CS10: The Beauty and Joy of Computing

The Internet: How it Works and How to Keep it That Way

In the news: flying file-sharing drones!

In the ongoing struggle between file-sharing organizations and lawmakers, the Pirate Bay recently announced its autonomous flying file-sharing drones that they hope to soon launch miles above Sweden.

Photo by Claus Langer

The Rise of the Internet

Computing has advanced extremely quickly during the past 40 – 50 years, and many of the most impactful inventions in recent years have been related to digital technology.

The Internet may have changed the world more than any other set of technologies.

Photo by Steve Rhode

The Internet + Us

The Internet is tightly integrated into our lives. If the Internet itself were to go down, we would feel it. Quickly.

Let’s take a look at how the Internet works, understand why it’s naturally difficult to disable, and look at some situations where part of it has been disabled recently.

The Young Internet

The foundation of the Internet came from research in the late 60's and early 70's. *The Internet has only been technologically feasible for 50 years, and popular for 20.*

Commercial ISP's began emerging in the late 1980's, and Internet adoption began exploding upward almost immediately afterward.

The Rapid Growth of the Internet

The Internet was successful for a number of reasons, many of them related to human factors.

There are at least two critical technical reasons as well:

- Decentralized architecture
- Open protocols

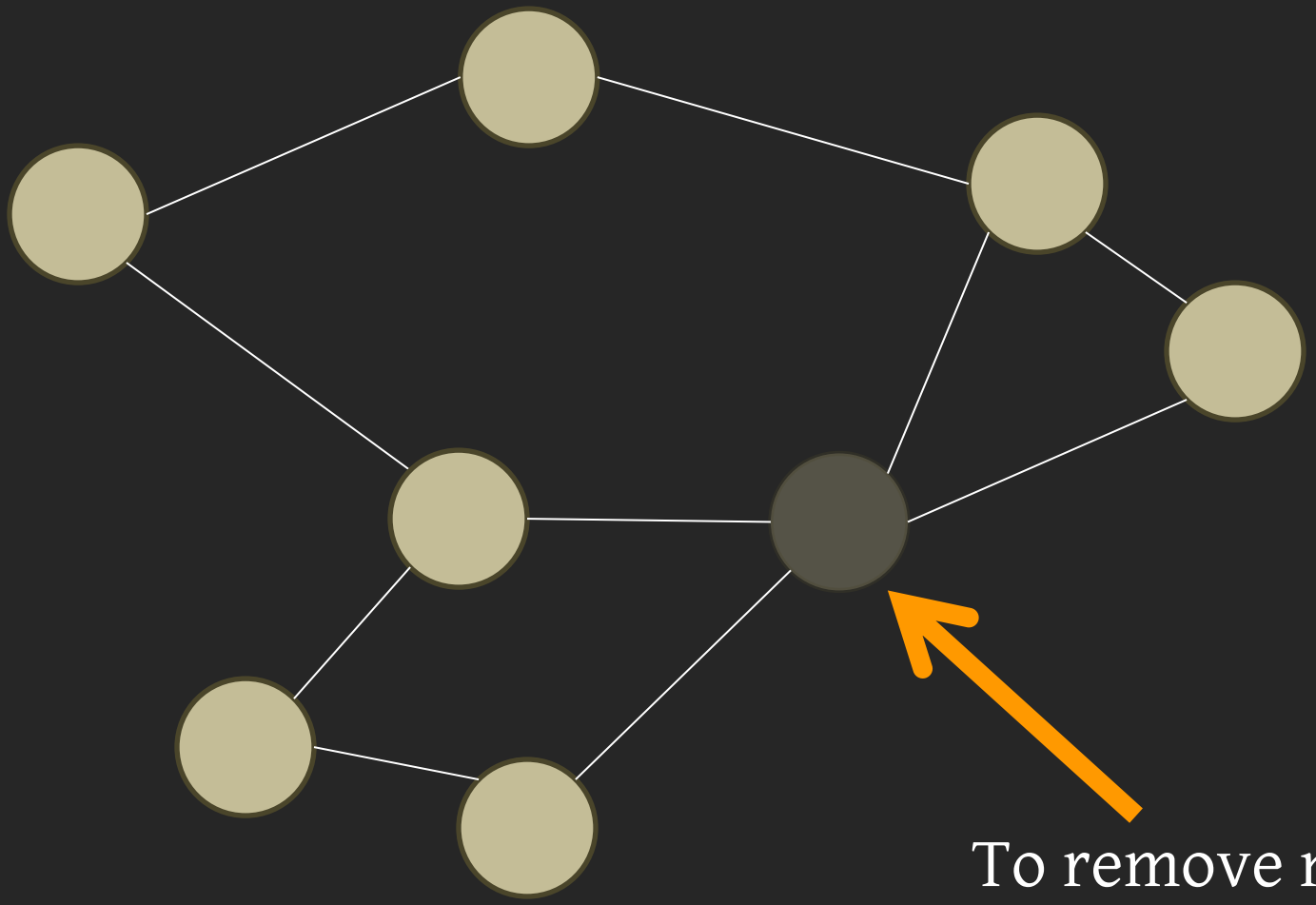
Technical Brilliance

Decentralized architecture

The Internet was designed as a decentralized system that wasn't owned by any single organization or government.

Open protocols

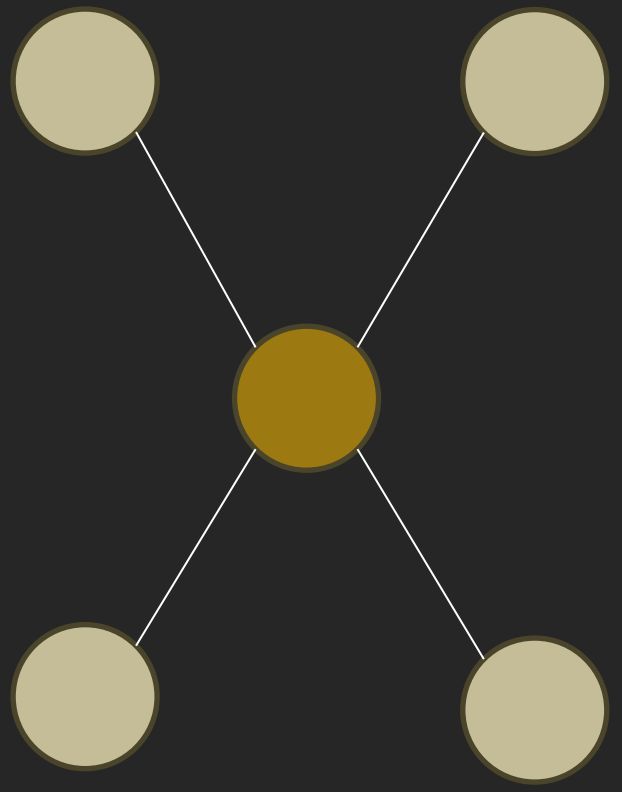
It used publicly available protocols that made it relatively easy for anyone to connect their devices or write new software that could communicate with others.


To remove nodes:
unplug them!

Routers: how information moves

Routers are a key piece of technology that makes the Internet so good at moving information.

They determine how to get information where it needs to go.



The Amorphous Internet

As you can see, the Internet is a pretty flexible thing!

Easy to add and remove new systems, and the architecture scales without limit.

