Routing vs. Forwarding

• **Routing:**
  – Computing paths the packets will follow
  – Routers talking amongst themselves
  – Individual router *creating* a forwarding table

• **Forwarding:**
  – Directing a data packet to an outgoing link
  – Individual router *using* a forwarding table
Inter-domain vs. Intra-domain Routing

• Today, an internet can be so large that one routing protocol cannot handle the task of updating the routing tables of all routers. For this reason, an internet is divided into autonomous systems.

• An autonomous system (AS) is a group of networks and routers under the authority of a single administration. Routing inside an autonomous system is called intra-domain routing. Routing between autonomous systems is called inter-domain routing.
AS Number Trivia

- **AS number is a 16-bit quantity**
  - So, 65,536 unique AS numbers
- **Some are reserved (e.g., for private AS numbers)**
  - So, only 64,510 are available for public use
- **Managed by Internet Assigned Numbers Authority**
  - Gives blocks of 1024 to Regional Internet Registries
  - IANA has allocated 39,934 AS numbers to RIRs (Jan’ 06)
- **RIRs assign AS numbers to institutions**
  - RIRs have assigned 34,827 (Jan’ 06)
  - Only 21,191 are visible in interdomain routing (Jan’ 06)
- **Recently started assigning 32-bit AS #s (2007)**
Autonomous System Numbers

AS Numbers are 16 bit values. 64512 through 65535 are “private”
Currently over 20,000 in use.

- Villanova: 10448
- MIT: 3
- Harvard: 11
- Yale: 29
- Princeton: 88
- AT&T: 7018, 6341, 5074, …
- UUNET: 701, 702, 284, 12199, …
- Sprint: 1239, 1240, 6211, 6242, …
- …

whois –h whois.arin.net as10448

ASNumber: 10448
ASName: VILLANOVA-UNIV
ASHandle: AS10448
RegDate: 1997-08-06
Updated: 2001-07-11
Ref: http://whois.arin.net/rest/asn/AS10448

OrgName: Villanova University
OrgId: VILLAN
Address: 800 Lancaster Avenue
City: Villanova
StateProv: PA
PostalCode: 19085
Country: US
RegDate: 1989-03-13
Updated: 2004-02-20
Ref: http://whois.arin.net/rest/org VILLAN
…
Example

Intra-Domain

Intra-domain routing protocol aka Interior Gateway Protocol (IGP), e.g. OSPF, RIP
Inter-Domain

Inter-domain routing protocol aka Exterior Gateway Protocol (EGP), e.g. BGP

Routing Protocols

Routing protocols

Intradomain

Distance vector
RIP

Link state
OSPF

Interdomain

Path vector
BGP
Intra-Domain Distance-Vector Routing (RIP)

Network as a Graph

- Nodes are routers
- Edges are network links
- Each edge has an associated cost
  - desirability of sending over that link
Bellman-Ford Idea

\[ D_{ij} = \text{minimum} \{ (c_{i1} + D_{1j}), (c_{i2} + D_{2j}), \ldots, (c_{iN} + D_{Nj}) \} \]

Legend
- \( D_{ij} \) Shortest distance between \( i \) and \( j \)
- \( c_{ij} \) Cost between \( i \) and \( j \)
- \( N \) Number of nodes

Bellman-Ford Algorithm

```c
Bellman_Ford()
{
    // Initialization
    for (i = 1 to N; for j = 1 to N)
    {
        if(i == j) \( D_{ij} = 0 \) \( c_{ij} = 0 \)
        else \( D_{ij} = \infty \) \( c_{ij} = \text{cost between } i \text{ and } j \)
    }
    // Updating
    repeat
    {
        for (i = 1 to N; for j = 1 to N)
        {
            \( D_{ij} \leftarrow \text{minimum} \{ (c_{i1} + D_{1j}), \ldots, (c_{iN} + D_{Nj}) \} \)
        } // end for
    } until (there was no change in previous iteration)
} // end Bellman-Ford
```
Distance Vector Routing Tables

DV – Initialization of Routing Tables
In distance vector routing, each node shares its routing table with its immediate neighbors periodically and when there is a change.

