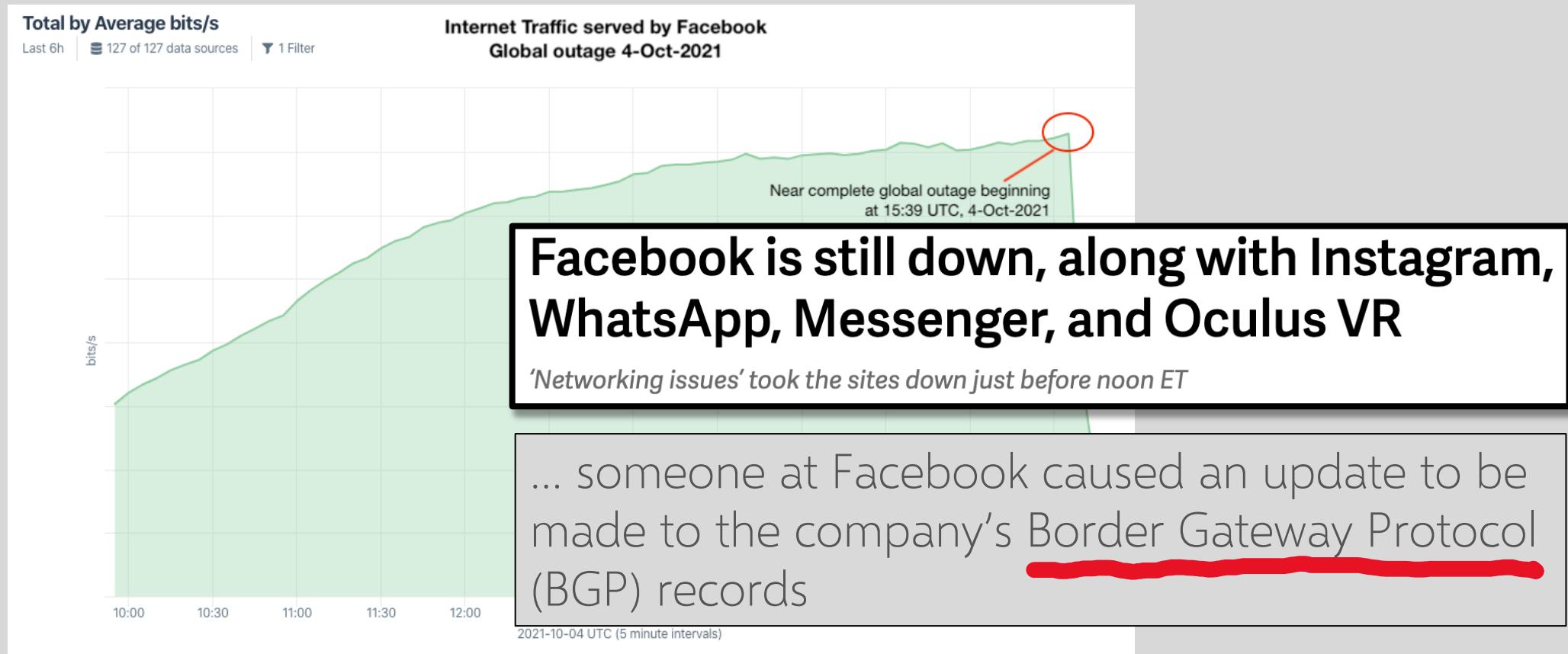




BGP SECURITY

VITALY SHMATIKOV

Facebook Outage in 2021



<https://krebsonsecurity.com/2021/10/what-happened-to-facebook-instagram-whatsapp/>

IP Routing

Routing of IP packets is based on IP addresses

- 32-bit host identifiers (128-bit in IPv6)

Routers use a forwarding table

- Entry = destination, next hop, network interface, metric
- Table look-up for each packet to decide how to route it

Routers learn routes to hosts and networks via routing protocols

- Host is identified by IP address, network by IP prefix

BGP (Border Gateway Protocol) is the core Internet protocol for establishing inter-AS routes

Distance-Vector Routing

Each node keeps vector with distances to all nodes

Periodically sends distance vector to all neighbors

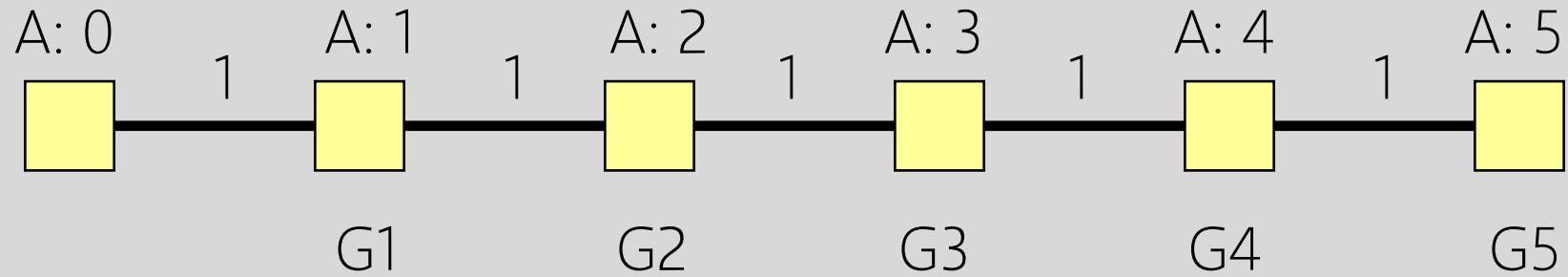
Neighbors send their distance vectors, too; node updates its vector based on received information

- Bellman-Ford algorithm: for each destination, router picks the neighbor advertising the cheapest route, adds his entry into its own routing table and re-advertises
- Used in RIP (routing information protocol)

Split-horizon update

- Do not advertise a route on an interface from which you learned the route in the first place!

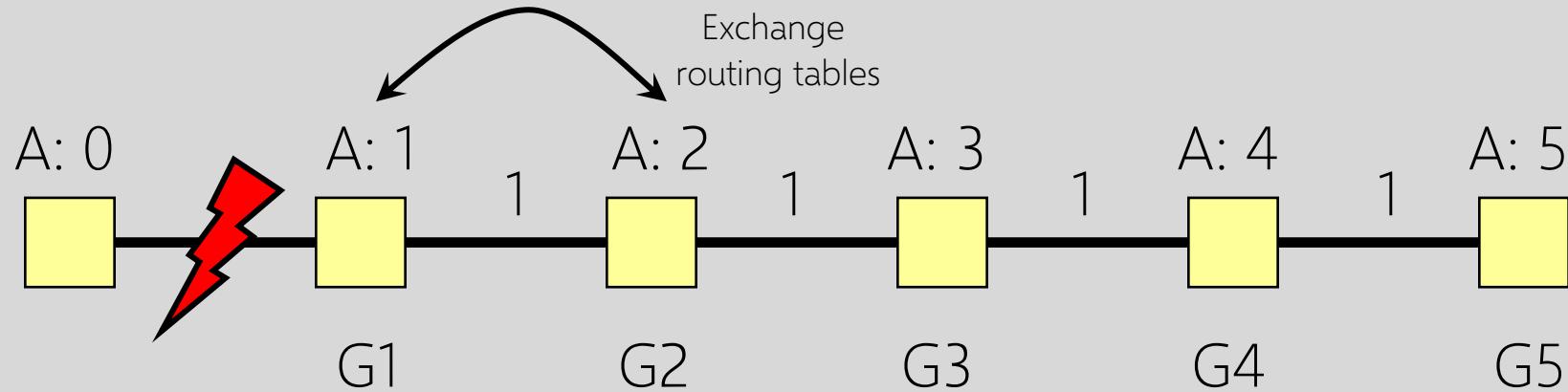
Good News Travels Fast



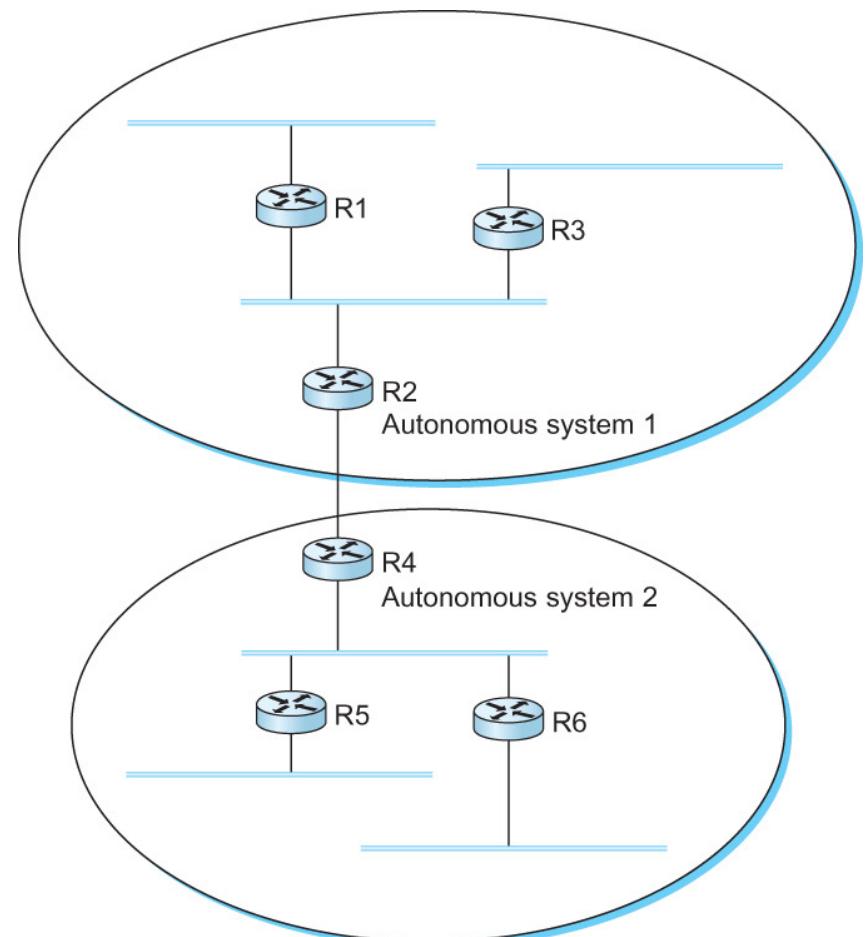
- G1 advertises route to network A with distance 1
- G2-G5 quickly learn the good news and install the routes to A via G1 in their local routing tables

