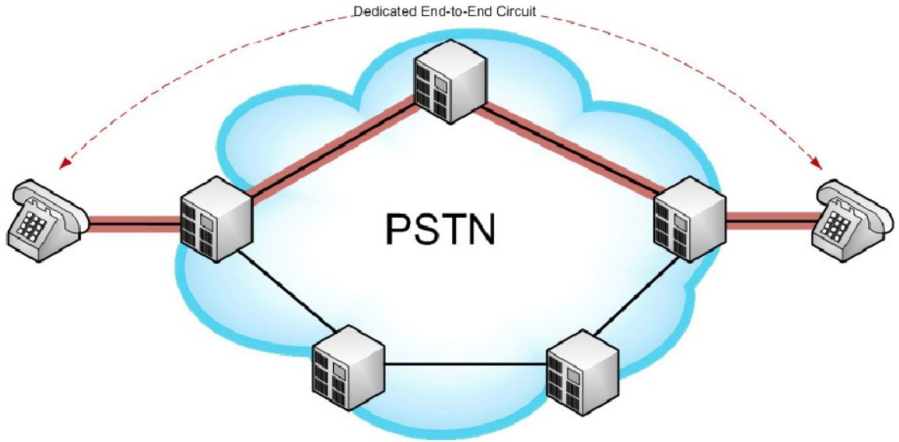
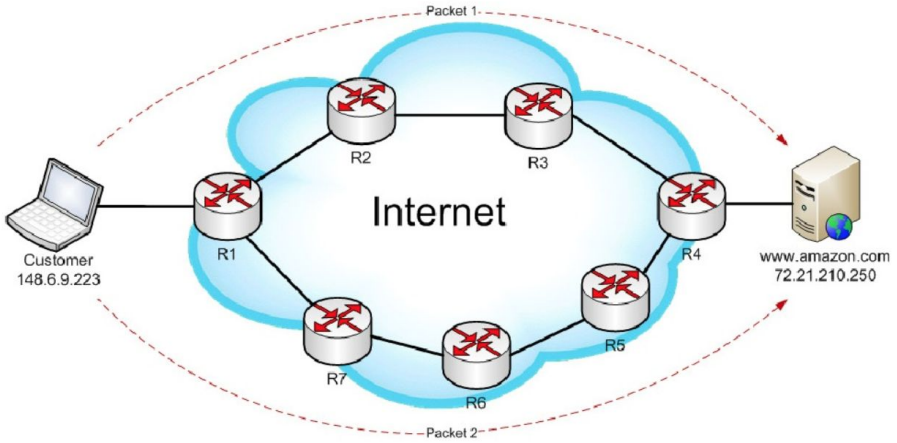
Internet Routing Security

Kenny Huang, Ph.D. 黃勝雄博士

CEO, Director of the board, TWNIC
huangk@twnic.net.tw
2018 Feb. 07

Circuit switching vs. packet switching



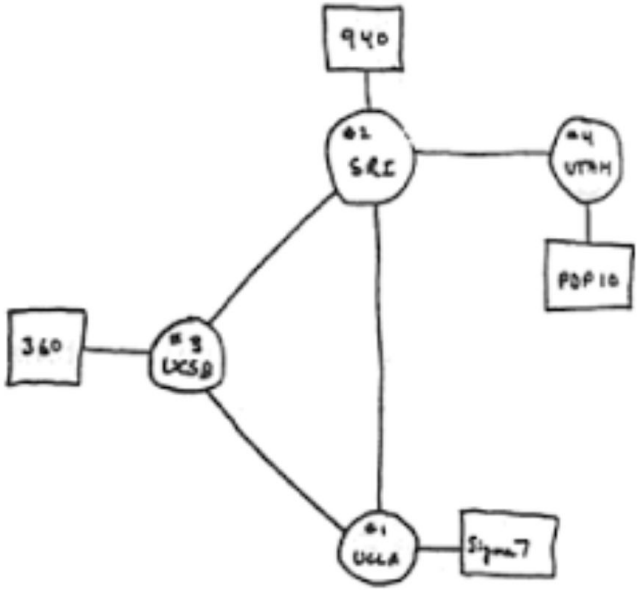


Forwarding decision ?

How do individual routers know how to perform the correct forwarding decision?

- Through knowledge of the topology state of the network
- The knowledge is maintained and distributed via routing protocols

ARPANET 1969

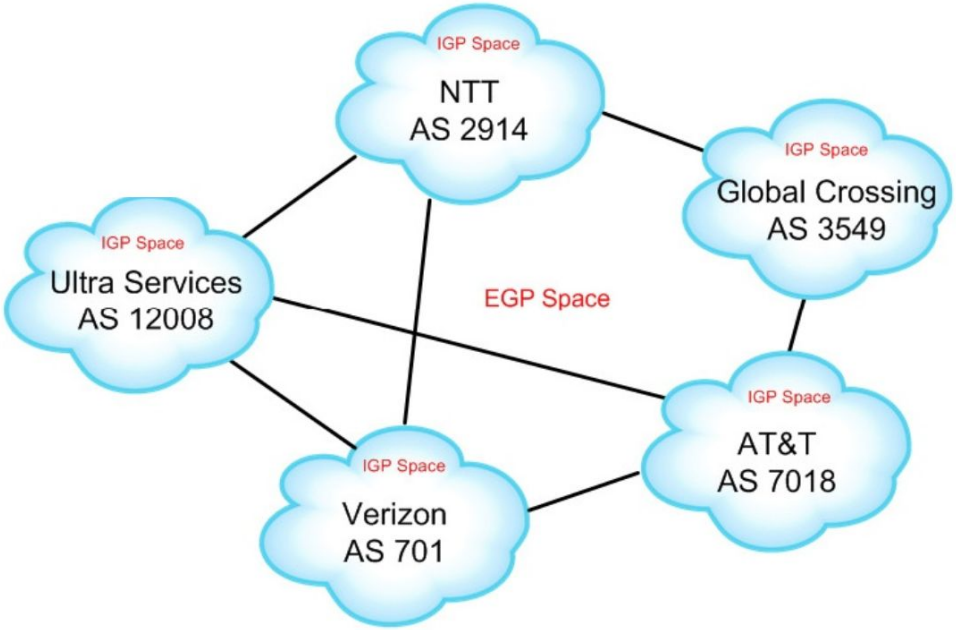


THE ARPA NETWORK

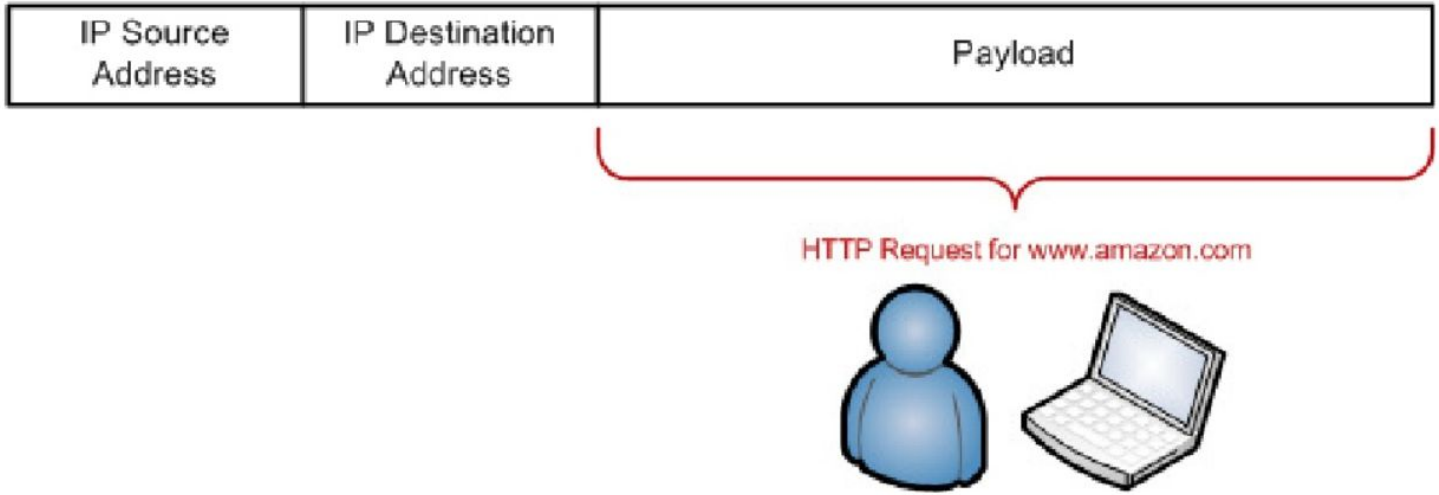
DEC 1969

Dynamic routing protocols

- Networks within an organizations are grouped into autonomous system (AS)
- An AS is a collection of IP routing prefixes under the control of a single administrative entity
- **IGPs** (Interior) are used to exchange routing information within a given AS
- **EGPs** (Exterior) are used to exchange routing information between routers bordering two networks

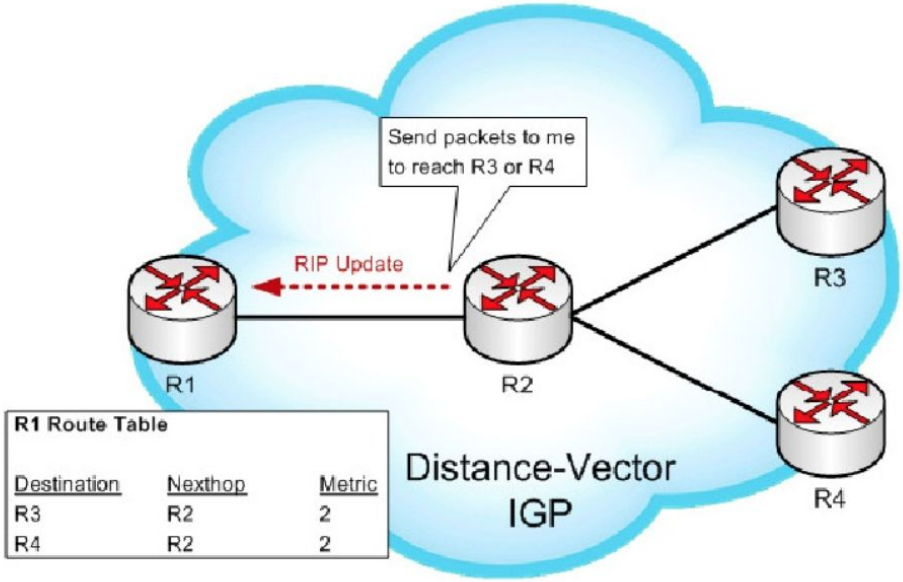


Internet Protocol



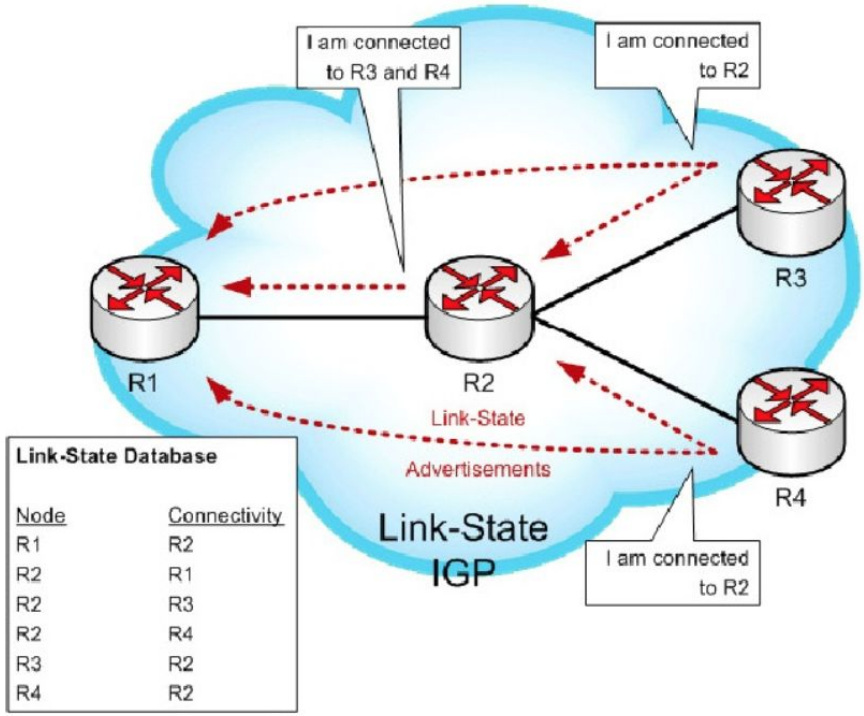
IGP - Distance Vector

- Distance vector routing protocols are referred as "**routing by rumor**" (e.g., RIP)
- Each routers knows very little about the overall network topology
- They only see their directly connected neighbors and "**trust**' that what is being advertised is reachable

