaching staff and I need a thousand feet of this cable. Well, one of the things you will do in an upcoming section, if you are local, is actually make your own Ethernet cable.

Now, that seems a bit weird, if I motivated some of this discussion of how to go buy such a cable from the store, but there's a remarkable sense again, of empowerment and also frustration, when you try to take the 8 little wires that are inside one of these cables, line them up just the right way so that you can slip on one of these half [Inaudible] RJ45 connectors. You slip those wires in and take a crimping tool that we will hand to you and squeeze it as hard as you can. You usually then pass it to the TF who will squeeze it even harder for you, and then at the end of the section, if all goes well, we will test out your cables on a working computer, and if we can bring up a web page, you've done well. And if you didn't, well... you will walk home with your head held low, perhaps. But it will be, for those of you who get it working, a wonderful souvenir to take home to friend or family.

In fact, let me take this moment, as we tend to do... in this class, despite how we try to teach you how to make computers and upgrade computers, we always seem to go through more computers and equipment than we actually retain. So, I'm going to take a knife to one end of this otherwise working Ethernet cable, and I'm going to cut off some of the plastic sheath, because even though this thing looks pretty thick, all this yellow thing is... is just a piece of plastic, or rubber that keeps all the little wires... and they are sort of tucked together and secure.

So, I'll pass this around now and any of you that have an internet connection at home that uses a cable modem or DSL modem, you have one of these in your home. If you have a wired connection at work, you have one of these in your office. So, take a look as this comes around, at the 8 wires that are now coming out of this thing. And again, in section, your goal will be to line them up in the right way and crimp them back down. But then notice too, on the other end of this thing, that those very same colors and those very same wires are lined up in a certain pattern and therein lay the challenge. You can't just plug in 8 wires on one end in any old order, plug the other end in any old order and expect them to work. They all have to be lined up so that what comes out this end goes in that end and so forth. So, I'll pass this around, and also one of these RJ45 connectors. I'll pass these around so you can see what's going on inside these cables.

This is not to say that, to be clear, that you should ever have a burning desire to crimp your won Ethernet cable. Every time I do it in my apartment, to run some new line, when I got a Tivo, when I've got a sling box, and I've just got too much free time, I end up banging my head against the wall. I sort of wonder why do we make the students do this. But, in the end, it is possible to do this.

Another way that that you can lay out a network is via a bus network. This too, is somewhat dated, at least for today's purposes, whereby you essentially run one wire or line in between all the computers and then connect each of your computers to it sort of like bus stops, or driveways on a local street. These are network topologies.

All right. So we know then how to connect computers locally. You take a switch, you take some Ethernet cables, and it takes the computers themselves. Thinking back to last week, if I draw a poorly formed building with a door around the computers, if just to put this into context... what did we call that kind of network, local let's say to one building? So... That was our LAN.

So, even though we are focusing on the details underneath the hood tonight, it's really the same tale as last week. Once you have a LAN, or a bunch of WAN, and you actually want to connect these things together... how do you go about connecting them together? Well, therein lays the design of

the internet. So, we said last week that the design of the internet grew out of US military project called ARPA net. One of whose design goals was to have a fairly well distributed and redundant work so that there never just one path, ideally, from A to B. You could go from A to C to B, from A to D to B, and so forth. In short, there was never just one wire connecting 2 important sites.

Well, what you have these days, in the more popular usage of the internet, what are generally known as internet backbones. So companies like MCI, and AT&T are the big fish, so to speak, with which you are familiar, largely from the telco days, in fact, own a lot of the underlying infrastructure which the internet ultimately is defined by. And by infrastructure, I mean really fast, really thick wires. They are thicker than these things with just 8 wires that carry not just local network's traffic, but can carry gigabytes of traffic, terabytes of traffic, theoretically. You have lines; not only in the US, but certainly you have transatlantic and transpacific cables hundreds of feet below the water that they, themselves, are carrying a whole bunch of cables. They, in spirit, work similarly sometimes to these kinds of cables which are copper wires. And you can do much better with the right kind of money and the right kinds of technology.

Some of these cables these days are fiber-optic. One of the other cables I'll pass around is this orange cable. Notice that as this thing goes around, those are just copper wires. If you pull off the plastic colored sheaths that are around the 8 individual wires, you will just see some bronze or copper colored wire. And it's just like a flow of electrons across the wire. Contrast that with something like a fiber optic cable, which (this is more expensive) and we are not going to clip the ends off of this one tonight, but how does data travel along a fiber optic cable? It's fundamentally different from a copper based cable like that.

STUDENT: Inaudible

Yeah, it's effectively traveling by way of light. So, it's not so much the flow of electrons across a metal wire, but rather flashes of light, if you will, across a medium that is conductive, and allows the passage of light. not just in a straight direction, but through a cable, in effect. You all might remember, say from grade school that even if you don't remember that light travels at what... 3×10^8 to the 8th meters/second... Well, you probably remember that light travels really fast. Whereas anything that tends to be based in electricity, or at least something that flows against resistance, like a metal wire, tends to be slower. So, fiber optic cable tends to be even faster.

And folks building homes these days, or are renovating homes, are building into them network capabilities so that they have jacks in the various rooms of their homes. They will sometimes lay CAT 5 cable these days but also some fiber optic cable, not so much because it is necessary or useful yet. But seeing how fast things move these days, it is probably useful and cheaper than pulling and breaking through the walls 10 years from now.

And just to put some more numbers and jargon, perhaps all of what you see or read something tech related, cat 5 cable is essentially capable of operating at up to 100 megabits/second; contrast that with fiber optic cables, with which can get into the gigabit range. You can also have cables that are in the 1000 megabit, or rather 1 gigabit range themselves. But, in short, if you are using that kind of wire that yellow wire we were passing around there, you can have flow 100,000,000 bits/second.