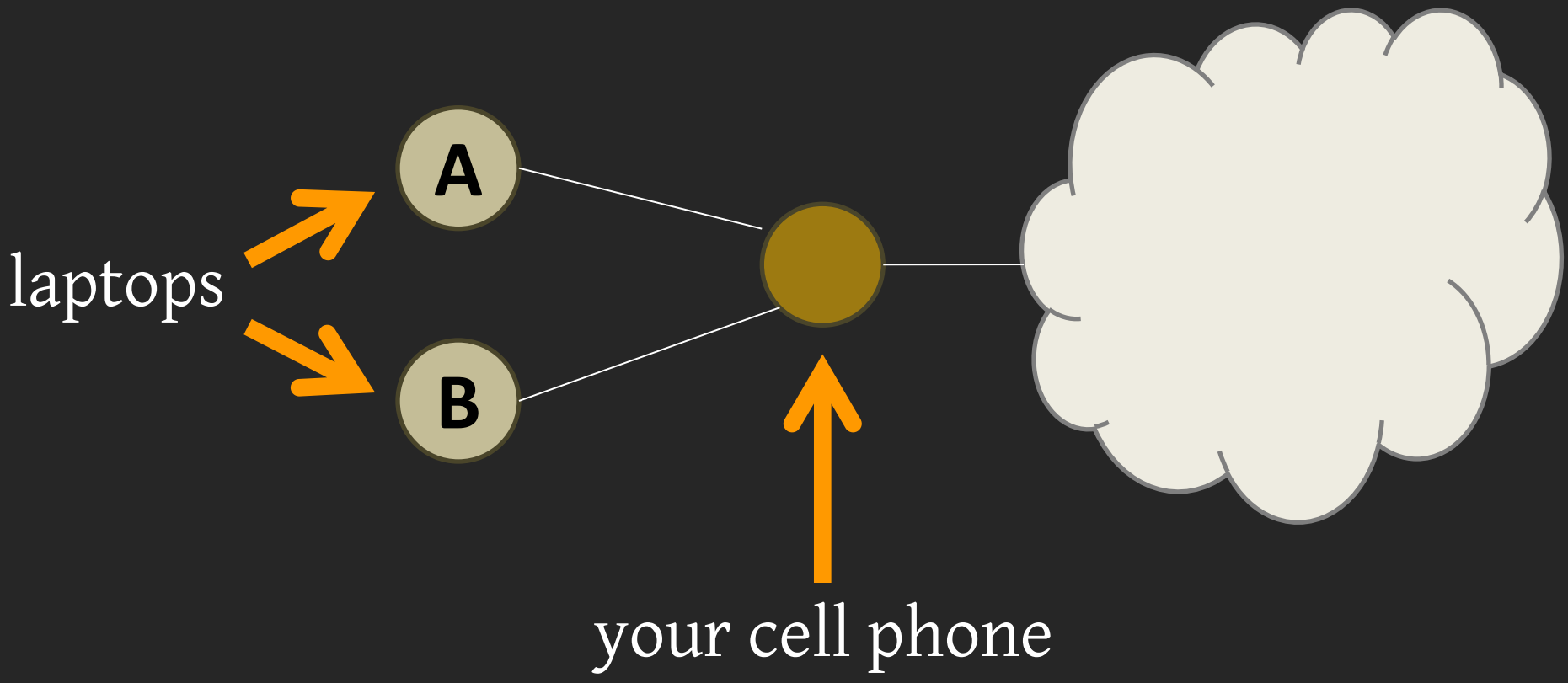
No single point where the Internet exists, and no single point where you can take the whole thing down for everyone in the world.

Weaknesses?

The only way that groups can realistically threaten the availability of the Internet is to attack

BOTTLENECKS.

A **bottleneck** is a circumstance where the performance of an entire system is restricted by a limited number of components or resources.


A
B
laptops
your cell phone

Which of the following are bottlenecks in the system to the right?

