

Logistics

No more extensions on labs

3pm was the deadline. No exceptions.

Final exam

Thursday December 6,
3:00-4:00pm
Same drill.

Final Exam Review

Doesn't count as credit for extratation

Same drill

Gates 5222, Sat/Sun 1-2pm



What happens when you do a google search?

A very casual in-depth introduction on this otherwise
harmless interview question

Context

- We ran out of ideas for presentations
- I actually got this problem 2 years ago
- Kinda gave a 2 minute answer
- Realized that there so much arcane stuff that isn't common knowledge that can be revealed by answering this question
- A pretty good “last lecture” to tie everything together
- <https://github.com/alex/what-happens-when#the-g-key-is-pressed>

Typical interview answer:

“You send an HTTP GET request to google.com. It does some magic backend stuff. You get the result back. Your browser displays the results.”

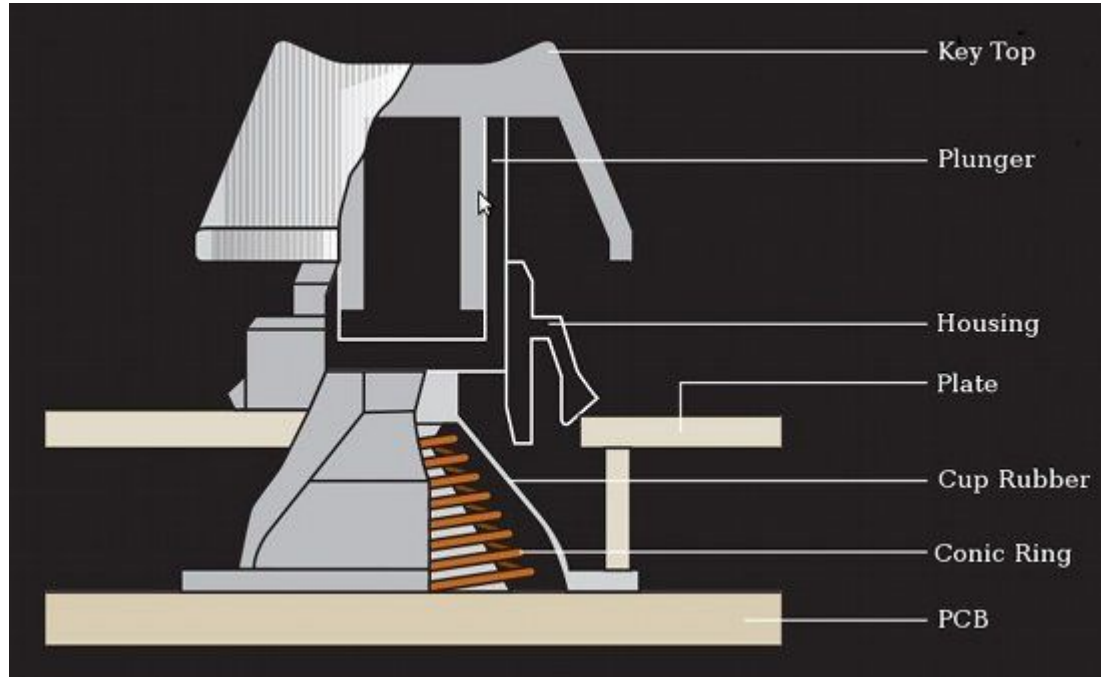


Our answer:

A bit more complicated than that!