Split-horizon update only
prevent two-node loops

Bad News Travels Slowly



- G1's link to A goes down
- G2 is advertising a pretty good route to G1 (cost=2)
- G1's packets to A are forever looping between G2 and G1
- G1 is now advertising a route to A with cost=3, so G2 updates its own route to A via G1 to have cost=4, and so on... slowly counting to infinity



Interdomain Routing

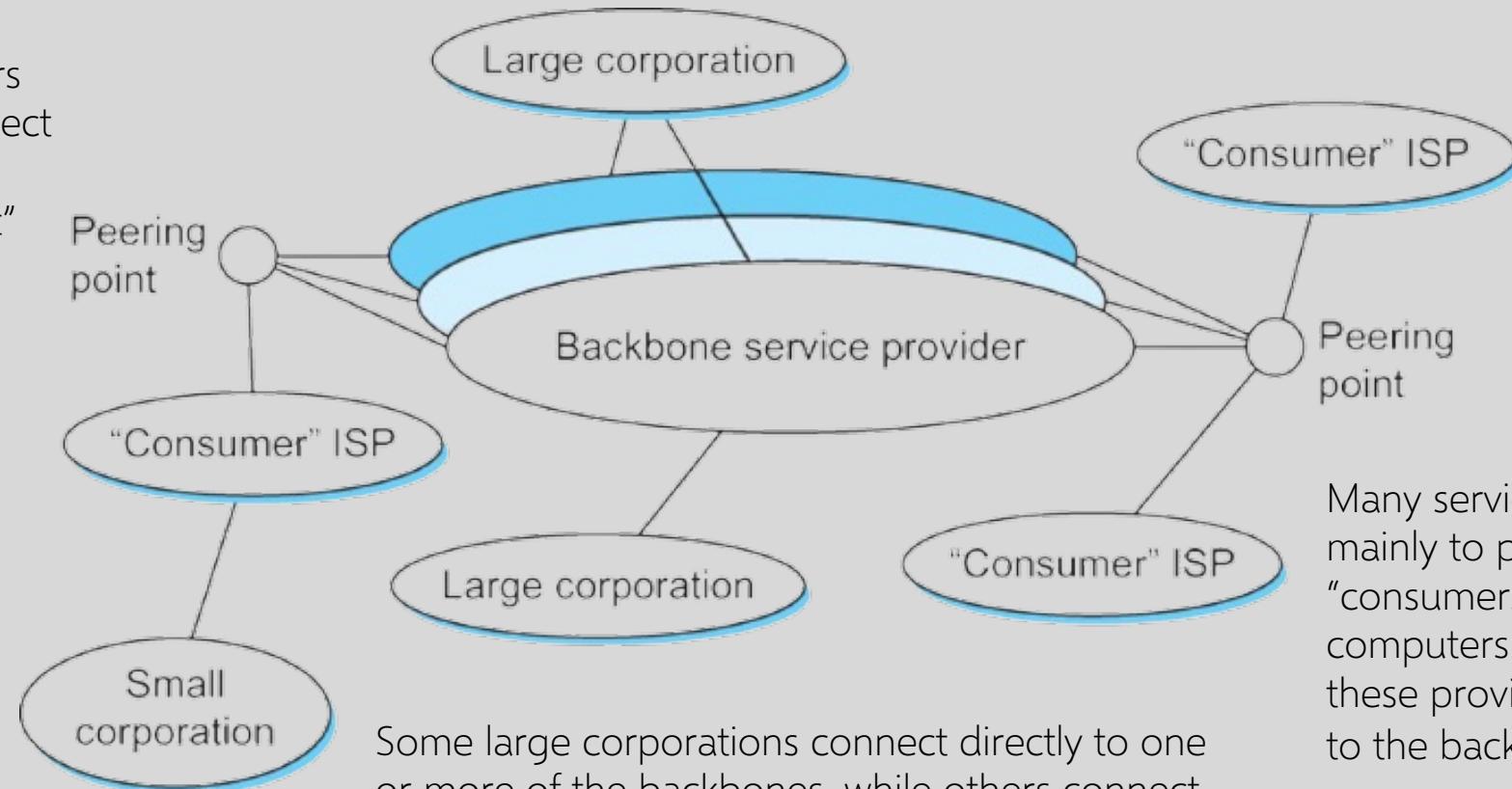
Internet is organized as autonomous systems (AS), each under the control of a single administrative entity

Autonomous System (AS) corresponds to an administrative domain

- Examples: university network, corporate internal network, backbone network, network of a single Internet service provider

Simple Model of Global Internet

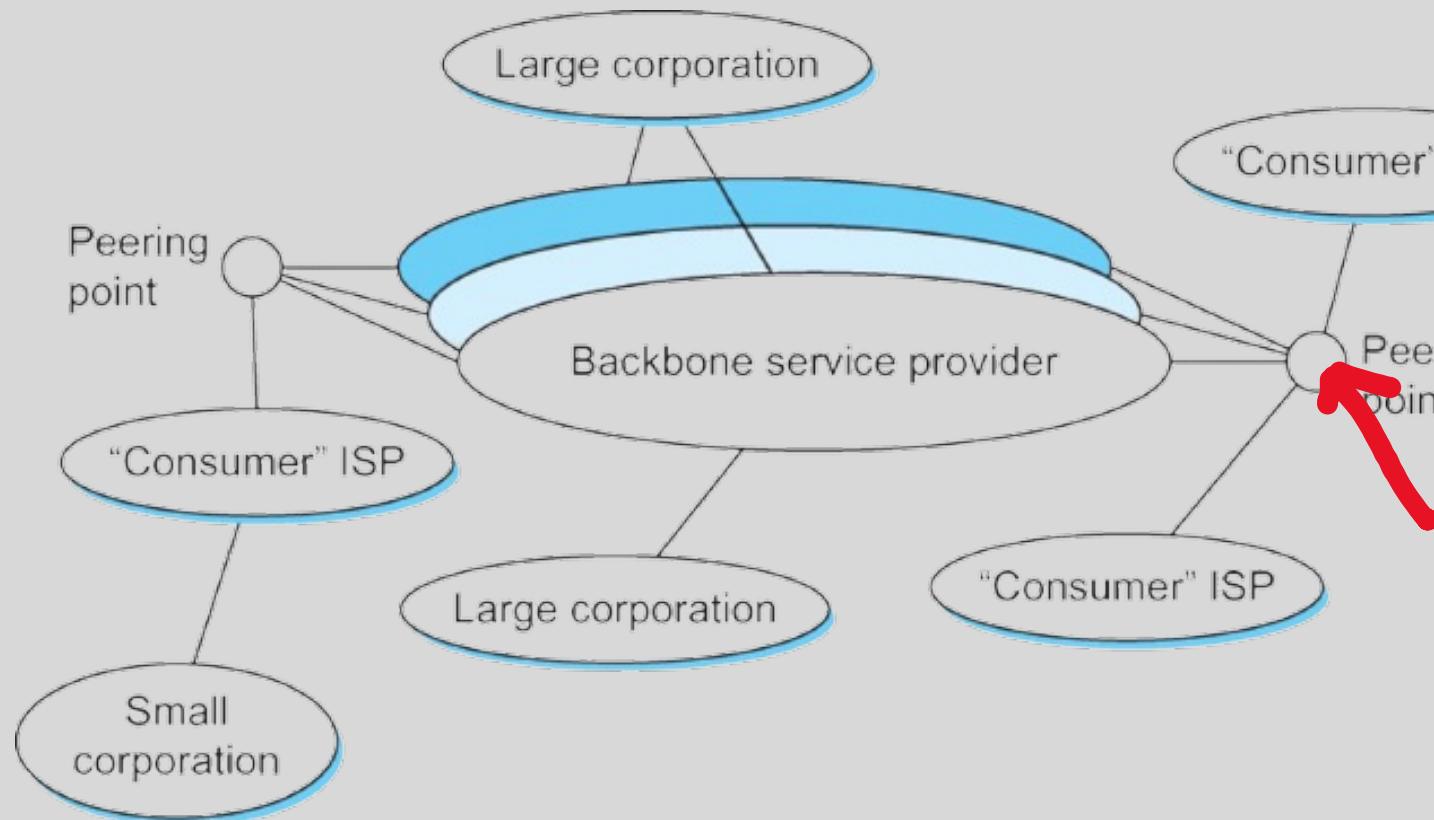
Often many providers arrange to interconnect with each other at a single “peering point”