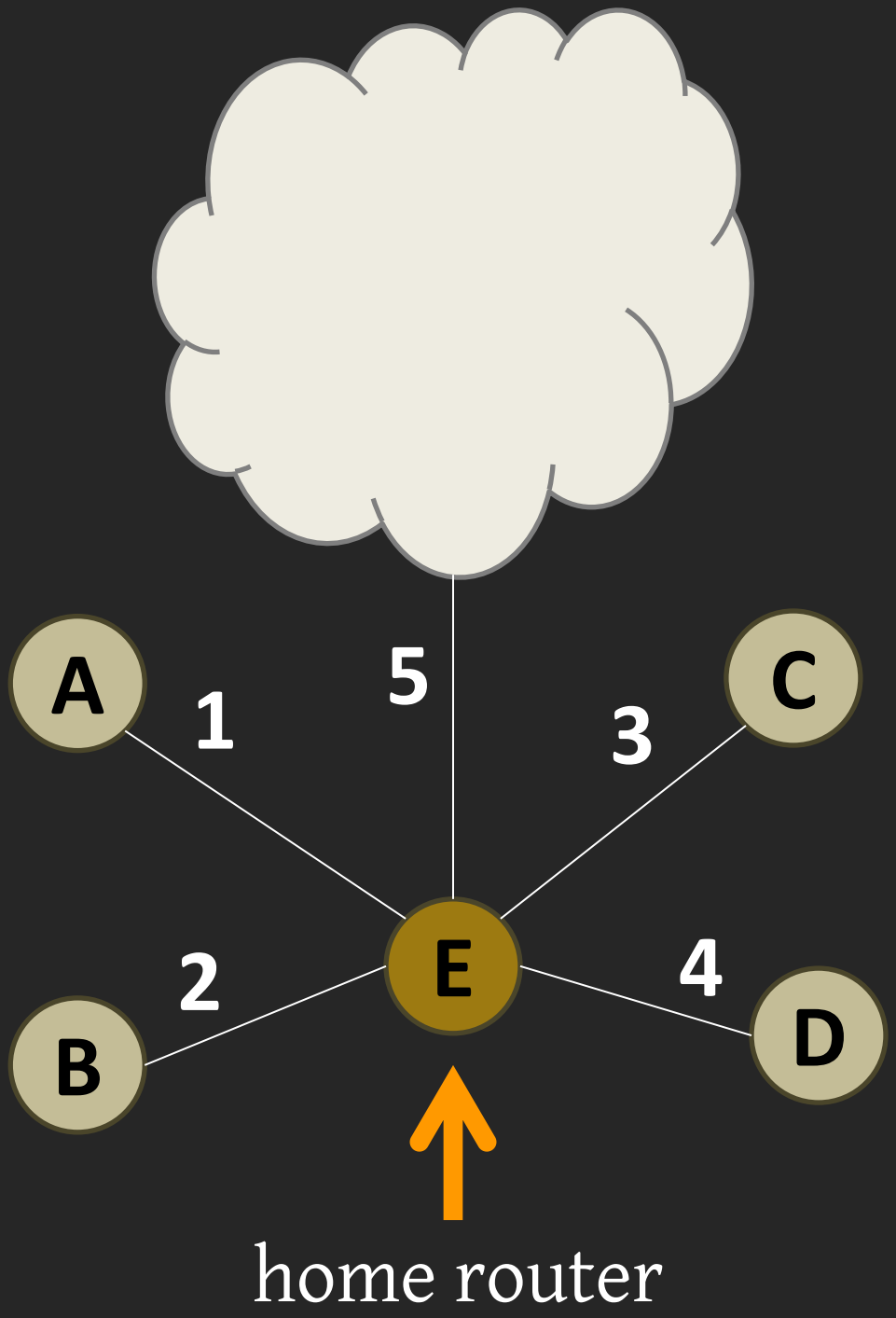
A) *there are no bottlenecks.*

B) *E is a bottleneck.*

C) *E and 5 are bottlenecks.*

D) *1, E, and 5 are bottlenecks.*

E) *Everything is a potential bottleneck!*



Networking: Under the Hood

Photo credit to Eric Kilby

Identifying Locations Online

We use IP (Internet Protocol) addresses in order to uniquely* reference machines online.

01110010.11010011.00101001.01001101

114.211.41.77

There are ~2^{32} of these available, and all addresses have now be allocated to regional owners.

* Not all IP addresses point to single systems now, partially because we've run out of addresses.