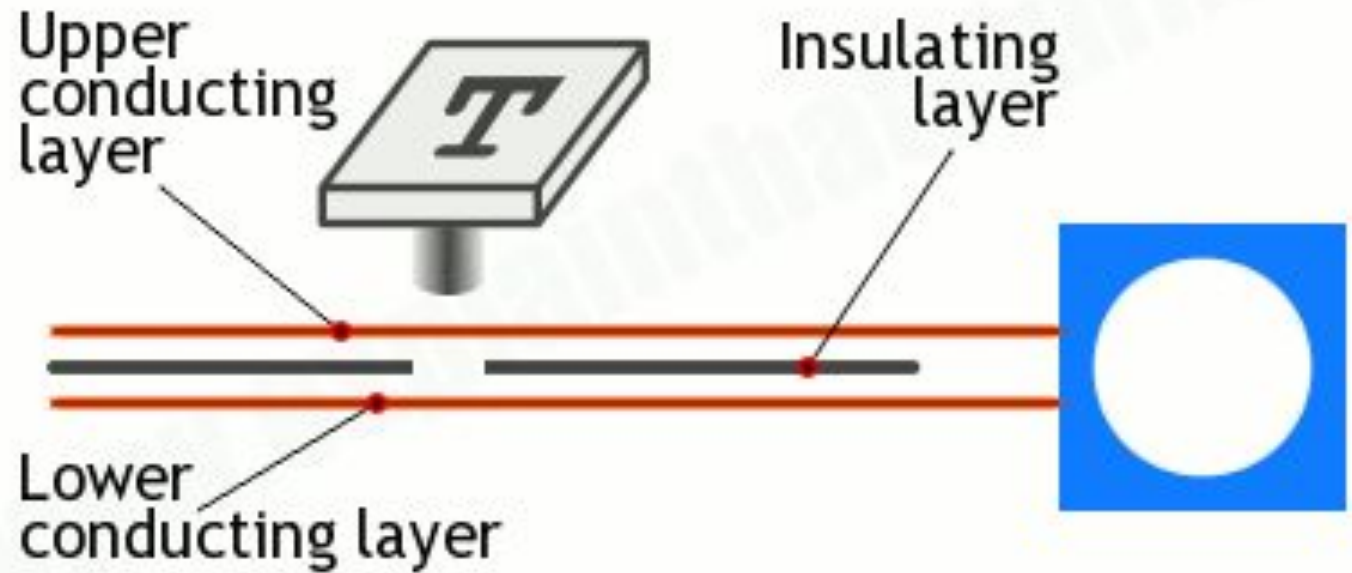
DISCLAIMER: MIGHT CONTAIN A LOT OF ACRONYMS

The enter key is pressed



Keyboard

www.explainthatstuff.com



Some electronic stuff happens on the USB level

- **“Universal Serial Bus”**
- The USB circuitry of the keyboard is powered by the 5V supply provided over pin 1 from the computer's USB host controller.
- The host USB controller polls that "endpoint" every ~10ms (minimum value declared by the keyboard), so it gets the keycode value stored on it.
- This value goes to the USB SIE (Serial Interface Engine) to be converted in one or more USB packets that follow the low level USB protocol.
- Those packets are sent by a differential electrical signal over D+ and D- pins (the middle 2) at a maximum speed of 1.5 Mb/s, as an HID (Human Interface Device) device is always declared to be a "low speed device" (USB 2.0 compliance).
- This serial signal is then decoded at the computer's host USB controller, and interpreted by the computer's Human Interface Device (HID) universal keyboard device driver. The value of the key is then passed into the operating system's hardware abstraction layer.
- **BLAH BLAH BLAH BLAH BLAH**

The Keydown event (some more stuff I copy pasted)

The interrupt signal triggers an interrupt event in the I/O Kit kext keyboard driver. The driver translates the signal into a key code which is passed to the OS X WindowServer process. Resultantly, the WindowServer dispatches an event to any appropriate (e.g. active or listening) applications through their Mach port where it is placed into an event queue. Events can then be read from this queue by threads with sufficient privileges calling the `mach_ipc_dispatch` function. This most commonly occurs through, and is handled by, an `NSApplication` main event loop, via an `NSEvent` of `NSEventTypeKeyDown`.

Your operating system/kernel

- In charge of multitasking and handling input/output drivers
 - Drivers provide the interrupt handling required for any necessary *asynchronous time-dependent hardware interface*.
- It checks if it is a system-defined keypress (like cmd-alt-del)
- It keeps track of the active application and “sends” the keypress event to that process
- You can learn how this works in **Operating Systems** (15-410)

Url is parsed

Parse URL (UNIVERSAL RESOURCE LOCATOR)

- The browser now has the following information contained in the URL (Uniform Resource Locator):
 - *Protocol "http"*
 - Use 'Hyper Text Transfer Protocol'
 - *Resource "/"*
 - Retrieve main (index) page



Domain names are often confused with URLs
but they are not the same.

Detour...what's the internet?

