

## More on IP and the Network Layer

Kai Shen

10/7/2009

CSC 257/457 - Fall 2009

1

## The Internet Protocol

- hierarchical addressing
- packet format
- hierarchical routing:
  - intra-AS (RIP, OSPF)
  - inter-AS (BGP)

10/7/2009

CSC 257/457 - Fall 2009

2

## IP Address Depletion (DHCP)

Address depletion  $\Rightarrow$  32-bit address space soon to be used up.

Observation: not every host is online at a given time.

**DHCP:** allow host to *dynamically* obtain its IP address from network server when it joins network

- can renew its lease on address in use
- allows reuse of addresses (only hold address while connected)

10/7/2009

CSC 257/457 - Fall 2009

3

## IP Address Depletion (NAT)

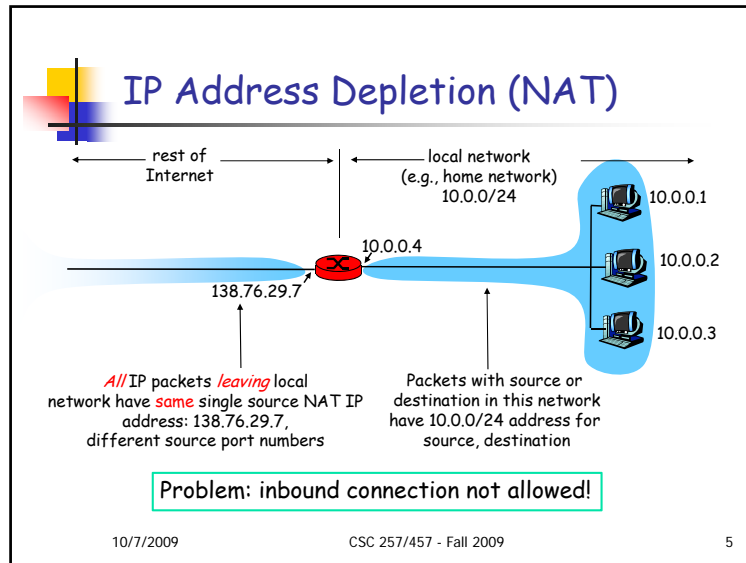
Observations:

- A lot of traffic is local
- Although IP addresses are few, possible (Ipaddr, port) tuples are more abundant and they can identify comm. end point

