Multipoint Communication over IP and ATM

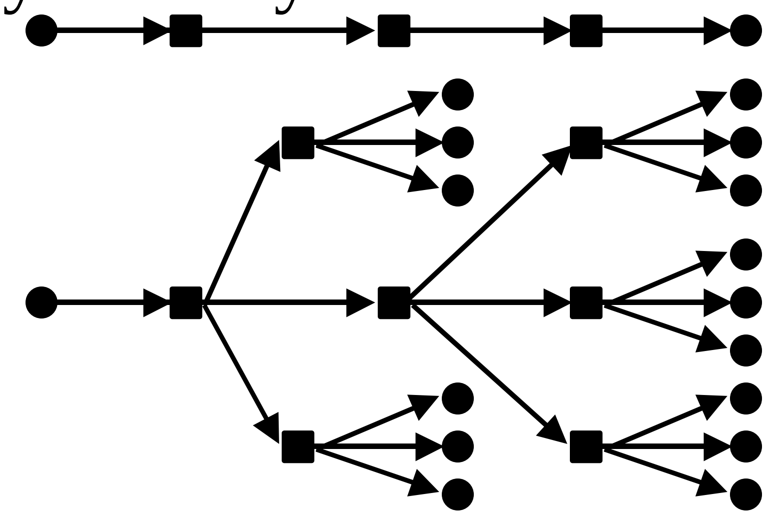
Raj Jain

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

- Why Multipoint?
- Multipoint Routing Algorithms
- Multipoint Communication in IP networks
- Multipoint Communication in ATM Networks

Multipoint Communication

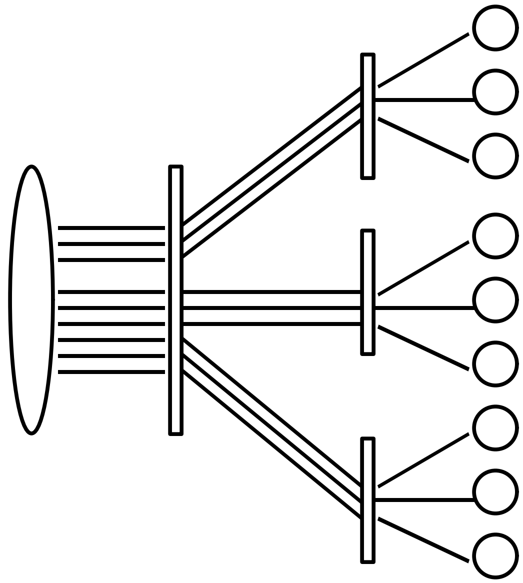
- Can be done at any layer
- Application Layer: Video Conferencing
- Transport Layer: ATM
- Network Layer: IP
- Datalink + Physical Layers: Ethernet



Multipoint Applications

- Audiovisual conferencing
- Distance Learning
- Video on Demand
- Tele-metering
- Distributed interactive games
- Data distribution (usenet, stock prices)
- Server synchronization (DNS/Routing updates)
- Advertising and locating servers
- Communicating to unknown/dynamic group

Application Layer Multipoint Comm.



- Problems: *n* times more processing/buffering/bandwidth overhead
- Applications need lower layers' help in handling unknown addresses

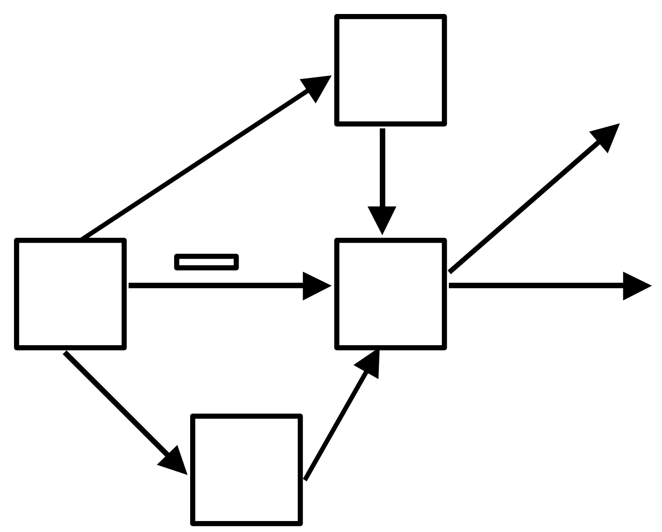
Multipoint Routing Algorithms

- Flooding
- Spanning Trees
- Reverse Path Forwarding
- Flood and Prune
- Steiner Trees
- Center-Based Trees, e.g., core-based trees

