

MPOA

Multiprotocol Over ATM

Raj Jain

Professor of Computer and Information Science

Raj Jain is now at
Washington University in Saint Louis
Jain@cse.wustl.edu
<http://www.cse.wustl.edu/~jain/>

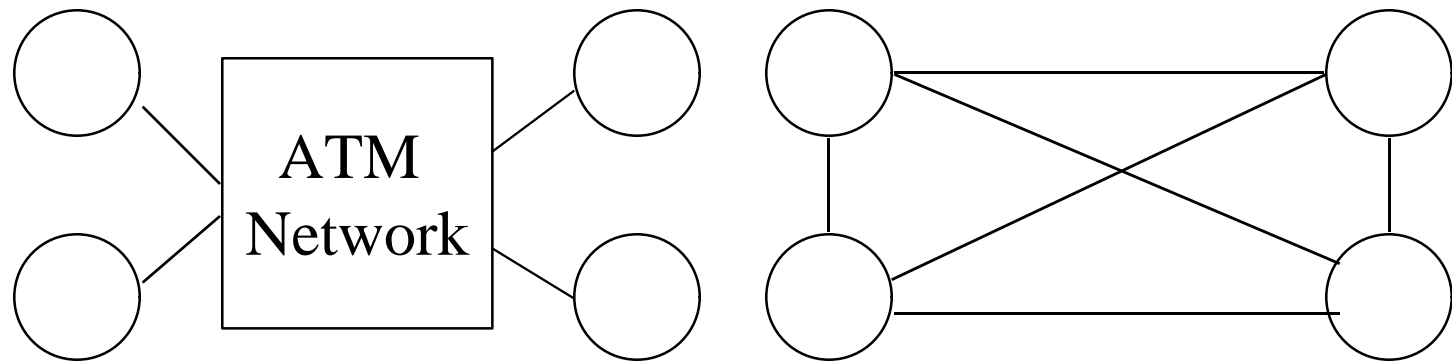


- q RFC 1483
- q Classical IP
- q Next Hop Routing Protocol NHRP
- q Multicast Address Servers MARS
- q MPOA
- q IP Switching

Multiprotocol over ATM (MPOA)

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

IP Over ATM: RFC 1483



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

Protocol Multiplexing

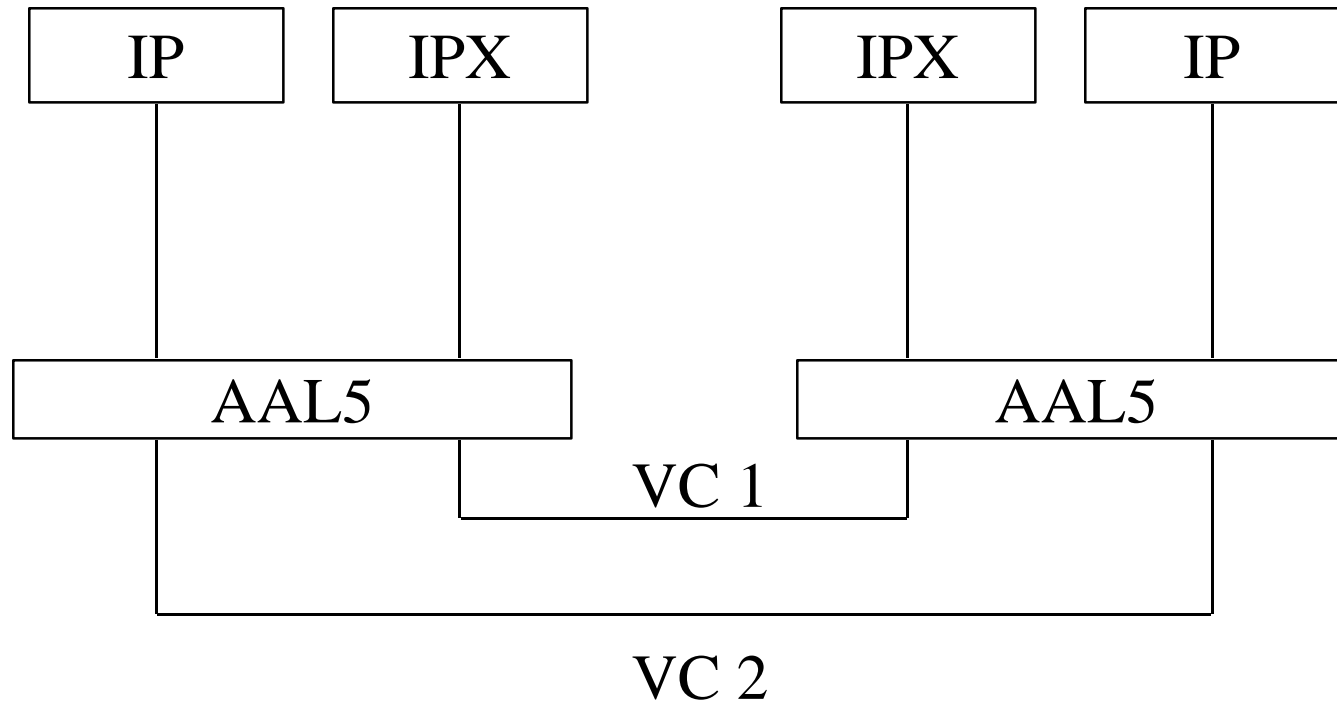
1. LLC Encapsulation:

