LAN Switching and Traffic Classes

Raj Jain
Professor of Computer and Information Sciences

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

MBone Instructions

- Handouts for the class are available on-line: http://www.cis.ohio-state.edu/~jain/cis788-97/index.html or http://www.netlab.ohio-state.edu/~jain/cis788-97/index.html or ftp://netlab.ohio-state.edu/pub/jain/cis788-97/
- The schedule keeps changing. Please always check current schedule at: http://www.cis.ohio-state.edu/~jain/cis788-97/schedule.html

Instructions (Cont)

- Please email your positive and negative feedback about the quality of the reception as well as the content with a subject field of "Feedback" to mbone@netlab.ohio-state.edu
- If you are not able to receive the program due to some technical difficulties, please email "Feedback" to mbone@netlab.ohio-state.edu
- Please email technical questions with the subject field "Question" to mbone@netlab.ohio-state.edu. We will try to answer selected questions live.

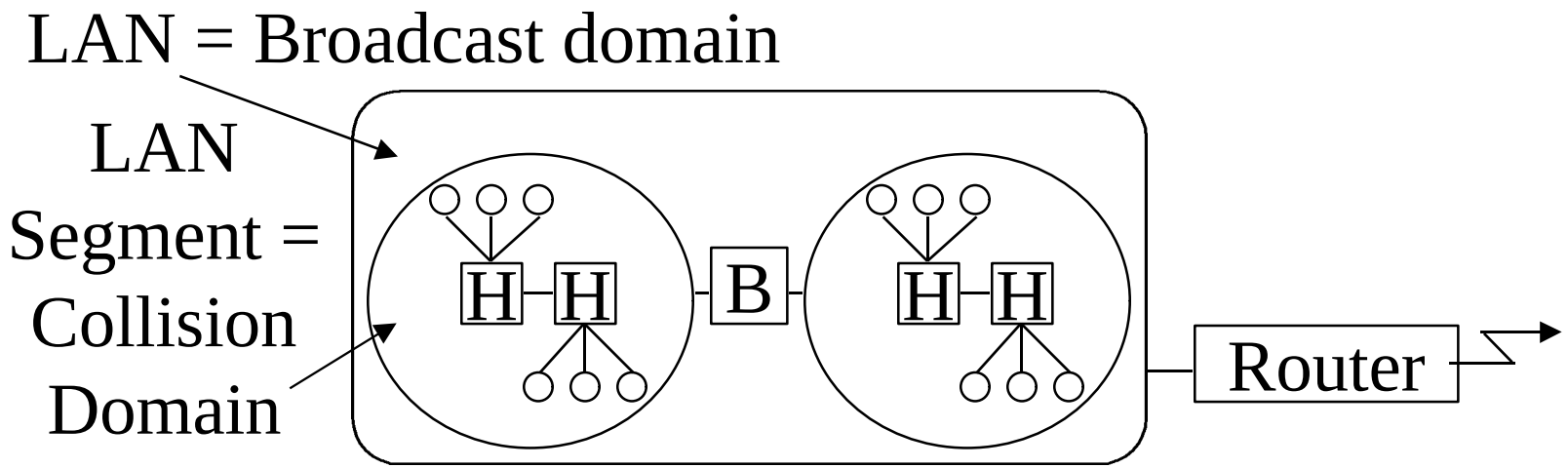
- Fundamentals: Hub, bridge, router, switch
- Full-duplex link
- Features of switches
- IEEE 802.1p standard on traffic classes in LANs and Dynamic multicast
- Generic Attribute Registration Protocol (GARP)

ATM vs LANs

- Quality of service
- Switching
- LAN emulation
  Ease of management
- Traffic management
- High Speed