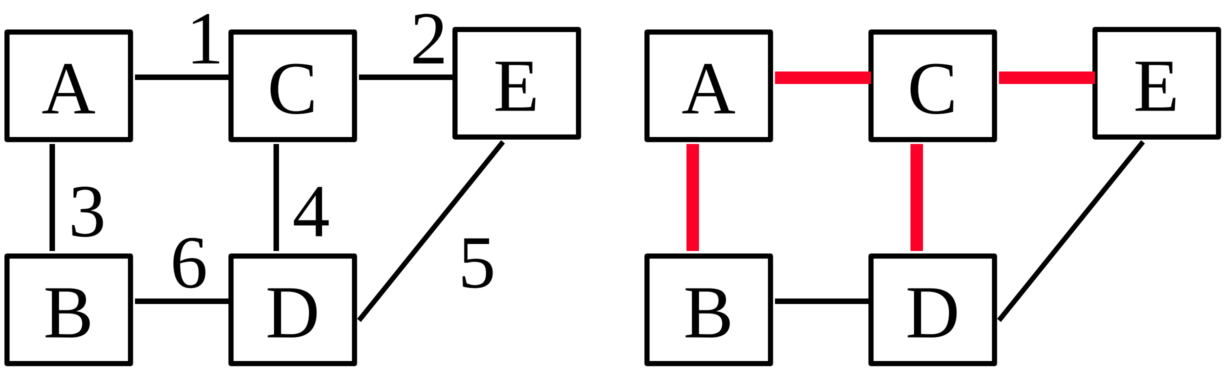
Most routing protocol standards are combination of these algorithms.

Flooding

- Used in usenet news
- Forward if first reception of this packet
  $\Rightarrow$ Need to maintain a list of recently seen packets
- Sometimes the message has a trace of recent path

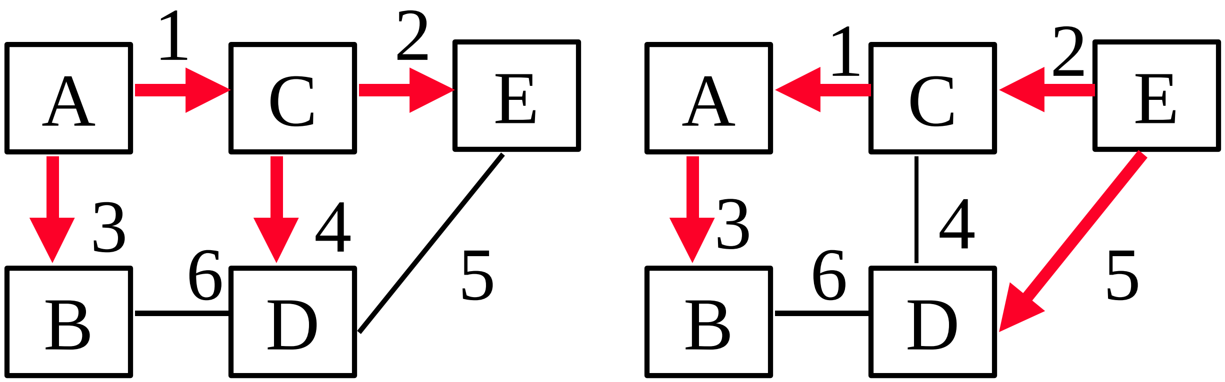


Spanning Tree



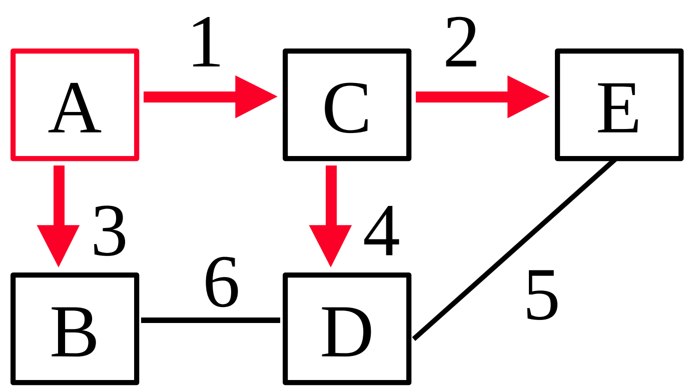
- Used by MAC bridges
- Packet is forwarded on all branches of the tree except the one it came on
- Problem:
  All packets from all sources follow the same path
  $\Rightarrow$ Congestion

