


Routing
Introduction

CSCI E 45a: The Cyber World – part A

1 Copyright © Scott Bradner & Ben Gaucherin 2015

Learning goals



- Understand the concepts of Internet Protocol routing
- Understand what IP routing can and cannot do
- Understand the different assumptions underlying enterprise and Internet service provider routing

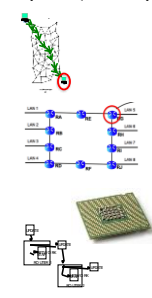
2 Copyright © Scott Bradner & Ben Gaucherin 2015

Introduction: this module

- This module deals with technology

3 Copyright © Scott Bradner & Ben Gaucherin 2015


Topics (all required)



- Routing concepts
Purpose & operation of IP routing
- Routing Basics
Basic technology in static and dynamic routing
- Routers
Parts of a router
- Enterprise (interior) routing
IP routing within a company or within an ISP

4 Copyright © Scott Bradner & Ben Gaucherin 2015

Topics (all required)



- Routing approaches
Geographic, vs. next hop vs source directed routing
- ISP (exterior) routing
Routing between ISPs

5 Copyright © Scott Bradner & Ben Gaucherin 2015

Image credits

All drawings and photos by Scott Bradner unless noted

Slide#	credit
4	network - Drawing from Baran 1962 paper cpu - https://commons.wikimedia.org/wiki/File:Intel_CPU_Pentium_4_640_Prescott_bottom.jpg

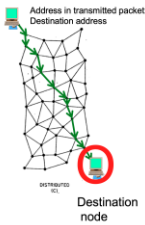
6 Copyright © Scott Bradner & Ben Gaucherin 2015

Routing
Routing Concepts

CSCI E 45a: The Cyber World – part A

1 Copyright © Scott Bradner & Ben Gaucherin 2015

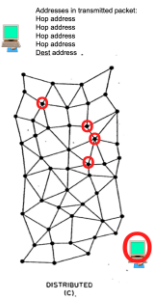
Purpose of routing



- Routing: “forward towards a destination”
Note not “to a destination”
- In IP, routing is hop by hop
Each hop determines the next hop
- Source directed path routing possible
Rarely used
Knowledge, security & trust issues

2 Copyright © Scott Bradner & Ben Gaucherin 2015

Source Directed Path Routing



- Source route
- Packet sender determines path the packet will take
- Puts list of routers to be visited in header
- IETF working group: Source Packet Routing in Networking (spring)
For MPLS & IPv6

3 Copyright © Scott Bradner & Ben Gaucherin 2015

Routing

- Forwarded from host -> router -> router ... -> host
- Each packet contains a destination IP address
- Router uses this destination IP address to decide:
 - Which network interface to send the packet out of
 - Which node to send the packet to next

4 Copyright © Scott Bradner & Ben Gaucherin 2015

Routing, contd.

- Uses table of known destination prefixes
Routing table
- Uses "longest match" to select specific entry
- Generally dynamically updated

prefix	port	Next hop	metric	timer
128.1030.0/6	4	196.43.6.38	5285	53
128.1030.0/6	5	196.43.7.42	1093	1033
0.0.0.0	4	196.43.6.38	1	9542


5 Copyright © Scott Bradner & Ben Gaucherin 2015

Hop-by-hop

- IP packet is end-to-end
IP Header modified in transit
- Encapsulated in link layer packet between devices
Header and CRC
New L2 header each hop

6 Copyright © Scott Bradner & Ben Gaucherin 2015

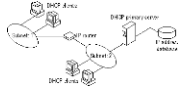
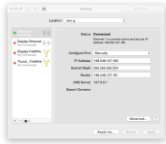
Host involvement in routing



- Host cannot know topology of the Internet
 - It only knows about the network it is directly connected to
 - A underlying problem for source routing
- Thus the host forwards packets to a router if destination is off-LAN
 - Rather than using source route

7 Copyright © Scott Bradner & Ben Gaucherin 2015

Host knowing the router(s)



- How hosts knows the address of the router:
 - Configured
 - DHCP
 - Router advertisements (IPv6)

8 Copyright © Scott Bradner & Ben Gaucherin 2015

Determining if the destination is off-LAN

- Compare “network parts” of destination and host LAN addresses
 - Use local network interface subnet mask
 - If results are identical then the destination is on local LAN

