

ECE 435 – Network Engineering

Lecture 13

Vince Weaver

<http://web.eece.maine.edu/~vweaver>

vincent.weaver@maine.edu

1 March 2022

Announcements

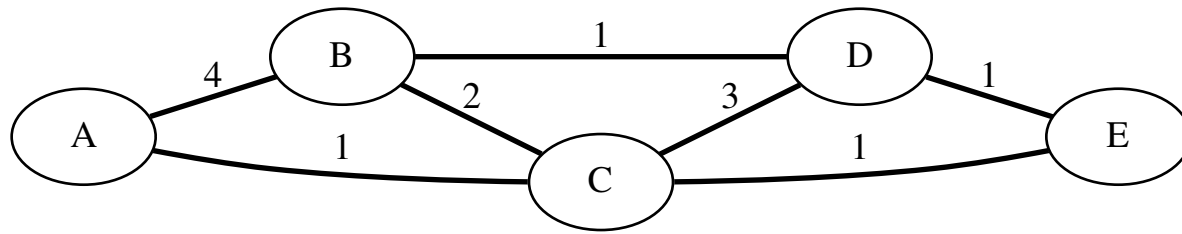
- Midterm on the 8th
- Project was posted
- Interesting IPv4 address Allocation link:
<https://blog.daknob.net/mapping-44net/>



(dynamic) Distance Vector Routing

- Used by ARPANET until 1979
- Asynchronous, distributed, uses local info
- Each router maintains a table (vector) giving best known distance to each destination and line to use to get there
- First line shows out starting info they all know.
Each iteration shows as the info from neighbors is passed on and the routing tables are updated.





DVR example

1. Start with what you know
2. Send routing table to neighbor
3. Update if find shorter route. This is all happening simultaneously
4. Should converge on Dijkstra.



DVR example

D=destination, C=cost, N=next hop

| | A | | | B | | | C | | | D | | | E | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | D | C | N | D | C | N | D | C | N | D | C | N | D | C | N |
| 1 | B | 4 | B | A | 4 | A | A | 1 | A | B | 1 | B | C | 1 | C |
| | C | 1 | C | C | 2 | C | B | 2 | B | C | 3 | C | D | 1 | D |
| | | | | D | 1 | D | D | 3 | D | E | 1 | E | | | |
| | | | | E | 1 | E | E | 1 | E | | | | | | |
| 2 | B | 3 | C | A | 3 | C | A | 1 | A | A | 4 | C | A | 2 | C |
| | C | 1 | C | C | 2 | C | B | 2 | B | B | 1 | B | B | 2 | D |
| | D | 4 | C | D | 1 | D | D | 2 | E | C | 2 | E | C | 1 | C |
| | E | 2 | C | E | 2 | D | E | 1 | E | E | 1 | E | D | 1 | D |
| 3 | B | 3 | C | | | | | | | A | 3 | E | | | |
| | C | 1 | C | | | | | | | B | 1 | B | | | |
| | D | 3 | C | | | | | | | C | 2 | E | | | |
| | E | 2 | C | | | | | | | E | 1 | E | | | |

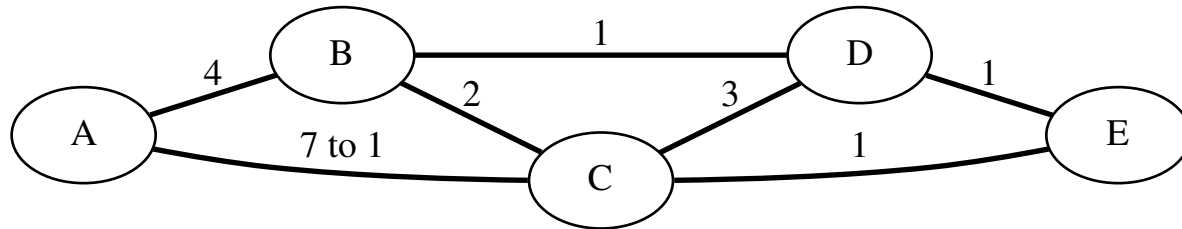


Problems with DVR

- Looping problems: packets can get stuck in loops.
- Good news travels fast, bad news travels slowly.