Reverse Path Forwarding



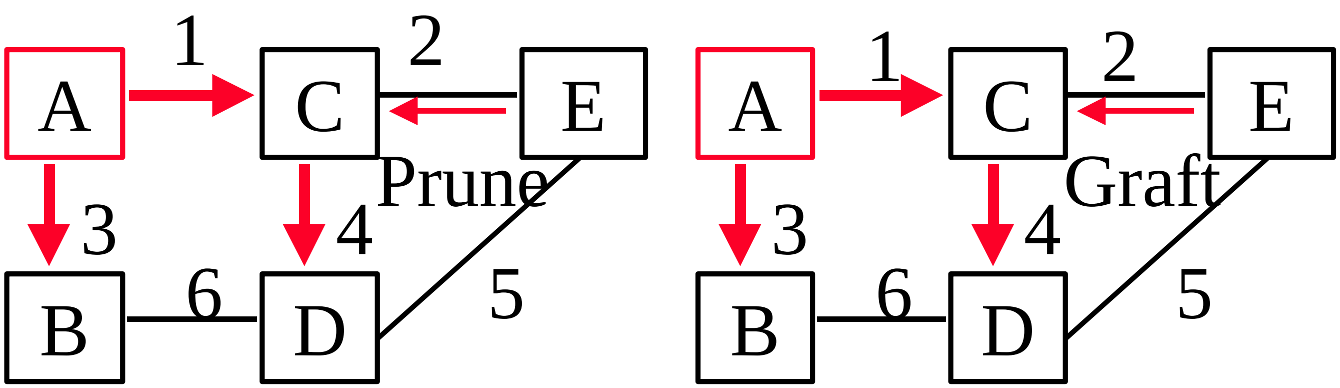
- Also known as reverse path broadcasting (RPB)
- Used initially in MBone
- On receipt, note source S and interface I
- If "I" belongs to shortest path towards S, forward to all interfaces except I
- Otherwise drop the packet

RPF (Cont)



- Optionally, check and forward only if the node is on the shortest path to the next node
- Implicit spanning tree. Different tree for different sources.
- Problem: Packets flooded to entire network

Flood and Prune



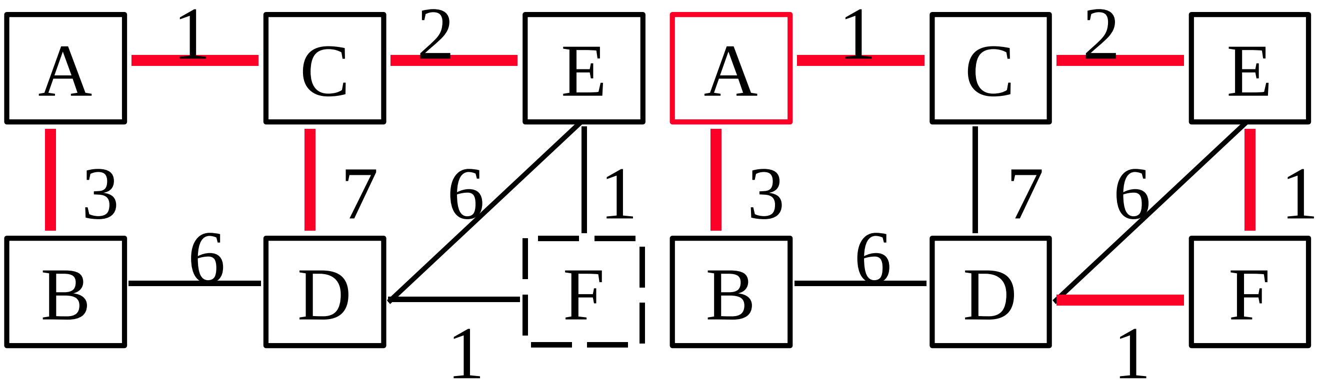
No listeners at E

Listeners at E

- Also known as reverse path multicasting (RPM)
- Used in MBone since September 1993
- First packet is flooded
- All leaf routers will receive the first packet