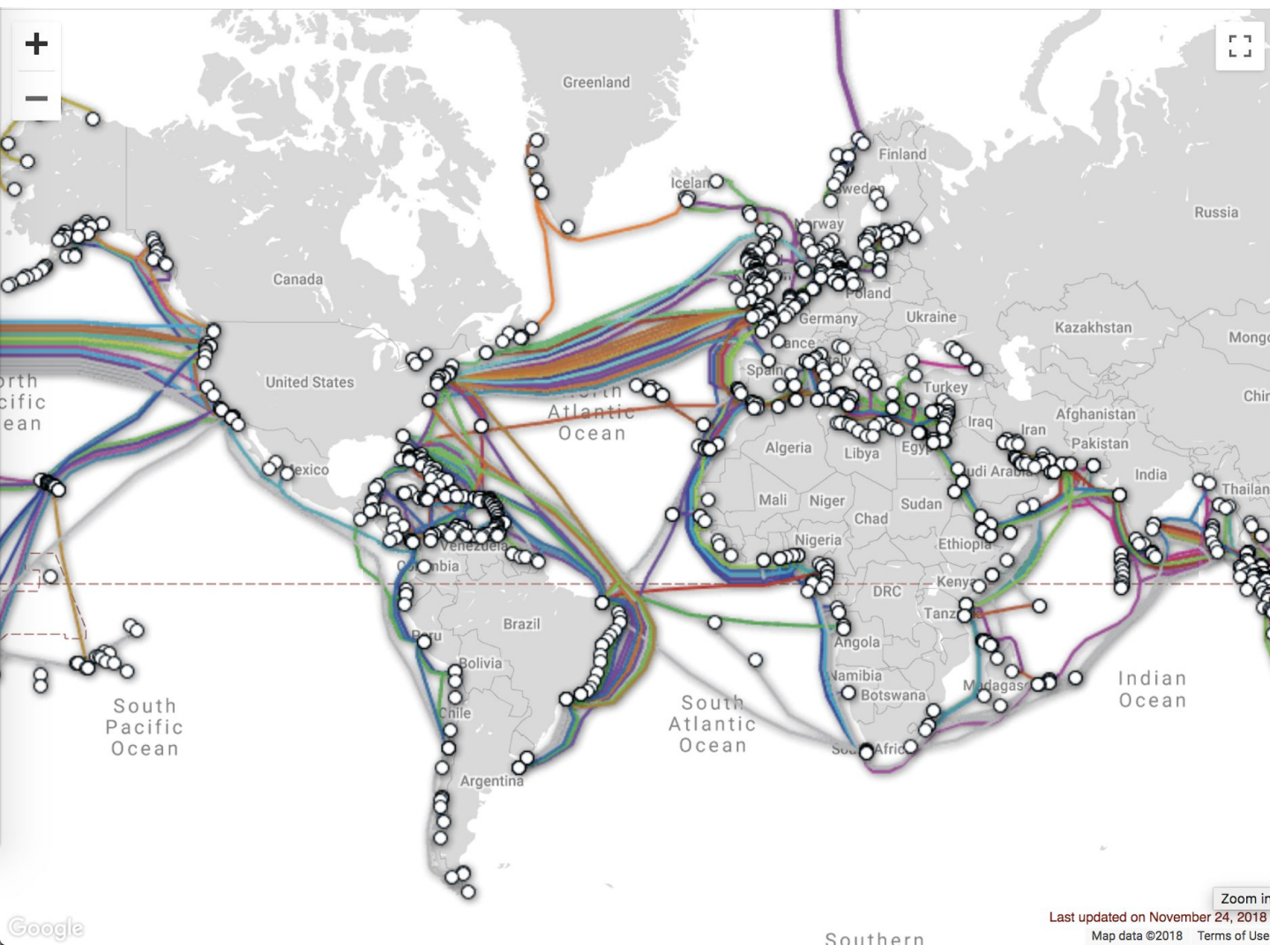
Ask the class what they think here



No seriously...the internet is a bunch of cables

This actually blew my mind





TeleGeography Submarine Cable Map

The Submarine Cable Map is a free and regularly updated resource from TeleGeography.

Got a question about how we make this map? Or about how submarine cables work? [Look no further.](#)

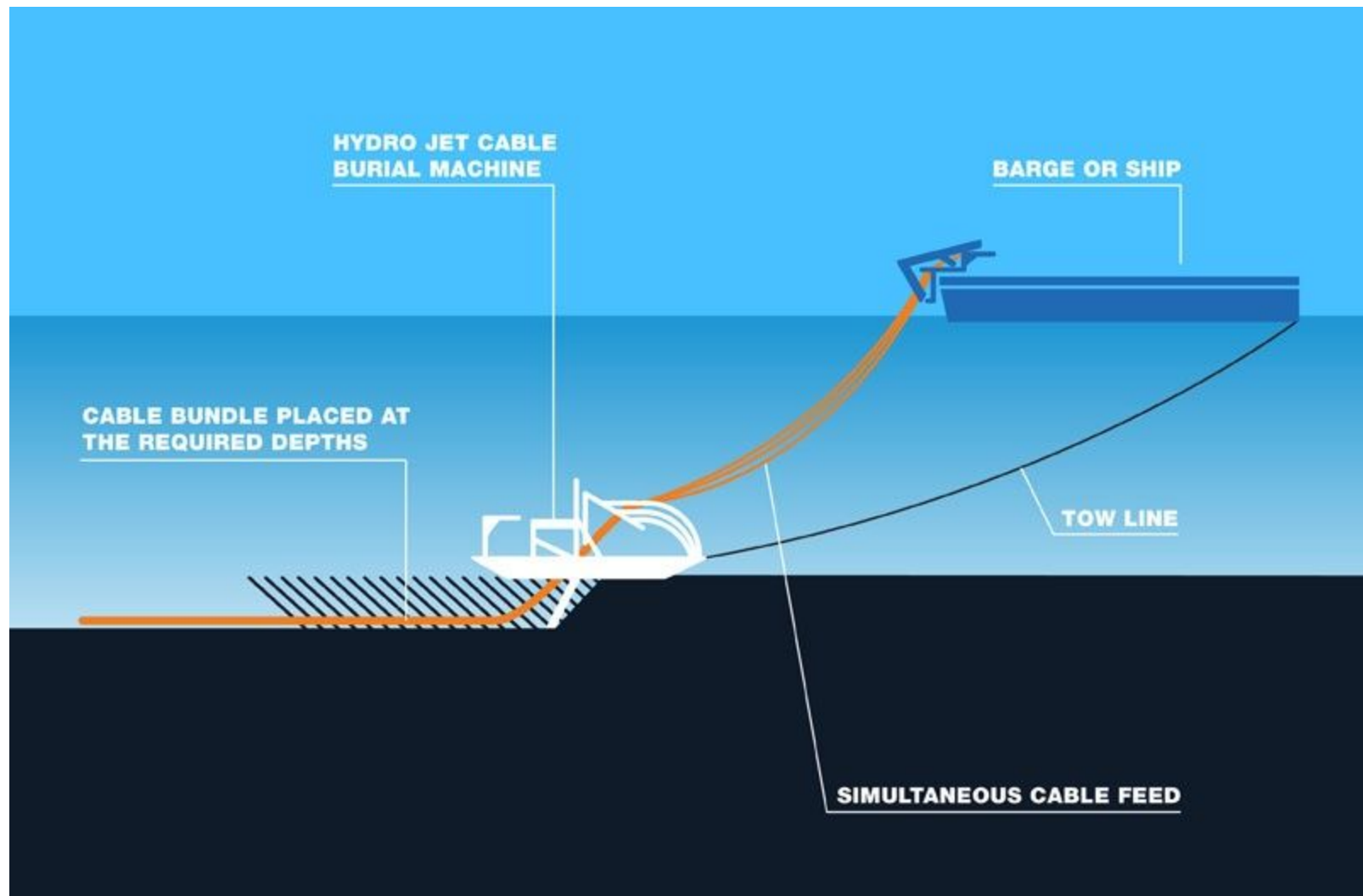
Feedback [t](#) [f](#) [github](#)