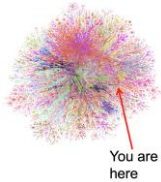
```
DA: 108.41.175.14 01101100.00101001.10101111.00001110
HA: 108.41.175.17 01101100.00101001.10101111.00001110
M: 255.255.255.0 11111111.11111111.11111111.00000000
```

Otherwise, destination is off-LAN

```
DA: 108.41.105.14 01101100.00101001.01101001.00011110
HA: 108.41.175.17 01101100.00101001.10101111.00011110
M: 255.255.255.0 11111111.11111111.11111111.00000000
```

9 Copyright © Scott Bradner & Ben Gaucherin 2015

(Inter)network topology



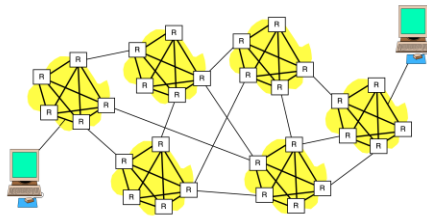
- Internet = interconnected networks running IP
- Internet uses internetwork device address unique throughout network specifies the LAN and the device on the LAN

10

Copyright © Scott Bradner & Ben Gaucherin 2015

(Inter)network topology, contd.

The Internet is a collection of ASs



11

Copyright © Scott Bradner & Ben Gaucherin 2015

Image credits

All drawings and photos by Scott Bradner unless noted

Slide# credit

2-4 Drawing from Baron's 1962 paper

8

https://www.microsoft.com/resources/documentation/windows/4/server/reskit/en-us/net/sur_dhcp.msp?mfr=true

10

<https://www.flickr.com/photos/44124348109@N01/916142/>

42/

12

Copyright © Scott Bradner & Ben Gaucherin 2015

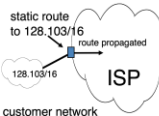
Routing

Routing basics

CSCI E 45a: The Cyber World – part A

1 Copyright © Scott Bradner & Ben Gaucherin 2015

Static routing

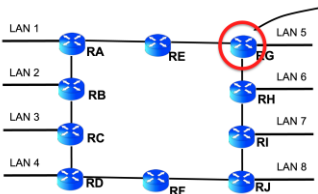


- Manually configure entry next-hop table
- Then propagated with rest of routing information
- Many ISPs use static routes to customers
- Reduces route flap
- Safer because do not have to trust routing information from customer

2 Copyright © Scott Bradner & Ben Gaucherin 2015

Dynamic routing

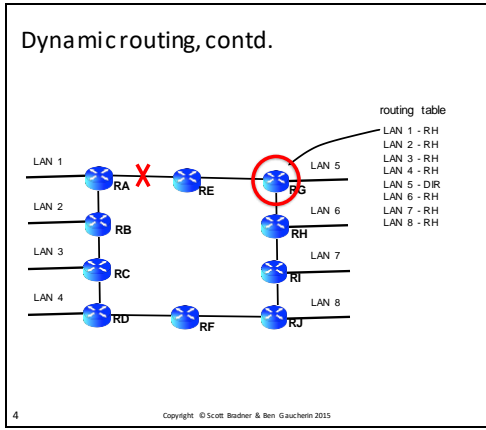
- Next-hop table updated with information from adjacent routers
- Updates sent when topology changes




LAN 1 - RE
LAN 2 - RE
LAN 3 - RE
LAN 4 - RE or RH
LAN 5 - DIR
LAN 6 - RH
LAN 7 - RH
LAN 8 - RH

assume that routing protocol run on all inter-router links

3 Copyright © Scott Bradner & Ben Gaucherin 2015




Routing protocols



- “Routing protocol” used to distribute information about current network topology between routers
- Used in just about all IP networks, enterprise and ISP
Other than isolated functions like default route & ISP customer routes

5 Copyright © Scott Bradner & Ben Gaucherin 2015

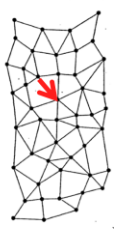
Dynamic routing, issues



- Can be complex between organizations
- Can be very complex between ISPs
Thousands of lines in a ISP router configuration

6 Copyright © Scott Bradner & Ben Gaucherin 2015

Next-hop

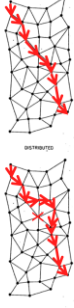


- IP routers, even if they calculate a full path to a destination, only send a packet to the next-hop on the way to that destination
 - The next-hop can be a router or the destination itself
- Each router does its own next hop calculation for each received packet

DISTRIBUTED ICL

7 Copyright © Scott Bradner & Ben Gaucherin 2015

Next-hop, contd.