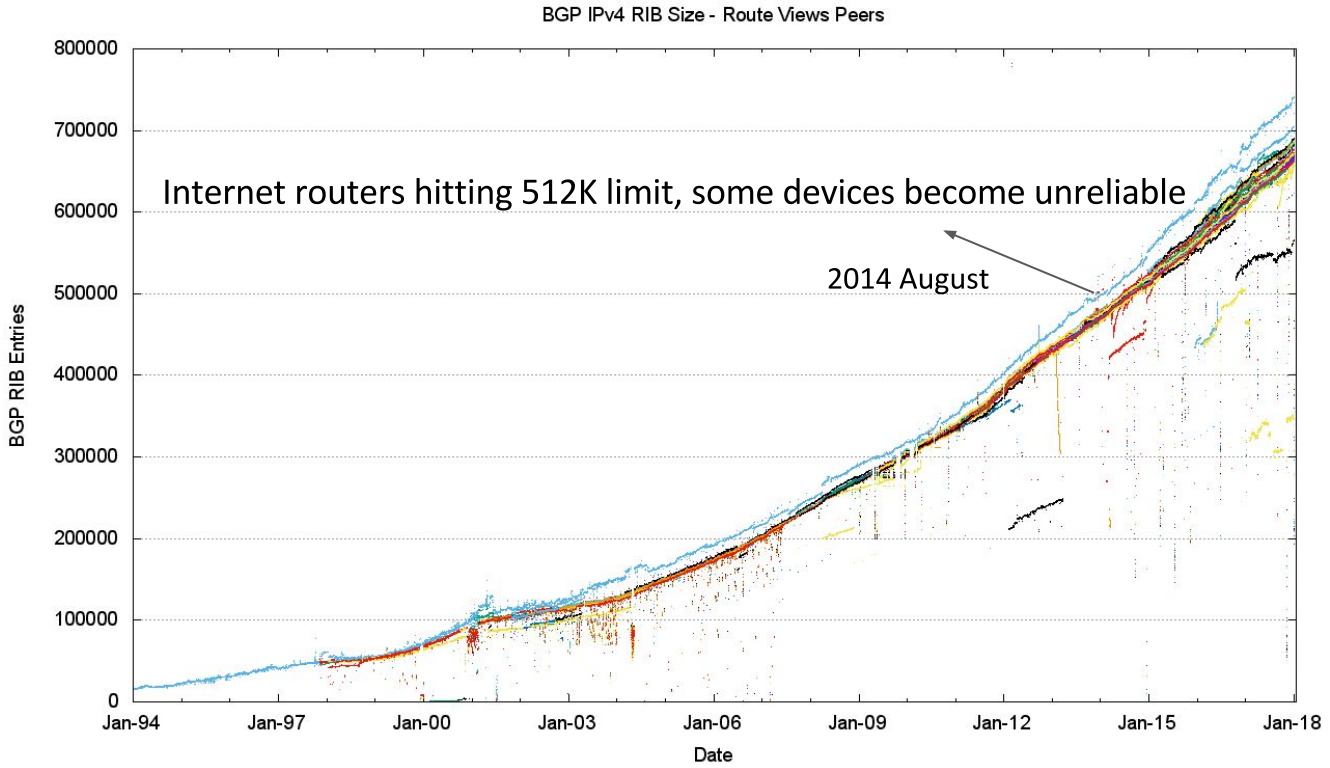


IGP - LINK STATE

- Topology information is flooded throughout the entire network
- Each router has their own conclusion as to what path is the best path (e.g. OSPF)

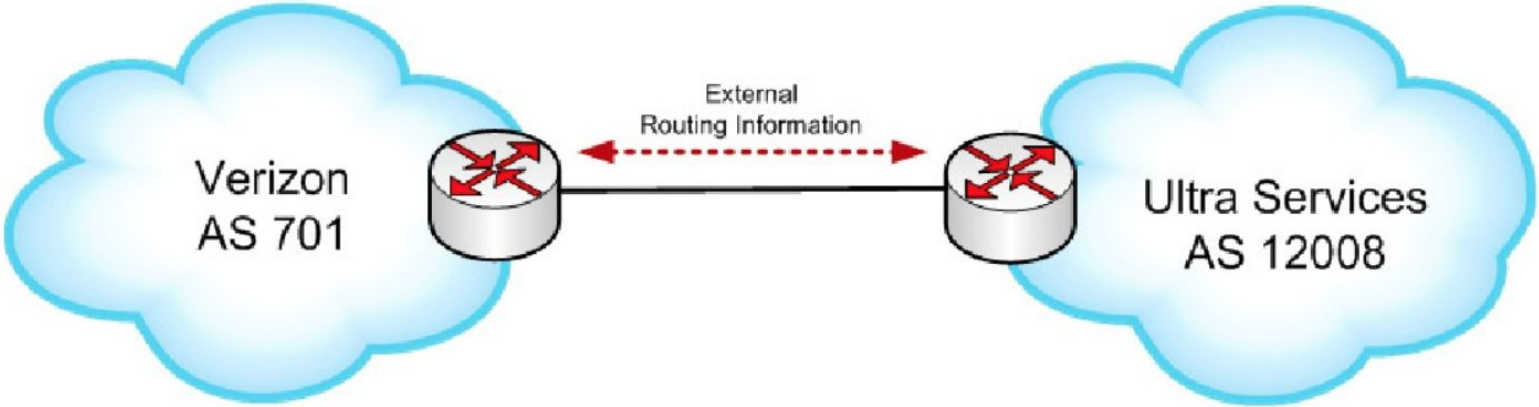


Growth of Internet routing table

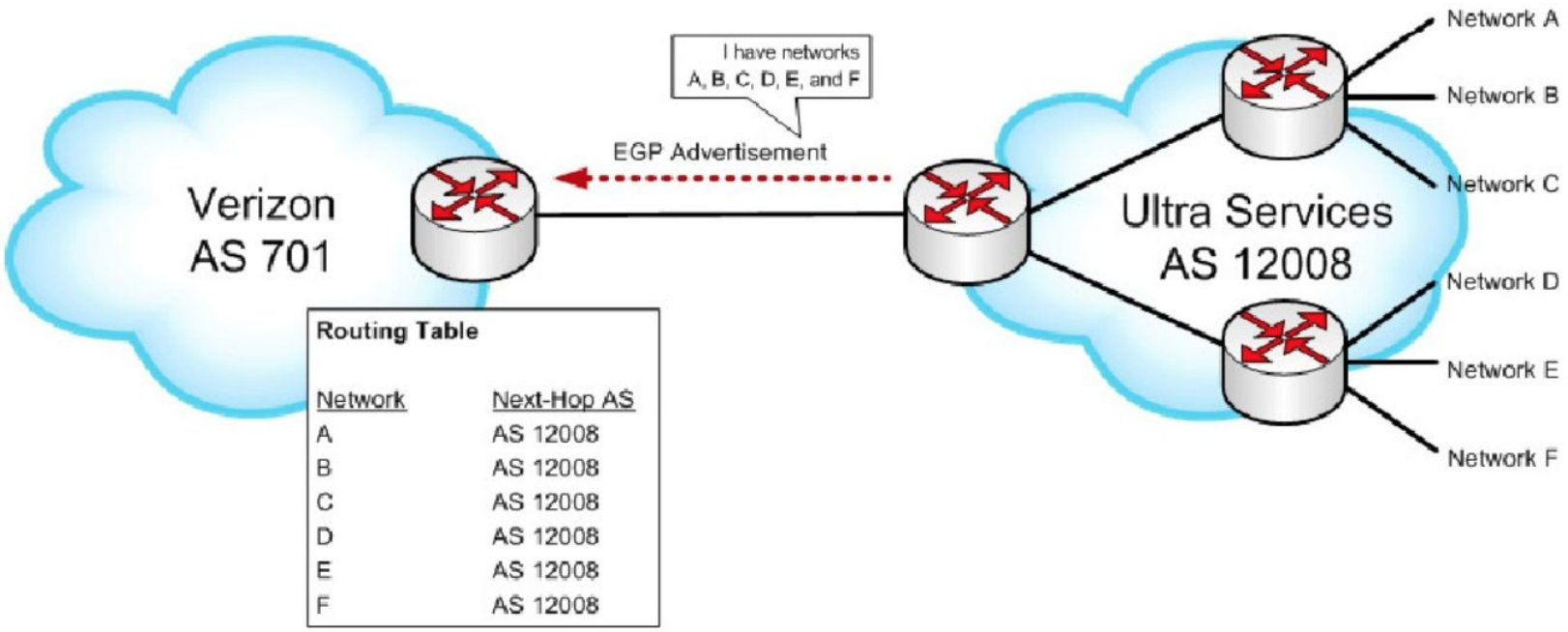


EGP

- used between Autonomous systems to covey external routing information
- Used to apply routing policy, e.g, policy based control routing vs. shortest cost path (e.g., EGP, BGP)

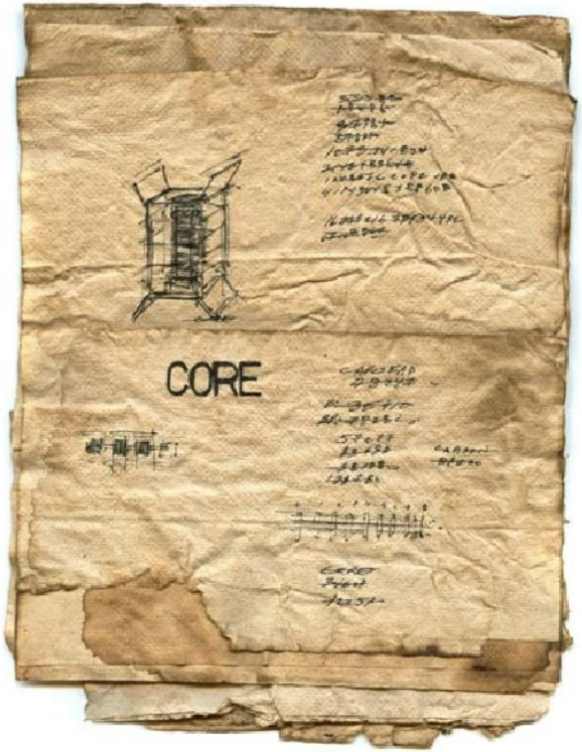


EGP


I have networks
A, B, C, D, E, and F
EGP Advertisement
Verizon
AS 701
Ultra Services
AS 12008
Network A
Network B
Network C
Network D
Network E
Network F
Routing Table
Network
Next-Hop AS
A
AS 12008
B
AS 12008
C
AS 12008
D
AS 12008
E
AS 12008
F
AS 12008

BGP - IETF12 (1989)

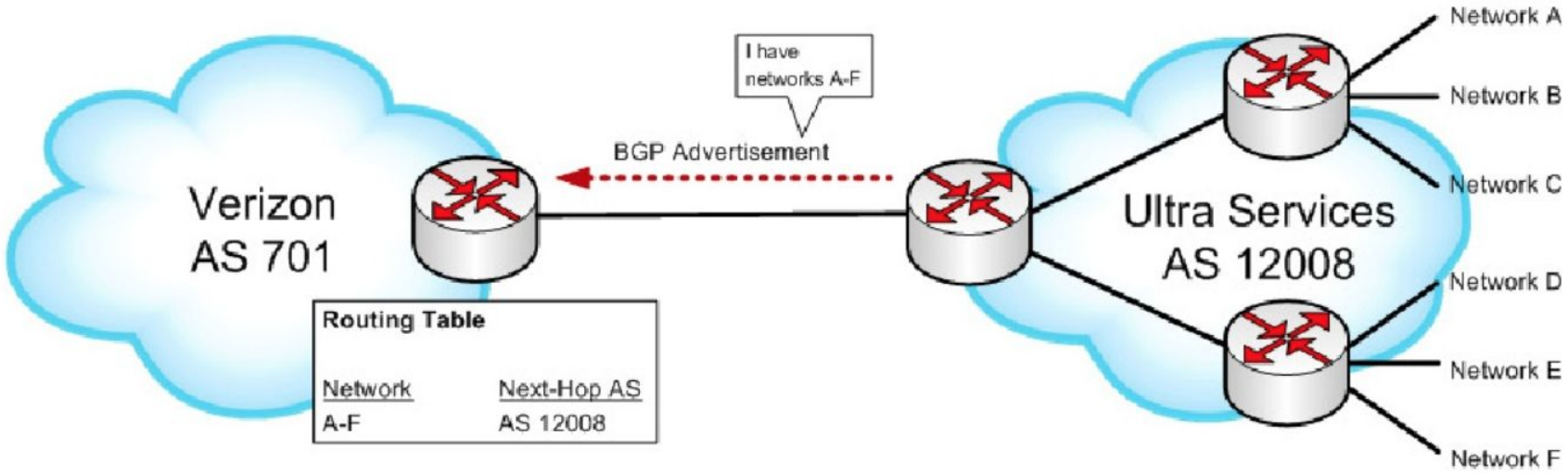
- Over lunch at IETF12 in 1989, Len Bosack, Krik Lougheed, Yakov Rekhter devised BGP
- The design goals behind BGP were to develop a protocol capable of providing policy control, loop detection, and the scalability required to support hundreds of thousands of networks through address aggregation



...the original idea for BGP was written on three napkins, giving BGP its unofficial title as the "Three Napkins Protocol"

BGP

- BGP provides loop avoidance, address aggregation
- BGP's most serious shortcoming is that it's up to network admins to check and filter information in route advertisements.

