MPOA
Multiprotocol Over ATM

Raj Jain

Professor of Computer and Information Science
The Ohio State University
Columbus, OH 43210
Jain@CIS.Ohio-State.Edu

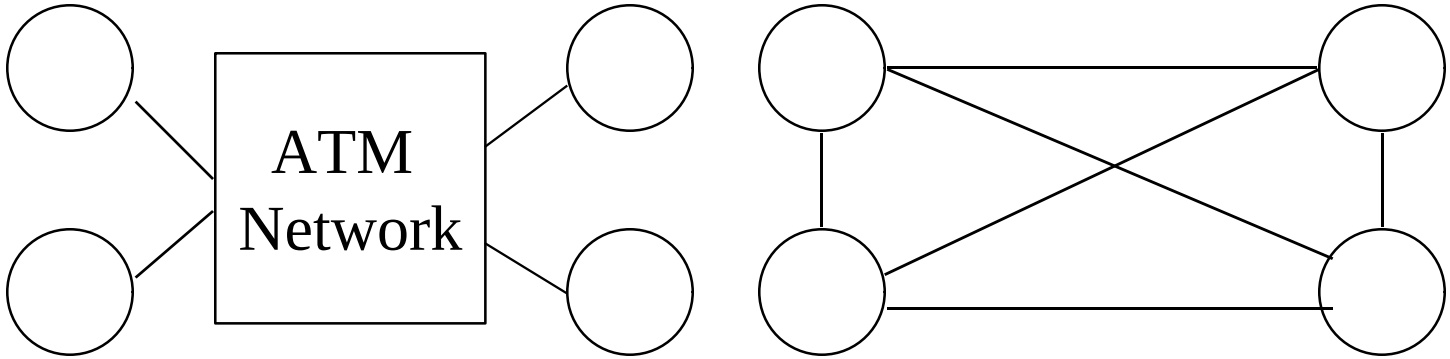
http://www.cis.ohio-state.edu/~jain/

- RFC 1483
- Classical IP
- Next Hop Routing Protocol NHRP
- Multicast Address Servers MARS
- MPOA
- IP Switching

Multiprotocol over ATM (MPOA)

- Problem: IP, IPX, DECnet routing, CLNP, AppleTalk DDP, IPv6, Vines, etc over ATM
  - Data transfer
  - Address resolution
  - Route determination
  - Multicasting
  - Multiple Server synchronization
- Solution: LAN emulation
  - IP over ATM
  - MPOA
- MPOA = LANE + IP over ATM + NHRP + MARS

IP Over ATM: RFC 1483



- Hosts set up VCs with other hosts/routers on ATM
- PVC ⇒ No need to know ATM addresses
- Treat each VC as a separate wire

Protocol Multiplexing

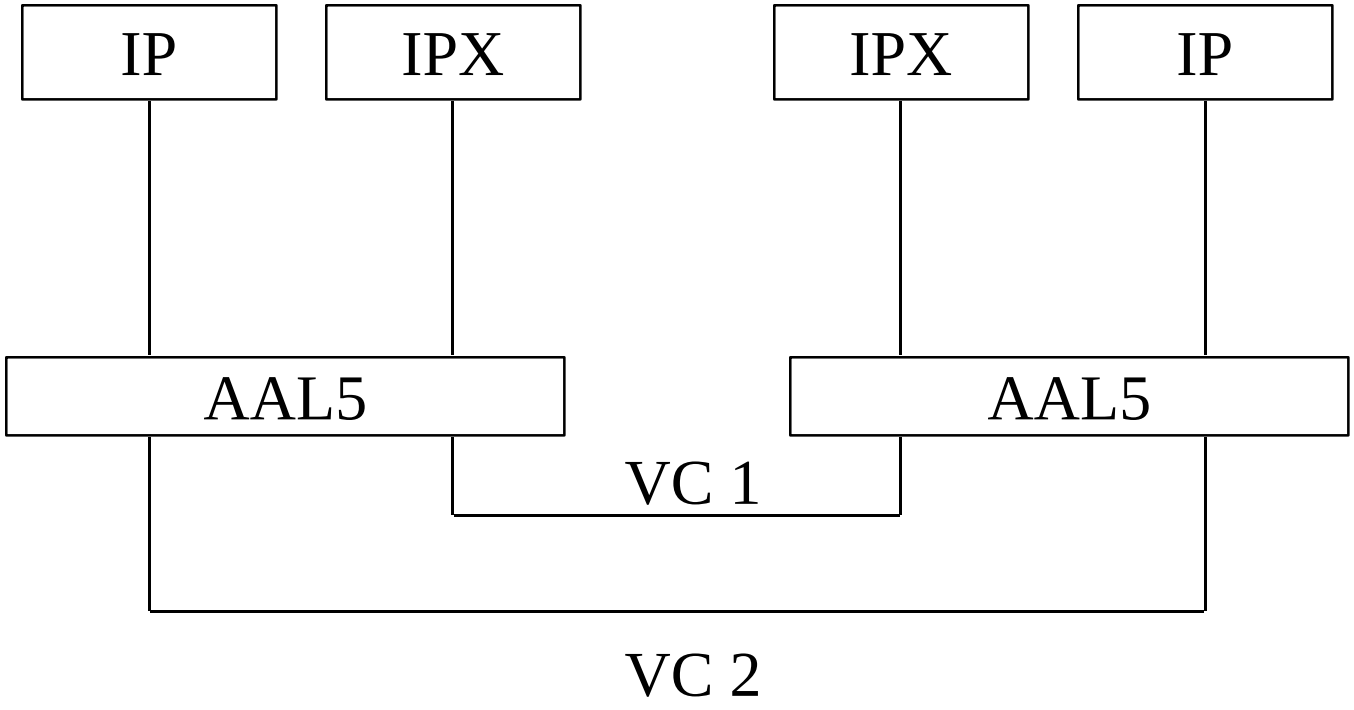
1. LLC Encaptulation:

- Multiple protocols share the VC using LLC SNAP/SAP:
- Subnetwork Access Protocol/Service Access Point
- Routed Protocols:
  - OUI = 00-00-00, EtherType = 08-00 ⇒ IP,
- Bridged Protocols:
  - OUI = 00-80-C2 ⇒ 802.1
  - Protocol ID = 0001 ⇒ 802.3/Ethernet w FCS
    0007 ⇒ 802.3/Ethernet w/o FCS

