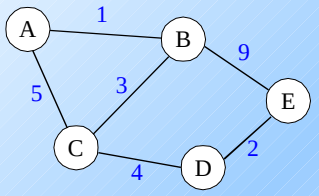
Routing Protocols

Bhaskaran Raman

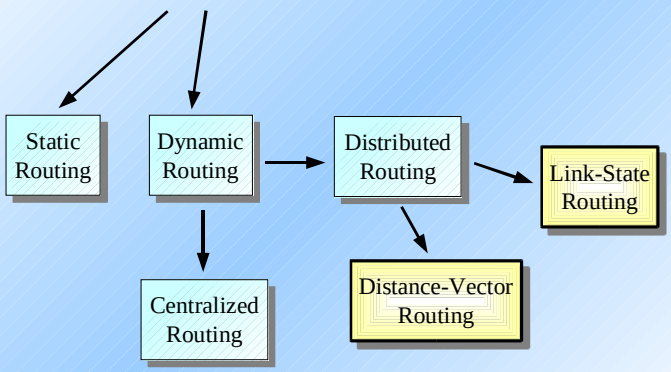
Department of CSE, IIT Kanpur

What is Routing?

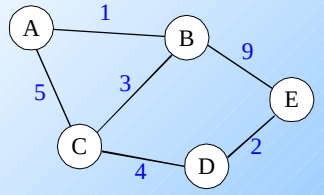


- Routing: process of filling the forwarding table
- View network as a graph
  - Metrics can be defined for links
- Why not use the spanning tree algorithm?
  - Does not scale
  - Specific to ethernet

Routing Protocols: Classification



Distance Vector (DV) Routing



Destn	Cost	NextHop
B	1	B
C	5	C
D	∞	-
E	∞	-

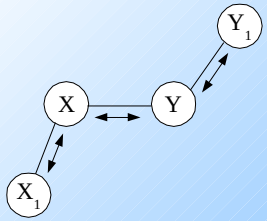
Initial Table at A

Destn	Cost	NextHop
B	1	B
C	4	B
D	8	B
E	10	B

Final Table at A

- Idea: each node maintains a distance vector
- Initial state: distance (cost) to neighbours is known
- Final state: distance (cost) to all nodes is known, and the next-hop

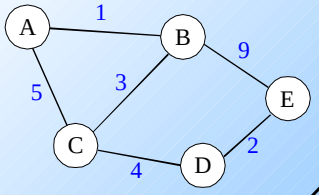
How does DV work?



- Each node sends its DV to its neighbours
- On receiving a DV from a neighbour, a node updates its DV
  - Choose minimum of current cost and new cost
  - Also update next-hop appropriately
- If there are no topology changes, the algorithm converges to a stable state

- After one message, each node knows about nodes two hops away
- After two messages, each node knows about nodes three hops away
- And so on...

An Example run of DV



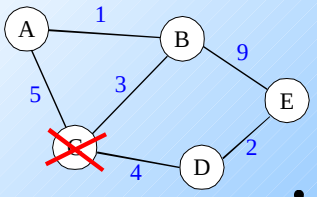
Destn	Cost	NextHop
B	1	B
C	5	C
D	∞	-
E	∞	-

Table at A

Destn	Cost	NextHop
B	1	B
C	4	B
D	9	C
E	10	B

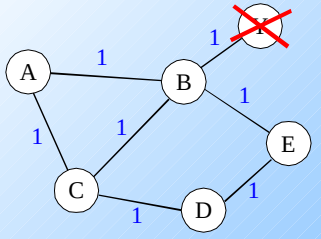
Destn	Cost	NextHop
B	1	B
C	4	B
D	8	B
E	10	B

Dynamic Updates in DV



- Periodic updates from each node X:
  - To tell others that X is still running
  - To update others' DV in case some route becomes invalid
- Triggered updates from each node X:
  - Whenever the DV at X changes
  - So that invalid routes are not used for long

Problems with DV



Distance to Y

A	B	C
2	∞	2
∞	∞	2
3	∞	∞
3	4	∞
∞	4	∞
∞	∞	5
...	...	...

- Counting to infinity
  - Make infinity small
  - Split horizon (with poisoned reverse)
  - Does not work for loops with more than 2 nodes
- Convergence is slow
- Alternate approach to routing: link-state

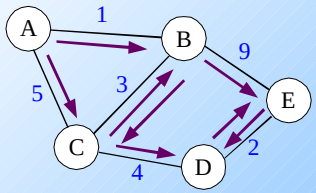
Some Remarks

- DV also known as the Bellman-Ford algorithm, after its inventors
- Routing Information Protocol (RIP) is the standard implementation of DV routing
- Path-vector routing is a variation of distance-vector
  - Each node sends to its neighbours not just the cost, but the entire path to the destination
  - Avoids the looping problem of DV

