

# TEU00311

## What is the Internet doing to me? (witidtm)

Stephen Farrell

[stephen.farrell@cs.tcd.ie](mailto:stephen.farrell@cs.tcd.ie)

<https://github.com/sftcd/witidtm>

<https://down.dsg.cs.tcd.ie/witidtm>

URLs accessed 20190907

# What're we here for?

- We all use the Internet all the time
- You may or may not know what's happening under the hood, and shouldn't need to know all the nitty-gritty detail
- But, to make better decisions as to what you do, it's good to know something about some of those details
- This module aims to help you learn enough to make better decisions about what you want, and how to get it, as you interact with the Internet
- I hope: you'll apply those lessons, tell others about it all and maybe agitate for a better Internet for a better society (but you won't fail the module if you don't agitate:-)

# Administrivia

# TCD Personnel/Contacts

- Lectures:
  - Dr. Stephen Farrell, [stephen.farrell@cs.tcd.ie](mailto:stephen.farrell@cs.tcd.ie), WR3.4, x2354
  - Dr. Dave Lewis TBD
- Teaching Assistant:
  - Christian Cabrera, [cabrerac@scss.tcd.ie](mailto:cabrerac@scss.tcd.ie)
- Demonstrators: TBD

# Schedule

- Tuesday 1000-1050, Synge 2, Hamilton building
- Thursday 1600-1750, M17, Museum building
- Weeks 3-7: Just lectures
- Weeks 8,10-14:
  - Themed lecture (Tue, 1000-1050)
  - Hackathon/lab (Thu, 1600-1750) – **venue moves to Regent's House**
- Reading week: Week 9, October 21
- September 10<sup>th</sup> is in week 3 - Why do week numbers start at 3? I've no idea;-)
- Details of labs are being worked out, stay tuned

# External Experts

- Most labs/hackathons will be lead by an external expert, most will be remote
- Getting busy people to help (for nothing:-) means being flexible so changes are likely as we go
- Will let you know as dates/topics are firmed up

# Assessment

- Blackboard tests – 45% (see next slide)
  - [https://tcd.blackboard.com/webapps/blackboard/execute/courseMain?course\\_id=\\_56807\\_1](https://tcd.blackboard.com/webapps/blackboard/execute/courseMain?course_id=_56807_1)
  - I'm not a huge fan of blackboard but we'll try...
- GDPR request and anonymised report – 25 %
- Group project (report) – 15%
- Individual project (report) – 15%
- No exam unless you flunk, then a supplemental for 100%, but don't flunk!

# Blackboard tests

- Initial survey (not marked)
- Week 2 - Small essay (5%)
- Week 3 - Multiple Choice Questions (MCQ) on content to date (5%)
- Week 4 - As with week 3:-) (5%)
- Hackathons: MCQ before each so you have enough background to (3 x 10%)
- For all BB tests – deadlines apply, late submission means high penalties, details will be posted for each
- I may change some of that depending how well the quiz thing works (or not!)



# Module Materials

- There is no book – feel free to recommend some if you like
- Materials will be linked to from, or on, the module web page:
  - <https://down.dsg.cs.tcd.ie/witidtm>
- Content of module web page is also in Github at:
  - <https://github.com/sftcd/witidtm>
- Clone that repo and/or visit that page often, as it will change during the runtime of the module!
  - Who knows what “clone that repo” means?
  - I’ll be happy to take PRs, if offered – if **very** good I might even give some marks of let someone off an assignment

# Style

- The module is designed to be interactive, so don't sit there and say nothing!
- It is entirely ok to ask what might appear to be dumb questions, e.g. "Who makes money from YouTube?" - supposedly naive questions can be good and the answers might be quite subtle
- It is entirely ok to comment on what we tell you, e.g. "That's nonsense – I use <foo> all the time and it's fine" - this is about you after all, so (dis)agreeing with us and one another is desirable (but don't be an ass, and do listen)
- If you don't comment or ask questions, we'll all be more bored and I'll get cranky!
- So, let's start...