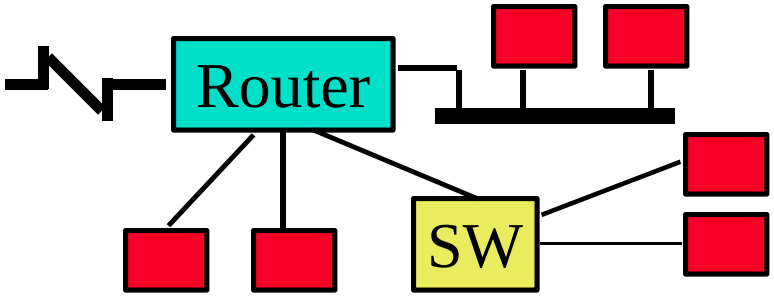
DSAP	SSAP	Control	OUI	Prot. ID	Payload
AA	AA	03	00-00-00	08-00	

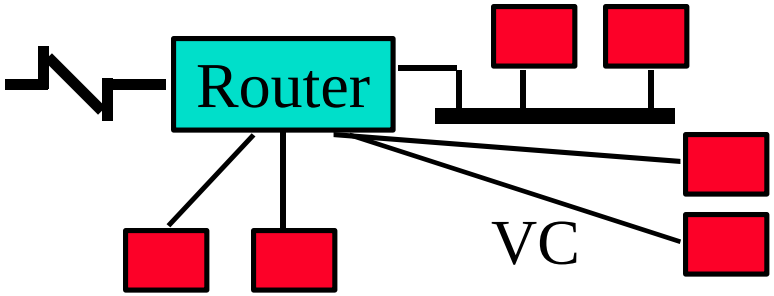
2. VC Multiplexing:

- Set up individual VCs for each protocol



Classical IP Over ATM

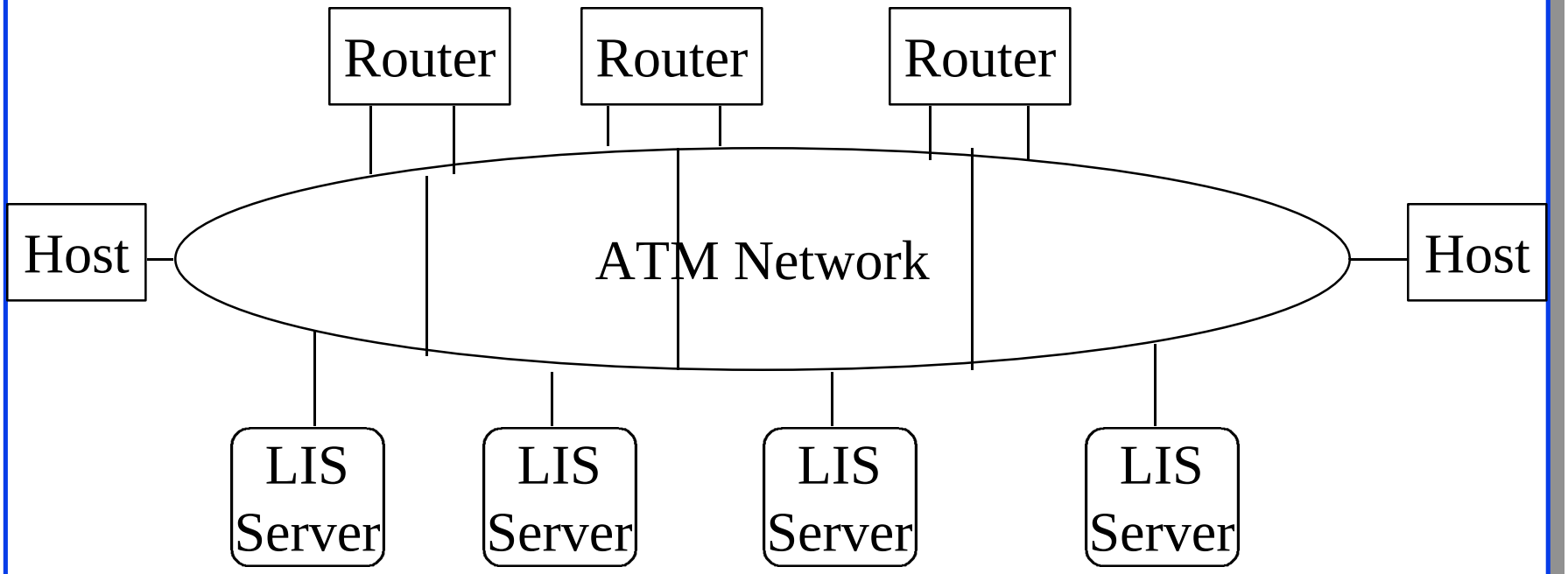




- ATM similar to point-to-point WANs.
  Simpler than LAN emulation
- IP address: 123.145.134.65
  ATM address:…1-614-999-2345-…
- Issue: IP Address ⇔ ATM Address translation
  - Address Resolution Protocol (ARP)
  - Inverse ATM ARP: VC ⇒ IP Address
- Solution: Logical IP Subnet (LIS) Server
- Ref: RFC 1577

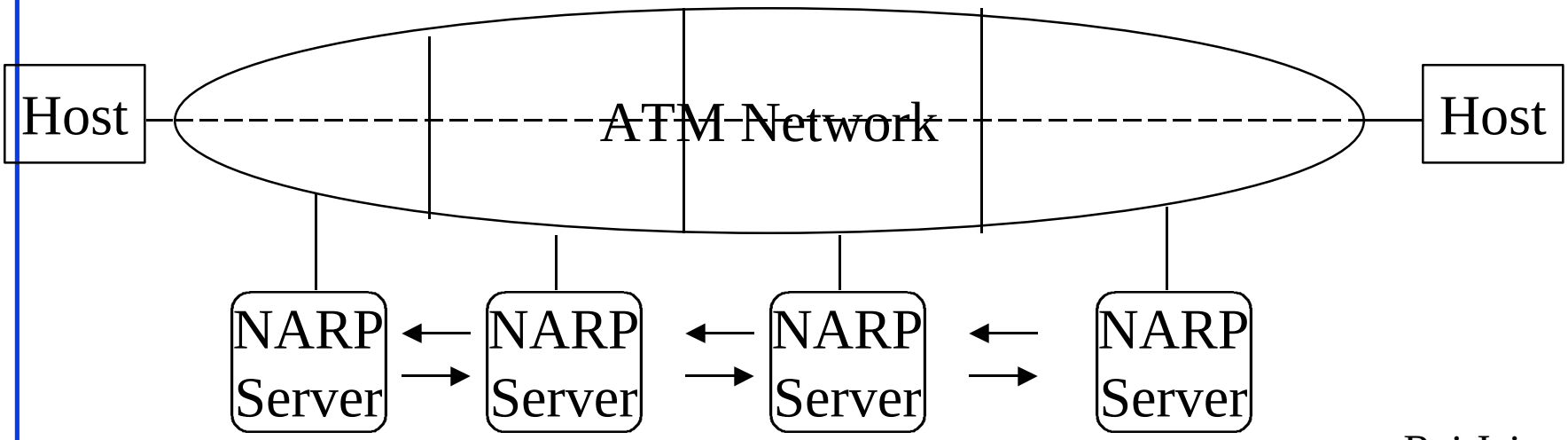
ARP over ATM

- Address Resolution Protocol
- Logical IP Subnet (LIS) server responds to ARP



NBMA ARP (NARP)

