PNNI:
Routing in ATM Networks

Raj Jain

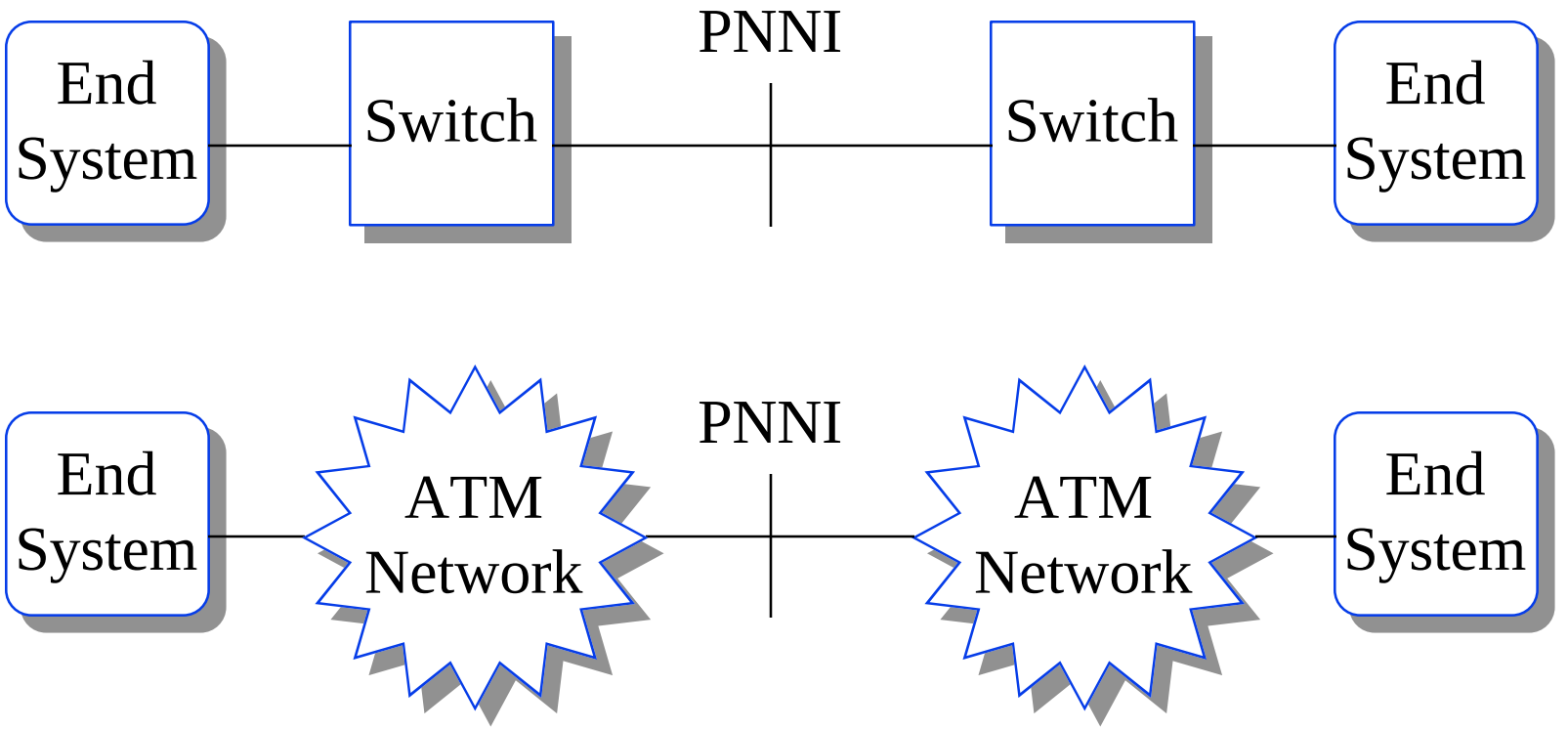
Professor of CIS

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

- Distribution of topology information
- Hierarchical groups
- Source routing ⇒ Designated Transit Lists
- Crankback and Alternate routing
- Addressing

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

PNNI



- Private Network-to-network Interface
- Private Network Node Interface

Features of PNNI

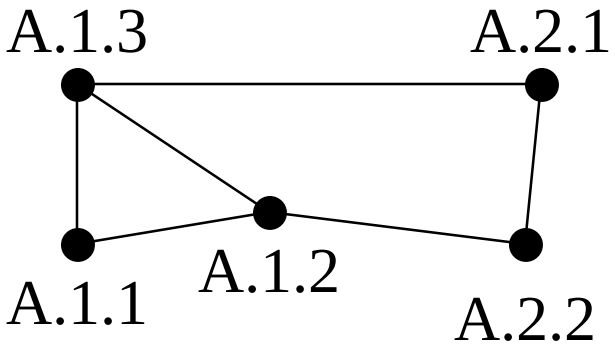
- Point-to-point and point-to-multipoint connections
- Can treat a cloud as a single logical link
- Multiple levels of hierarchy ⇒ Scalable for global networking.
- Reroutes around failed components at connection setup
- Automatic topological discovery ⇒ No manual input required.
- Connection follows the same route as the setup message (associated signaling)
- Uses: Cost, capacity, link constraints, propagation delay
- Also uses: Cell delay, Cell delay variation, Current average load, Current peak load
- Uses both link and node parameters
- Supports transit carrier selection
- Supports anycast