Submarine Cables

- [ACS Alaska-Oregon Network \(AKORN\)](#)
- [Aden-Djibouti](#)
- [Adria-1](#)
- [AEC-1](#)
- [Africa Coast to Europe \(ACE\)](#)
- [Africa-1](#)
- [Alaska United East](#)
- [Alaska United Southeast](#)
- [Alaska United Turnagain Arm \(AUTA\)](#)
- [Alaska United West](#)
- [ALBA-1](#)
- [Aletar](#)
- [Alonso de Ojeda](#)
- [ALPAL-2](#)
- [America Movil Submarine Cable System-1 \(AMX-1\)](#)
- [American Samoa-Hawaii \(ASH\)](#)
- [Americas-I North](#)
- [Americas-II](#)
- [Amerigo Vespucci](#)
- [Antillas 1](#)
- [APCN-2](#)
- [Aphrodite 2](#)
- [Apollo](#)
- [Aqualink](#)
- [ARBR](#)

Last updated on November 24, 2018
Map data ©2018 Terms of Use

All content © 2018 PriMetrica, Inc.

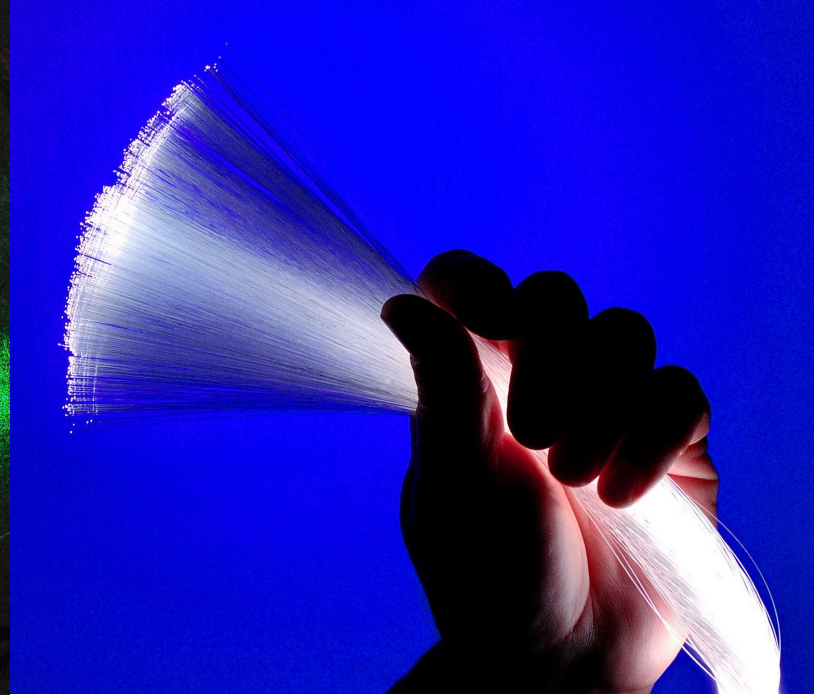
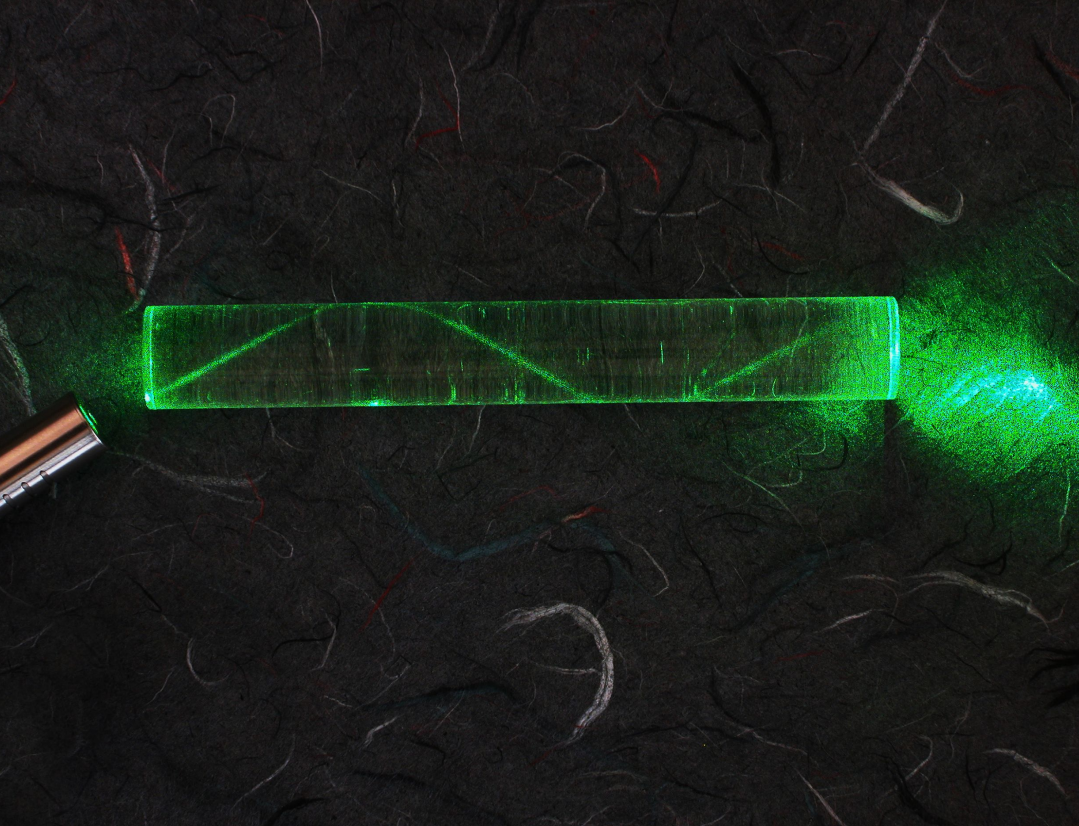