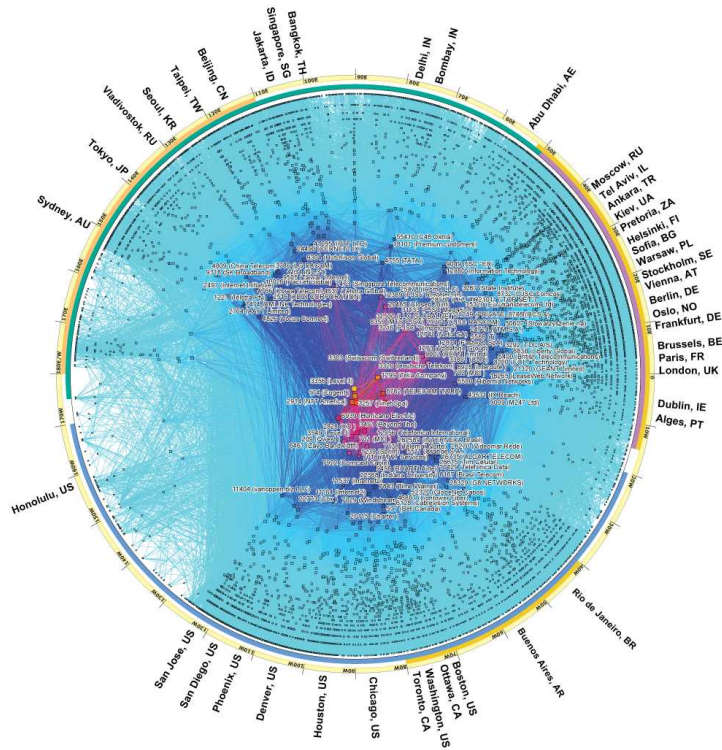
Is the Internet a network?

Is the Internet a network?  
(hint: the answer is “no”:-)

