

Asynchronous Transfer Mode ATM - Part 2

Dr. Andreas Steffen

©1999-2001 Zürcher Hochschule Winterthur

A. Steffen, 18.11.2001, KSy_ATM_2.ppt 1

ABR Rate-Based Congestion Control

- Congestion control loop using resource management cells
- The use of the payload type field in the ATM header

ATM Addressing and Routing Schemes

- Public ATM network addresses
- Private ATM end system addresses
- Hierarchical ATM address space and prefix-based routing protocols

LAN Emulation (LANE)

- Methods of ATM interworking
- LAN emulation clients
- LAN emulation service components
- Control and data connections

Multiprotocol over ATM (MPOA)

- IP layer 3 routing versus ATM layer 2 switching
- Edge devices and MPOA servers
- Virtual router concept
- ATM shortcuts
- Coexistence of LANE and MPOA

Multiprotocol Label Switching (MPLS)

- Label switched routers and label switched paths
- The label stack, virtual private networks, tunneling
- The shim header

- **ABR Rate-Based Congestion Control**
 - „Computer Networks“, Chapter 5.6.7 Rate-Based Congestion Control, pp. 469-471, Andrew S. Tanenbaum, Prentice Hall
- **ATM Addressing and Routing Schemes**
 - „ATM Internetworking“, Chapter 3.0 ATM Signaling and Addressing, pp. 6-10, Anthony Alles, Cisco Systems, Inc.
- **LAN Emulation (LANE)**
 - „ATM Internetworking“, Chapter 5.0 LAN Emulation, pp. 24-32, Anthony Alles, Cisco Systems, Inc.
- **Multiprotocol over ATM (MPOA)**
 - „MPOA - Multiprotocol over ATM“, pp. 1-29, Applied Technology Group's Communications & Networking Technology Guide Series, www.techguide.com
- **Multiprotocol Label Switching (MPLS)**
 - RFC 3031 – „Multiprotocol Label Switching Architecture“
 - RFC 3032 – „MPLS Label Stack Encoding“



**Available Bit Rate (ABR)
Service Class
Rate-Based Congestion Control**

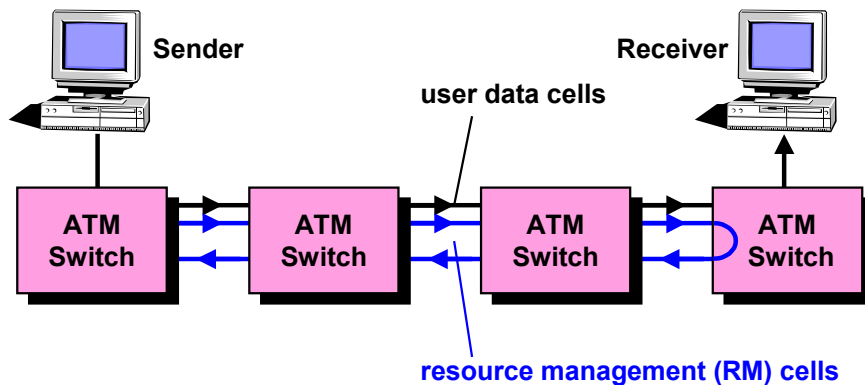
Parameter	CBR	rt-VBR	nrt-VBR	ABR	UBR
CLR	✓	✓	✓	✓	—
CTD	max	max	mean	—	—
CDV	✓	✓	—	—	—
PCR & CDVT	✓	✓	✓	✓	✓
SCR & MBS	—	✓	✓	—	—
MCR	—	—	—	✓	—
Congestion Control	—	—	—	✓	—

Features of the Available Bit Rate (ABR) Service Class

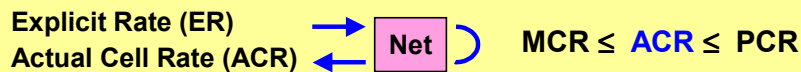
The ABR service class has two special features, that differentiates it from the other service classes:

- a minimum cell rate (MCR) can be specified, which the network is trying to guarantee.
- an active control loop prevents congestion in the network while at same time maximizing the bit rate available to the individual user.