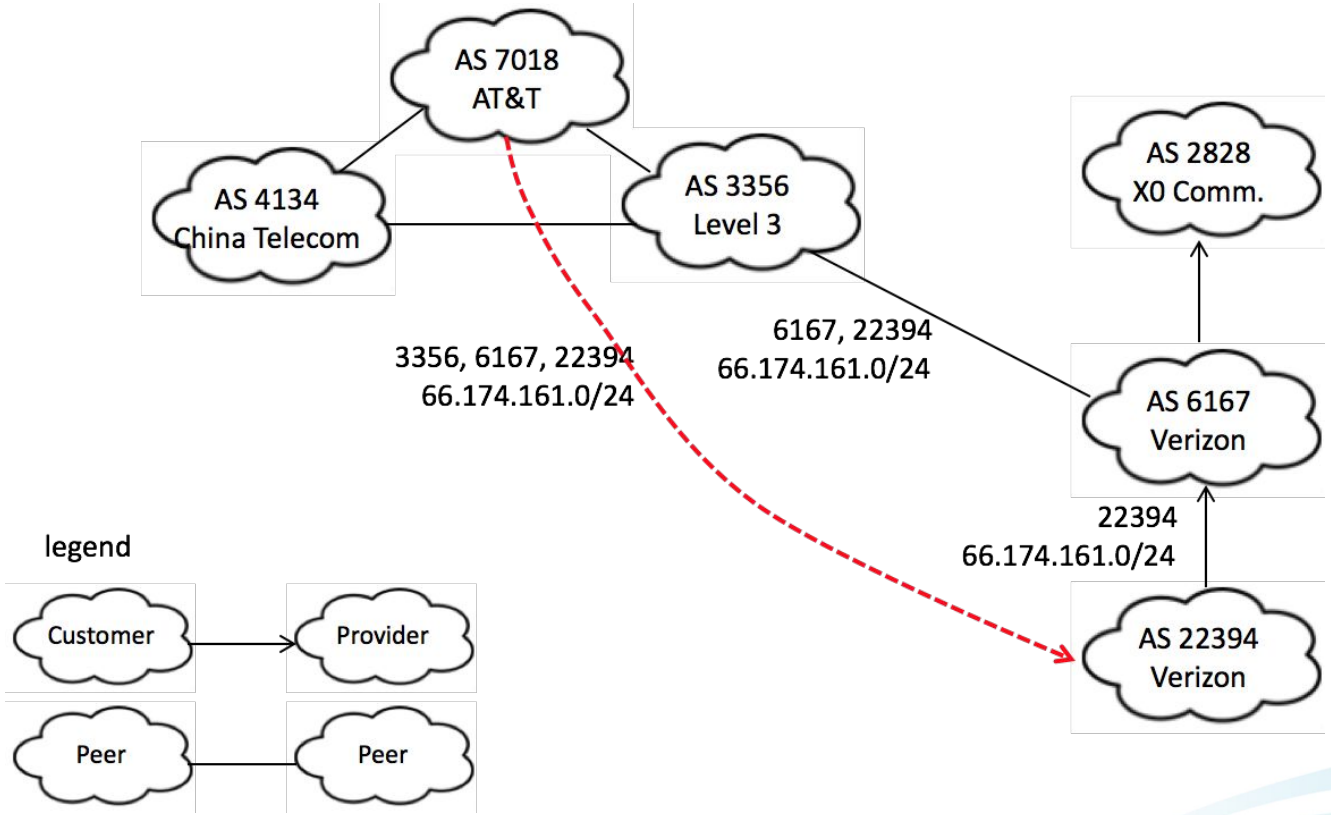


Internet routing security - detour

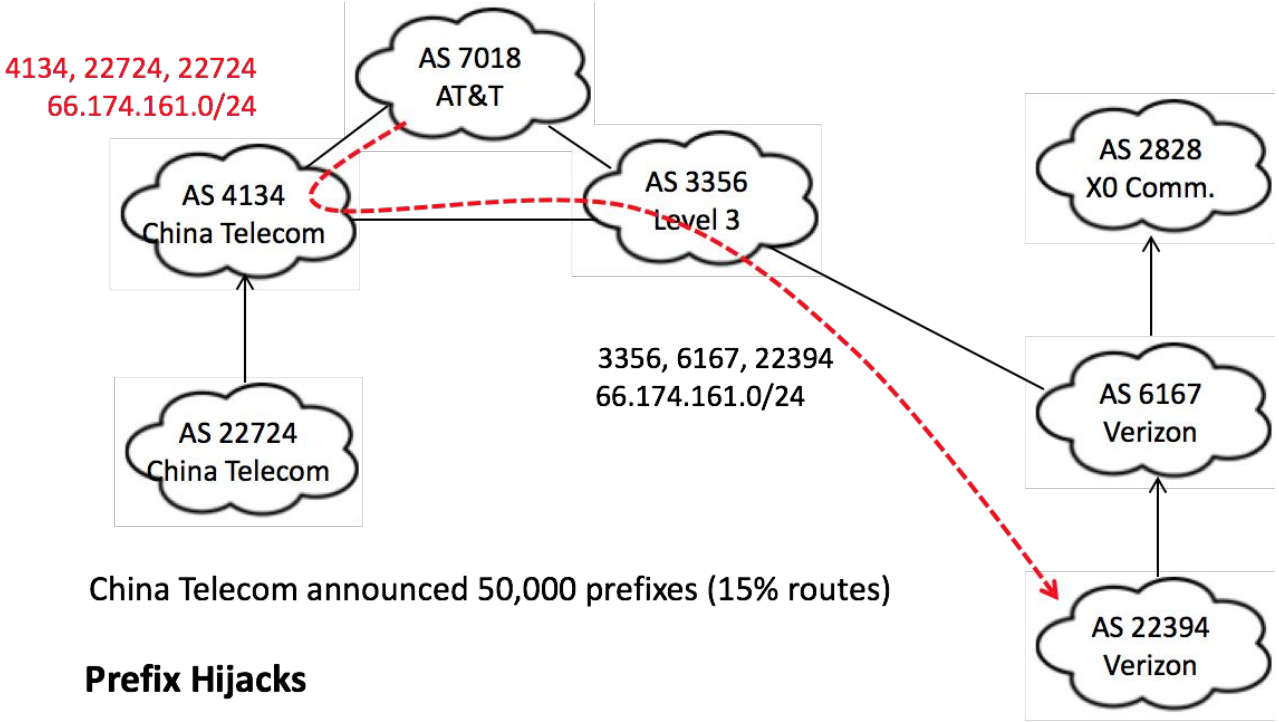
Detour Example

A path that originates in one country, cross international boundaries and returns back to origin country

BGP ROUTING



China telecom hijacks verizon wireless

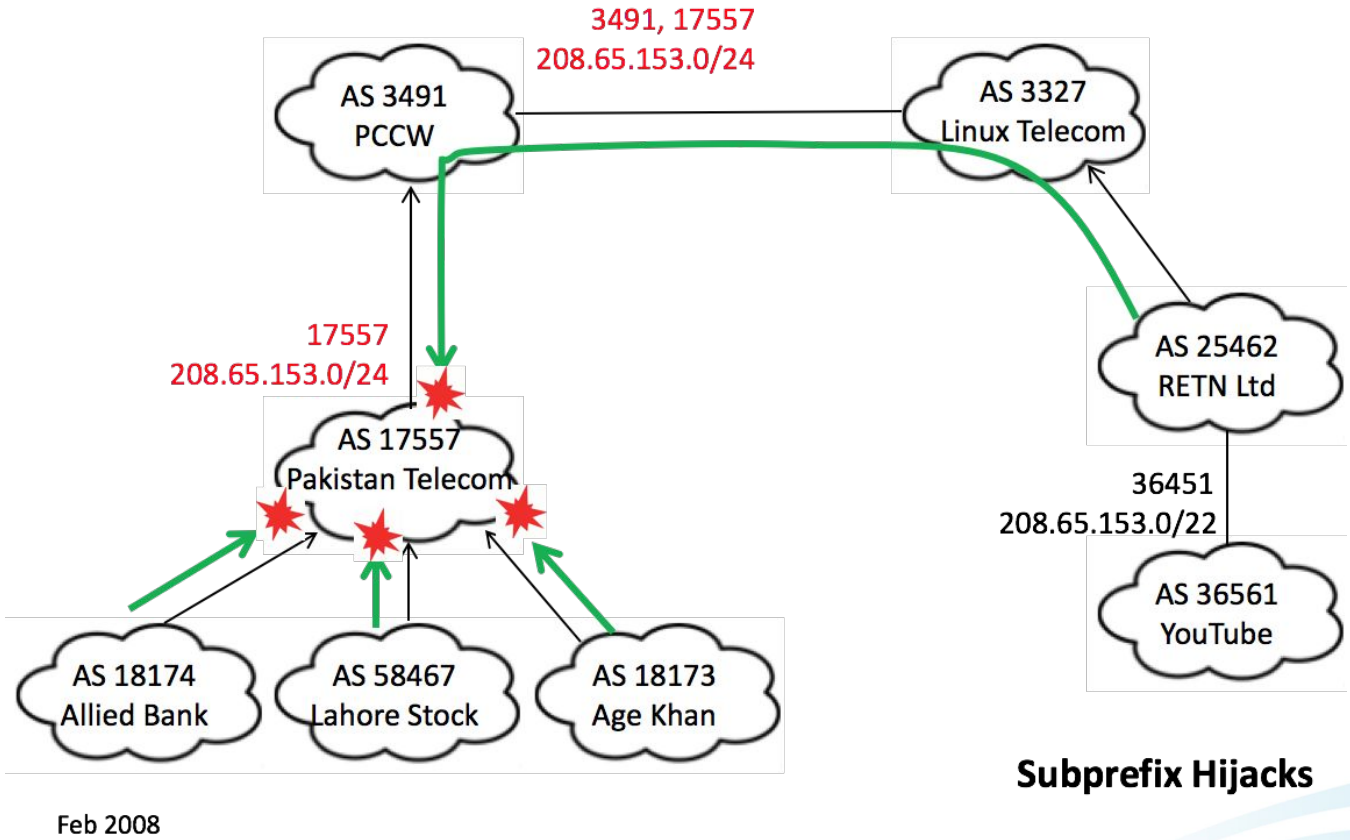


China Telecom announced 50,000 prefixes (15% routes)