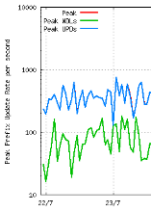


- Thus, system can dynamically route through network changes

DISTRIBUTED ICL

8 Copyright © Scott Bradner & Ben Gaucherin 2015

Routing protocol goals

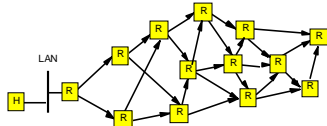


- Lowest "cost" path to destination
 - Cost could be: hops, bandwidth, time, load, \$\$
- Topology can change often (many times per second)
- Must deal with scaling
- Maximum stability

9 Copyright © Scott Bradner & Ben Gaucherin 2015

Change notification

- Router adjacent to change informs its neighbors of change
- Information propagates throughout network
Routers update next-hop tables



10

Copyright © Scott Bradner & Ben Gaucherin 2015

Convergence



- Time it takes for whole network to agree on a topology after a change (e.g. link break)
- ‘The Internet never converges’
Can take minutes to propagate a routing change across the Internet but have multiple updates per second
- The Internet never agrees on what the Internet looks like

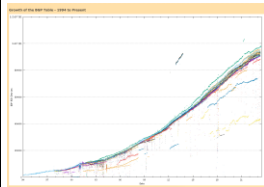
Big Idea!

11

Copyright © Scott Bradner & Ben Gaucherin 2015

Convergence, contd.

- Convergence problem is worse with bigger routing tables
Bigger tables mean longer processing
And bigger tables also mean more chance of change

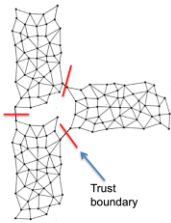


12

Copyright © Scott Bradner & Ben Gaucherin 2023

UPDATED

Side note on reality

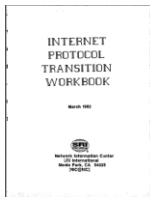


- No central control over ISPs
- Wide range in clue density
- Must be defensive
BGP is mostly programming language to describe level of paranoia between ISPs
- Keep in mind CIDR & forced renumbering
- Note that multiple vendors must support any new routing protocol

13

Copyright © Scott Bradner & Ben Gaucherin 2015

Side note on reality, contd.



- Can not have a “flag day” to move to a new routing protocol
Remember the lesson of IDRP



14

Copyright © Scott Bradner & Ben Gaucherin 2015

What IP Routing Does Not Do



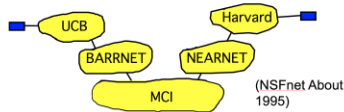
- Converge instantly
- Respond to congestion
- Help QoS
- Find lowest actual cost path

15

Copyright © Scott Bradner & Ben Gaucherin 2015

Default route

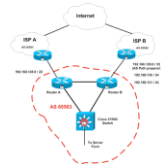
- IP routing permits the use of a default route
- Implies that another router might know more networks
- Permits routers to carry less than full routing tables
- Required in current Internet



16

Copyright © Scott Bradner & Ben Gaucherin 2015

Autonomous system (AS)



- “Routers under common management”
- AS number is value indicating authority (source) for routing information
- Required in some protocols e.g. BGP & OSPF
- BGP (& an AS) is required for multi-homed sites
- Get AS numbers through RIR

17

Copyright © Scott Bradner & Ben Gaucherin 2015

Routing protocol types




- Distance vector - e.g. RIP
What does the network look like from here?
- Link state - e.g. OSPF
Building & parsing a map
- Path vector - e.g. BGP4
Island hopping

18

Copyright © Scott Bradner & Ben Gaucherin 2015

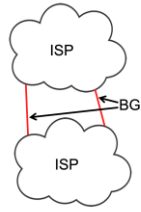
Interior vs. exterior



- Interior Gateway Routing Protocol (IGP)
- For routing within an AS
e.g. RIP, OSPF, IS-IS
- Designed for environments under a single authority and trust environment (e.g., within an enterprise)

19 Copyright © Scott Bradner & Ben Gaucherin 2015

Interior vs. exterior, contd.