Addressing

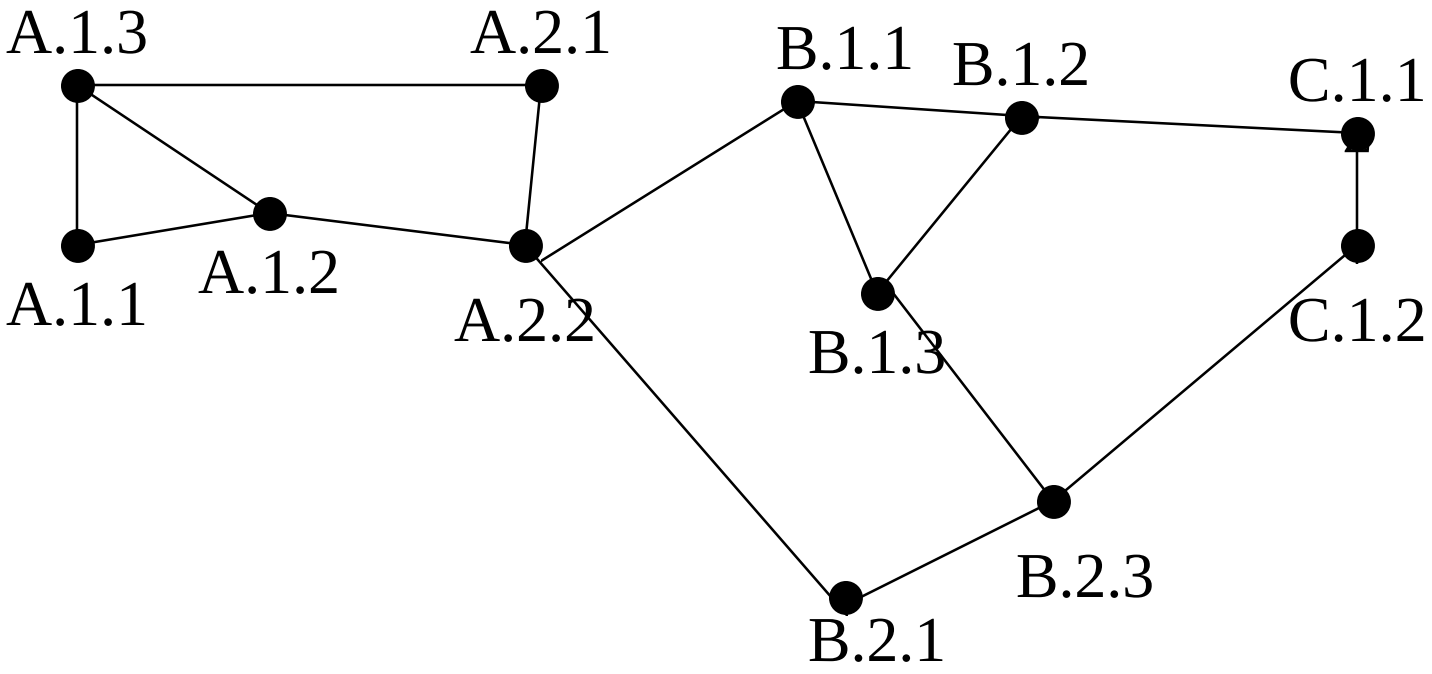
- Multiple formats.
- All 20 Bytes long addresses.
- Left-to-right hierarchical
- Level boundaries can be put in any bit position
- 13-byte prefix ⇒ 104 levels of hierarchy possible