- Priority
- Switching
- VLANs
- Flow control
- Gigabit Ethernet

Interconnection Devices



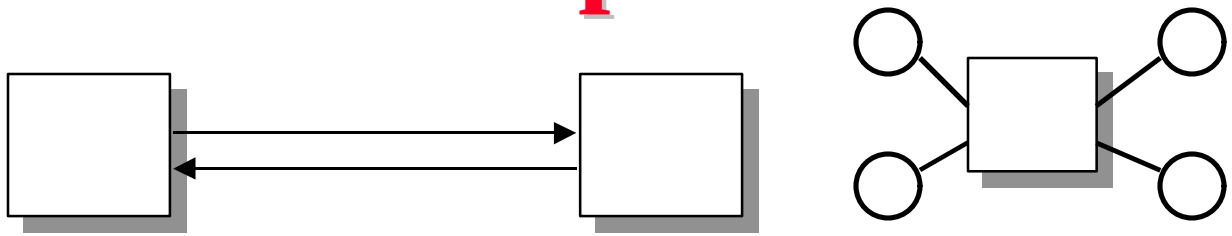
Application	Gateway	Application
Transport		Transport
Network	Router	Network
Datalink	Bridge/Switch	Datalink
Physical	Repeater/Hub	Physical

Interconnection Devices

- **Repeater**: PHY device that restores data and collision signals
- **Hub:** Multiport repeater + fault detection and recovery
- **Bridge:** Datalink layer device connecting two or more collision domains. MAC multicasts are propagated throughout "extended LAN."
- **Router:** Network layer device. IP, IPX, AppleTalk. Does not propagate MAC multicasts.
- **Switch**: Multiport bridge with parallel paths

These are functions. Packaging varies.

Full-Duplex LANs



- Uses point-to-point links between TWO nodes
- Full-duplex bi-directional transmission
  Transmit any time
- Not yet standardized in IEEE 802
- Many switch/bridge/NICs with full duplex
- No collisions ⇒ 50+ Km on fiber.
- Commonly used between servers and switches or between switches

Frame Switches: Features

- **Forwarding Type**: Cut-through, Store and forward
  - Cut-through gives lower latency but erroneous/partial frames get forwarded
  - Collision fragments (runt)
    ⇒ Adaptive Cut-through (after 64 bytes)
- **Switch Matrix**: Cell vs Frame switching
  - Frame switching mostly
- **Buffer Sharing**: Static or dynamic (based on usage), Input or output buffer

Switch Features (Cont)

- **Flow Control**: Switch jams the input port
  - Some switches jam all traffic
  - Others selectively jam packets only if they are going to congested port
- **Number of MAC addresses per port**: Small in workgroup switches
- **VLAN support**: by port, by MAC address, by subnets
  - Some allow stations to be in multiple VLANs, others don't.

Switch Features (Cont)

- **Routing**: Some switches route IP, IPX, and/or AppleTalk internally. Others require external routers between VLANs.
- **Fault Tolerance**:
  - Hot swappable media, power, uplinks, and fans.
  - Redundant port, power, mgmt processor, fans

Switch Features (Cont)

- **Manageability**:
  - Proprietary, SNMP and/or RMON support
  - Traffic monitoring using mirror ports.
  - In some switches, single mirror port can monitor multiple ports.
- **Types of LANs Supported**: Ethernet, Fast Ethernet, FDDI, Token ring, 100VGAnyLAN
- **Switch Matrix Location**: Centralized vs distributed (on each port)

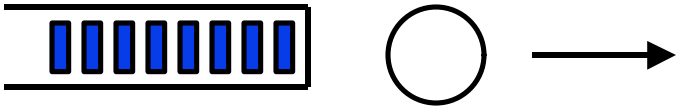
Traffic Classes in LANs

- IEEE 802.1p working group
- Goal: To support time-critical (continuous media) traffic
- Method:
  - 1. Prioritization of traffic
  - 2. Efficient support of multicasting
- Bridge filtering database for each port indicates whether any members of the group exist on the port
  ⇒ Need Group registration protocol

What's in a Name?