- Exterior Gateway Routing Protocol (EGP) For routing between ASs
e.g. BGP4
- Designed for multiple authorities and trust environments (e.g. the Internet)

20 Copyright © Scott Bradner & Ben Gaucherin 2015

Image credits

All drawings and photos by Scott Bradner unless noted

Slide#	credit
5	http://www.dailykos.com/story/2009/01/22/687428/-Iraq-Ruins-and-Rubble-w-personal-photos
6	http://www.cs.cmu.edu/~kbe/minenet05_full.pdf
7, 8 & 13	Baran 1962 paper
9	http://bgpupdates.potaroo.net/instability/bgpupd.html
12	http://bgp.potaroo.net/
14	www.dtic.mil/dtic/tr/fulltext/u2/a153607.pdf
15	http://www.m2mnow.biz/2014/07/23/22894-iot-failure-rate-hit-30-cellular-network-congestion-says-isv-yet-analysts-reply/
17	https://ccieat21.wordpress.com/category/bgp/


21 Copyright © Scott Bradner & Ben Gaucherin 2015

Routing
Routers

CSCI E 45a: The Cyber World – part A

1 Copyright © Scott Bradner & Ben Gaucherin 2015

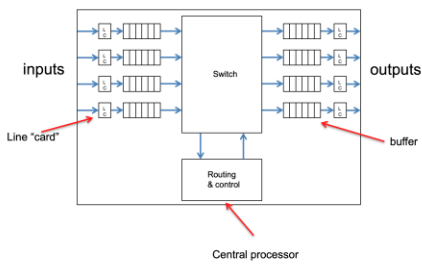
Router



- Originally, software on a general purpose computer with more than one network interface
- Now generally, special purpose devices used to forward IP packets between networks

2 Copyright © Scott Bradner & Ben Gaucherin 2015

Router diagram



3 Copyright © Scott Bradner & Ben Gaucherin 2015

Parts of a router



- Interfaces

Connections to “directly connected” networks
 “hardware forwarding” in interfaces in most high-end routers



- Central processor (CPU)

Processor – runs routing, management & control software
 Also, serves as forwarding engine in some cases

E.g., when the line card does not have the right route or if packet has options

4

Copyright © Scott Bradner & Ben Gaucherin 2015

Parts of a router, contd.

- Routing table

Table of reachable prefixes – includes:

- Prefix
- Output port & next hop address
- Route metrics or path information
- Timers

prefix	port	Next hop	metric	timer
128.103.0.0/16	4	196.43.6.18	5285	53
128.103.0.0/16	5	196.43.7.42	1093	1033
0.0.0.0	4	196.43.6.18	1	9542

5

Copyright © Scott Bradner & Ben Gaucherin 2015

Image credits

All drawings and photos by Scott Bradner unless noted

Slide# credit

4 top <http://www.eircomicdirect.ie/cisco-13e3-interface-card-expansion-module-ports-113e3-p-12838.html>

bottom

https://commons.wikimedia.org/wiki/File:Intel_CPU_Pentium_4_640_Prescott_bottom.jpg

6

Copyright © Scott Bradner & Ben Gaucherin 2015

Routing
Enterprise (interior) routing


CSCI E 45a: The Cyber World – part A

1 Copyright © Scott Bradner & Ben Gaucherin 2015

Routing information protocol version 2 (RIPv2)

XEROX

Berkeley
UNIVERSITY OF CALIFORNIA

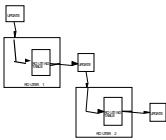


Chuck Hedrick

- Distance vector protocol
- First IP routing protocol
- Updated for classless operation (v2)
- No RIPv2-specific configuration required
- Automatically finds neighboring RIP routers & trusts routing information
- Useful in multi-subnet homes & small enterprises

2 Copyright © Scott Bradner & Ben Gaucherin 2015

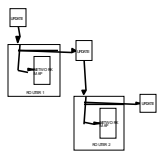
RIP operation



- Router maintains table of its view of the network
- Modifies table based on received updates
 - Replace entry if update with a lower cost is received
- Sends copy of table or changed parts of table as updates

3 Copyright © Scott Bradner & Ben Gaucherin 2015


Link-state routing



- Router maintains map of network topology
- Modifies map based on received updates
- Sends status of changed links as updates
- Router calculates paths through network for each destination

7 Copyright © Scott Bradner & Ben Gaucherin 2015

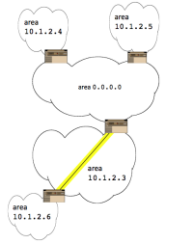
OSPF Hierarchical routing



- Can subdivide network into areas
- Within area topology can be hidden
- Within area topology changes often hidden
- As long as reachability stays the same

8 Copyright © Scott Bradner & Ben Gaucherin 2015

OSPF Hierarchical routing, contd.