- If no group member on the subnet,
  the router sends a "prune"
- If all branches pruned, the intermediate router sends a "prune"
- Periodically, source floods a packet
- Problem: Per group and per source state

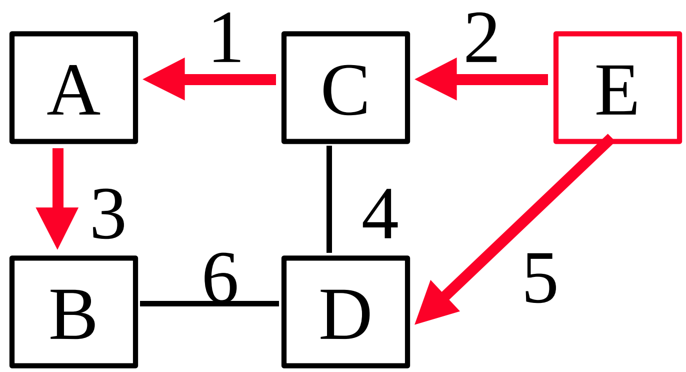
Steiner Trees



(a) F is not a member (b) F joins the group

- Centralized algorithm to compute global optimal spanning tree given all listeners
- Applies only if links are symmetric
- NP Complete ⇒ Exponential complexity ⇒ Not implemented
- Tree varies with the membership ⇒ Unstable

Center-Based Trees



- Aimed at multiple senders, multiple recipients
- Core-based tree (CBT) is the most popular example
- Choose a center
- Receivers send join messages to the center (routers remember the input interface)
- Senders send packets towards the center until they reach any router on the tree

CBT (Cont)

- Possible to have multiple centers for fault tolerance
- Routers need to remember one interface per group (not per source) $\Rightarrow$ More scalable than RPF
- Problem: Suboptimal for some sources and some receivers

Multipoint Routing Protocols

- Reverse Path Forwarding (RPF)
- Distance-vector multicast routing protocol (DVMRP): Flood and prune
- Multicast extensions to Open Shortest-Path First Protocol (MOSPF): Source-based trees (RPF)
- Protocol-Independent Multicast - Dense mode (PIM-DM): Flood and prune
- Protocol-Independent Multicast - Sparse mode (PIM-SM): Core-based trees

IP Multicast: Design Principles

- Single address per group
- Members located anywhere
- Members can join and leave at will
- Senders need not be aware of memberships
  Like a TV channel ⇒ Scalable
- Sender need not be a member
- Soft connections ⇒ periodic renewal

IP vs ATM

Category	**IP/RSVP**	**ATM UNI 3.0**
Orientation	Receiver based	Sender based
State	Soft	Hard
QoS Setup time	Separate from route establishment	Concurrent with route establishment
Directionality	Unidirectional	Unidirectional multicast
Heterogeneity	Receiver heterogeneity	Uniform QoS to all receivers

- UNI 4.0 adds leaf initiated join.