That's pretty good, in fact, think back... actually, did we... we will come back to that in just a moment... exactly where the bottle neck might be in your own home.

Yeah?

STUDENT: Inaudible

Good question. I will punt on the first questions, since I'm not sure off hand, but a quick Google search could probably turn up the cost for us per square foot of fiber optic cable. How soon will it be useful? It's already useful. Much of Harvard yard, for instance, is already wired with fiber optic cable, because even the dorms and the buildings themselves are only wired with say, 100 megabit wire, Cat 5 cable. Well, once you start connecting buildings together that suggests that you are connecting maybe a few dozen or a few hundred people in one building to the few hundred people in the other building, which suggests you need more bandwidth. That is the ability to have more bits of information flowing in between them. So, you see fiber optic cable being used certainly, in larger networks. But in the home...

STUDENT: Inaudible

It's a good question. So, if you are renovating a residential condo, is it worth wiring with fiber optic cable? I'll go out on a limb and say it's probably slightly premature. Consider, after all, a point we will get to tonight, which is that Cat 5 is capable of 100 megabits/second. We can jump ahead here because it is an interesting question. How fast is your typical internet modem... cable modem or DSL modem?

STUDENT: Inaudible

30, 30 what? 30 megabits/second... maybe, but although if it is I'm quite jealous, because that's somewhat high. Cable modem is typically these days, another thought, or another proposal?

STUDENT: Inaudible

So, a cable modem tends to be, I would say, between 1.5 megabits/second to... Comcast, I think offers up to 8 megabits/second. A DSL modem meanwhile offers a speed of what?

STUDENT: Inaudible

20 mega... wow, you guys are getting much better service than I am. So, usually not... I'll say not with DSL, at least. I'd say typically, you would see slower speeds with DSL. You are seeing a convergence these days with Verizon, for instance in Boston, and Comcast become more competitive. I'll say that generally and again these numbers are just typical, they are not necessarily upper and lower boundaries. You might get something like 768 kilobits/second to a few megabits/second. These days Verizon in Boston, if you pay for DSL modem service, which for our purposes, is different only in so far as DSL modems come in over your phone lines, cable modems use your coaxial cable lines, beyond that, that is the physical difference. The real world difference is that these days, at least in the Boston area, cable modems tend to be faster. In fact, you can sort of spoil the answer, what's the upside of DSL in the Boston area and elsewhere, certainly?

STUDENT: Inaudible

Comcast is better in that it tends to be faster, it tends to be more expensive. You see that same tradeoff when we were talking about RAM and hard disk space and so forth. For instance, Comcast, frankly, is a little insane. I think you pay between \$40-\$60/month for Comcast service, depending if you have cable bundled with it and phone and all that. It's pretty high. Whereas, DSL, even though you can get introductory rates, for like \$14.95 or \$19.95, I think the current rate is \$25ish or \$35 max. So, I'll say something like \$15-\$35. Take this with a grain of salt. It does tend to be less. For the uninformed, this isn't such a problem. It is very hard, if not impossible, on say, Verizon's website to find out what the speed is of their cable modem sever, of their DSL modem server is. What they will typically do these days, which is a bit misleading, is say that it is 40 times faster than dial up. Well, who cares? How much slower is it than cable modem service?

Notice too that these numbers are somewhat misleading. Because, in fact, even though I, at home, get 8 megabits/second download usually, or a lot of the time, your upload speed is much slower. It is an asymmetric medium, whereby you will get up to 8 megabits/second. What is my maximum upload speed? Only I know this, but take a guess.

STUDENT: Inaudible

Much slower. Good guess. I'll take that as an E-1 answer. Let's push a little harder.

STUDENT: Inaudible

1-3. yes, actually it's sadder than that... 768 kilobits/second. This, mind you, is because I pay \$10 more a month to double my upload speed. It's a finny thing. So, with Comcast, again, you will pay \$40-\$60/month. The typical speed they quote you, at least in the metro Boston area, is up to 6 megabits/second, and up to 364 kilobits/second upload. That's ok though. Why does it tend to be ok for the typical user to have this asymmetry whereby uploads are much slower than downloads?

45 minutes 16 seconds

STUDENT: Inaudible

Yes, you tend, as a typical user, to download far more than you actually upload. What are you doing when you are on the internet? You are pulling down web pages. You are sending very little requests. You are saying give me today's news. And then all of CNN's news gets sent back to your computer. If you are downloading MP3s. you are sending a very small request. Give me the latest MP3, but then you are downloading 3 megabytes of that file. If you are downloading movies, you are downloading hundreds of megabytes of files. How often do you, say the typical E-1 student, upload a movie that's 200 megabytes somewhere else? So, this is where you start to see the distinction between us, won her and people like you. What we do every day, or often, is in the other direction, sending data up for research purposes, for course purposes. Every time, frankly, we upload the latest videos to the podcast. It's the biggest headache, because it zaps my internet connection for an hour

or two, usually every Wednesday night. Fortunately lost is on Wednesday nights, so it's not such a problem.

But the point is you have this asymmetry and you have to read the fine print sometimes, to know the difference. You also have to understand what your needs are so as to know does this matter to me. DSL will also have this asymmetry. It will sometimes be quoted, but again, you have to be aware. Verizon, particularly, has dumbed things down enough in a marketing sense, where again, they say 40 times faster than dial-up service. But even if you get a service person, on the phone, they will tell you that it takes no more than 53 seconds to download a movie. Well, what does that mean? How big is the movie? And what are we comparing it to? So, you have to bear that in mind.

