Feb 11 Lecture Notes

Jiahe Tan, Sihan Ding
[jt7gu, sd3rf]@virginia.edu

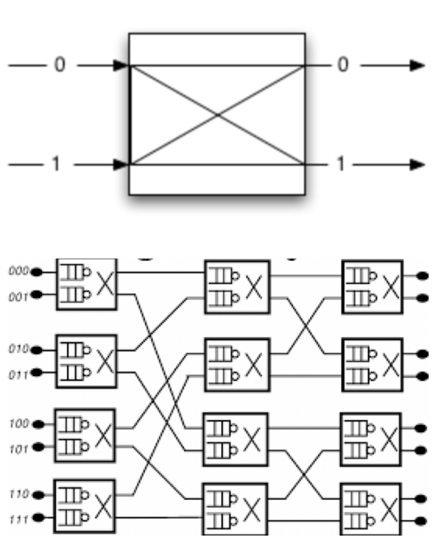

Figure 1. 2x2 Banyan Switch

1. Banyan Switch

A 2x2 Banyan switch has two inputs numbered 0 and 1 and two outputs numbered 0 and 1. The Banyan switch has a switching control built into the switch itself. If the switch reads the bit and it has value 0, it sends the packet to its upper output (which is marked with 0 in Figure 1). If the switch reads the bit and it has value 1, it sends the packet to its lower output (which is marked with 1 in Figure 1).By connecting these simple Banyan switching elements in series and parallel, it is possible to route packets in more complicated ways depending on the desired routes to establish. An example is shown in Figure 1.

2. Distance Vector Algorithm

$D_x(y)$ = Estimate of least cost from x to y .

$C(x, v)$ = Node x knows cost to each neighbor v.

$D_x = [D_x(y) : yN]$ =Node x maintains distance vector.

Node x also maintains its neighbors' distance vectors – For each neighbor v, x maintains $D_v = [D_v(y) : yN]$.

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 it receives a distance vector from a neighbor containing different information than before.

Bellman Ford: Each router maintains a Distance Vector table containing the distance between itself and ALL possible destination nodes. Distances,based on a chosen metric, are computed using information from the neighbors' distance vectors.

From time-to-time, each node sends its own distance vector estimate to neighbors.

When a node x receives new DV estimate from any neighbor v, it saves v's distance vector and it updates its own DV using B-F equation: $D_x(y) = \min C(x, v) + D_v(y), D_x(y)$ for each node $y \in N$

2.1. Link Cost Changes

Node detects local link cost change

Updates routing info, recalculates distance vector

if DV changes, notify neighbors

"Good news travels fast": If link costs decreases, the distance vectors of each nodes will be quickly updated.

"Bad news travels slow": If link costs increases, the distance vectors of each nodes will needs many iterations before algorithm stabilizes. See Figure 2.

The link cost between X and Y changed from 4 to 60. After this change, Y will update $D_y(x) = c(y, z) + D_z(x) = 2 + 3 = 5$. From the equation, we can see that this update used the outdated link cost value and the outdated $D_z(x) = 3$. This outdated value updating problem will continues for 44 iteration before the algorithm stabilizes. This is the infinity problem.

"Poisoned Reverse": If Z routes through Y to get to X : Z tells Y its (Z's) distance to X is infinite, so Y won't route to X via Z, and count-to-infinity problem does not occur.

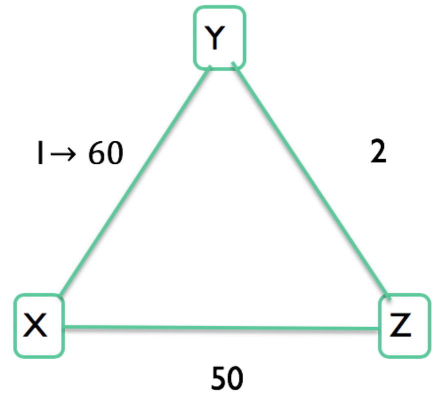

Figure 2. Link Costs Increase

However, can have problems when more than 2 routers are involved.

If Z tells Y that $D_z(x) = \infty$, then Y will not route to X through Z, and $D_y(x)$ will be updated to the new value 60. The outdated value problem will not occur any more.

3. Internet approach to scalable routing

Aggregate routers into regions known as "autonomous systems" (AS) (a.k.a. "domains")

3.1. Interconnected ASes

Forwarding table configured by both intra- and inter-AS routing algorithm. Intra-AS routing determine entries for destinations within AS. Inter-AS intra-AS determine entries for external destinations

3.2. intra-AS routing

Routing among hosts, routers in same AS ("network").

All routers in AS must run same intra-domain protocol.

Routers in different AS can run different intra-domain routing protocol.

Gateway router: at "edge" of its own AS, has link(s) to router(s) in other AS'es

Intra-AS Routing: OSPF (Open Shortest Path First)

The OSPF protocol is a link-state routing protocol, which means that the routers exchange topology information with their nearest neighbors. The topology information is flooded throughout the AS, so that every router within the AS has a complete picture of the topology of the AS. This picture is then used to calculate end-to-end paths through the AS, normally using a variant of the Dijkstra algorithm. Therefore, in a link-state routing protocol, the next hop address to which data is forwarded is determined by choosing the best end-to-end path to the eventual destination.

Security: all OSPF messages authenticated (to prevent malicious intrusion)

Multiple same-cost paths allowed (only one path in RIP)

Integrated uni- and multi-cast support: Multicast OSPF (MOSPF) uses same topology data base as OSPF

Hierarchical OSPF in large domains.

3.3. inter-AS routing

Inter-AS routing: BGP(Border Gateway Protocol)

BGP provides each AS a means to:

1. eBGP: obtain subnet reachability information from neighboring ASes

2. iBGP: propagate reachability information to all AS-internal routers.

3. Determine "good" routes to other networks based on reachability information and policy

BGP allows subnets to advertise its existence to rest of Internet: "I am here"

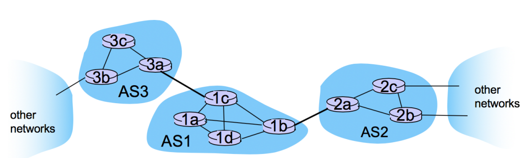

Figure 3. Autonomous Systems

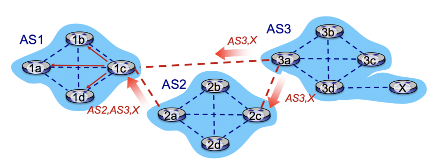

Figure 4. BGP Example