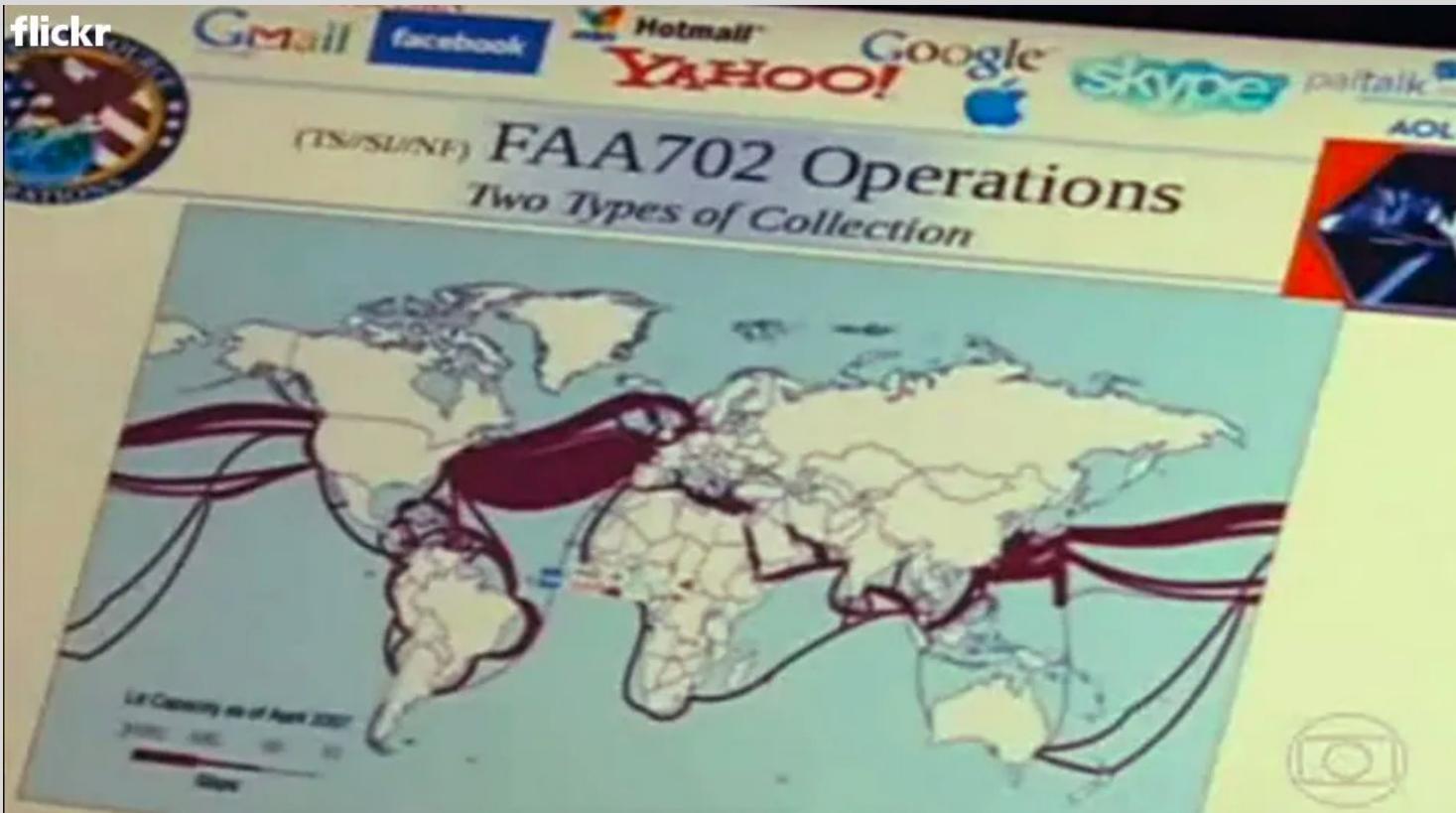
Some large corporations connect directly to one or more of the backbones, while others connect to smaller, non-backbone service providers

Many service providers exist mainly to provide service to “consumers” (individuals with computers and devices), and these providers must connect to the backbone providers

A Little Side Note About Peering Points



ATT peering facility
811 10th Avenue



<https://www.dailydot.com/debug/nsa-spy-prgrams-prism-fairview-blarney/>

TOP SECRET//SI//ORCON//NOFORN



Gmail
facebook

Hotmail

YAHOO!

Google

skype

paltalk.com
YouTube

AOL mail

(TS//SI//NF) FAA702 Operations *Two Types of Collection*



Upstream

- Collection of communications on fiber cables and infrastructure as data flows past.
(FAIRVIEW, STORMBREW, BLARNEY, OAKSTAR)

You
Should
Use Both

PRISM

- Collection directly from the servers of these U.S. Service Providers: Microsoft, Yahoo, Google Facebook, PalTalk, AOL, Skype, YouTube Apple.

TOP SECRET//SI//ORCON//NOFORN



<https://www.dailydot.com/debug/nsa-spy-prgrams-prism-fairview-blarney/>

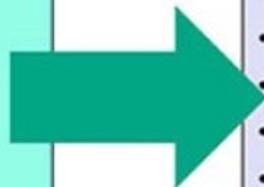
TOP SECRET//SI//ORCON//NOFORN



(TS//SI//NF) PRISM Collection Details

Current Providers

- Microsoft (Hotmail, etc.)
- Google
- Yahoo!
- Facebook
- PalTalk
- YouTube
- Skype
- AOL
- Apple



What Will You Receive in Collection
(Surveillance and Stored Comms)?
It varies by provider. In general:

- E-mail
- Chat – video, voice
- Videos
- Photos
- Stored data
- VoIP
- File transfers
- Video Conferencing
- Notifications of target activity – logins, etc.
- Online Social Networking details
- **Special Requests**

Complete list and details on PRISM web page:

Go PRISMFAA

TOP SECRET//SI//ORCON//NOFORN



<https://www.dailydot.com/debug/nsa-spy-prgrams-prism-fairview-blarney/>

Goals of BGP

The goal of Inter-domain routing
is to find any path to the intended
destination that is **loop-free**

Reachability rather than optimality!

- Finding a path anywhere close to optimal
is considered a great achievement (why?)

BGP Challenges

Scalability: a backbone router must be able to forward any packet to anywhere in the Internet

- Routing table must have a match for any valid IP address

Full autonomy of the domains

- Impossible to calculate meaningful cost for a path that crosses multiple ASs
- A cost of 1000 might imply a great path from one provider and unacceptably bad from another provider

Providers may not trust each other

- Providers may not believe each other's route advertisements

Overview of BGP



Yakov Rekhter

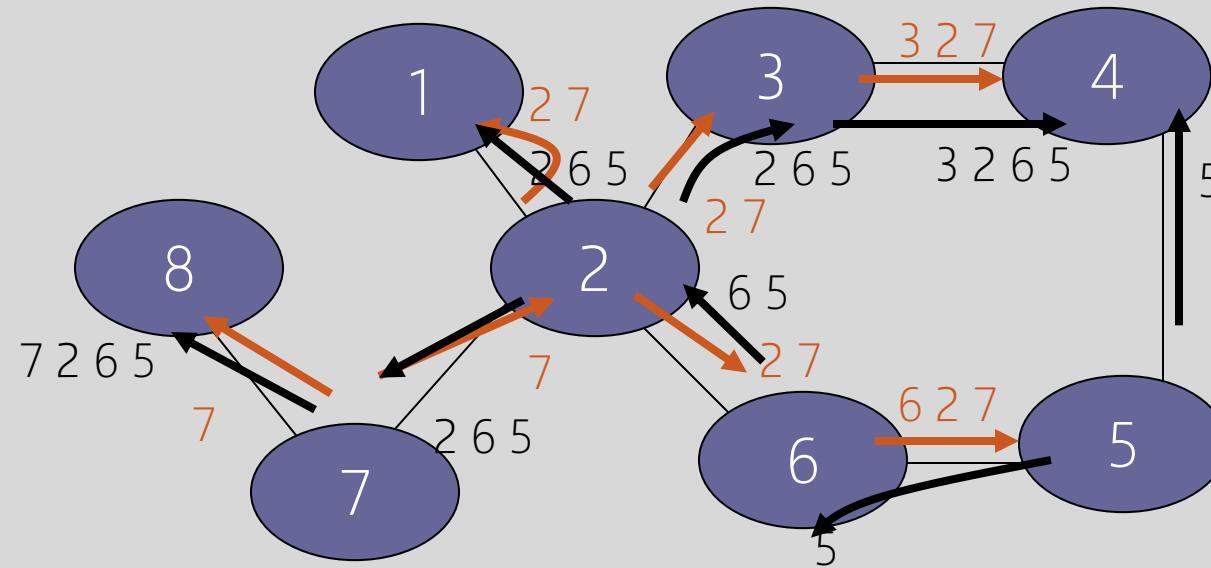
BGP is a **path-vector** protocol between ASes

- Just like distance-vector, but routing updates contain an actual path to destination node, ie, the list of traversed ASes and the set of network prefixes belonging to the first AS on the list

Each BGP router receives update messages from neighbors, selects one “best” path for each prefix, and advertises this path to its neighbors

- Can be the shortest path, but doesn’t have to be (“hot potato” vs. “cold potato”)
- Always route to the **most specific prefix** for a destination

BGP Example



AS 2 provides **transit** for AS 7

- Traffic to and from AS 7 travels through AS 2

Some (Old) BGP Statistics

BGP routing tables contain about 125,000 address prefixes mapping to about 17-18,000 paths