Typically, if you want really fast speeds, cable modems tend to be better, but again, here there is a caveat. I pay, again, \$10 more each month to double my upload speed from 364 which is normal for Comcast, up to 768. For me that is huge. It literally halves the amount of time it takes to upload one of E-1's podcast videos. Or if you recall from our first lecture "slingbox", when I think we were showing press your luck on the overhead screen, or some TV show. Slingbox being the device that streams my cable TV connection this way. I'll admit, I like to tell myself that I pay \$10 more a month so that I can work more productively, but frankly, it also has the fringe benefit of doubling the quality of watching TV remotely. But, you pay \$10 more a month and you go from 6 megabits/second download to 8 megabits. Now, that sounds like a lot more. Right? That's 33% more for only \$10. But you have to ask yourselves what question at that point?

STUDENT: Inaudible

Do you need it? So, yes, it's faster. But your internet connection is not the only bottleneck in the world. The person you are downloading that file from might not have as good an internet connection, especially if it is one of these peer-to-peer file sharing programs. If you are downloading even from a big server like iTunes, or even E-1's podcast, the server is configured to send you the file somewhat slowly. The presumption being, we want to be able to satisfy lots of downloads at once. You don't really need that video right now; you are probably ok with waiting 20 minutes instead of 15 minutes to download that file. So, the server might be artificially delaying on the other end as well.

The problem is I don't need more download speed, 6 megabits is pretty darn fast, at least as far as typical computing needs go today. It's this where they really get you, and it's really hard to go beyond this speed unless you want to pay much more for it. However, those of you who are commuting in from Boston, or who are watching this from outside of Boston, from folks like Verizon have the advantage of an up and coming technology known as fi-os. If you have this then I am genuinely jealous. This is a fiber optic based technology for delivery of internet connectivity to the home and it does tend to be much faster than cable modems. I don't remember values off hand, maybe the TFs know, I want to say something like 10 or 20 megabits/second. It might be a little pricier. Some of you in the outer areas of Boston can get this already. So I would go to Verizon.com, and search around for fi-os, fiber optic and so forth, if only out of curiosity for yourselves; to see if you can get it.

So, in answer to your question, is it worth wiring a condo (one's own condo) for fiber, honestly, probably not because even though you already have the ability to wire your home pretty cheaply, pennies per foot with cat 5 or cat 6, which is slightly better insulated essentially, look at the incoming internet speed. So, you would have a really zippy home network, but where's the bottleneck? It's in the jack in the wall, where your home interfaces with the outside world. So the only advantage to wiring one's home or apartment with something better than say cat 5, 100 megabit/second, would be for what kind of situations? What kinds of people would want to wire their home with fiber optic cable?

STUDENT: Inaudible

So, if you upload a lot it doesn't matter, actually, because this is symmetric. You do get 100 megabits down and up. You only hit that asymmetry once the data starts to leave your home, or your office. Why would you want a really fast connection locally, even though it won't speed up your connection to the outside world? (Because again, the bottleneck is at the jack in the wall.)

STUDENT: Inaudible

Yeah, to share data among your own computers. So, companies and Harvard University's campuses will wire their local networks with faster technology so that you can have peer-to-peer file transfers happen more quickly. Students on campus, for instance, could stream this course to videos at a much higher resolution than outsiders could using the general internet. And if you run your own file server, or your own back up server, even though this might seem a little geeky these days to imagine that I might indeed be one of the people in his apartment with a fileserver in his apartment, the more the world talks about having home media centers where you store all your movies and music in a central computer, essentially in your entertainment center. Sharing data within your home has become a more plausible and compelling feature. So, you will see that in the future. Again, I think you would be overpaying at this point, but again, this is conjecture. You can indeed, save money by buying (long story short) special cables that have pretty much everything inside of them, like planning for a rainy day. Coaxial cable, cat 5 cable, fiber optic cable, it just depends on the cost, if you think it's worth it to you. But as to when that day arrives where it is useful and fast, and necessary it is tough to say.

That's a long answer to one question, but we were going there any way.

Other questions?

Yes.

STUDENT: Inaudible

Yes, so municipal wireless... how many minutes do we have on tape?

STUDENT: Inaudible

So let's go here now before our break. Let's relate this same conversation to wire less, because it is very much in the same vein. So, municipal wireless is being touted in a number of major and minor cities. Let me come back to that in just a moment. Let me say first, for your local network, this is the kind of technology you are using for a local network if you have a yellow wire that you are running around your apartment; or if you are in your office and you are plugging into the data jack, that's probably the kind of wires you have there. But if in your home you have wireless access, or in your office, or in Harvard, you have wireless access, notice (and it might be obscured by this big light) but there is this little box here with 2 little antennas that's plugged in with one of these yellow cables to the wall. But that is allowing all of the FTs, for instance, and some of you to use Harvard's network wirelessly. Well, what is that technology called? This is not an internet providing technology as much as it is a local area technology. You have technologies like 80211 B and 80211 G. If you have a laptop or even a desktop that has wireless capabilities odds are you have a piece of hardware in there that supports one or both of these technologies. The speed of your internet connection, locally, if you have 80211 B... does anyone know?

STUDENT: Inaudible

It is indeed 11 megabits/second. However, that is signal strength dependant. So ideally, you have 11 megabits/second in a through put within your own apartment. Is that ok? Well, odds are that's plenty if you are trying to use the internet in your own home, why? 11 seems much smaller than 100, but why is that probably ok, and not cause for worry, or be a cause that you skimped?

