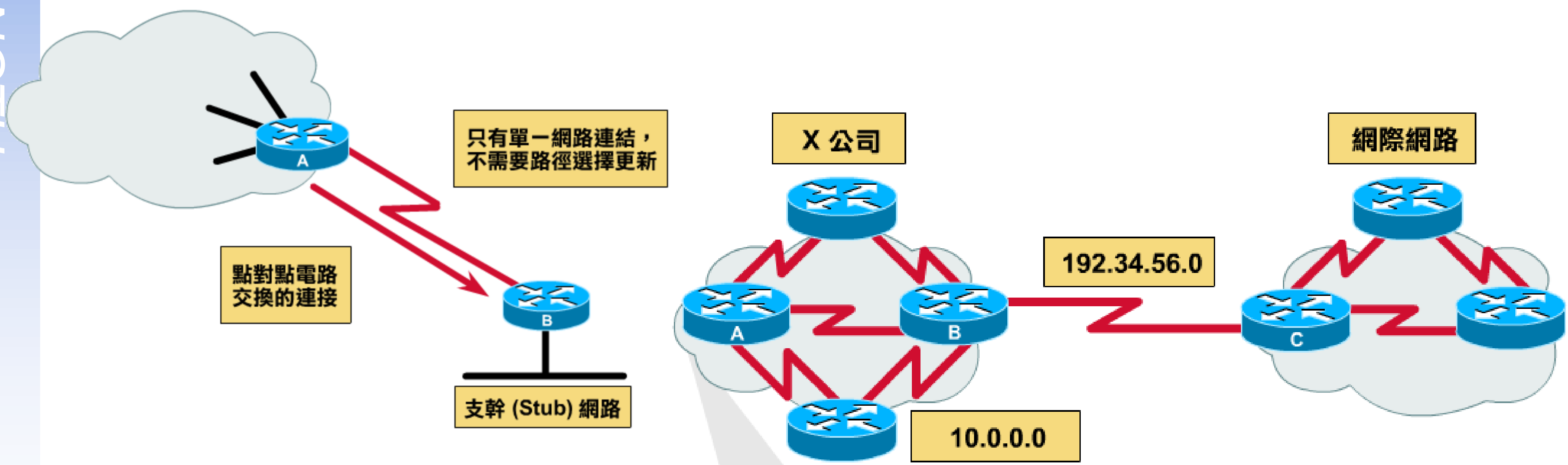
Routing

- Dynamic Route
- Routing Protocol

Why dynamic route ? (1)

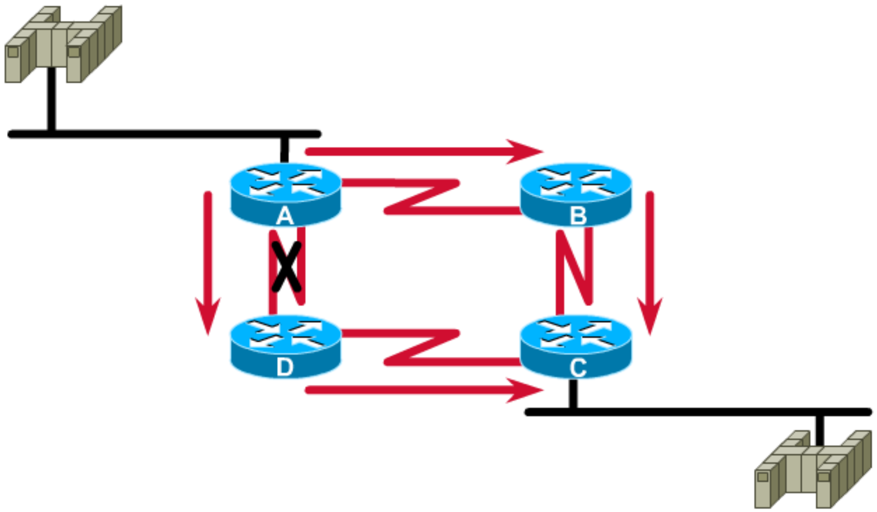
- Static route is ok only when
  - Network is small
  - There is a single connection point to other network
  - No redundant route



Why dynamic route ? (2)

❑ Dynamic Routing

- Routers update their routing table with the information of adjacent routers
- Dynamic routing need a routing protocol for such communication
- Advantage:
  - They can react and adapt to changing network condition

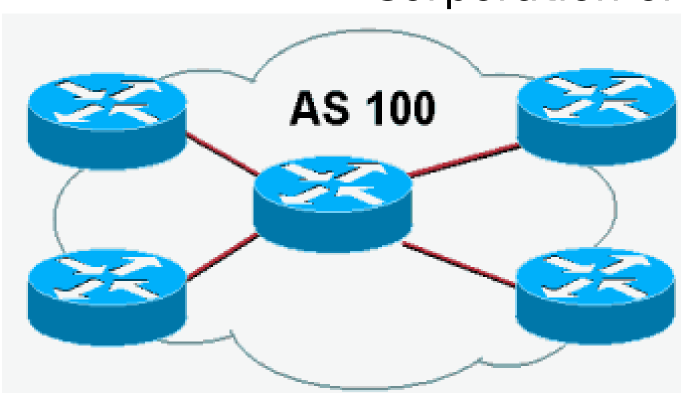


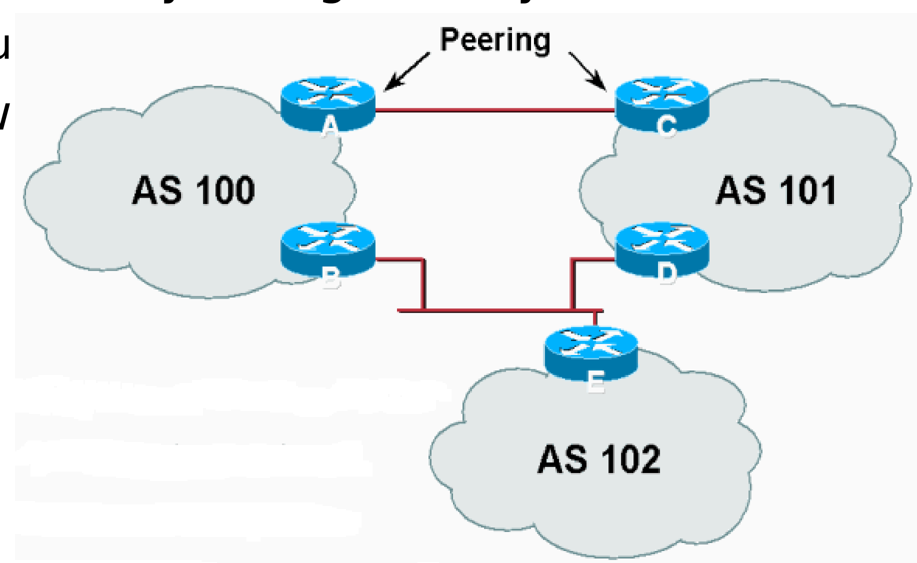
Routing Protocol

- Used to change the routing table according to various routing information
  - Specify detail of communication between routers
  - Specify information changed in each communication,
    - Network reachability
    - Network state
    - Metric
- Metric
  - A measure of how good a particular route
    - Hop count, bandwidth, delay, load, reliability, …
- Each routing protocol may use different metric and exchange different information

Autonomous System

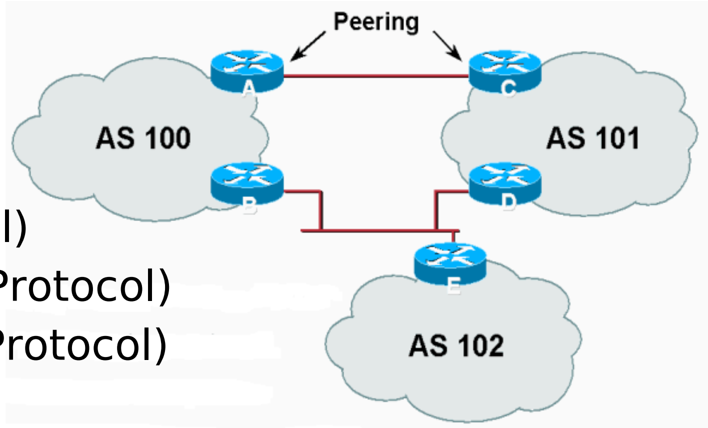
- Autonomous System (AS)
  - Internet is organized into a collection of autonomous system
  - An AS is a collection of networks with same routing policy
    - Single routing protocol
    - Normally administered by a single entity
      - Corporation or u