- The “p” in 802.1p is lower case.
- Uppercase letter ⇒ Base standard
- Lowercase letter ⇒ supplement
- 802.1p is a supplement to 802.1D bridge standard
- 802.1Q is a base VLAN standard
- 802.3z is a 1000 Mbps supplement to Ethernet Standard

Priority



- Total delay = Queueing delay + Access delay
- Access priority: As in token rings
- Queueing priority: Even in Ethernets
- User priority: Requested by the originator
  Used to determine both queueing and access priorities

IEEE 802.1p: Features

- Allows up to 8 traffic classes (priorities)
- Priority ⇒ Both queueing and access
- Allows queueing priority on LANs that have no access priorities, e.g., Ethernet
- Different number of priorities on different ports
- Allows dynamic multicast filtering
- Applies to all 802 MAC protocols + FDDI
- **802 MAC Protocols**: 802.3 (Ethernet), 802.4 (Token Bus), 802.5 (Token Ring), 802.6 (DQDB), 802.9 (Integrated Services), 802.12 (Demand Priority)

Number of Priorities

- Up to 8 traffic classes (0 through n-1).
  0 = Normal service = Low priority.
- Different ports/bridges may have different number of traffic classes
  $\Rightarrow$ Low-speed ports need priorities first
- Recommended four priorities:
  - Time and safety critical
  - Time critical
  - Non-time critical, loss sensitive
  - Non-time critical, loss insensitive

How is Priority Set?

- Priority may be set by user, destination address, input port, output port, access priority, or by VLAN
- A priority may be assigned for a port
  $\Rightarrow$ For a source station connected to a switch
- In some LANs, priority can be encoded in frames.
- In some LANs, priority cannot be encoded in frames. 802.1p does not have a mechanism to communicate priority in such LANs.
  - It has to be regenerated locally using local database, or use 802.1Q VLAN tags.