Link State Routing

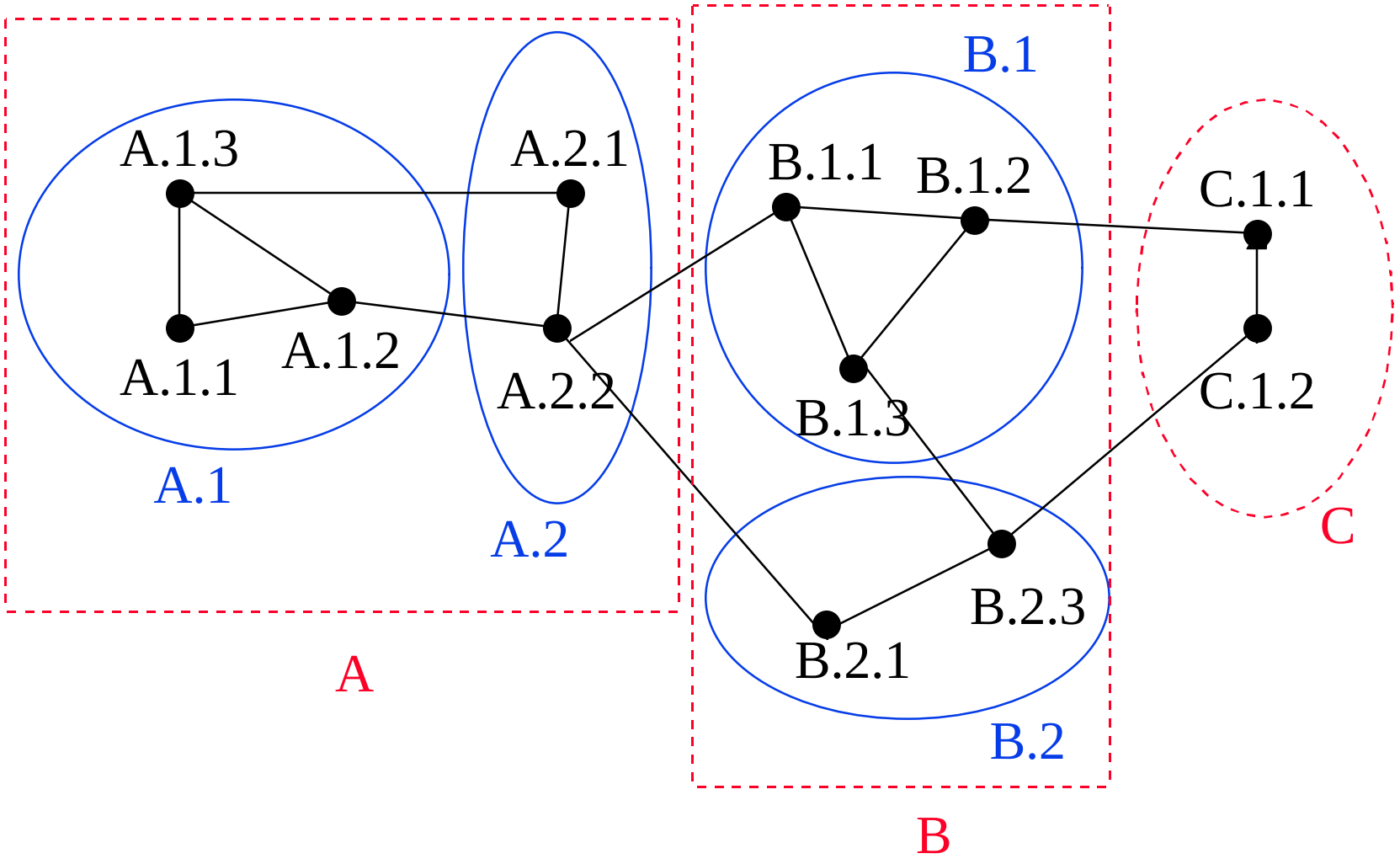
- Each node sends "Hello" packets periodically and on state changes.
- The packet contains state of all its links
- The packet is flooded to all nodes in the network