The **US\$300 million FASTER cable system**, backed by Google and East Asian telecom firms, will have a peak capacity of 60 Tbps (terabits per second) when it's ready next year.

About 6,000 kilometers of the cable has been loaded onto the cable ship René Descartes, which is docked at a submarine cable factory in Kitakyushu, southern Japan, operated by NEC group firm OCC. It will be the first time for the vessel, owned by French telecom firm Orange, and her crew to lay some 9,000 km of cable in one go.

These cables are pretty hi-tech



Whats an IP (Internet Protocol) Address?

- Numerical label assigned to each device connected to the internet
- Used for interface identification
 - Ex: 172.16.254.1
- Before we communicate, we need to find the IP address of our website

DNS Lookup

- Stands for Domain Name Service
- Phone book for the internet
- Maps a domain name -> IP address
- Uses **ARP** -> Address resolution protocol

Yo where's google.com



Your Computer



Root TLD
(top level domain)
name server

Idk man but I do know
a buddy that can help
you



Your Computer



Root TLD name server

Idk man but I do know
a buddy that can help
you



Your Computer



Root TLD name server



.com name server

peace



Your Computer



Root TLD name server



.com name server

Yo where's google.com



Your Computer

Hmmm...I actually
might know...the
location of the
authoritative name
server



.com name server

Sick I'm going to **cache**
this for later



Your Computer



.com name server

What happens when you register a domain?

- Verisign -> for profit company that owns .com
- You buy from a *registar*
- Registrars are like high street shops – they actually get the domains from wholeseller-equivalents, but provide a convenient way for customers to buy and manage domains.
- Godaddy/Namecheap/etc.
- You pay money for the governing bodies to set a custom dns server for your own “domain.”

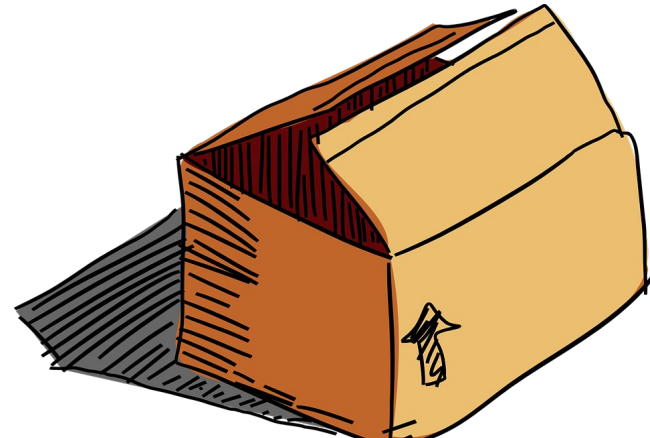


Okay...we now have a the IP address...now what?

216.58.217.142

It gets pretty damn complicated.

Data is transmitted through the internet through these little things called packets



Some stuff that I copy pasted

Once the browser receives the IP address of the destination server, it takes that and the given port number from the URL (the HTTP protocol defaults to port 80, and HTTPS to port 443), and makes a call to the system library function named `socket` and requests a TCP socket stream - `AF_INET/AF_INET6` and `SOCK_STREAM`.

- This request is first passed to the Transport Layer where a TCP segment is crafted. The destination port is added to the header, and a source port is chosen from within the kernel's dynamic port range (`ip_local_port_range` in Linux).
- This segment is sent to the Network Layer, which wraps an additional IP header. The IP address of the destination server as well as that of the current machine is inserted to form a packet.
- The packet next arrives at the Link Layer. A frame header is added that includes the MAC address of the machine's NIC as well as the MAC address of the gateway (local router). As before, if the kernel does not know the MAC address of the gateway, it must broadcast an ARP query to find it.

