Misc.

Homework due 4/28 on class webpage this evening
Project due 4/10; project 3 available same day
Second exam next Thursday 4/14
Final will be Friday evening before final's week (5/6), time/room TBA

Today, finish multicast, IP v6

SUNY-BINGHAMTON – CS428 SPRING '05 LEC. #16

Reverse Path Broadcast (RPB)

- When a multicast packet comes in:
  - If it came from the NextHop for the source of the multicast, send it out on all links

- Better than a flood? Not really, but allows us to improve later with RPM

- Improvement: TRPB – don’t forward to leafs with no membership

- Improvement: Reverse Path Multicast (RPM): prune full subtrees, not just leaves
  - First packet uses Reverse Path Broadcast
  - Use NMR packets to tell upstream routers not to forward packets
  - “Upstream” defined by the first RPB packet (parent is your next hop to the source)

Link State Multicast

- Recall: link-state – exchange information with all routers about immediate neighbors

- Supporting Multicast –
  - Idea: information about what multicast groups hosts on your network belong to part of the state
  - The “expanded” state information is exchanged in LSAs
  - Routers figure out the shortest multicast tree depending on source and forward packets accordingly (a tree per source per group)

- Protocol is known as Multicast OSPF (MOSPF); RFC 1584

SUNY-BINGHAMTON – CS428 SPRING '05 LEC. #16
Core Based Trees (CBT)

- Idea, a single distribution tree per multicast group; also called “shared tree” approach
  
- A multicast message is picked up by a router
  - If it is part of the tree, it floods it along the tree
  - If not, it unicasts to the root of the tree (the core) which floods the tree
  
- Joining the tree
  - Node informs its router using IGMP of interest in joining the group
    - If the router is not already part of the tree, it unicasts a join request to the “core”
    - As each intermediate router receives the request, it sets up a transient join state
      - If router part of the CBT, it sends a Join ACK; otherwise
      - It forwards the join request on to the core

Protocol Independent Multicast (PIM)

- Multicast most expensive when the group is sparse

- PIM distinguishes between sparse and dense multicast

- In the sparse mode
  - Similar to CBT – a shared distribution tree
  - Hosts join/leave groups explicitly using join/prune protocol messages
  - Where to send this message?
    - Every group is assigned a rendezvous point (RP) using a distributed algorithm
      - The RP collects information about members in the group and builds a shared multicast tree
      - It may elect later (“if traffic becomes heavy”) to build a source specific tree that is more optimal

- RFC 2362
**PIM Operation**

1. Choice of RP important
   - Single point of failure
   - Can cause suboptimal operation
2. Problem: both receivers and senders need to know addresses of RPs for every group
   - Configured or can extend IGMP to provide this information
3. How do we make it real? Multicast across different AS’s
4. Multicast requires router state
   - Adds significant complexity and serious scaling issues

**Hierarchical Multicast**

- Use hierarchy to scale algorithms to large internetworks
- Also allow multi-cast protocol independence across AS's?
- IETF working group dmrp working on inter-domain multicast

**PIM Discussion**

- Treat subdomains (subnetworks) as links in routing algorithms
  - Problem: all intra-domain routers need to be in groups
- Solution: create a ‘wild’ group to which all inter-domain routers join and modify routing algorithm to send to wild card instead of next-hop router for multicast traffic
- Only needed for transit subdomains, leaves only need one inter-domain router

**MBone**

- Mbone (Multicast Backbone)
  - Large scale experiment in supporting multicasting in the internet
  - Collection of Islands supporting multicasting (overlay network)
  - Each island has a multicast router (e.g., a host running mrouted)
  - Routers connected via tunnels (IP-in-IP)
  - DVMRP has been the routing protocol/being replaced by PIM

- Sample applications: video conferencing (vic); shared whiteboard (wb)

- Highlights the difficulty of adding functionality to the internet
  - Qbone, 6bone

- Book on Mbone: http://www.savetz.com/mbone/
IPv6 – why

- Address depletion is the main driver; but
  - **Difficult to add features post-facto to IPv4** (e.g., Multicast)
  - Since we are taking the painful decision to upgrade IP, might as well fix everything that is wrong with it
  - Address routing table growth (approx. 100,000 entries in backbone)
  - Easier to configure/use
  - Simplify packet processing
  - Multicast
  - Security
  - Quality of Service
  - Real time traffic support
  - Flow identification
  - IP billing
  - ...
- A lot of what is “optional” in IPv4 is required in IPv6

IPv6 – complex header

IPv6: simple header; how to implement fragmentation, authentication, etc...?

