PNNI:
Routing in ATM Networks

Raj Jain

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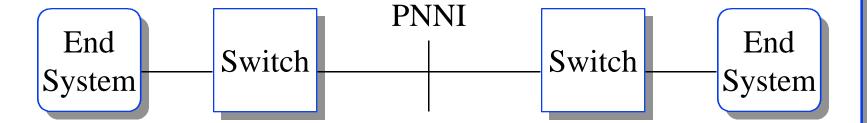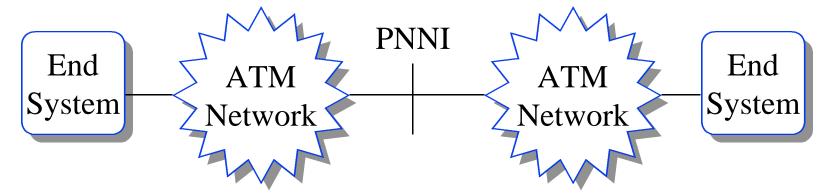
Overview

- q Distribution of topology information
- q Hierarchical groups
- q Source routing $\Rightarrow$ Designated Transit Lists
- q Crankback and Alternate routing
- q Addressing

Ref: ATM Forum 94-0471R9, "PNNI Draft Specification (Phase 1)"

PNNI

q Private Network-to-network Interface

q Private Network Node Interface

Features of PNNI

q Point-to-point and point-to-multipoint connections

q Can treat a cloud as a single logical link

q Multiple levels of hierarchy $\Rightarrow$ Scalable for global networking.

q Reroutes around failed components at connection setup

q Automatic topological discovery $\Rightarrow$ No manual input required.

q Connection follows the same route as the setup message (associated signaling)

q Uses: Cost, capacity, link constraints, propagation delay

q Also uses: Cell delay, Cell delay variation, Current average load, Current peak load

q Uses both link and node parameters

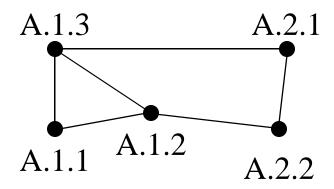q Supports transit carrier selection

q Supports anycast

Addressing

q   Multiple formats.

q   All 20 Bytes long addresses.

q   Left-to-right hierarchical

q   Level boundaries can be put in any bit position

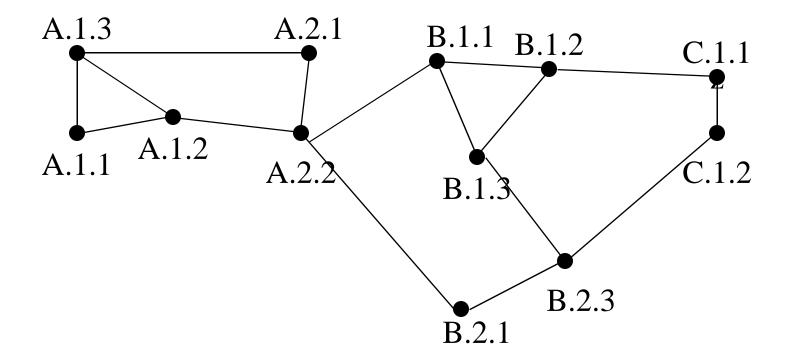q   13-byte prefix $\Rightarrow$ 104 levels of hierarchy possible