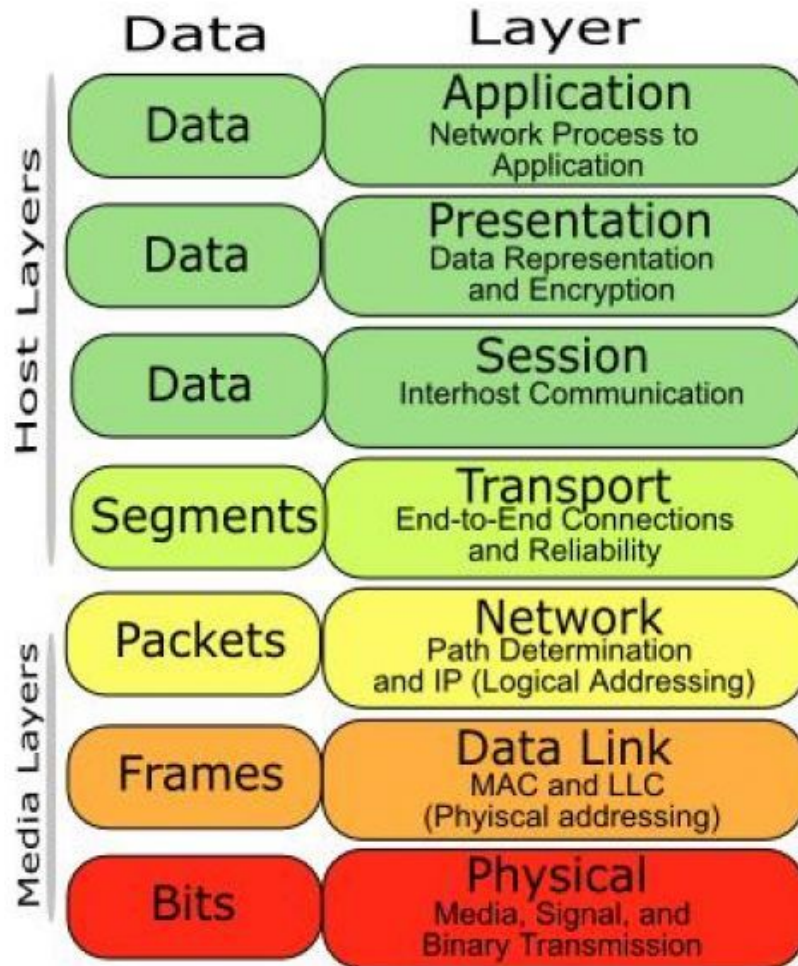
But we open a socket

Basically a socket is an open communications channel between two machines. In this case, the two machines talking are your computer and the server.

Implemented/handled by the operating system



OSI Model



HTTP Request

HTTP means HyperText Transfer Protocol.

HTTP is the underlying protocol used by the World Wide Web and this protocol defines how messages are formatted and transmitted, and what actions Web servers and browsers should take in response to various commands.

It's the “top-level” layer of a packet





Transmission

- These packets are then passed through a modulator/demodulator...
 - Sounds familiar?

Transmission

- These packets are then passed through a modulator/demodulator...
 - Translates digital discrete signals to real-world electronic waves

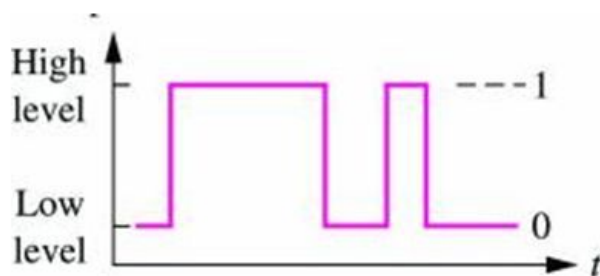
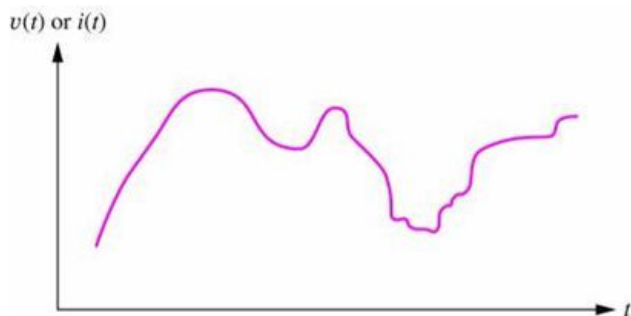
modulator/demodulator



Signals

- Each packet can be represented/encoded by a series of 1s and 0s
- This is then converted into **analog** signals
- Good thing we're not ECE majors! We don't have to really learn this
- Mad *wavey*

Analog vs. Digital Signals



***Digital signals themselves
are really analog signals!***

- ❖ Analog signals take on continuous amplitude values – EEs typically use charge, voltage, or current (current = charge flow per unit time).
- ❖ Analog Signals represent the physical world.
e.g. Electrocardiogram (ECG) signals are analog signals:



- ❖ Digital signals take on discrete amplitude levels. Typically, we use binary signals which utilize only two levels. One level is referred to as *logical 1*, the other level is referred to as *logical 0*.
- ❖ Modern computers and data communications systems process digital signals

These get passed through cables close to the speed of light....to a server?