- All areas must connect to a backbone area if more than one area
- Area 0.0.0.0
- Can tunnel connection through another area with virtual links
- Conceptual hub and spoke topology

9 Copyright © Scott Bradner & Ben Gaucherin 2015

Router Link State Announcement (LSA)

- Router discovers its neighbors & their addresses
- Builds **Router Link State Announcement (LSA)** with information about its links to its neighbors
- Sends LSA to its neighbors
 - Reliably sends updates on any changes
- Reliably forwards any received LSAs

LSA details for R1:
 Router ID: R1
 Link 1: R4
 Link 2: R6

LSA details for R5:
 Router ID: R5
 Link 1: R1
 Link 3: R3

10 Copyright © Scott Bradner & Ben Gaucherin 2015

Other LSAs

- **Summary LSA**
LSA from area boundary router
Summarizes area contents
- **Default routes**
Indicates exit from stub area
- **External LSA**
LSA from AS border router
Includes external routes

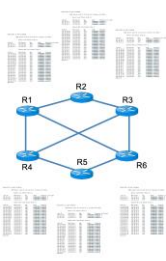
11 Copyright © Scott Bradner & Ben Gaucherin 2015

Topology database

- Router builds database of network based on received LSAs and its own links
- Computes routes to all destinations (advertised prefixes) with itself as the root for each update

12 Copyright © Scott Bradner & Ben Gaucherin 2015

Link-state update process



- Aim is to maintain identical copies of topology databases in all nodes
- Really bad if topology database gets corrupted e.g. routing loops ("Byzantine failures")

13 Copyright © Scott Bradner & Ben Gaucherin 2015

Link-state update process, contd.


LS Age	
Options	LS Type
Link State ID	
Advertising Router	
LS Sequence Number	
LS Checksum	
Length	

LSA Header

- Lots of protections against corruption
 - Hop-by-hop acknowledgements
 - Updates have sequence numbers
 - Checksum protection for updates
 - Discards entry if no update within timeout
 - Password & crypto protection for OSPF updates

14 Copyright © Scott Bradner & Ben Gaucherin 2015

Link-state route calculation



- Walk through topology
- Build set of paths with "self" as the start
 - One for each prefix in topology database
 - Can have multiple paths if equal cost
 - Only uses the next-hop in path
- Results in a tree with the router as the root

15 Copyright © Scott Bradner & Ben Gaucherin 2015

OSPF neighbor discovery

- On multi-access broadcast networks – e.g., Ethernet
- Automatically finds neighbor OSPF routers
- Sequence:
 - Each router periodically sends a multicast Hello packet
 - Hello packet includes
 - network mask, hello interval, router priority, dead router interval, designated router (if any), backup designated router (if any), known neighbors
- Routers form adjacencies and exchange LSAs

OSPF Packet Header

Ver #	1	Packet Len
Router ID		
Area ID		
Checksum	Auth type	
Authentication		
Authentication		
Network Mask		
Hello Int	Options	Rtr Pri
Deadint		
Designated Router		
Backup Designated Router		
Neighbor		
Neighbor		
...		

Hello Packet Header

16 Copyright © Scott Bradner & Ben Gaucherin 2015

OSPF designated router

- Designated router used to reflect LSAs
- Instead of a full mesh of interconnection between routers

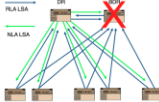

17 Copyright © Scott Bradner & Ben Gaucherin 2015

OSPF designated router, contd.

- Elected by Hello packets
- Multicast on subnets that support multicast
- Hello includes router IDs of all known routers
- Uses Router Priority to select DR & BDR
- Router ID used to break ties
- BDR takes over if DR stops sending Hello packets
- New BDR elected

18 Copyright © Scott Bradner & Ben Gaucherin 2015


IS-IS



- Intermediate System to Intermediate System routing protocol
- Developed and maintained by ISO
 - IP extensions developed by IETF
- Link-state protocol
- Basic operation same as OSPF
- A few differences
 - No backup designated router

19 Copyright © Scott Bradner & Ben Gaucherin 2015

IS-IS, contd.