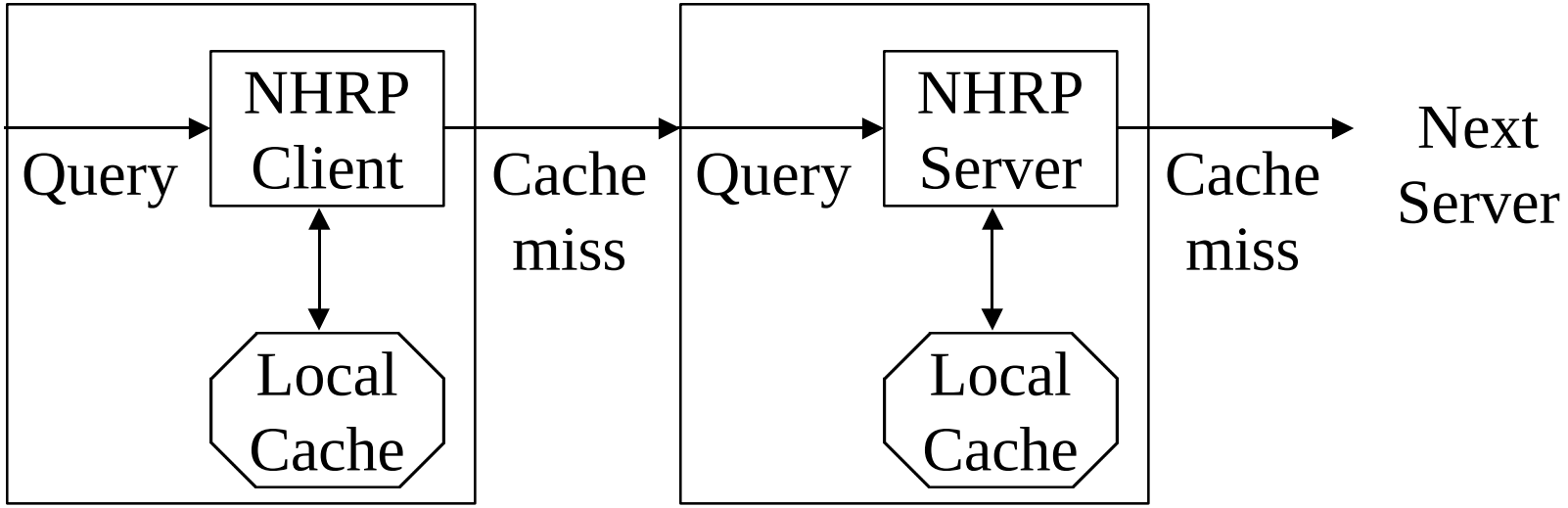
- NBMA = Non-broadcast Multiple Access
- Examples: X.25, Frame-relay, SMDS, ATM
- Problem with RFC 1577 Approach: Data needs to go through routers even if on the same NBMA
- Like going to the airport just to go to next block
- Solution: ARP servers of different LIS's communicate



- NARP servers cache the results
- NARP replies can be non-authoritative or authoritative
- NARP requests can be non-authoritative or authoritative
- Authoritative requests generally issued after failures.
- NARP works only if destination is on the same NBMA.
- NHRP, a superset of NARP, works even otherwise.

NHRP

- Next Hop Routing Protocol
- Developed by Routing over Large Clouds (ROLC) group
- Provides the next hop towards the destination.
- Next Hop Servers (NHSs)



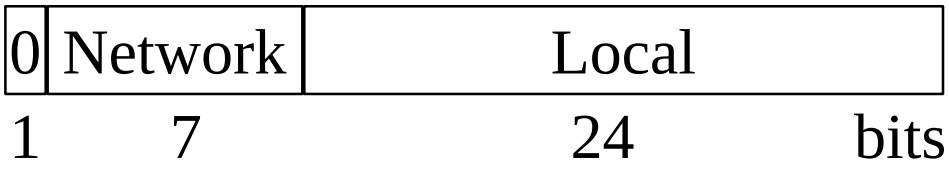
- Hosts are configured with the address of local NHS
- NHRP servers cache the results
- NHRP replies can be non-authoritative or authoritative
- NHRP requests can be non-authoritative or authoritative
- Authoritative requests generally issued after failures.
- While waiting for NHRP shortcut, data may be forwarded along the routed path.
- NHS learns about hosts via manual configuration or registration

NHRP Features

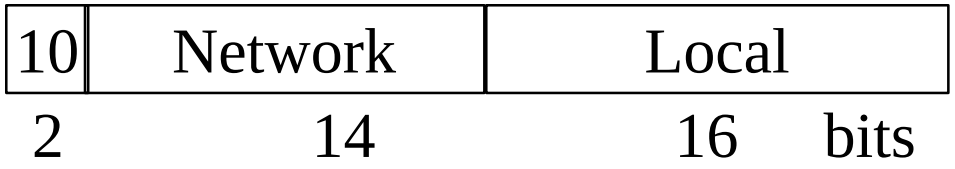
- Recording forward/reverse/both NHSs traversed
- Allows indicating QoS of path desired. Syntax TBD.
- Aggregated info: Range of addresses reachable via x using prefix
- Cached information has an associated holding (expiry) time
- Purge request can be sent after a reconfiguration
- Allows authentication information in NHRP packets
- NHRP Domino effect = Every router on the path of a packet sends an NHRP request and forwards the packet without waiting for the reply
  $\Rightarrow$ Only originating (not transit) routers should issue NHRP requests