What's a server?

A server, much like a browser, is a process continuously running on someone else's computer.

Optimized!

Makes thousands of simultaneous connections to browsers

The NGINX logo is displayed in a bold, green, sans-serif font. The letters are stylized, with the 'i' and 'x' having unique shapes. The logo is centered within a light gray rectangular background.

What does an HTTP server do?

- *The server breaks down the request to the following parameters:*
 - HTTP Request Method (either GET, HEAD, POST, PUT, DELETE, CONNECT, OPTIONS, or TRACE). In the case of a URL entered directly into the address bar, this will be GET.
 - Domain, in this case - google.com.
 - Requested path/page, in this case - / (as no specific path/page was requested, / is the default path).
- The server verifies that the client is allowed to use this method (by IP, authentication, etc.).
 - Is the user authenticated?
- The server goes to pull the content that corresponds with the request, in our case it will fall back to the index file, as "/" is the main file (some cases can override this, but this is the most common method).
- And gives you a response code... **404 is bad**
- **You implement your own multi-threaded web server in 15-440 (Distributed Systems)**

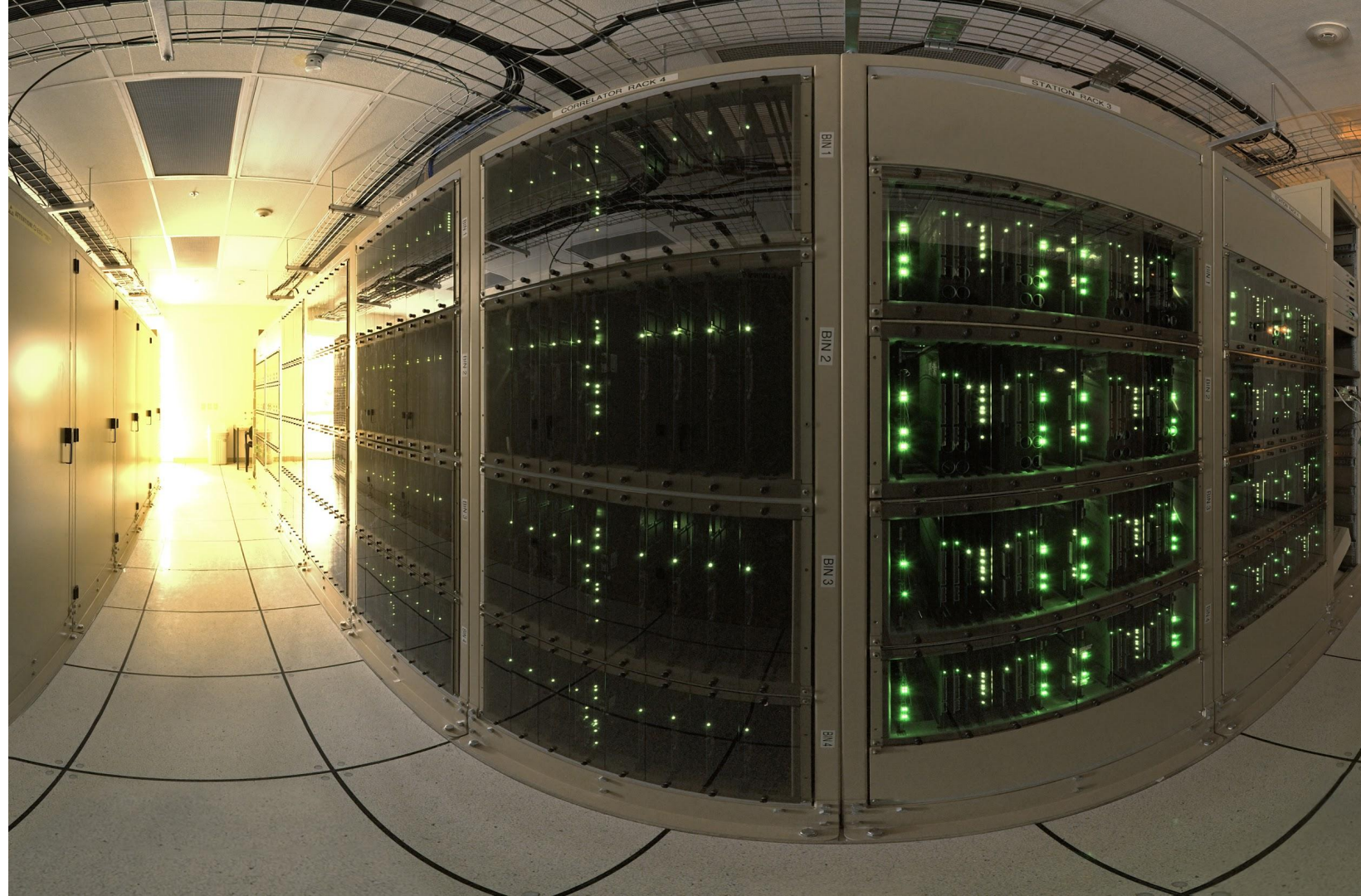
Server backends dont just serve webpages

- What if you have content you have to store, such as checking if a user exists?
- Can't just store in a massive text-file....extremely ineffecient
- Your server access a **database** for this information
- Servers also run scripts, such as using the CGI protocol, to dynamically generate content
 - PHP, Node.js, .NET, Java Server Pages

How is the hardware optimized?