# A network of networks

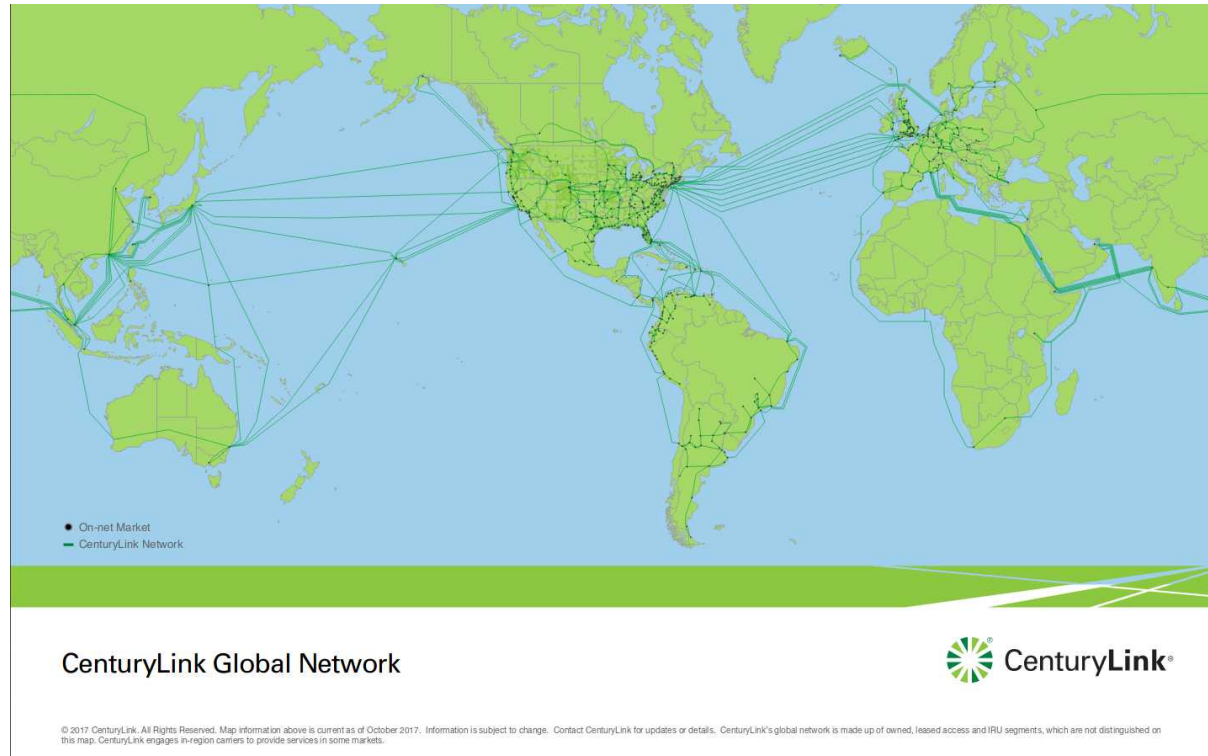
- The Internet is made up of about 65,428 Autonomous Systems (ASes)
  - [https://en.wikipedia.org/wiki/Autonomous\\_system\\_%28Internet%29](https://en.wikipedia.org/wiki/Autonomous_system_%28Internet%29)
  - 65,428 number as of Aug 19 (<https://www.cidr-report.org/as2.0/>)
- Think of these as the set of Internet Service Providers (ISPs, like Eircom), other networks (e.g. HEANET which is TCD's ISP), big companies (e.g. Google, FB) and oddities like Internet eXchange Points (IXPs, like INEX)
- Each is (in principle and often in practice) an independent network (or set of networks) and their operators can do whatever they want
  - They're essentially defined by sets of numbers: Static: AS number (ASN); Dynamic: sets of IP address prefixes
- They interact using Internet protocols (like IP, TCP, BGP)
  - IP: Internet Protocol; TCP: Transmission Control Protocol; BGP: Border Gateway Protocol
- We'll delve more into all that later, but first... some pretty pictures

# CAIDA Map of ASes



- CAIDA (Center for Applied Internet Data Analysis) is a UC San Diego Internet measurement organisation
  - You can measure **a lot** of what happens on the Internet as it happens!
- This is a 2017 map of the ASes as they were then  
[https://www.caida.org/research/topology/as\\_core\\_network/2017/](https://www.caida.org/research/topology/as_core_network/2017/)
- More central => more connected, serving more people
  - You can see Dublin at about 3 o'clock
  - In the middle (if you zoom) you could see level3 and cogent which are large ASes

# Level3 is one of those (a BIG one)



<https://www.level3.com/~media/files/maps/en-network-services-level-3-network-map.ashx>