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

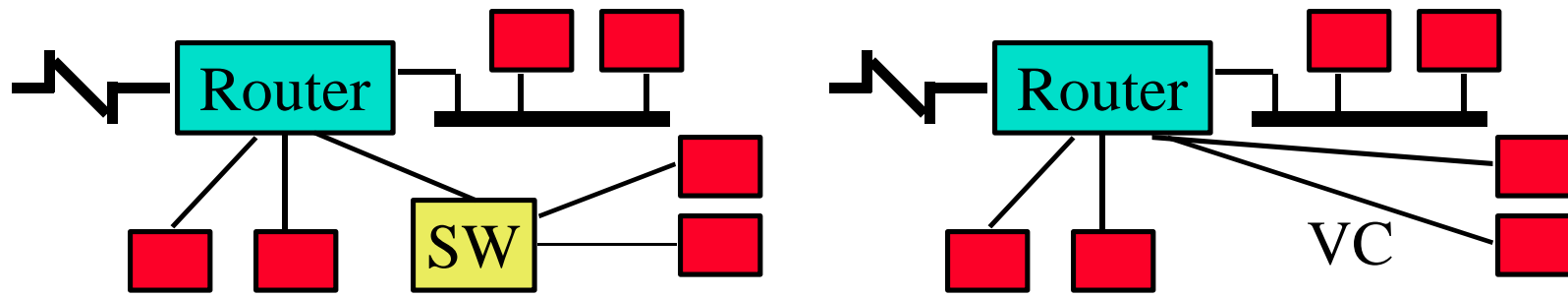
DSAP	SSAP	Control	OUI	Prot. ID	Payload
AA	AA	03	00-00-00	08-00	