IP Address

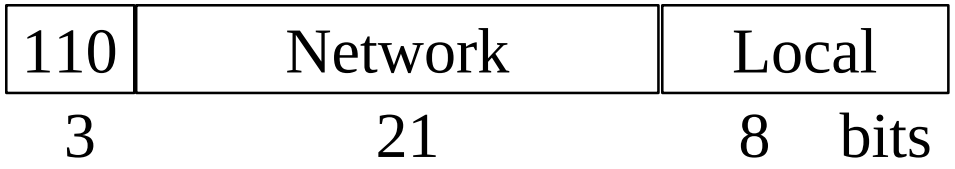
- Class A:



- Class B:

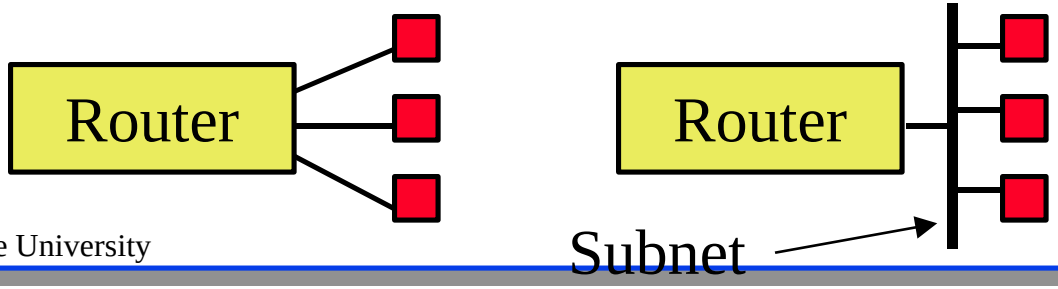


- Class C:



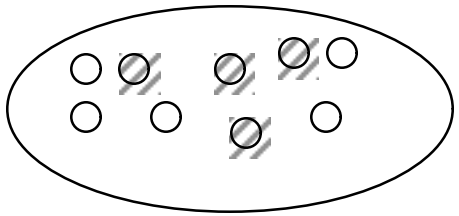
- Class D:

- Local = Subnet + Host (Variable length)



IP Multicast over ATM

- Need to resolve IP multicast address to ATM address list
- Multicast Address Resolution Servers (MARS)
- Each MARS serves a cluster (LIS) of IP hosts
- Each LIS contains only one cluster
- Old LIS members not using MARS are not in the cluster
- Internet Group Multicast Protocol (IGMP)
- Hosts are configured with MARS address
- Multicast group members send IGMP join/leave messages to MARS

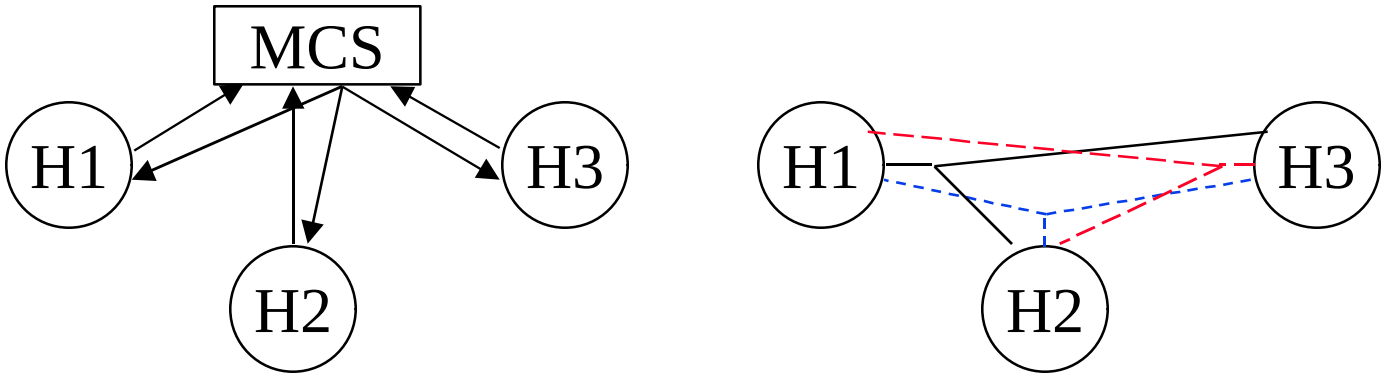


- LIS members
- Cluster members

- All cluster members join 224.0.0.1 (All local cluster members)
- Hosts wishing to send a multicast send a resolution request to MARS
- Senders do not need to be members of the multicast group
- IGMP queries to a multicast address can be used to get list of all members of the group
- Queries to 224.0.0.1 ⇒ "What are your multicast groups and who are their members?"
- Responses are sent to respective group addresses so that other members of the same group hear it and do not respond.

Multicast Data Flow

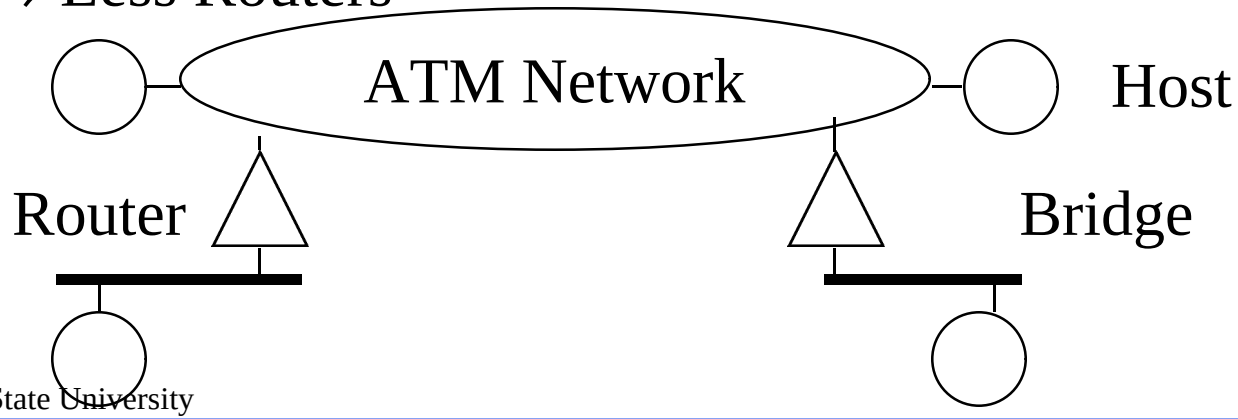
- Multicast Servers (MCS): Retransmit packets to multicast members on a single point-to-multipoint VC (or multiple unicast VCs)
- VC Mesh
- Both models allowed. Each group can use different method.
- MARS keeps a point-to-multipoint VC with all cluster members and a p-t-m VC with all MCSs