# Cogent similarly

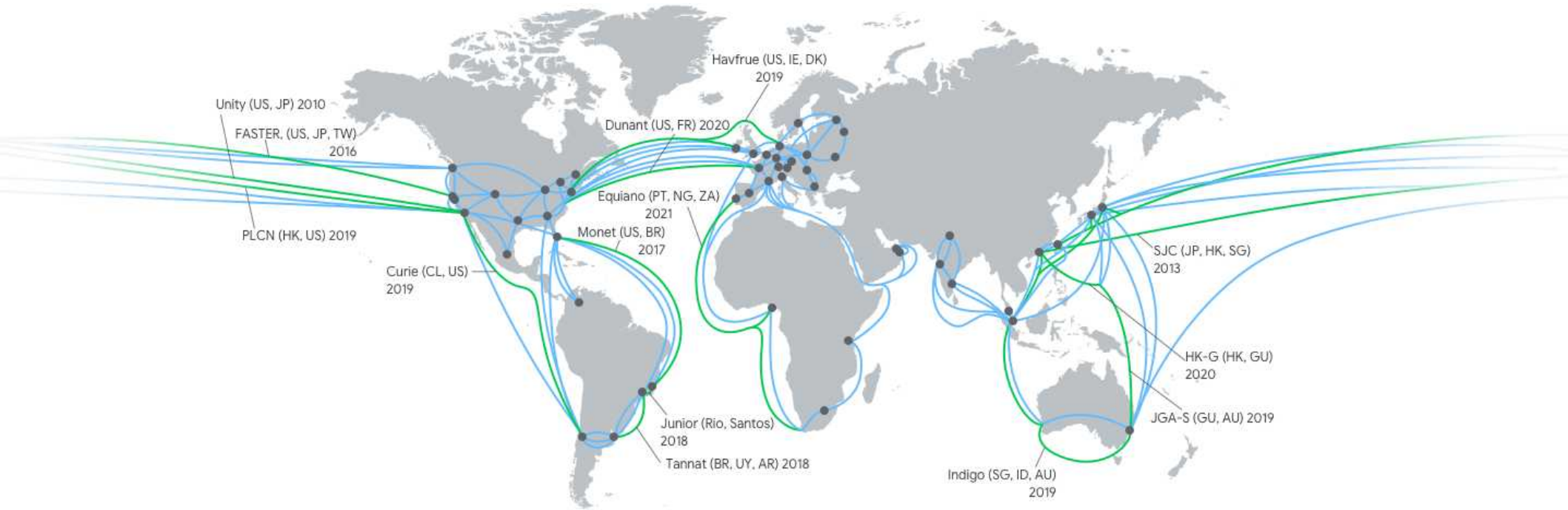


Cogent's network stretches over 204 markets throughout **43 countries** in North America, Europe and Asia, with over **57,400 route miles of long-haul fiber** and more than **34,100 miles of metropolitan fiber** serving over **840 metropolitan rings**. Our end-to-end optical transport network consists of IP-over-WDM fiber links running up to **2,120 Gbps** intercity capacity and **2,000 Gbps** on metropolitan rings, located in Cogent's major markets throughout North America, Europe and Asia. On the IP layer, Cogent's Tier 1, IPv6 and MPLS enabled network has direct IP **connectivity to more than 6,760 AS** (Autonomous System) networks around the world with over **185 Tbps** internetworking capacity.