STUDENT: Inaudible

Look at the incoming connection speed. If this thing is only coming into your home at 8 megabits/second, all you need is 8 megabits/second within the home. And in fact, it's maximally 11, it does depend somewhat on signal strength, so if you have lots of cement or very thick walls, or brick walls, or you are on a different floor of your home, you are not going to get as strong a signal. Those of you with Windows laptops will see (in the bottom right corner of your screen, if you are using a wireless connection) excellent strength, very good strength. That's sort of a nice way of saying you are getting 11 megabits/second, you're not getting 11 megabits/second, but not all operating systems are precise as to what your actual throughput is, so fortunately, you can do a little bit better. This is more common these days. G gives you what throughput?

STUDENT: Inaudible

54 megabits/second. This too is also signal strength dependant. Odds are that even if you are only getting a half-quality signal, that's still 27 megabits/second (speaking off the cuff) and that's more likely to guarantee you this throughput here. And in fact, just to put it into perspective, when I finally did set up this fileserver in my apartment so that I would have a central place for my data and backups, I actually bought a NIC (Network Interface Card) off E-bay that would fit in this specific Sony laptop, but whose technology was G instead of B. Because of sharing files in my apartment, when I load big files from my file server onto my laptop, this is 5 times faster, in theory. That means a big speed up for me locally. So again, you have to consider what your needs are. What are your metrics for actually giving a good connection verses a bad connection.

So, finally, in answer to your question, these wi-fi, (oh, and this stuff... the marketing speak, wi-fi is wireless fidelity) you have something called wi-max, or other variants of this for municipalities. I just read recently that one subset of Boston would be rolling this out sometime in the next year or two as an experiment. These speeds... do any of the TFs know what they are quoted at? I'm not sure. Let me check during break. I will come back to you. So, I will hold you in suspense. It is an interesting question. For one, just to put this into context, you now have free internet service potentially competing with the likes of Verizon, and Comcast, and frankly, I don't mind so much that you will see more of that kind of competition. It is an interesting question. I would actually wager that, coming from the government, huge expenditures of money and wireless infrastructure, I don't think these things will last, at least in their proposed form for very long. We have all seen, in this course, right... what we did last week is out of date already. So, you can only imagine spending millions or hundreds of thousands of dollars on wireless infrastructure... it's not going to endure for very long. I question the interest of public service in keeping up the expenses of these kinds of services. I think it's compelling to have wireless service. It's not clear to me if it is going to be sustainable coming to the world for free, at least from local government. But it is a nice experiment, if nothing else. It's going to push the envelope in terms of wireless speed. And frankly, \$40-\$60/month... that's one of the biggest utilities per month, and I use it every day, but there's only so much I want to pay to watch TV outside my apartment.

Let's take a 5 minute break.

58 minutes 1 second

58 minutes 10 seconds

All right. Welcome back, so what have we gotten done so far? So we talked about how a LAN works, really, underneath the hood. We talked about how you can connect your own LAN or your own local computer to the rest of the world by way of some internet connection via any of these technologies, Cable modem or DSL.

Rei was kind enough to look up the current speeds for Fi-os and maximally you can get 30 megabits/second download and up to 5 megabits/second upload. So... not too shabby at all. You can check out Verizon.com to see if it is available to you. Odds are it's not because it has not been rolled out terribly so, but it is being rolled out more in the suburbs than in the local city.

So, once you've got your connection to the internet, what does it look like? Well, again, this is where we began our discussion. So, you have these internet backbones. They are essentially these really fast computers and really fast wires that interconnect major points on the internet. You might hear those points called peering points. A peering point generally (or a POP which is a point of presents) tends to be areas of the internet where major backbones connect, where Sprint connects to MCI, or to AT&T, and so forth. So, these very popular nodes, one in Chicago, Los Angeles, Dallas, and so forth, are major peering points, or major points of presents; the internet connections, big ones flow in to and out of.

Some of the interesting things that the macroscopic scale big companies like Sprint and MCI need to think about is the cost of running their own network. And often you will have arrangements

between these big providers whereby they pay each other to take traffic from one network onto their own, or to put traffic of their own onto another. One of the curious things about the internet is we already know that data doesn't necessarily have to travel from A to B, via one fixed route. That was the fundamental design early on in a network like the internet, but it turns out that the fastest way from point A to point B is not even the proverbial straight line, but rather, you will have data hopping around cities, around within cities for either technological reasons because the data is simply flowing along the fastest possible link, not necessarily the shortest possible link. Because of the various peering arrangements and the various ISP arrangements it's their choice. They want to route your data this way then back that way, simply because it might save the backbone money, rather than having to put your data on some other network. For you, the user, none of this is even that relevant, because, at the end of the day, you only care about where your data, or rather that your data arrives, either at your computer or at it's destination.

But on a higher scale, there are some really interesting network-related topics that you can actually explore in other courses here and elsewhere, if you take a fancy to this kind of topological discussion. But if you have such complexity manifest in diagrams like this, how in the world does data get from my computer even get to CNN's server somewhere in the world, or Google's server? Well, it turns out that every server on the internet has an address; just like you assuming that you receive and send postal mail have a US postal service address in this country that uniquely identifies you or your mail box, or your PO box, or your building in the United States; so do computers have a address that uniquely identifies them on the internet. This address, this IP address, internet protocol address is a number that is 32 bits long, which is to say, every computer on the internet has an address that is 32 bits long. Thinking back to lecture one, how many computers does this naively suggest can be on the internet at one possible time? Each computer must have an IIP address, each IP address is 32 bits long, how many possible IP addresses are there?

STUDENT: Inaudible

I'll give the E-1 answer to this... A lot, but there was one number I begged you to remember in lecture 1. 256 would be unfortunate if that many IP addresses in the world and on the internet at the same time. It is much more than that

STUDENT: Inaudible

Bigger, yes.