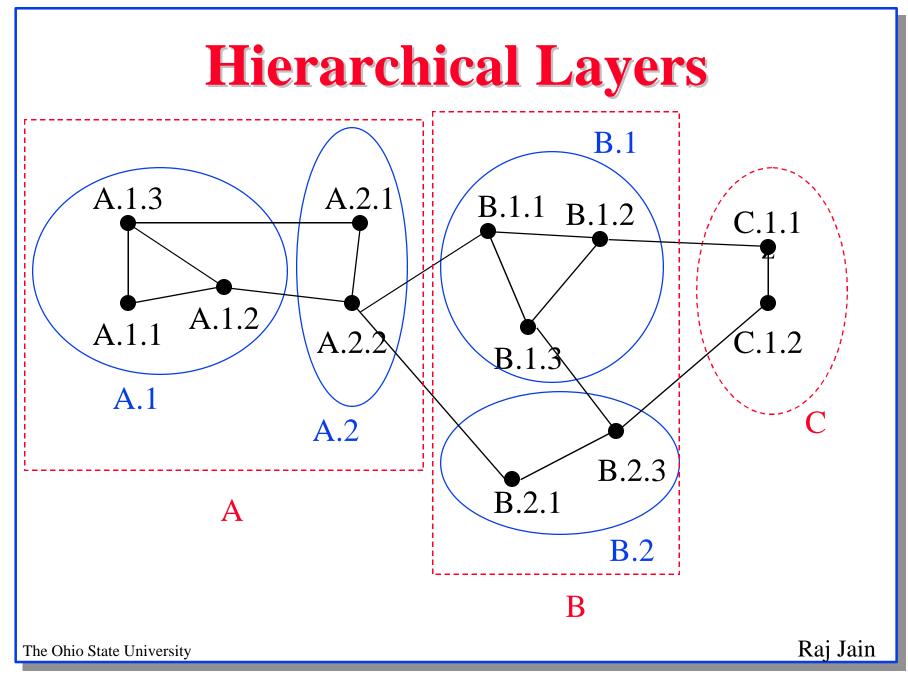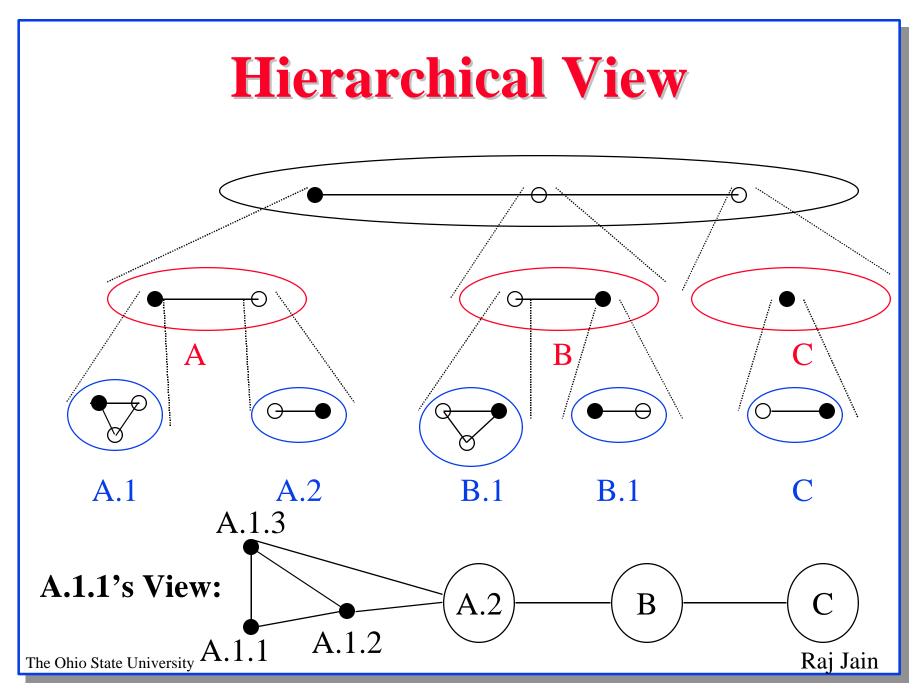| Level 1	Level 2	Level 3	Level 4

Link State Routing

q Each node sends "Hello" packets periodically and on state changes.

q The packet contains state of all its links

q The packet is flooded to all nodes in the network

A.1.3   A.2.1

A.1.1   A.1.2   A.2.2

Very Large Networks

A.1.3　　　A.2.1　　B.1.1　B.1.2　　C.1.1

A.1.1　A.1.2　　A.2.2　　B.1.3　　C.1.2

B.2.1　　B.2.3

Hierarchical Layers

Hierarchical View

A

B   C

A.1   A.2   B.1   B.1   C

A.1.3

A.1.1's View:

A.2 — B — C

A.1.1   A.1.2

Terminology

q Peer group: A group of nodes at the same hierarchy

q Border node: one link crosses the boundary

q Logical group node: Representation of a group as a single point

q Logical node or Node: A physical node or a logical group node

q Child node: Any node at the next lower hierarchy level

q Parent node: Logical group node at the next higher hierarchy level

q Logical links: links between logical nodes

Raj Jain

q Peer group leader (PGL):
Represents a group at the next higher level.
Node with the highest "leadership priority" and
highest ATM address is elected as a leader.
Continuous process $\Rightarrow$ Leader may change any time.

q PGL acts as a logical group node.
Uses same ATM address with a different selector value.

q Peer group ID: Address prefixes up to 13 bytes

Topology State Information

q Metric: Added along the path, e.g., delay

q Attribute: Considered individually on each element.