<https://www.cogentco.com/en/network/network-map>

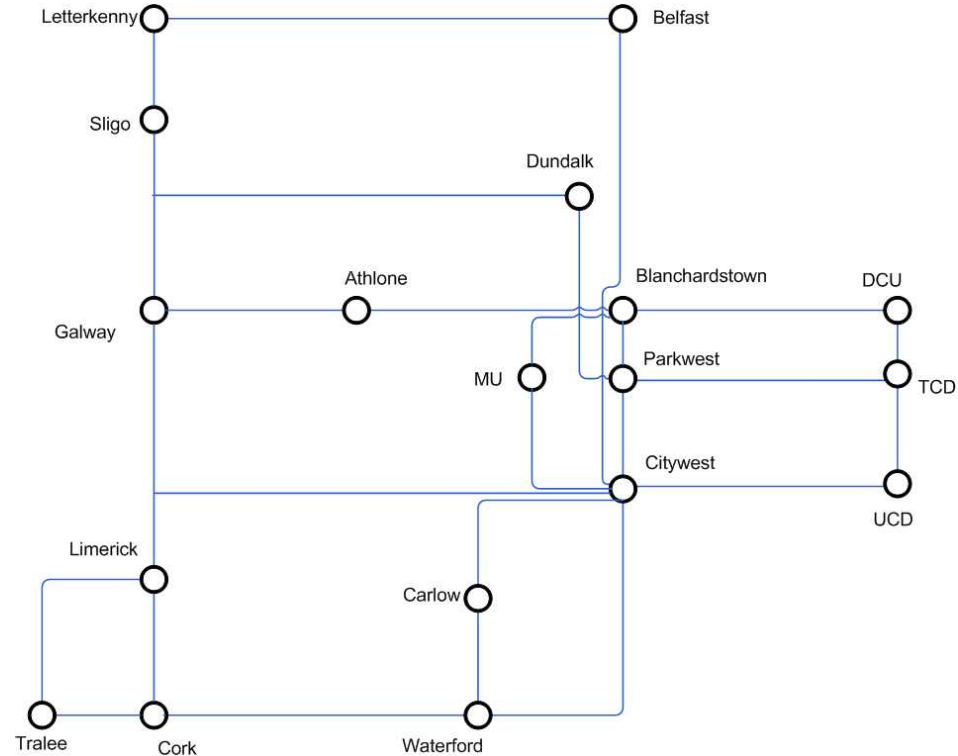


# Google cloud



<https://cloud.google.com/about/locations/#network-tab> note this is just google cloud, not all their stuff

# Heanet national n/w

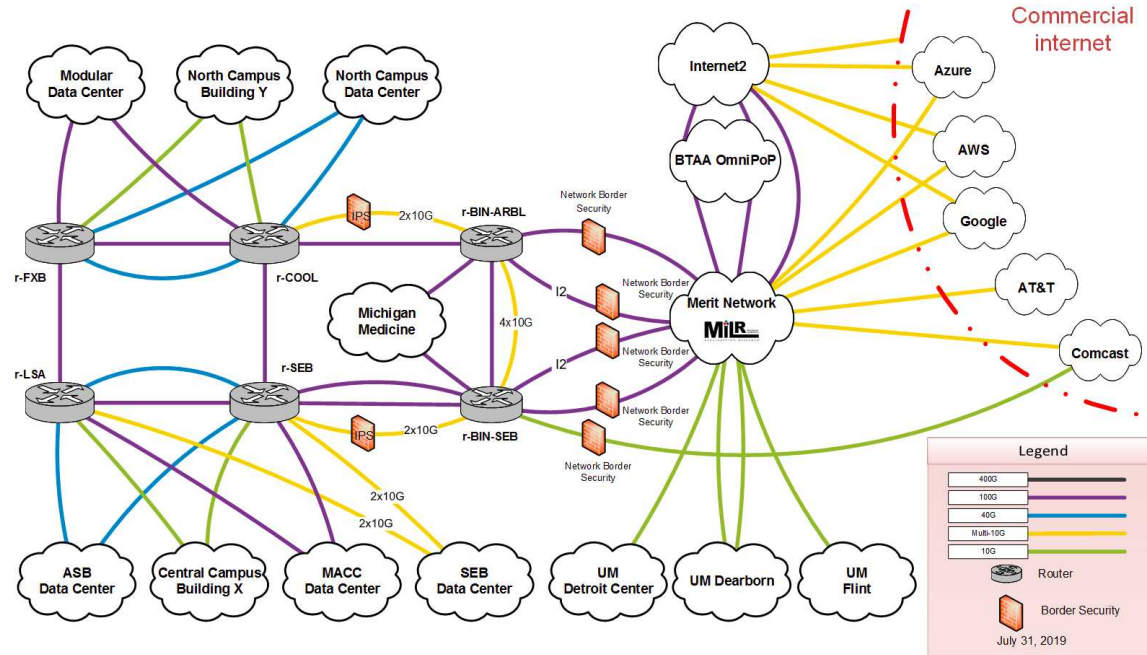


# The TCD network

- Me: Could you send me a n/w diagram to give students an indication of the complexity of a campus n/w?
- Them: No.... <mumble> ... security ;-(
- Ah well, they're mostly similar so...

