Has Pakistan Stolen Your Traffic Lately?

Threats to Internet Routing and Global Connectivity

20th Annual FIRST Conference
Vancouver, British Columbia Canada
June 2008

Earl Zmijewski, Renesys Corp

The Internet – For Better or For Worse

Has the following characteristics ...

- *voluntarily association*
- *no governing body*
- *cooperative trust-based system*
- *commodity service*

which have given the Internet …

- its *explosive growth* **and**
- many of its *major problems*.

An ungoverned, trust-based Internet has given us …

- Spam
- Viruses
- Malware
- Phishing
- Spyware
- Worms
- Trojans
- DDoS attacks
- Wide-spread identity theft and fraud
- …

We will not talk about any of these.

Instead, we'll talk about threats to the Internet infrastructure …

- Physical problems

  (Physical Infrastructure: Natural, accidental or intentional destruction)
  - Earthquakes, Anchors/Backhoes, Hurricanes
- Routing Vulnerabilities

  (Logical Infrastructure: if routers cannot direct traffic appropriately, the Internet is broken.)
  - Misconfigurations, hijacks, attacks
- Business Conflicts

  (Competitors might not want to exchange traffic.)
  - De-peerings

These threats can …

- Break global connectivity
  - Disrupting Internet-dependent businesses
  - Impacting the global economy
- Be difficult to diagnose and troubleshoot
  - People don't often think about routing
  - Or don't have good visibility into global routing
- Be extremely difficult to fix
  - Physical destruction: weeks to months
  - Routing vulnerabilities: hours to days
  - Business conflicts: weeks to forever
- Lack sufficient business drivers to fix
  - In a commodity market, it's difficult to justify added expense for redundancy, monitoring, security
  - Interconnections are voluntary!

Threats to the Internet infrastructure

- Physical problems

  (Physical Infrastructure: Natural, accidental or intentional destruction)
  - Earthquakes, Anchors/Backhoes, Hurricanes
- Routing Vulnerabilities

  (Logical Infrastructure: if routers cannot direct traffic appropriately, the Internet is broken.)
  - Misconfigurations, hijacks, attacks
- Business Conflicts

  (Competitors might not want to exchange traffic.)
  - De-peerings

Physical Problems – Recent History

- Earthquakes
  - Taiwan quakes (2006 – 2007)
- Anchors – The Backhoes of the Sea
  - Mediterranean & Gulf cable breaks (2008)
- Hurricanes
  - Katrina (2005)

Physical failures can also be the result of malicious activity, e.g., 9/11

Physical Problems

- Earthquakes
  - Taiwan quakes
- Anchors – The Backhoes of the Sea
  - Mediterranean & Gulf cable breaks
- Hurricanes
  - Katrina

But first, a bit of terminology

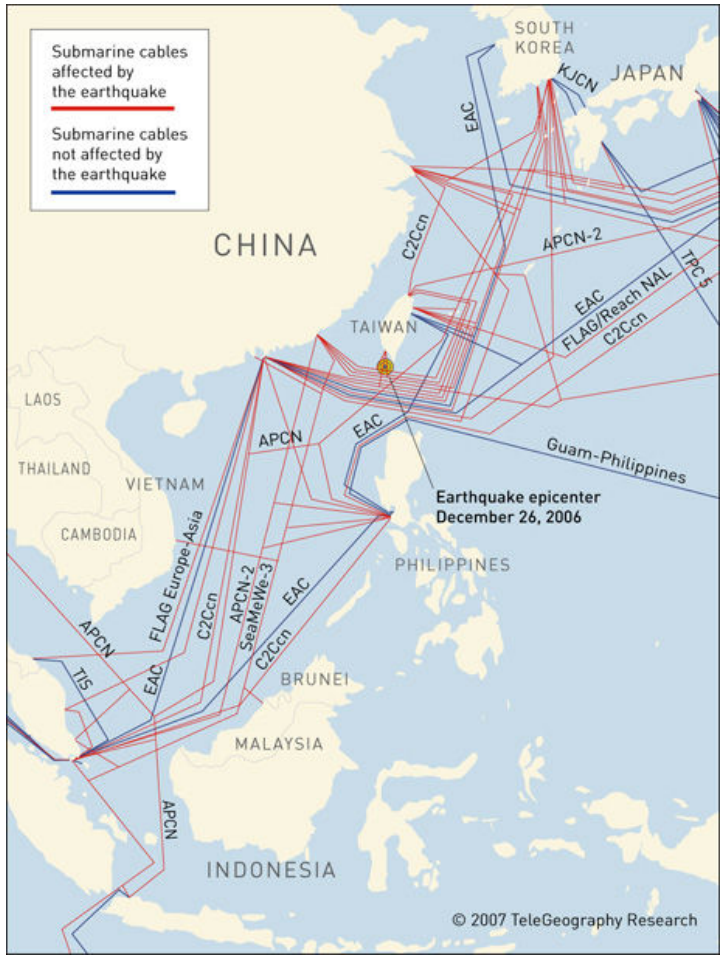
- **Network prefix** is a range of contiguous IP addresses:
  - **11.1.18.0/24** contains addresses **11.1.18.0, …, 11.1.18.255**
- **Prefix length** equals the number of network address bits:
  - **11.1.18.0/24 = 00001011.00000001.00010010.00000000**
    24 network address bits 8 host IP address bits
  - **11.1.16.0/20 = 00001011.00000001.0001**0000.00000000
    20 network address bits 12 host IP address bits
- **More-specific prefix** means …
  - more network address bits
  - fewer host IP addresses
  - **11.1.18.0/24** is more specific than **11.1.16.0/20**.

Taiwan Earthquakes – December 2006

- Large earthquakes hit Luzon Strait, south of Taiwan on 26 December 2006
- 7 of 9 cables in the strait were severed
- 7901 prefix outages (28% of region)
- 15,780 unstable prefixes (36% of region)
- All cables reported repaired 51 days later!

(source: 14 February 2007, Office of the Telecommunications Authority of Hong Kong)