2. VC Multiplexing:

- q Set up individual VCs for each protocol



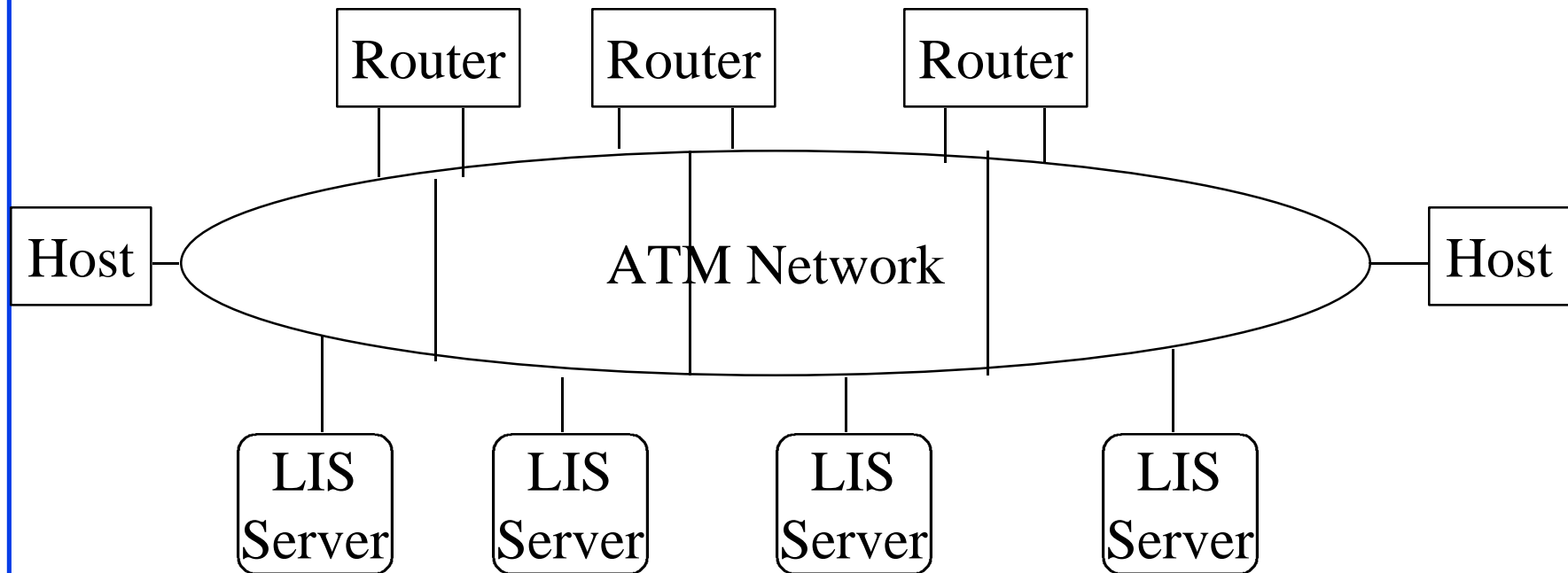
Classical IP Over ATM



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

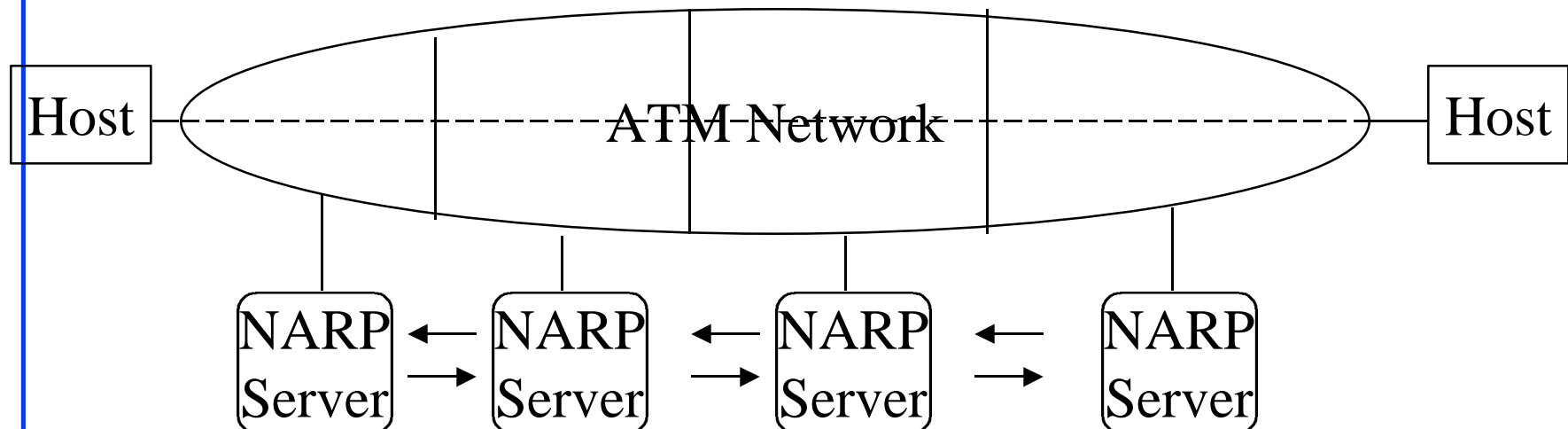
ARP over ATM

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



NBMA ARP (NARP)

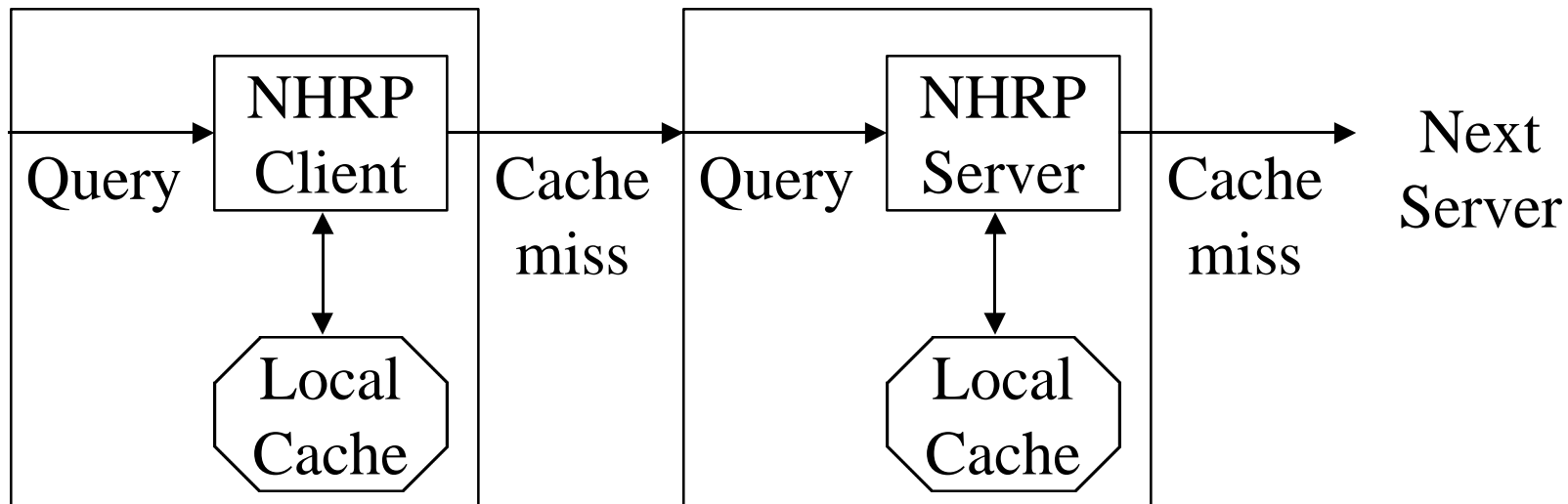
- q NBMA = Non-broadcast Multiple Access
- q Examples: X.25, Frame-relay, SMDS, ATM
- q Problem with RFC 1577 Approach: Data needs to go through routers even if on the same NBMA
- q Like going to the airport just to go to next block
- q Solution: ARP servers of different LIS's communicate