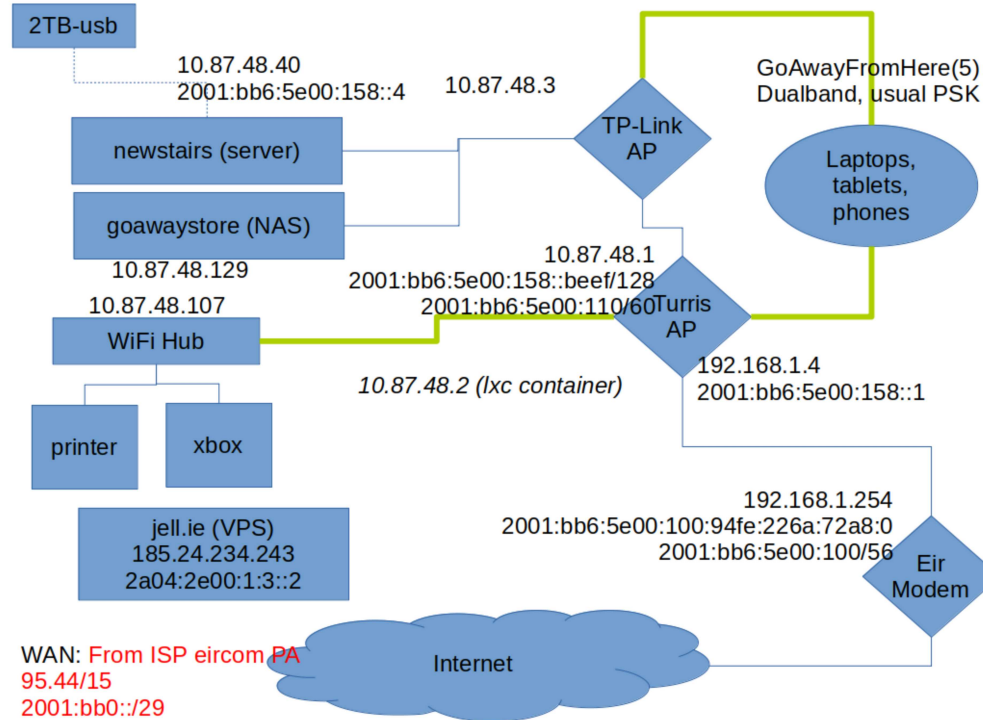
# Umich n/w diagram

## University of Michigan Network And connections to the World



<https://its.umich.edu/enterprise/wifi-networks/campus-network-diagram-description>

# My home network



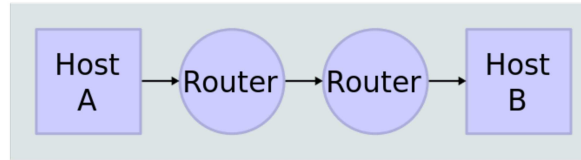
That's from a few years ago – numbers are out of date:-)

# Interoperability

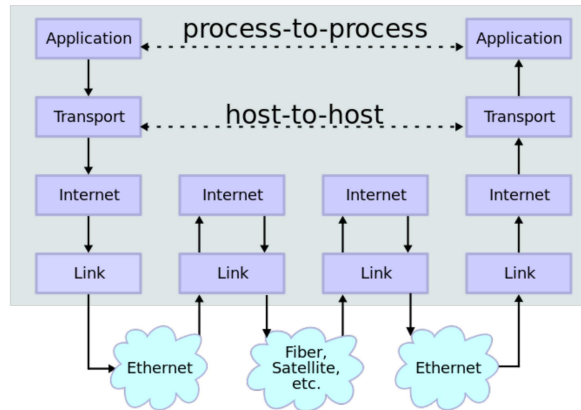
- To make the Internet work we need to agree on how to **interoperate** for some basic/minimal set of things
  - That means defining/agreeing on Internet Protocols
  - Where we need to agree on how to interoperate, a lot of that is done by the Internet Engineering Task Force (IETF) and other Internet standards bodies (IEEE SA, W3C)
    - I'm quite involved with IETF stuff, so consider me biased there:-)
- But we do not aim to agree about everything in everyone's network
  - So an awful lot happens at the “application layer” in code written by people and organisations
  - And yet more happens when people configure services that use that code