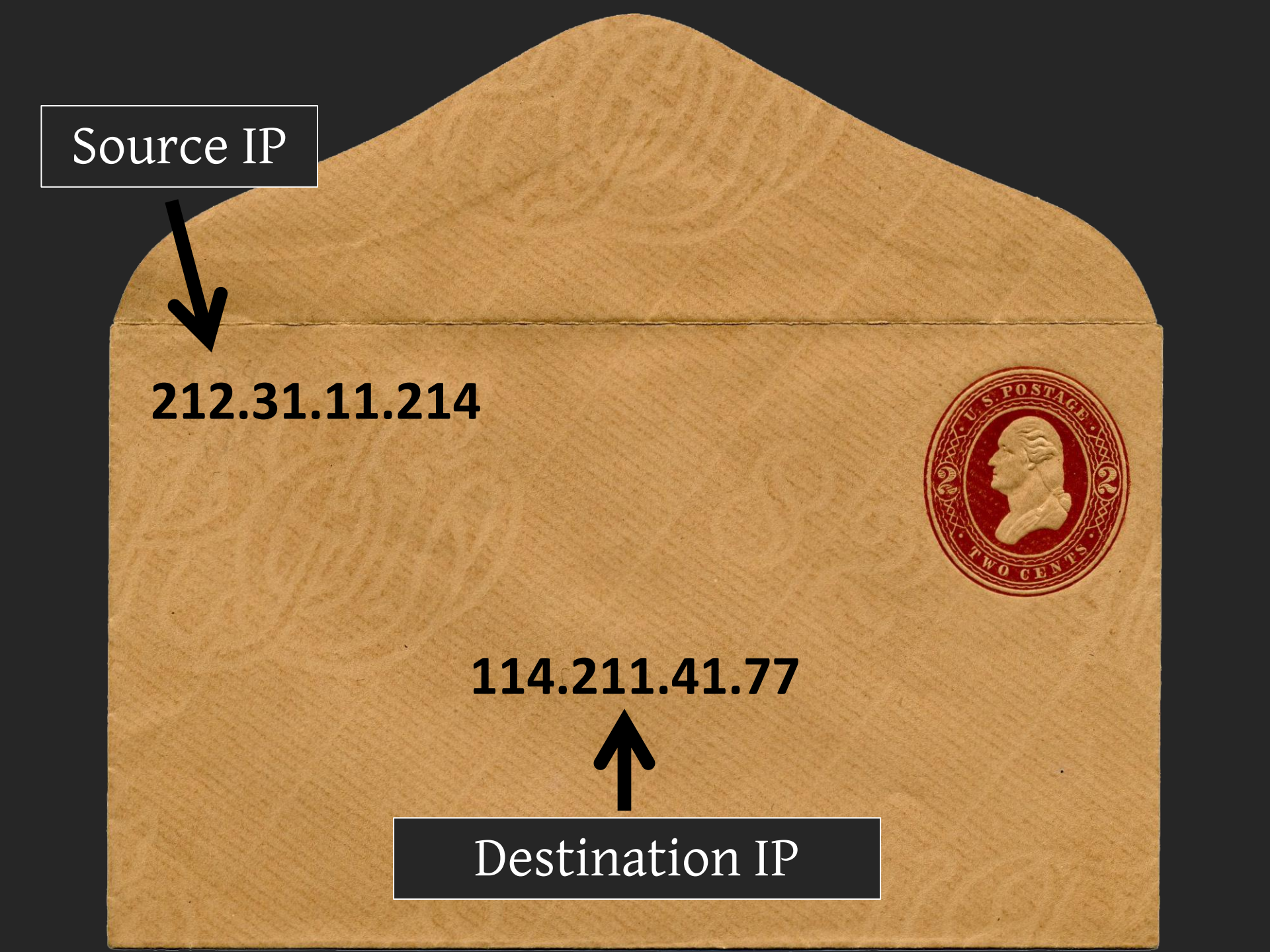
“Owning” IP addresses

In order to keep IP addresses unique, they are allocated to different organizations, who are the only ones that can use them publicly.

GE: 3.x.x.x

IBM: 9.x.x.x

USPS: 56.x.x.x


Source IP
212.31.11.214
U.S. POSTAGE
2
TWO CENTS
114.211.41.77
Destination IP

IP Everywhere

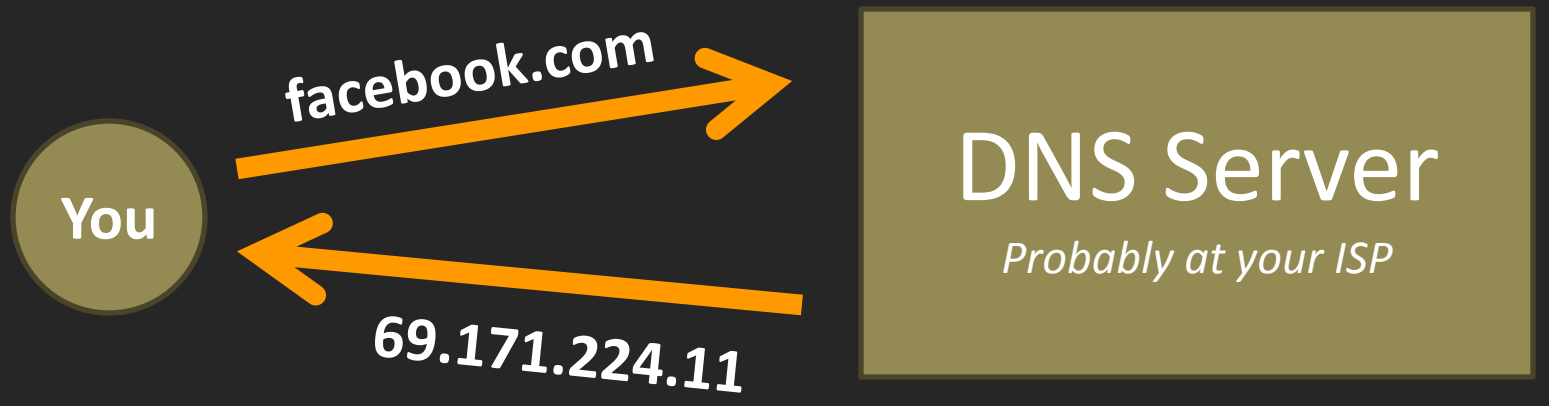
IP addresses exist to make it possible for machines to uniquely identify different machines in the vast expanses of the Internet.