Submarine cables in East Asia

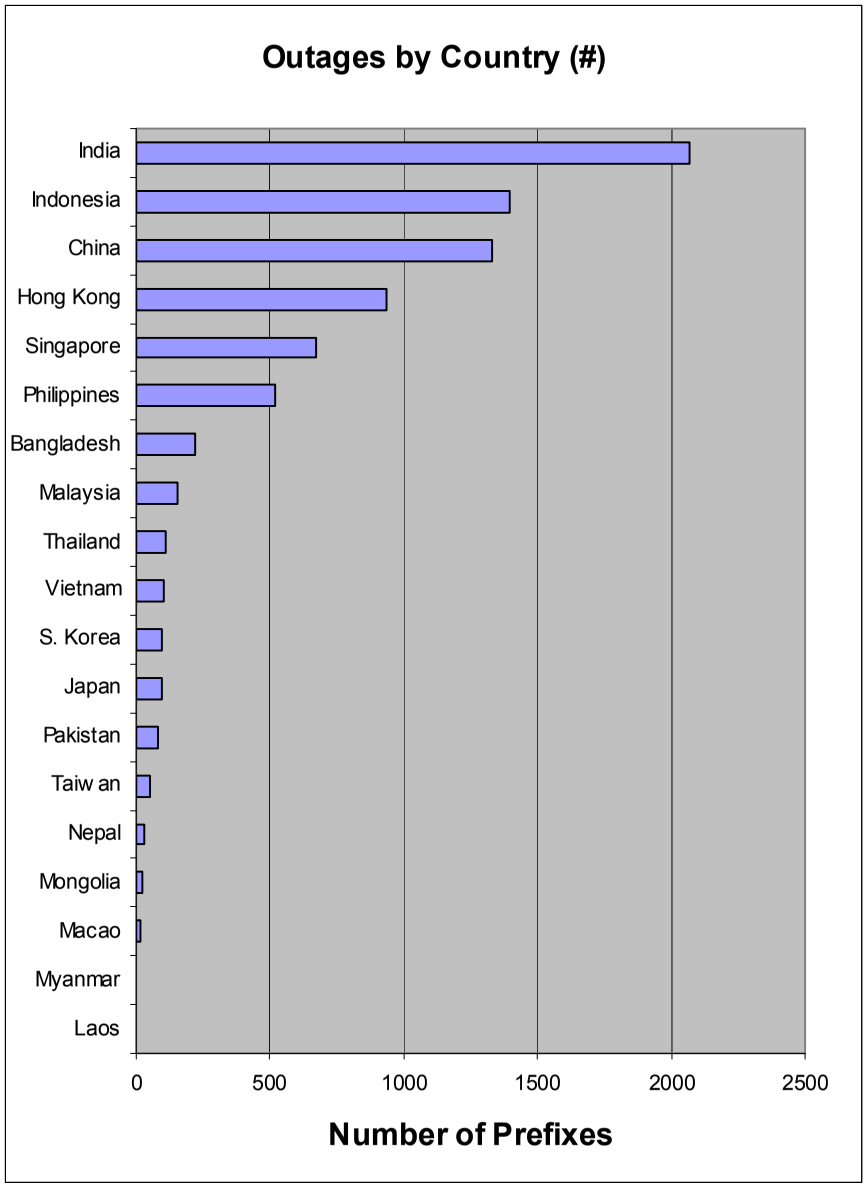


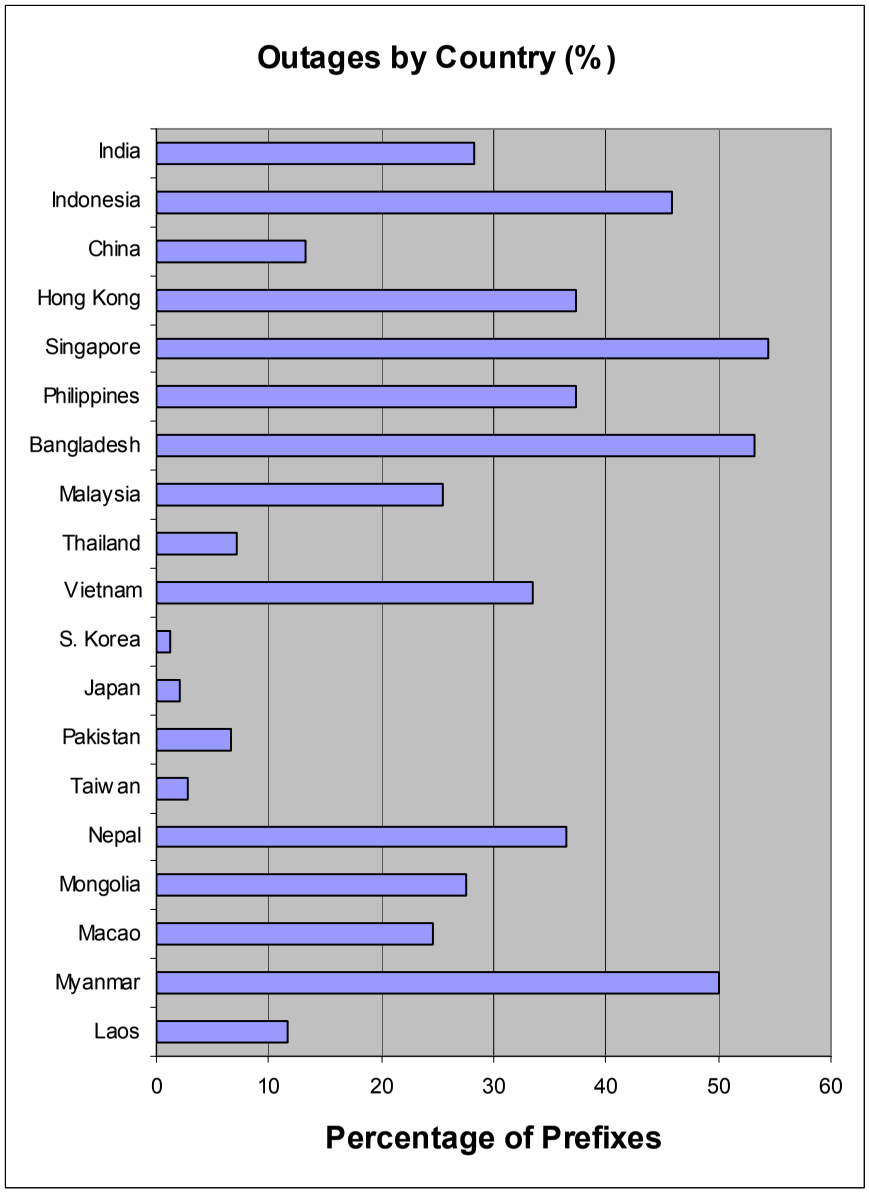
- Damaged cables
  - China-US Cable Network
  - SEA-ME-WE 3
  - Asia-Pacific Cable Network 2
  - FLAG Europe Asia
  - FLAG North Asia, and others
- Only two **not** impacted
  - Asia Netcom's EAC
  - Guam-Philippines Cable

Different ways to impact

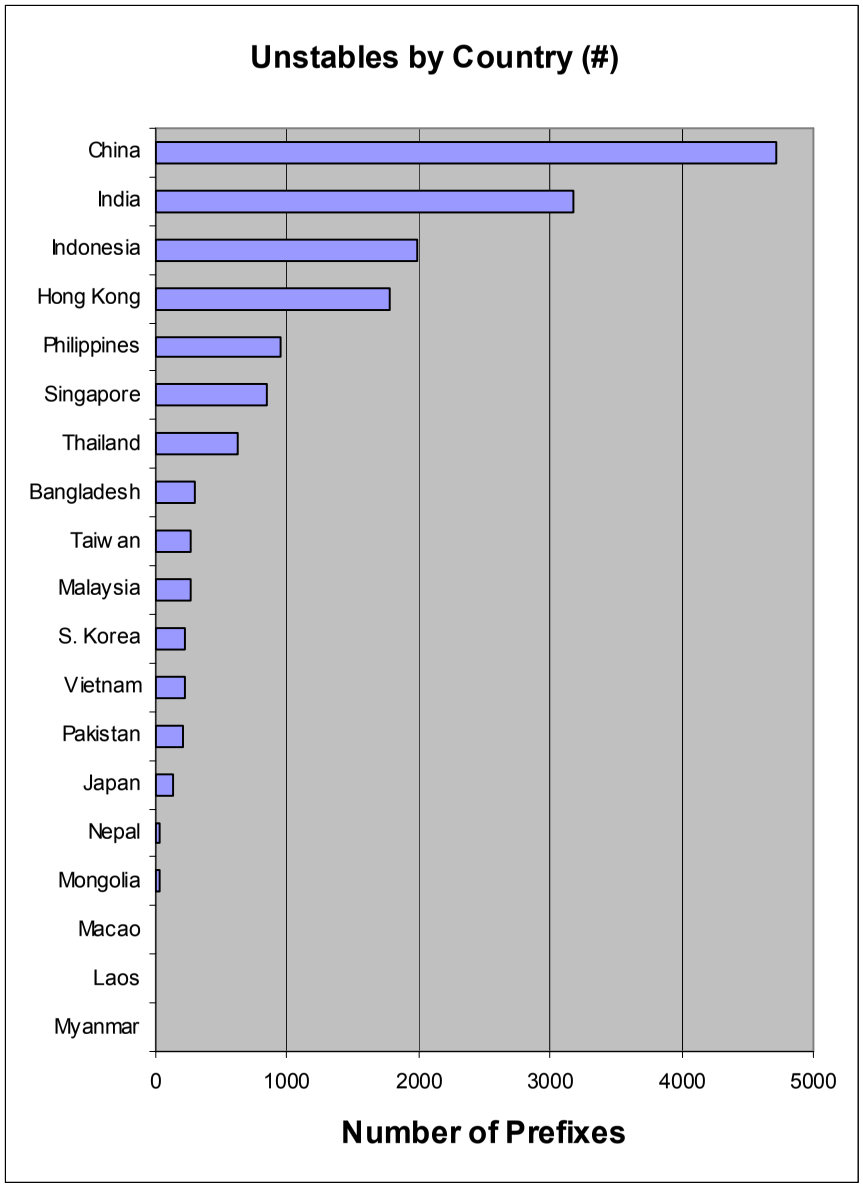
- Prefix Outages
  - By Country / Population
    - Total outages
    - Percentage of outaged prefixes
- Unstable Prefixes
  - By Country / Population
    - Total unstable prefixes
    - Percentage of unstable prefixes

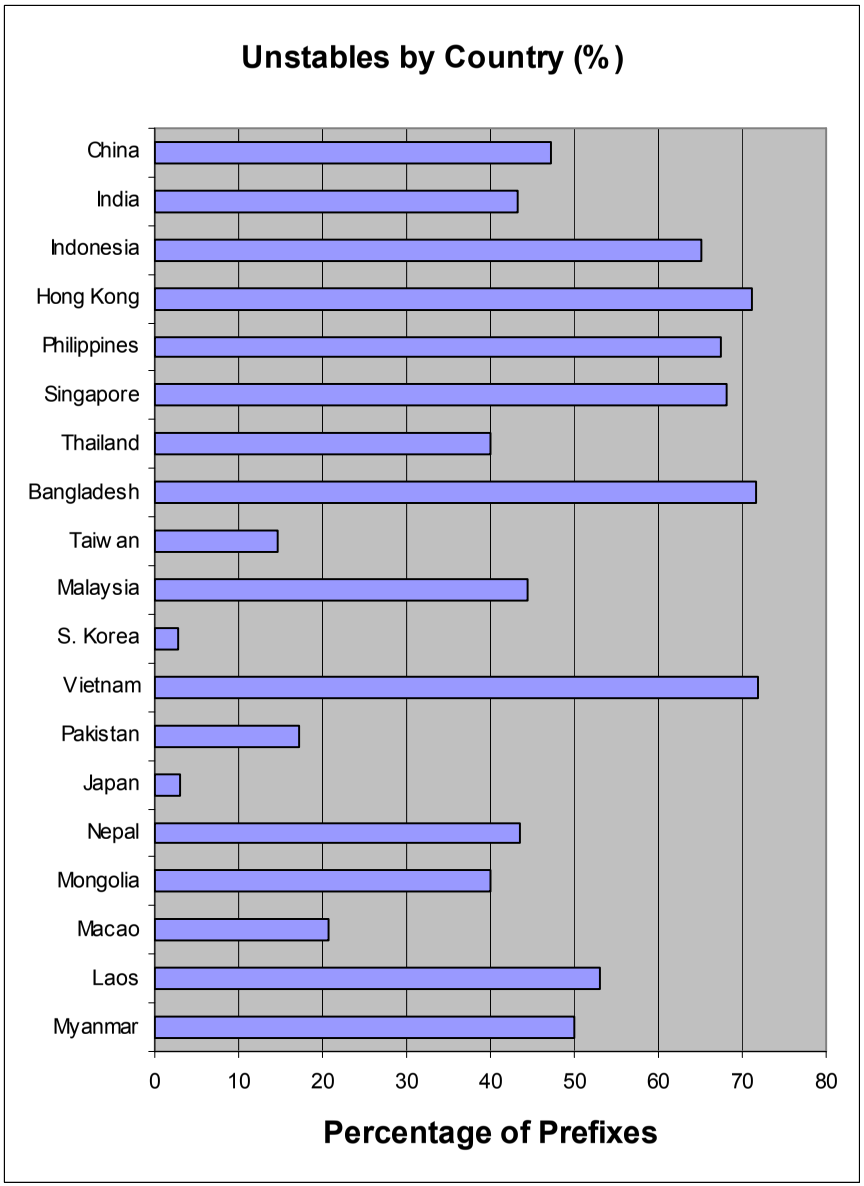
Outages by Country: Total and Percentage





Unstables by Country: Total and Percentage





Taiwan Earthquakes – Overall Impact

- Worst Impacted (> 60% Outages + Unstables)
  (with large populations)
  - India, China/Hong Kong, Indonesia, Philippines
- Least Impacted (< 5% Outages + Unstables)
  - Korea, Japan
- Modest Impact on Taiwan
  - ~ 3% Outages + ~14% Unstables
- Significant Impact on India
  - Numerous outsourced operations were hurt

Why India?