10/7/2009

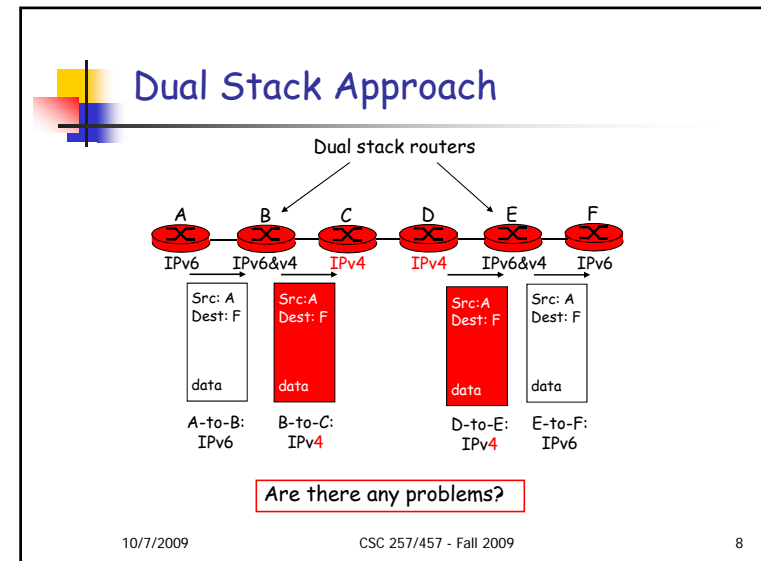
CSC 257/457 - Fall 2009

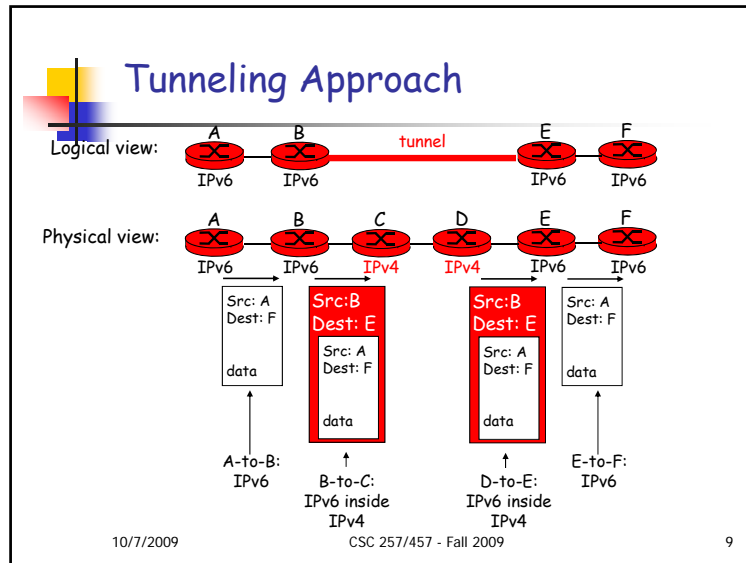
4



- ### IPv6
- **A key motivation:** address depletion
    - 128-bit addresses
  - **Additional motivation:**
    - simplification help speed processing/forwarding
      - fixed-length 40 byte header
      - no fragmentation allowed in intermediate routers
    - header changes to facilitate QoS
      - flow label
    - new "anycast" address: route to one of a set of several servers
- 10/7/2009 CSC 257/457 - Fall 2009 6

- ### Transition From IPv4 To IPv6
- It is difficult to change the network layer protocol
    - The large number of routers and administrative autonomy makes simultaneous change impossible.
    - Routers speaking different protocols do not naturally work together.
  - Two proposed approaches:
    - **Dual Stack:** some routers with dual stack (v6, v4) can "translate" between formats
    - **Tunneling:** IPv6 carried as payload in IPv4 packet among IPv4 routers
- 10/7/2009 CSC 257/457 - Fall 2009 7





### IPv6 Deployment

IPv6 deployment is slow:

- resistance to change ⇔ lack of critical mass
- v4/v6 co-existence requires backward compatibility (means additional overhead and complexity)

Address depletion is delayed by incremental techniques like DHCP and NAT.

10/7/2009 10  
CSC 257/457 - Fall 2009

### Multicast

- **Multicast:** act of sending packet to multiple receivers with single "transmit" operation
- Applications??

10/7/2009 11  
CSC 257/457 - Fall 2009

### Multicast via Multiple Unicasts

**Multicast via unicast**

- source sends N unicast packets, one addressed to each of N receivers
- no change on routers
- problem?

10/7/2009 12  
CSC 257/457 - Fall 2009

## Ideal Network Multicast

Router actively participate in multicast, making copies of packets as needed and forwarding towards multicast receivers

Multicast routers (red) duplicate and forward multicast packets

10/7/2009 CSC 257/457 - Fall 2009 13

## Internet Multicast Service Model

Multicast group concept: use of indirection

- hosts "join" multicast group
- hosts send IP packets with the multicast group address
- routers forward multicast packets to hosts that have "joined" that multicast group

10/7/2009 CSC 257/457 - Fall 2009 14

## Local and Wide-Area Multicast Management

- local: host informs local mcast router of desire to join group
  - IGMP (Internet Group Management Protocol)
- wide area: local router interacts with other routers to receive mcast packet flow

10/7/2009 CSC 257/457 - Fall 2009 15

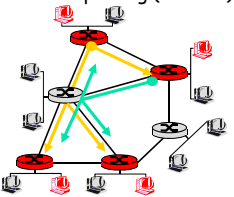
## IGMP: Internet Group Management Protocol

- host: sends IGMP report when application joins mcast group
- local router: sends IGMP query at regular intervals
  - host belonging to a mcast group must reply to query
- host need not explicitly "unjoin" group when leaving