Link-State (LS) Routing

- Initial state: each node knows the cost to its neighbours
- Final state: each node knows the entire graph
- Depends on two primitives:
  - Reliable flooding: each node sends its link-state to everyone reliably
  - Route calculation: after knowing the graph, each node computes its route to every other node

Reliable Flooding



An LSP flooded from A

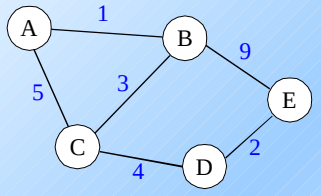
An LSP from node X contains:
- The id of node X
- The link-state of X: cost of each links of X
- An LSP sequence number
- A Time-To-Live (TTL)

- Each node floods a Link-State Packet (LSP)
- Reliability through retransmission on each link
- A node forwards an LSP on all links except the one from where the LSP was received
- Only "newer" LSPs are forwarded: based on LSP sequence number

More on Flooding

- LSP floods are generated:
  - Periodically, as well as on a link-state change
- After each node's flood has propagated through the network, each node has the entire graph information
- How to find the shortest-path to a given destination?
  - Route calculation using Dijkstra's algorithm

Dijkstra's Algorithm for Route Calculation



Node	Cost	NextHop	Confirmed?
A	0	A	Yes
B	1	B	No
C	5	C	No
D	∞	-	No
E	∞	-	No

Choose min-cost among unconfirmed, and expand

Node	Cost	NextHop	Confirmed?
A	0	A	Yes
B	1	B	Yes
C	4	B	No
D	∞	-	No
E	10	B	No

Choose min-cost among unconfirmed, and expand

Node	Cost	NextHop	Confirmed?
A	0	A	Yes
B	1	B	Yes
C	4	B	Yes
D	8	B	No
E	10	B	No

Two more such steps

Node	Cost	NextHop	Confirmed?
A	0	A	Yes
B	1	B	Yes
C	4	B	Yes
D	8	B	Yes
E	10	B	Yes

Rules for Dijkstra's Calculation

- Dijkstra calculation at node X:
  - Initialize table with X having 0 cost and confirmed
  - And X's neighbours with appropriate costs, but unconfirmed
  - At each step:
    - Choose the min-cost node M from the unconfirmed list
    - Mark M as confirmed
    - Expand M: update costs for each of M's unconfirmed neighbours
  - Stop when all nodes are marked confirmed

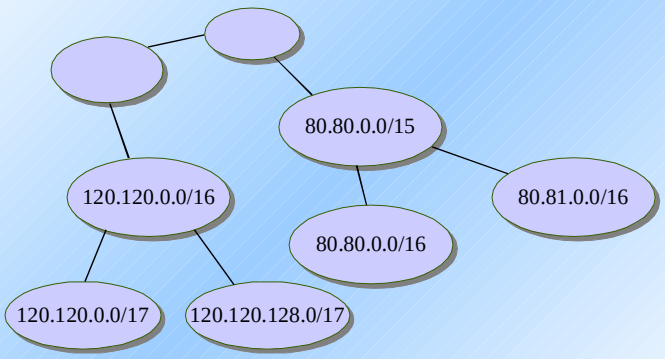
Some Remarks

- No problem of looping since each node has global information
- Convergence is fast
- But, scaling problems due to:
  - Flooding, computation, amount of information storage required at each node
- Open Shortest Path First (OSPF) is the standard implementation of LS routing

How to Scale Routing?

- Number of Internet hosts ~ 350 million
- Hierarchical routing: information is not global
- Requires hierarchical naming!
- Internet addresses are hierarchical
  - Class A, Class B, Class C addresses
  - Class-less Inter-Domain Routing (CIDR)
  - Address aggregation

Hierarchical Addressing & Routing



Autonomous Systems (AS)

- The Internet is divided into various ASes
  - An AS is also called a domain
- Each AS has an AS number, and participates in inter-domain routing
  - Border Gateway Protocol (BGP) is a path-vector protocol used for inter-domain routing
  - Within each domain, RIP or OSPF may be used
- O(10,000) ASes on the Internet today

Internet Service Providers (ISPs)

- A business entity which provides *Internet connectivity* to its customers
- Example: VSNL, AT&T, Sprint
- Customer gets address(es) and connectivity
- Hierarchy of ISPs exist
  - Higher levels in the hierarchy have an AS number

ISP Relationships

- Business relationships between ISPs:
  - Peer-Peer
  - Provider-Customer (Transit relationship)
- ISP relationships are NOT transitive
  - Connectivity is not transitive

The ISP or AS Hierarchy

- Autonomous System
  - An ISP in most cases
  - A large ISP may consist of many ASes
- Hierarchy of ISP relationships
  - Tier-1, Tier-2, Tier-3...
  - Dense-core, Transit-core, Outer-core, Regional-ISPs, Customers

Summary

- Routing: process of forming forwarding tables
- Distance-Vector versus Link-State
- Hierarchical routing and addressing
- Autonomous systems and ISPs on the Internet