DV – Updating A’s Routing Table
Updating the Routing Table

- If sender and next-hop are different
  - Choose the path with smaller cost
  - If there is a tie, the old one is kept

- If sender and next-hop are the same
  - Update with the new distance, even if higher

When to Share

- Periodic Update
  - Send distance vector to neighbors periodically
    (normally 30 seconds)

- Triggered Update
  - Send distance vector to neighbors when it changes
    1. After updating the routing table, or
    2. Detects some failure in the neighboring links
**Distance Vector: Two-node Instability**

**Counting to Infinity Instability**

- **Defining Infinity**
  - Most implementations define 16 as infinity

- **Split Horizon**
  - If B routes through A to get to X, B does not advertise this piece of information to X
Counting to **Infinity** Instability

- **Split Horizon with Poison Reverse**
  - Include all routes in a routing update, but set to infinity distances learned from the destination
  - If B routes through A to get to X, B tells A that its distance to X is infinite (“poison” the entry)

*If the instability is between three nodes, stability cannot be guaranteed*
Routing Information Protocol (RIP)

- Implements distance vector routing directly
  - All link costs are 1 (minimum hop route)
  - RIP routers send their advertisements every 30 seconds
  - Infinity represented as 16
- RIP is limited to fairly small networks
  - No more than 16 hops between any two routers

*RIP uses the services of UDP on well-known port 520.*
Intra-Domain Link-State Routing (OSPF)

Concept of Link State Routing
Link-State Routing

- Different philosophy from that of DV routing
- Each node collects information on the entire topology of the domain—the list of nodes and links, how they are connected including the type, cost (metric), and the condition of the links (up or down)
- Each node uses Dijkstra’s algorithm on the local network view to build a routing table.

Link State Knowledge
Building Routing Tables at Each Node

- Create a Link State Packet (LSP):
  - (router ID, list of links, sequence number, TTL)
    - first two items used to calculate the route
    - sequence number used to identify the most recent LSP copy
    - TTL used to prevent infinite flooding loops
- **Flood** the LSP to every other router
- Run **Dijkstra** to build a shortest path tree (SPT)
- Calculate a routing table based on the SPT

Link State Advertisement (LSA)

- The LSA of router A is as follows:
  - Advertising Router: 10.10.10.1
  - Number of links: 4 *(3 links plus router itself)*
  - Description of Link 1: Link ID = 10.1.1.1, Metric = 5
  - Description of Link 2: Link ID = 10.1.2.1, Metric = 2
  - Description of Link 3: Link ID = 10.1.3.1, Metric = 3
  - Description of Link 4: Link ID = 10.10.10.1, Metric = 0
- **Our notation:**
  LSP(A) = {(B, 5), (C, 2), (D, 3)}
LSA Flooding

- Node sends link-state information out its links
- And then the next node sends out all of its links
- … except the one where the information arrived

When to Initiate Flooding

LSP (router ID, list of links, sequence number, TTL)

- Topology change
  - Link or node failure / recovery
  - Link cost change

- Periodically
  - Refresh the link-state information
  - No actual need for this type of flooding (60 mins / 2 hours)
  - Corrects for possible corruption of the data
Dijkstra’s Algorithm

Start
Set root to local node and move it to tentative list

Tentative list is empty?
Yes
No
Among nodes in tentative list, move the one with the shortest path to permanent list.
Add each unprocessed neighbor of last moved node to tentative list if not already there. If neighbor is in the tentative list with larger cumulative cost, replace it with new one.