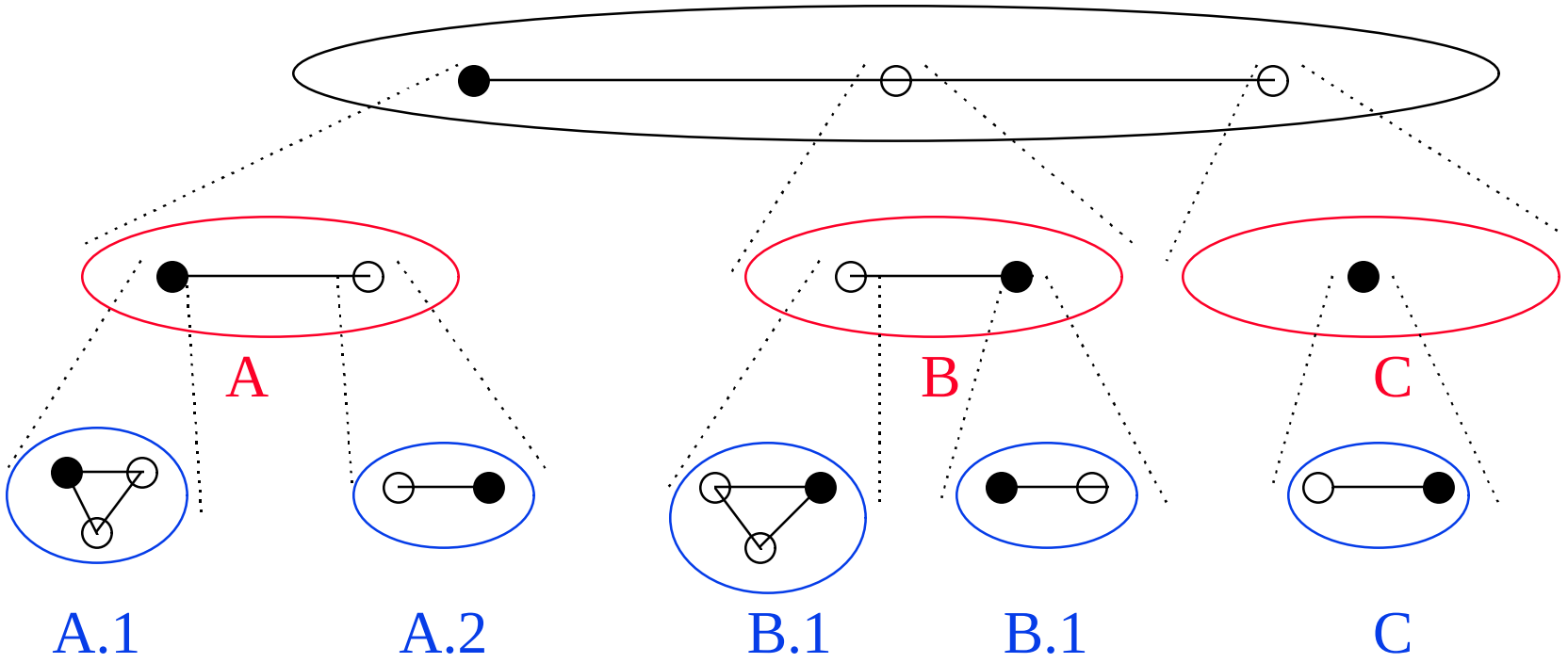
Very Large Networks



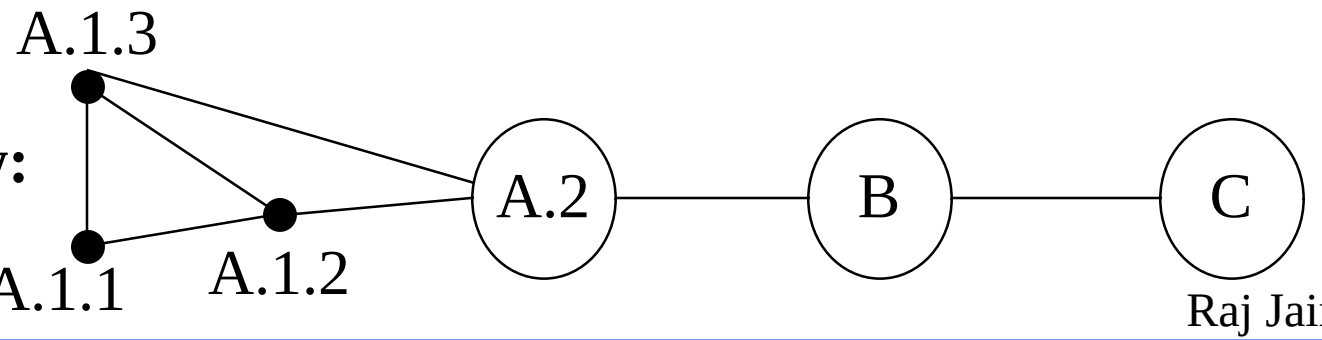
Hierarchical Layers



Hierarchical View



A.1.1's View:



Terminology

- Peer group: A group of nodes at the same hierarchy
- Border node: one link crosses the boundary
- Logical group node: Representation of a group as a single point
- Logical node or Node: A physical node or a logical group node
- Child node: Any node at the next lower hierarchy level
- Parent node: Logical group node at the next higher hierarchy level
- Logical links: links between logical nodes

- 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 ⇒ Leader may change any time.
- PGL acts as a logical group node.
  Uses same ATM address with a different selector value.
- Peer group ID: Address prefixes up to 13 bytes

Topology State Information

- Metric: Added along the path, e.g., delay
- Attribute: Considered individually on each element.
  - Performance, e.g., capacity or
  - Policy related, e.g., security
- State parameter: Either metric or attribute
- Link state parameter. Node state parameter.
- Topology = Link + Nodes
- Topology state parameter: Link or node state parameter
- PNNI Topology state element (PTSE):
  Routing information that is flooded in a peer group
- PNNI Topology state packet (PTSP): Contains one PTSE

Topology State Parameters

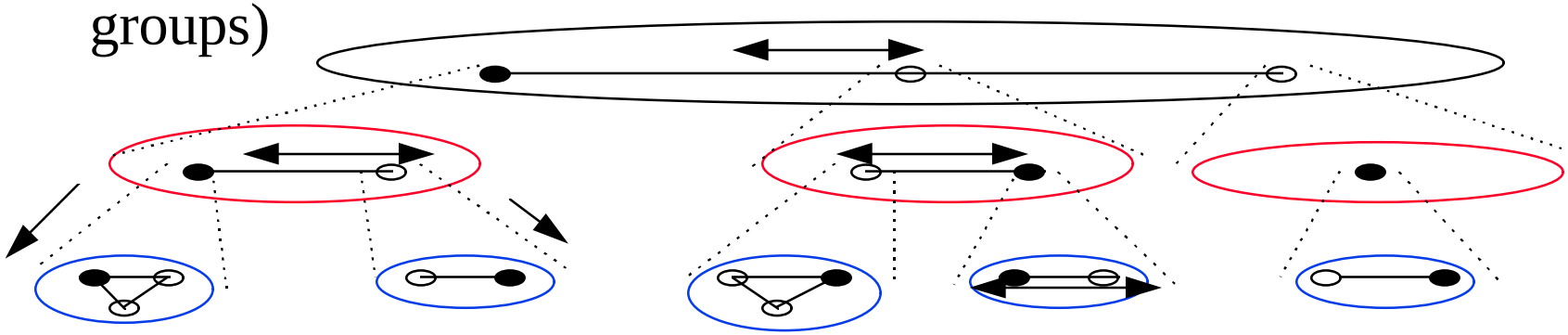
- Metrics:
  - Maximum Cell Transfer Delay (MCTD)
  - Maximum Cell Delay Variation (MCDV)
  - Maximum Cell Loss Ratio (MCLR)
  - Administrative weight
- Attributes:
  - Available cell rate (ACR)
  - Cell rate margin (CRM) = Allocated - Actual
    First order uncertainty. Optional.
  - Variation factor (VF) = CRM/Stdv(Actual)
    Second order uncertainty. Optional.
  - Branching Flag: Can handle point-to-multipoint traffic
  - Restricted Transit Flag: Supports transit traffic or not

Database Synchronization and Flooding

- Upon initialization, nodes exchange PTSE headers (My topology database is dated 11-Sep-1995:11:59)
- Node with older database requests more recent info
- After synchronizing the routing database, they advertise the link between them
- The ad (PTSP) is *flooded* throughout the peer group
- Nodes ack each PTSP to the sending neighbors, update their database (if new) and forward the PTSP to all *other* neighbors
- All PTSEs have a life time and are aged out unless renewed.
- Only the node that originated a PTSE can reissue it.
- PTSEs are issued periodically and also event driven.