Available Bit Rate (ABR) Service Class Rate-Based Congestion Control



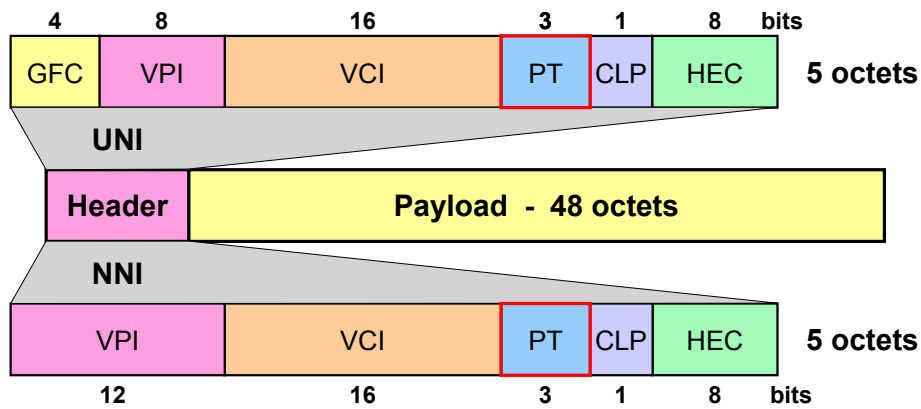
Congestion Control Loop



Congestion Control Loop

- ABR congestion control is based on the idea that each sender has a current rate, the **Actual Cell Rate (ACR)**, that falls between MCR and PCR. When congestion occurs, ACR is reduced (but not below MCR). When congestion is absent, ACR is increased (but not above PCR).
- After every k data cells, the sender inserts a special **Resource Management (RM)** cell containing the rate at which the sender currently would like to send. This value is called **Explicit Rate (ER)** and must fulfill the condition $MCR \leq ER \leq PCR$.
- As the RM cell passes through the various switches on the way to the receiver (travelling along the same path as the data cells), each switch experiencing congestion can lower the ER value, but not below the minimum cell rate (MCR).
- At the receiver side, the RM cell is looped back to the sender. When the sender gets the RM cell back, it can then see what the minimum acceptable rate is according to all the switches along the path and can then adjust its Actual Cell Rate (ACR), to bring it into line with what the slowest switch can handle.
- The rate-based congestion control algorithm is quite robust and fully reliable, since even a lost RM cell will be noticed by the sender when it fails to return within the expected time interval. The sender then can take action e.g. by lowering its ACR.

ATM Cell Header - Payload Type (PT) field



UNI User-Network Interface
GFC Generic Flow Control
VPI Virtual Path Identifier
VCI Virtual Channel Identifier

NNI Network-Network Interface
PT Payload Type
CLP Cell Loss Priority
HEC Header Error Control

The Payload Type (PT) Field

- The payload type field in the ATM cell header consists of three bits. Besides designating various payload types these bits have found a number of additional uses that will be explained in detail in the following slide.

ATM Cell Payload Type Field

PT	Meaning
000	
001	1. Bit: User data cell
010	2. Bit: Explicit Forward Congestion Control (EFCI)
011	3. Bit: AAL-indicate (last AAL5 cell)
100	Operation and Management (OAM) cell: Segment
101	Operation and Management (OAM) cell: End-to-End
110	Resource Management (RM) cell: ABR Congestion Control
111	Reserved for future use

User Data Cells

- A user data cell has its first PT bit set to '0'.
- The second PT bit is used for Explicit Forward Congestion Control (EFCI). The sender initially sets this bit to '0'. Any switch on the way experiencing congestion can set this bit to '1'. The receiver can then send a congestion warning embedded in a RM cell back to the sender.
- In violation of the strict separation of hierarchical communication layers, the third PT bit is used to indicate the last cell belonging to a next higher ATM Adaptation Layer 5 Protocol Data Unit (AAL 5 PDU).

Operation and Management (OAM) Cells

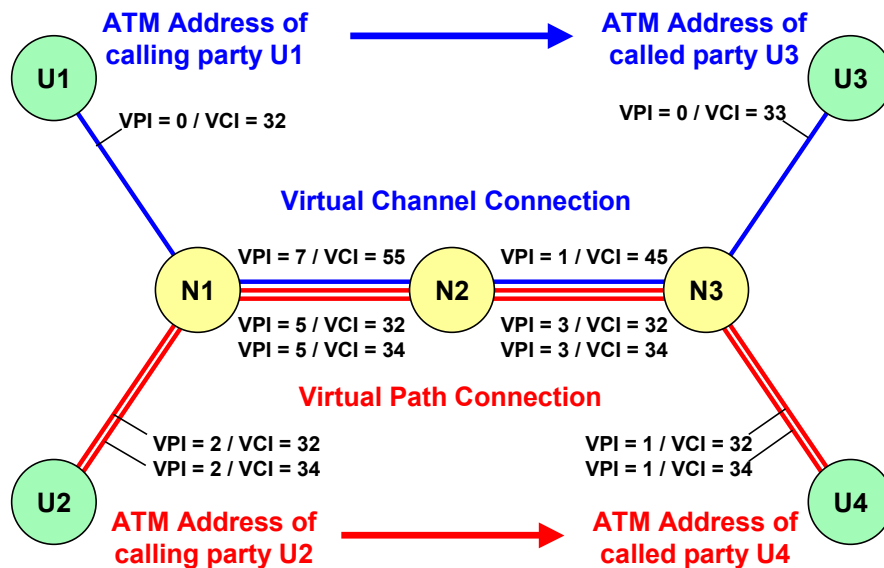
- An OAM cell has its first two PT bits set to '10'.
- If the third PT bit is set to '0', then the destination of the OAM cell is just the next hop in the virtual connection, whereas with the bit set to '1' the OAM cell will travel end-to-end.
- OAM cells are used for various management purposes. For details see ITU-T I.610.

