E6998-02: Internet Routing

Lecture 9
OSPF

John Ioannidis
AT&T Labs – Research

ji+ir@cs.columbia.edu

Announcements

Lectures 1-9 are available.

Homework 3 is out. Due 10/10 at 3am.

Submit only plain ASCII or PDF!

No class on Thursday 10/3 (inauguration).

Still looking for a TA.

Have you been thinking about your project?

OSPF

- More accurately: OSPFv2.
  - v1 was never really deployed.
- Link-state IGP, “open”, based on Dijkstra’s SPF algorithm.
- RFC2328 (and many others).
- Recommended IGP, esp. in a multivendor environment.
- Several features in common with other LS protocols
  - IS-IS, NLSP, PNNI.
  - We may look into IS-IS if time permits.
  - We’ll point out some things that other protocols do better.

- Basis for other IETF LS protocols:
  - MOSPF.
  - OSPF for IPv6.

OSPF Properties

- Reduced LSA distribution overhead.
  - ***Areas*** limit the extent of flooding.
  - Multicast limits impact on broadcast networks.
  - OSPF goes quiet when there are no route changes.
- 16-bit dimensionless metric.
- Equal-cost load balancing.
- Route aggregation.
  - CIDR, VLSM, etc.
- Route tagging.
- Authentication.

OSPF Overview

- Neighbor discovery:
  - ***Hello*** packets sent on all OSPF-enabled interfaces.
  - ***Neighbors***: routers on same link that agree on certain hello parameters.
  - ***Adjacencies***: virtual point-to-point links between certain neighbors.
- Link State Advertisements (***LSA***s):
  - Multiple LSA types.
  - Sent over all adjacencies.
  - List all of router's interfaces and the state of all links.
  - Flooded throughout an area.
  - Recorded in Link State Database and forwarded to neighbors.

OSPF Overview (cont'd)

- Designated Router / Backup Designated Router.
  - Two of the routers on a multiaccess link.
  - Used to reduce overall traffic on the link.
- When LSDB is complete:
  - Shortest Path Tree is computed on each router
    - (using Dijkstra's SPF algorithm).
  - Forwarding table built from SPT.
- Keep quiet:
  - Hellos are exchanged as keepalives.
  - LSAs are retransmitted every 30 minutes.

OSPF Network Types

- Point to point links.
  - High and low speed PPP links.
- Broadcast networks.
  - Ethernet-like.
- Non-Broadcast Multiple Access (NBMA) networks.
  - ATM, Frame Relay, X.25, (tunnels).
- Point-to-multipoint.
  - Really, special configuration of NBMA networks.
- Virtual links.
  - OSPF-specific meaning of the term.
  - Effectively, unnumbered point-to-point links.

Some Multicast Addresses

- `224.0.0.5` `AllSPFRouters OSPF-ALL.MCAST.NET`
- `224.0.0.6` `AllDRouters OSPF-DSIG.MCAST.NET`

- `FF02::5` and `FF02::6`, respectively for OSPFv3.

- While we are at it:
- `224.0.0.1` `ALL-SYSTEMS.MCAST.NET`
- `224.0.0.2` `ALL-ROUTERS.MCAST.NET`
- `224.0.0.9` `RIP2-ROUTERS.MCAST.NET`
- `224.0.0.10` `IGRP-ROUTERS.MCAST.NET`

- Look up some more (with `dig -x` *`address`*).