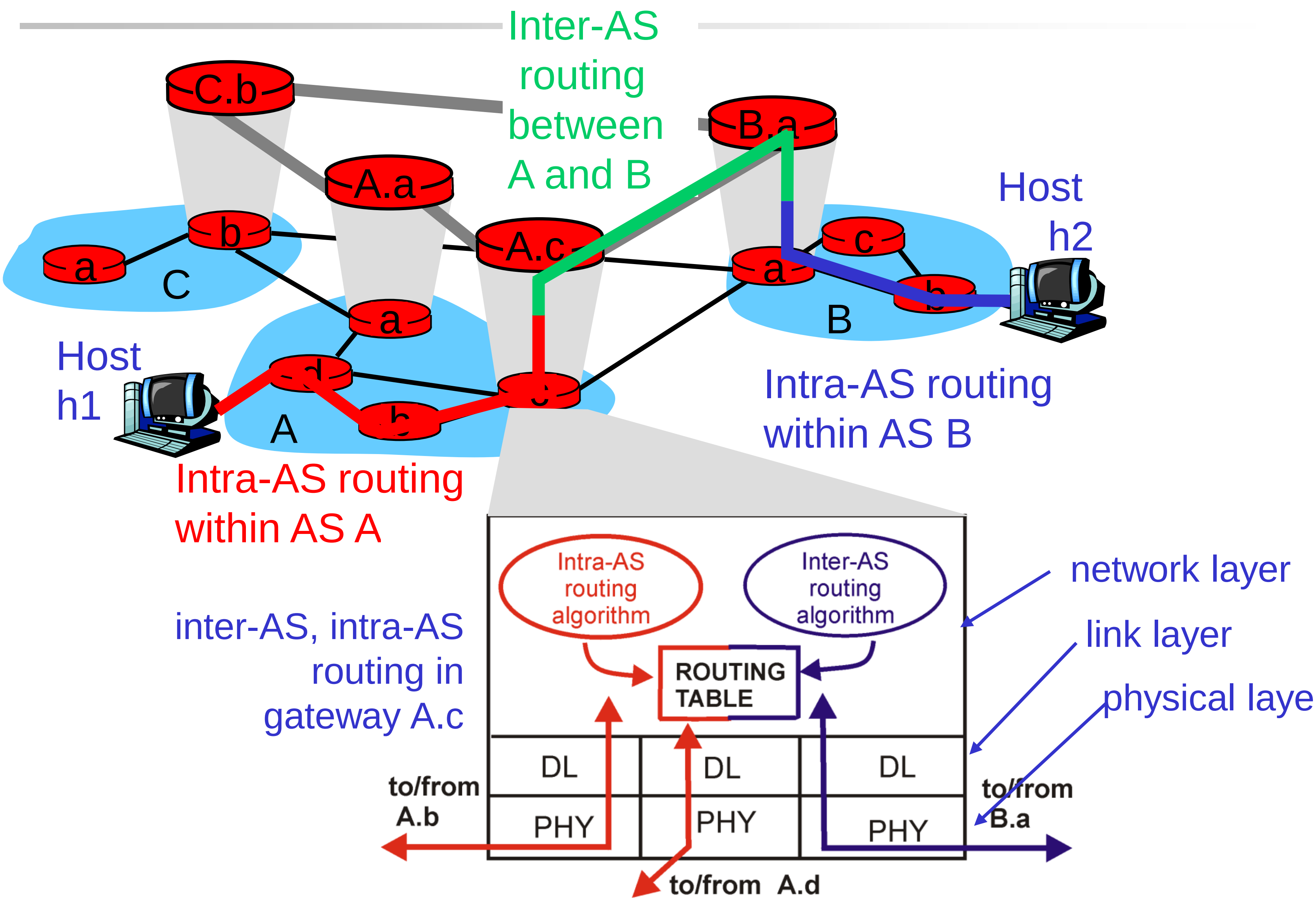


Category of Routing Protocols – by AS

- AS-AS communication
  - Communications between routers in different AS
  - Interdomain routing protocols
  - Exterior gateway protocols (EGP)
  - Ex:
    - BGP (Border Gateway Protocol)
- Inside AS communication
  - Communication between routers in the same AS
  - Intradomain routing protocols
  - Interior gateway protocols (IGP)
  - Ex:
    - RIP (Routing Information Protocol)
    - IGRP (Interior Gateway Routing Protocol)
    - OSPF (Open Shortest Path First Protocol)