STUDENT: Inaudible

So, we are talking if you have e32 bits, again for those of you that love math, that's 2 to the 32nd power. I didn't expect you to remember the number, but I did say, please remember that this is roughly... give me an order of magnitude...

STUDENT: Inaudible

People who are watching from home are kicking the screen right now, that no one is answering the question. It is a lot. It is about 4 billion. So that means that you can have about 4 billion computers

on the internet using this IP addressing schema. 256, incidentally, if IP addresses were only 8 bits long, you would have 2 to the 8th power, or 256 possible IP addresses. Fortunately, we have way more. It turns out that even though it might seem like plenty, the way IP addresses are allocated through out the world, actually, some of the times, it is increasingly becoming tight... the address space, and some of the IP addresses go wasted because one person owns them, but not necessarily uses them all, and so forth. But there are ways of sharing IP addresses, it turns out. If you have ever wondered, if only tonight, how your home network uses one internet connection to share internet services to multiple computers, in effect what you are getting when you using up with cable modem providers or DSL modem providers, is your are getting a physical connection to the internet. You are more technically getting one IP address. You are being assigned a number that is of this form that uniquely identifies you (your computer) on the internet. By this form, I mean it's a number dot number dot number dot number. And each of those numbers themselves can be from zero to 255.

So, how do you see this? Let's make this a bit more real. Those of you with windows, and again I'm windows biased in lectures only because this is what I bring to lectures, if in windows you go to your network connections and choose your appropriate network connection, and click properties. This is the kind of detail you, ideally, never have to deal with these days. But if you scroll down you will notice something called TCP/IP. It did flash on the screen a while ago. You actually heard that spoken by a big booming voice earlier tonight. (He's TCP/IP.)

That was the teaser for tonight. Well, TCP/IP (I thank you for not laughing at me when I mimicked the video here tonight.) TCP/IP is the language that is spoken on the internet. Any computer that is connected to the internet speaks the language or more technically, the protocol, known as TCP/IP. So, anything you do on the internet, for the most part, is carried across the backbone we saw depicted earlier using a protocol called TCP/IP. How does this relate to your home connection? Well, notice when I double click on TCP/IP in that window, I get this box.

And for 9.5 out of you, you never need to look at this, or change this. A few years ago, this kind of tinkering was common. Notice that the top bubble there says obtain IP address automatically. The bottom one says obtain DNS server address automatically. What that essentially means is that when I turn on my computer, and am connected to my cable modem or my DSL modem (via wireless or wire) I am going to automatically ask my ISP for that IP address.

It's not hard coded into your computer. Years ago it was not uncommon for a Comcast rep or the like to show up at your home with a printed piece of paper and say type in this number as your IP address. It came printed on a piece of paper. It doesn't work very well when you want to be able to change things and reconfigure things on Comcast then. So, things are almost always dynamic these days, at least for public networks and for utilities like Comcast and DSL providers and so forth. What you get, ultimately, is an IP address.

Well, what does this mean? Well, if you request a webpage and you receive a response, well, what is in that response? Well, in the example we often use, it is today's news. Right? We said, a week or so ago that web pages are written in a language that is called HTML. Well, what does it mean then if web pages are written in HTML, and what language do web browsers and web servers speak?

STUDENT: Inaudible

What do you type when you type any URL?

HTTP, hyper text transport protocol is the protocol that web browsers and web servers use to transmit data. I seem to be changing our story every week. Web pages are written in HTML. Web pages are transmitted via HTTP. Anything on the internet is spoken in TCP/IP. It turns out (and this is a very clever design fundamental) that most everything in the internet, so far as the implementation goes, is layered. So we began talking tonight about Ethernet. That's a physical layer. Once you have a physical connection, you can do things with that connection. And with that physical connection, you can put on top of that this language (this protocol) called TCP/IP. That gets the data from A to B. What data do you want to send?

Well, if it's a web page, or a request thereof, you want to send an HTTP piece of data. So, you have Ethernet, you have TCP/IP, and then you have HTTP on top of that. What is in an HTTP response? Well, if it's for a web page, you get back HTML. So, it's not that we keep changing our story; it's that we are looking at different layers of this puzzle that is the internet. And again, at the ground floor, you have the physical connection, something like Ethernet. On top of that you have TCP/IP, which for right now, assume that that just gets data from A to B. Well, what's that data?? If it's a web page, it's a chunk of bits that are stored in HTTP's format. Well, if that HTTP's response contains a web page, inside of that is HTML.

Let's try to make this more real. Suppose I have made a request to (I need a volunteer from the back of the room. If someone could just call out their name from the back of the room. You don't have to come down to the front of the room.)

What's your name?

STUDENT: Inaudible

Abdul, so I wish to send data to Abdul. I'm going to write this data (let's say in English of some sort) let's say it's an e-mail. Here's my e-mail. What I am going to do is hit this send button on my computer. What my computer essentially does when I hit send is figure out that it is e-mail. (That's an internet service.) I'm going to send that out over the internet. Well, what language do I use to get anything out on the internet? I use TCP/IP. One of the features of TCP/IP, is in the interest of efficiency and guaranteeing that all your data gets from A to B, is it breaks it up into smaller chunks. So I [inaudible] tonight's paper but imagine your computer essentially, without you even caring about these details, it breaks your message into 4 smaller messages. Each of these messages goes inside a TCP/IP packet. The packet will be sealed, and it will be numbered. This is packet 1 of 4. Well, just like the US postal service, I'm going to send data to Abdul, I have to address this packet, this virtual envelope, so I'm going to write Abdul on the "to" field. And just like you would in the real world, I'm going to say David in the "return" field; thereby addressing the envelope to and from. It is sort of a simple example to belabor, but what the computer is doing is precisely that. My computer supposes that Abdul actually has a computer that I'm sending this message via. Well, what's going to go here instead of "Abdul" based on what we just learned?