Standardized so that any device can use them (mobile phones, for example) and communicate with the rest of the network.

Standards can also make it a bit hard to move away once it's widely accepted.

Domain Name System

Used for converting names that are useful for *humans* into names that are useful for *machines.*

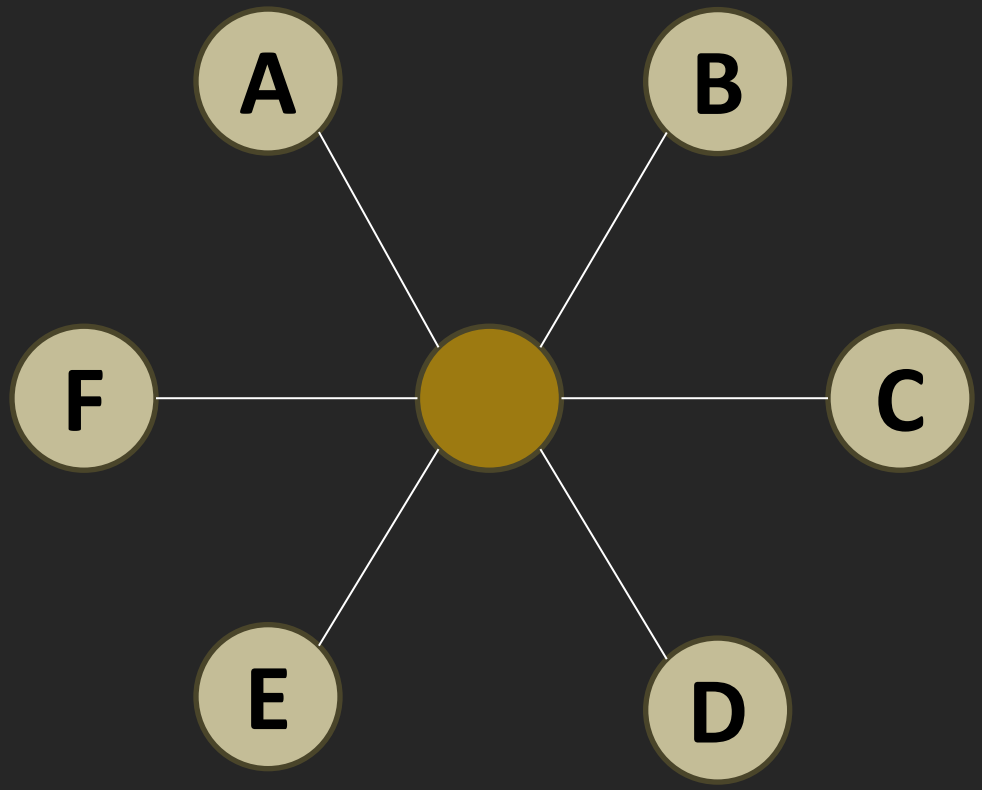


Finding bottlenecks

The core protocols for the internet were designed in relatively non-adversarial times, and as a result (in addition to a desire for simplicity and extensibility) ***the internet is strongly based on the idea of trust between nodes.***

Computers can send whatever information they want to over the network, including ***false source addresses*** (incorrectly identifying themselves).