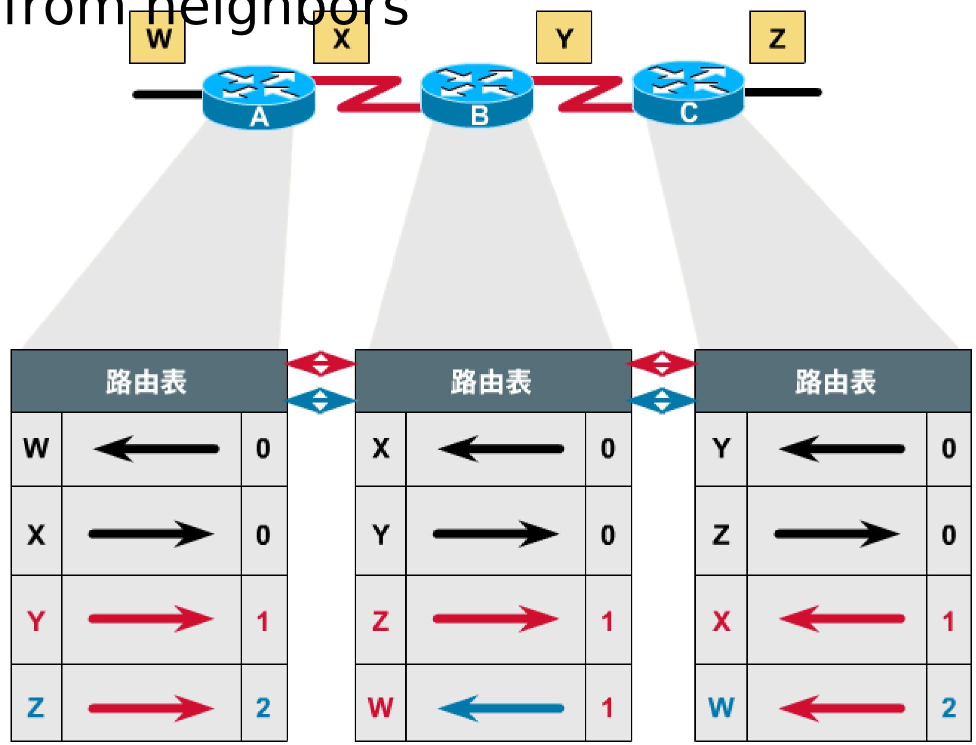


Intra-AS and Inter-AS routing



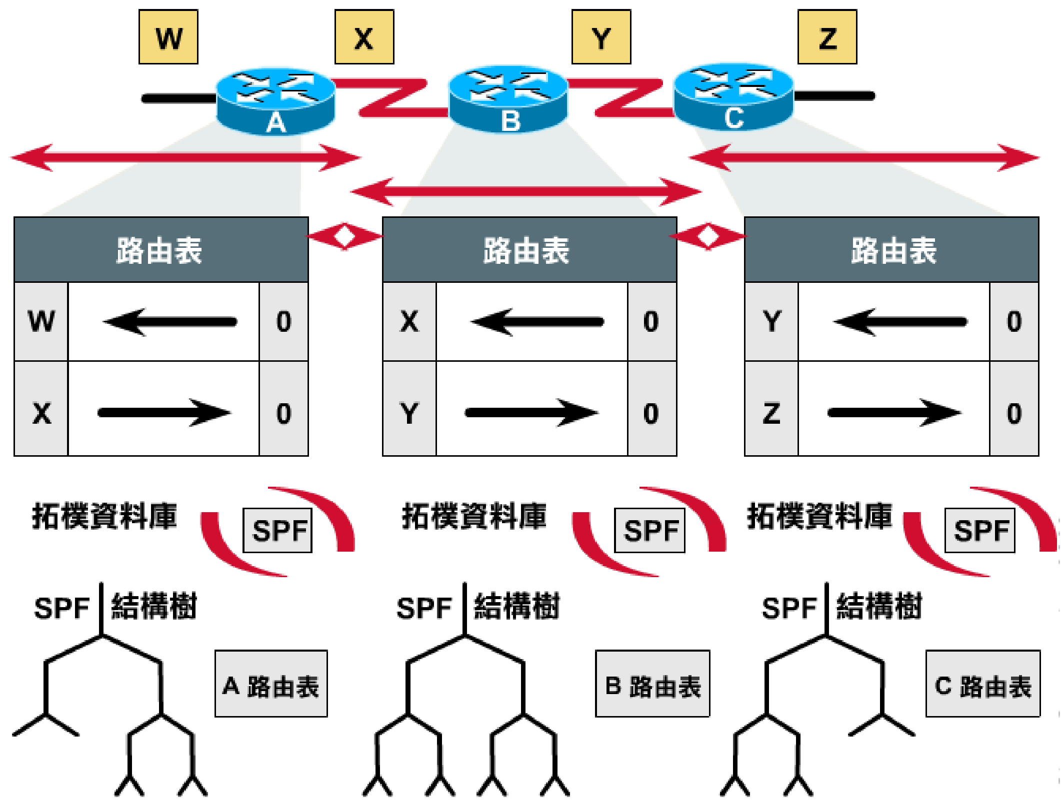
Category of Routing Protocols – by information changed (1)

- Distance-Vector Protocol
  - Message contains a vector of distances, which is the cost to other network
  - Each router updates its routing table based on these messages received from neighbors
  - Protocols:
    - RIP
    - IGRP
    - BGP



Category of Routing Protocols – by information changed (2)

- Link-State Protocol
  - Broadcast their link state to neighbors and build a complete network map at each router using Dijkstra algorithm
  - Protocols:
    - OSPF

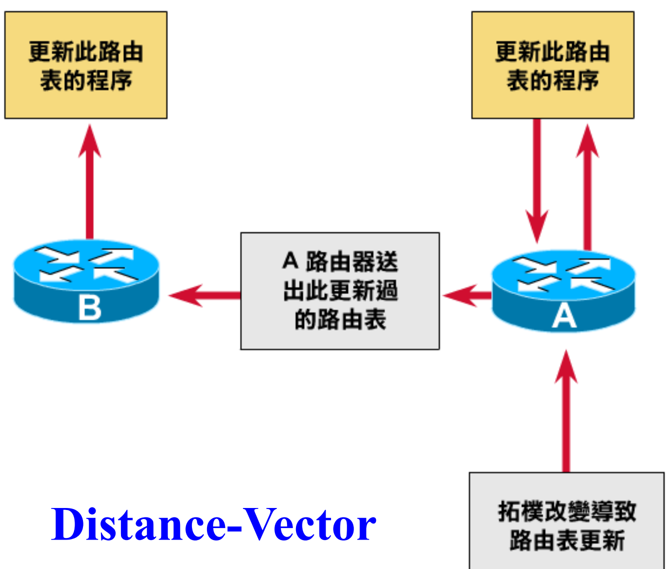


Difference between Distance-Vector and Link-State

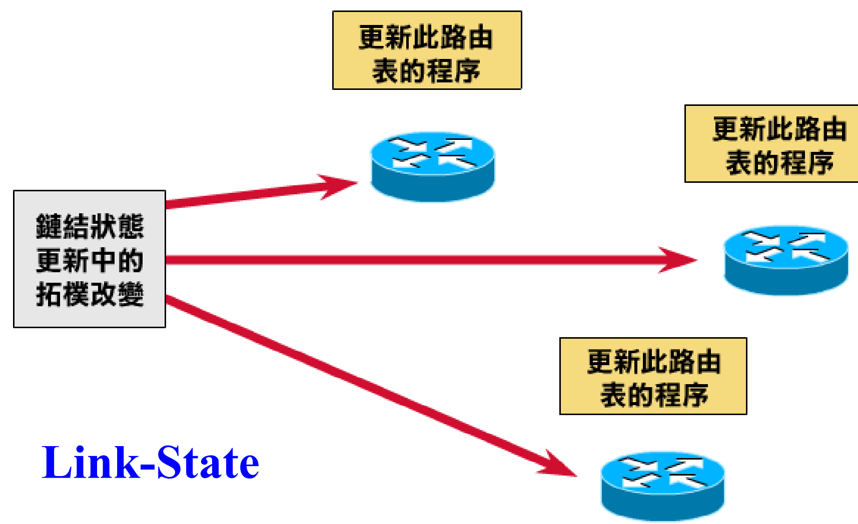
- Difference

	Distance-Vector	Link-State
Update	updates neighbor (propagate new info.)	update all nodes
Convergence	Propagation delay cause slow convergence	Fast convergence
Complexity	simple	Complex

- Information update sequence



Distance-Vector



Link-State

Routing Protocols

RIP	IGP,DV
IGRP	IGP,DV
OSPF	IGP,LS
BGP	EGP

RIP

- RIP
  - Routing Information Protocol
- Category
  - Interior routing protocol
  - Distance-vector routing protocol
    - Using "hop-count" as the cost metric
- Example of how RIP advertisements work