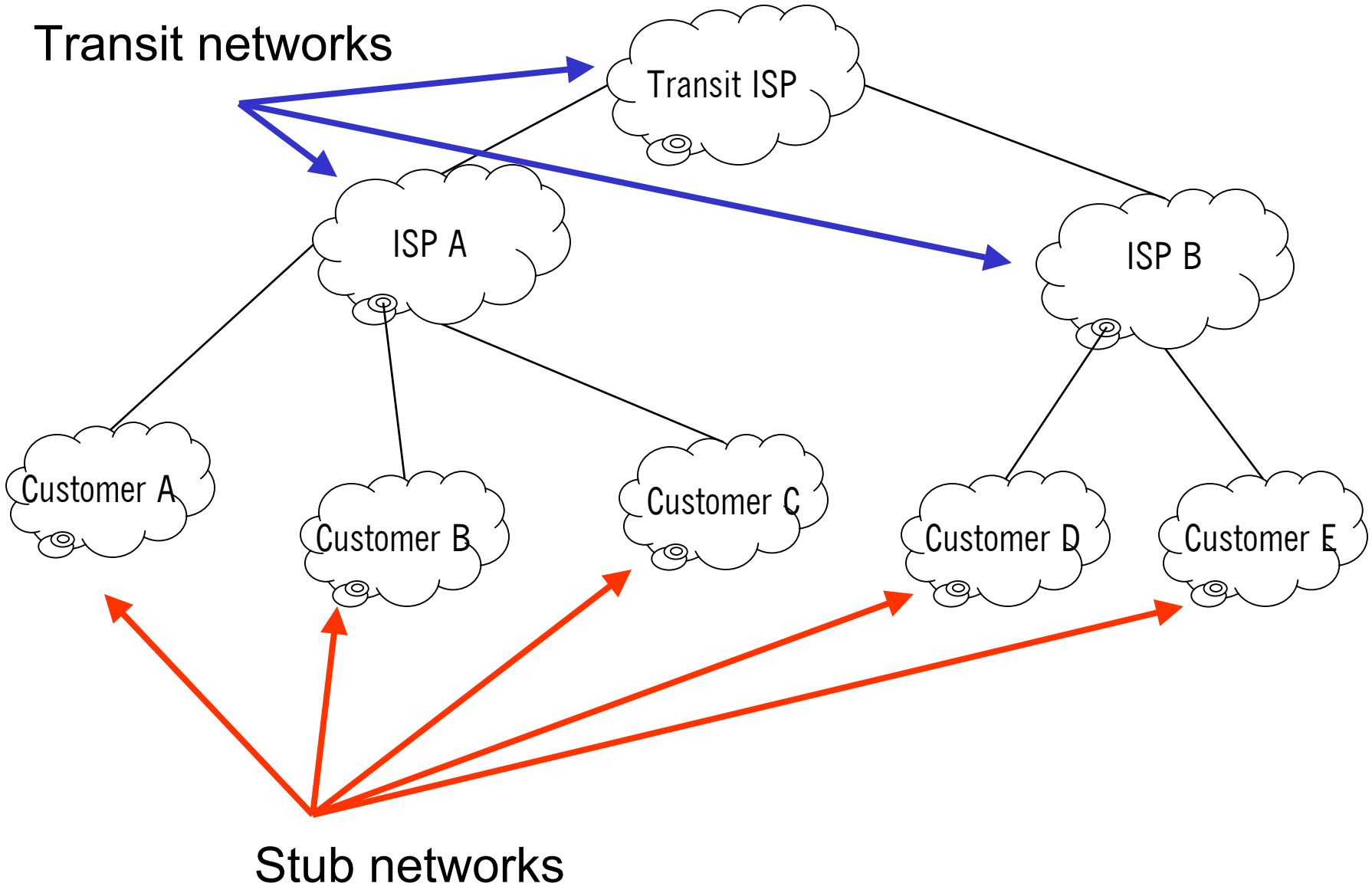
Destination Addresses Used

- On point-to-point networks:
  - No need to elect a DR.
  - Neighbors always become adjacent.
  - All OSPF packets except retransmitted LSAs sent to `AllSPFRouters` (224.0.0.5).
- On Broadcast networks:
  - DR and BDR are elected.
  - Packets sent to `AllSPFRouters` (224.0.0.5):
    - Hello packets.
    - All packets originating from the DR and BDR.
  - Packets sent to `AllDRouters` (224.0.0.6):
    - All packets sent by the rest of the routers.
    - Since these should only go to the DR/BDR.