Bottlenecks: Smurf attacks

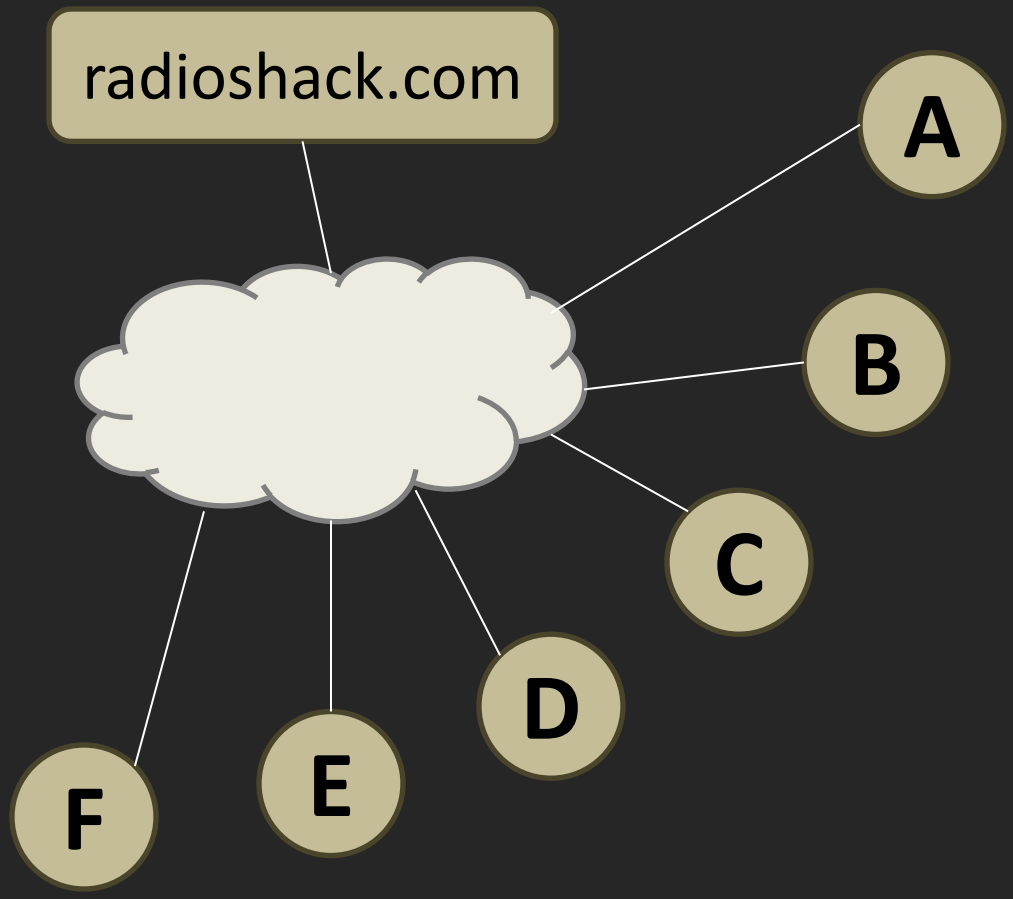


Step 1: Node A sends out a broadcast message, called a *ping*, to everyone nearby to see which computers are online.

Step 2: Node A sets the source address in it's request to the IP address of node C.

Step 3: Repeat millions of times per second.

Bottlenecks: Denial of Service attacks



Step 1: Node A (and 20,000 of his closest friends) send continuous web requests to radioshack.com.

Step 2: RadioShack's systems get overwhelmed and are unable to answer all requests.

Step 3: Legitimate users visit radioshack.com and get a message indicating that the site is not responding.

iClicker Question

Which of the following countries has *not* shut down its access to the Internet in the past five years?