Approx. 10,000 BGP routers

Approx. 2,000 organizations own AS

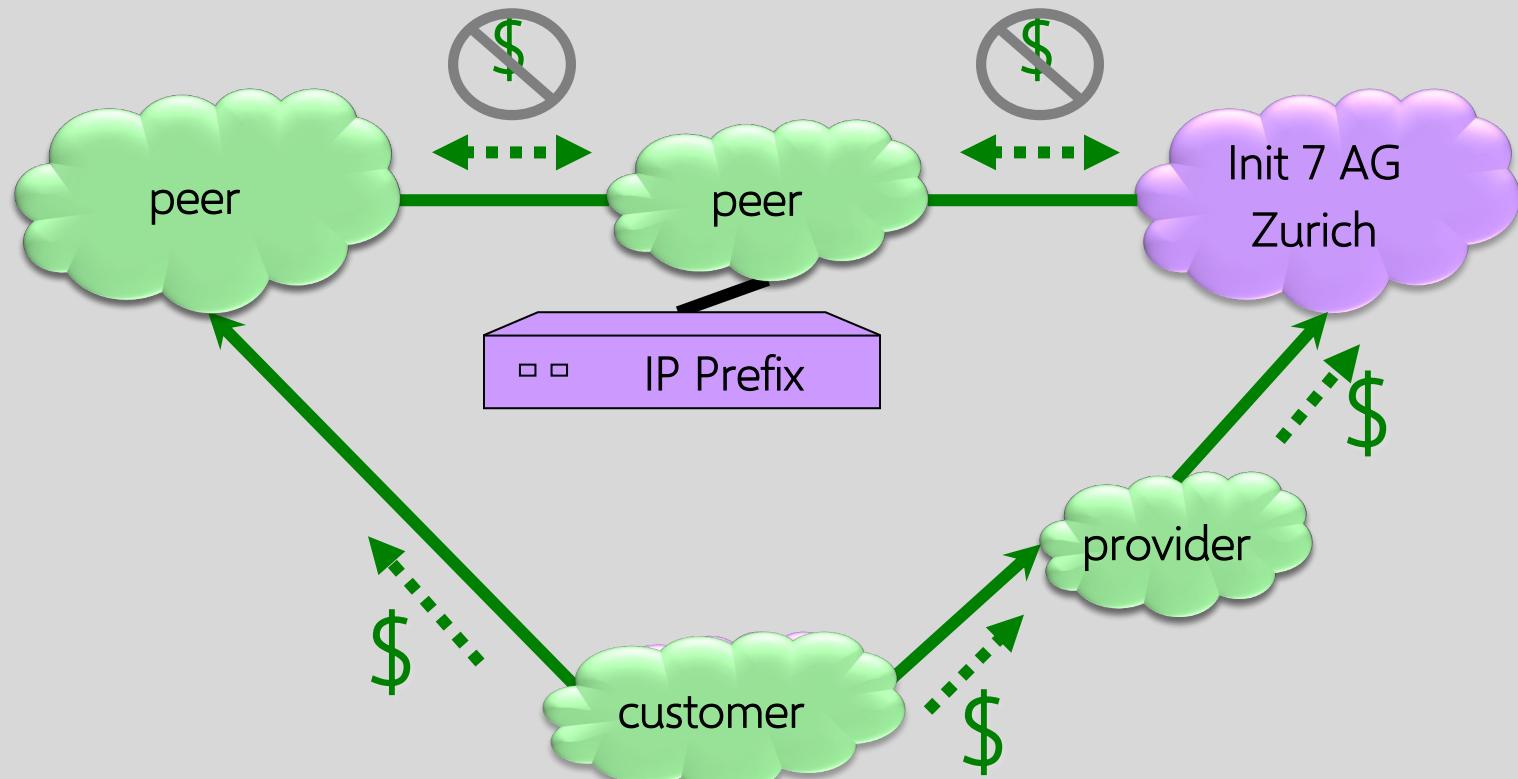
Approx. 6,000 organizations own prefixes

Average route length is about 3.7

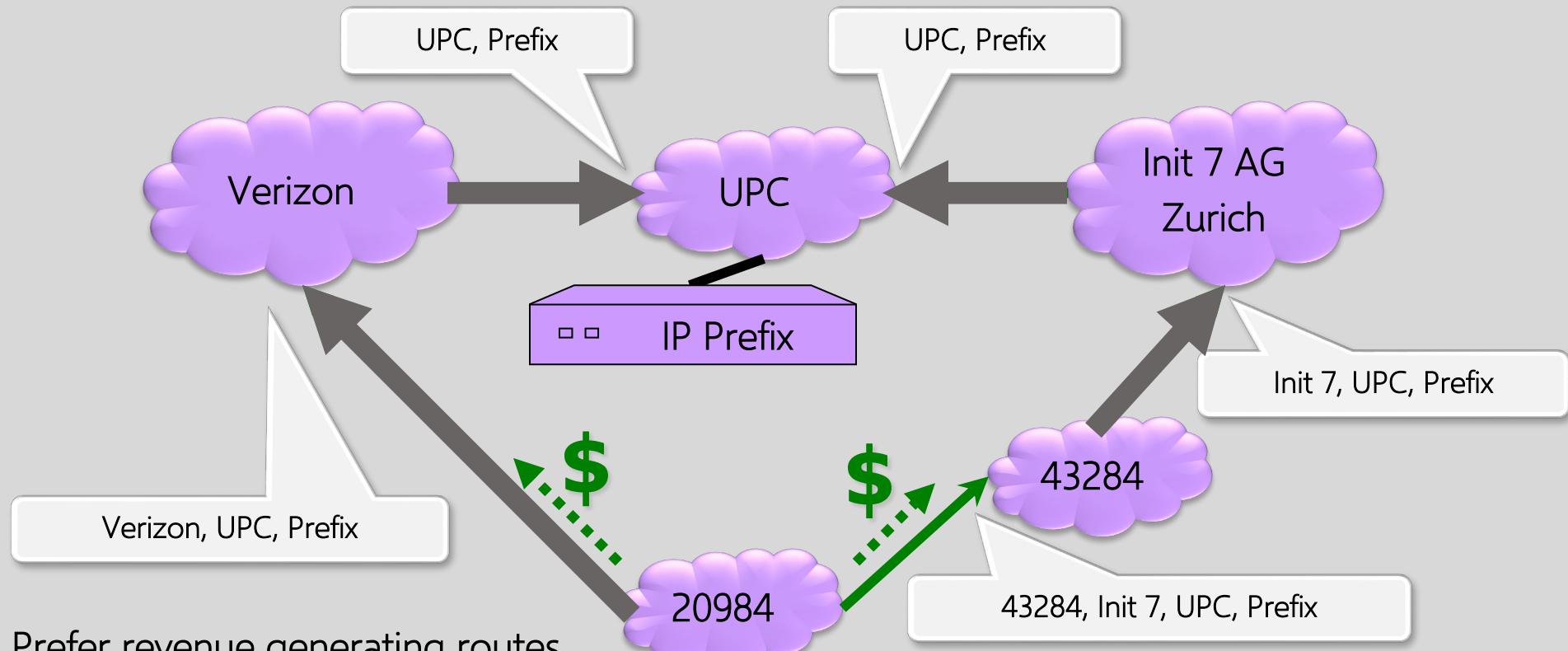
50% of routes have length less than 4 ASes

95% of routes have length less than 5 ASes

BGP Illustration

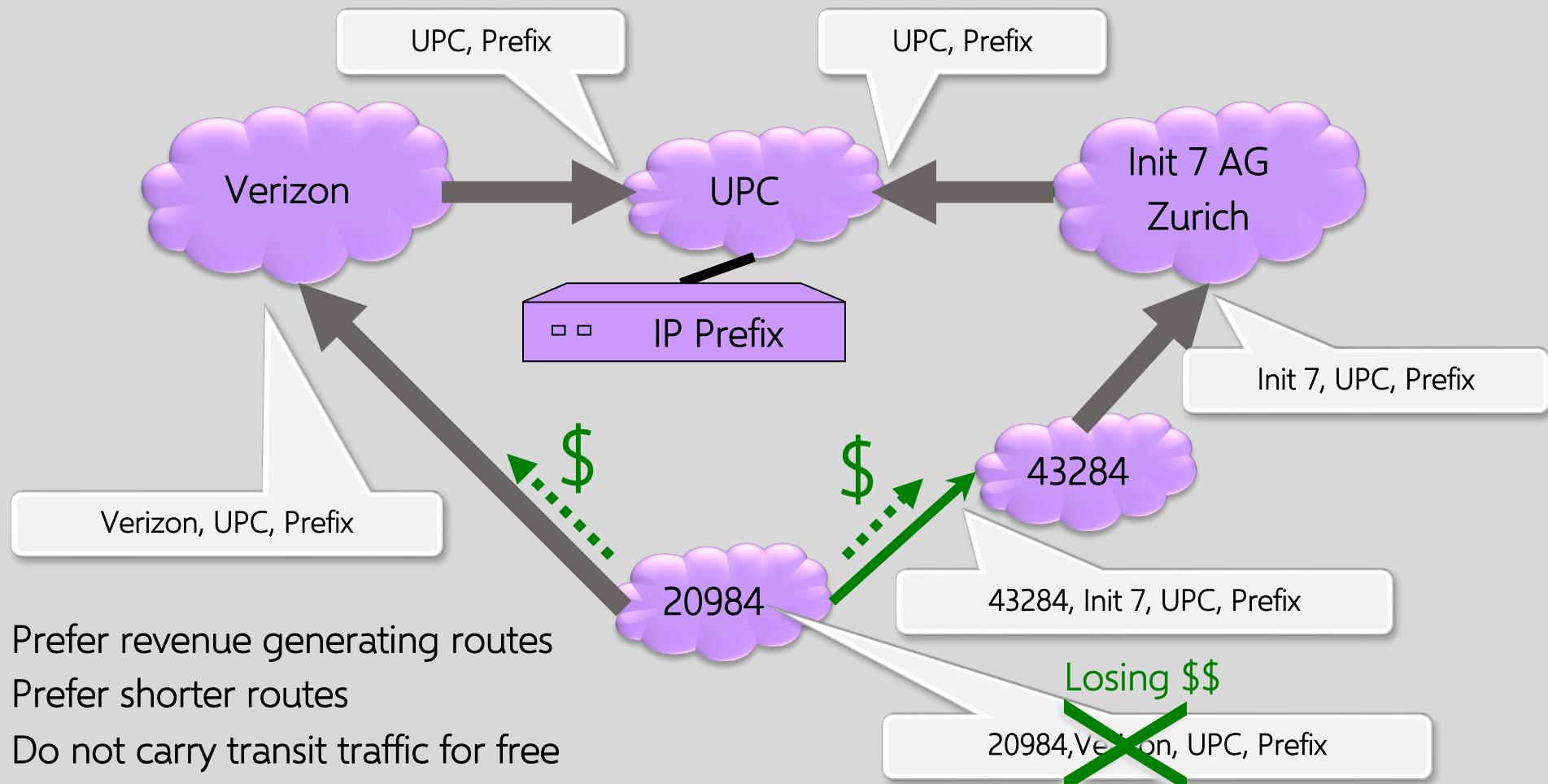


Routing with BGP



- 1) Prefer revenue generating routes
- 2) Prefer shorter routes

Routing with BGP



BGP Issues

Convergence problems

- Protocol allows policy flexibility
- Some legal policies prevent convergence
- Even shortest-path policy converges slowly