Major subcontinent bandwidth heads East

Image credit: Asia Netcom

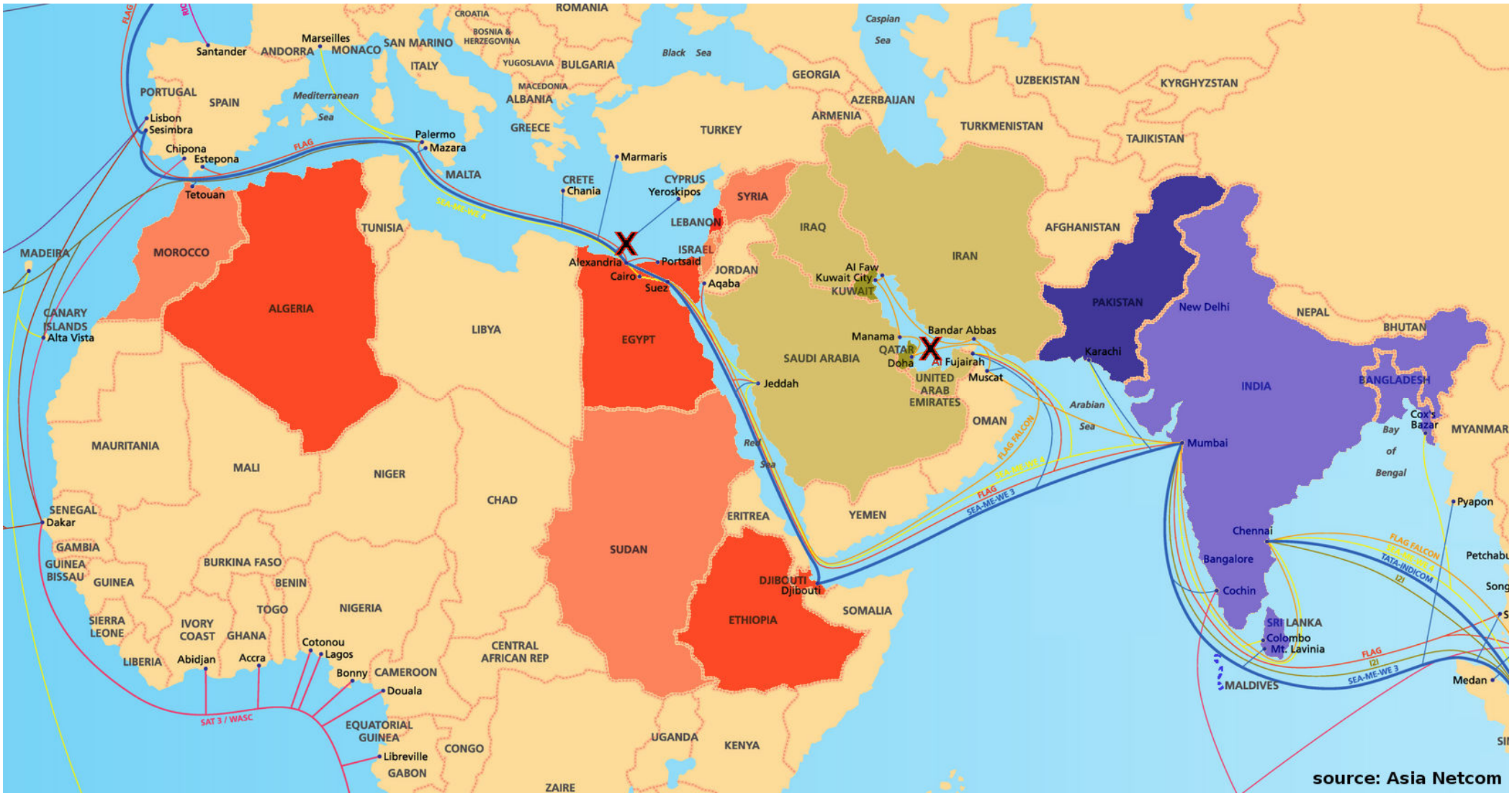
Physical Problems

- Earthquakes
  - Taiwan quakes
- Anchors – The Backhoes of the Sea
  - Mediterranean & Gulf cable breaks
- Hurricanes
  - Katrina

Middle East Cable Breaks – Jan/Feb 2008

- Several cables in the Mediterranean and the Persian Gulf were damaged from around January 30 to February 2
  - All repairs completed within 14 days.
- Impacted regions include …
  - Middle East / North Africa (65% outaged prefixes)
    (not including Israel since they weren't impacted)
  - Persian Gulf (45% outaged prefixes)
  - Indian Subcontinent (32% outaged prefixes)
- 6856 prefixes from 23 countries suffered outages