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

NHRP

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



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

NHRP Features

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

IP Address

q Class A:

0	Network	Local
1	7	24 bits

q Class B:

10	Network	Local
2	14	16 bits

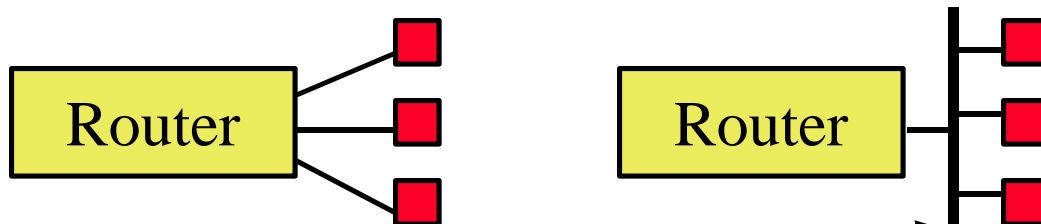
q Class C:

110	Network	Local
3	21	8 bits

q Class D:

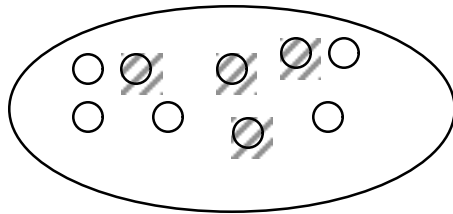
1110	Host Group (Multicast)
4	28 bits

q Local = Subnet + Host (Variable length)



IP Multicast over ATM

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

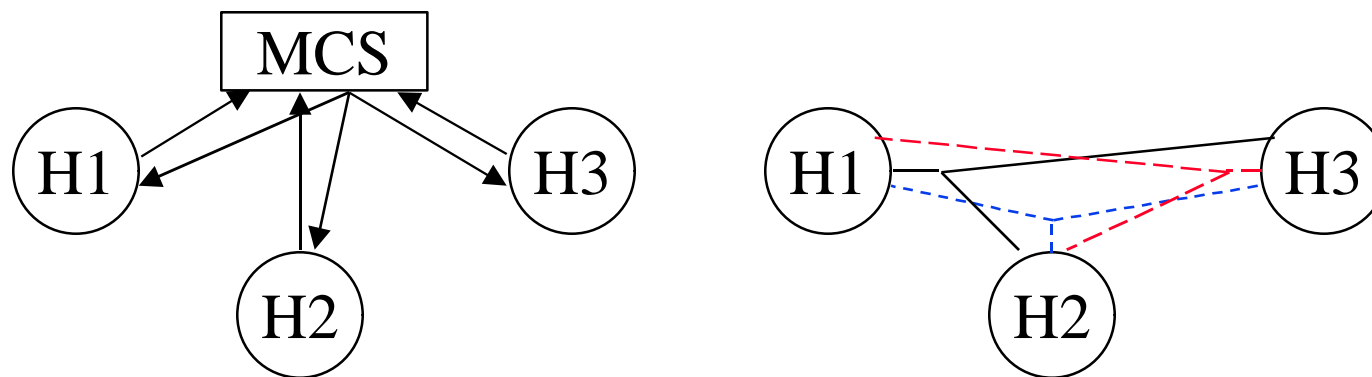


- LIS members
- ◐ Cluster members

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

Multicast Data Flow

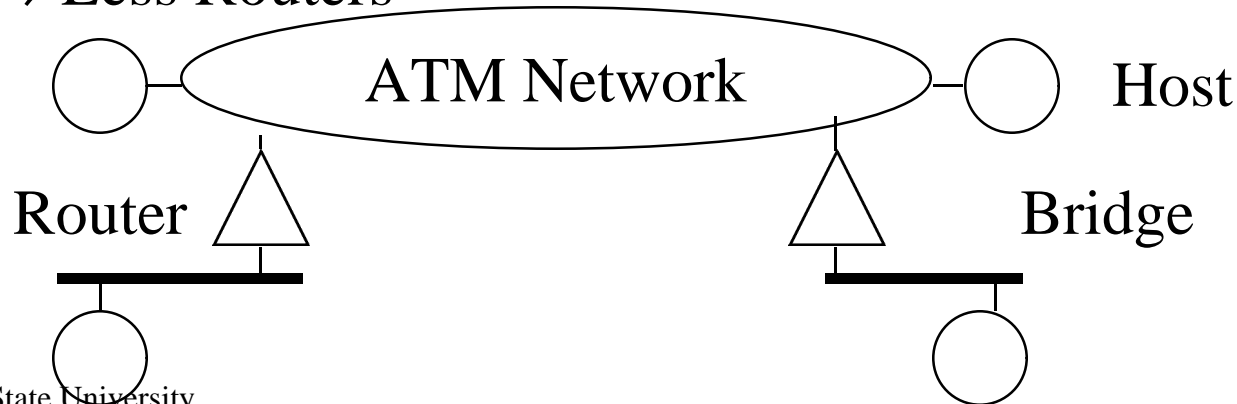
- q Multicast Servers (MCS): Retransmit packets to multicast members on a single point-to-multipoint VC (or multiple unicast VCs)
- q VC Mesh
- q Both models allowed. Each group can use different method.
- q MARS keeps a point-to-multipoint VC with all cluster members and a p-t-m VC with all MCSs



