Routing
Figure 7.6 Metropolitan Area Network (MAN)

- Gateway
- To internet or wide area network
- Backbone
- Organization Servers
- Departmental Server
Wide Area Network (WAN)

Interdomain level

Border routers

Autonomous system or domain

LAN level

Intradomain level

Border routers

Internet service provider

Leon-Garcia & Widjaja: Communication Networks

Figure 7.7
Advanced Computer Networks: Routing

National service provider A

Network Access Point

National service provider B

National service provider C

(b)

LAN

Route server

Copyright ©2000 The McGraw Hill Companies

Leon-Garcia & Widjaja: Communication Networks
Datagram Packet Switching
## Datagram Network Routing Table

<table>
<thead>
<tr>
<th>Destination address</th>
<th>Output port</th>
</tr>
</thead>
<tbody>
<tr>
<td>0785</td>
<td>7</td>
</tr>
<tr>
<td>1345</td>
<td>12</td>
</tr>
<tr>
<td>1566</td>
<td>6</td>
</tr>
<tr>
<td>2458</td>
<td>12</td>
</tr>
</tbody>
</table>
Virtual Circuit Network

Packet

Packet
## Virtual Circuit Network Routing Table

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Output port</th>
<th>Next identifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>13</td>
<td>44</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>23</td>
</tr>
<tr>
<td>27</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>58</td>
<td>7</td>
<td>34</td>
</tr>
</tbody>
</table>

Entry for packets with identifier 15
Routing

Routing algorithm:: that part of the Network Layer responsible for deciding on which output line to transmit an incoming packet.

Remember: For virtual circuit subnets the routing decision is made ONLY at set up.

Algorithm properties:: correctness, simplicity, robustness, stability, fairness, optimality, and scalability.
Routing Classification

Adaptive Routing
- based on current measurements of traffic and/or topology.

1. centralized
2. isolated
3. distributed

Non-Adaptive Routing
- routing computed in advance and off-line

1. flooding
2. static routing using shortest path algorithms
Flooding

• *Pure flooding* :: every incoming packet to a node is sent out on *every outgoing line*.
  – Obvious adjustment – do not send out on arriving link (assuming full-duplex links).
  – The routing algorithm can use a hop counter (e.g., TTL) to *dampen the flooding*.

• *Selective flooding* :: only send on those lines going “approximately” in the right direction.
Shortest Path Routing

1. Bellman-Ford Algorithm [Distance Vector]
2. Dijkstra’s Algorithm [Link State]

What does it mean to be the shortest (or optimal) route?

Choices:

a. Minimize the number of hops along the path.
b. Minimize mean packet delay.
c. Maximize the network throughput.
Metrics

- Set all link costs to 1.
  - Shortest hop routing.
  - Disregards delay.
- Measure the number of packets queued.
  - Did not work well.
- Timestamp $\text{ArrivalTime}$ and $\text{DepartTime}^*$ and use link-level ACK to compute:
  \[
  \text{Delay} = (\text{DepartTime} - \text{ArrivalTime}) + \text{TransmissionTime} + \text{Latency}
  \]

* Reset after retransmission
Metrics

- Unstable under heavy link load.
- Difficulty with granularity of the links.

• Revised ARPANET routing metric:
  - Compress dynamic range of the metric
  - Account for link type
  - Smooth variation of metric with time:
    • Delay transformed into link utilization
    • Utilization was averaged with last reported utilization.
    • Hard limit set on how much the metric could change per measurement cycle.
Figure 4.22 Revised ARPANET routing metric versus link utilization
Dijkstra's Shortest Path Algorithm

Initially mark all nodes (except source) with infinite distance.
working node = source node
Sink node = destination node
While the working node is not equal to the sink
  1. Mark the working node as permanent.
  2. Examine all adjacent nodes in turn

    If the sum of label on working node plus distance from working node to adjacent
    node is less than current labeled distance on the adjacent node, this implies a
    shorter path. Relabel the distance on the adjacent node and label it with the node
    from which the probe was made.

  3. Examine all tentative nodes (not just adjacent nodes) and
    mark the node with the smallest labeled value as permanent.
    This node becomes the new working node.