STUDENT: Inaudible

The IP address of Abdul's computer or really, Abdul's mail server. It may be g-mail or hotmail, or whatever my computer figures out based on the domain name in Abdul's e-mail address whose IP address needs to go there (g-mail, hotmail, or MSN's or someone else's.) Meanwhile, we have 3 other parts to this message, so I'm going to go ahead and slip part 2 in the envelope. The computer is going to go ahead and seal it. It is going to become part 2 of 4. Abdul's name goes in the to field: my name goes in the from field. That's ready to go. I'm going to take part 3 and it's going to go into the virtual envelope. Fortunately computers do this much faster. This is going to be TCPIP packet 3 or 4 from David to Abdul. And finally, packet 4 goes in its own virtual envelope. It gets sealed up and it is 4 of 4. Maybe this is an obvious question, why am I bothering to include the 1 of 4, 2 of 4, 3 of 4, 4 of 4? Put on one's engineering hat so to speak.

STUDENT: Inaudible

Right. You kind of have to know what order to reassemble these things into. To Abdul's computer this is going to look like bits. Well, together, they are all going to paint some picture of an e-mail. But if you don't arrange them in the right way, it could end up being gibberish. So, the ordering is important. So, that data is also included in these packets. And now my computer (and sometimes it feels like it takes this long) is ready to send these packets, this e-mail, off to Abdul. My computer (Windows, Mac OS) speaks TCPIP so the computer out of the box just knows how to do these things these days.

Well, where does this go? My laptop is connected wirelessly to Harvard's network. So, somehow or other via 80211 G these packets are going to be sent to these antennas up here. They are then going to travel through the yellow wire in the walls in this building. That wire is connected to what piece of hardware somewhere in the building?

STUDENT: Inaudible

A server, but we replaced server in our discussion earlier with a switch, which is just a big wiring closet somewhere in this hallway, or in the basement. That switch, in turn does in fact connect to a device called a router. A route is a more intelligent computer that doesn't just spit data out ports based on what's physically connected to it, but a router, rather, looks at the "to" field, and says hum... Abdul. Well, a router is going to have multiple connections coming into it and leaving it. Just for simplicity think of a router as having connections going north, south, east, and west. A router is smart. A router understands TCPIP. A switch only understands Ethernet. A router understands TCPIP, and says Abdul. Inside every router is a table of sorts, like an excel spread sheet, or an access database. Essentially rows and columns where the router can look up and say where is Abdul? There's going to be a row, in a sense, in the router's table saying that Abdul is to the north. So, what the router is going to effectively do is send these 4 packets, as they arrive, out its northbound port.

Well, where are those packets going to end up? Odds are that Abdul is pretty far away from me right now. There's no direct connection between me and Abdul. In fact, for the sake of discussion, think of yourselves as routers. I can reach Rei, but I can't reach Abdul. So, it turns out that on the internet, just as our diagram a moment ago suggested, these circles that have all these points coming in and out of them... guess what those are physically?

STUDENT: Inaudible

Routers. So, fortunately, Rei is connected to another router via some link (maybe to the north, south, east or west), but each of these routers has a table that either automatically, that is dynamically, or because humans manually inputted this kind of information, each router knows the next best hop to send the packet. It may not be physically the closest, but it is the preferred next hop.

Well, how does a router know? Well, obviously packets aren't literally stamped with Abdul's, that is the beauty of using numbers as schema it is because routers can look at the first few numbers in an IP address, and say all IP addresses starting with 140 go that way. And all IP addresses that start with 141 go that way. And that's why numbers are useful. So, when I then hit send, my computer figures all that out. It sends it to the access point; goes to the switch; the switch leads it to the router; maybe that router is connected to another one via fiber optic cable or something similar. Eventually that router is connected via some other medium, maybe a transatlantic cable, maybe microwaves, maybe some physical land line connection, but in short, through one or more other routers is there a connection to Abdul from me.

Just to hammer home this point, why don't we go ahead and I'll send them (in order) to Rei--- to Rei--- and to illustrate our point, don't necessarily send each packet to the same way; because, again, at any point in time, the router is getting busy. A lot of internet traffic, at some point in the country and so a router decides I'm going to route around this congestion to another router. Hence all those multiple paths between A and B. Hopefully, those have soon arrived at Abdul, and there is a real demo here. If you wouldn't mind opening up those envelopes you have received and reassemble them and reading them aloud, if you could, pretty loud, what the message is that I sent you.

STUDENT: Inaudible

I guess it doesn't work to whisper across the crowd. So you have all the envelopes that you have received. That goes without saying. They were numbered, so read them 1, 2, 3, 4.

silence

Note to self... no self-adhering envelopes next year. What do you have for us?

STUDENT: Inaudible

The winner of the iPod shuffle is... it turns out that just as in the internet, TCPIP packets can be dropped, literally, and I don't know if a few of you saw me try to be clever, but I physically dropped this one on the floor, and that can in fact happen in reality. WE aren't talking about dropping something physical, but rather bits or packets, but TCPIP is designed, or rather routers are designed... If they get too congested, let the sender retransmit. And again, it's an exaggeration, but a router will sometimes drop your data, but that's why they are numbered. Built into TCPIP in any computer that supports it, is a realization that... if Abdul's computer, thinking to itself, well, I got 1 of 4, 2 of 4, 3 of 4, where's 4 of 4? He waits a second or two seconds and nothing had arrived, what

Abdul's computer is designed to do, because it supports TCPIP, as does Windows and Mac OS, and Linux, he will request that I retransmit at least that packet that he hasn't received. And so what will next happen is my computer will say, wait a minute, Abdul has informed me he didn't get this packet. I'll retransmit that same packet and hopefully he will, now get it this time. And just to I'll save you the routing of this. Just suppose I did send this final packet, recall the sentence that was read in great dramatic fair her was, "the winner of the iPod shuffle, for having submitted and scored higher than 75% on problem set 1 is a Mr. Timothy Conant." Timothy, are you here? Deafening silence. What a wonderful dramatic surprise this is for Timothy who will hear about this in about 48 hours, apparently. So, as much as I would like to just give it to Abdul, I'm afraid that we are bound by the rules of this contest to give it to Timothy who is also selected pseudo-randomly. If we could, it will be preserved on camera, a round of applause for Timothy and for Abdul.