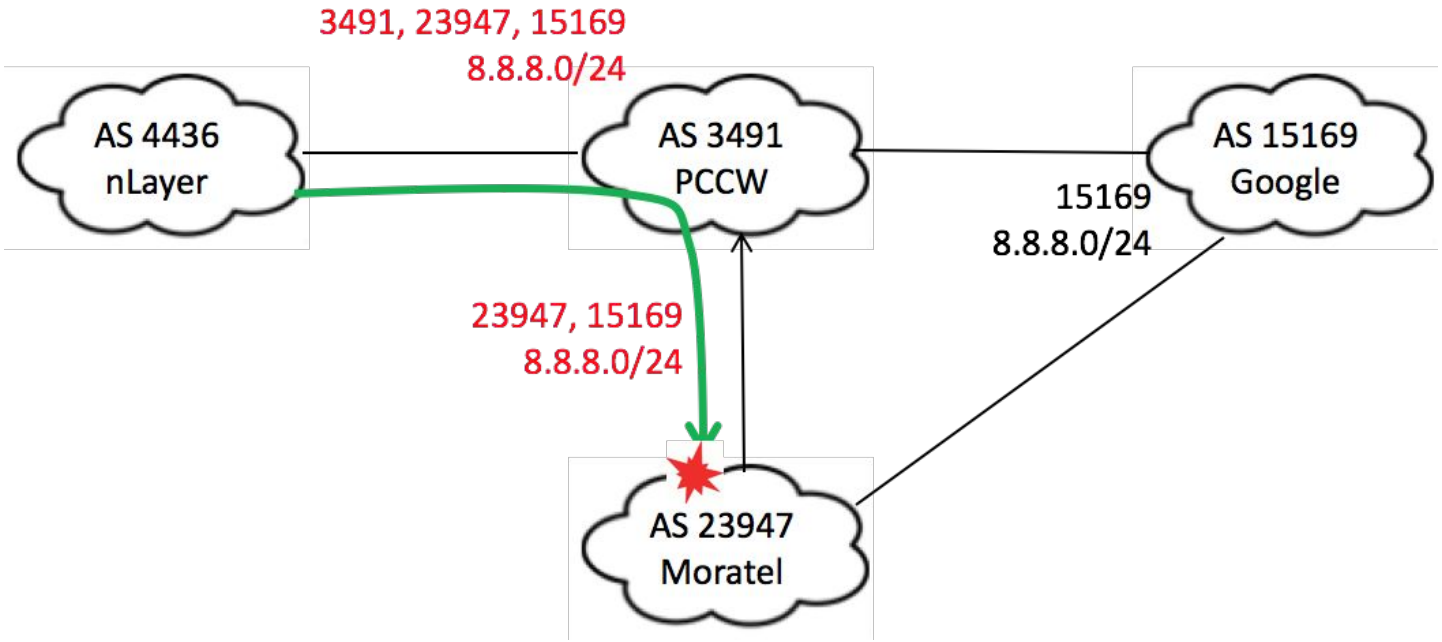
Prefix Hijacks

Apr, 2010

Pakistan telecom hijacks youtube

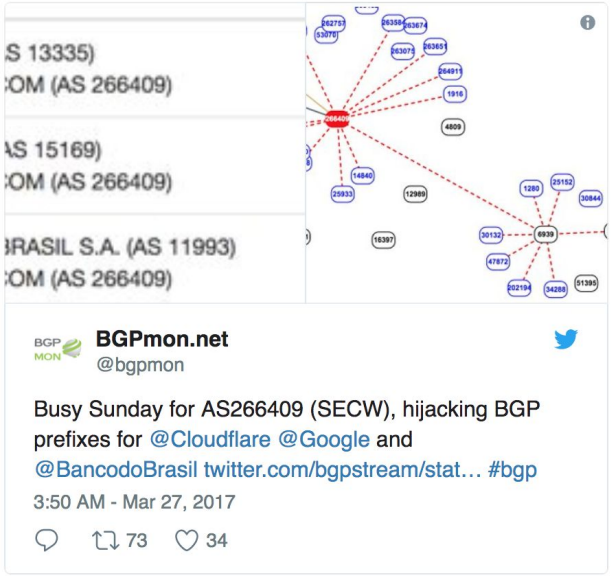


Moratel leaks a route to PCCW



2017 Routing security

- March 2017
  - **SECW Telecom** in Brazil hijacked prefixes of **Cloudflare, Google, and BancoBrazil** causing some outage for these services in the region.



Source : BGPmon, APNIC

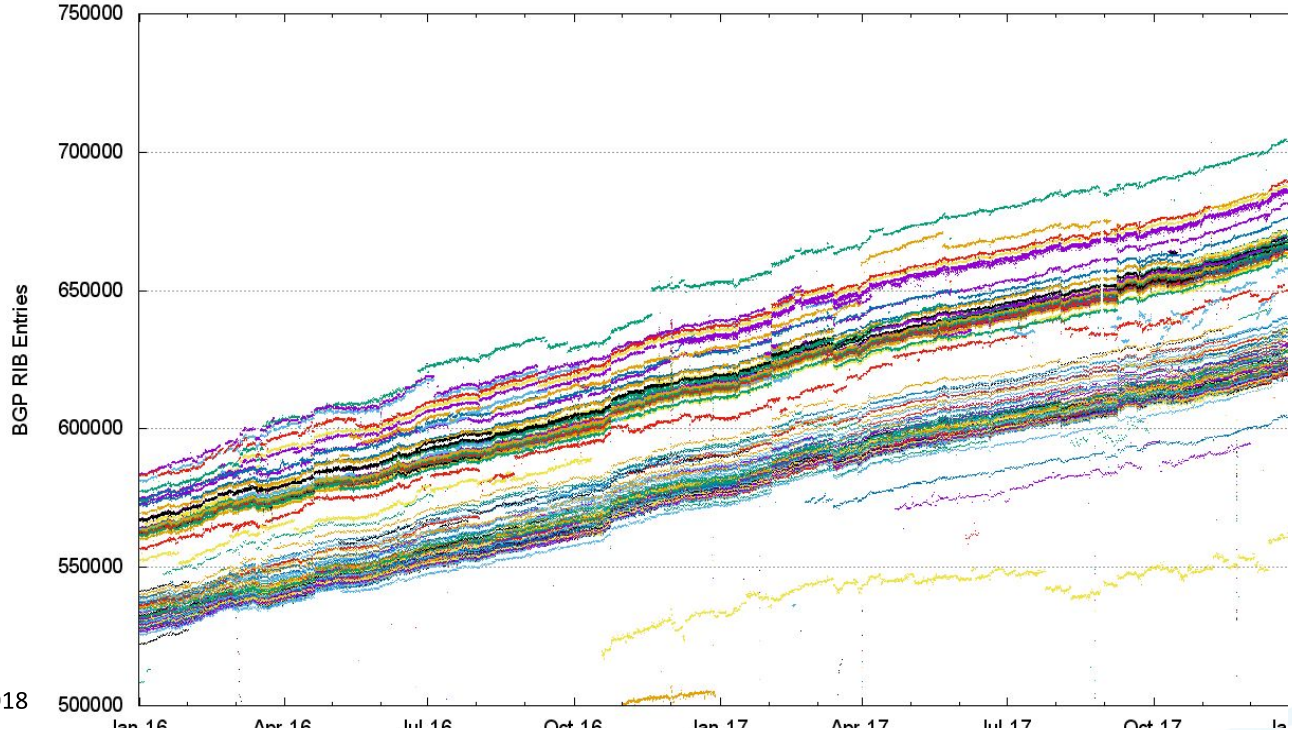
2017 Routing security

- April
  - **MasterCard, Visa, and more than two dozen other financial services** companies were briefly routed through a Russian telecom
- August
  - **Google** leaked 135,000 routes to Verizon, the result of which was traffic from Japanese giants like NTT and KDDI was sent to Google to transit.
  - **Nintendo** website crashed, same as the following
  - **Resona Bank, Saitama Resona Bank and Kinki Osaka Bank** Internet banking system
  - **JR East** Suica card payment system
  - **Twitter** japan
  - SecuAvail Inc. stock market price jump 10.07%

2017 Routing security

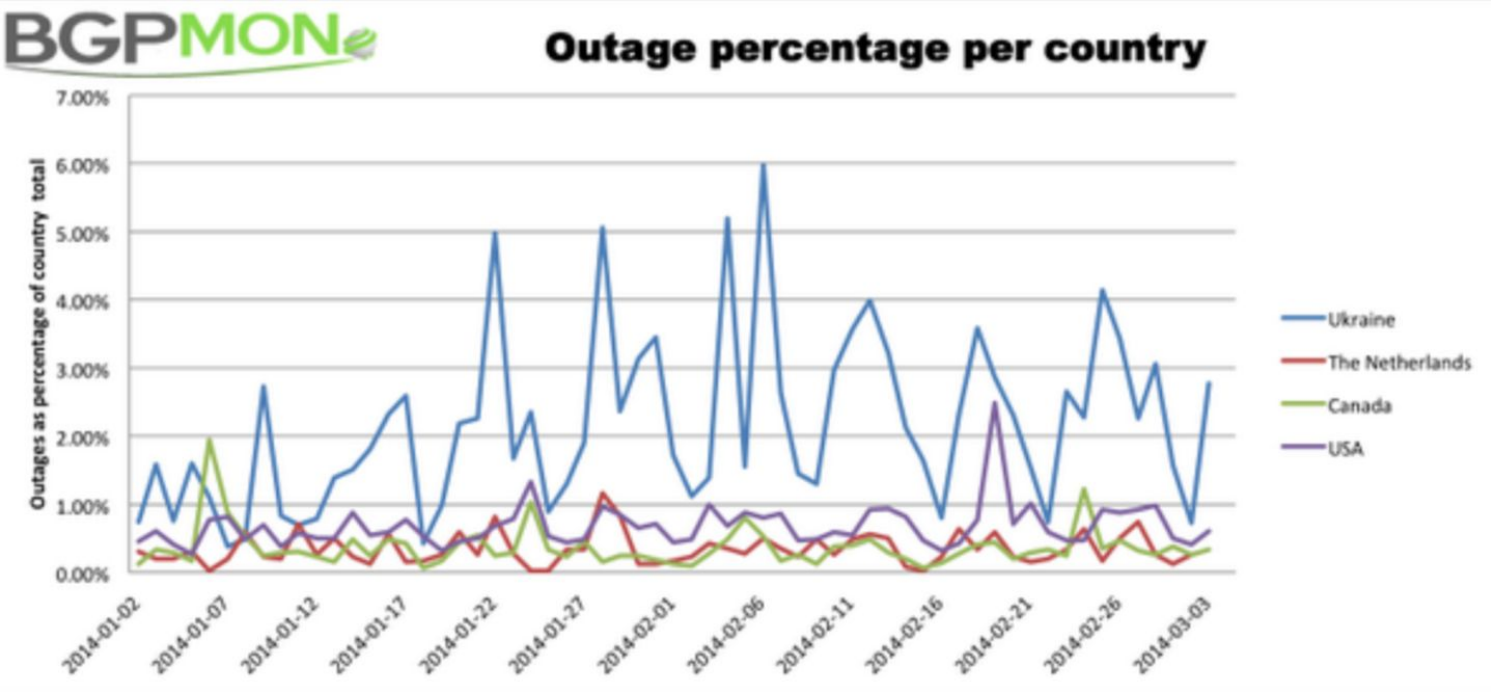
- October
  - US based network AS33362, would have sent traffic to **Google** via 6939 (**HE**) to 16735 (**Algar Telecom**, Brazil), to 263361 infovale telecom
  - Twitter, Google, and others large CDNs was rerouted through Brazil for more than 20 minutes
- November
  - BGP route leaks between **Level 3** and **Comcast** caused large scale network outage in North America for more than 90 minutes
- December
  - **Google, Apple, Facebook, Microsoft, Twitch, NTT** were rerouted to a previously unused Russian Autonomous System

- No single authoritative view of the Internet's inter-domain routing table.
- The collective management of the routing system as a single entity could be seen as an instance of a "tragedy of the commons"



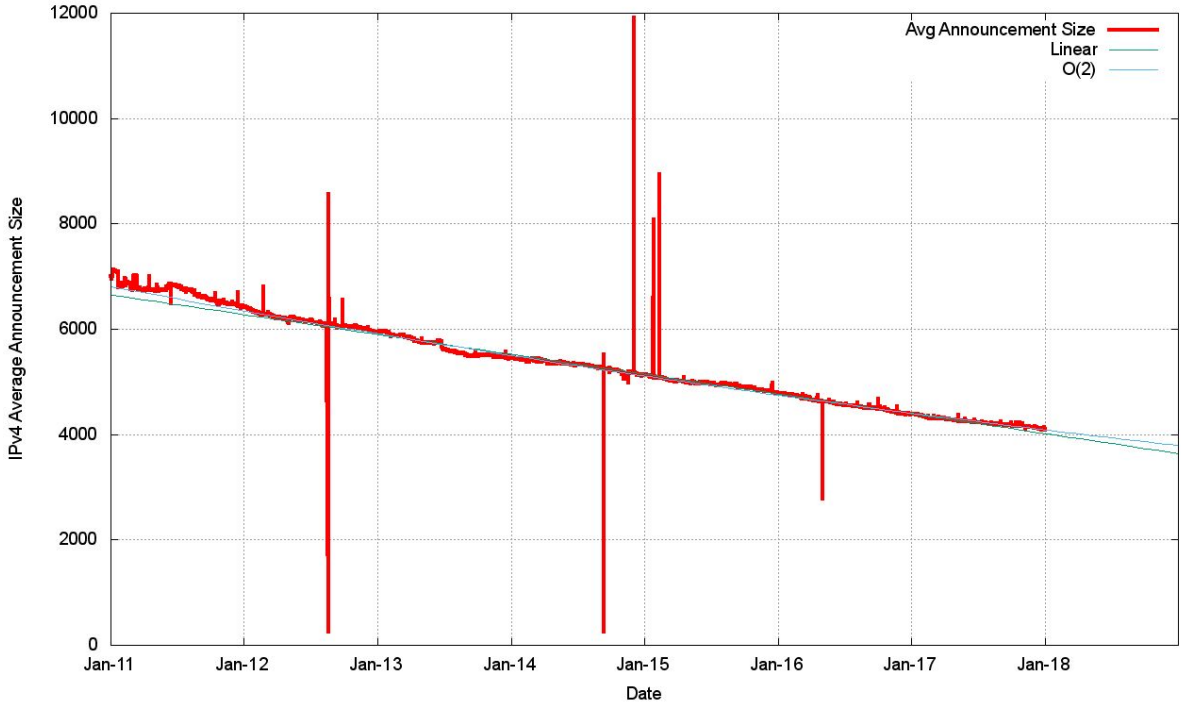
Source : APNIC, 2018

On August 8th, 2014, the global Internet routing table had passed 512,000 routes. It caused certain aging hardware platforms exceed the default routing TCAM allocations limit.