- Used by some big ISPs for historical reasons
 - Cisco Systems IS-IS implementation was far better than their OSPF implementation in the late 1980s
- Once established, too hard to change

20 Copyright © Scott Bradner & Ben Gaucherin 2015

Image credits

All drawings and photos by Scott Bradner unless noted

Slide#	credit
2	xerox logo - http://www.logodesignlove.com/tom-geismar-interview berkeley logo - http://brand.berkeley.edu/identity/ hedrick photo - http://toolbox.rutgers.edu/~hedrick/
4	fireflies http://www.smithsonianmag.com/arts-culture/beautiful-flight-paths-fireflies-180949432/?no-ist
5	rfc 2453
6	https://capaocho8.com/configuracion-de-ospf-con-ipv4-e-ipv6-para-ios-cisco-xr-parte-1/
10	https://blog.apnic.net/2015/05/12/ospf-topology-database-design-optimization-principle-of-isp-igp-routing/


21 Copyright © Scott Bradner & Ben Gaucherin 2015

Routing
Routing approaches & address allocation

CSCI E 45a: The Cyber World – part A

1 Copyright © Scott Bradner & Ben Gaucherin 2015

Geographic addressing

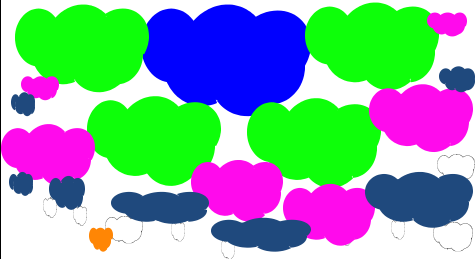


- **Partial geographic**
e.g., high-order bits of address define country (ITU-T) (all politics) but maybe not a big routing table impact
- **Full geographic**
Whole address defines physical location
Not network location
Not enough local connectivity to make this work

2 Copyright © Scott Bradner & Ben Gaucherin 2015

Hierarchical routing

- The Internet is a rough hierarchy



3 Copyright © Scott Bradner & Ben Gaucherin 2015

Hierarchical routing, contd.



Len Kleinrock



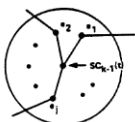
Farouk Kamoun

- Kleinrock & Kamoun – 1977
Hierarchical routing is the only way to deal with large networks
Else routing table gets too large
- Packet forwarded towards router connecting the LAN with the address prefix that includes the target address to the Internet

4

Copyright © Scott Bradner & Ben Gaucherin 2015

Hierarchical routing, contd



- There could be shorter paths but hierarchy reduces the size of the routing table
Listing every LAN in the Internet would be far too big
Lose some performance in order to scale better
- Current Internet routing approach

5

Copyright © Scott Bradner & Ben Gaucherin 2015

IP address assignment



Internet Assigned Numbers Authority

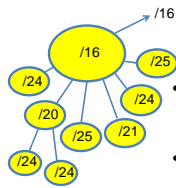


- Root assignments by IANA function in ICANN
- Regional assignments by 5 regional IP registries (RIRs):
ARIN (NA)
RIPE (Europe)
LATNIC (SA)
APNIC (AP)
AFRINIC (Africa)
- Exclusive territories

6

Copyright © Scott Bradner & Ben Gaucherin 2015

IP address assignment, contd.



- Big (enough) ISPs get allocations
"Big enough" means can justify ~ 4,000 hosts
Else must get addresses from upstream ISP
- Sub allocate to their customers (some of which are ISPs)
- ISP customers sub allocate to their customers
ISP customers sub allocate to their customers, etc.
- Appears as single entry in routing table

7

Copyright © Scott Bradner & Ben Gaucherin 2015

IPv4 address market



- IANA ran out of IPv4 addresses
Jan 31, 2011
RIRs ran out later
- Plenty of IPv6 addresses
- Now market in IPv4 addresses
Or, specifically, the right to use IPv4 addresses
- Value varies
2019 example \$20/address
- Still must show need
- Routing table impact

8

Copyright © Scott Bradner & Ben Gaucherin 2015

Image credits

All drawings and photos by Scott Bradner unless noted

Slide#	credit
4	Kleinrock - http://www.nae.edu/29075.aspx Kamoun - https://en.wikipedia.org/wiki/Farouk_Kamoun
5	http://www.cs.bu.edu/fac/matta/Teaching/cs556/archive/Kamoun-HRouting.pdf
6	www.iana.org https://www.nro.net/policies/getting-internet-number-resources
8	www.iana.org http://www.hultprize.org/mit-Logo/