- q MARS broadcasts all join/leave messages to members so that they can adjust their p-t-m VCs accordingly
- q These messages have sequence numbers. Gap \Rightarrow Request update from MARS.
- q Clusters are connected via IP multicast routers
- q Endpoints (e.g., IP multicast routers) may join all (or ranges) of multicast groups
- q Multiple backup MARSs allowed.
- q 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
 - \Rightarrow Allows cluster members to send the packet only to MCS

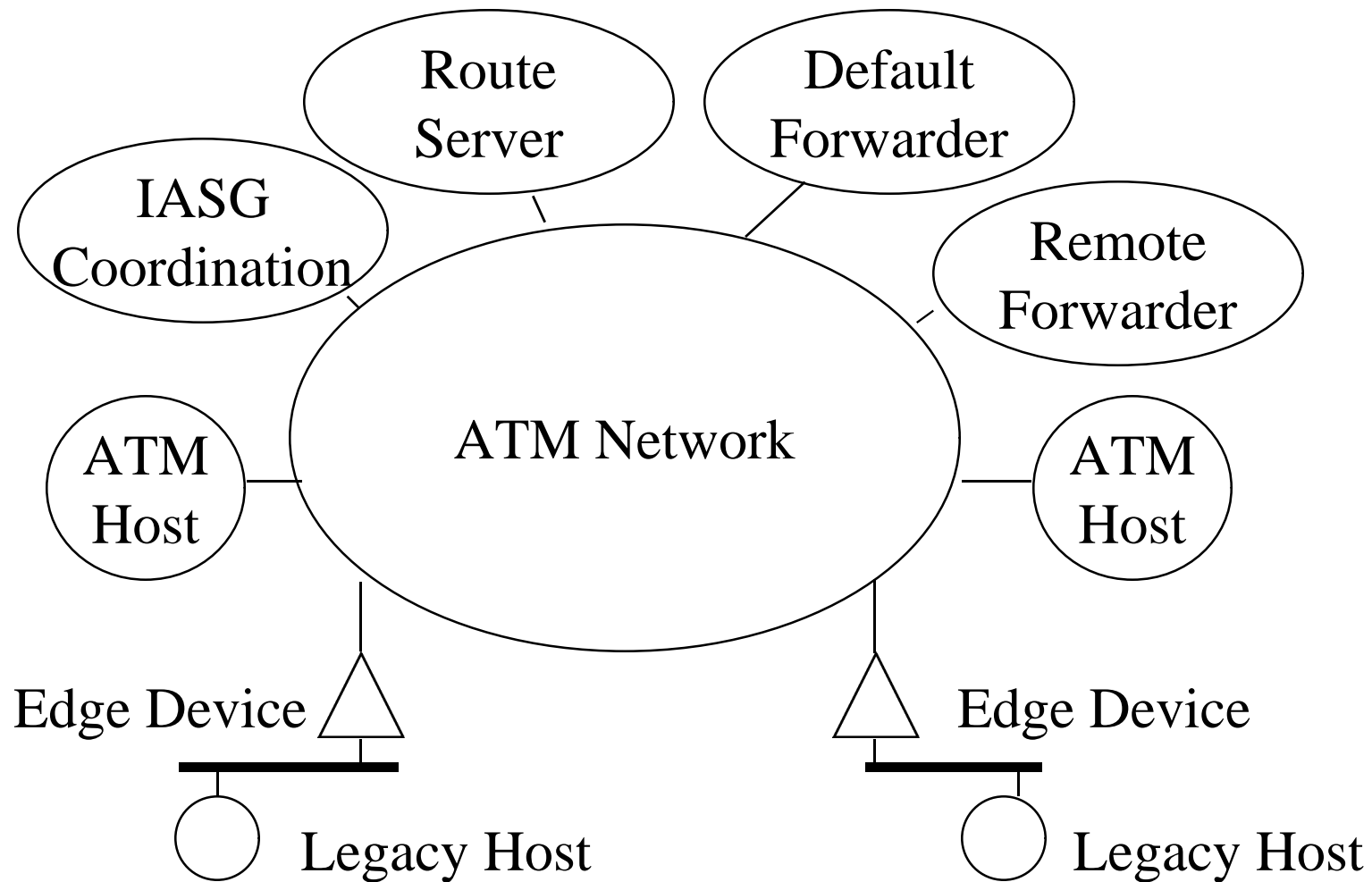