Destination Addresses Used (cont'd)

- On NBMA networks:
  - DR and BDR are elected.
  - Extra configuration is needed to acquire neighbors.
  - All packets are unicast (no point in multicasting them).
- On Point-to-Multipoint networks:
  - These are treated as a collection of point-to-point links.
  - No DR/BDR are elected.
    - no need to.
  - Packets are multicast.
    - This way you don't have to find the address of the machine on the other side of the link.
- Virtual Links:
  - Packets are multicast.

Reminder: Transit *vs.* Stub Networks



Hello Protocol

- Sent every *HelloInterval* (default: 10s).
- Neighbor discovery.
- Parameter announcement/discovery.
  - No negotiation!
- Used as keepalive.
  - Dead after *RouterDeadInterval* (default: 4**HelloInterval*).
- Establishes bi-directional communication.
- On broadcast and NBMA networks:
  - Elects DRs and BDRs ([Backup] Designated Routers).

Hello Packet Contents

- ***Router ID*** of originating router (32 bits):
  - Highest IP address on loopback interfaces.
  - If no lb, highest IP address on regular interfaces.
  - Unchanged even if interfaces go down.

The rest of the fields pertain to the originating *interface*.

- ***Area ID*** (32 bits):
  - Area ID 0 is the ***backbone*** area.
- Checksum (16 bits).
- Authentication type (16 bits) and information (64 bits).
  - None, cleartext (bad!), or keyed hash.
  - The hash is appended to the packet and is not considered part of the packet for checksumming purposes.

Hello Packet Contents (cont'd)

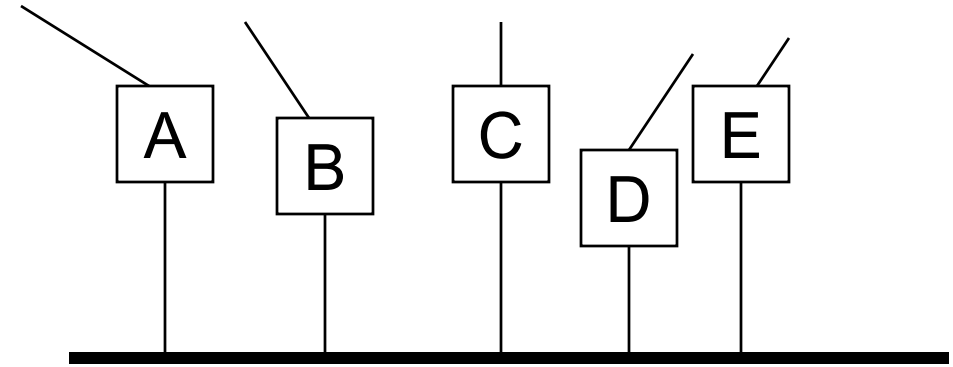
- ***HelloInterval*** (16 bits).
- ***RouterDeadInterval*** (32bits).
- Options (5 of 8 bits).
- Router Priority (8 bits).
- DR and BDR (32 bits each).
- List of neighbors.

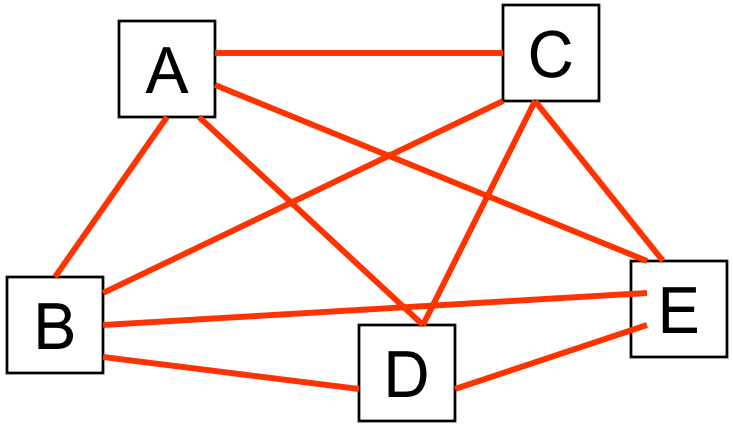
Hello Packet Processing

- Receiving routers (on same link) check:
  - AreaID, Authentication, Netmask, HelloInterval, RouterDeadInterval, and Options.
  - If they don't match its own, packet is dropped.
- If RouterID is known to the receiving interface:
  - RouterDeadInterval timer is reset.