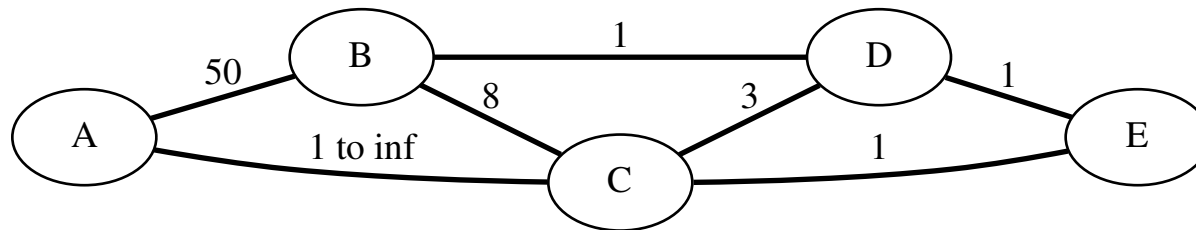
Good News Travels Fast



- Converges in two steps.



Bad News Travels Slowly



- A to C line goes down.
- Have bad timing. (Note: really need better description here)
- C thinks fastest to A is E, E thinks fastest to A is still C.
- C tells B+E cost to A is inf. E (old) tells C cost to A still 2.



- C updates with this info, path to A is 3 if go via E
- E updates path to A is 4 if go via C
- slowly loop, “counting to infinity”



Solutions to Counting to Infinity

- Split horizon – a router should not tell neighbor back the least cost it just got from that neighbor
- Poison Reverse – instead of not telling back, should say the cost back to itself is infinity
- These only work for two hop loops. Other options to send additional “next hop” data, or have a “hold down timer” that lets things settle before updating info



(dynamic) Link State Routing

- Problems with DVR: did not take delay into account, took too long to converge
- Instead, send entire routing table to everyone. Each node then rebuilds own.
- Each router must:
 1. Discover neighbors and learn network address
 2. Measure delay or cost of each neighbor
 3. Construct a packet telling all it learned
 4. Send a packet to all other routers



5. Compute the shortest path to all other routers
 - Learning about neighbors: sends HELLO packet at boot out all links
 - Measure line cost: Send special ECHO packet and measure return. Take into account load?
 - Building link-state packets
 - Distributing
 - Computing new routes



Hierarchical Routing

- At some point not possible for every router to know about every other
- Split into regions
- Example?



Internetworking

- Robert Metcalfe (one of inventors of Ethernet)
Metcalfe's Law: networks value is the square of the nodes
- Joining networks together of different types
- Might have to convert packets at boundaries
- Or tunnel
- What if packets too big for size limit?
 - Fragmentation (difficult)
 - Path MTU (Maximum Transmit Unit) discovery



The Internet Protocol v4

- RFC791
- Network of “autonomous systems” interconnected
- Transport layer takes data and breaks into dataframes of up to 64kB. Sent through Internet (possibly broken up) and when get to other side reconstructed by network layer and passed up to transport layer.
- Global and unique address.
- Need hierarchical structure to locate IP address globally



IPv4 Addresses

- Each IP address is 32-bits and has network address and host ID
- Can write many ways: decimal, hex, (all equivalent) but most common is dotted decimal (i.e. 12.34.56.78)
- Unique to *interface* not necessarily to *host*.



Who Hands these Out?