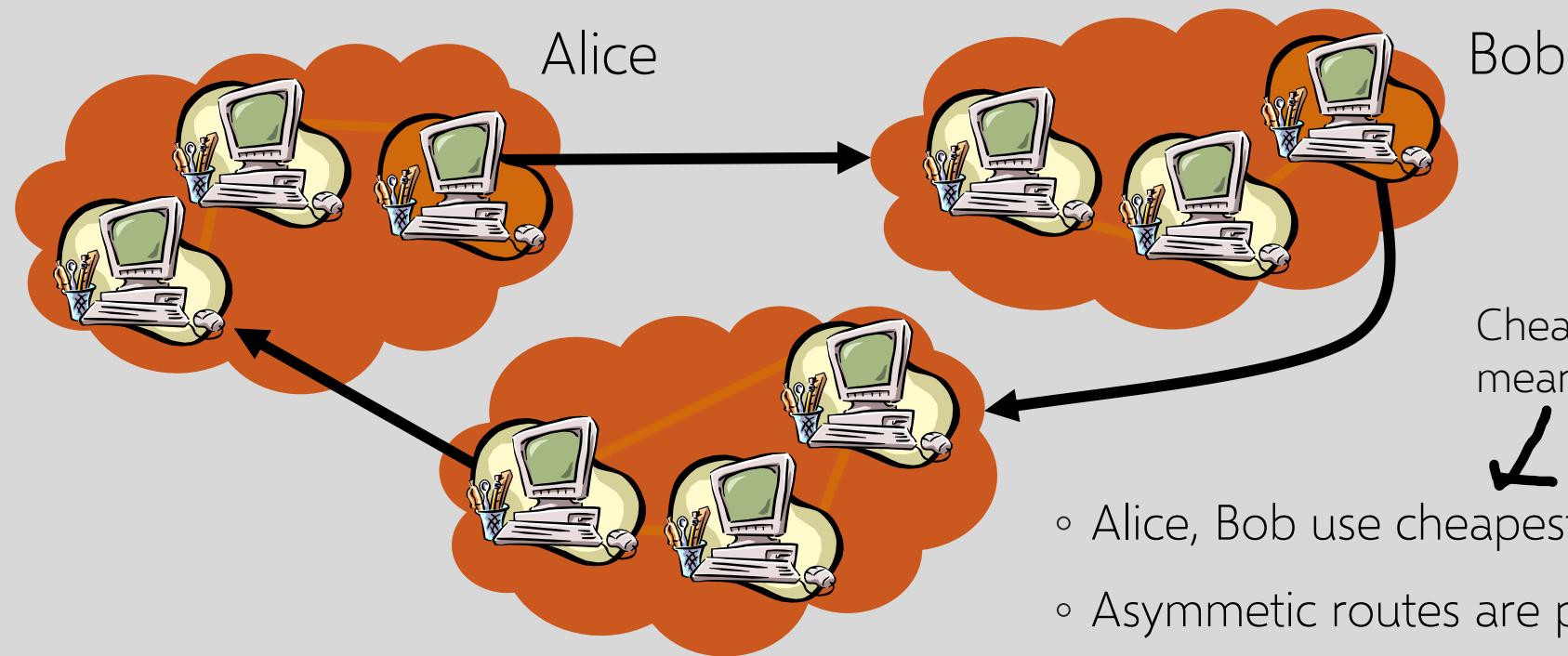
Incentive for dishonesty

- ISP pays for some routes, others free

Security problems

- Potential for disruptive attacks

Evidence: Asymmetric Routes



Cheapest does not always mean the shortest!

- Alice, Bob use cheapest routes to each other
- Asymmetric routes are prevalent
 - AS asymmetry in 30% of measured routes
 - Finer-grained asymmetry far more prevalent

BGP Misconfiguration (Route Leak)

Domain advertises good routes to addresses it does not know how to reach

- Result: packets go into a network “black hole”

April 25, 1997: AS7007 (Florida Internet Exchange) de-aggregated the BGP route table and re-advertised all prefixes as if it originated paths to them

- In effect, AS7007 was advertising that it has the **best route to every host on the Internet**
- Huge network instability as incorrect routing data propagated and routers crashed under traffic

BGP Incident:

Summary: Routing Leak briefly takes down Google

Start: 2015-03-12 08:58:00

End: Unknown

Details: This morning, users of Google around the world were unable to access many of the company's services due to a routing leak in India. Beginning at 08:58 UTC Indian broadband provider Hathway (AS17488) incorrectly announced over 300 Google prefixes to its Indian transit provider Bharti Airtel (AS9498). Bharti in turn announced these routes to the rest of the world, and a number of ISPs accepted these routes including US carriers Cogent (AS174), Level 3 (AS3549) as well as overseas incumbent carriers Orange (France Telecom, AS5511), Singapore Telecom (Singtel, AS7473) and Pakistan Telecom (PTCL, AS17557). Like many



<http://www.securerouting.net/incident/20/>

<https://dyn.com/blog/routing-leak-briefly-takes-google/>

BGP Has No Security Model

BGP update messages contain no authentication or integrity protection

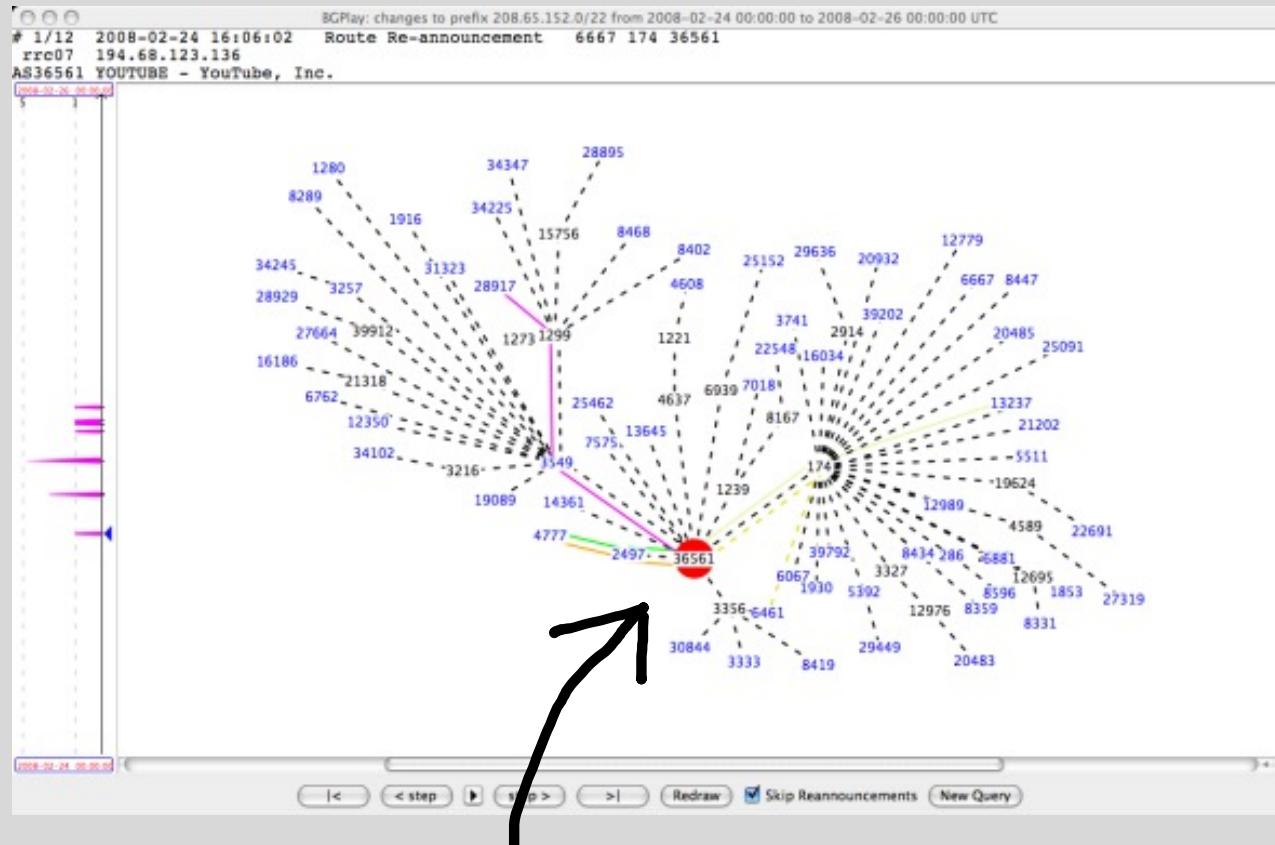
Anyone can advertise any route

- Change the AS path
 - Either attract traffic to attacker's AS, or divert traffic away
 - Interesting economic incentive: an ISP wants to dump its traffic on other ISPs without routing their traffic in exchange
- Hijack IP addresses
 - AS announces it originates a prefix it shouldn't
 - AS announces it has shorter path to a prefix
 - AS announces more specific prefix

In-Class Exercise

What attacks / problems could result from BGP insecurity?