Applause

Thank you very much for partaking. Atop the slide here, what you see is a more technical layout of what we have been talking about. It's not so important for tonight's purposes, as to what the bits look like that are being sent across the wire via these routers. But just to put things into perspective, even though we keep calling it TCPIP, TCPIP is actually 2 different protocols. They are just so commonly used together that people just usually refer to them as a pair. But this, it turns out is what TCP datagram or TCP segment, looks like. Well, what does that mean? Well, any piece of information that is sent by TCP is laid out in this fashion. You don't really have to absorb all of this. All this means is that you have the first bit, second bit, third bit, fourth bit, fifth bit essentially laid out from left to right, top to bottom.

Embedded in this TCP packet then is some interesting stuff. It should make intuitive sense. Notice this, sequence number, and acknowledgment number. Well, what was I doing in the top right of the envelope? I was saying 1 of 4, 2 of 4. That was the sequence number. So, embedded in each of those envelopes, or TCP packets, is precisely that number, using in this case how many bits? How many bits are used to represent a sequence number?

STUDENT: Inaudible

If this is zero... we have 10 bits, so each of these dashes represents a bit. So, how many bits? 31, or counting from zero, 32 bits are used for the sequence number. Now this might seem silly, but how often am I going to send Abdul an e-mail that requires 4 billion packets to be sent? Well, your computer is doing a lot of other things at once, so these sequence numbers aren't just for one e-mail, but for all the internet traffic that your computer is sending. So, it is useful to have so many options.

Source port and destination port... these I think came up briefly last week, though I tend to confuse semesters, so help me out. What port does HTTP use?

STUDENT: Inaudible

Am I remembering? Am I in 2006?

STUDENT: Inaudible

80, does 80 sound familiar? Yes, no? so let's talk about ports. I thought we did this last week, but you know that a computer can do multiple things at once. It can provide multiple internet servers. One computer can be an e-mail server, it can be a web server, it can be a SSH server (which is a technology you will use in section, if you haven't already), in short a computer can do multiple things. But if a computer has only one IP address, well, that sort of begs the question... how does your computer distinguish whether some bits are arriving or that reply from Abdul, or that web page from CNN... well, associated with every internet service... and by that I mean HTTP, that is web, e-mail, Google talk, AOL instant messenger, all the popular programs you can think of... Skype last week, is a number. It is a relatively small number like 22 or 80 or 1,000, but relatively small that uniquely identifies the type of service that that packet includes. So, if I wanted to be really particular, on each of those envelopes I sent to Abdul, I would have written the number 25. That is the port the number that uniquely identifies e-mail on the internet (AKA SMTP, the sending of E-mail.)

Well, inside this TCP packet, are those numbers, so when Abdul receives that envelope his computer can say, not only do I know this is for me, and from David, it's an e-mail which means I should route it to outlook, or to Eudora, instead to Internet Explorer, or Firefox, or a browser. So, those are ports in a nutshell.

Meanwhile, it is in IP (do I have this picture here?) it's in IP, just to give you a glimpse of this, that we actually store the source address, and destination address. But again, we would defer to a networking course to tease apart more of these features. But we can see them in action.

What I have here is a window. It's a little small, but hopefully you can see it when it fills the screen. What I'm going to run in a second is a program called trace route. So, I am connected to one of Harvard's computers in the electrical engineering computer science building. So it's a few buildings away. That's fine. Think of it as being on my laptop.

What I am going to do is ask the computer to show me the route between my computer and CNN.com. So, I'm going to run trace route (as the program is called.) www.cnn.com and here are the results.

There's a lot going on, some of it pretty fast. We are up to 16, now notice some of these won't work for us for various technical reasons, but notice that the very first line just reminds us the trace route to cnn.com... what's this in parentheses?

STUDENT: Inaudible

The address of cnn.com's computer, good... notice that this first line here starts with 14024760.1 This is a useful hint that is not always true, but is often true in the home (with your home routers) a routers IP address is usually ends in dot one by convention, but it is not a requirement. And one of these conventions useful [inaudible] recognizes things. So, what does 1 represent, so you think? What does line 1 represent?

STUDENT: Inaudible

Land line is the right idea, but not quite right though. Keep looking ahead, so we are talking in terms of hops again. What does each of these rows represent? Take a guess.

STUDENT: Inaudible

Server of some sort, but why type of server?

STUDENT: Inaudible

A backbone of sorts, but what do each of these hops (again I'm trying to use the same nomenclature) hops represent?

STUDENT: Inaudible

A switch is just a local thing, for a local network, a node yes... looking for a word that starts with R... so routers... so this has been the buzz word for most of tonight, A router is just a device that gets data in and figures out that this IP address goes that way... let me send it along to the next router... well, who's the next router? Well, from my computer, the first router is that guy depicted in line one. Who is that router, in turn connected to via some physical connection or wireless connection?