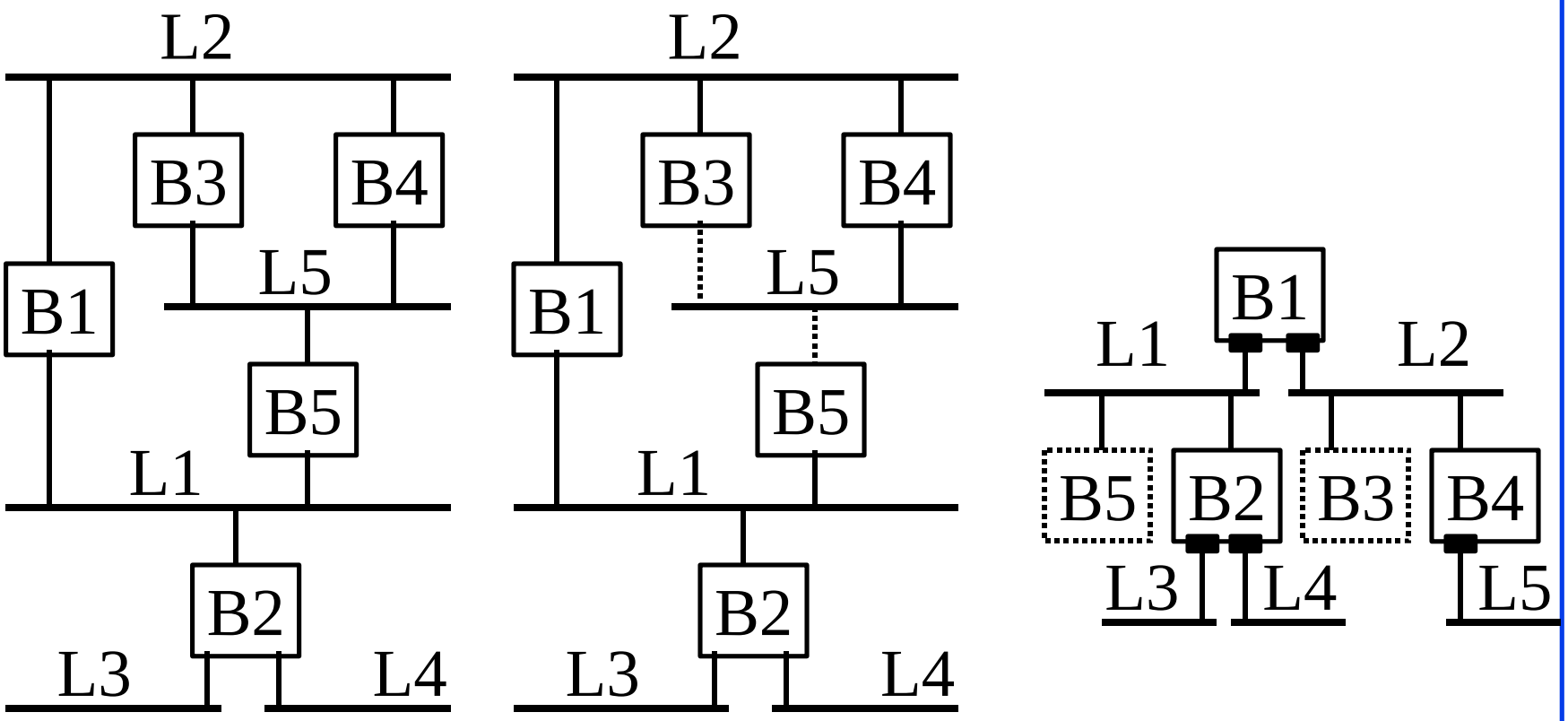
Scheduling

- Separate queue for each priority
- higher priority first (strict priority).
- No reordering of frames for a given priority and a given source and destination address pair

Multicast: Today

- Bridges forward multicast on all active ports
- A spanning tree is formed to avoid loops

Spanning Tree



(a) Original Network

(b) Active Network

(c) Spanning Tree

Spanning Tree (Cont)

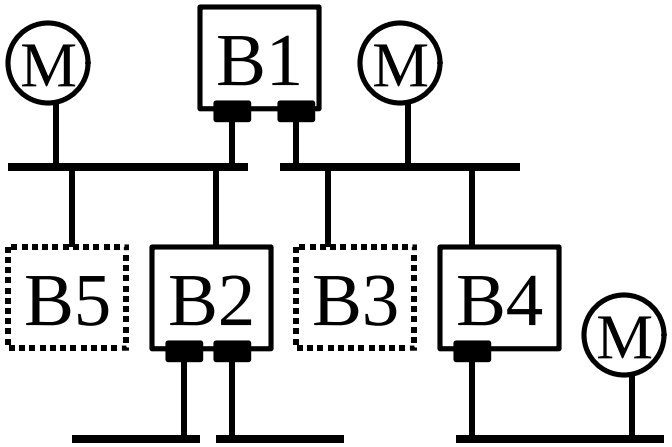
- Unique path from each source is ensured by spanning tree
- Each tree has a root bridge
- Each LAN has a designated bridge
- The port connecting the LAN towards the root is the designated port for the LAN
- The bridge containing the designated port is the designated bridge for the LAN

Dynamic Multicast Filtering

- Send multicast frames only on LANs where receivers exist
- Multicast address registration: Join/leave a group
- Legacy multicast addresses: Unregistered
- Join/leave “all groups” (Used on legacy segments)
- Join/leave “all unregistered groups” (For coexistence of legacy and new stations during migration.)
- Static entries can exclude some multicast addresses from "all groups"
- Membership information is forwarded to other bridges

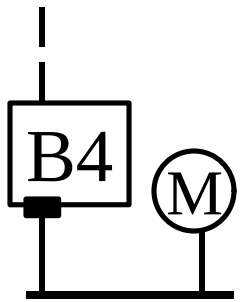
Dynamic Filtering (Cont)

- A directed subtree of the spanning tree is formed for each group
- Multicast frames are forwarded along the directed graph



Source Pruning

- Sources can stop transmission if there are no members
- Helps save local LAN bandwidth
- Implementation optional



Filtering Database

1. Static entries set by management
2. Dynamic entries:
   - a. Learning Process: Observe the source addresses of frames received on each port. Aged out.
   - b. Registration

- Static entries may specify: Forward, filter, use dynamic info to forward or filter

Input Port	Dest Address	Output Port
1	AA-01-03-44-56-78	2
2	09-12-34-56-78-88	Filter

Filtering Database (Cont)

- Priority-aware bridges use all three types of entries
- Priority-unaware bridges use only static and learned entries
- Static port entries may specify:
  - Forward all groups
  - Forward unregistered groups
  - Filter unregistered groups.