- MARS broadcasts all join/leave messages to members so that they can adjust their p-t-m VCs accordingly
- These messages have sequence numbers. Gap ⇒ Request update from MARS.
- Clusters are connected via IP multicast routers
- Endpoints (e.g., IP multicast routers) may join all (or ranges) of multicast groups
- Multiple backup MARSs allowed.
- If a group has both MCS and non-MCS members, MARS supplies MCS as group list if asked by members and member list if asked by MCS
  ⇒ Allows cluster members to send the packet only to MCS

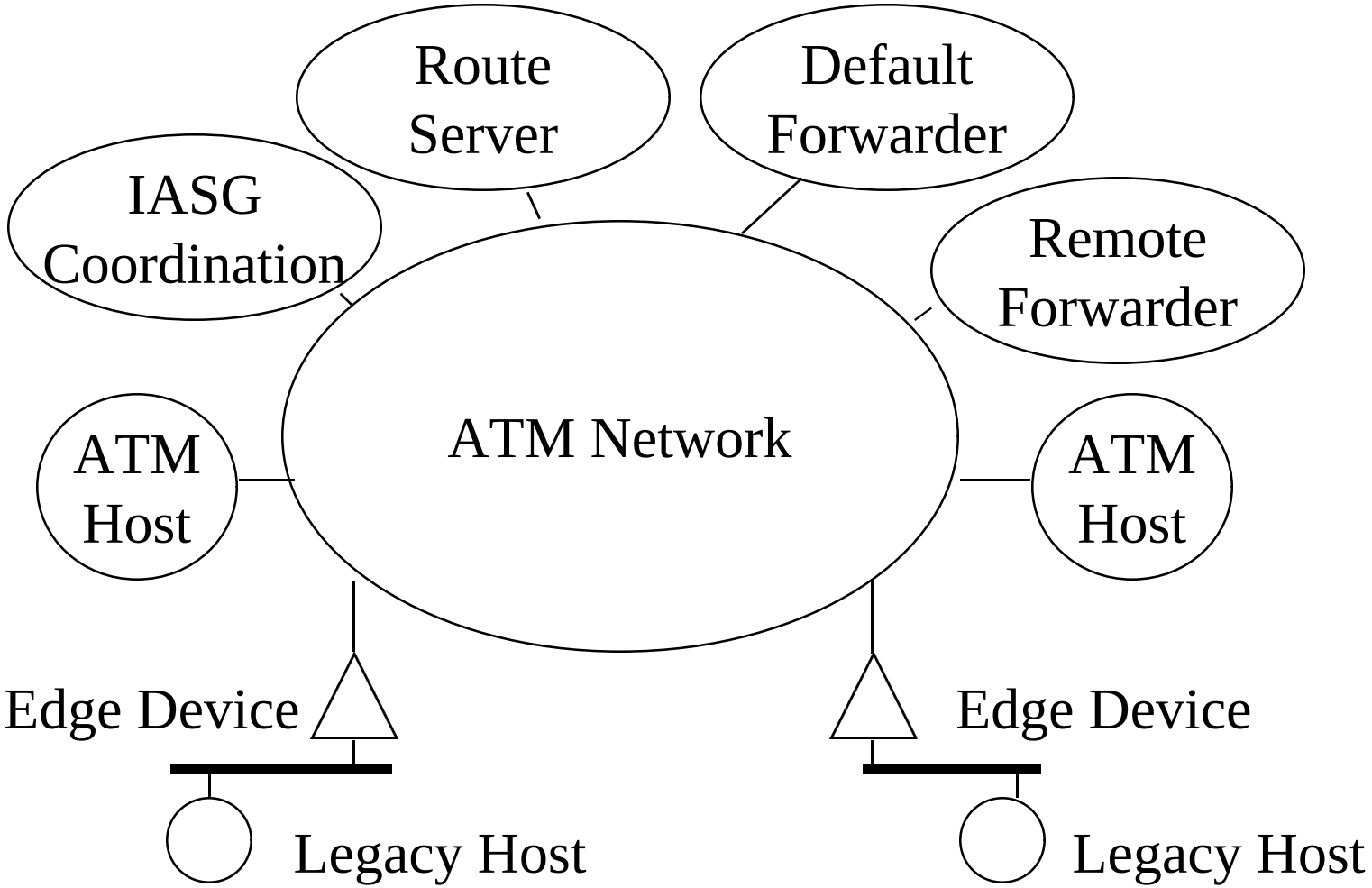
MPOA: Overview

- LANE hides QoS
- Both LANE and RFC 1577 need routers even in the same ATM network
- MPOA provides direct layer 3 connectivity across an ATM fabric
- For legacy hosts, MPOA provides direct layer 3 connectivity to/from ATM-to-legacy bridges/routers
  $\Rightarrow$ Less Routers



- Layer 3 protocol runs directly over ATM ⇒ Can use ATM QoS
- LANE operates at layer 2
- RFC 1577 operates at layer 3
- MPOA operates at both layer 2 and layer 3
  ⇒ MPOA can handle non-routable as well as routable protocols
- MPOA uses LANE for its layer 2 forwarding
- MPOA = Unified approach for all layer 3 protocols over ATM
- MPOA servers can be replicated. Clients are not aware of replication.