# What's a network protocol?

Network Topology



Data Flow



[https://en.wikipedia.org/wiki/File:IP\\_stack\\_connections.svg](https://en.wikipedia.org/wiki/File:IP_stack_connections.svg)

# “Permissionless innovation”

- One important point is: in principle each network operator can do whatever it wants so long as it interoperates “nicely” with others (and even when it doesn’t act particularly nicely;-)
  - That also applies to your home network (if you want and are able)
  - There are no protocol police
- This is one of the main reasons why the Internet has been so successful
- Related: the classic “End-to-end argument” paper
  - Salzer, Reed. Clark, “End-to-end arguments in system design” ACM ToCS (1984).
    - <https://web.mit.edu/Saltzer/www/publications/endtoend/endtoend.pdf>
  - I recommend a read of that!
  - Don’t consider it as gospel though – it’s the end-to-end **argument** and not really the end-to-end **principle** even though it gets called the latter a lot



# “Tussles”

- Repeating: we do not aim to agree about everything in everyone’s network...
  - So an awful lot happens at the “application layer” in code written by people and organisations
  - And yet more happens when people configure services that use that code
- When the “policies” reflected in those collide then “fun” follows;-)
  - If protocols or application code constrains what operators can do then people complain
  - If what n/w operators are doing breaks (esp changes to) applications then people complain
  - In both cases people often complain at the wrong place;-)
- Another paper:
  - Clark, David D., et al. "Tussle in cyberspace: defining tomorrow's Internet." ACM SIGCOMM Computer Communication Review. Vol. 32. No. 4. ACM, 2002.
    - <https://www2.cs.duke.edu/courses/compsci514/cps214/spring09/papers/p347-clark.pdf>
  - Same “Clark”, but older:-) Interestingly, the 2002 paper is IMO far more dated (and wrong!) than the 1984 paper!
- We’ll consider a “live” example later (Search for “DNS over HTTPS” if you want to check it out); don’t worry if this issue seems a bit too hard to grok at this point

# Aside: Cyber<blah>

- Be wary of anyone who uses a term like “Cyber<blah>”
- ~90% of the time, that’s a strong indicator that they don’t really know what they’re talking about (if they did, they’d use a more precise and well-defined term)
- Sadly, about 10% of the time (and increasing) such terms are used because “the industry” keeps on doing it and people just repeat stuff thoughtlessly
- Don’t be afraid to ask someone to define “Cyber<blah>” if they use such a term, and don’t be surprised if they find that hard!
  - E.g. “Does cyberspace include a person driving a car that’s had it’s license plate automatically scanned?” or “What’s not included in cybersecurity?”

# The Internet is not the web

- Another important point!
- The web is (roughly) the set of computers that speak the HTTP protocol
  - HTTP == HyperText Transfer Protocol
- Email doesn't use HTTP, but rather (mostly) the Simple Mail Transfer Protocol (SMTP) which is a couple of decades older than HTTP
- Mobile network internals (3G, 4G, 5G...) mostly run over IP using a bunch of protocols you'd prefer to never have to know about
- But lots of our interactions with the Internet are via the web
- We'll look more at the web later, meanwhile back to you...

How do you interact with the Internet?

Do you care about your, my, or all of our,  
security on the Internet?

If so, what do you care about most?

Do you care about your, my, or all of our  
privacy?

That's all for now, we'll be returning to those questions as we go.

But... what other questions should we be considering?

# Some of the things we'll do later...

- Understand what happens under the hood when your browser loads a web site
- Get an overview of advertising networks
- Learn how to watch network traffic
- Build a home network
- Be sneaky on a home network



# What else?

- Your topic here... what'd you like to cover?
- 
- 
- 

(we can revisit this multiple times)