   q Performance, e.g., capacity or

   q Policy related, e.g., security

q State parameter: Either metric or attribute

q Link state parameter. Node state parameter.

q Topology = Link + Nodes

q Topology state parameter: Link or node state parameter

q PNNI Topology state element (PTSE):
Routing information that is flooded in a peer group

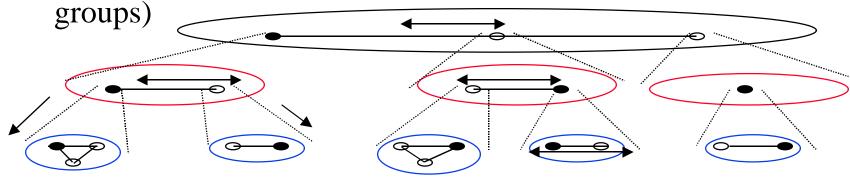q PNNI Topology state packet (PTSP): Contains one PTSE

Topology State Parameters

q Metrics:
- q Maximum Cell Transfer Delay (MCTD)
- q Maximum Cell Delay Variation (MCDV)
- q Maximum Cell Loss Ratio (MCLR)
- q Administrative weight

q Attributes:
- q Available cell rate (ACR)
- q Cell rate margin (CRM) = Allocated - Actual
  First order uncertainty. Optional.
- q Variation factor (VF) = CRM/Stdv(Actual)
  Second order uncertainty. Optional.
- q Branching Flag: Can handle point-to-multipoint traffic
- q Restricted Transit Flag: Supports transit traffic or not

Database Synchronization and Flooding

q   Upon initialization, nodes exchange PTSE headers (My topology database is dated 11-Sep-1995:11:59)

q   Node with older database requests more recent info

q   After synchronizing the routing database, they advertise the link between them

q   The ad (PTSP) is *flooded* throughout the peer group

q   Nodes ack each PTSP to the sending neighbors, update their database (if new) and forward the PTSP to all *other* neighbors

q   All PTSEs have a life time and are aged out unless renewed.

q   Only the node that originated a PTSE can reissue it.

q   PTSEs are issued periodically and also event driven.
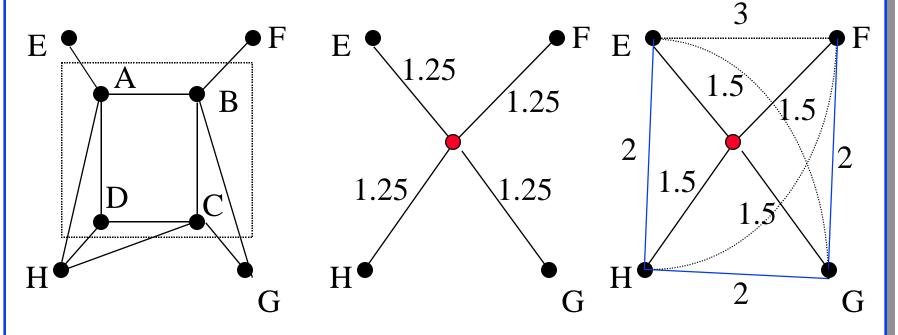
Information Flow in the Hierarchy

- q Information = Reachability and topology aggregation
- q Peer group leaders *summarize* and circulate info in the parent group
- q A raw PTSE never flows upward.
- q PTSEs flow horizontally through the peer group and downward through children.
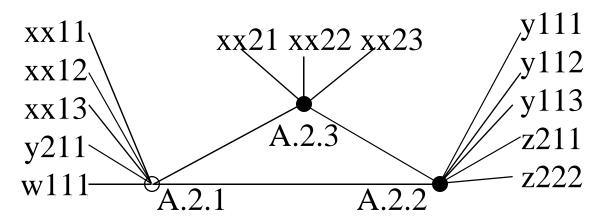- q Border nodes do not exchange databases (different peer groups)

Topology Aggregation

q   Get a simple representation of a group

q   Alternatives: Symmetric star (n links) or mesh ($n^2/2$ links)

q   Compromise: Star with exceptions

Address Summarization

xx11, xx12, xx13, y211, w111 — A.2.1 (open node)
xx21 xx22 xx23 — A.2.3 (node)
y111, y112, y113, z211, z222 — A.2.2 (node)

q Summary = All nodes with prefix xxx, yyy, ...
        + foreign addresses

q Native addresses = All nodes with prefix xxx, yyy, ...