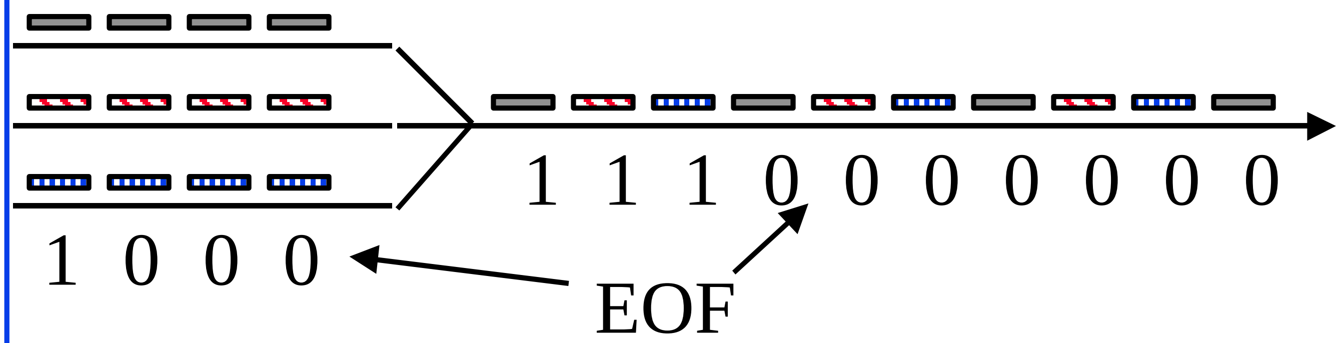
Multiway Communication on ATM

- ATM Forum Multiway BOF formed in June 1996 after marketing studies indicated high user interest
- ITU Study group 13 on ATM based multiway communications technologies
- ITU Study group 11 on Signaling requirements for Capability Set 3 (Multimedia) specifies 4 types of multipoint connections.

Multiway on ATM (Cont)

- Type 1: point-to-point
- Type 2: Point-to-multipoint
  - Unidirectional
  - Bi-directional with nonzero return bandwidth
- Type 3: Multipoint-to-point
- Type 4: Multipoint-to-Multipoint
- Variegated VCs
  $\Rightarrow$ Receivers with different bandwidth
  Applications: Video distribution, stock market

Key Issues



- Routing and packet multiplexing
- Packet multiplexing not allowed in AAL5
- AAL 3/4 has a 10-bit multiplexing ID in each cell payload ⇒ 1024 packets can be intermixed

ATM Multiway Methods

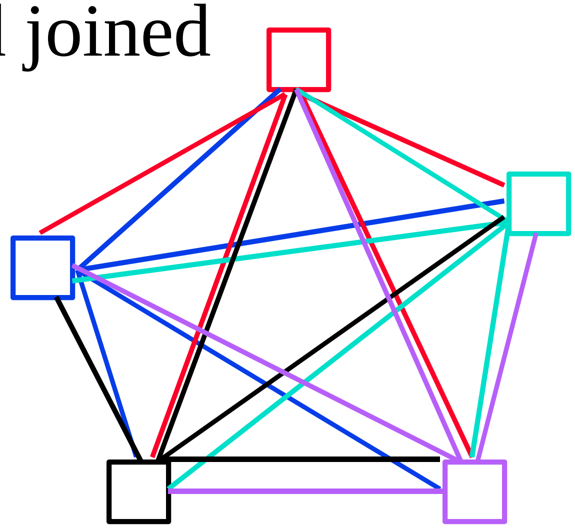
1. LAN Emulation
   ⇒ Broadcast and Unknown Server (BUS)
2. MPOA
   ⇒ Multicast Address Resolution Server (MARS)
3. VC Mesh: Overlaid pt-mpt Connections
4. Multicast Server (MCS)
5. SEAM
6. SMART
7. VP Multicasting
8. Subchannel multicasting

IP Multicast over ATM

- Need to resolve IP multicast address to ATM address list
  $\Rightarrow$ Multicast Address Resolution Servers (MARS)
- Multicast group members send IGMP join/leave messages to MARS
- Hosts wishing to send a multicast send a resolution request to MARS

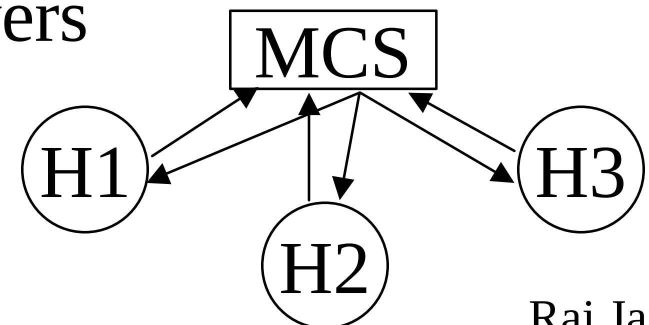
Overlaid pt-mpt Connections

- Also known as VC Mesh
- Each sender in the group establishes a pt-mpt connection with all members
- Problem: VC explosion, new members should be advertised and joined