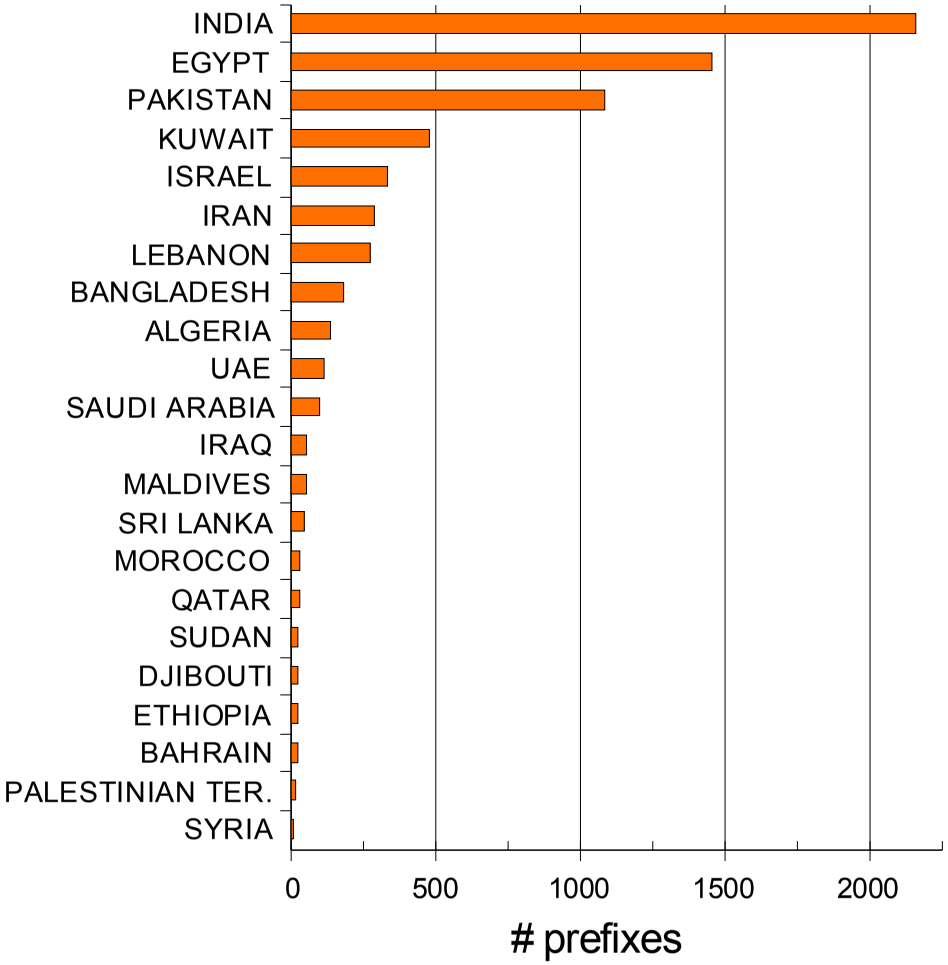
Impacted Countries



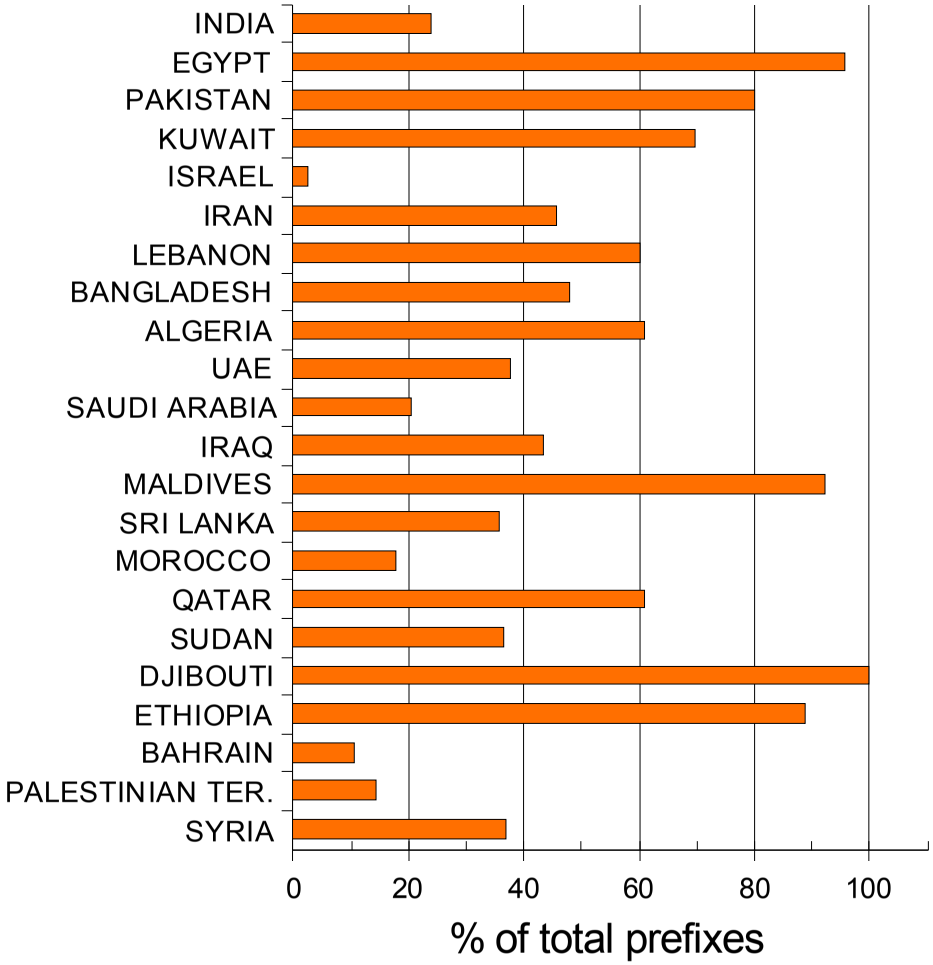
Darker colors in *each region* represent countries with prefix outages of at least 50%.

Outages by Country: Total and Percentage

Outages per country (# pfxs)



Outages per Country (%)



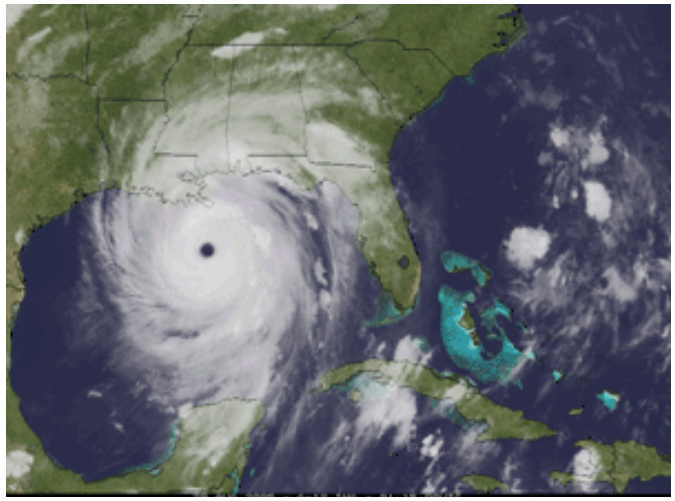
Middle Cable Breaks – Overall Impact

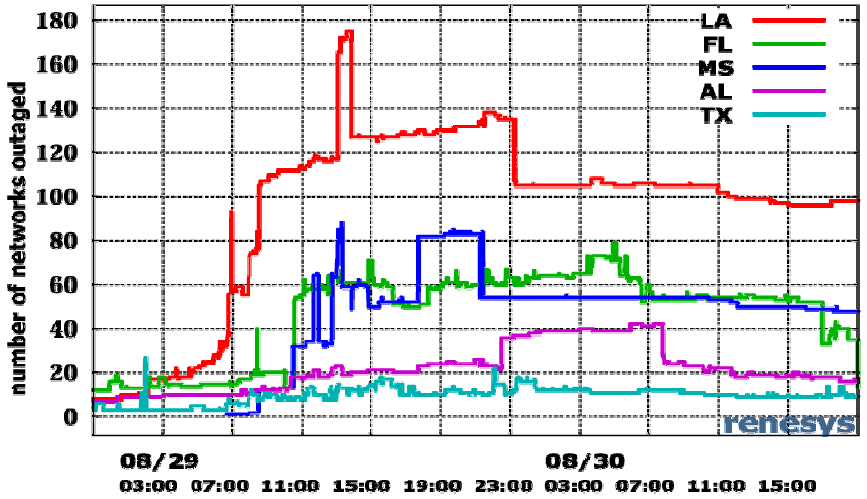
- Most Impacted (≥ 80% Outages)
  - Djibouti, Egypt, Ethiopia, Maldives, Pakistan
- Least Impacted (≤ 10% Outages)
  - Israel, Bahrain
- Significant Impact on India (again!)

Physical Problems

- Earthquakes
  - Taiwan quakes
- Anchors – The Backhoes of the Sea
  - Mediterranean & Gulf cable breaks
- Hurricanes
  - Katrina

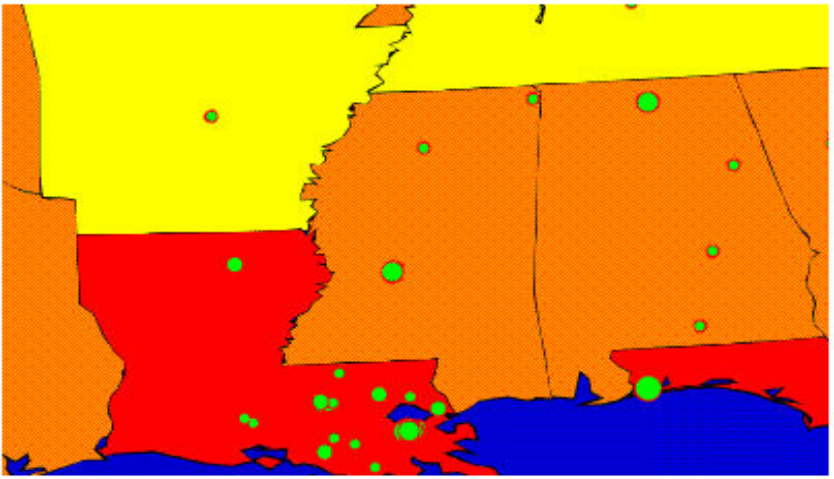
These threats aren't faraway problems





Mississippi Report: Network Outages over the past 2 hours

19:33:53 UTC 08 Sep 2005



Country	Network	State	Zip
	CommuniGroup of Jackson MS MS 39201 (65.183.96.0/20)	MS	39201
	WATER VALLEY INTERCHANGE MS 38965 (64.49.18.0/24)	MS	38965
	TriState Education initiative MS 38852-4375 (192.149.138.0/24)	MS	38852-4375
	Arch Communications MS 39157 (208.251.18.0/24)	MS	39157
	AIR2LAN Inc MS 39216 (216.212.214.0/24)	MS	39216
	AIR2LAN Inc MS 39216 (216.212.215.0/24)	MS	39216
	AIR2LAN Inc MS 39216 (216.212.216.0/24)	MS	39216
	AIR2LAN Inc MS 39216 (216.212.217.0/24)	MS	39216
	AIR2LAN Inc MS 39216 (216.212.218.0/24)	MS	39216
	AIR2LAN Inc MS 39216 (216.212.219.0/24)	MS	39216
	AIR2LAN Inc MS 39216 (216.212.220.0/24)	MS	39216
	AIR2LAN Inc MS 39216 (216.212.221.0/24)	MS	39216
	AIR2LAN Inc MS 39216 (216.212.222.0/24)	MS	39216
	AIR2LAN Inc MS 39216 (216.212.223.0/24)	MS	39216