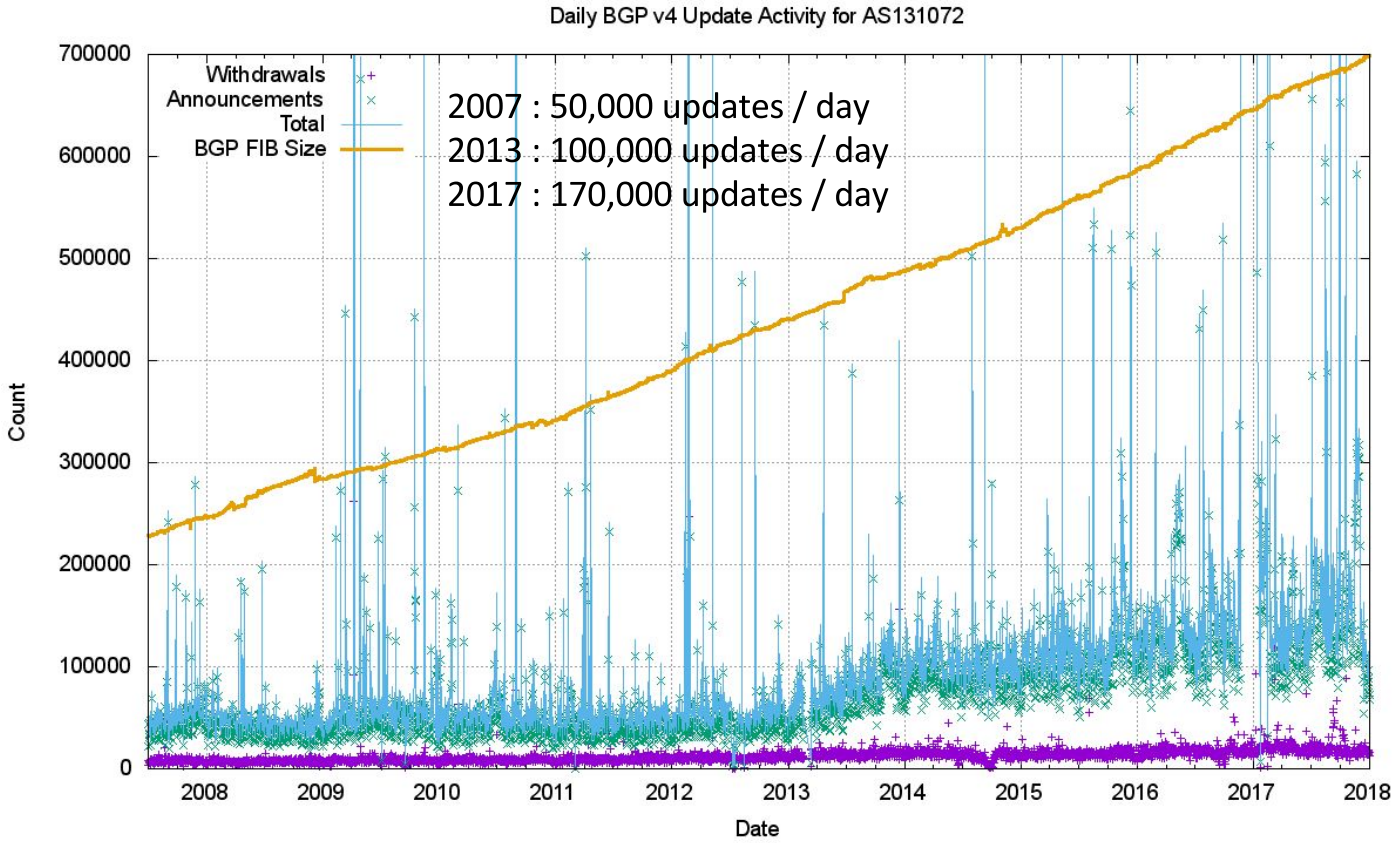


Source : BGPMON, 2014

The average BGP announcement size has dropped from 7,000 host addresses at the start of 2011 to 4,000 addresses at the end of 2017



Source : APNIC, 2018



Source : APNIC, 2018

	IPv4 Table	IPv4 Prediction
Jan 2014	488,011	
Jan 2015	529,806	
Jan 2016	586,879	
Jan 2017	645,974	
Jan 2018	698,714	693,000
Jan 2019		744,000
Jan 2020		795,000
Jan 2021		846,000
Jan 2022		897,000
Jan 2023		948,000

	IPv6 Table	IPv6 Prediction (linear)	IPv6 Prediction (exponential)
Jan 2014	16,158		
Jan 2015	20,976		
Jan 2016	27,241		
Jan 2017	37,469		
Jan 2018	45,388	40,842	47,314
Jan 2019		46,876	62,771
Jan 2020		52,910	83,278
Jan 2021		58,961	110,569
Jan 2022		64,995	146,690
Jan 2023		71,029	194,611

Source : APNIC, 2018

Trust for Internet Interconnection

- Internet trust behavior
  - Identification-based trust among networkers revolves around the idea that the Internet is “a people thing”
  - Interconnecting networks is equated with a personal relationship.
- Evaluation factors for trust
  - People
  - Expertise
  - Demonstrated quality (24x7)
- Social connection as a currency
  - The Internet is 40,000 competitors. But if they don’t work together, then none of them has a product
  - Damaging connectivity is against every operator’s self-interest. This reasoning evokes a very basic level of confidence in each other. It may not be high, but it is widely shared.

Trust and distrust dual-stack model

- Internet interconnection / peering
  - personal relationships count as substitutes for the exchange of money if an interconnection is established without a contractual relationship.
  - the best way of keeping that relationship is by having some kind of a social connection which acts as a currency
- Trust and distrust
  - You can trust your spouse, not the peering partner.
  - If you betray the trust, if you lie and say: ‘Well, I am YouTube’, then the rest of the Internet is going to come down on you
  - It is yet unclear how common the use of such vigilante-justice sanctioning mechanisms is, but this quote shows how easily identification-based trust can turn into distrust.

Current practice for interconnection relationship

- Build trust before commercial arrangement
- Do we need a contract for interconnection?
  - Case : “A company “
  - Peering with more than 1,000 networks
  - Less than 10 peering agreements have been signed
  - 99% peering relationship is based on trust

Coping with distrust

- Detect irregularity
  - Operators engage heavily in monitoring their networks to detect irregularities
- Contractual relationship
  - initiating contracts and service level agreements, at least for meaningful interconnections.
  - Contracts allow networkers to create bounded transactions despite distrust.
- Unclear power of peering contracts
  - the power of peering contracts is unclear. Because no one is paying the other in free peering, there is probably no way to uphold what is in the contract in a court of law. Contracts certainly have the effect of solidifying a relationship.

The value of distrust

- Value of distrust
  - Distrust eliminate informal, personal aspects from Internet interconnection.
- Enforceability of distrust
  - distrust can become a productive force for the good of the Internet
  - Distrust needs to be transformed into improvements , such as RPKI and BGPsec
  - These ongoing routing security innovations are actually materialised expressions of distrust that ultimately are able to generate trust.

constant innovation

Applications:

Internet infrastructure:

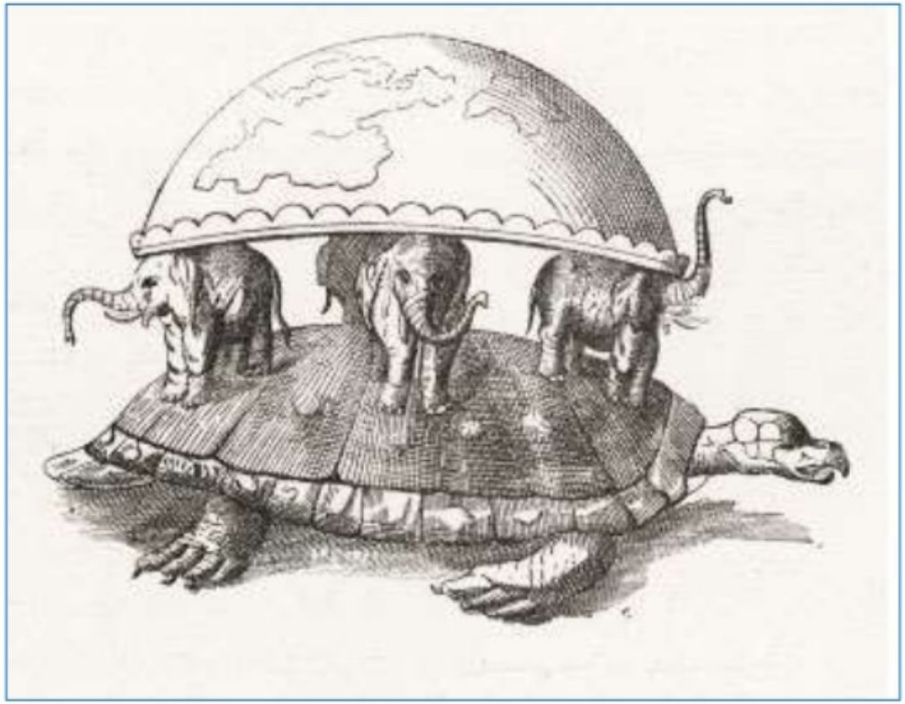
routing, congestion control, naming, …

(TCP/IP, BGP, DNS, OSPF, ECMP,…)

stagnant!

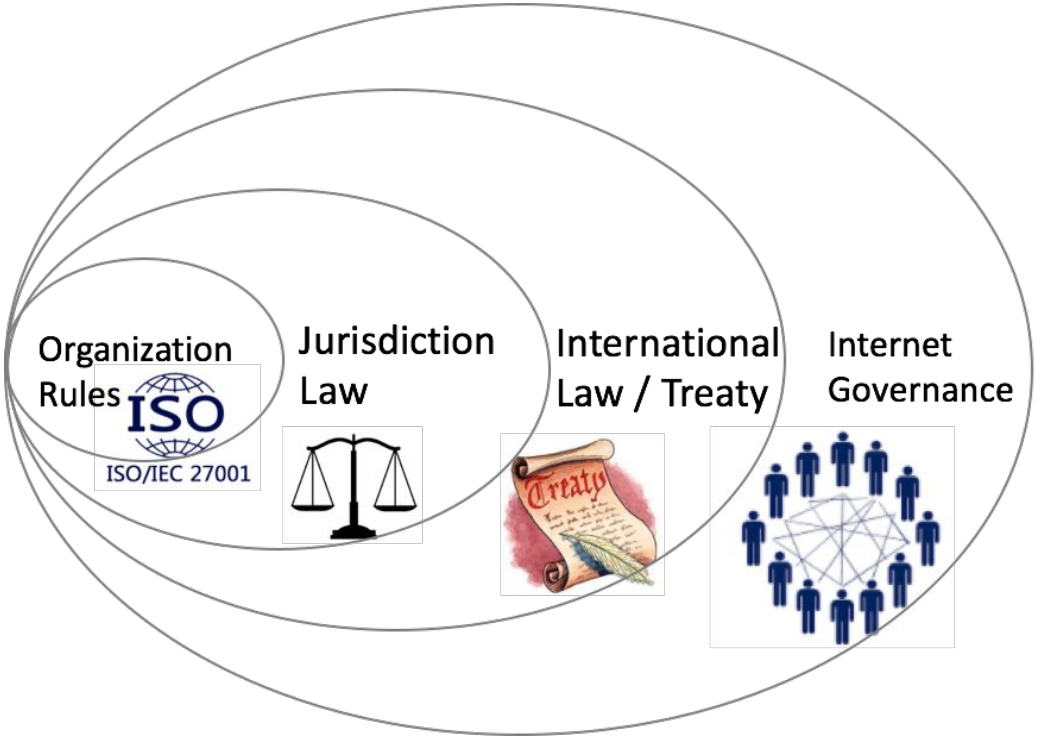
Technologies:

constant innovation



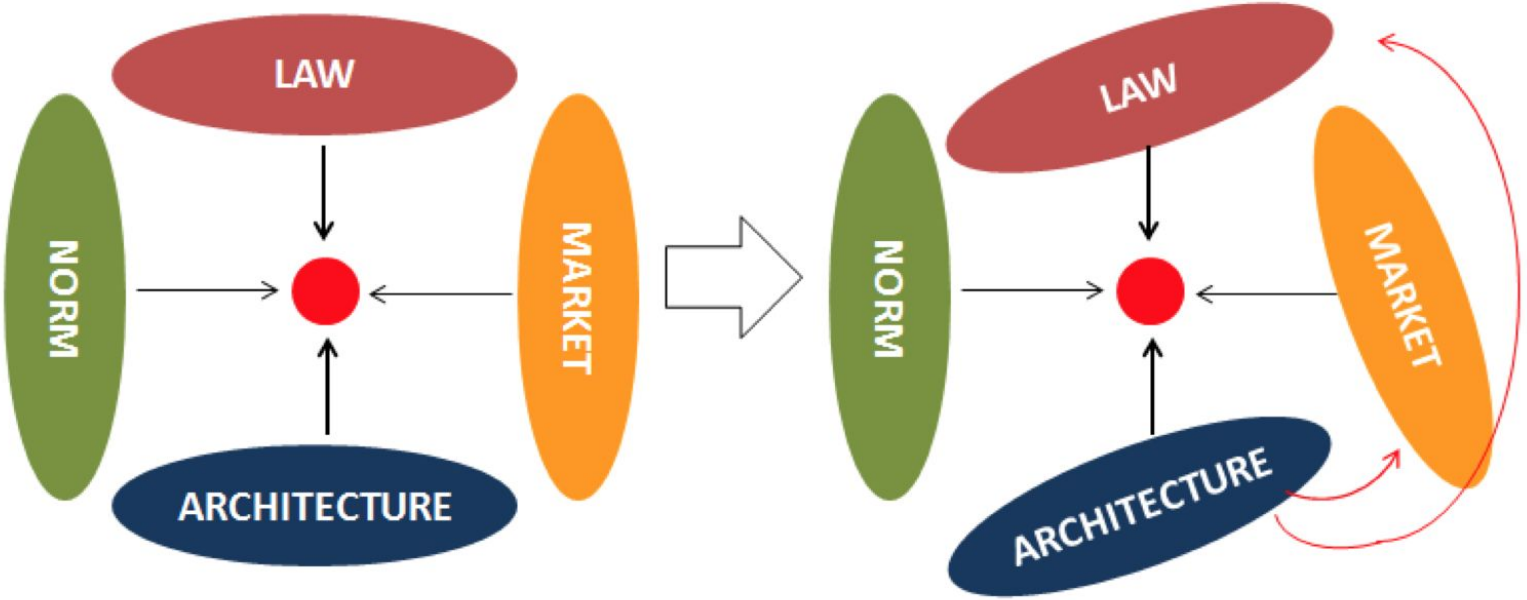
Source : Michael Schapira, 2017

Internet governance and cybersecurity



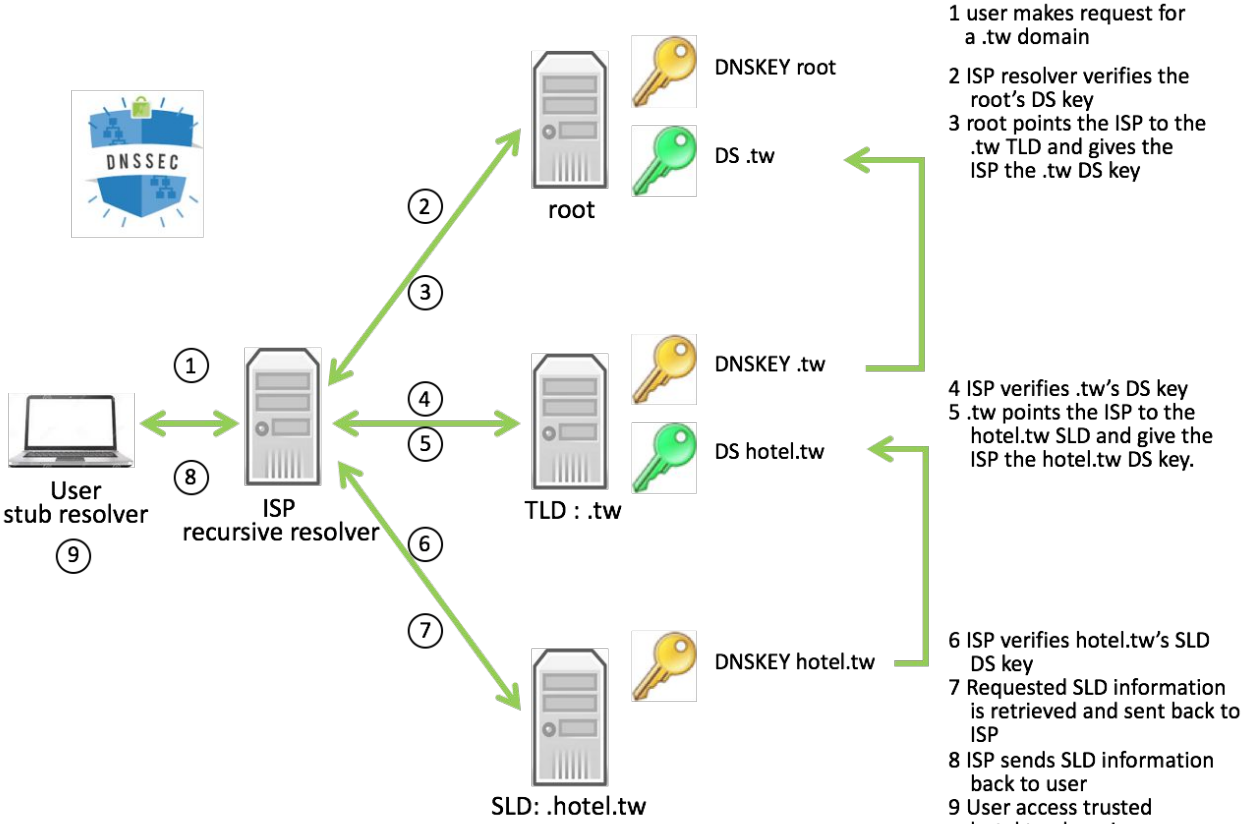
IG Regime
Multistakeholder
Standard
Technology
Architecture
Policy
Procedure
Best Practices
Cooperation
Coordination

Code is law



(Lawrence Lessig, 2000)

Secure namespace - DNSSEC



Secured Communication

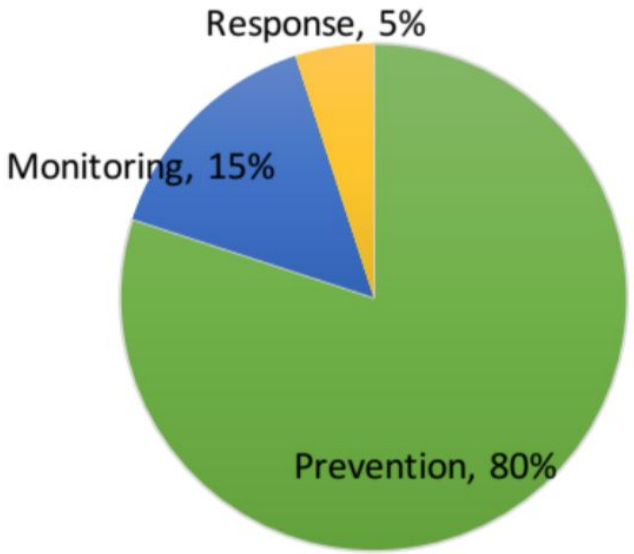
RFC 2409	The Internet Key Exchange (IKE)
RFC 3526	More Modular Exponential (MODP) Diffie-Hellman groups for Internet Key Exchange (IKE) Proposed DH1024
RFC 4307	Cryptographic Algorithm for Use in the Internet Key Exchange Version 2 (IKEv2) Proposed DH 2048

Remove support for DH1024

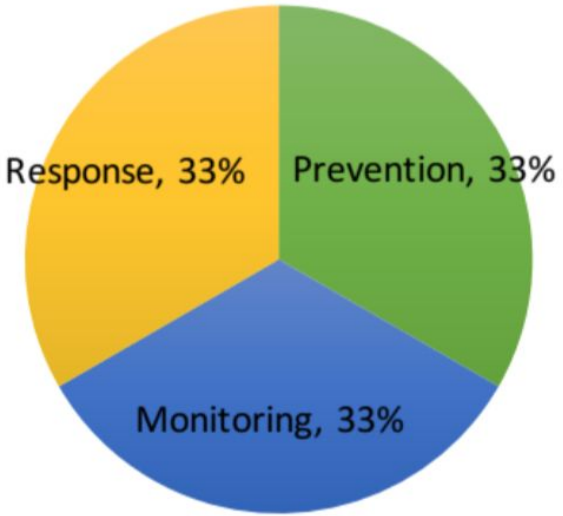
RFC 7258	Pervasive Monitoring Is an Attack
RFC 7457	Summarizing Known Attacks on Transport Layer Security (TLS) and Datagram (DTLS)
RFC 7525	Recommendations for Secure Use of Transport Layer Security (TLS) and Datagram Transport Layer Security (DTLS)

Evolution of Cybersecurity Strategy

Now



Future



Source : RSA Conference 2016 Singapore

Internet Routing Registry

- Routing registries are queried by upstream providers for
  - Route filters updates, ensuring stability and consistency of routing information shared via BGP
  - Better control on BGP traffic, example to avoid BOGONS.
- If ISPs don't have objects in a routing registry, then he need to create new objects to avoid being filtered by upstream providers
  - Based on the planned routing policy, other objects need to be created (AS-SET, ROUTE-SET)
  - For routing purposes not all objects are needed. It depends on the situation and routing policy

Commercial Routing registry database

- RADB
  - Routing Assets Database (Routing Arbiter Database)
  - Many upstream providers requested RADB registration
- Fee (Merit RADB)
  - For-profit Maintainer agrees to pay a fee of Four Hundred Ninety Five United States Dollars (USD$495) for each maintainer object to be registered in the RADb.
  - Non-profit Maintainer agrees to pay a fee of Three Hundred Ninety Five United States Dollars (USD$395) for each maintainer object to be registered in the RADb.

RPSL

- RPSL
  - Routing Policy Specification Language RFC2622
- Purpose of RPSL
  - RPSL was designed for Internet Service Providers (ISPs) to publish their routing policies. Since the introduction of RPSL, many ISPs publish their policies in public Internet Routing Registries. Such as RIR Routing Registry.

Fragmentation within RPSL

- Two primary sources: RIPE WHOIS and RADB
  - A kind of ‘europe’ / ‘rest of the world’
  - Content can conflict. Which one is right?
- Other sources, APNIC, AfriNIC, JPNIC … less globally applied
  - Content is visible in several sources
  - ISP specific (e.g. NTT) with automated customer-AS routing
  - National scope (JPIRR) with strong checks
- Lack of visible cohesion. What determines ground-truth
  - If IRR conflict ?
  - If IRR are incomplete?
  - If IRR include data with no visible linkage to origin assigning registry?