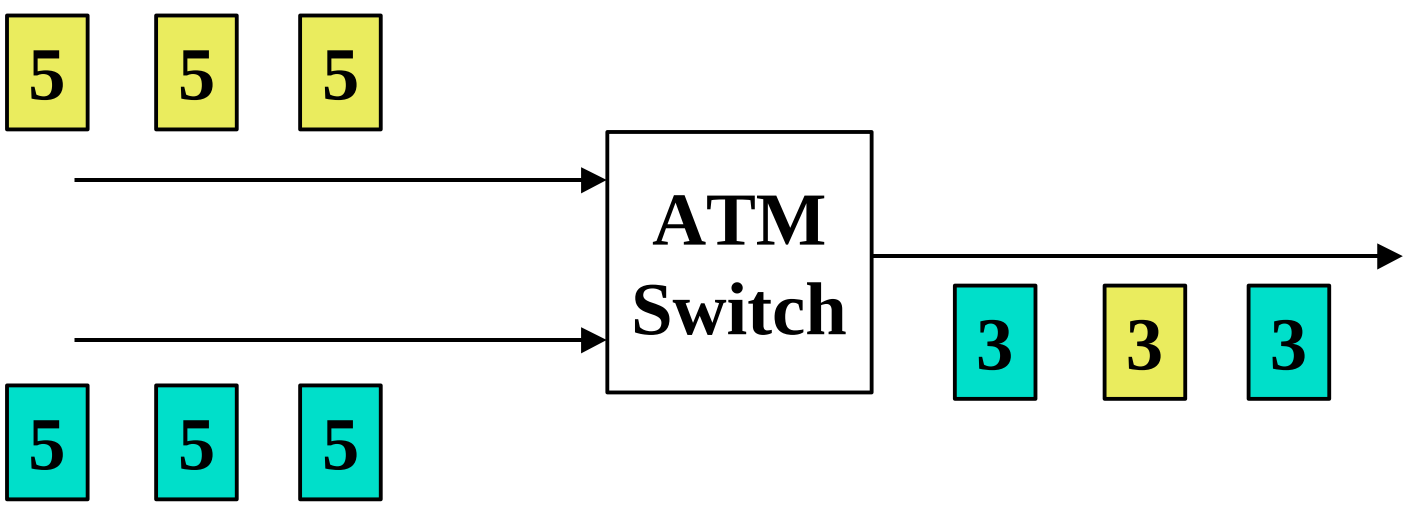
Multicast Server (MCS)

- All hosts send to MCS
  MCS has a single mpt VC to all members
- MCS serializes the packets ⇒ Does not intermingle cells of packets from different incoming VCs
- Problems with MCS:
  - Reflected packets
  - Single point of congestion
- Better for dynamic set of receivers

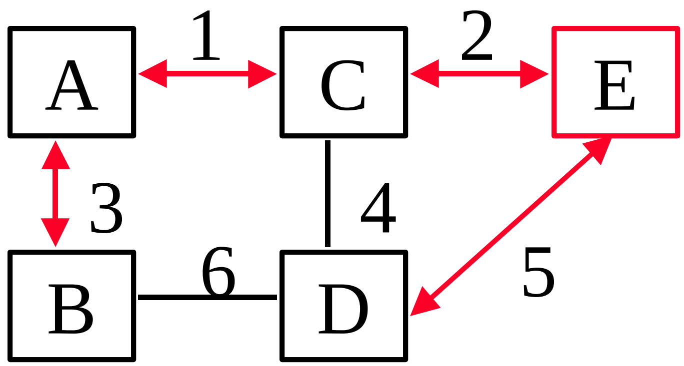


VC Merge

- Allows multipoint to point flow
- All cells of one source are switched until the last cell of the packet
- Cells from other sources on the same VC wait



SEAM



- Scalable and Efficient ATM Multipoint-to-multipoint Communication
- Uses core-based tree
- At merging points, switches have to store all cells of a packet (reassembly is not required)
  $\Rightarrow$ Packet switching (Authors call it "cut through")
- Ref: M. Grossglauser and K.K. Ramakrishnan, ATM Forum/96-1142, August 1996.