Submarine Cable Systems by Capacity

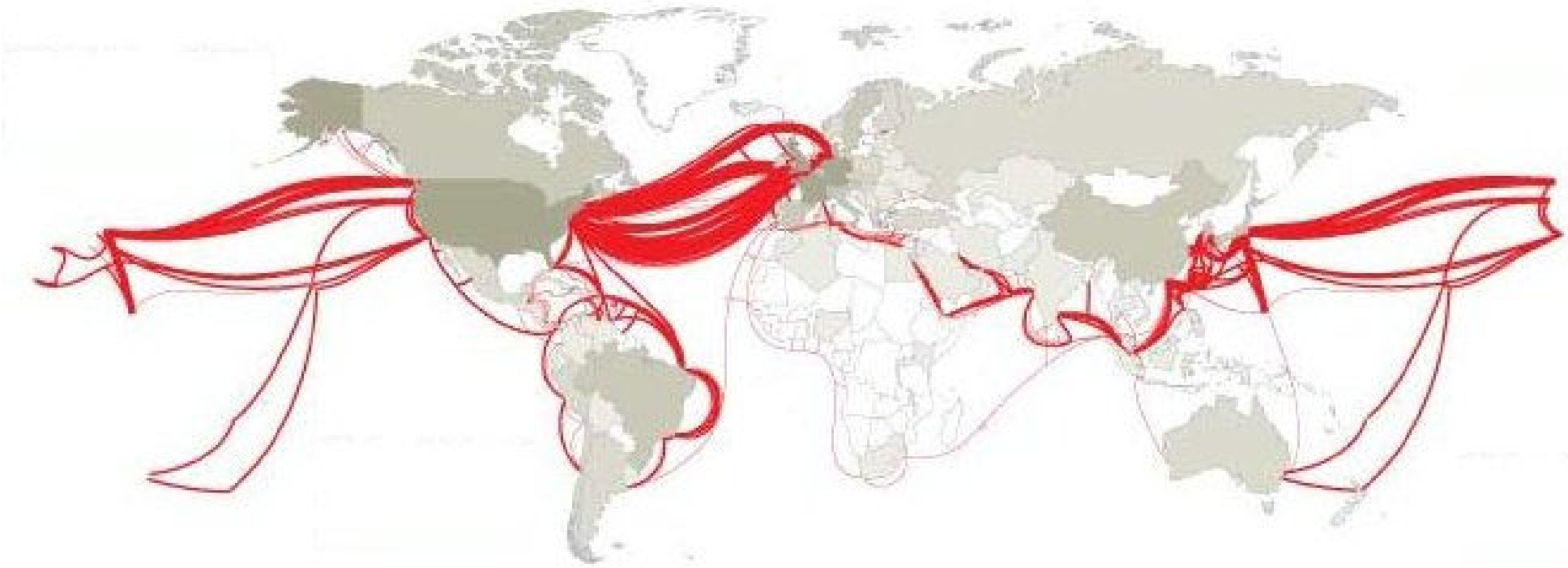


Image credit: Telegeography

Bandwidth-limited areas and choke points are obvious

Lessons learned from physical failures

- You get what you pay for
  - Natural trade-off: cost, performance/latency vs. reliability
- Entropy happens
  - Cables break in the Atlantic all the time, nobody notices
- Geography plays an important role
  - Cables break in the Taiwan Straits or Suez Canal, entire geographic regions lose connectivity
- Intelligence is essential for disaster planning and recovery
  - For organizations to select providers
  - For local ISPs to select new upstream providers
  - For a global ISP to acquire new customers

Physical Problems: Prognosis is Good!

- Recent failures have gotten the attention of businesses and governments
  - Businesses want providers with redundant capacity
  - Governments want control of their critical infrastructure
    - Saudi Arabia and Egypt have announced new cable deals
- Mini cable-building boom in progress
  - At least 25 new cables, costing approximately USD 6.4 billion, will be constructed between 2008 and 2010

Source: Telegeography

What can I do?

- Operations
  - Understand capacity, location and vulnerabilities
  - Diversify existing capacity
  - Get more capacity on diverse physical paths
- Policy
  - Influence national communications policy to encourage investment

Threats to the Internet infrastructure

- Physical problems
  (Physical Infrastructure: Natural, accidental or intentional destruction)
  - Earthquakes, Anchors/Backhoes, Hurricanes
- Routing Vulnerabilities
  (Logical Infrastructure: if routers cannot direct traffic appropriately, the Internet is broken.)
  - Misconfigurations, hijacks, attacks
- Business Conflicts
  (Competitors might not want to exchange traffic.)
  - De-peerings

Routing Vulnerabilities

- Hijacks
  - YouTube (2008)
  - Review of other notable incidents (1997 – present)
  - DOD (2008)
- DNS misappropriation
  - Root name server identity theft (2007 – 2008)

Autonomous System Numbers (ASNs)

Each organization announcing a routing policy on the Internet is assigned:

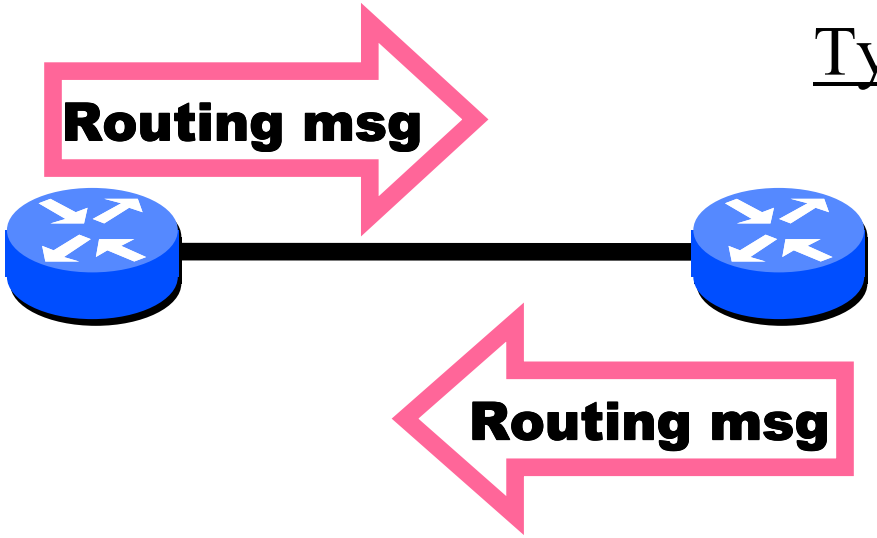
- A unique ASN (integer)
- One or more prefixes (range of IP addresses)

Example ASNs:

- Sprint: **1239**
- AT&T: **7018**
- IRS: **30313**
- YouTube: **36561**

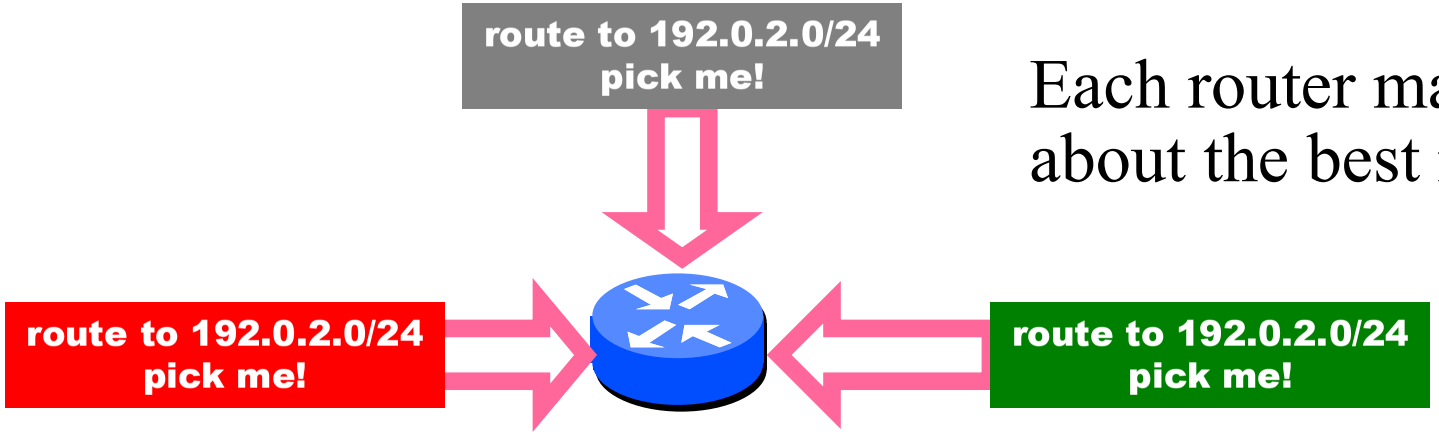
Routers talk to neighboring routers via BGP

(Border Gateway Protocol) that's how global routing is established



Typical messages

- "I have a route to prefix A"
- "I no longer have a route to prefix B"



Each router makes local decisions about the best route to each prefix

Routing Vulnerabilities

- Hijacks
  - YouTube (2008)
  - Review of other notable incidents (1997 – present)
  - DOD (2008)
- DNS misappropriation
  - Root name server identity theft (2007 – 2008)

YouTube – February 2008

Like most content providers, YouTube has no need for massive amounts of IP space.

- They are assigned only 5 small prefixes
- One prefix is 208.65.152.0/22
  - This contains the more-specific 208.65.153.0/24
  - This /24 *used* to contain all of YouTube's
    - DNS Servers (have since moved)
    - Web Servers
  - YouTube announced only the /22

Overview of 24 February 2008 Hijack

- Pakistan's government decides to block YouTube (Posted video is deemed "offensive".)
- Pakistan Telecom apparently black holes 208.65.153.0/24 on their *internal* network
- So far, this is only impacting Pakistan and their ability to reach YouTube
- This is **not** uncommon for some governments

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.

Overview of Hijack (Continued)

- Pakistan Telecom then mistakenly announces this critical YouTube prefix to their upstream Asian carrier PCCW as if it was their own.
- PCCW propagates this route globally.
- On the Internet, most specific route wins!
- Most of the Internet goes to Pakistan for YouTube for 2 hours and gets nothing!
- Eventually, PCCW turns off Pakistan Telecom

We've been here before, but on a larger scale …

Apr 1997	MAI Network Services (AS 7007) (70K bogus routes)
Dec 2004	TTNet (AS 9121) (100K bogus routes)
Jan 2006	Con Edison (AS 27506) (dozens of bogus routes)

Each of these providers announced parts of the Internet not under their control, resulting in various degrees of bedlam.

But do hijacks really occur with any regularity?

Examine two US DoD networks and their more specifics

DoD owns but does not announce 7.0.0.0/8, 11.0.0.0/8, 30.0.0.0/8 and others. These networks are "free for the taking" without any impact on DoD.