Destination network	Next router	# of hops to destination
1	A	2
20	B	2
30	B	7

Routing table in router before Receiving advertisement

Destination network	Next router	# of hops to destination
30	C	4
1	--	1
10	--	1

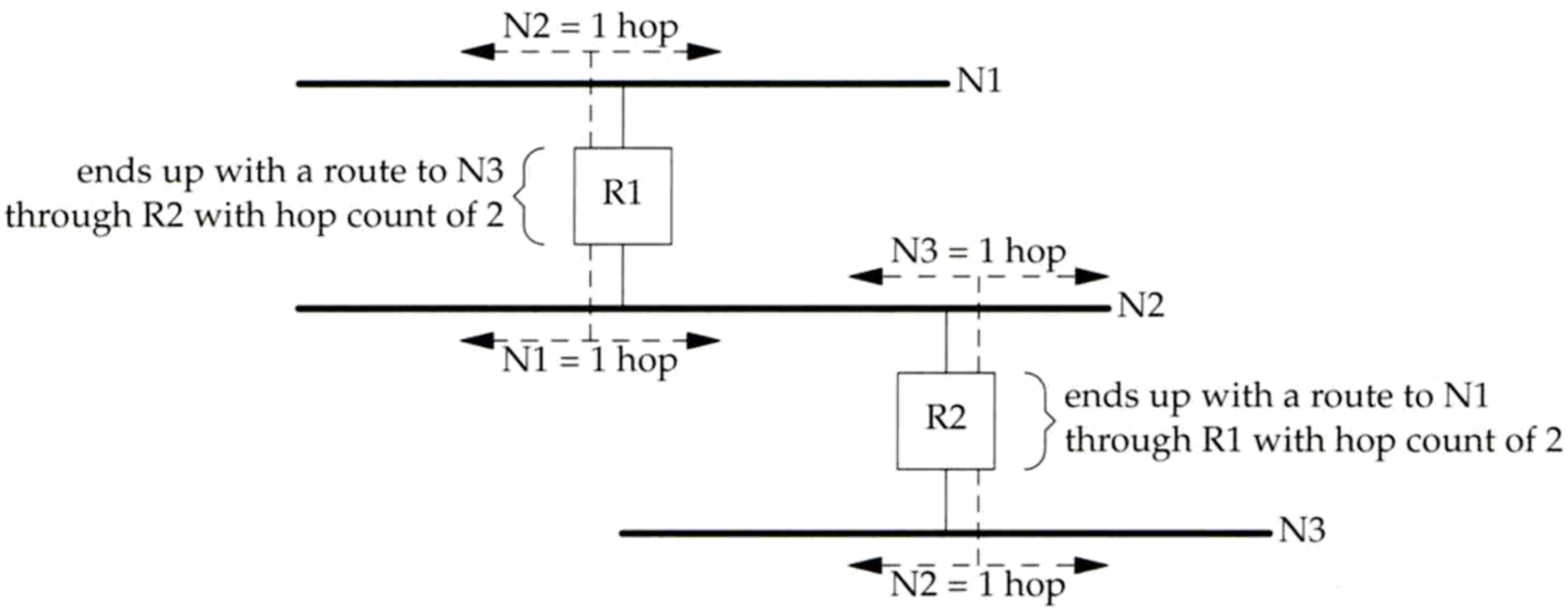
Advertisement from router A

Destination network	Next router	# of hops to destination
1	A	2
20	B	2
30	A	5

Routing table after receiving advertisement

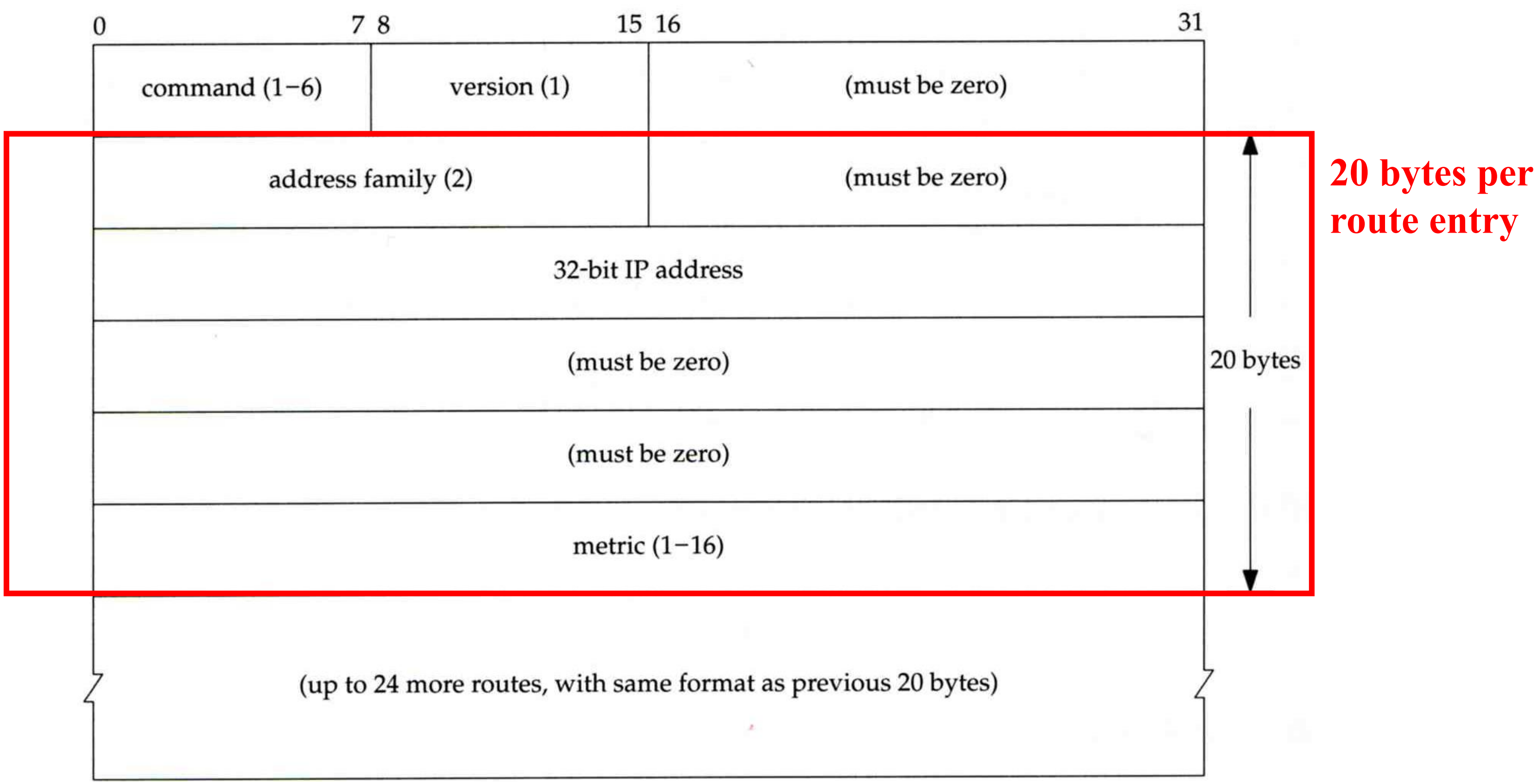
RIP
– Example

- Another example



RIP – Message Format

- RIP message is carried in UDP datagram
  - Command: 1 for request and 2 for reply
  - Version: 1 or 2 (RIP-2)

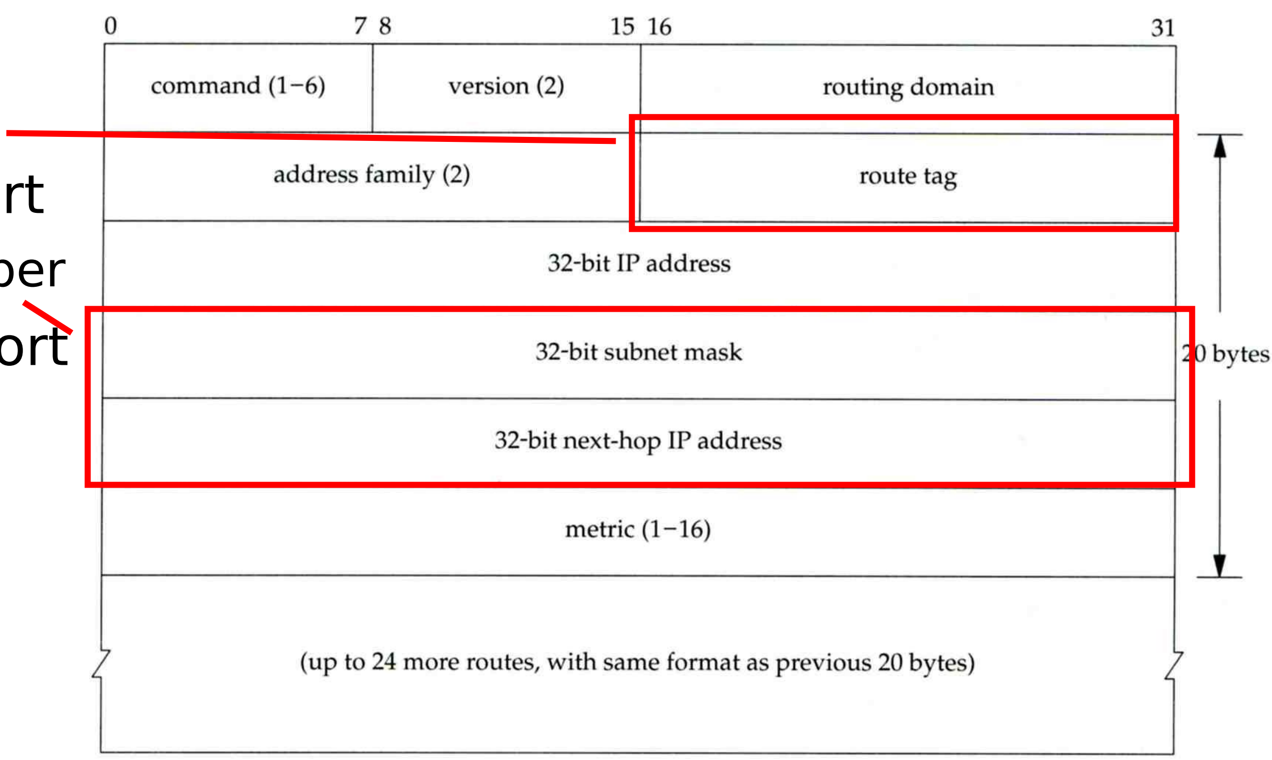


RIP – Operation

- routed – RIP routing daemon
  - Operated in UDP port 520
- Operation
  - Initialization
    - Probe each interface
    - send a request packet out each interface, asking for other router's complete routing table
  - Request received
    - Send the entire routing table to the requestor
  - Response received
    - Add, modify, delete to update routing table
  - Regular routing updates
    - Router sends out their routing table to every neighbor every 30 minutes
  - Triggered updates
    - Whenever a route entry's metric change, send out those changed part routing table