MPOA Reference Model

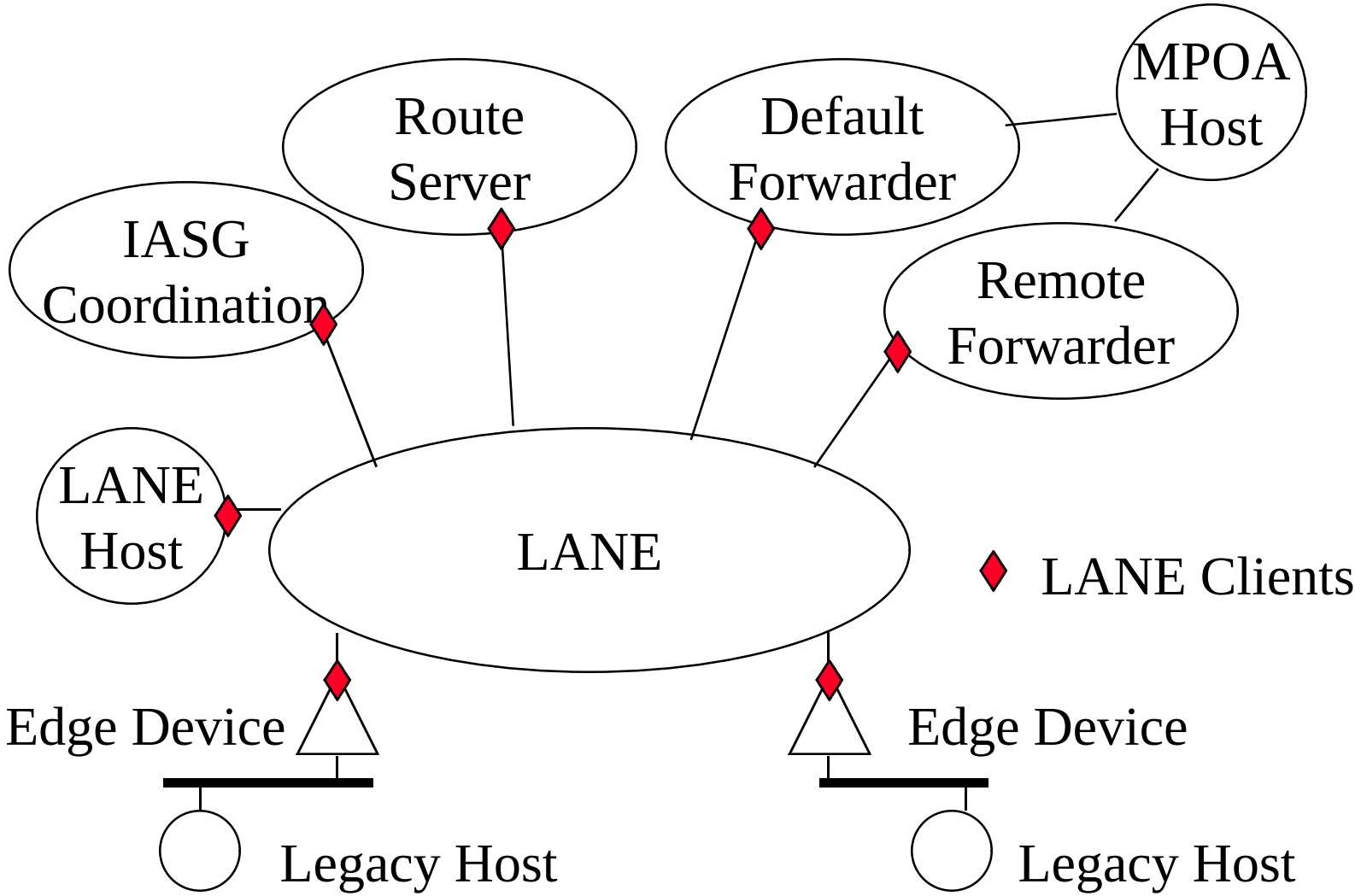


MPOA Components

- Internet Address Sub-Group (IASG)
  ≈ Subnet ≈ LIS ≈ Cluster
- IASG Coordination functions (ICFG)
  ≈ LIS Server: Host registration, etc.
- Internet Broadcast Sub-Group (IBSG) ≈ Cluster
- Route Servers = Run routing protocols and determine routes (does not forward packets)
- Edge Devices = ATM to legacy bridge/router
- Multicast Servers: Layer 3 multicast within an IASG
- Default Forwarders: For intra-IASG traffic in the absence of direct client-to-client connectivity. Also acts as multicast server.

- Remote Forwarder: Between IASG's
- MPOA Client = ATM-attached Host (AHFG) or Edge device (EDFG)
- MPOA Server = IASG CF (ICFG), Route Server (RSFG), or Intra-IASG broadcast/unknown server (IBUFG)

MPOA Architecture



ICFG

- IASG Coordination Functional Group
- Coordinates IASG distributed over multiple legacy LANs and ATM stations
- For example, a single IPv4 subnet across multiple Ethernets
- Hosts find IASG CF's address from configuration server
- Hosts register protocols and addresses with ICFG
- LANE may be used to reach hosts on the same IASG
- IASG is layer-3 protocol specific
- If a host operates two layer-3 protocols, it is a member of two IASG's

Route Servers

- Legacy router = Route server + packet forwarder
- Route server functions exchanges connectivity information and determines route
- Packet forwarding function forwards packets in the right direction
- In an ATM-to-ATM router, packet reassembly and forwarding can be avoided.
- Only route server function is needed.
- Packet forwarding is needed only in edge devices

- Virtual Routers = Route servers and packet forwarders in different boxes
- Conventional Routers =Route servers and packet forwarders in one physical device