9

Copyright © Scott Bradner & Ben Gaucherin 2015

Path vector protocol

- Router maintains table of paths of ASs to prefixes
- Modifies table based on received updates
 - Replace entry if update with a better path is received
- Sends copy of table or changed parts of table as update

4 Copyright © Scott Bradner & Ben Gaucherin 2015

Path vector operation

- Export list of within-AS prefixes prepended with AS# to neighbors
 - Update: prefix, AS path, attributes
- When update received
 - Put into AS path table
 - Prepend own AS# and send to neighbors
- Filtered AS path table put into forwarding table

5 Copyright © Scott Bradner & Ben Gaucherin 2015

BGP4, updates

- Routes stored in Routing Information Base (RIB)
- Conceptual picture - 3 tables
 - Adj-RIBs-In - all received routes
 - Loc-RIB - locally used routes
 - Adj-RIBs-In data filtered by policies
 - Adj-RIBs-Out - routes to be advertised to peers
 - Loc-RIB data filtered by policies

6 Copyright © Scott Bradner & Ben Gaucherin 2015

BGP4 operation

Marker		
Length	Tp: 1	Ver
My AS	Hold Time	
BGP ID		
Optional parameters		

- OPEN messages exchanged at start
- Keepalive messages exchanged if no updates within *hold time* (negotiated at start)
- If no update or keepalive within hold time or if link goes down
 - **Discard all paths learned through that link**

Marker		
Length: 19	Tp: 4	

7 Copyright © Scott Bradner & Ben Gaucheri 2015

BGP4 operation, contd.

```

521215 network entries using 8800380 bytes of memory
14632697 path entries using 2466445440 bytes of memory
2434252/89607 BGP path/bestpath attribute entries using 836207144 bytes of memory
3098608 BGP AS-PATH entries using 159700130 bytes of memory
2 BGP ATTR_SET entries using 80 bytes of memory
175053 BGP community entries using 23970884 bytes of memory
2636 BGP extended community entries using 194906 bytes of memory
...
BGP using 3712194040 total bytes of memory
    
```

- Accumulates information about
 - What prefixes are in what ASs
 - Connections between ASs
- Creates paths to all prefixes - AS paths
 - Paths are a series of ASs in route back to AS with prefix
 - Actual path selected by length and weight or other policy
 - Next hop recalculated at each hop
- All paths kept in case of failure of a non-adjacent link
 - Quick move to alternate path

8 Copyright © Scott Bradner & Ben Gaucheri 2015

BGP operation, contd.

- Routing scale

```

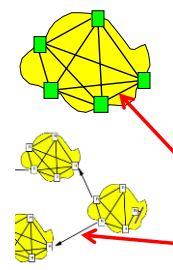
08/10/21 snapshot of a backbone router

909982 network entries using 225675536 bytes of memory
20553712 path entries using 2466445440 bytes of memory
3371803/157232 BGP path/bestpath attribute entries using 836207144 bytes of memory
3098608 BGP AS-PATH entries using 159700130 bytes of memory
2 BGP ATTR_SET entries using 80 bytes of memory
175053 BGP community entries using 23970884 bytes of memory
2636 BGP extended community entries using 194906 bytes of memory
...
BGP using 3712194040 total bytes of memory
    
```

UPDATED

9 Copyright © Scott Bradner & Ben Gaucheri 2021

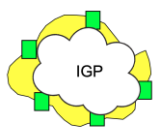
iBGP



- All BGP border routers for an AS must be full-mesh connected
With iBGP over logical (TCP), not physical, links
- BGP between routers in same AS = iBGP
- BGP between routers in different ASs = eBGP

10 Copyright © Scott Bradner & Ben Gaucherin 2015


iBGP, contd.