RIP – Problems of RIP

- Issues
  - 15 hop-count limits
  - Take long time to stabilize after the failure of a router or link
  - No CIDR
- RIP-2
  - EGP support
    - AS number
  - CIDR support



IGRP (1)

- IGRP – Interior Gateway Routing Protocol
- Similar to RIP
  - Interior routing protocol
  - Distance-vector routing protocol
- Difference between RIP
  - Complex cost metric other than hop count
    - delay time, bandwidth, load, reliability
    - The formula

$$\left(\frac{bandwith\_weight}{bandwith * (1 - load)} + \frac{delay\_weight}{delay}\right) * reliability$$

  - Use TCP to communicate routing information
  - Cisco System’s proprietary routing protocol

IGRP (2)

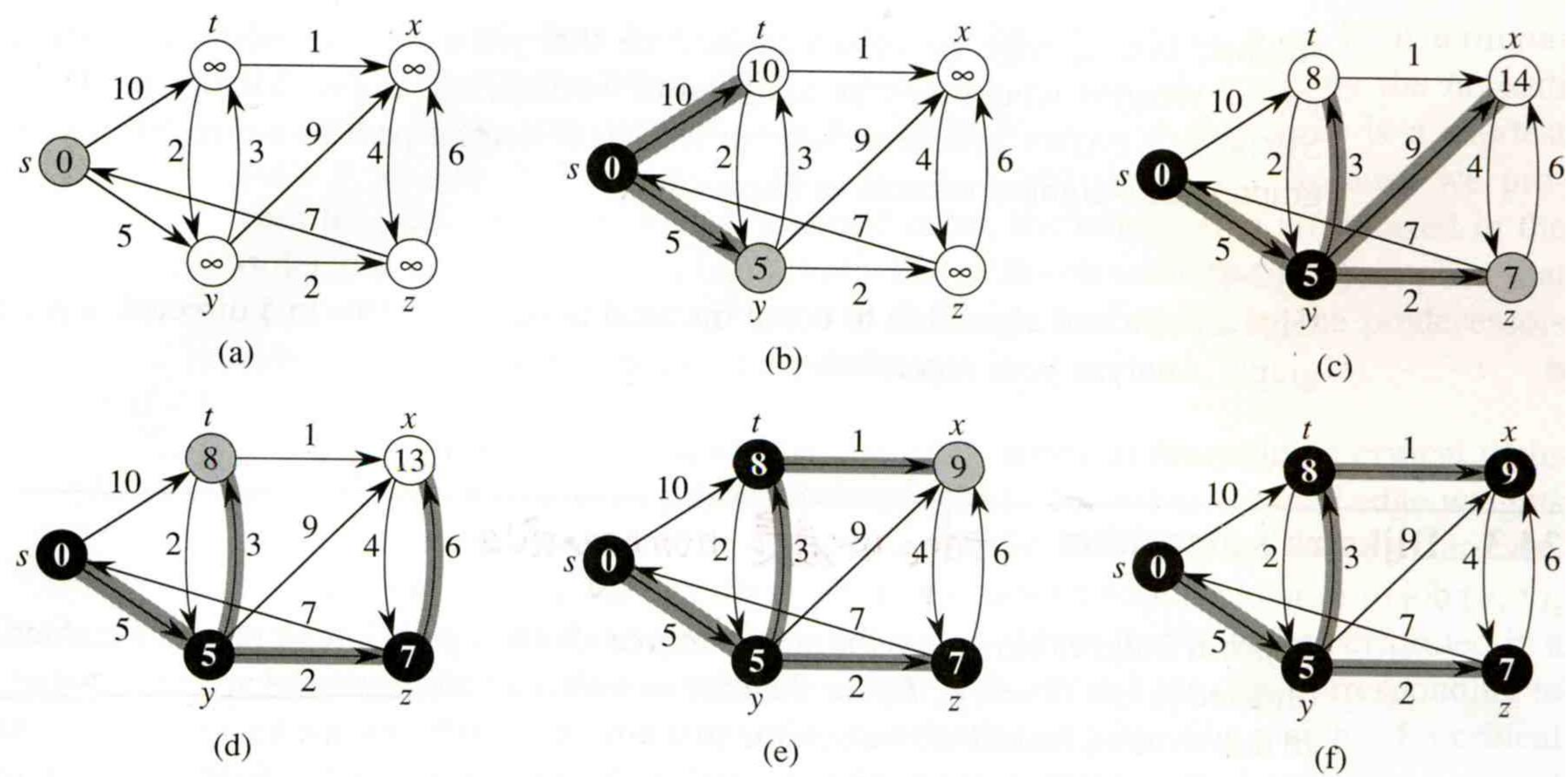
- Advantage over RIP
  - Control over metrics
- Disadvantage
  - Still classful and has propagation delay
  - Vendor dependency

OSPF (1)

- OSPF
  - Open Shortest Path First
- Category
  - Interior routing protocol
  - Link-State protocol
- Each interface is associated with a cost
  - Generally assigned manually
  - The sum of all costs along a path is the metric for that path
- Neighbor information is broadcast to all routers
  - Each router will construct a map of network topology
  - Each router run Dijkstra algorithm to construct the shortest path tree to each routers

OSPF
– Dijkstra Algorithm

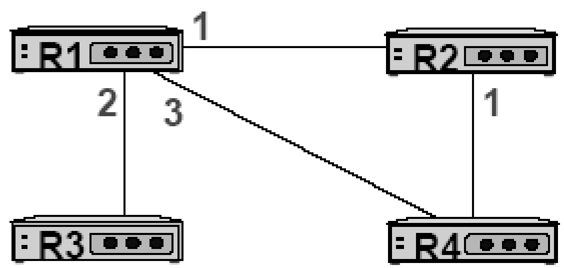
- Single Source Shortest Path Problem
  - Dijkstra algorithm use "greedy" strategy
  - Ex:



OSPF
– Routing table update example (1)

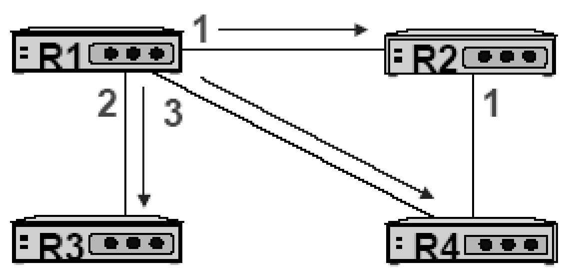
R1

D	Path	M
R1	?	
R2		
R3		
R4		



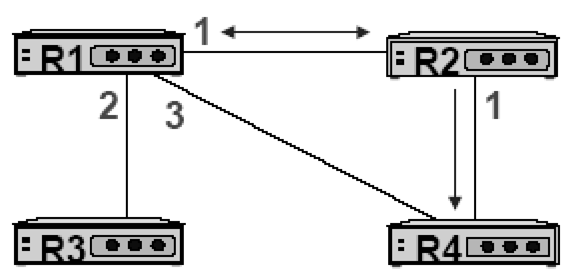
R1

D	Path	M
R1	direct	0
R2	*R1-R2*	*1*
R3	*R1-R3*	*2*
R4	*R1-R4*	*3*



R1

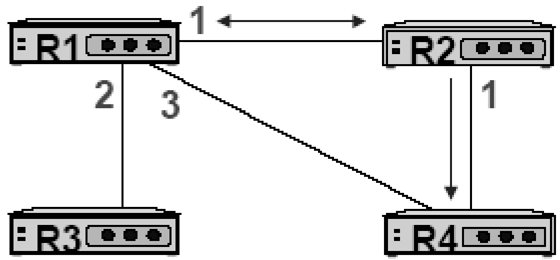
D	Path	M
R1	direct	0
R2	R1-R2	1
R3	*R1-R3*	*2*
R4	*R1-R4*	*3*