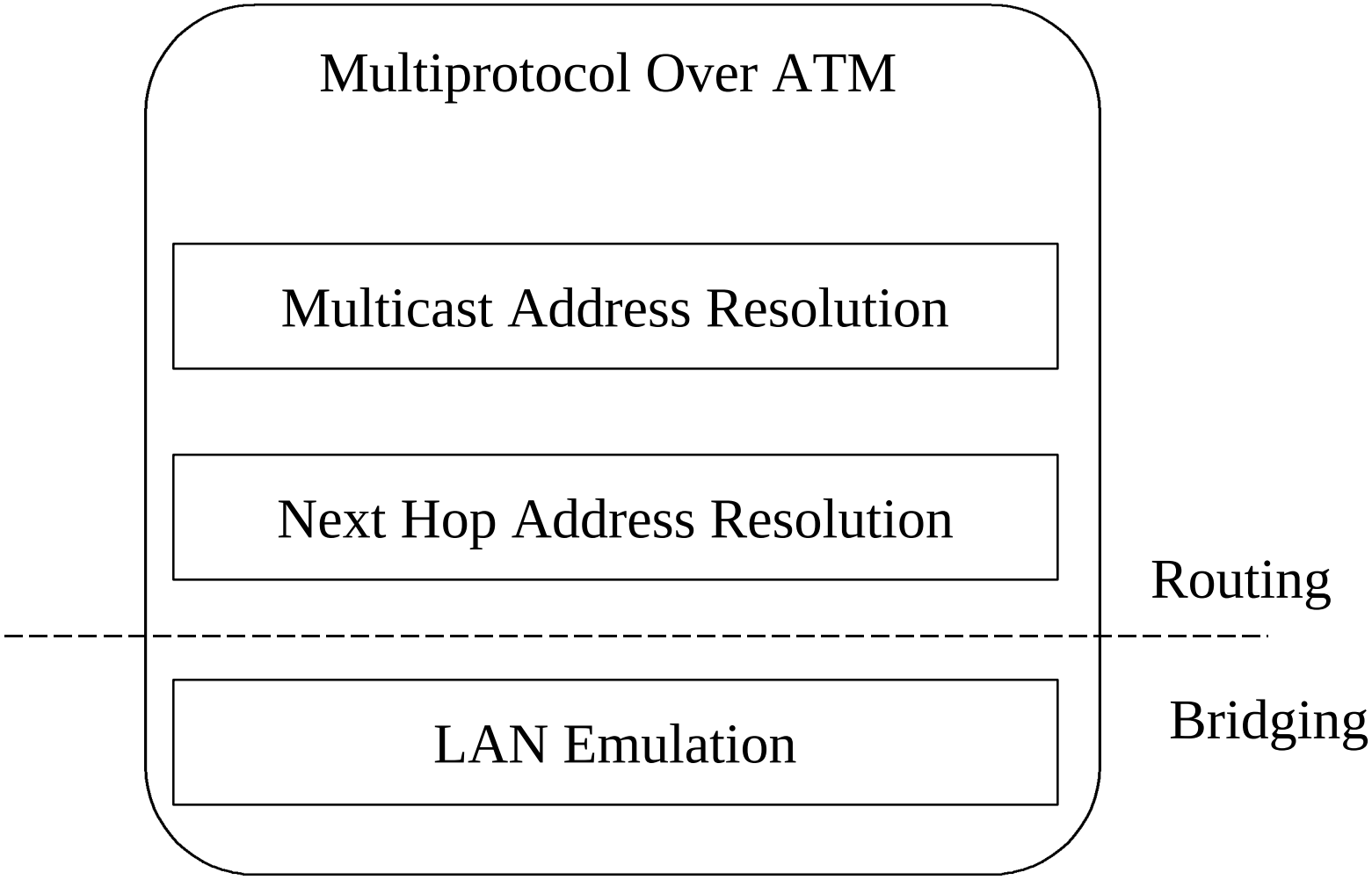
Address Resolution

- Hosts:
  - Host queries IASG CF about a layer 3 address
  - ICFG responds with the ATM address
  - ICFG may send the query to route server
  - Route servers use NHRP to determine the next-hop
- Edge Devices: If the destination is
  - A legacy host behind another edge device, it forwards the ARP to that device
  - An ATM host, it may answer ARP for the host.
  - Unknown, it broadcasts the ARP on ELAN.

Data Flow

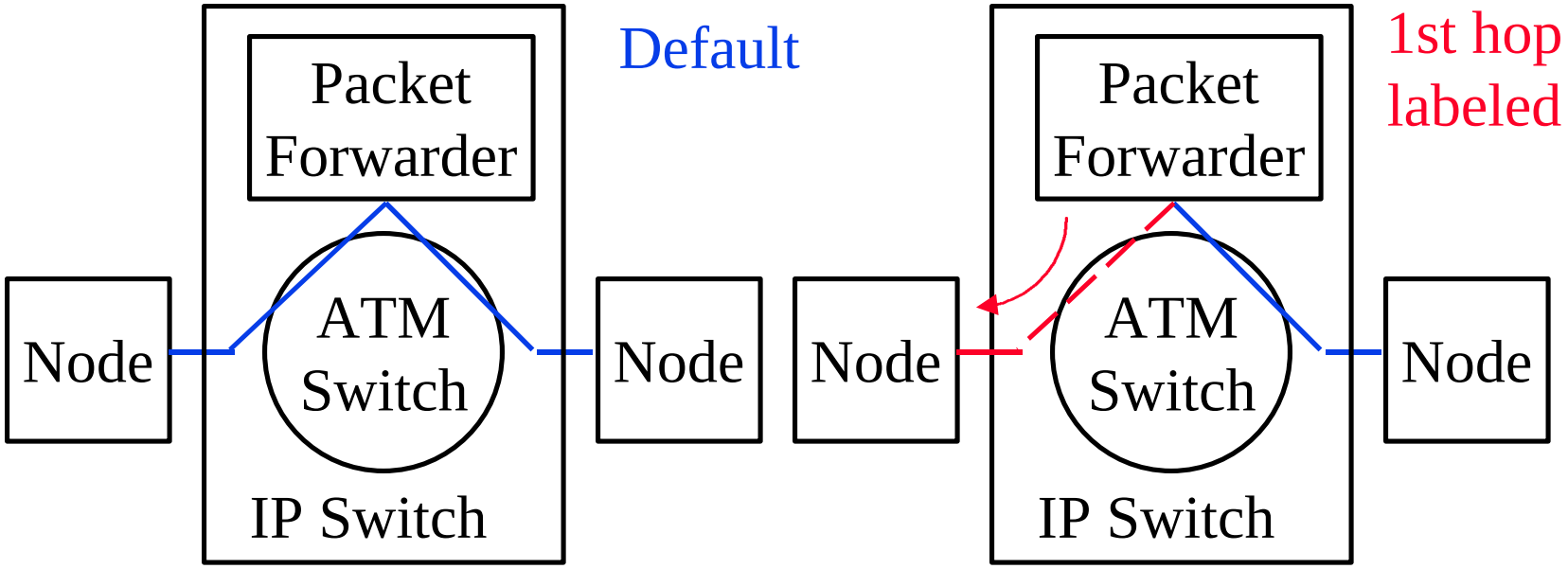
- Default flow:
  - Used when short cut does not exist.
  - AHFG sends to Remote forwarder or default forwarder
  - Edge devices use LANE for delivery based on MAC addresses
  - Dual-mode hosts use LANE for intra-IASG traffic and use AHFG for inter-IASG traffic
  - Default forwarder will send to remote forwarder if necessary
- Shortcut Flow:
  - Send directly on a VC to the destination or edge device