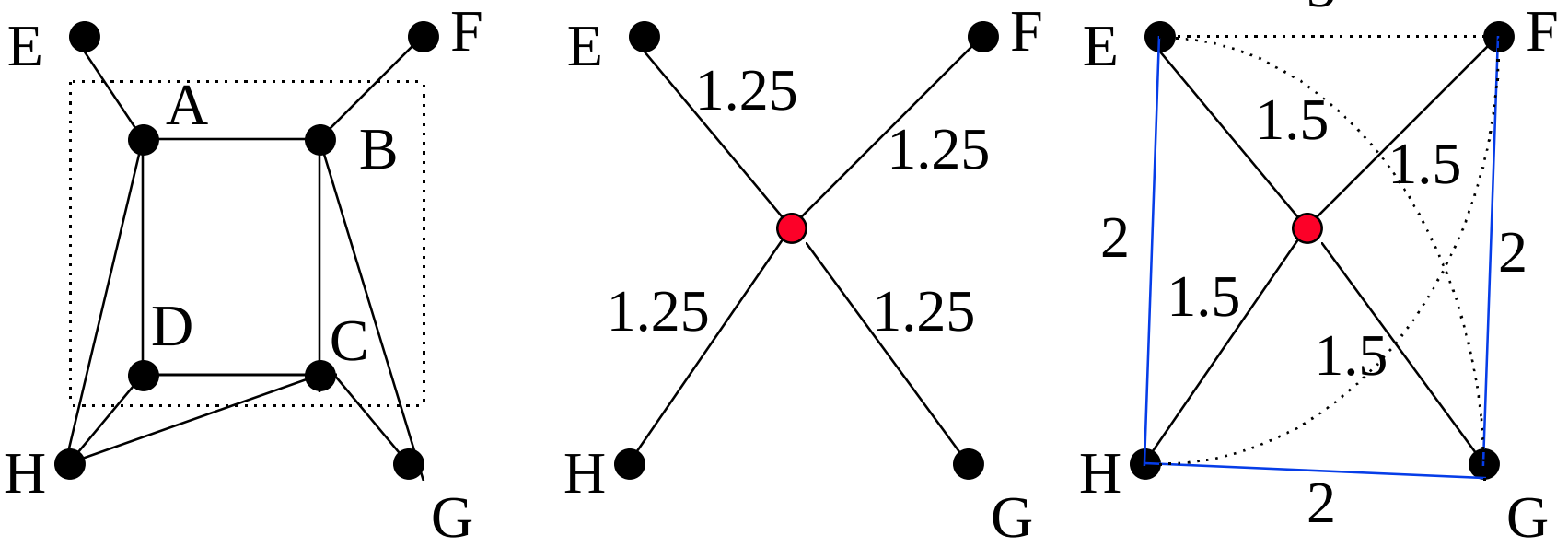
Information Flow in the Hierarchy

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

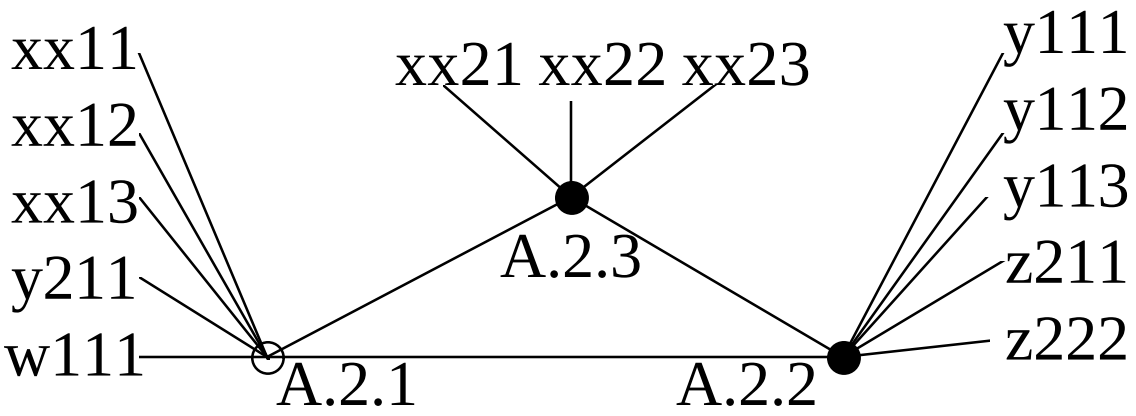


Topology Aggregation

- Get a simple representation of a group
- Alternatives: Symmetric star (n links) or mesh ($n^2/2$ links)
- Compromise: Star with exceptions