Resource Management (RM) cells

- A RM cell has its PT bits set to '110'.
- RM cells are used to signal congestion in a network. Any switch experiencing congestion can generate a RM cell and send it back to the sender of a virtual connection, with the possibility of course that the RM cell gets dropped on the way and will never arrive at its destination. As we have seen, only the rate-based congestion control mechanism employed by the ABR service class uses RM cells in a fully reliable way.
- For details on the use of RM cells see ITU-T I.371.

Source: ITU-T Recommendation I.361 (02/99) „B-ISDN ATM layer specification“

**ATM Addressing and
Routing Schemes**



Setup of Switched Virtual Connections based on ATM End-System Addresses

- In order to automatically set up a Switched Virtual Connection (SVC), each ATM end system must be uniquely identified by an ATM address. In a similar way as in an ISDN Q.931 call setup used by the Public Switched Telephone System (PSTN), the addresses of both the calling and called parties are contained in the B-ISDN Q.2931 connection setup messages.

Public ATM Addresses

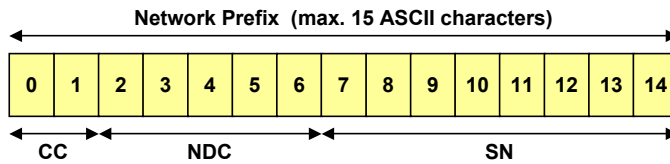
- These are defined by the ITU-T and are used in public ATM networks

Private ATM Addresses

- These are defined by the ATM Forum and are used in private ATM networks

ITU-T Recommendation E.164

International Public Telecommunication Numbering Plan



CC Country Code (1-3 digits)

NDC National Destination Code

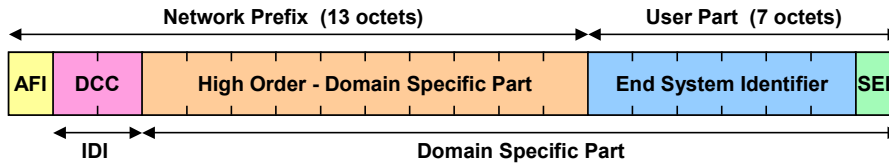
SN Subscriber Number

ATM Public Network Addresses

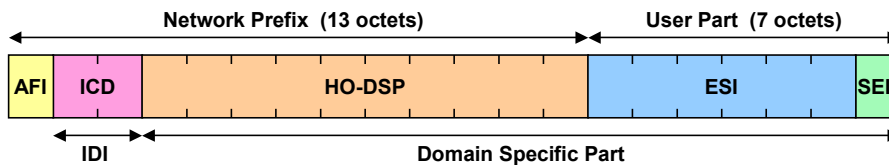
- ATM Public Network Addresses have the same format as normal public telephone numbers that comply with the ITU-T E.164 International Public Telecommunication Numbering Plan.

Source: ITU-T Recommendation E.164 (05/97) „International Public Telecommunication Numbering Plan“

DCC AESA Format (AFI = 0x39, DCC for USA: 0x840F)



ICD AESA Format (AFI = 0x47, ICD by British Standards Institute)



AFI Authority and Format Indicator

IDI Initial Domain Identifier

SEL Selector

DCC Data Country Code

ICD International Code Designator

A. Steffen, 18.11.2001, KSy_ATM_2.ppt 11

Private ATM Addresses

- ATM End System Addresses (AESA) are defined by the ATM Forum to have a constant length of 20 octets. They come in many types and they differ mostly in what authority assigns them. Four variants appear in ATM Forum specifications.

AFI - Authority and Format Indicator

- The first octet in an AESA determines the type and the format of the address.

DCC - Data Country Code (AFI = 0x39)

- In this case the IDI is a unique Data Country Code (DCC); these identify particular countries, as specified in ISO 3166. Each country is free to decide the structure and the rules for assignment of the **Domain Specific Part (DSP)**.