  else
  - RouterID is added to the table of known neighbors.
- If receiving router sees its own ID in the list of of neighbors in the hello packet, it knows that it has bi-directional communication with the sender.
- Adjacencies may now be formed, if appropriate.
  - Depends on network type.

Adjacencies on Broadcast Networks





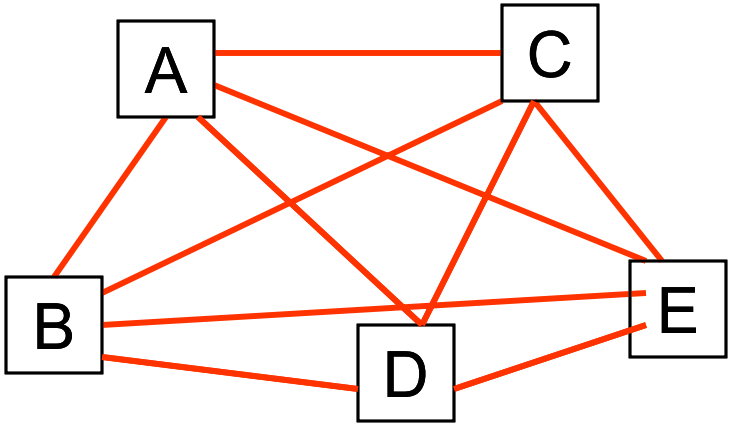
- If n routers are on a bc link, n(n-1)/2 adjacencies could be formed.
- n^2 LSAs would be originating from this network (why?).

Adjacencies, cont'd

- If routers formed pairwise adjacencies:
  - Each would originate (n-1)+1=n LSAs for the link.
  - Out of the network, n*n LSAs would be emanating.
- Routers would also send received LSAs to their adjacencies.
  - Multiple (n-1) copies of each LSA present on the network.
  - Even with multicast, (n-1) responses would still result.

- To prevent this, a Designated Router is elected.
  - Routers form adjacencies only with DR.
  - Link acts as a (multi-interface) virtual router as far as the rest of the area is concerned.

Adjacencies, cont'd



- One router is selected as the DR.
- Actually, another is selected as the BDR.
  - If the DR fails, we want the BDR to take over within RouterDeadInterval rather than go over a new election.
    - During which no traffic would be forwarded.
- Routers form adjacencies with both DR and BDR.
- DR and BDR also form adjacencies with each other.

DR Election

- When router joins in:
  - Listen to hellos; if DR and BDR advertised, accept it.
    - This is the case if all Hello packets agree on who the DR and BDR are.
    - Unlike IS-IS, status quo is not disturbed!
- If there is no elected BDR, router with highest priority becomes BDR.
- Ties are broken by highest RouterID.
  - RouterIDs are unique (IP address of lb if).
- If there is no DR, BDR is promoted to DR.
- New BDR is elected.

DR Election Details

- Routers who believe can be BDRs or DRs put their own IDs in their Hello packets.
- Once 2-way communication has been established, all routers know who the candidates are.
- They can now all pick a BDR.
  - Highest priority, then Router ID.
- And then a DR.
- If only one router claims he's the DR, he becomes the DR.
- First two routers to come up become the DR and BDR.

OSPF Interface Data Structure

- IP Address and Mask
- Area ID
- Router ID
- Network Type
- Cost
- Interface Transit Delay
- State
- Priority
- DR
- BDR
- Hello Interval
- Hello Timer
- Router Dead Interval
- Wait Timer
  - Before DR selection
- Rxmit Interval
  - Ack packets
- Neighbors
- Auth type
- Auth key

OSPF Interface State Machine