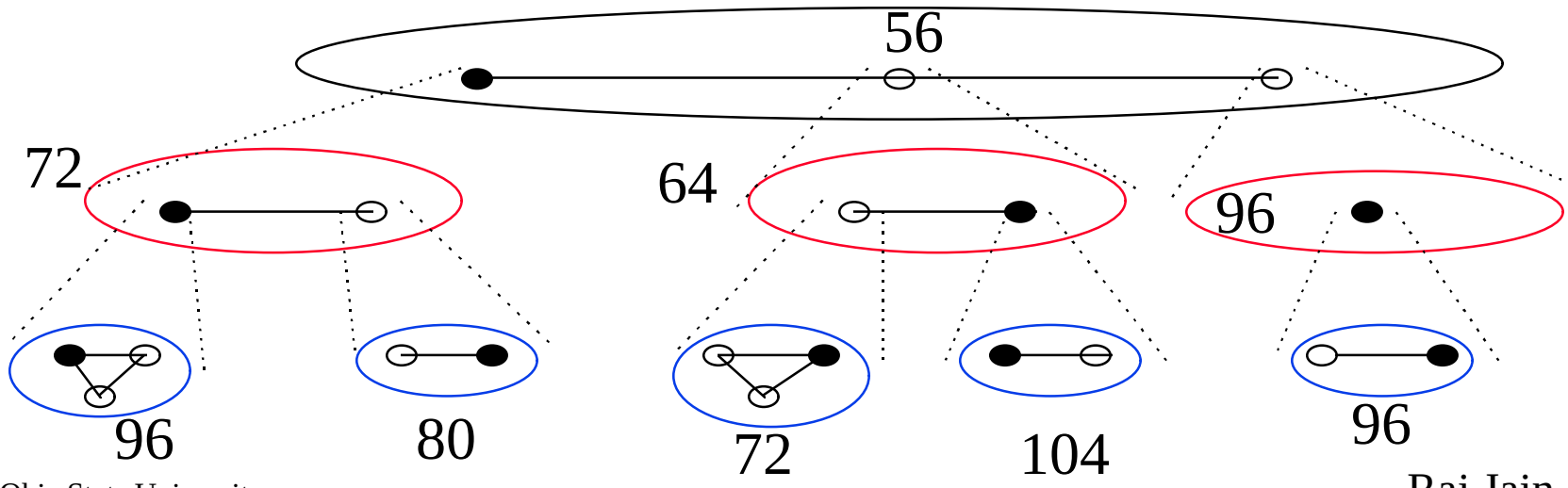
Address Summarization



- Summary = All nodes with prefix xxx, yyy, ...
  + foreign addresses
- Native addresses = All nodes with prefix xxx, yyy, ...
- **Example**:
  - A.2.1 = XX1*, Y2*, W111 A.2.2 = Y1*, Z2*
  - A.2.3 = XX2*
  - A.2 = XX*, Y*, Z2*, W111. W111 is a foreign address

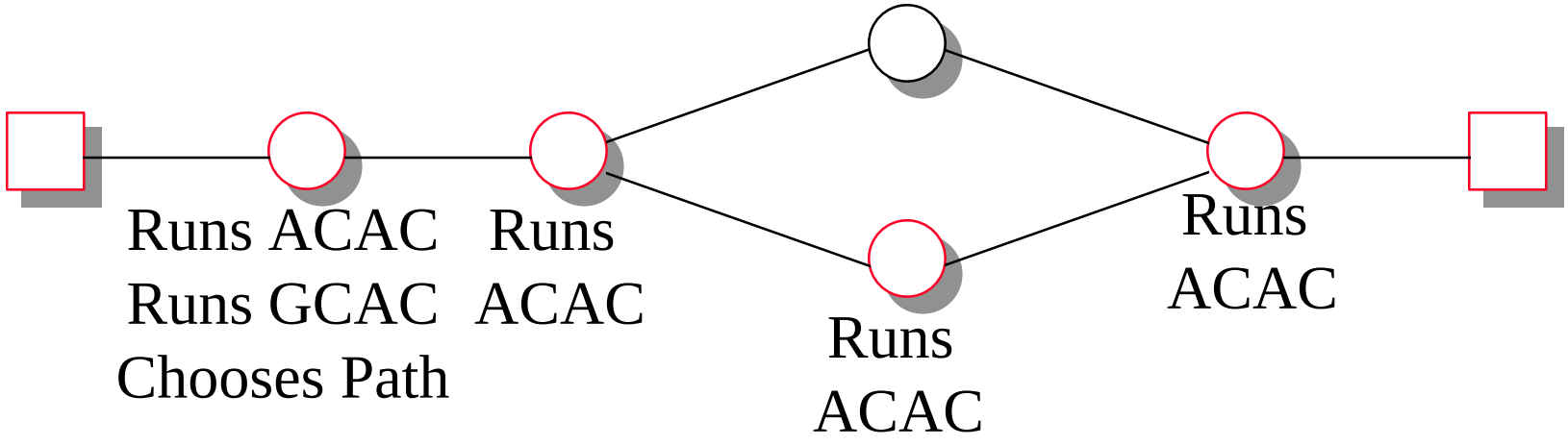
Address Scope

- Upward distribution of an address can be inhibited, if desired.
  E.g., Don't tell the competition B that W111 is reachable via A.
- Each group has a level (length of the shortest prefix).
- Each address has a scope (level up to which it is visible).