A Sampling of Hijacks in 2008					
Prefix	**Date(s)**	**Origination (AS)**	**Country**	**Avg Time per Peer**	**Max Peers**
11.11.11.0/24	May 17	Teknoas (AS 42075)	Turkey	6.5 min	232
11.11.11.0/24	May 10	INDO Internet (AS 9340)	Indonesia	2.1 min	155
30.30.30.0/24	April 30	Telefonica (AS 10834)	Argentina	40 min	241
11.0.0.0/24	April 25 – 26	ITC Deltacom (AS 6983)	US	16 hours	244
7.7.7.0/24	March 7	Posdata (AS 18305)	S. Korea	16 min	227
11.1.1.0/24	March 5 – 29	Helios Net (AS 21240)	Russia	3.5 weeks	248
11.11.11.0/24	January 5	Hutchinson (AS 9304)	Hong Kong	1.1 hours	207

Every announcement in this assigned, but unused, space is a hijack.

Threat is Poorly Understood: Memorable Quotes

- Full technical details published 24 February at www.renesys.com/blog

- "We are not hackers. Why would we do that?" Shahzada Alam Malik, head of the Pakistan Telecommunication Authority, told Associated Press Television News. YouTube's wider problems were likely caused by a "malfunction" elsewhere, he said.

— International Herald Tribune, 27 February 2008

- Attempts to log on to the Google-owned site typically timed out. Keynote* is unable to uncover the causes of an outage, said Shawn White, Keynote's director of operations, but he added that he would be shocked if one country had the ability to bring down YouTube globally.

— CNET, 24 February 2008

* Keynote Systems, Inc. is “The Mobile and Internet Performance Authority”

Solutions?

- Replace BGP (go ahead, I'll wait)
  - Limited value unless everyone does it
- Filter announcements
  - Difficult to implement and maintain
  - No single authoritative source of who owns what
  - Security may not be compatible with resiliency
- Monitor networks you care about
  - Increases costs
  - Reactive, not Proactive
    - Do you know anyone at Pakistan Telecom?
    - You are hijacked, then what?

Routing Vulnerabilities

- Hijacks
  - YouTube (2008)
  - Review of other notable incidents (1997 – present)
  - DOD (2008)
- DNS misappropriation
  - Root name server identity theft (2007 – 2008)

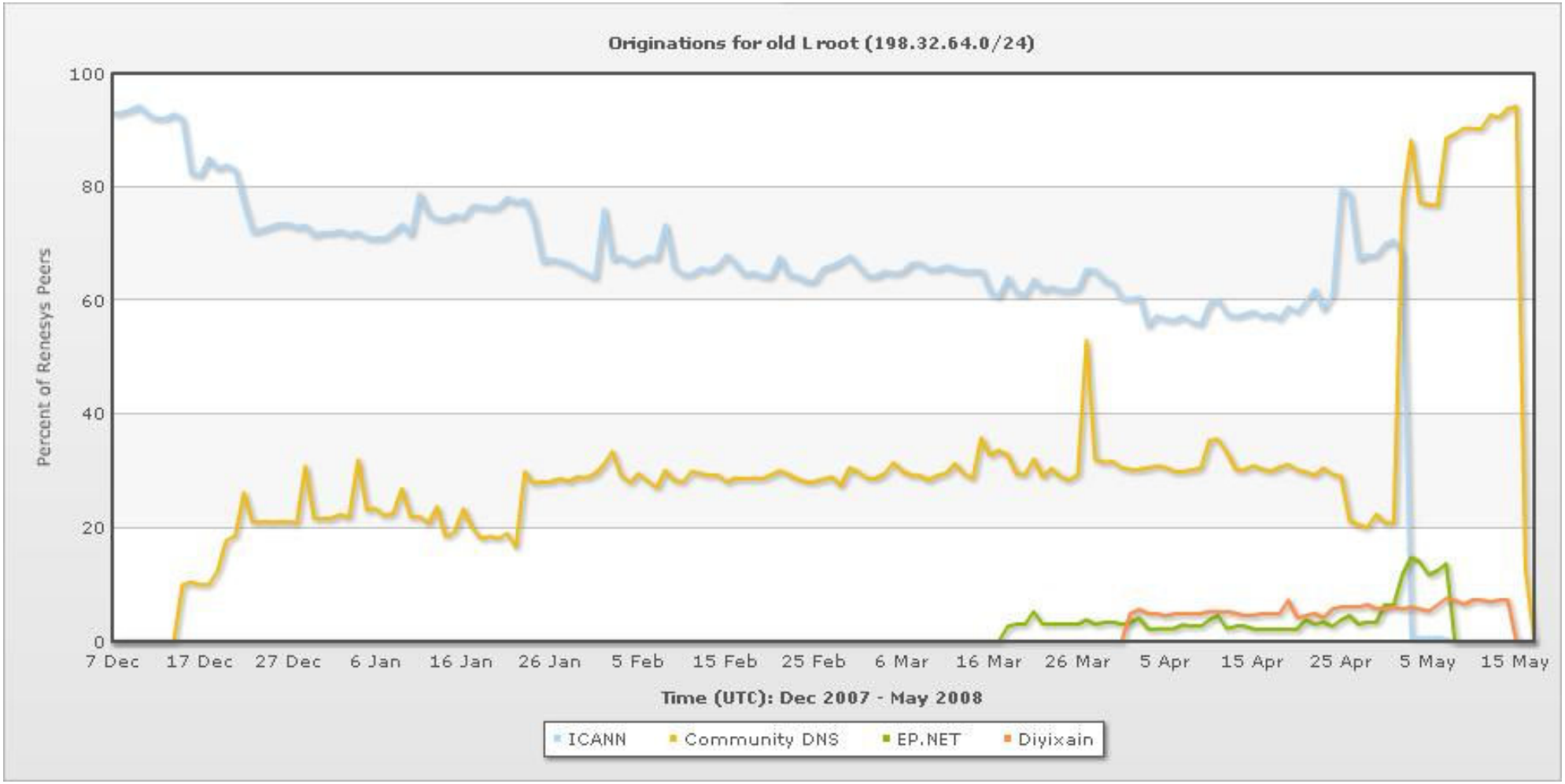
DNS – Root Name Servers

- 13 root name server IP addresses
- Named by single letters: A, B, C, … M.
- L root is run by ICANN
  - Old IP: 198.32.64.12
    - From 1997 until 2007
    - Registered to Bill Manning, ep.net
  - New IP: 199.7.83.42
    - Effective 1 November 2007
    - Registered to ICANN

Old L Root Name Servers: 2007 – 2008

- ICANN runs the old L root for additional 6 months
- New unauthorized root servers start appearing
  - Dec 15th – Community DNS (England)
  - Mar 18th – EP.NET (US – Bill Manning)
  - Apr 1st – Diyixian (Hong Kong)
- May 2nd – ICANN turns off its own old L root
- May 16th – All bogus old L root servers turned off under pressure from ICANN

Timeline for old L root servers



Lots of unanswered questions

- Why was ICANN not using their own space?
  - Was it difficult to get IP space in 1997?
- Why the change after 10 years?
  - 11.7 million DNS servers worldwide (source: Infoblox)
  - How many of those do you think were updated?
- Why wasn't the space given to ICANN?
  - ARIN shows Manning has five /16s and a /22
  - This is equivalent to 1284 /24s (only 44% routed)
  - ICANN only needs a single /24 for the L root
- Why all the bogus L root servers?

Is this much ado about nothing?

- What could **you** do with a root name server?
  - Provide *updated* list of **all** the root name servers
  - Provide *updated* NS records for **all** TLDs
  - Set TTL = 0 for your answers
  - Perform recursion by default
  - Log everything
  - Censor, misdirect
- No evidence of any of this in this case
  - But the duration of the event, the potential for mayhem, the complete absence of *any* safeguards are the cause for concern

Lessons learned from routing vulnerabilities

- No one is minding the store. No one.
- No mechanism in place to handle root operators or ISPs who go *rogue*
- No single authoritative source of who should be doing what

Routing Vulnerabilities: Prognosis is Dismal

- Hijacking has been going on for over 10 years!
- No incremental, comprehensive solutions
- Lacking economic drivers
  - Doesn't happen daily and universally
  - Avoiding negative publicity is not necessarily compelling
- Miscreants are actively hijacking now
  - To send spam from "clean" IP blocks
  - To cover their tracks
    - What good are your firewall/IDS logs now?
    - Need historical global routing data

What can I do?

- Operations
  - Monitor prefixes you care about
    - Maintain alerts
    - Establish procedures for handling hijacks quickly
  - Develop procedures for dealing with hijacks in concert with your providers
- Policy
  - Build awareness of routing-based attacks
  - Influence positive policy (tricky, evidence not good)

Threats to the Internet infrastructure

- Physical problems
  (Physical Infrastructure: Natural, accidental or intentional destruction)
  - Earthquakes, Anchors/Backhoes, Hurricanes
- Routing Vulnerabilities
  (Logical Infrastructure: if routers cannot direct traffic appropriately, the Internet is broken.)
  - Misconfigurations, hijacks, attacks
- Business Conflicts
  (Competitors might not want to exchange traffic.)
  - De-peerings

Interconnections on the Internet

- Two forms of interconnection exist
  - **Transit**
    - " … transit service is typically priced per megabit per second per month …"