MPOA: Overview

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



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

MPOA Reference Model

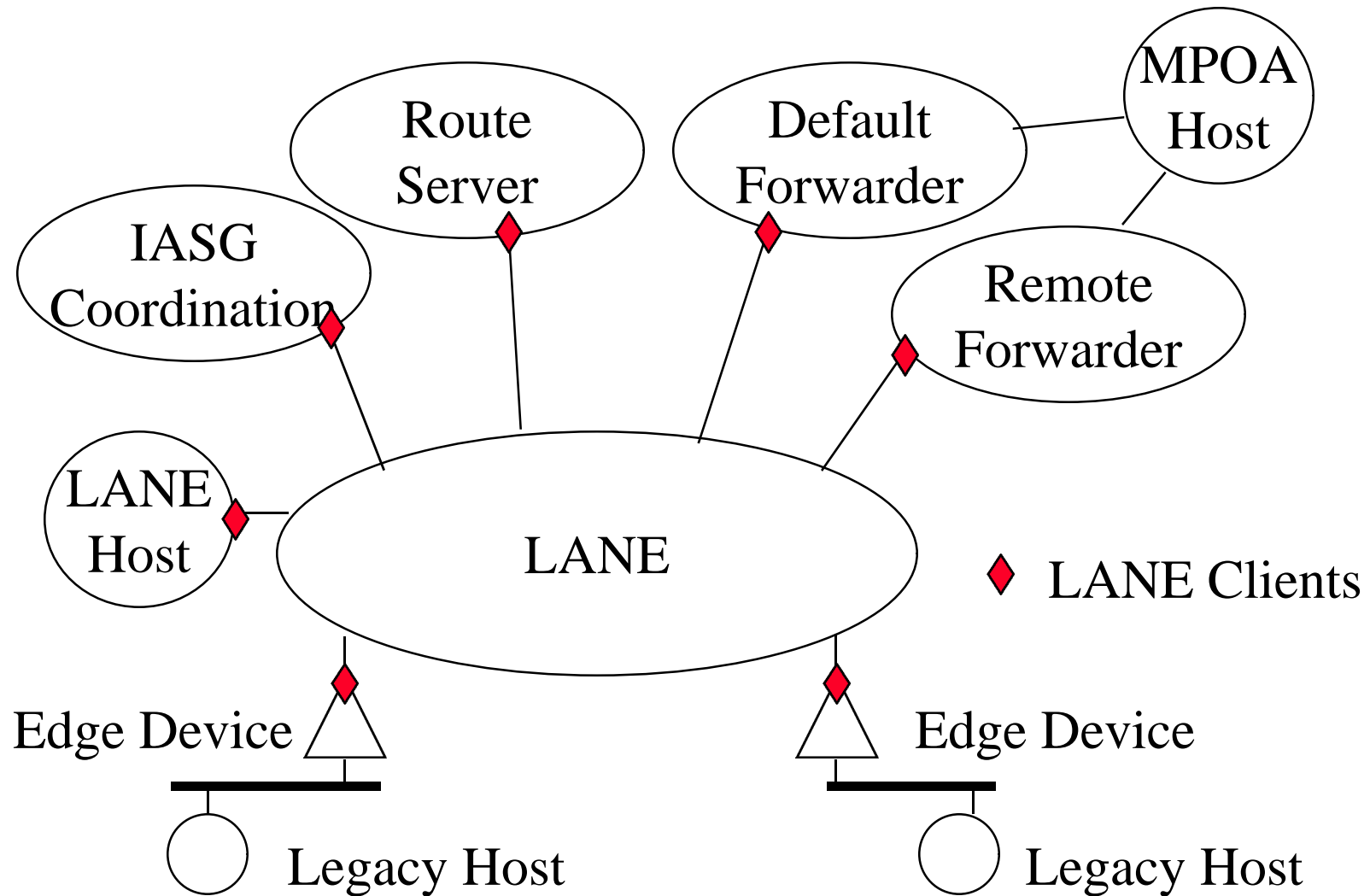


MPOA Components

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

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

MPOA Architecture



ICFG

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

Route Servers

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

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