- ICANN and various regional authorities Internet Corporation for Assigned Names and Numbers Internet Assigned Numbers Authority (IANA)
- Regional Internet Registrars
 - AfriNIC (Africa)
 - ARIN (N America),
 - APNIC (Asia-pacific)
 - LACNIC (latin america),
 - RIPE NCC (Europe and rest)



Subnets

- Number of hosts available can be larger than possible
- Divide network into subnets
- All hosts on subnet have the same prefix (left bits)
- Use subnet mask indicating the leftmost bits to use as subnet
- Can look like 255.255.255.0 meaning only bottom 8 bits are for host

Alternately can write this as 192.168.8.0/24 (24 is number of leading binary 1s in mask)



Classful IP Routing (No Longer Used)

- Routers just shifted right for A, B, and C class. Looked up A and B in table, C in hash table to find where to send
- Has a routing entry for each Class A (128), an entry for each class B (16k). Class C (2 million) a bit much, so hash table.
- Why so simple? In 80s memory and processors were expensive!
- Original classful addressing scheme (not used since 1993)



- Class A: 8 bit network (high bit 0) (24 bits of hosts) 0.0.0.0 to 127.255.255.255
- Class B: 16 bit network, (high bits 10) 128.0.0.0 to 191.255.255.255
- Class C: 24 bit network (high bits 110) 192.0.0.0 to 223.255.255.255
- Class D: multicast (high bits 1110) 224.0.0.0 to 239.255.255.255
- Class E: reserved (high bits 1111) 240.0.0.0 to 255.255.255.255



Reserved IP Ranges

- Private Networks
 - 10.0.0.0/8 private network (RFC 1918)
 - 172.16.0.0/12 private network (RFC 1918)
 - 192.168.0.0/16 Private Network (RFC 1918)
- Loopback
 - 127.0.0.0/8 loopback (RFC 6890)
- 0.0.0.0/8 reserved for current network (RFC 6890)
- 100.64.0.0/10 shared address space (RFC 6598)
- 169.254.0.0/16 link-local (RFC 3927)



- 192.0.0.0/24 IETF (RFC 6890)
- 192.0.2.0/24 test (RFC 5737)
- 192.88.99.0/24 IPv6 to IPv4 relay (RFC 3068)
- 224.0.0.0/4 IP Multicast (class D) (RFC 5771)
- 240.0.0.0/4 Reserved (class E) (RFC 1700)
- 255.255.255.255 Broadcast (RFC 919)



Other IPv4 Conventions

- .0 represents a subnet
See <https://lwn.net/Articles/850374/> really old UNIX treated .0 (or all host bits 0) as another broadcast, there's a push to reclaim it as unicast
- .1 is often (but not always) a router
- it all host bits 1, broadcast for that subnet
- 255.255.255.255 is broadcast for device that doesn't know own IP yet (DHCP)



Classless Inter-Domain Routing (CIDR)

- Running out (have run out) of network addresses
- For many groups, Class-A too big, Class-C too small (three bears problem?)
- Merge neighboring class-C together
- RFC 1519
- Scalability problem: each network takes up space in routing table
- Solution, group neighboring class Cs together
- With CIDR bit more complex.



- Triplet with IP address, subnet mask, outgoing line.
- In theory has to scan all. If multiple matches, one with longest mask is used.
- There are algorithms to make this go faster.



Local IP Routing

- If on same subnet, send packet directly to destination
- Otherwise, send on outgoing. See Linux route command. Often a “default router” 0.0.0.0/0. If doesn't match any other, sent out over default route
- Algorithm: if to same host, skip network. If to same subnet, deliver directly (Ethernet, etc) otherwise, send to default router
- If multiple network interfaces: If to this machine, deliver it, If to directly connected subnet, directly deliver, else



deliver to next hop router

- How do we know if on network? If $((\text{hostIP XOR destip}) \& \text{subnetmask}) == 0$
- If local, how do we map IP to MAC? We'll see that in a minute.
- Due to CIDR, longest prefix matching. If match both a /21 and /24 then 24 is the one to send to as it's the longest.
- Data structures. Hashes? Trie?
 - Linux: two level hashing
 - BSD - trie (prefix tree)



Linux/UNIX routing setup

- Was `route` command, has been replaced by `ip route`
- `route add default gateway sets default gateway (router) for packets leaving the local network`
- also set up local subnets you are on, those packets don't need a router
- more complicated if you are configuring your Linux box to **be** a router