A) Syria

B) China

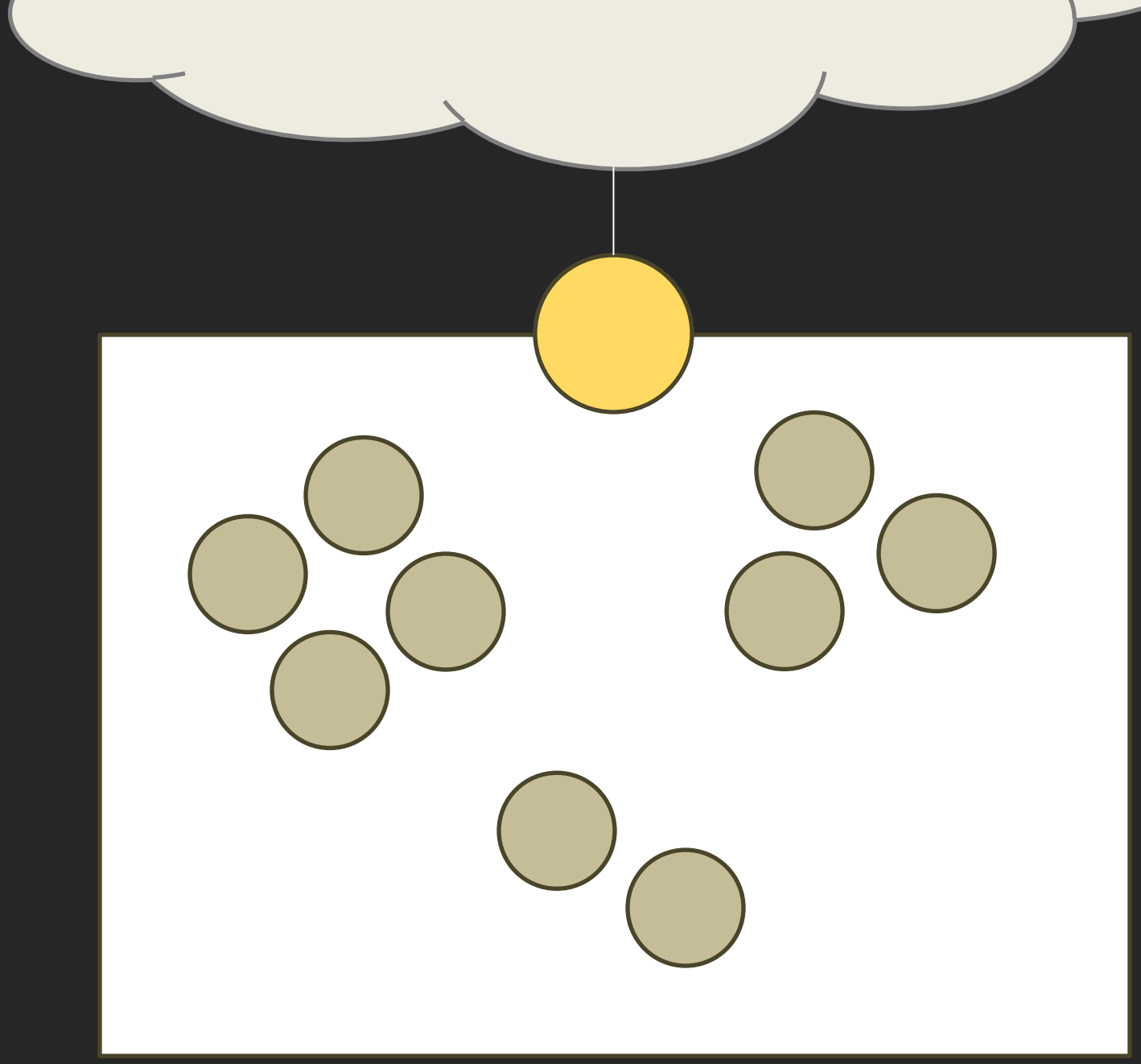
C) Egypt

D) Libya

Internet-scale Bottlenecks

We've seen some large-scale regional shutdowns recently (Egypt, Libya, Syria) and many people are aware of national establishments like the "Great Firewall" in China.

These techniques rely on artificially created bottlenecks, usually state-owned ISP's.



Internet-scale Bottlenecks

ISP's are likely still one of the most vulnerable bottlenecks in the US.

Many of them are (at least mostly) privately owned and its somewhat unclear whether the government could require a country-wide internet shutdown, both technically and legally.

Internet-scale Bottlenecks

DNS (Domain Name System) is another potential bottleneck for the world wide web. URL's would no longer work if DNS could be taken offline for week or so.

There have been several threats against the "root" DNS nodes recently, but most security researchers agree that there are too many resources available to these nodes to keep them offline for long.

The Internet: Summary

The Internet was designed to be very difficult to shut down and can cope with reasonably large amounts of network failure.

Taking down the whole Internet is essentially impossible because it is decentralized. The only way to take large parts of it offline is to target bottlenecks.

IP and other related protocols are responsible for a large part of the Internet's scalability and have made it easy for different governments, companies and individuals to innovate and connect to the Internet.