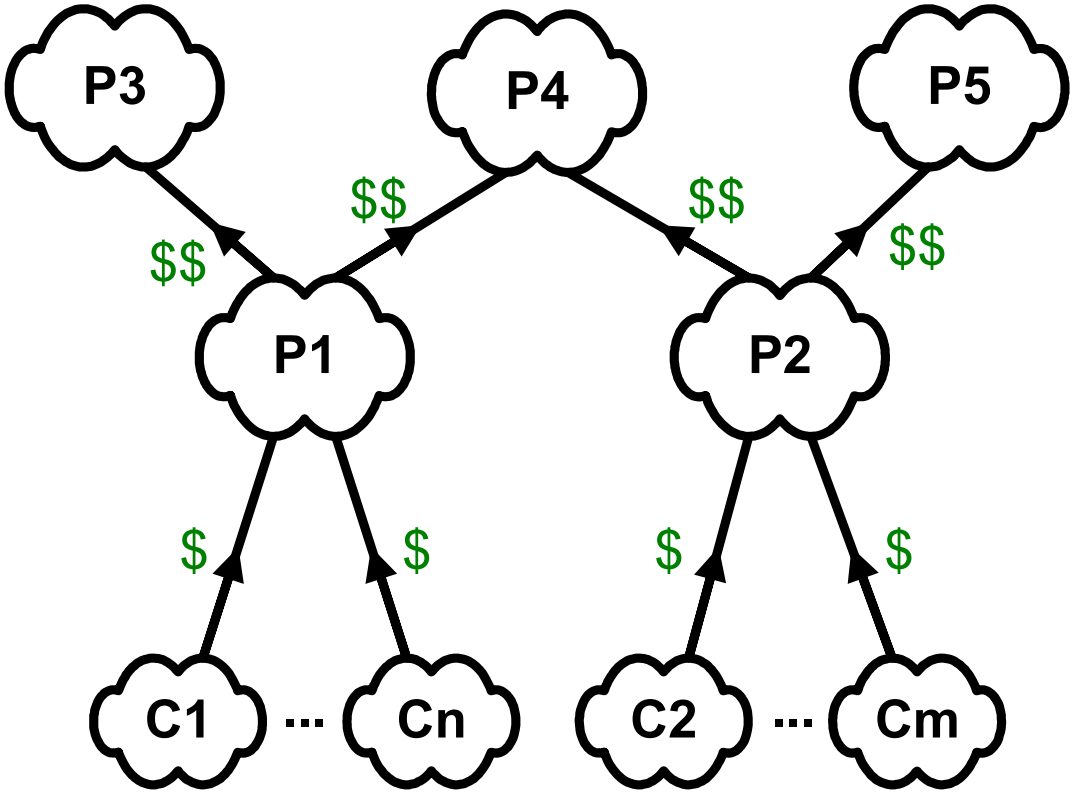
      Source: Wikipedia

      - single homed: one provider
      - multihomed: more than one provider
  - **Peering**
    - "Peering is voluntary interconnection of administratively separate Internet networks" … where "neither party pays the other for the exchanged traffic."

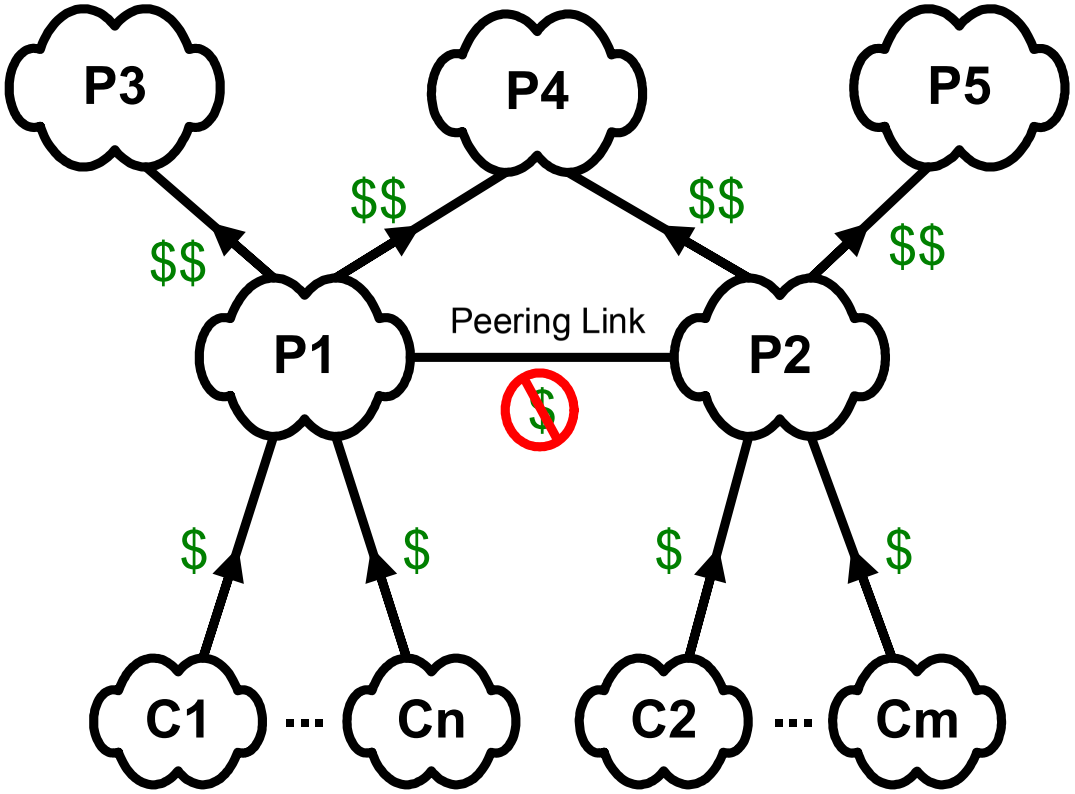
Business relationships:
Easy to observe, difficult to classify

- We observe over 8.7 million distinct AS paths.
- These paths are comprised of over 93,000 unique AS-AS adjacencies.
- Each adjacency represents a business relationship, e.g.,
  - Customer → Provider
  - Provider → Customer
  - Peer → Peer
  - Transit swap
  - AS cluster (multiple ASes now part of the same organization)
- These relationships can be compute algorithmically with a high degree of confidence.

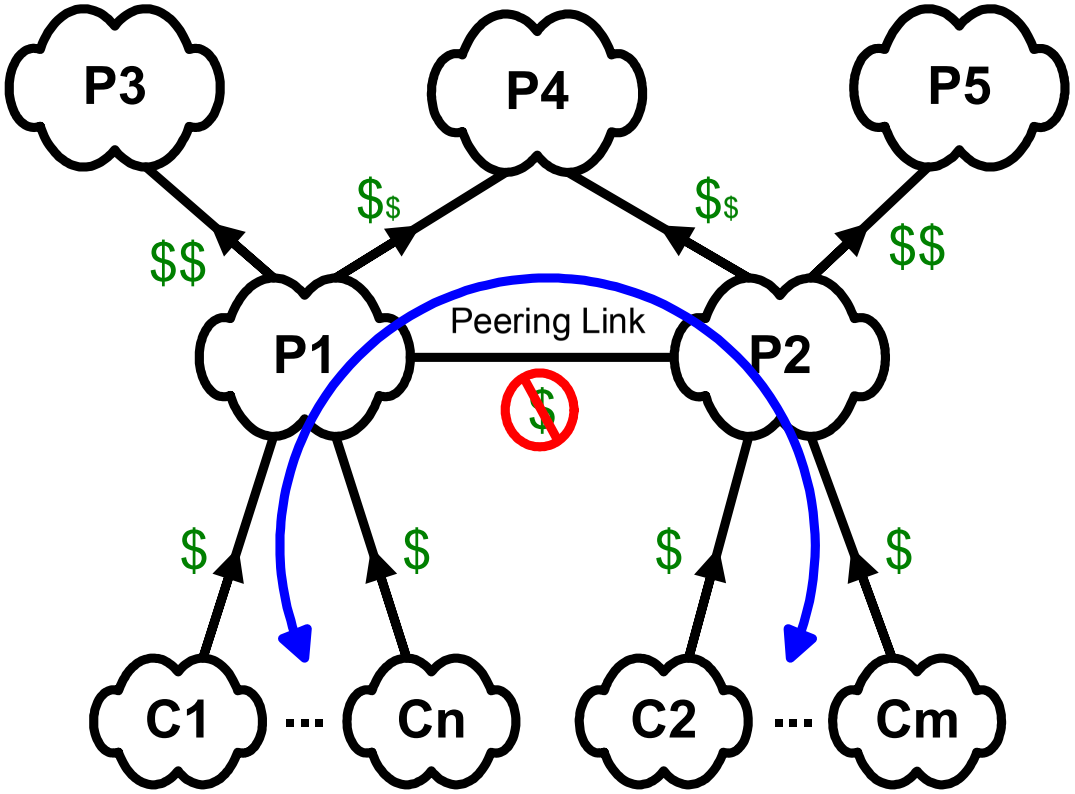
Peering Goal – Maximize $ In, Minimize $ Out



Reduce Transit Costs by Peering with Competitors



Customers of P1 and P2 exchange traffic without either provider incurring transit costs