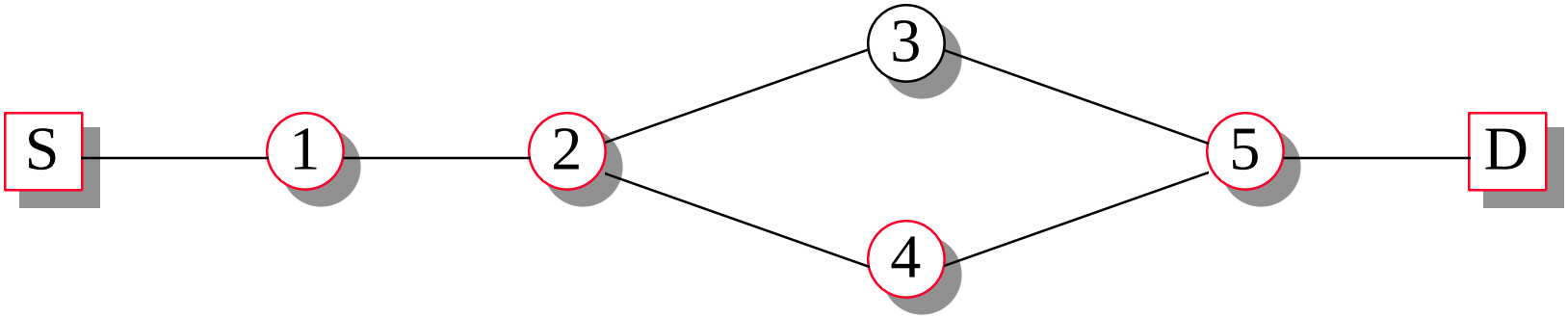
Call Admission Control

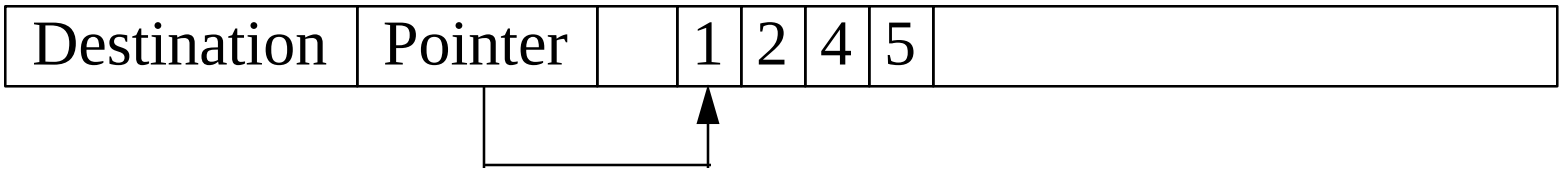
- Generic Call Admission Control (GCAC)
  - Run by a switch in choosing a source route
  - Determines which path can probably support the call
- Actual Call Admission Control (ACAC)
  - Run by each switch
  - Determines if it can support the call