q **Example**:

   q A.2.1 = XX1*, Y2*, W111   A.2.2 = Y1*, Z2*
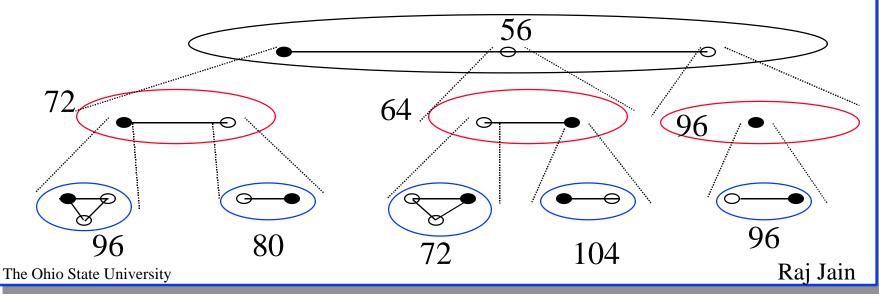
   q A.2.3 = XX2*

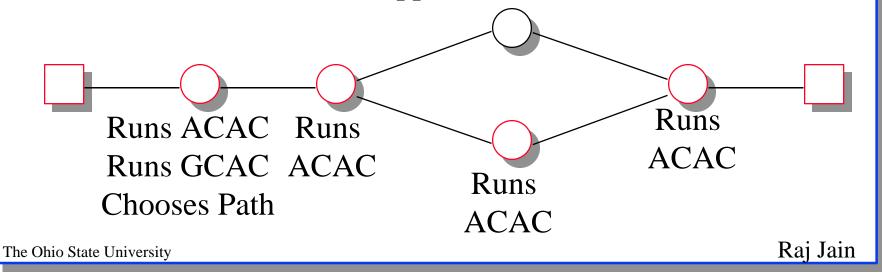   q A.2 = XX*, Y*, Z2*, W111. W111 is a foreign address

Address Scope

q   Upward distribution of an address can be inhibited, if desired.
E.g., Don't tell the competition B that W111 is reachable via A.

q   Each group has a level (length of the shortest prefix).

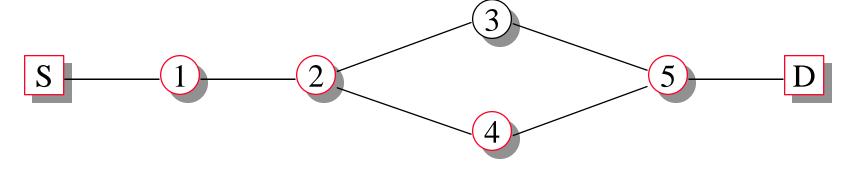q   Each address has a scope (level up to which it is visible).

56

72   64   96

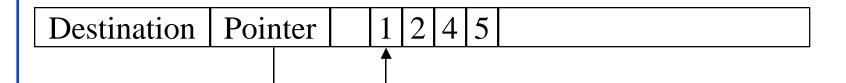96   80   72   104   96

Call Admission Control

q Generic Call Admission Control (GCAC)

   q Run by a switch in choosing a source route

   q Determines which path can probably support the call

q Actual Call Admission Control (ACAC)

   q Run by each switch

   q Determines if it can support the call

Runs ACAC
Runs GCAC
Chooses Path

Runs ACAC

Runs ACAC

Runs ACAC

Source Routing

- q Used in IEEE 802.5 token ring networks
- q Source specifies all intermediate systems (bridges) for the packet