SMART

- Shared Many-to-many ATM Reservations
- Needs only one VCC but allows using multiple VCCs for performance and reliability
- Limits to one transmitter at a time. Token holder (root) can transmit.
- Anyone wishing to transmit data, must request the token from current root and become new root.
- Ensures that there only one transmitter in the tree $\Rightarrow$ No cell interleaving
- Ref: E. Gauthier, et al, IEEE JSAC, April 1997

SMART (Cont)

- Data blocks delineated by RM cells
- Not scalable for very large ATM networks or for small interactions

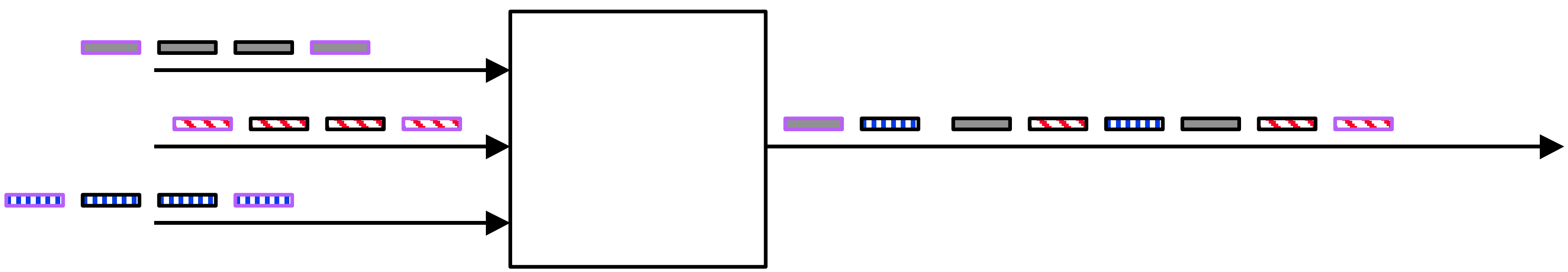
VP Multicasting

- A single VP is setup connecting all nodes
- Each source is given a unique VCI within the VP
- Problem: Size limited
- VPs are used by carriers for other purposes

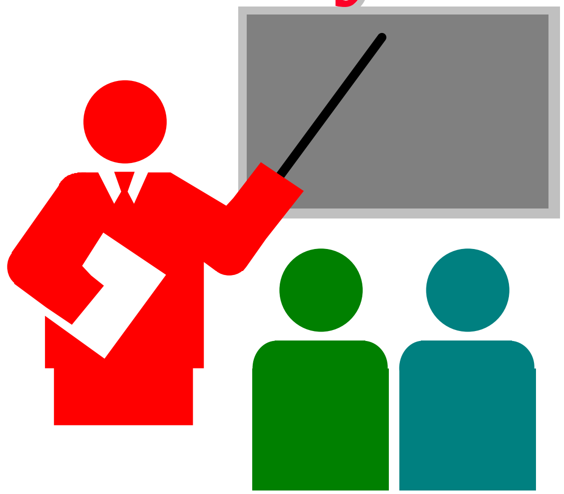
Subchannel Multicasting

- Used in Washington University's Giga Switch
- Use GFC to provide 15 subchannels for each VC (FF indicates idle subchannel)
- Each burst is preceded and followed by "Start" and "End" RM cells.
- Subchannel is allocated on the first RM cell and released on the last.
- Subchannel IDs are changed at every switch (just like VC IDs)

- Allows multiplexing up to 15 simultaneous packets at each switch port per VC.
- If a Start RM cell is received and no subchannel is available, the burst is lost.
- Jon Turner claims the loss probability is less than 10^{-12}



Summary



- Multipoint communication is required for many applications and network operations
- Network and transport support
- Internet community has developed and experimented with many solutions for multipoint communication
- ATM solutions are being developed

Key References

- See http://www.cis.ohio-state.edu/~jain/refs/mul_refs.htm for further references.
- C. Huitema, "Routing in the Internet," Prentice-Hall, 1995
- T. Maufer and C. Semeria, "Introduction to IP Multicast Routing," March 1997, http://www.internic.net/internet-drafts/draft-ietf-mboned-intro-multicast-02.txt

References (Cont)

- S. Fahmy, et al, "Protocols and Open Issues in ATM Multipoint Communications," http://www.cis.ohio-state.edu/~jain/papers/mcast.htm
- C. Diot, et al, "Multipoint Communication: A Survey of Protocols, Functions, and Mechanisms," IEEE JSAC, April 1997, pp. 277-290.