Source Routing

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

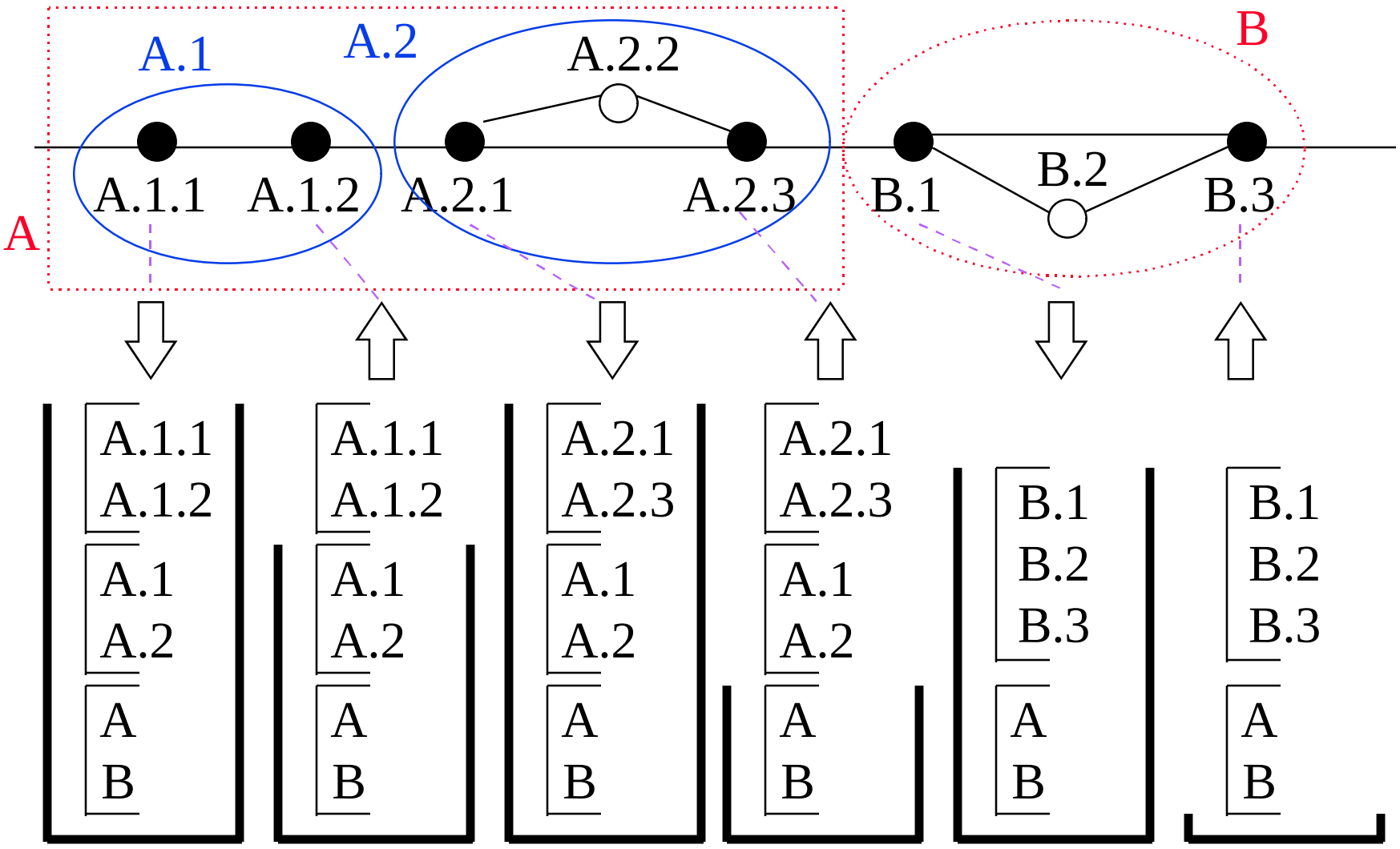




Designated Transit Lists

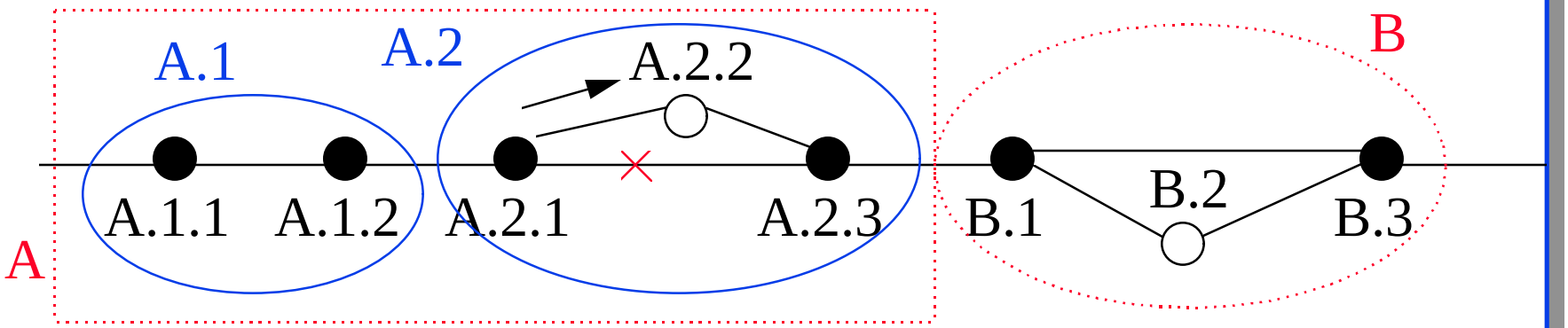
- DTL: Source route across each level of hierarchy
- Entry switch of each peer group specifies complete route through that group
- Entry switch may or may not be the peer group leader
- Multiple levels ⇒ Multiple DTLs
  Implemented as a stack

DTL: Example

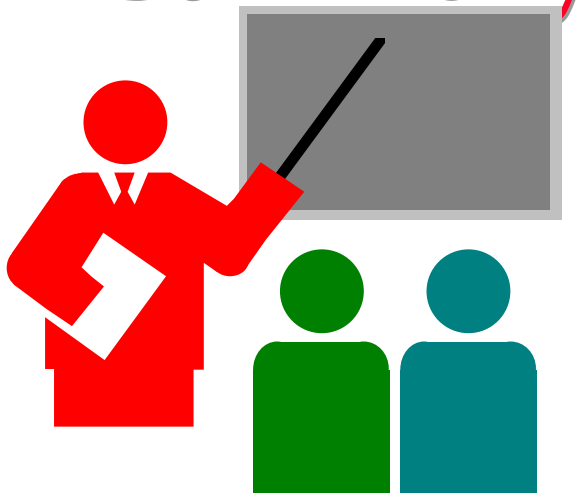

A
A.1
A.2
A.2.2
B
A.1.1
A.1.2
A.2.1
A.2.3
B.1
B.2
B.3
A.1.1
A.1.2
A.1
A.2
A
B
A.1.1
A.1.2
A.1
A.2
A
B
A.2.1
A.2.3
A.1
A.2
A
B
A.2.1
A.2.3
A.1
A.2
A
B
B.1
B.2
B.3
A
B
B.1
B.2
B.3
A
B

Crankback and Alternate Path Routing

- If a call fails along a particular route:
  - It is *cranked back* to the originator of the top DTL
  - The originator finds another route *or*
  - Cranks back to the generator of the higher level source route



Summary



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

Abbreviations

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

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

- PNNI Private network-network interface
- PVCC Permanent virtual channel connection
- PVPC Permanent virtual path connection
- RD Routing domain
- SAAL Signaling ATM adaptation layer
- SNPA Subnetwork point of attachment
- TIG Topology information group
- TLV Type, length, value
- VF Variance factor

References

- B. Dorling, D. Freedman, C. Metz, and J. Burger, "Internetworking over ATM: An Introduction," Prentice Hall, 1996.
- 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
- A. Alles, “Routing and Internetworking in ATM Networks,” Networld+InterOP, March 1995
- W.C. Lee, “Topology Aggregation for Hierarchical Routing in ATM Networks,” Computer Communication Review, April 1995.

References (cont)

- J.M. Halpern, "ATM Call Routing," Communication Systems Design, pp. 30-35.
- R. Callon, et al, "Issues and Approaches for Integrated PNNI," ATM Forum 96-0355, April 1996.