GARP

- Generic Attribute Registration Protocol
- General purpose registration/distribution protocol
- The information is propagated, if necessary, to all GARP-aware bridges
- Attribute numbers have been standardized
- GARP is used for multicast and VLAN registration.
- Registrar: Records registrations by other participants on the segment. Does not send any messages.
- Applicant: Sends registration requests and queries

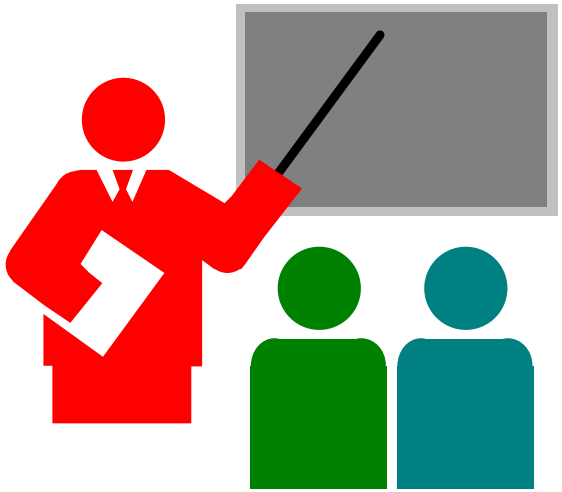
GARP Messages

- If two other stations have joined a group on your segment, you do not need to join. Protocol works even if one message is lost.
- JoinIn: I know that one other station is listening to this group. I want to join too.
- Leave: I want to leave.
- LeaveAll: "Everybody! This attribute will be de-registered soon. Rejoin if you want."
- Empty: Are there any members of this attribute?
- JoinEmpty: I have not seen any other station join this group. I am the first one to Join.

GMRP

- Group Multicast Registration Protocol
- A GARP application
- Attributes:
  - 1 = Group address registration
  - 2 = Service requirement registration
    - 0 = All groups
    - 1 = All unregistered groups

Summary



- LAN switches = Multiport bridges
- Traffic classes and dynamic multicast on LANs to allow multimedia
- IEEE 802.1p allows 8 priorities
- Distributed multicast registration protocol

References

- For a detailed list of references, see http://www.cis.ohio-state.edu/~jain/refs/lsw_refs.htm
- IEEE 802.1 Email list: p8021-request@hepnrc.hep.net
  Mail archive: http://www.hep.net/mail/p8021.html
- Traffic Class Expediting and Dynamic Multicast Filtering, IEEE P802.1p/D6, April 28, 1997.
- Other Related Standards
  - 802.1D MAC bridges
  - 802.1G Remote MAC Bridging
  - 802.1H Ethernet V2.0 and 802 bridging

Current Schedule

7/17/97 Priority and Multicasting on LANs

7/22/97 No Class

7/24/97 Virtual LANs

7/29/97 Gigabit Ethernet

7/31/97 Quiz 2 (No MBone transmission)

8/5/97 Residential broadband: Cable Modems, xDSL

8/7/97 Multimedia: Compression Standards

8/12/97 Multimedia over IP: RSVP, RTP

8/14/97 Wireless LANs and WANs

8/19/97 Quiz 3 (No MBone transmission)

Credits

The MBone transmission of this lecture was made possible by:

- Mark Fullmer, OSU/UTS
- Mike Iverson, OSU/UTS
- Mike Douglas, OSU/UTS
- Jayaraman Iyer, OSU/CIS
- Sohail Munir, OSU/CIS