ICD - International Code Designator (AFI = 0x47)

- In this case, the IDI is an International Code Designator (ICD); these are allocated by the ISO 6523 registration authorities (the British Standards Institute). ICD codes identify particular international organizations.

HO-DSP - High Order Domain Specific Part

- Used to support flexible, multi-level addressing hierarchies for prefix-based routing protocols (similar to hierarchical IP subnet addresses).

ESI - End System Identifier

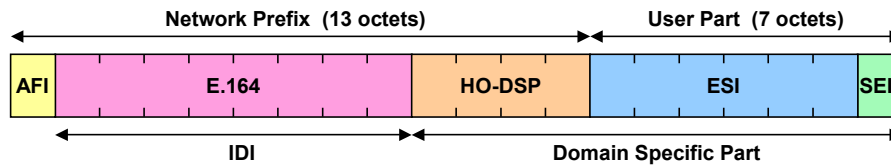
- The ESI field is specified to be a 48-bit MAC address, as administered by the IEEE. This facilitates the support of LAN equipment, which is typically hardwired with such addresses.

SEL - Selector

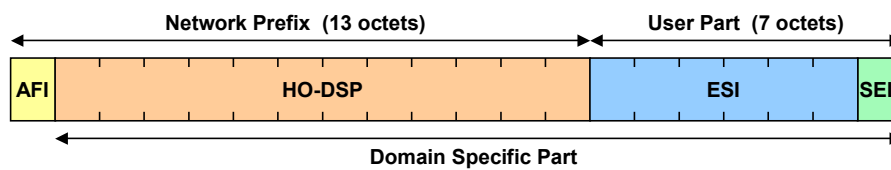
- Used for local multiplexing within end-stations. Has no network significance

Source: ATM Forum Addressing: User Guide Version 1.0. AF-RA-0105.000, Jan 1999

E.164 AESA Format (AFI = 0x45)



Local AESA Format (AFI = 0x49)



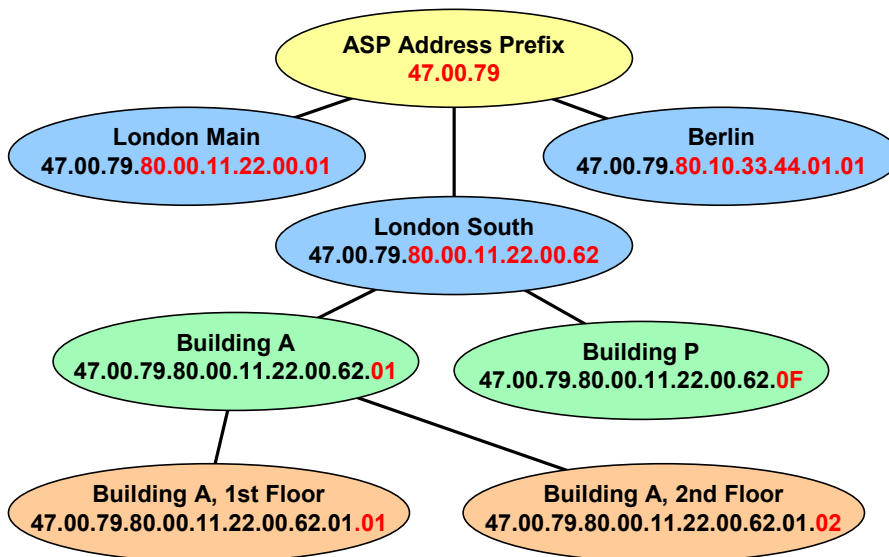
E.164 AESA Format (AFI = 0x46)

- Used to embed a native E.164 number. Usually the HO-DSP, ESI and SEL fields are set to zero.

Local AESA Format (AFI = 0x49)

- No Initial Domain Identifier (IDI) field. Can be freely structured and used by anyone in a private network (similar to 10.x.x.x private network IP addresses).