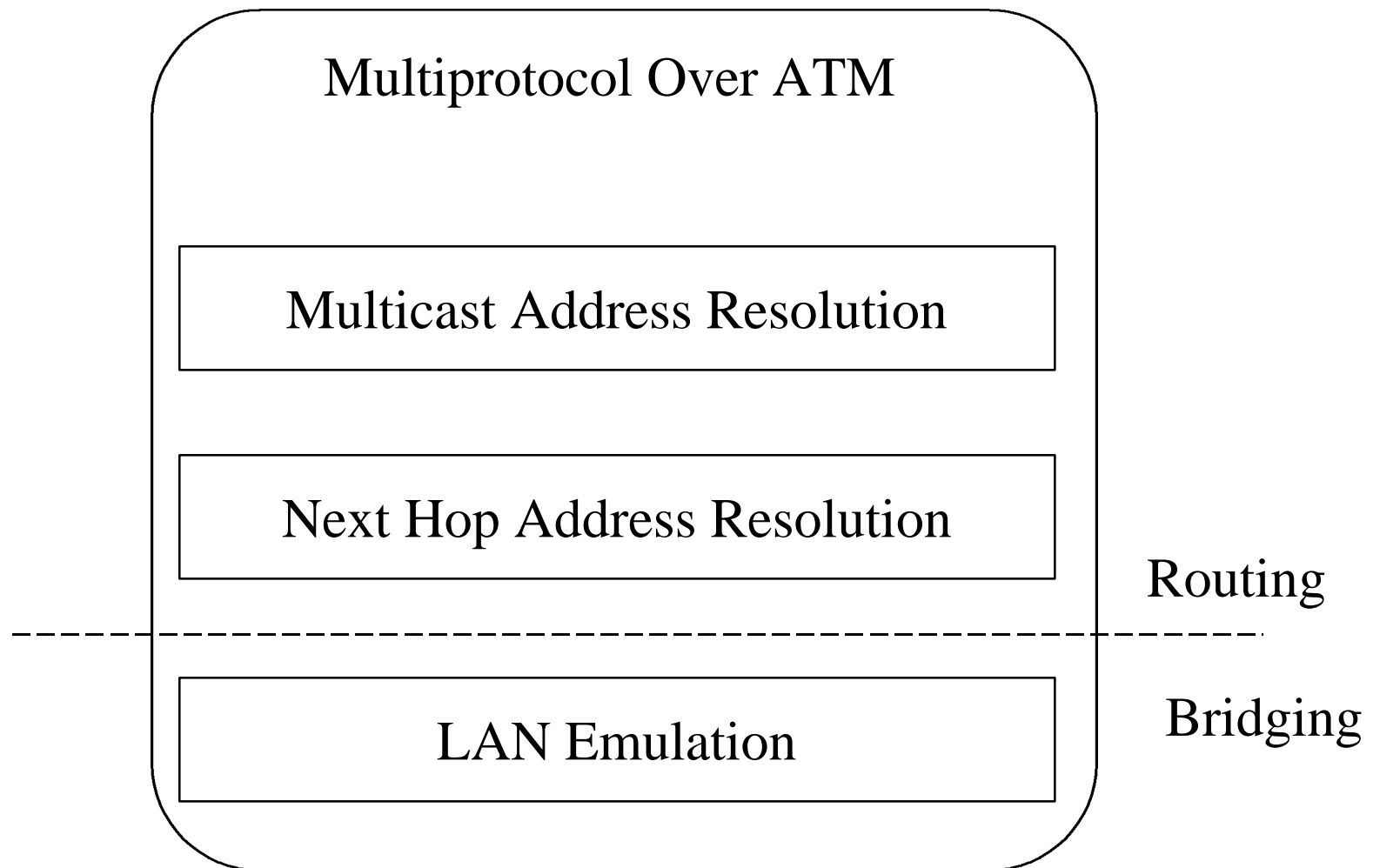
Address Resolution

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

Data Flow

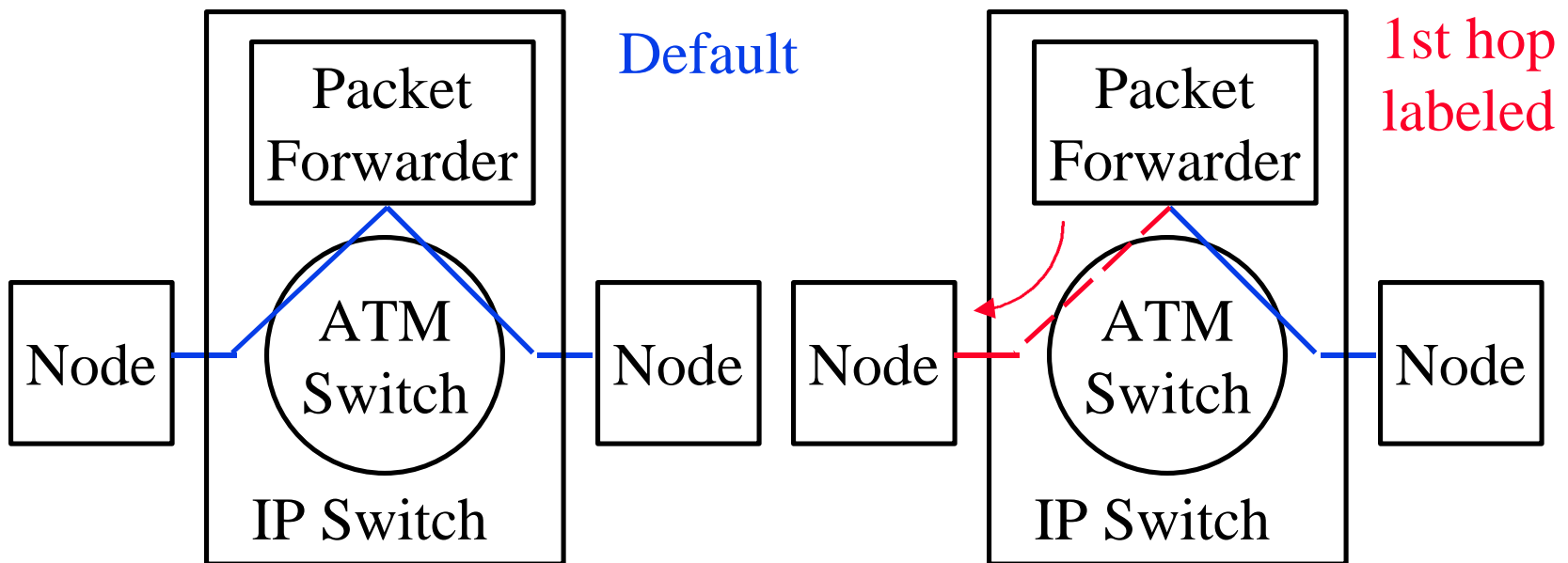
- q Default flow:
 - q Used when short cut does not exist.
 - q AHFG sends to Remote forwarder or default forwarder
 - q Edge devices use LANE for delivery based on MAC addresses
 - q Dual-mode hosts use LANE for intra-IASG traffic and use AHFG for inter-IASG traffic
 - q Default forwarder will send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