YouTube (Normally)



AS36561 (YouTube) advertises 208.65.152.0/22

February 24, 2008



Corrigendum- Most Urgent

**GOVERNMENT OF PAKISTAN
PAKISTAN TELECOMMUNICATION AUTHORITY
ZONAL OFFICE PESHAWAR**

Plot-11, Sector A-3, Phase-V, Hayatabad, Peshawar.
Ph: 091-9217279-5829177 Fax: 091-9217254

www.pta.gov.pk

NWFP-33-16 (BW)/06/PTA

February 2008

Subject: Blocking of Offensive Website

Reference: This office letter of even number dated 22/02/2008

I am directed to request all ISPs to immediately block access to the following website:

URL: <http://www.youtube.com/watch?v=o3s8jtvvg00>

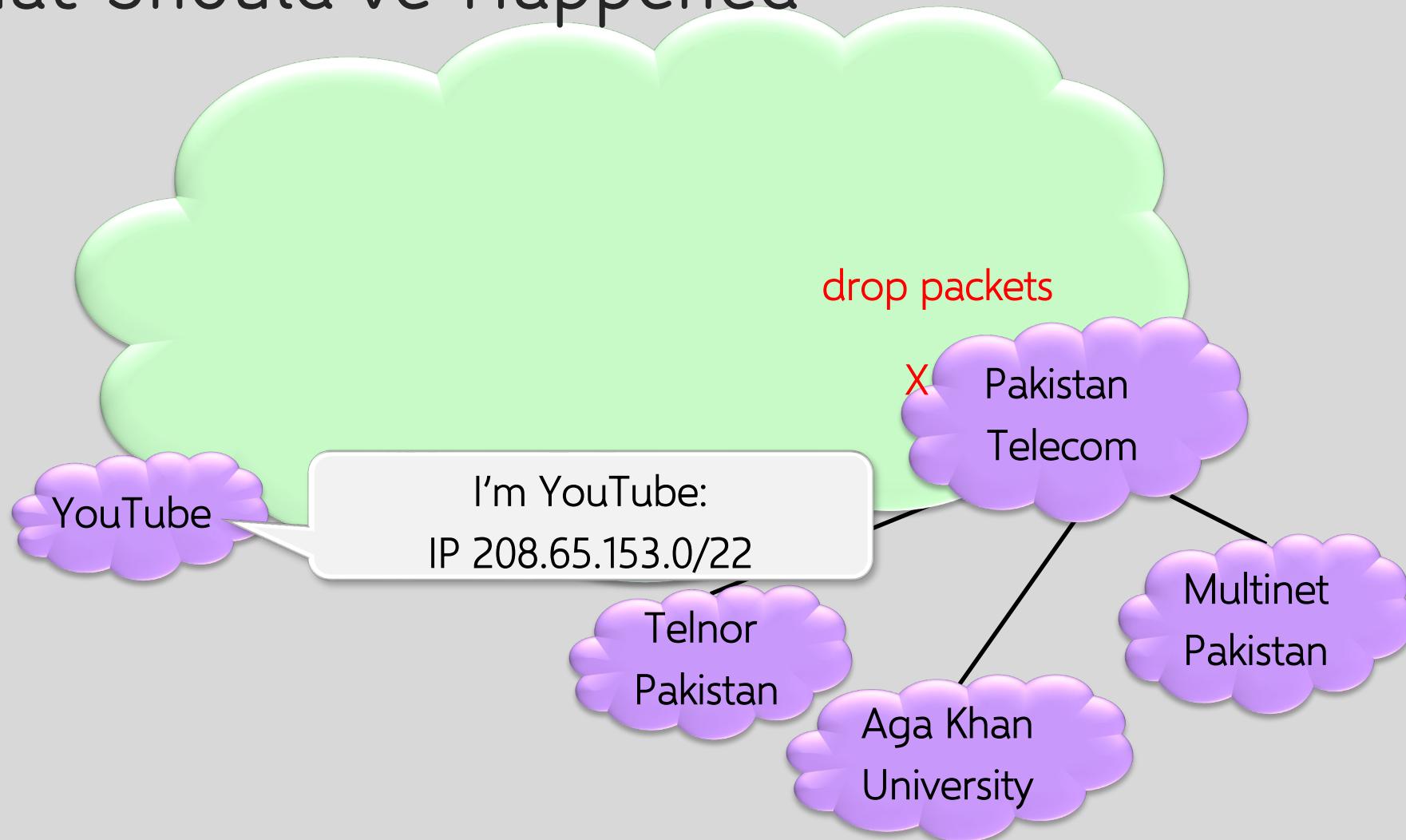
IPs: 208.65.153.238 208.65.153.253 208.65.153.251

Compliance report should reach this office through return fax or at email peshawar@pta.gov.pk today please.