Protocol Stack

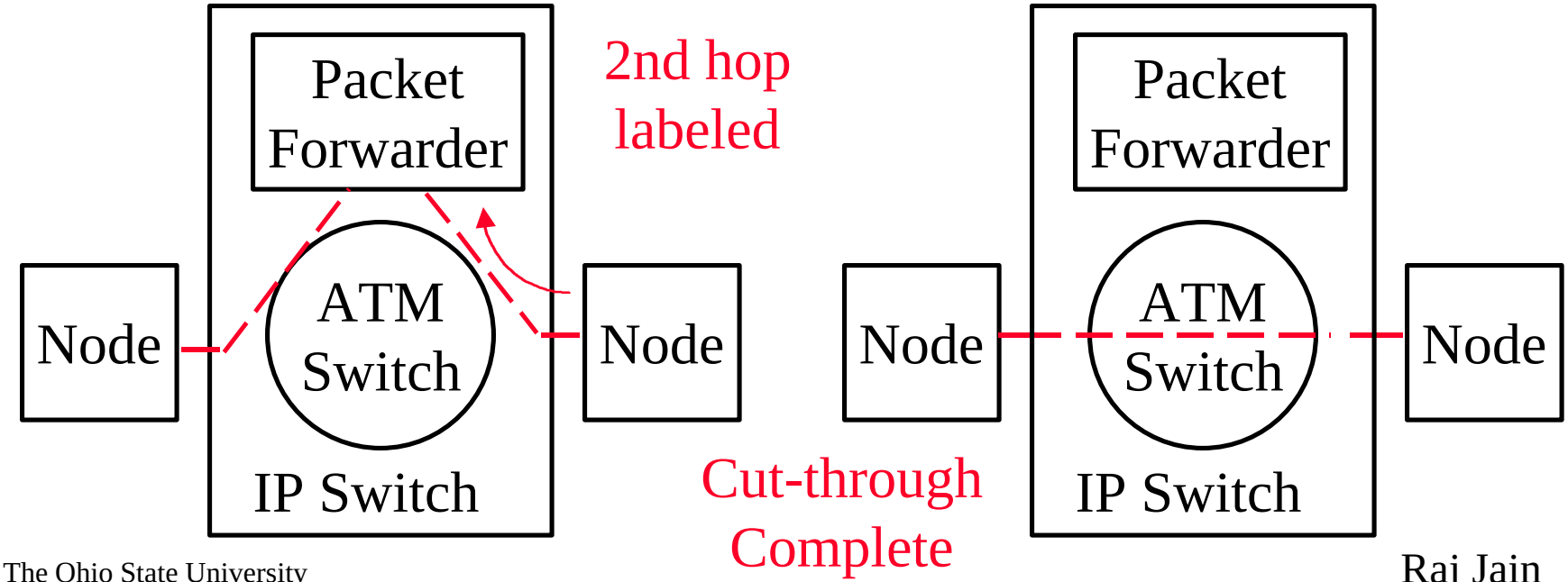


IP Switching

- Normally the packets are reassembled and forwarded in the router. Segmentation and reassembly in the forwarder.
- If a flow is deemed to be "flow oriented", previous node is told to set up a new VC. Forwarder uses cached info.



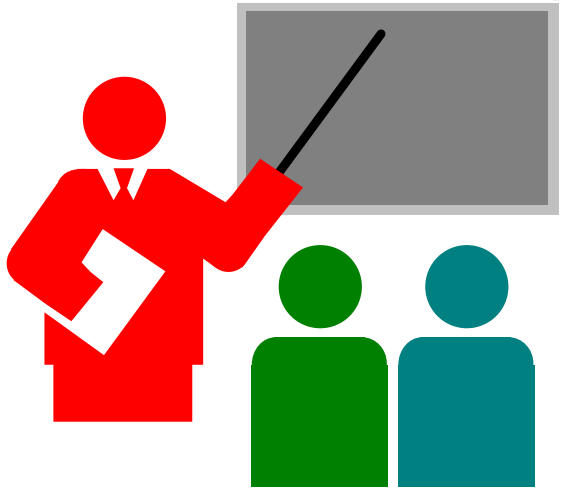
- Downstream nodes may also ask for a new VC. The switch then makes a mapping for cut-through
- Flow-oriented traffic: FTP, Telnet, HTTP, Multimedia
- Short-lived Traffic: DNS query, SMTP, NTP, SNMP, request-response
- Ipsilon claims that 80% of packets and 90% of bytes are flow-oriented.



IP Switch Protocol Stack

- Generic Switch Management Protocol (GSMP)
  - Controls the ATM Switch Engine
  - Set up, tear down VCs
- Ipsilon Flow Management Protocol (IFMP)
  - Enables communication between nodes/switches
  - Associates flows with VCs
  - Defines flow redirect messages and acks
  - Implemented in ATM end stations and edge routers
- Implemented as a software layer over an ATM switch
  - Ipsilon claims GSMP to be 2000 lines, IFMP 10000 lines of code

Summary



- Classical IP allows ARP using LIS servers
- NHRP allows shortcut between ATM hosts and edge devices
- MARS allows multicast
- MPOA combines LANE, NHRP, and MARS
- IP switching selectively changes some flows to switched path

References: IP over ATM

- RFC 1483, "Multiprotocol Encaptulation over ATM Adaptation Layer 5," July 1993.
- RFC 1735, "NBMA Address Resolution Protocol (NARP)", December 1994
- D. Katz, "NBMA Next Hop Resolution Protocol (NHRP)," Internet Draft, December 1995.
- G. Armitage, "Support for Multicast Over UNI 3.0/3.1 based ATM Networks," Internet Draft, February 1996.
- RFC 1755, "ATM Signaling support for IP over ATM," February 1995
- RFC 1932, "IP Over ATM: A Framework Document," April 1996

- RFC 1926, "An Experimental Encapsulation of IP Datagrams on Top Of ATM," April 1996
- RFC 1821, "Integration of Real-time Services in an IP-ATM Network Architecture," August 1995
- RFC 1680, "IPng Support for ATM Services," August 1994.
- RFC 1626, "Default IP MTU for use over ATM AAL5," May 1994.

References: MPOA

- G. Armitage, "Multicast and Multiprotocol Support for ATM Based Internets," Computer Communications Review, April 1995(?)
- C. Brown, "Baseline text for MPOA," ATMF/95-0824R6, February 26, 1996.

References: IP Switching

- Ipsilon, "IP Switching: The intelligence of Routing, the Performance of Switching," February 1996.
- P. Newman, et al, "Flow Labeled IP: Connectionless ATM under IP," Networld+Interop, April 1996.
- RFC 1953, "Ipsilon Flow Management Protocol Specification for IPv4, Version 1.0," May 1996.
- RFC 1954, "Transmission of Flow Labeled IPv4 on ATM Data Links, Ipsilon Version 1.0," May 1996.

Quiz

T F Please check True/False

1. ❑ ❑ MPOA reduces the number of routers required in a network.
2. ❑ ❑ MARS is a satellite-based long distance communication method
3. ❑ ❑ NHRP allows a shortcut path between two ATM devices on different subnets.
4. ❑ ❑ VC multiplexing means several applications sharing a VC.
5. ❑ ❑ NHRP gives authoritative answers only if requested so.
6. ❑ ❑ VC mesh among n nodes requires n^2 VCs.
7. ❑ ❑ Edge devices connect ATM network to legacy networks.
8. ❑ ❑ IP switch switches all IP packets.
9. ❑ ❑ With MPOA, you do not LANE.
10. ❑ ❑ MPOA supports IPX.