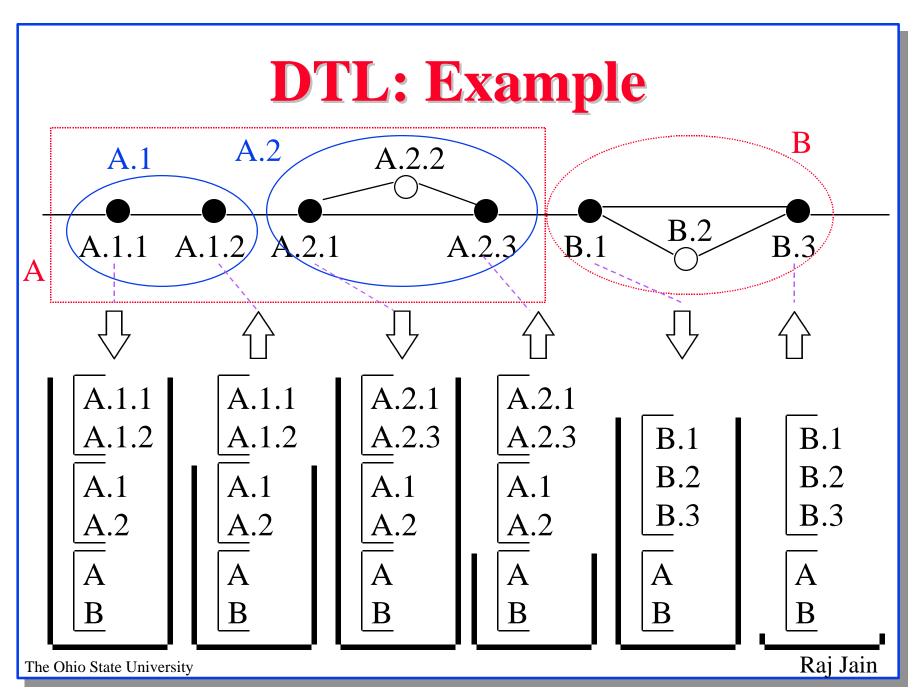| Destination	Pointer		1	2	4	5	

Designated Transit Lists

q DTL: Source route across each level of hierarchy

q Entry switch of each peer group specifies complete route through that group

q Entry switch may or may not be the peer group leader

q Multiple levels $\Rightarrow$ Multiple DTLs
Implemented as a stack

DTL: Example
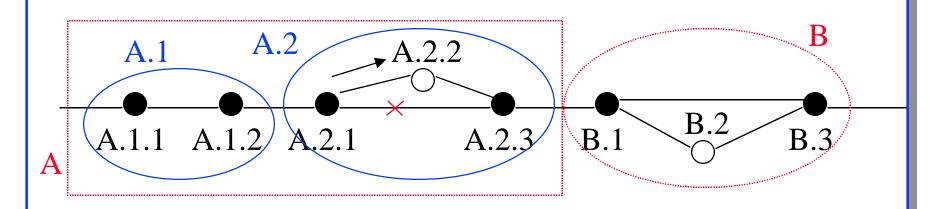
Crankback and Alternate Path Routing

q If a call fails along a particular route:

  q It is *cranked back* to the originator of the top DTL

  q The originator finds another route *or*

  q Cranks back to the generator of the higher level source route

Summary

- q Database synchronization and flooding
- q Hierarchical grouping: Peer groups, group leaders
- q Topology aggregation and address summarization
- q Designated transit lists
- q Crankback

Abbreviations

- q AFI   Authority and format identifier
- q BIS   Border intermediate system
- q BISPDU   Border intermediate system protocol data unit
- q CAC   Connection admission control
- q CNR   Complex node representation
- q CRM   Cell rate margin
- q DSP   Domain specific part
- q DTL   Designated transit list
- q ES   End system
- q ESI   End system identifier
- q GCAC   Generic connection admission control

- q IDI   Initial domain identifier
- q IDP   Initial domain part
- q IS   Intermediate system
- q LGN   Logical group node
- q LSAP   Link service access point
- q MaxCR   Maximum cell rate
- q NPDU   Network protocol data unit
- q NSAP   Network service access point
- q OSPF   Open shortest path first
- q PG   Peer group
- q PGL   Peer group leader
- q PTSE   PNNI Topology state element
- q PTSP   PNNI Topology state packet

q PNNI   Private network-network interface

q PVCC   Permanent virtual channel connection

q PVPC   Permanent virtual path connection

q RD   Routing domain

q SAAL   Signaling ATM adaptation layer

q SNPA   Subnetwork point of attachment

q TIG   Topology information group

q TLV   Type, length, value

q VF   Variance factor

References

q B. Dorling, D. Freedman, C. Metz, and J. Burger, "Internetworking over ATM: An Introduction," Prentice Hall, 1996.

q D. Dykeman and M. Goguen, "PNNI Draft Specification V1.0," af-pnni-0055.000, March 1996. ftp://ftp.atmforum.com/pub/approved-specs/af-pnni-0055.000.ps

q A. Alles, "Routing and Internetworking in ATM Networks," Networld+InterOP, March 1995

q W.C. Lee, "Topology Aggregation for Hierarchical Routing in ATM Networks," Computer Communication Review, April 1995.

References (cont)

q J.M. Halpern, "ATM Call Routing," Communication Systems Design, pp. 30-35.

q R. Callon, et al, "Issues and Approaches for Integrated PNNI," ATM Forum 96-0355, April 1996.