OSPF
– Routing table update example (2)

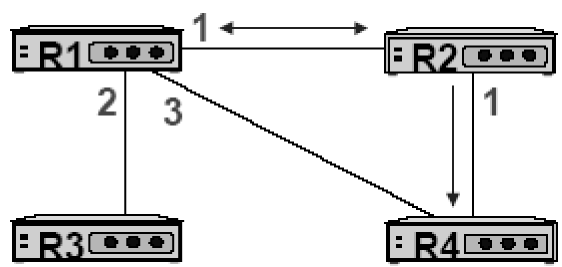
R1

D	Path	M	
R1	direct	0	✓
R2	R1-R2	1	✓
R3	*R1-R3*	*2*	
R4	*R1-R4*	*3*	



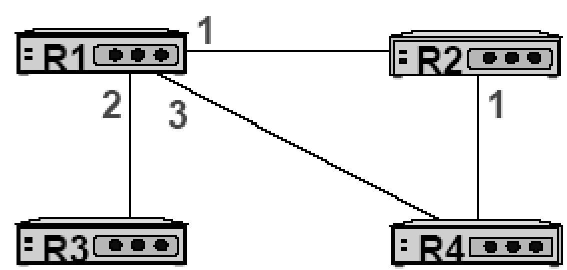
R1

D	Path	M	
R1	direct	0	✓
R2	R1-R2	1	✓
R3	*R1-R3*	*2*	
R4	*R1-R2-R4*	*2*	



R1

D	Path	M	
R1	direct	0	✓
R2	R1-R2	1	✓
R3	R1-R3	2	✓
R4	R1-R2-R4	2	✓



OSPF
– Summary

- Advantage
  - Fast convergence
  - CIDR support
  - Multiple routing table entries for single destination, each for one type-of-service
    - Load balancing when cost are equal among several routes
- Disadvantage
  - Large computation

ISIS (1)

- ISIS
  - Intermediate System to Intermediate System
- Category
  - Interior routing protocol
  - Link-State protocol
- Each interface is associated with a cost
  - Generally assigned manually
  - The sum of all costs along a path is the metric for that path
- Neighbor information is broadcast to all routers
  - Each router run Dijkstra algorithm to construct the shortest path tree to each routers
- Rides directly above layer two
  - I/IS-IS runs on top of the Data Link Layer

Comparing ISIS and OSPF (1)

- Same
  - Interior routing protocol (IGP)
  - Link-State protocol
  - Classless Inter-Domain Routing (CIDR)
  - Variable Subnet Length Masking (VLSM)
  - Authentication
  - Multi-path
  - IP unnumbered links

Comparing ISIS and OSPF (2)

- OSPF
  - Host
  - Router
  - Link
  - Packet
  - Designated Router (DR)
  - Backup DR (BDR)
  - Link-Stats Advertisement (LSA)
  - Hello packet
  - Database Description(DBD)

- ISIS
  - End System(ES)
  - Intermediate System(IS)
  - Circuit
  - Protocol Data Unit (PDU)
  - Designated IS (DIS)
  - N/A
  - Link-State PDU (LSP)
  - IIH PDU
  - Complete sequence number PDU (CSNP)

Comparing ISIS and OSPF (3)

- OSPF
  - Area
  - Non-backbone area
  - Backbone area
  - Area Border Router(ABR)
  - Autonomous System Boundary Router (ASBR)
- ISIS
  - Sub domain (area)
  - Level-1 area
  - Level-2 Sub domain (backbone)
  - L1L2 router
  - Any IS

BGP (1)

- BGP
  - Border Gateway Protocol
- Exterior routing protocol
  - Now BGP-4
  - Exchange network reachability information with other BGP systems
- Routing information exchange
  - Message:
    - Full path of autonomous systems that traffic must transit to reach destination
    - Can maintain multiple route for a single destination
  - Exchange method
    - Using TCP
    - Initial: entire routing table
    - Subsequent update: only sent when necessary
    - Advertise only optimal path
- Route selection
  - Shortest AS path

BGP (2)

- Incremental Updates
- Many options for policy enforcement
- Classless Inter Domain Routing (CIDR)
- Widely used for Internet backbone
- Autonomous systems

```
140.113.0.0/16 *[BGP/170] 1w1d 02:30:41, localpref 200, from 62.115.128.39 AS path: 9505 18185 9916 I
```

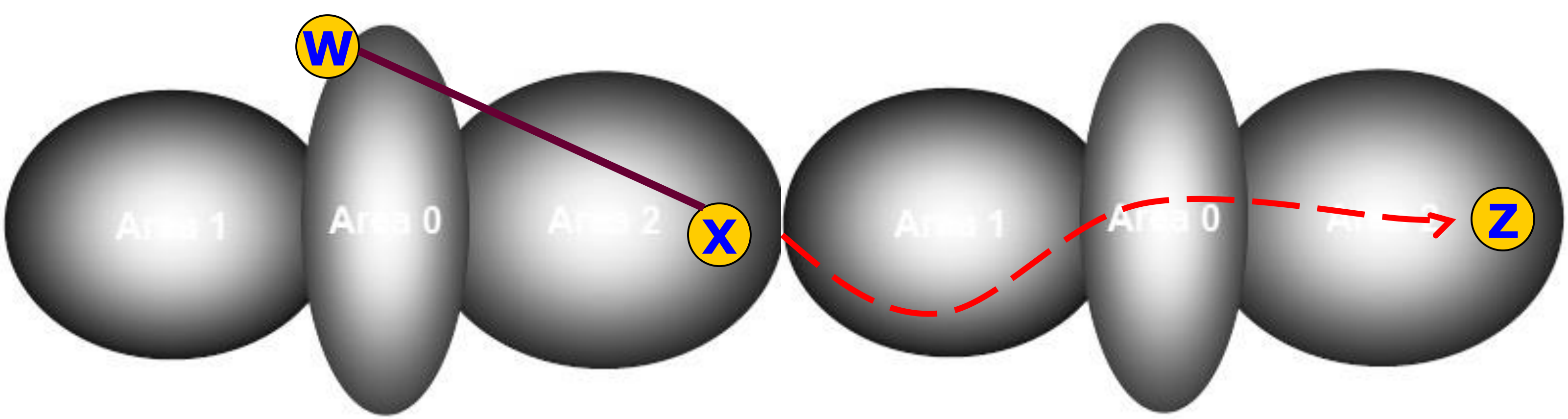
https://nsrc.org/workshops/2016/senix-ixp/presentations/00-BGP-Introduction.pdf