Stop

Routing Table for A

<table>
<thead>
<tr>
<th>Node</th>
<th>Cost</th>
<th>Next Router</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>—</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>—</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>—</td>
</tr>
<tr>
<td>D</td>
<td>3</td>
<td>—</td>
</tr>
<tr>
<td>E</td>
<td>6</td>
<td>C</td>
</tr>
</tbody>
</table>
Intra-Domain Link-State Implementation
Open Shortest Path First Protocol (OSPF)

OSPF

• Intra-domain routing protocol based on link state routing.
• Its domain is also an autonomous system.
• Adds a number of features to the basic LS algorithm

OSPF packets are encapsulated in IP datagrams.
OSPF Features

- Authentication of routing messages
- Additional hierarchy
  - Partition a domain into areas
  - Necessary to make the system more scalable
- Load balancing
  - Maintain multiple routes of identical cost
  - Distribute traffic evenly over these routes

Areas in an Autonomous System
Areas in OSPF

• A collection of networks with area ID
• Routers inside an area flood the area with routing information (regular Link State routing)
• Area border routers summarize the information about each area and send it to other areas
• Backbone area and backbone routers
  – All areas in AS must be connected to the backbone

Example
Example

Summary LSAs passed by ABRs from Area 0
Summary of LS and DV Routing

- Distance vector and link state routing are **both interior routing protocols**. They can be used inside an AS.
- Both of these routing protocols become **intractable** when the domain of operation becomes large.
- Distance vector routing is subject to **instability** if there is more than a few hops in the domain of operation.
- Link state routing needs a **huge amount of resources** to calculate routing tables. It also creates heavy traffic because of flooding.
- There is a need for a third routing protocol which we call **path vector routing**.

Inter-Domain
Path Vector Routing
Border Gateway Protocol (BGP)
Autonomous Systems

- The basic unit of hierarchy in the Internet
  - Within an AS, the owner decides how routing is done
  - Between ASs, must use BGP (Border Gateway Protocol)

```
traceroute -a <destination> reports AS numbers
```

Interdomain vs. Intradomain

- Intradomain routing (IGP – Interior Gateway Protocols)
  - Routing is done based on metrics
  - Routing domain is one autonomous system

- Interdomain routing (EGP – Exterior Gateway Protocols)
  - Routing is done based on policies
  - Routing domain is the entire Internet
Multiple Routing Processes on a Router

- OSPF Process
  - OSPF Routing tables
- BGP Process
  - BGP Routing tables

Forwarding Table Manager

Forwarding Table

Customers and Providers

- Configure route 192.0.2.0/24 pointing to customer

- Provider obtains access to the Internet through the provider.
- Customer pays provider for service.
Internet Structure

- **Hierarchical AS-level topology**
  - Large, tier-1 providers form nationwide backbone
  - Edges represent business relationships

Customer-Provider Relationship

- Transit network that is a provider may be a customer for another network
The Peering Relationship

Peers provide transit between their respective customers
Peers do not provide transit between peers
Peers (often) do not exchange $$$

Peering Provides Shortcuts

Note that peering reduces upstream traffic
InterDomain Routing Goals

- Privacy
  - ASs don’t want to divulge internal topologies
  - … or their business relationships with neighbors

- Policy
  - No Internet-wide notion of a link cost metric
  - Need control over where you send traffic
  - … and who can send traffic through you

AS Terminology

- Local traffic = traffic with source or destination in AS
- Transit traffic = traffic that passes through the AS
- Stub AS = has connection to only one AS, only carry local traffic
- Multihomed AS = has connection to >1 AS, but does not carry transit traffic
- Transit AS = has connection to >1 AS and carries transit traffic
Non-Transit vs. Transit AS

Traffic NEVER flows from ISP 1 through NET A to ISP 2 (At least not intentionally!)

Multihomed/Nontransit AS might be a corporate or campus network.

Selective Transit

NET A provides transit between NET B and NET C and between NET D and NET C

NET A DOES NOT provide transit Between NET D and NET B

Stub AS carries only local traffic

This example shows a routing policy at A.