Well, the router in line 2... Well, what's its name? The host's name of router # 2, to borrow last week's lingo, is this rather cryptic core-1-gw-v1415, but thank God, something familiar, .fas.harvard.edu. So, this is one of the faculties of Art and Science's routers. "GW", is a common nickname for gateway, AKA router. A gateway is a router and vice versa.

What's in line 3? And I'm going to stop it before it scrolls too far for us. What's line 3? It looks like yet another router. Notice one buzzword in there. It is useful to at least see this stuff in practice, WAN, so some kind of wide area network, but these host names are completely the prerogative of Harvard sys admin. They mean nothing in and of themselves. They have just been chosen by some tech guys.

4 doesn't look like it has a host name, or domain name per say, and that's ok. Not every computer on the internet needs a domain name that is human memorable, because who cares what the name of a router is, frankly. You just care about your own computer or the destination.

But moving on, take a look and make an inference here. Where is my data going from the moment it leaves my computer, on its way to CNN? What happens in step 5?

Where are we?

STUDENT: Inaudible

So, we are in Boston, on Quest's backbone. Quest is a big ISP of sorts, it is a big backbone provider, locally. So it looks like our data hops around Quest's network for quite a bit. But it turns out that between lines 6 and 7, where does our data go geographically?

STUDENT: Inaudible

Good.

STUDENT: Inaudible

So, good, in spirit, what's often happening is that your data is leaving your computer going to faster and bigger computers and connections, and eventually on the tail end, the process will eventually tend to reverse. At least if you are visiting a somewhat small site. But again, the question on the table here is what looks like? Tell me something interesting about where my data is going between hop 6 and 7. Use your airport codes.

STUDENT: Inaudible

I know this one. I grew up in the tri-state area. Hint-hint... there's a 33% chance of getting this one right.

STUDENT: Inaudible

Newark... EWR is the airport code for Newark. This, again, is convention. Not all routers are named with airport codes, but sysadmins tend to use it because it's useful convention to obey. Literally, between points 6 and 7, assuming they are not fudging the host names, there is some kind of physical connection between Boston and a router in Newark, New Jersey. How much time does it take for my data to go from my computer o Newark? Well, that's what all the other numbers on these rows are indicating.

1 hour 29 minutes 46 seconds

In parentheses is each router's IP address, but to the right of each routers IP address, say right here, 6.983ms, and 6.700ms... what is that do you think?

STUDENT: Inaudible

Milliseconds... and that denotes the number of milliseconds it took for my packets to literally go from my laptop to Newark, New Jersey. It takes 5-7 hours to drive there; it takes 6 point some odd milliseconds for those virtual envelopes I sent to Abdul to get to Abdul in Jersey if he were there.

So, when you start talking about speed of light and electricity, and fiber optic cable, traveling up to the speed of light maximally, 6.7 milliseconds to get form here to Jersey, that's pretty darn fast. It certainly blows away any sort of physical transport mechanism we have yet invented. Yes?

STUDENT: Inaudible

Good question. Why are there 3 of these numbers? Well, this program, in particular, just to give you a sense of averages tries to send the data 3 times. So you can look at it and all right, 6, 6, 6 that's a pretty good estimate, because sometimes, again, you will get congestion. And if you were just to see something like 15 milliseconds, you might infer that there is a bottleneck there, but it may have been very temporal. It might disappear the next time. And what this program is doing, it's not sending e-mail. It's not sending web pages. It's just sending empty envelopes, but in such a way that it is sending the amount of time it took to get there. Another question?

STUDENT: Inaudible

So, why does it go faster from Boston to Newark? So, what you are actually seeing, these times from my computer to that router. These aren't the times between routers. So, what you hope to see, and what we are generally seeing is that as you look from top to bottom, these numbers are in fact getting bigger and bigger and that is consistent with the distance getting slightly farther away. So, you read them from start all the way to the router. Let's take a look at... I don't know what DCX is, it might be some where in D.C. I'm not familiar with it if it is an airport code. The others are a bit more cryptic to me, and sometimes you get these stars, that means that the routers are not set up to respond to our kinds of requests.

But let's try another one. Let's say, I think we said `cnn.co.uk` is the UK version of CNN's website. Let's give this a try. Ok... here we go. Notice there are similarities in the path, which you might expect, because the data is routed through some common points. Notice, you can see it here, notice for some reason, notice line 11, the numbers are all over the place, but it looks like it was a temporary congestion, or something. Notice it took us over... 185 milliseconds to get to hop 11, but clearly we can get past that more recently. It looks like the routers along this path just aren't being cooperative.

So, let's try something in the other direction. So, `.co.jp` for a country code of Japan. Let's see what we get this time. We got the first few. So, now we can start using our airport codes again. So we are going from Boston to where? New York, JFK city code. SVL and PAX, and guesses

STUDENT: Inaudible

It is in this country. You can infer it from the time. Let me ask you this. Tell me between which routers is there a really expensive, really thick, really long transpacific cable, or transatlantic?

STUDENT: Inaudible

Take a look at that. You can read something interesting into these numbers. To hop 11 it takes us 75 milliseconds, to hop 12 from here takes 200 milliseconds. That suggests a really big something. In the middle, odds are that, in fact, is the ocean. If I was better at inferring where, by the host names, where they are, I could tell you if it is going to the east or the west. Since it is going to... PAX, we can look this up later... but there, that one made it all of the way. So, to go from here to Japan, which in a plane would take many hours, with Google earth even, it would take a few seconds, with

the internet itself, 200 milliseconds. You can get there and back in less than half a second, literally. It's kind of a mind-blowing thing.

So, how does all of this fit together? Well, let's take this as promised...

video

And on that note, we will leave the ending as a surprise for next time. We will see you next week, when we will dive in to new material all together.