Extension headers: headers are chained; the next header field holds the protocol id

Currently Defined Headers

- Hop-by-hop options
  - Options examined by every hop
  - Example: jumbogram (32-bit payload)

- Routing header
  - Similar to source routing

- Fragmentation header

- Authentication header

- Encapsulation security

- Destination options header

IPv6 Addresses

- Address space 128-bits:
  340282366920938463463374607431768211456
  globally addressed devices! 1.6million addresses per cubic mile of the solar system
  - billions of addresses per person; overkill?
  - routing?

- Address Notation
  - Separate 4 figures of hexadecimal using : (8 figures total)
    * Example:
      2001:0000:0000:0000:0000:0000:0000:0001
    - Preceding 0’s can be omitted:
      2001:0000:0000:0000:0000:0000:0000:0001
    - Continuous 0 piece can be replaced by :: (at most once):
      ff::55 to mean
      00ff:0000:0000:0000:0000:0000:0000:0055
    - IPv4 addresses shorthand ::128.226.123.1
    - IPv4 address for IPv6 capable host ::ffff:128.226.123.1
IPv6 Routing

- IPv4 backbone has big routing table size; headache for backbone operators
- IPv6 addressing specification restricts the number of routing table entries by using architecture-enforced “routing aggregation”
- Hierarchical routing
  - Geographically
  - Provider based (change of provider = change of address)
- How does NAT compare in terms of addressing? in terms of changing provider?

Other Routing Features

- Supports source routing
- Multicast
  - Scope identifier to limit how far a multicast propagates
- Anycast
  - Like Multicast, several nodes have the same address
  - Like Unicast, messages sent go to only one of them (which one?)
  - Why?

Aggregatable Global address
- Top level – 13-bits (3 bits to indicate aggregatable unicast addresses)
- Next level – 48-bits. Site level – 16-bits

Aggregatable Global address
- Only 8192 entries in the top most level
- CIDR from the beginning, so the divisions are not too important

Autoconfiguration

- Network configuration is difficult
  - How to obtain a globally unique address?
- Obtaining a locally unique address
  - Put your physical (ethernet) address in low order 48-bits
  - Attach yourself to the link-local address space (beginning with 11111110)

- Obtaining a globally unique address
  - At a minimum need to set IP number, subnet mask and nameserver
  - DHCP helps but requires the presence of a DHCP server
  - Would like true plug-and-play, “stateless” configuration
One of the basic requirements is to make the transition from IPv4 to IPv6 easy—why?

- RFC 1933
- Internet is IPv4
- Most nodes are IPv4
- IPv6 nodes use their equivalent IPv4 compatible addresses
- IPv6 enabled nodes run both IPv4 and IPv6 stacks
- Similar to Mbone, 6bone spans 50 countries

In the early stages (today)

- IPv6 nodes use their equivalent IPv4 compatible addresses
- IPv6 enabled nodes run both IPv4 and IPv6 stacks

Use tunneling when crossing IPv4 network

SUNY-Binghamton – CS428 Spring ’05 Lec. #16

Alternative – Getting through NAT

But is IPv6 on the way?

- Not clear
  - Military has definite timeline; all military machines support IPv6 as of past October, and will completely switch over to IPv6 by 2008
  - But it's much less sure in the commercial world (which ends up determining what happens anyway)
    * They are motivated by $$: risks and opportunities
    * Risk not high enough: NAT and solutions based on it can overcome the address shortage problem— if we can solve the peer to peer addressability problem
    * Opportunity not immediately available: Not clear they can make money out of the additional features in IPv6

Late stages – IPv6 most everywhere

- Some IPv4 nodes
- Not enough IPv4 addresses for all nodes
  - Must rely on translation
- IPv4 relegated to an option within the IPv6 stack
- Not there yet; it remains to be seen what will happen as we approach address depletion
Discussion: IP/Network Layer

- So far
  - Discussion of digital communication (basics; no details)
  - Directly connected networks – point-to-point and shared medium
  - Switched networks (bridges vs. routers)
    * Global addressability
    * Heterogeneity (fragmentation, address translation)
    * Forwarding model (datagram, VC, source routing...)
    * Routing
    * Making it scale
    * Multicast
    * Limitations/the future – IPv6

- IP is the workhorse of the Internet