Reconstruct the path backwards from sink to source.
Internetwork Routing

- Adaptive Routing
  - Centralized
    - [RCC]
  - Distributed
    - [IGP] Intradomain routing
      - [RIP] Distance Vector routing
    - [BGP, IDRP] Interdomain routing
      - [OSPF, IS-IS, PNNI] Link State routing
  - Isolated
    - [ECP] Exterior Gateway Protocols

Advanced Computer Networks: Routing
Adaptive Routing

Basic functions:
1. Measurement of pertinent network data.
2. Forwarding of information to where the routing computation will be done.
3. Compute the routing tables.
4. Convert the routing table information into a routing decision and then dispatch the data packet.
Adaptive Routing

Design Issues:

1. How much overhead is incurred due to gathering the routing information and sending routing packets?

2. What is the time frame (i.e., the frequency) for sending routing packets in support of adaptive routing?

3. What is the complexity of the routing strategy?
Distance Vector Routing

• Historically known as the old ARPANET routing algorithm {or known as Bellman-Ford algorithm}.

Basic idea: each network node maintains a Distance Vector table containing the distance between itself and ALL possible destination nodes.

• Distances are based on a chosen metric and are computed using information from the neighbors’ distance vectors.

Metric: usually hops or delay
Distance Vector Routing

Information kept by DV router:

1. each router has an ID
2. associated with each link connected to a router, there is a link cost (static or dynamic).

Distance Vector Table Initialization

Distance to itself = 0
Distance to ALL other routers = infinity number
Distance Vector Algorithm
[Perlman]

1. A router transmits its **distance vector** to each of its neighbors in a routing packet.

2. Each router receives and saves the most recently received **distance vector** from each of its neighbors.

3. A router **recalculates** its distance vector when:
   a. It receives a **distance vector** from a neighbor containing different information than before.
   b. It discovers that a link to a neighbor has gone down (i.e., a topology change).

   The DV calculation is based on minimizing the cost to each destination.
Distance Vector Routing

![Diagram of a subnet with routers and lines between nodes A, B, C, D, E, F, G, H, I, J, K, and L.]

**Figure 5-9.**

(a) A subnet. (b) Input from A, I, H, K, and the new routing table for J.

**Table 5-9.**

<table>
<thead>
<tr>
<th>To</th>
<th>A</th>
<th>I</th>
<th>H</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>24</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>B</td>
<td>12</td>
<td>36</td>
<td>31</td>
<td>28</td>
</tr>
<tr>
<td>C</td>
<td>25</td>
<td>18</td>
<td>19</td>
<td>36</td>
</tr>
<tr>
<td>D</td>
<td>40</td>
<td>27</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>E</td>
<td>14</td>
<td>7</td>
<td>30</td>
<td>22</td>
</tr>
<tr>
<td>F</td>
<td>23</td>
<td>20</td>
<td>19</td>
<td>40</td>
</tr>
<tr>
<td>G</td>
<td>18</td>
<td>31</td>
<td>6</td>
<td>31</td>
</tr>
<tr>
<td>H</td>
<td>17</td>
<td>20</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td>I</td>
<td>21</td>
<td>0</td>
<td>14</td>
<td>22</td>
</tr>
<tr>
<td>J</td>
<td>9</td>
<td>11</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>K</td>
<td>24</td>
<td>22</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>L</td>
<td>29</td>
<td>33</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

**New estimated delay from J**

- 8 A
- 20 A
- 28 I
- 20 H
- 17 I
- 30 I
- 18 H
- 12 H
- 10 I
- 0 –
- 6 K
- 15 K

**New routing table for J**

Vectors received from J's four neighbors:
- JA delay is 8
- JL delay is 10
- JH delay is 12
- JK delay is 6
Routing Information Protocol (RIP)

- RIP had widespread use because it was distributed with BSD Unix in “routed”, a router management daemon.
- **RIP** is the most used Distance Vector protocol.
- Sends packets every 30 seconds or faster.
- Runs over UDP.
- Metric = hop count
- BIG problem is max. hop count = 16
  - RIP limited to running on small networks!!
- Upgraded to RIPv2
(network_address, distance) pairs

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Command</td>
<td>Version</td>
<td>Must be zero</td>
</tr>
<tr>
<td>Family of net 1</td>
<td>Address of net 1</td>
<td></td>
</tr>
<tr>
<td>Address of net 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance to net 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family of net 2</td>
<td>Address of net 2</td>
<td></td>
</tr>
<tr>
<td>Address of net 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance to net 2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4.17 RIP Packet Format**
1. Each router is responsible for meeting its neighbors and learning their names.