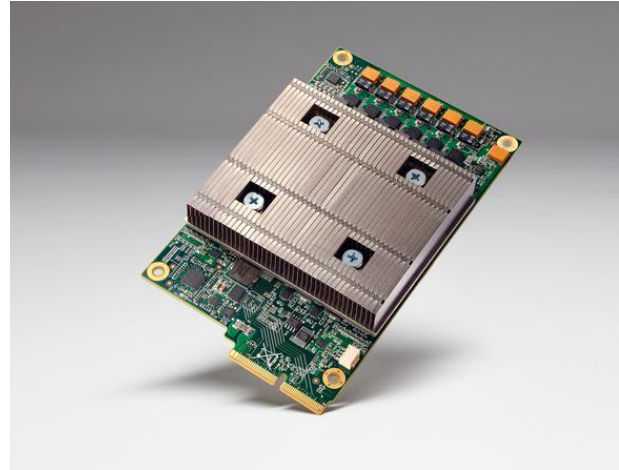
None of this uses supercomputers; Google basically runs on commodity Intel hardware stuck in shipping containers in areas with low electricity costs.





Well, okay, also the TPUs

Tensor Processing Units are special-purpose processors for running Machine Learning workloads. Google developed them and uses them to answer search queries.



Browser (in this case chrome)



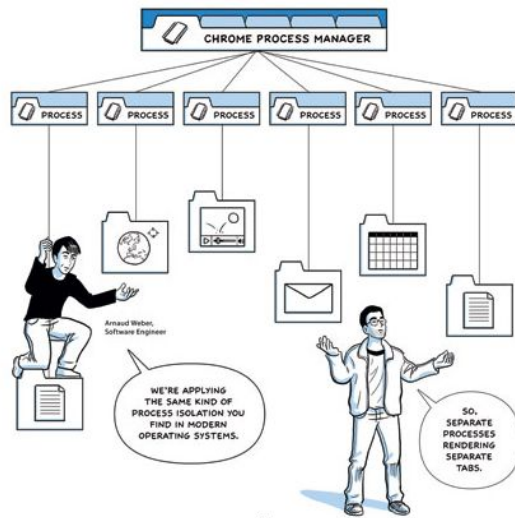
What is a browser?

Browsers render “web content” accessed via URLs.

Typically, web content is Javascript/HTML/CSS over HTTP(S).

But browsers also support, or have supported, weirder things like web assembly, RSS, FTP, Gopher, NNTP, Archie.

Side note: The Google Chrome comic



4

<https://www.google.com/googlebooks/chrome/>

The browser's functionality is to present the web resource you choose, by requesting it from the server and displaying it in the browser window. The resource is usually an HTML document, but may also be a PDF, image, or some other type of content. The location of the resource is specified by the user using a URI (Uniform Resource Identifier).

The way the browser interprets and displays HTML files is specified in the HTML and CSS specifications. These specifications are maintained by the W3C (World Wide Web Consortium) organization, which is the standards organization for the web.

Overview of Browser Internals

- User interface: various
- Extension API: WebExtensions (Chrome, Firefox)
- Browser engine: Webkit, Blink, Gecko, Servo, Trident
- Graphics engine: Skia
- Networking: Berkeley Sockets API
- UI backend: GTK/QT
- JavaScript engine: V8/Spidermonkey/Chakra
- Data storage: Cookies and local things

How a browser renders web content

- Download and decompress web content
- Render HTML using layout engine
- Render CSS using layout engine
- Compile and execute javascript

Everything compiles to the DOM

The DOM is the Document Object Model. It's the HTML tree structure and the common denominator for rendered web content.

The final rendered page is a reflection of the DOM.



And finally...

A white rectangular search input field with a thin gray border and a small vertical cursor line on the left side.

Google Search

I'm Feeling Lucky


becomes...





[All](#) [Images](#) [News](#) [Videos](#) [Books](#) [More](#) [Settings](#) [Tools](#)

About 9,980,000 results (0.50 seconds)

To **quit** the vi editor without saving any changes you've made: If you are currently in insert or append mode, press Esc. Press : (colon). The cursor should reappear at the lower left corner of the screen beside a colon prompt. Enter the following: q! Oct 1, 2013



[keyboard shortcuts - Exit vim more quickly - Unix & Linux Stack ...](https://medium.freecodecamp.com/keyboard-shortcuts-exit-vim-more-quickly-unix-linux-stack-overflow)
<https://unix.stackexchange.com/questions/93144/exit-vim-more-quickly>

 About this result  Feedback

People also ask

How do I save and quit in vi?	▼
How do I stop recording in Vim?	▼
How do I close vim help?	▼
How do I save and exit changes in Vim?	▼

Feedback

Conclusion

- Yeah....computers are complicated
- There are **so many** moving parts that have to work seamlessly behind the scenes for such a simple task
- After 4 years at CMU you come to appreciate the little things in life
 - And understand how they work!
- Goodbye!
- Hope you learned something at GPI!