APNIC Internet Routing Registry (WHOIS)

- Whois
  - The APNIC Whois Database can be used to publish information about the routing of Internet number resources
- Maintain route filters
  - Using IRR to manage route filters typically requires use of software like the irrtoolset or rtconfig.
- Routing policies
  - Simple routing assertion can be made by creation of inetnum/inet6num/aut-num and associated route objects with no optional data
  - Recording network routing policy by using the policy specification language (RPSL). (RFC2622; RFC2650)

Benefit of APNIC WHOIS

- Free
- Easy maintenance
  - Same set of objects are used (aut-num, maintainer ..etc)
- Security
  - Route objects are tied to aut-num; created only by APNIC Hostmasters
  - Only “holder” of prefixes can create route objects for given inetnum
  - Considerable reduced risk of hijacking

What is RPKI

- Chain of resource allocation certification
- APNIC allocates address resources to NIR (TWNIC). TWNIC allocates address resources to ISPs or IP members.
- With RPKI, APNIC certifies for TWNIC. TWNIC certifies for ISPs or IP members

Who

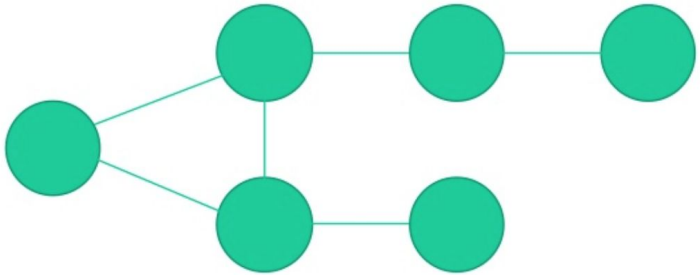
- IETF (SIDR WG, IDR WG), routing and security area
  - Technical specifications, researchers, implementers
- RIR, TWNIC and ISP
  - Database, registration, cryptographic key management
  - Resource certification for resource holder
- BGP routing operators, researchers, web service operators
  - Input security requirements into IETF

When

Year	chronicle
1997	The basic concept of RPKI and Secure BGP
1999	Technical draft (I-D) for resource certification
2004	RFC3779
2006	The first SIDR WG meeting
2008	YouTube incident
2009	Some of RIR started resource certification
2012	RPKI capable BGP routers

Where

- Routing table in BGP routers
  - Misused IP address information is propagated between AS.
  - Depends on network topology
- Affected area : reachability
  - E.g., Global web service



Why

- Why IP address is misused?
  - Faking users of DNS server, web server, etc.
  - For sending SPAM from temporary IP address
- Why resource certification?
  - Detecting allocated/assigned IP address prefix + finding correct AS who should use the IP address = misused IP address is detected.
- Why RPKI is good to deploy
  - Misused IP address can be found with origin validation. It is good inputs for BGP operation avoiding receiving malicious traffic (SPAM. etc).
    - BGP operators can find own IP address faked without RPKI. But with RPKI, for other also can be.

How

- How misused IP address in BGP routing are conducted?
  - By configuring BGP routers
    - Fat finger, mis-typing
    - Intentional use of victim's IP address or unused IP address
- How RPKI is used?
  - Registry issues resource certificate and LIR creates ROA
  - BGP operators compare ROA and route.
  - BGP operators change their router configuration

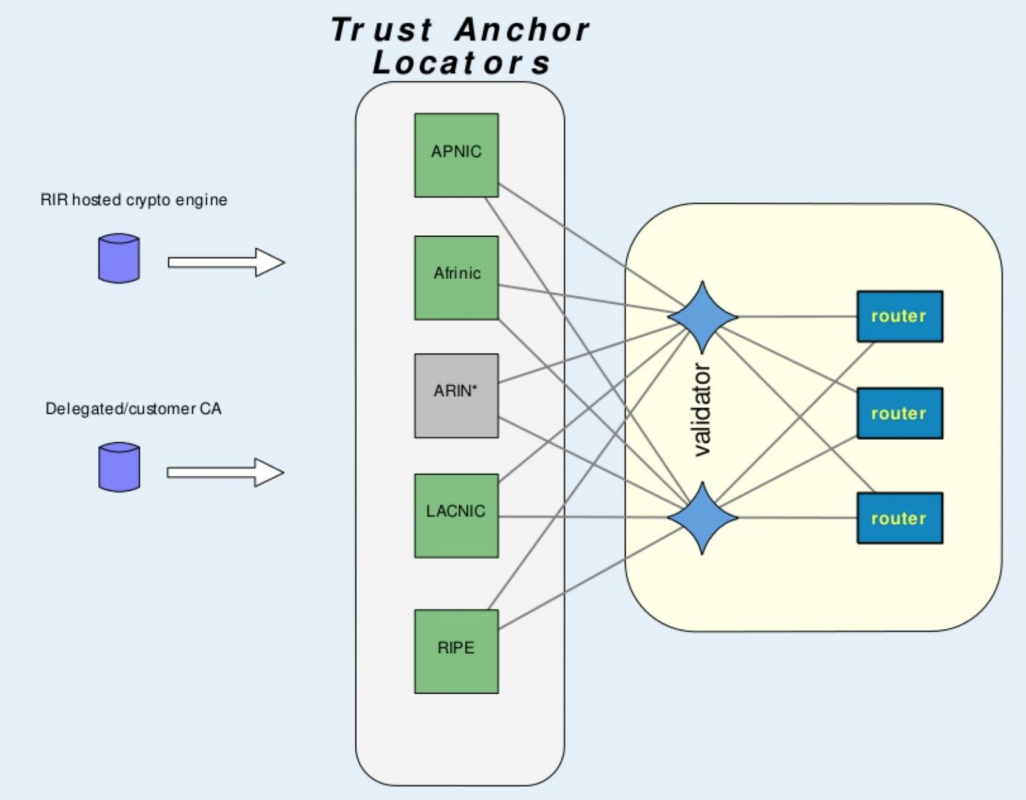
How much

- Does misused IP address have any economic value ?
  - Yes, recent example: an advertisement area in a website has paybacks when unreachable from Internet. Untraceable zombie hosts for SPAM email sending.
- The free of charge for getting resource certificate and ROA currently
  - Operational cost in the registries, for giving certificates to provide information integrity
- Operational cost in BGP routing
  - When BGP operators face a misused IP address, they do their incident action. Investigation, testing, negotiation with their peer.

ROA Content

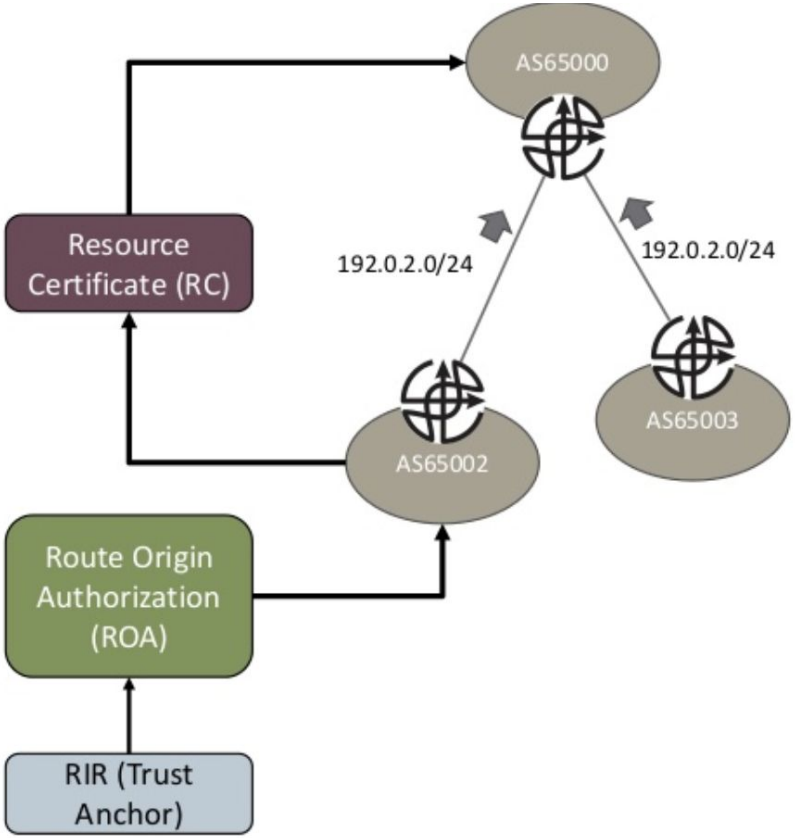
- It is able to be used to find misused IP address in the Internet
  - Origin validation = validation of BGP routes
  - whether IP address prefix is used properly in Internet routing
- ROA content
  - Original autonomous system number
  - Prefix
  - Validity dates
  - When a ROA is signed, it has a cryptographically provable chain to the source of authority allowing that IP to be advertised by that ASN

What it look like



Origin Authentication

- The RIR authorizes AS65002 to originate 192.0.2.0/24
- AS65002 creates a Resource Certificate (RC) signed with a private key and any additional parameters
  - Can be a longer length, etc.
  - Longest prefix within this block
- AS65002 places this in the RPKI database
- AS65000 can verify the origin AS against the RPKI database

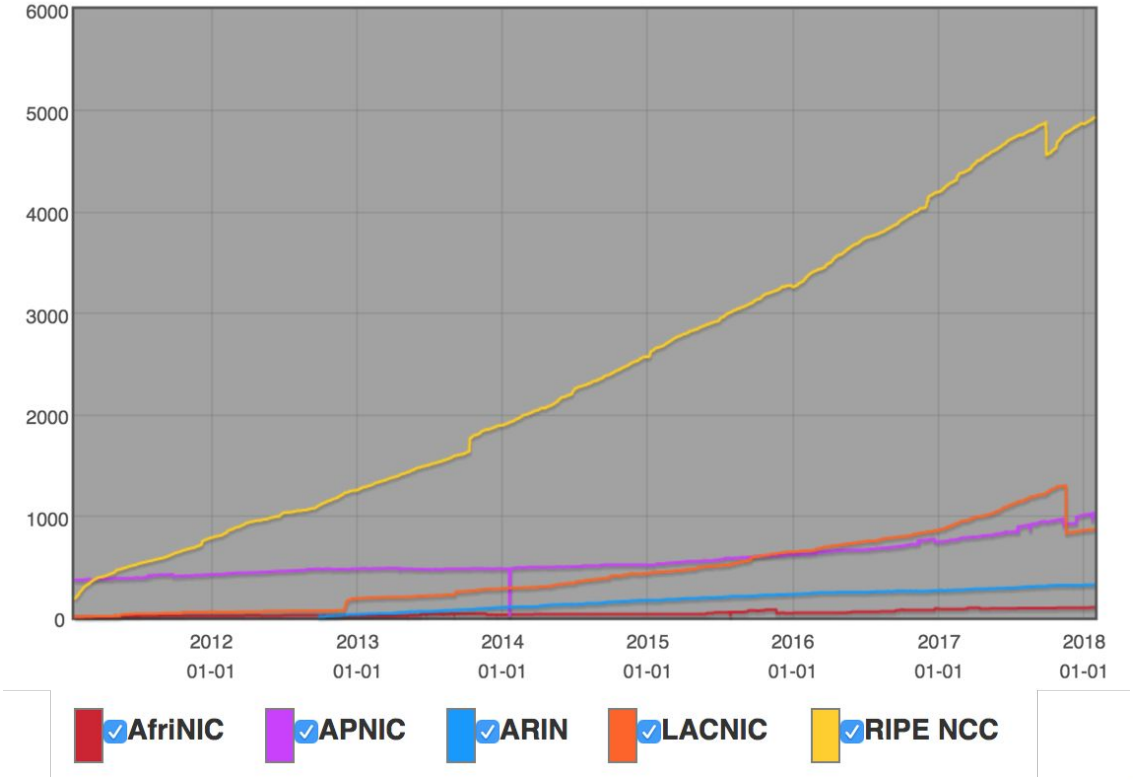


Coverage for RPKI and RPSL

Technology	% covered in BGP (apnic)	% Overlap RPSL/RPKI (apnic)
RPSL v4	75% (81%)	93% (97%) rpki in rpsl
RPSL v6	91% (67%)	94% (82%) rpki in rpsl
RPKI v4	12% (2%)	12% (15%) rpsl in rpki
RPKI v6	9% (1%)	10% (1%) rpsl in rpki

- Not all BGP is in RPSL, but much more is in RPSL than in RPKI
- Most RPSL is outside of RPKI
- Most RPKI is covered in RPSL

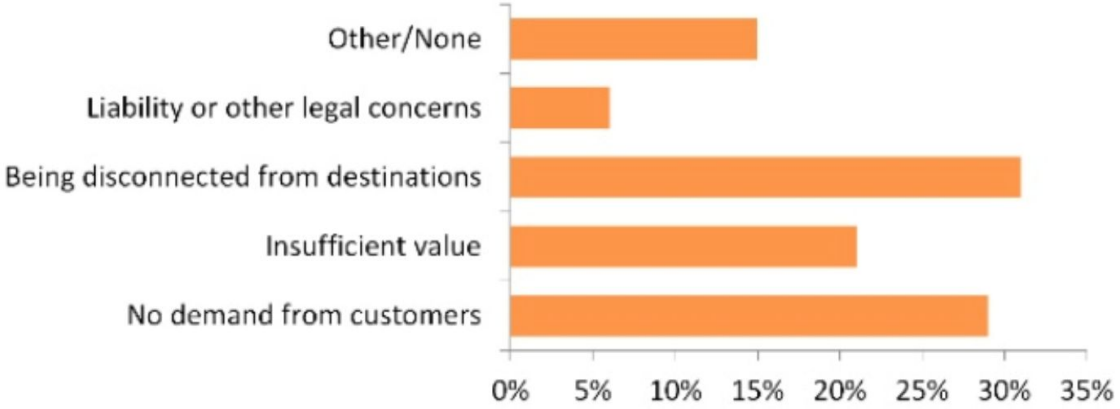
RPKI deployment status



Source: RIPE

Challenge to Deployment

What are your main concerns regarding executing RPKI-based origin authentication in your network?