2. Each router constructs a link state packet (LSP) which consists of a list of neighbors’ names and the cost to reach each neighbor.

3. The LSP is transmitted to ALL other routers. Each router stores the most recently generated LSP from each other router.

4. Each router uses complete information on the network topology to compute the shortest path route to each destination node.
Figure 4.18 Reliable LSP Flooding
Reliable Flooding

• The process of making sure all the nodes participating in the routing protocol get a copy of the link-state information from all the other nodes.

• **LSP** contains:
  – Sending router’s node ID
  – List connected neighbors with the associated link cost to each neighbor
  – Sequence number
  – Time-to-live
Reliable Flooding

• First two items enable route calculation
• Last two items make process reliable
  – ACKs and checking for duplicates is needed.
• Periodic Hello packets used to determine the demise of a neighbor.
• The sequence numbers are not expected to wrap around.
  – => field needs to be large (64 bits)
• Provides for authentication of routing messages.
  – 8-byte password designed to avoid misconfiguration.

• Provides additional hierarchy
  – Domains are partitioned into areas.
  – This reduces the amount of information transmitted in packet.

• Provides load-balancing via multiple routes.
Open Shortest Path First (OSPF)

Figure 4.32 A Domain divided into Areas

Backbone area
Open Shortest Path First (OSPF)

- OSPF runs on top of IP, i.e., an OSPF packet is transmitted with IP data packet header.
- Uses Level 1 and Level 2 routers
- Has: backbone routers, area border routers, and AS boundary routers
- LSPs referred to as LSAs (Link State Advertisements)
- Complex algorithm due to five distinct LSA types.
OSP F Terminology

Internal router :: a level 1 router.
Backbone router :: a level 2 router.
Area border router (ABR) :: a backbone router that attaches to more than one area.
AS border router :: (an interdomain router), namely, a router that attaches to routers from other ASs across AS boundaries.
OSPF LSA Types

1. Router link advertisement [Hello message]
2. Network link advertisement
3. Network summary link advertisement
4. AS border router’s summary link advertisement
5. AS external link advertisement
### Figure 4.21 OSF Type 1 Link-State Advertisement

<table>
<thead>
<tr>
<th>LS Age</th>
<th>Options</th>
<th>Type=1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Link-state ID**
- **Advertising router**
- **LS sequence number**

<table>
<thead>
<tr>
<th>LS checksum</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- **Number of links**

<table>
<thead>
<tr>
<th>Link ID</th>
<th>Link data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Link type</th>
<th>Num_TOS</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optional</td>
<td>TOS information</td>
<td></td>
</tr>
<tr>
<td>More links</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Indicates LSA type**
- **Indicates link cost**

---

Advanced Computer Networks: Routing 35
OSPF Areas

Area 0.0.0.0
R8
N7
Area 0.0.0.1
N1 R1 N2 R2 N3
Area 0.0.0.2
N5 R6 N4 R7 N6
Area 0.0.0.3

[AS Border router]
To another AS
R = router
N = network

Advanced Computer Networks: Routing
Copyright ©2000 The McGraw Hill Companies
Leon-Garcia & Widjaja: Communication Networks
Figure 8.33

36
Figure 5-65. The relation between ASes, backbones, and areas in OSPF.
Border Gateway Protocol (BGP)

• The replacement for EGP is BGP. Current version is BGP-4.

• BGP assumes the Internet is an arbitrary interconnected set of AS’s.

• In **interdomain routing** the goal is to find ANY path to the intended destination that is loop-free. The protocols are more concerned with **reachability** than optimality.
Source Routing

A G B C D E F H J