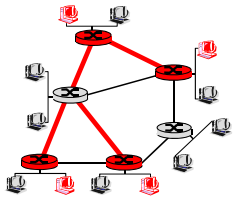
10/7/2009 CSC 257/457 - Fall 2009 16

## Wide-area Multicast Routing

- **Goal:** find path/paths (tree) connecting routers having local mcast group members
- **source-based tree:** one tree per source
  - shortest path trees, reverse path forwarding
- **group-shared tree:** group uses one tree
  - minimal spanning (Steiner) , center-based trees



Source-based trees

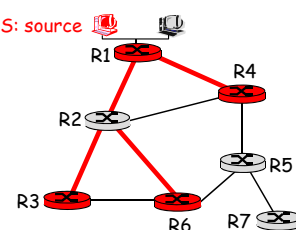


Shared tree

10/7/2009
CSC 257/457 - Fall 2009
17

## Shortest Path Tree

- mcast forwarding tree: tree of shortest path routes from source to all receivers
  - Dijkstra's algorithm



S: source

10/7/2009
CSC 257/457 - Fall 2009
18

## Reverse Path Forwarding

What if we don't have global network information?

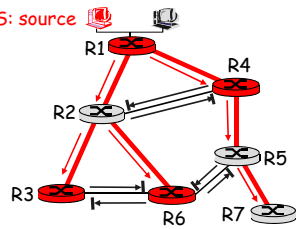
Goal is to forward along shortest paths. **Heuristic:**

- not to forward when packets clearly have deviated from the shortest paths.
- rely on router's knowledge of unicast routing table

**if** (mcast packet received on incoming link on shortest path back to source)  
**then** flood packet onto all outgoing links  
**else** ignore packet





10/7/2009
CSC 257/457 - Fall 2009
19

## Reverse Path Forwarding: Example



S: source

**LEGEND**

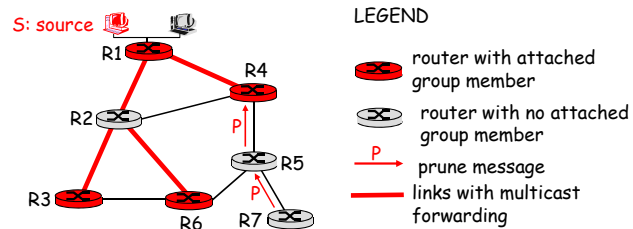
-  router with attached group member
-  router with no attached group member
-  packet will be forwarded
-  packet will not be forwarded

- result is a source-specific *reverse* SPT
- problem: reaching routers that are not in the group

10/7/2009
CSC 257/457 - Fall 2009
20

## Reverse Path Forwarding: Pruning

- forwarding tree contains subtrees with no mcast group members
  - "prune" msgs sent upstream by router with no downstream group members



10/7/2009

CSC 257/457 - Fall 2009

21

## Shared-Tree Multicast

- Single delivery tree shared by all
- Assume complete global network information is available, derive minimum cost tree connecting all relevant routers
  - problem is NP-complete
  - heuristic approximation exists

10/7/2009

CSC 257/457 - Fall 2009

22

## A Distributed Solution: Center-based Tree

- One router identified as "*center*" of tree
- To join:
  - edge router sends unicast *join-msg* addressed to center router
  - join-msg* "processed" by intermediate routers and forwarded towards center
  - join-msg* either hits existing tree branch for this center, or arrives at center
  - path taken by *join-msg* becomes new branch of tree for this router

10/7/2009

CSC 257/457 - Fall 2009

23

## Comparison


Compare reverse path forwarding with center-based tree

- Better multicast delivery along shortest paths produced by reverse path forwarding
- Too much pruning in a "sparse" system with relatively few interested routers

10/7/2009

CSC 257/457 - Fall 2009


24



## Internet Multicasting Routing

- **DVMRP**
  - distance vector multicast routing protocol, based on reverse path forwarding and pruning
- **PIM**
  - in a **sparse** mode where DVMRP doesn't do well, it uses Center-based Trees
- IP multicast deployment?

10/7/2009 CSC 257/457 - Fall 2009 25



## Disclaimer

- Parts of the lecture slides contain original work of James Kurose, Larry Peterson, and Keith Ross. The slides are intended for the sole purpose of instruction of computer networks at the University of Rochester. All copyrighted materials belong to their original owner(s).

10/7/2009 CSC 257/457 - Fall 2009 26