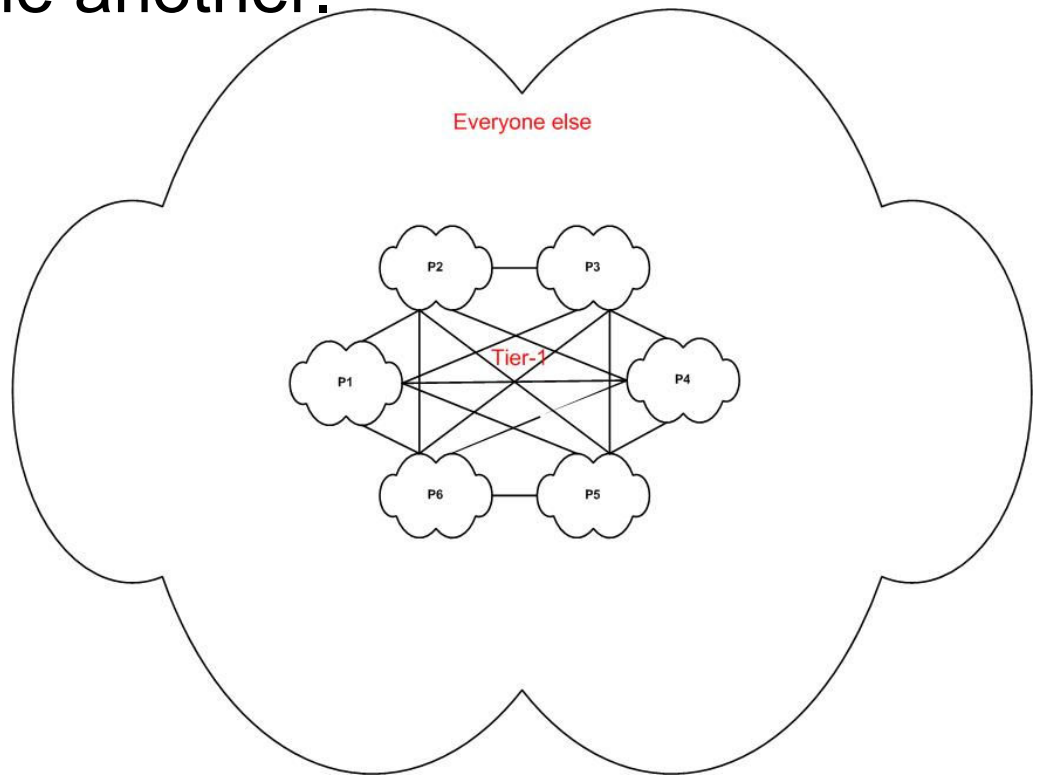
Reasons to Peer

- Reduce transit costs (happier providers)
- Reduce latencies (happier customers)
  - Increased billable traffic to customers
- Enhance operational stability
  - Localize connectivity
- Roughly equal mutual benefit

Top of the Internet Food Chain: Tier-1

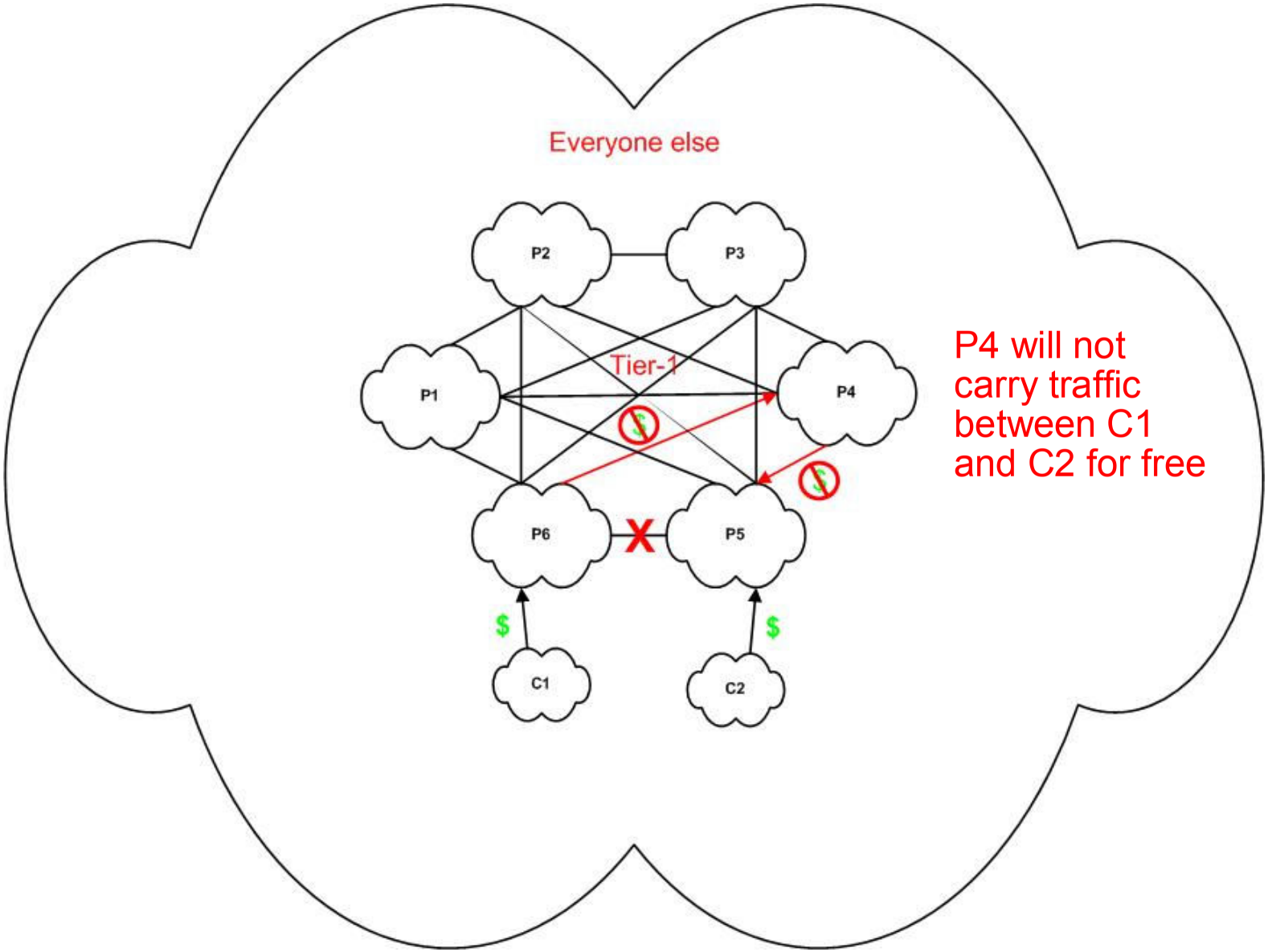
A **Tier-1 network** has no transit providers – only peers and customers. ("Tier-1" term is inconsistently used.)

To maintain *global connectivity*, the Tier-1 providers must all peer with one another.



Follow the Money –

if P5 & P6 de-peer, their single-homed customers can't communicate



Who are the Tier-1 providers?

- Sprint, AT&T, Level 3, NTT, Savvis, etc.
- Tier-1 providers act like a cartel and have no incentive to add members
- “Near” Tier-1 providers can try to buy their way into the club (via “paid peering”).

Cogent and Telia want to be Tier-1

- Cogent and Telia peer
- Cogent gets transit only from NTT to reach AOL
- Telia appeared to get transit from Verizon to reach certain networks
  - On February 27th, we stopped seeing evidence of transit
  - Renesys promotes Telia to “Tier-1” – no known providers
  - Telia could still be paying for some of these interconnections

Overview of Cogent-Telia Peering Dispute

- **March 13th:** Cogent de-peers Telia, claiming breach of contract
- The Internet is partitioned
  - Single homed customers behind Cogent and Telia could not reach one another.
- **March 28th:** Peering link is restored.
- After two weeks, the Internet is once again whole.

Which regions were impacted?

Telia cannot reach Cogent

Country	# Prefixes
US	1868
Canada	232
France	98
Spain	41
Germany	31
UK	27
Others	86

Cogent cannot reach Telia

Country	# Prefixes
Sweden	444
Finland	322
Russia	153
Poland	113
US	73
Latvia	62
Bulgaria	52
Spain	40
Denmark	35
Norway	30
Others	249

Why did this happen? Who knows?

- Peering disputes with Cogent tend to be about traffic ratios
  - Imbalanced ratios tend to imply that one party is carrying the other's traffic longer distances (distance = money)
- Did Cogent and Telia get into a dispute over ratios?
- Did Cogent view Telia as in a weaker position?
  - How many European customers want to reach Cogent hosted content?

What changed?

- After the re-peering
  - Telia reaches 2934 **more** prefixes via Cogent
  - Cogent reaches 635 **fewer** prefixes via Telia
- Result of re-engineering by both parties
- Sure seems like this was about traffic ratios

Lessons Learned from De-peerings

- Being a (near) Tier-1 is not easy
  - You depend on everyone else in the cartel
  - You will be punished if you are perceived to be in a position of weakness.
- Peering relationships are tricky
  - Depend on both objective measures and perceptions
  - Disputes can take a long time to resolve. The only driver is market pressure.
- Being single-homed is dangerous

De-peerings: Prognosis is Improving

- An increasing number of businesses and organizations are now multi-homed.
- Many more peering relationships, lessening the impact of any one loss
- True global Tier-1s would probably never consider de-peering
  - Impact would be wide ranging
  - Would invite government involvement
- Tier-1s are becoming less relevant, as smaller players peer around them

What can I do?

- Operations
  - Select multiple diverse providers
  - Investigate providers carefully with respect to connectivity & peering
- Policy
  - Build awareness of the issue
  - When peering disputes occur, make a stink
    - No stink → No action
  - Influence positive policy (tricky, evidence not good)

Conclusions

- Physical problems
  - Prognosis is good.
  - But US will become less relevant.
- Routing Vulnerabilities
  - Prognosis is dismal.
  - No clear path forward. Band-Aids only.
- Business Conflicts
  - Prognosis is improving.
  - Awareness and public pressure is key.

Thank You

Martin Brown	mabrown@renesys.com
James Cowie	cowie@renesys.com
Andy Ogielski	ato@renesys.com
Alin Popescu	alin@renesys.com
BJ Premore	bj@renesys.com
Todd Underwood	todd@renesys.com
Earl Zmijewski	earl@renesys.com