Source : Michael Schapira, 2017

Three Route States

- Valid
  - Prefix is covered by a valid ROA
- Unknown
  - No ROA exists for this prefix
- Invalid
  - Unauthorized announcement
    - Mismatch between authorized ASN and originating ASN, split origin
    - More specific announcement that valid ROA allows
    - Expired ROA

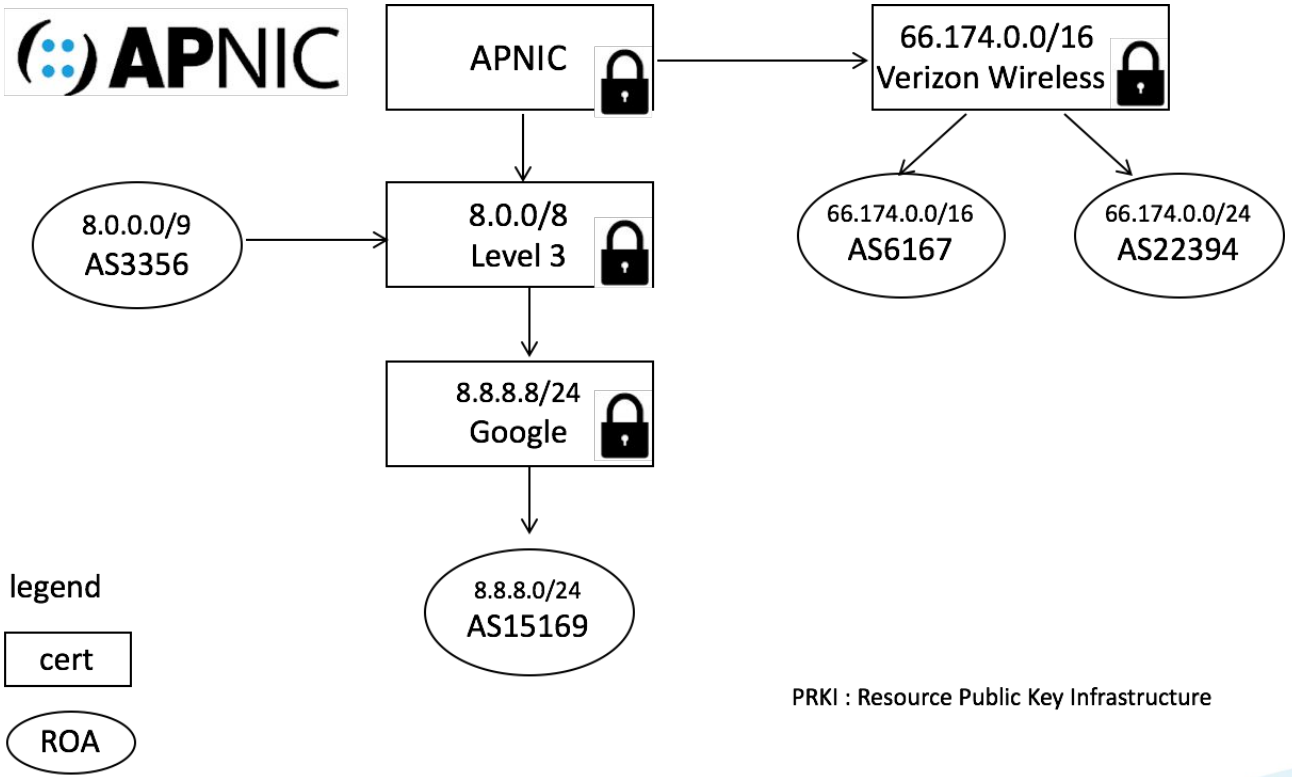
What to do with this data now

- With 95% of the table in the unknown state, probably nothing
- In a fully deployed RPKI environment, do you
  - Reject unknown, invalid routes?
  - Set LOCALPREF low ?
  - Set Community, put in a VRF?
- Still under operational development
- Study RFC6483

Public Resources

```
rpki@lr1.ham1.de> show route 72.52.2.0/24

inet.0: 511848 destinations, 511849 routes (511848 active, 0 holddown, 0 hidden)
+ = Active Route, - = Last Active, * = Both

72.52.2.0/24        *[BGP/170] 1d 11:54:16, localpref 90, from 79.141.168.1
                       AS path: 33926 32787 45757 I, validation-state: invalid
                     > to 193.34.50.1 via em0.0



agallo@foghorn:~$ whois -h whois.bgpmon.net " --roa 33926 72.52.2.0/24"
2 - Not Valid: Invalid Origin ASN, expected 32787
agallo@foghorn:~$ whois -h whois.bgpmon.net " --roa 32787 72.52.2.0/24"
0 - Valid
------------------------
ROA Details
------------------------
Origin ASN:       AS32787
Not valid Before: 2012-09-25 04:00:00
Not valid After:  2022-09-25 04:00:00  Expires in 7y256d16h34m26.2000000178814s
Trust Anchor:     rpki.arin.net
Prefixes:         209.200.128.0/18 (max length /32)
                  72.52.0.0/18 (max length /32)
                  2606:6c00::/32 (max length /64)
```

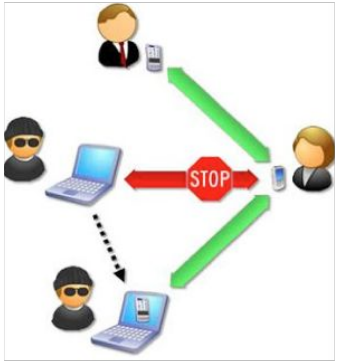
Secure internet routing - RPKI



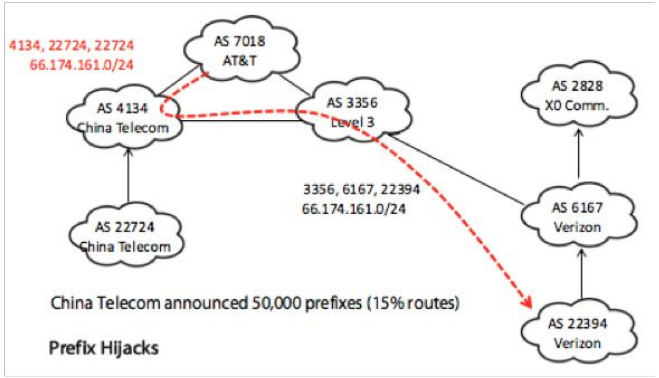
PRKI : Resource Public Key Infrastructure

Problem solved

	Confidentiality	Stakeholders 1.RiRs (e.g. APNIC) 2.ISPs 3.IETF 4.LEA (Law Enforcement Agent)
X	Integrity	
	Availability	



IP spoofing



Route hijacking

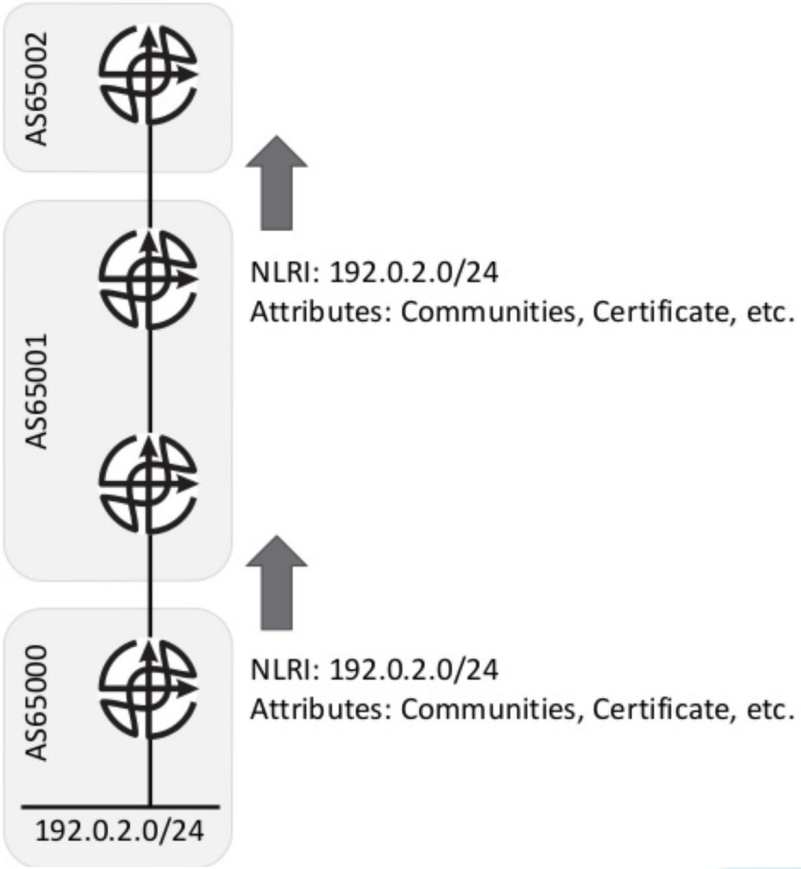
When

- When does TWNIC begin to provide RPKI service ?
  - Study team has been established
  - Lab operation will be completed in June 2018
  - Planned soft launch by the end of 2018

BGPsec - RFC8205

Operation

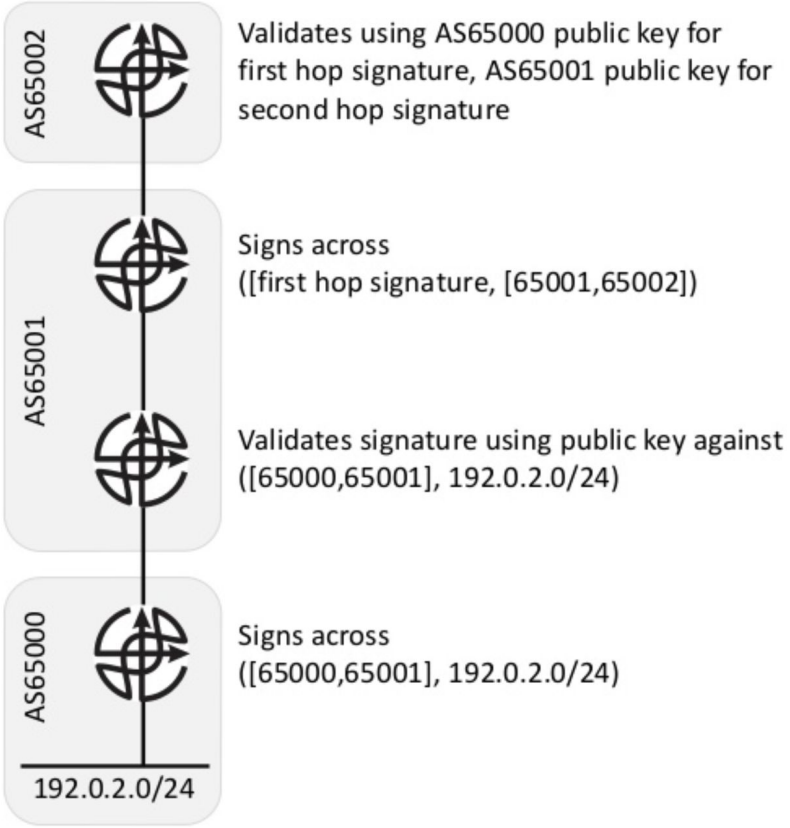
- Certificate added as BGP attribute
  - Signature (hash created using private key)
  - Private key retrieval information
  - Valid lifetime
  - Other information…
  - Certificate is around 256 octets
- Exchange of certificates attributes is negotiated at initial peering



BGPSEC

Operation

- Receiving Speaker
  - Calculates hash based on retrieved public key for each AS in the AS Path
  - Compares calculated hash against hash carried in the certificate within the update
- Sending Speaker
  - Adds a new certificate containing the sending AS and receiving AS
  - Calculates a hash across any existing certificates and this new information
  - Inserts the certificate into the attributes



What does BGPsec offer over RPKI

- Road to BGPsec full-deployment is very unpredictable
- Not fully deployed BGPsec provides only small benefits over RPKI