- iBGP is to exchange routing between BGP routers
- Also need a IGP for internal routing
And to tell edge routers what the internal prefixes are
- Two protocols in use
OSPF & IS-IS

11 Copyright © Scott Bradner & Ben Gaucherin 2015

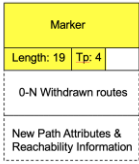
BGP4 policy



- BGP enforces policy, e.g.,:
 - Selectively rejecting routes through specific ASs
 - Selectively refusing to be a transit provider
 - Selectively prefer specific routes for a prefix
- Policy enforced by:
 - Controlling redistribution of routing information
 - Controlling selection of paths between multiple alternative paths

12 Copyright © Scott Bradner & Ben Gaucherin 2015

BGP4 updates

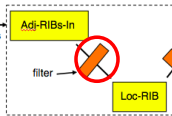


- Updates include
 - List of N withdrawn routes
 - 0 or one feasible route
 - List of prefixes & masks for the route
 - Path attributes for the route
- Path attributes are path-specific information, e.g.,
 - AS-path: sequence of ASs back to origin
 - next_hop: IP address of router to be used for next hop

13

Copyright © Scott Bradner & Ben Gaucherin 2015

Filtering input routes

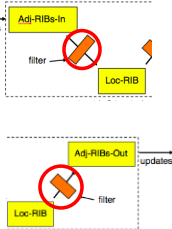


- ISPs deal with routes from
 - Customers, peers & internal
- Filter input routes
 - Reject input if not IRR compliant
 - Reject routes from customers with peer AS in their paths
 - Reject input if same as ISP internal routes

14

Copyright © Scott Bradner & Ben Gaucherin 2015

Filtering output routes



- Mark entries with community attributes
 - “Peer”, “customer”, & “internal” communities
 - Used in filtering output routes
- Use community attribute
 - Send peers only routes with “customer” attribute
 - Send customers routes with “customer” or “peer” attributes

15

Copyright © Scott Bradner & Ben Gaucherin 2015

BGP path determination

Example path selection sequence:

- Path with the highest LOCAL_PREF value
- Path with the shortest AS path
- Path with the lowest origin type
- Path with the lowest MED
 - if the same AS advertises the possible paths
- Prefer eBGP over iBGP paths
- Oldest path
- Path through the neighbor with the lowest router ID
- Path through the neighbor with the lowest IP address.

16 Copyright © Scott Bradner & Ben Gaucherin 2015

BGP path determination, contd.

17 Copyright © Scott Bradner & Ben Gaucherin 2015

Asymmetric data paths

- Also asymmetric load
- Nearest exit routing - get it off my net ASAP
 - a.k.a., "hot potato routing"
- Influences peering agreements
 - "Other" ISP must have a good backbone

18 Copyright © Scott Bradner & Ben Gaucherin 2015

Image credits

All drawings and photos by Scott Bradner unless noted

Slide# credit

5 <http://www.digitaltut.com/route-bgp-questions>

16 <http://routing-bits.com/2010/01/07/rb-route-selection/>

18 – web server - <https://openclipart.org/detail/163741/web-server>

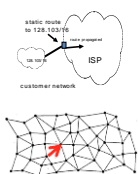
Routing
Conclusion

CSCI E 45a: The Cyber World – part A

1 Copyright © Scott Bradner & Ben Gaucherin 2015

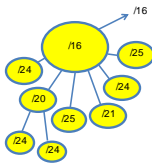
Routing summary

- Dynamic routing is the way to go
Heals the network when needed
- But there are times when static routing is better
E.g. customer links in ISPs
- Routing generally just gets you to the next hop
Source routing rare



2 Copyright © Scott Bradner & Ben Gaucherin 2015

Routing summary, contd.



IETF

- Recommended enterprise routing protocol is OSPF
Can be complex to set up because of its flexibility
- IGPs are for nets where trust exists
- Default routing makes the net work
- Hierarchical addressing and routing are required

3 Copyright © Scott Bradner & Ben Gaucherin 2015

Routing summary, contd.



- EGP needed between networks with multiple connections
ISP & enterprise
- BGP4 is the recommended RGP
- BGP is a programming language for mistrust
- Asymmetric data paths are the norm for the Internet

4

Copyright © Scott Bradner & Ben Gaucherin 2015

Routing summary, contd.



- The Internet never converges
- Routing does not recognize congestion

5

Copyright © Scott Bradner & Ben Gaucherin 2015

Image credits

All drawings and photos by Scott Bradner unless noted

Slide# credit

2 Drawing from Baron's 1962 paper

3 & 4 – IETF – www.ietf.org

5 <http://www.m2mnow.biz/2014/07/23/22894-iot-failure-rate-hit-30-cellular-network-congestion-says-isv-yet-analysts-reply/>

6

Copyright © Scott Bradner & Ben Gaucherin 2015