Routing Protocols Comparison

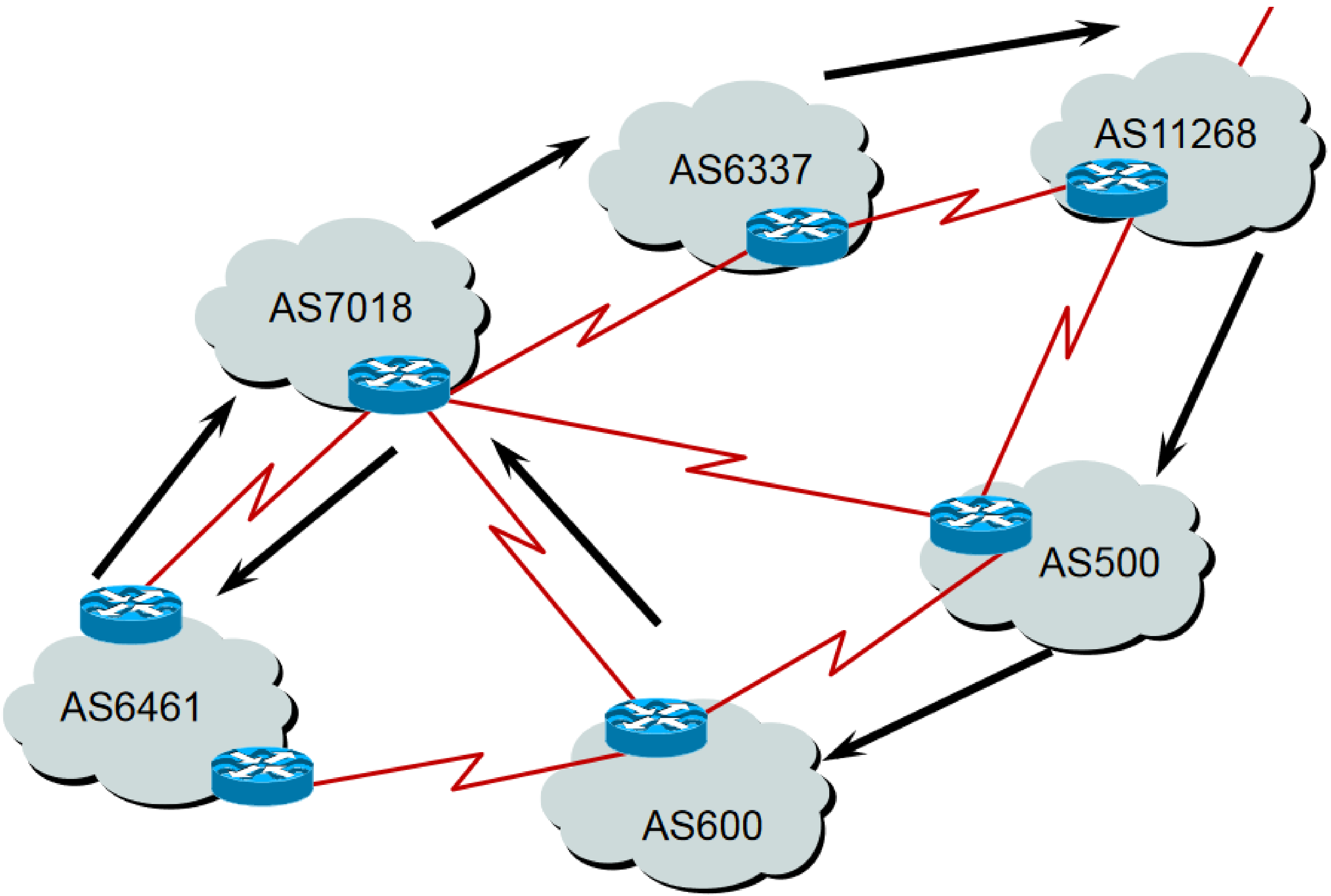
	RIP	IGRP	OSPF	BGP4
DV or LS	DV	DV	LS	Path Vec
TCP/UDP & Port	U - 520	IP - 9	T - 89	T - 179
Classless	No	No	Yes	Yes
Updates	Per.	Per.	Both	Trig.
Load Balance	No	Yes	Yes	No
Internal / External	Int.	Int.	Int.	Ext.
Metric	Hop Count	Load Errors Delay Bdwth	Sum of Int. Cost	Short. AS Path

BGP
– Operation Example

- How BGP work
  - The whole Internet is a graph of autonomous systems
  - X→Z
    - Original: X→A→B→C→Z
    - X advertise this best path to his neighbor W
  - W→Z
    - W→X→A→B→C→Z



BGP - Path Vector Protocol



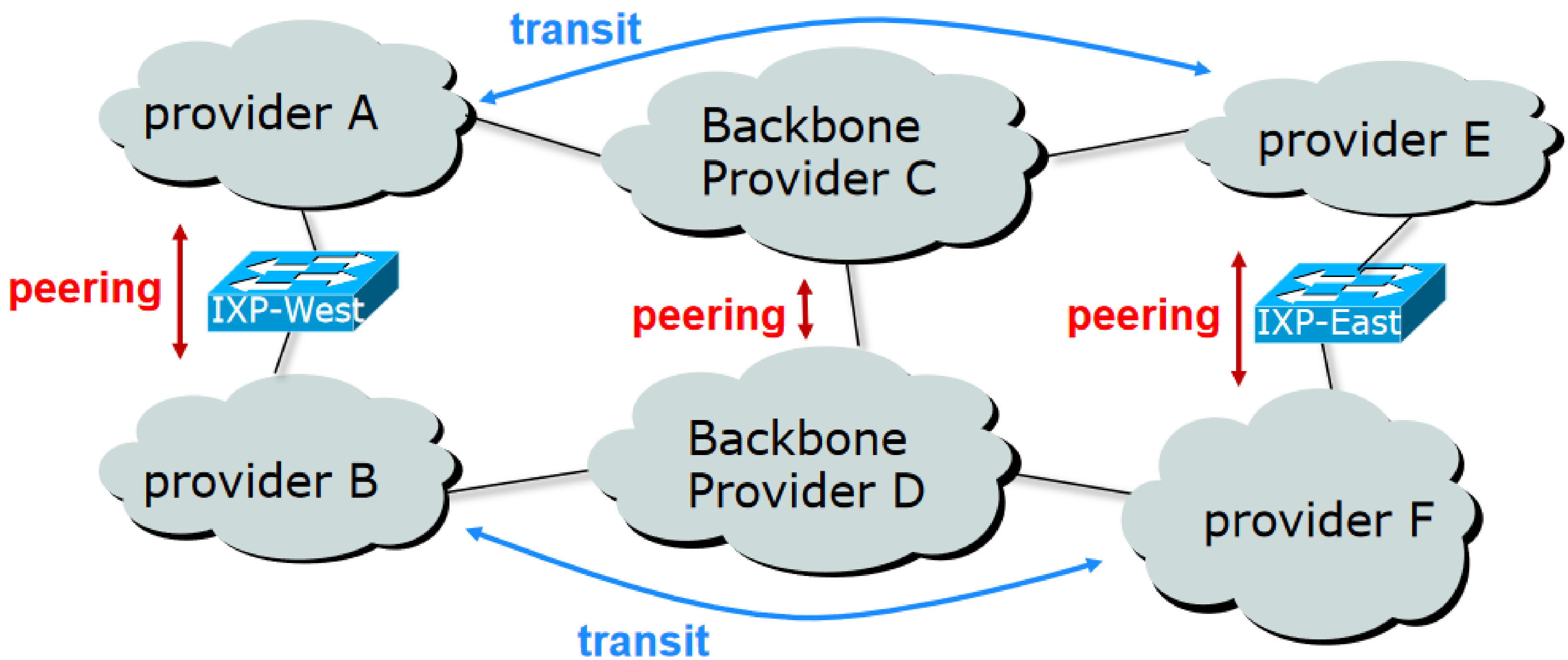
https://nsrc.org/workshops/2016/senix-ixp/presentations/00-BGP-Introduction.pdf

BGP - Definitions

- Transit
  - carrying traffic across a network
  - (Commercially: for a fee) but in Taiwan…
- Peering
  - exchanging routing information and traffic
  - (Commercially: between similar sized networks, and for no fee) but in Taiwan…
- Default
  - where to send traffic when there is no explicit match in the routing table