Protocol Stack

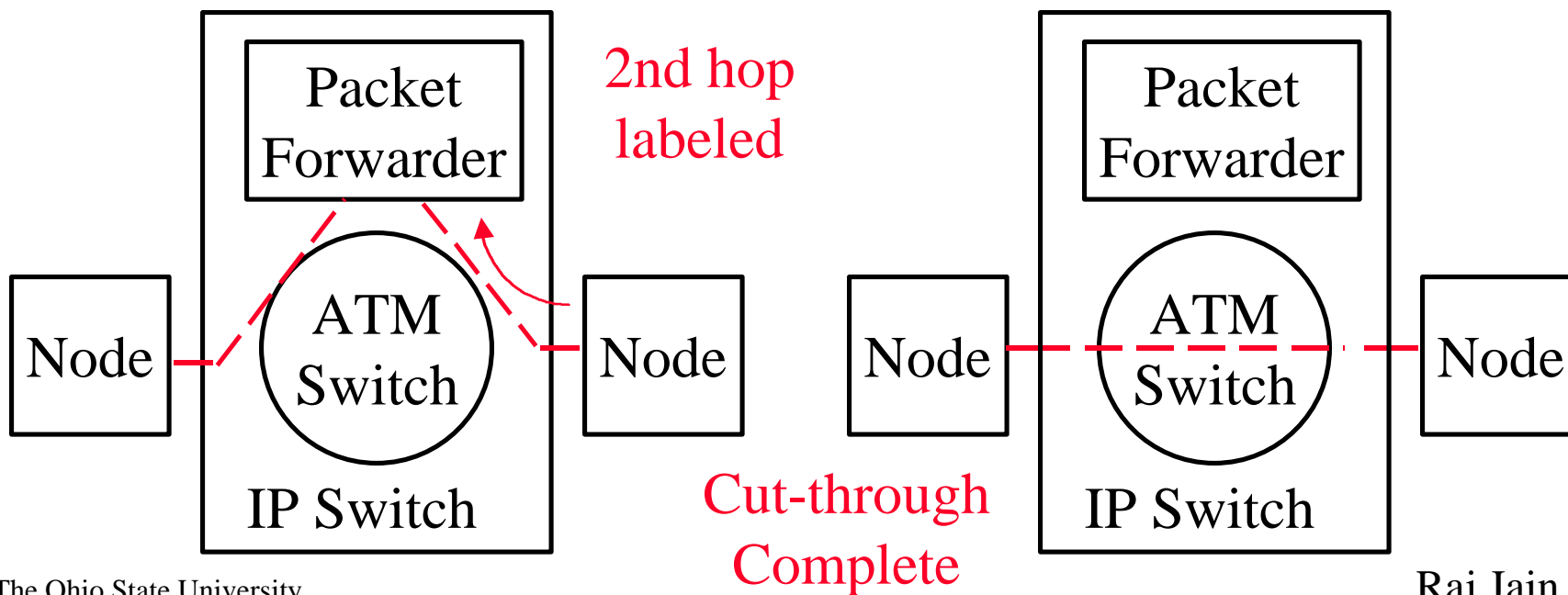


IP Switching

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



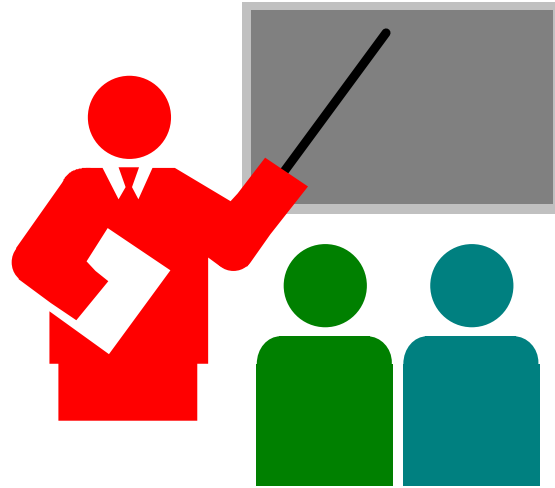
- q Downstream nodes may also ask for a new VC.
The switch then makes a mapping for cut-through
- q Flow-oriented traffic: FTP, Telnet, HTTP, Multimedia
- q Short-lived Traffic: DNS query, SMTP, NTP, SNMP, request-response
- q Ipsilon claims that 80% of packets and 90% of bytes are flow-oriented.



IP Switch Protocol Stack

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

Summary



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

References: IP over ATM

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

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

References: MPOA

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

References: IP Switching

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