Pakistan government wants to block YouTube

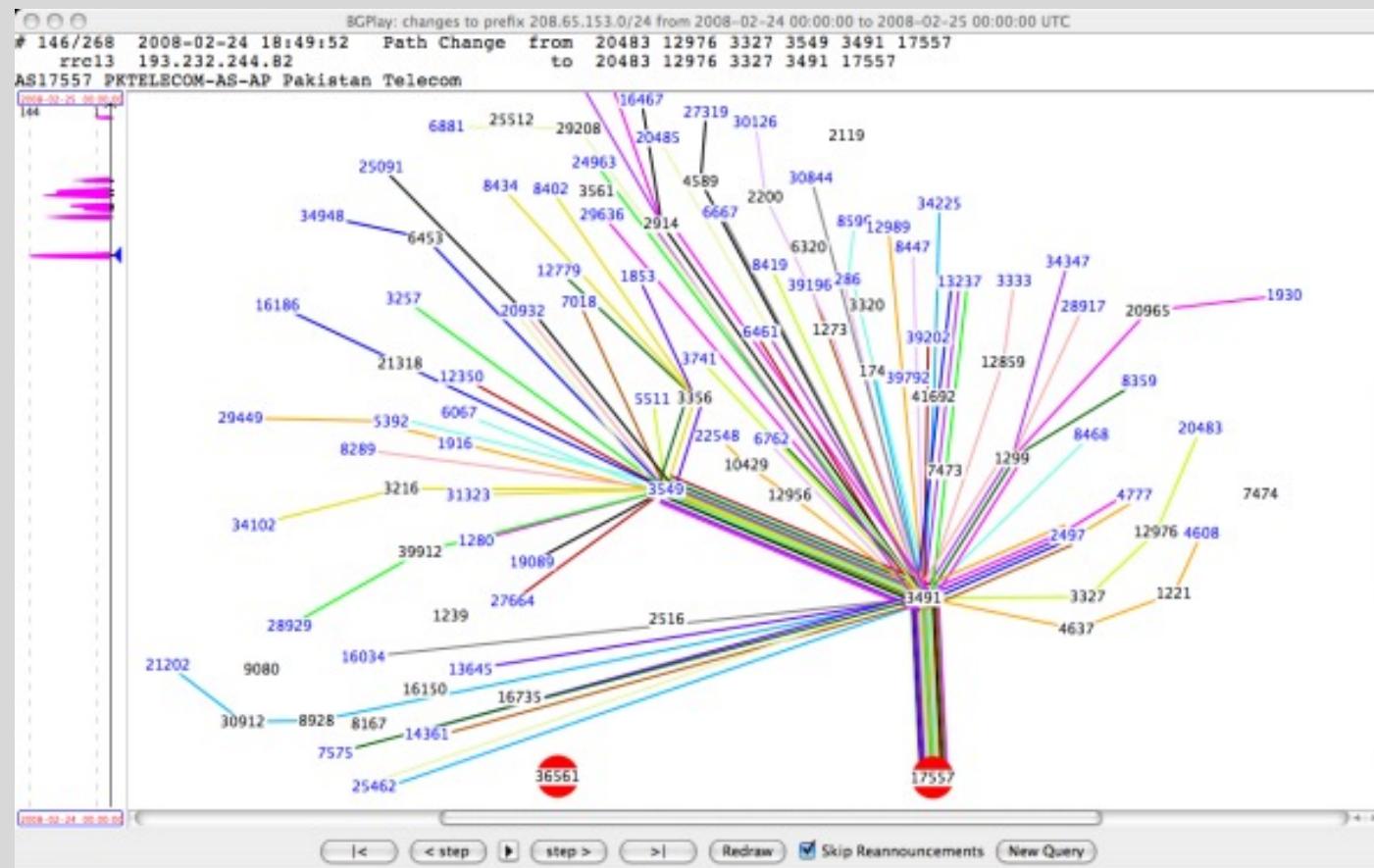
What Should've Happened



What Did Happen



Two-Hour YouTube Outage

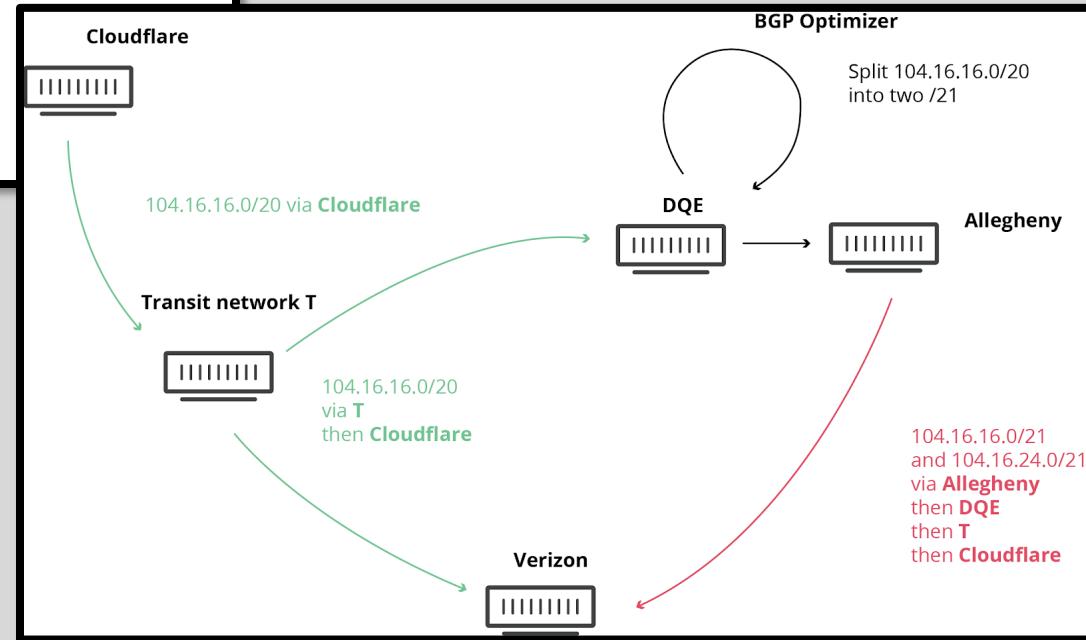


How A Small ISP in Pennsylvania Tanked a Big Chunk of the Web on Monday

June 2019

And how Verizon apparently made it much worse.

- DQE Communications advertised (“originated”) route to 2400 networks
 - Tiny ISP serving 2000 buildings in Pittsburgh
- Verizon propagated these routes
- Result: multi-hour outage at Cloudflare (approx. 1.5 million websites), AWS, Reddit, Overcast, Discord, Twitch...



This time, the damage was seen worldwide. What exacerbated the problem today was the involvement of a “BGP Optimizer” product from [Noction](#). This product has a feature that splits up received IP prefixes into smaller, contributing parts (called more-specifics)

For example, our own IPv4 route 104.20.0.0/20 was turned into 104.20.0.0/21 and

<https://slate.com/technology/2019/06/verizon-dqe-outage-internet-cloudflare-reddit-aws.html>

<https://blog.cloudflare.com/how-verizon-and-a-bgp-optimizer-knocked-large-parts-of-the-internet-offline-today/>

<https://arstechnica.com/information-technology/2018/11/strange-snafu-misroutes-domestic-us-internet-traffic-through-china-telecom/>

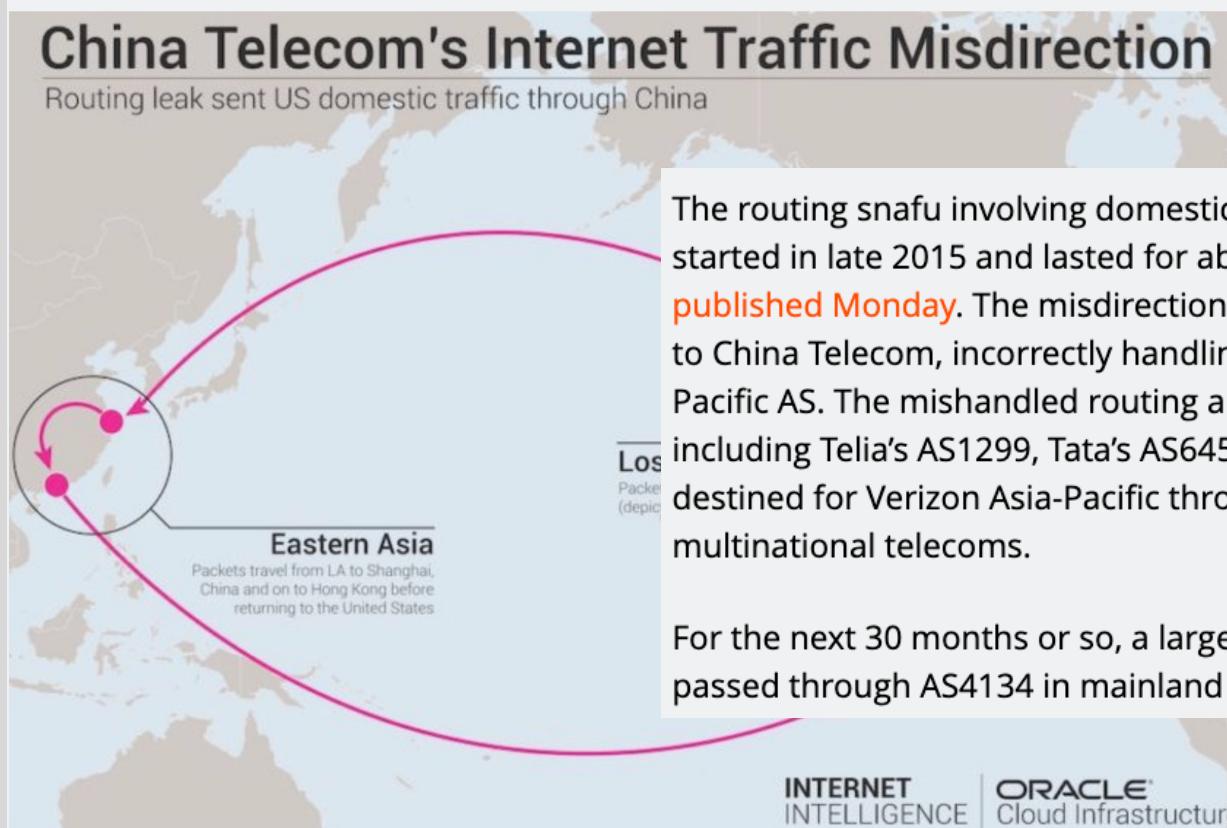
China Telecom Incidents

Telecom with ties to China's government misdirected traffic for two and a half years.

DAN GOODIN - 11/6/2018, 9:05 AM

China Telecom's Internet Traffic Misdirection

Routing leak sent US domestic traffic through China



The routing snafu involving domestic US Internet traffic coincided with a larger misdirection that started in late 2015 and lasted for about two and a half years, Madory said in a [blog post published Monday](#). The misdirection was the result of AS4134, the autonomous system belonging to China Telecom, incorrectly handling the routing announcements of AS703, Verizon's Asia-Pacific AS. The mishandled routing announcements caused several international carriers—including Telia's AS1299, Tata's AS6453, GTT's AS3257, and Vodafone's AS1273—to send data destined for Verizon Asia-Pacific through China Telecom, rather than using the normal multinational telecoms.

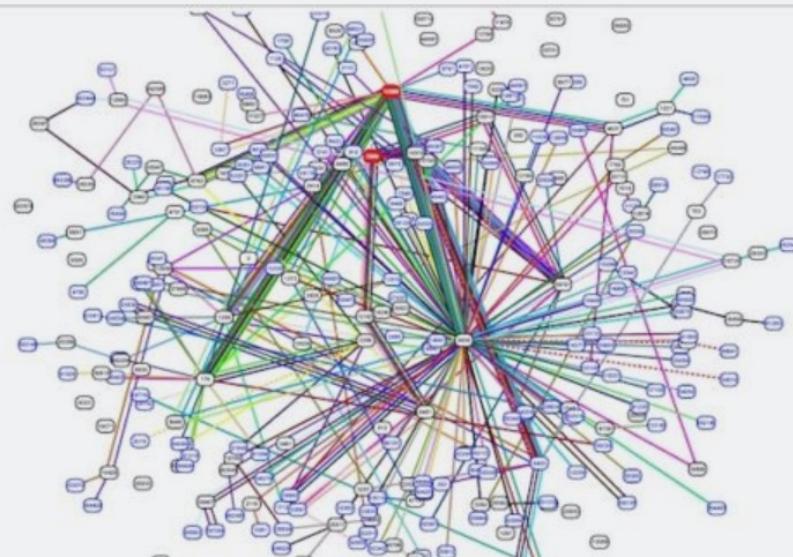
For the next 30 months or so, a large amount of traffic that used Verizon's AS703 improperly passed through AS4134 in mainland China first. The circuitous route is reflected in the following

Rostelecom Incident

Russian-controlled telecom hijacks financial services' Internet traffic

Visa, MasterCard, and Symantec among dozens affected by "suspicious" BGP mishap.

DAN GOODIN - 4/27/2017, 4:20 PM



36 networks affected,
incl. Symantec, EMC,
many banks

IP Address Ownership and Hijacking

IP address block assignment

- Regional Internet Registries (ARIN, RIPE, APNIC)
- Internet Service Providers

Proper origination of a prefix into BGP

- By the AS who owns the prefix
- ... or, by its upstream provider(s) in its behalf

However, what's to stop someone else?