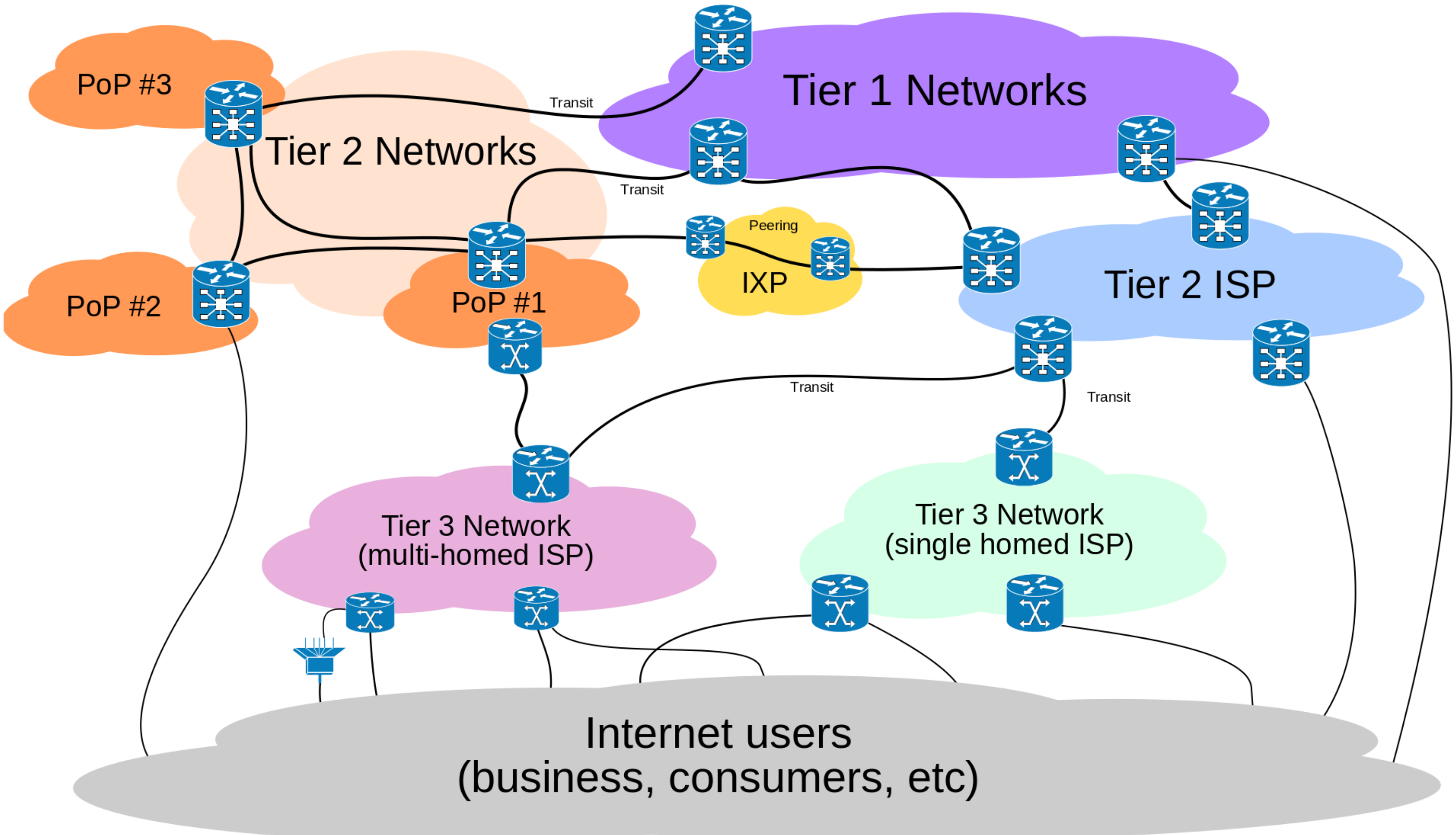
https://nsrc.org/workshops/2016/senix-ixp/presentations/00-BGP-Introduction.pdf

BGP - Peering and Transit example



A and B peer for free, but need transit arrangements with C and D to get packets to/from E and F

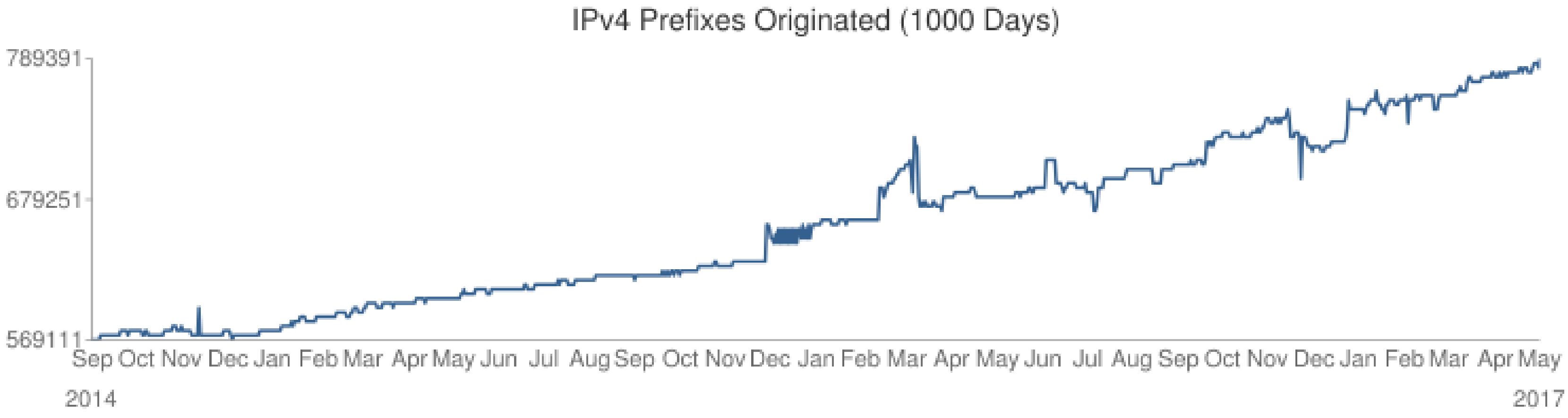
BGP – World Wide (1)



https://en.wikipedia.org/wiki/Tier_1_network

BGP – World Wide (2)

- Default route
  - End of full routing table
- Full route
  - Transit from other ISP / IXP
  - 789K – IPv4
  - 58K – IPv6



http://bgp.he.net/report/prefixes#_prefixes

BGP – Full Route

IPv4 Announced Prefix Count by CIDR and Growth over 1000 Days

CIDR:	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Count:	18	13	39	105	291	565	1,097	2,025	14,550	9,421	16,089	30,495	46,639	52,838	97,659	81,419	436,129
Change:	2	1	9	17	26	55	57	232	1,063	1,397	3,041	3,539	7,518	10,249	32,380	27,633	136,893
Percent:	12%	8%	30%	19%	9%	10%	5%	12%	7%	17%	23%	13%	19%	24%	49%	51%	45%

IPv6 Announced Prefix Count by CIDR and Growth over 1000 Days

CIDR:	16	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Count:	14	2	10	3	5	4	20	5	89	16	79	1,730	155	121	9,633
Change:	3	0	3	0	1	0	5	0	75	2	19	1,229	72	66	3,323
Percent:	27%	0%	42%	0%	25%	0%	33%	0%	535%	14%	31%	245%	86%	120%	52%

CIDR:	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
Count:	631	420	363	1,528	217	674	216	2,392	210	394	73	2,728	181	1,770	550	21,408
Change:	393	259	161	720	165	505	146	1,465	122	252	41	1,834	88	1,452	395	11,938
Percent:	165%	160%	79%	89%	317%	298%	208%	158%	138%	177%	128%	205%	94%	456%	254%	126%

CIDR:	49	50	51	52	54	55	56	57	58	59	60	61	62	63	64
Count:	45	29	5	67	25	10	4,641	2	3	6	13	1	4	3	8,166
Change:	15	22	3	46	23	9	4,459	1	2	5	6	0	3	2	7,245
Percent:	50%	314%	150%	219%	1150%	900%	2450%	100%	200%	500%	85%	0%	300%	200%	786%

BGP Route Hijacking

- ❑ Bad?
- ❑ Good?
- ❑ Neutral?

BGP Route Hijacking Howto

- BGP normally
  - Exchange “reachability” information between each other
  - Advertises the block of addresses to neighboring BGP
- IF someone
  - Advertise the addresses that does not belong to you
  - Your neighboring BGP announce to others
- BGP hijack explained
  - https://www.youtube.com/watch?v=9NBv7lKrG1A

BGP Route Hijacking

- DDoS mitigation to clean center
- BGP anycasting
  - Like 168.95.1.1, 8.8.8.8, 8.8.4.4