- Prefix hijacking: another AS originates the prefix
- BGP does not verify that the AS is authorized
- Registries of prefix ownership are inaccurate

Hijacking is Hard to Debug

The victim AS doesn't see the problem

- Picks its own route
- Might not even learn the bogus route

May not cause loss of connectivity

- E.g., if the bogus AS snoops and redirects
- ... may only cause performance degradation

Or, loss of connectivity is isolated

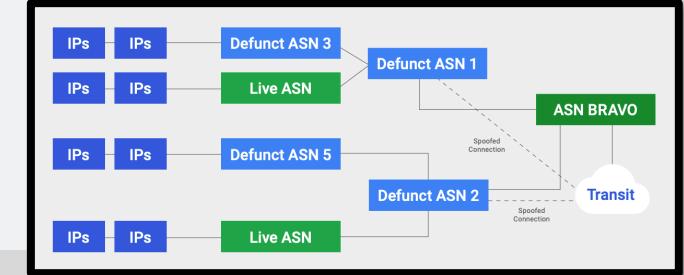
- E.g., only for sources in parts of the Internet

Diagnosing prefix hijacking

- Analyzing updates from many vantage points
- Launching traceroute from many vantage points

How 3ve's BGP hijackers eluded the Internet—and made \$29M

3ve used addresses of unsuspecting owners—like the US Air Force.



Eastern European cybercriminals set up several legitimate-looking ISPs + defunct ISPs to advertise routes to unused IP addresses

- Later started impersonating legitimate ISPs

1 million compromised IP addresses

IP addresses used for ad fraud: host fake sites impersonating premium websites, send bots to “view” ads on these websites

- 700K bots / compromised Windows machines
- 3B+ daily ad bid requests

Who Can Advertise Routes?

IRR (Internet Route Registry) system records which AS can originate/announce which routes... but

The hijacker, complainant Brian Rak reported in the thread, had misappropriated the IP addresses in part by creating fraudulent entries in the RIPE Internet routing registry (IRR), the organization's definitive database that maps what IP space belongs to various ASes. Rak said he published the post after Bravo's primary transit provider allowed the hijacking to continue even after he privately reported the fraudulent route announcements.

Rak later went on to identify the carrier providing service to Bravo as Telia Company of Sweden. When he brought the hijacking to Telia's attention in an attempt to get the fraudulent announcements removed, Rak said Telia officials cited the unauthorized entry in the RIPE IRR as proof Bravo had a legitimate right to announce the IP addresses in dispute. "I couldn't really facepalm hard enough after I got that email," Rak wrote.

Two International Cybercriminal Rings Dismantled and Eight Defendants Indicted for Causing Tens of Millions of Dollars in Losses in Digital Advertising Fraud

Global Botnets Shut Down Following Arrests

A 13-count indictment was unsealed today in federal court in Brooklyn charging Aleksandr Zhukov, Boris Timokhin, Mikhail Andreev, Denis Avdeev, Dmitry Novikov, Sergey Ovsyannikov, Aleksandr Isaev and Yevgeniy Timchenko with criminal violations for their involvement in perpetrating widespread digital advertising fraud. The charges include wire fraud, computer intrusion, aggravated identity theft and money laundering. Ovsyannikov was arrested last month in Malaysia; Zhukov was arrested earlier this month in Bulgaria; and Timchenko was arrested earlier this month in Estonia, all pursuant to provisional arrest warrants issued at the request of the United States. They await extradition. The remaining defendants are at large.

In-Class Exercise

How would you prevent IP hijacking?

Preventing Prefix Hijacking

Origin authorization

- Secure database lists which AS owns which IP prefix

soBGP

- Digitally signed certificates of prefix ownership

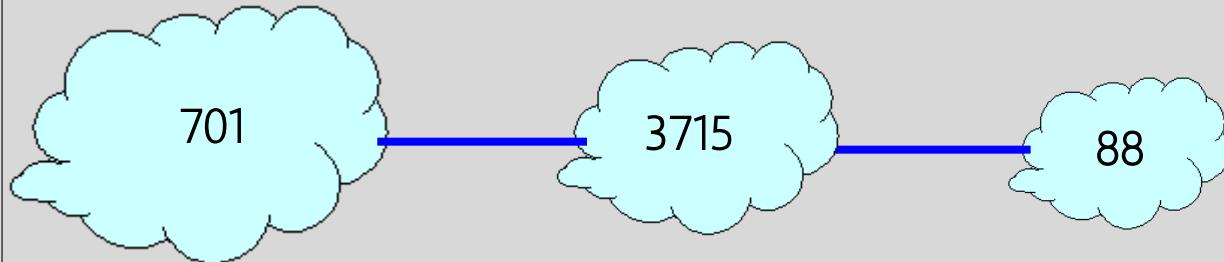
Prefix hijacking is not the only threat... in general, BGP allows ASes to advertise bogus paths

- Remove another AS from a path to make it look shorter, more attractive, get paid for routing traffic
- Add another AS to a path to trigger loop detection, make your connectivity look better

Bogus AS Paths

Remove an AS from the path

For example, advertise 701 88 for this path

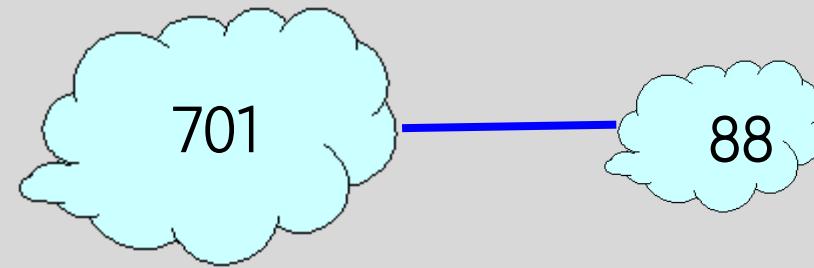


Possible motivations:

- Make AS path look shorter than it is
- Attract sources that try to avoid 3715

Add an AS to the path

For example, advertise 701 3715 88 for this path



Possible motivations:

- Trigger loop detection in AS 3715
- Make your AS look like it has richer connectivity

RPKI

<https://blog.cloudflare.com/rpki/>

IRR (Internet Route Registry) system records which AS can originate/announce which routes

- Unauthenticated, contains invalid data

RPKI cryptographically signs ROA (“Route Origin Authorization”) records

- Verifies that an owner of IP address block authorized a particular AS to originate routes to those addresses

Who has the authority to sign these records?

- TAs (“trust anchors”): currently RIRs (Regional Internet Registries)
- Somewhat similar to Web certificates

Mutually Agreed Norms for Routing Security



Now, more than ever, we need a more resilient Internet. Mutually Agreed Norms for Routing Security (MANRS) is a global initiative, supported by the Internet Society, that provides crucial fixes to reduce the most common routing threats.

Protecting BGP

Simple authentication of packet sources and packet integrity is not enough

Before AS advertises a set of IP addresses, the owner of these addresses must authorize it

- Goal: verify path origin

Each AS along the path must be authorized by the preceding AS to advertise the prefixes contained in the UPDATE message

- Goal: verify propagation of the path vector

S-BGP Protocol

Address attestation

- Owner of one or more prefixes certifies that the origin AS is authorized to advertise the prefixes
- Need a public-key infrastructure (PKI)
 - X.509 certificates prove prefix ownership; owner can then delegate his “prefix advertising rights” to his ISP

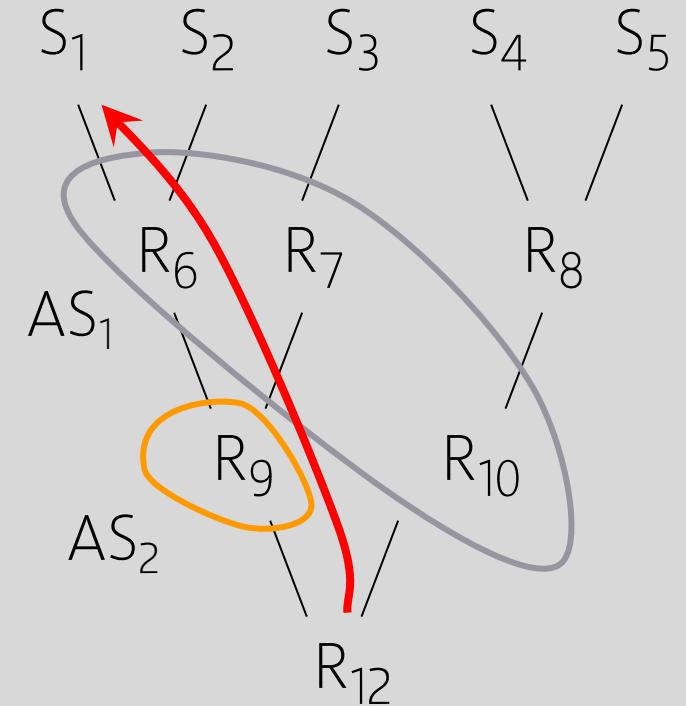
Route attestation

- Router belonging to an AS certifies (using digital signatures) that the next AS is authorized to propagate this route advertisement to its neighbors
- Need a separate public-key infrastructure
 - Certificates prove that AS owns a particular router

S-BGP Update Message

An update message from R_9 advertising this route must contain:

- Ownership certificate certifying that some X owns IP address S_1
- Signed statement from X that AS₁ is authorized to advertise S_1
- Ownership certificate certifying that AS₁ owns router R_6
 - If AS is represented by a router
- Signed statement from R_6 that AS₂ is authorized to propagate AS₁'s routes
- Ownership certificate certifying that AS₂ owns router R_9
- Lots of public-key operations!



Securing BGP

Dozens of proposals, various combinations of cryptographic protocols and anomaly detection

- Example: Secure BGP (S-BGP)
 - Origin authentication + entire AS path digitally signed
 - Can verify that the route is recent, no ASes have been added or removed, the order of ASes is correct
- Also: IRV, SPV, psBGP, PGBGP, PHAS, Whisper...

How many of these have been deployed?

None

- No complete, accurate registry of prefix ownership
- Need a public-key infrastructure
- Cannot react rapidly to changes in connectivity
- Cost of cryptographic operations
- Not deployable incrementally