Hierarchical WAN Enterprise Network using ATM Service Provider (ASP) Addresses



Hierarchical Address Space

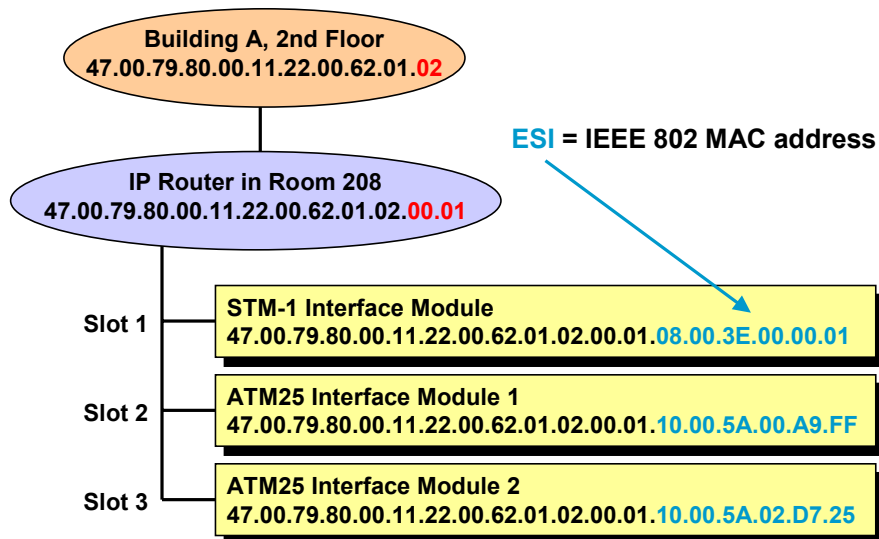
- Private ATM addresses are usually organized hierarchically. This makes it possible to use prefix-based routing algorithms in order to set up ATM network connections
- In our example an ATM service provider has a three octet hexadecimal ICD Address Prefix of '47.00.79'.
- A corporate customer with three major sites: London Main, London South, and Berlin gets three address prefixes from its ASP, each a length of 9 octets, which are
'47.00.79.80.00.11.22.00.01' for London Main
'47.00.79.80.00.11.22.00.62' for London South
'47.00.79.80.10.33.44.01.01' for Berlin
- The company is now free to allocate the remaining 4 octets of the HO-DSP according to its own internal numbering plan. It chooses one octet to designate the building at a given site and the next octet the floor of the building. The last two octets of the HO-DSP uniquely identify the ATM end system on that floor.

Prefix-Based Routing Protocols

- Each ATM switch maintains a routing table that is continuously updated using any of the various ATM routing protocols. When a Q.2931 connection setup message arrives, the switch matches the called party's ATM number against all the entries (prefixes) in its routing table. The entry with the **longest prefix match** wins and the signalling message is forwarded to the ATM switch designated by that prefix.

ATM End System Addresses III

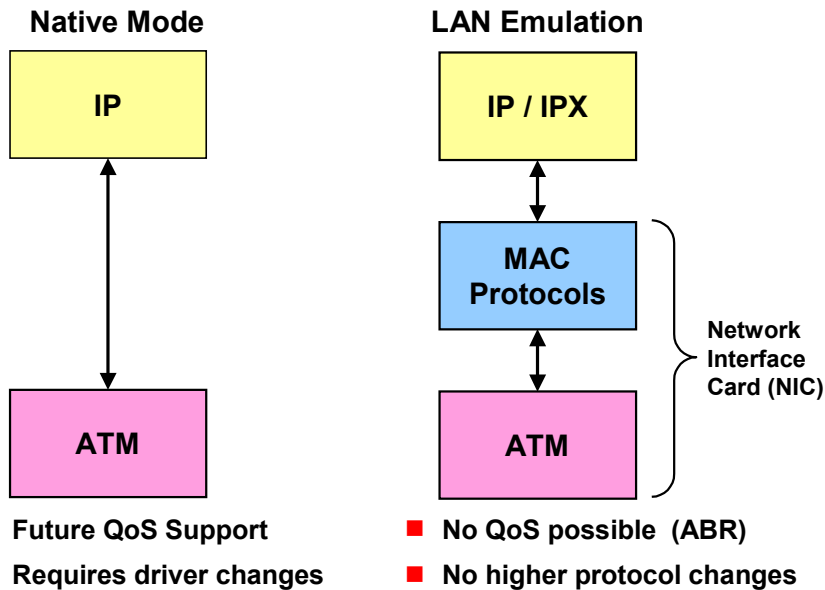
ATM Interface Modules



ATM Address of Interface Module

- In our example the ATM end system in room 208 on the 2nd floor of building A in London South is an IP router equipped with three ATM interface modules. Each physical port module is uniquely identified by its End System Identifier (ESI) that is nothing more than a 48-bit IEEE 802 MAC address stored in non-volatile memory on the interface card.
- If for example a virtual connection from Berlin terminates on the STM-1 interface module, then the full ATM end system address in London South is
'47.00.79.80.00.11.22.00.62.01.02.99.01.08.00.3E.00.00.01'
with a total length of 19 octets.

LAN Emulation (LANE)



A. Steffen, 18.11.2001, KSy_ATM_2.ppt 16

IP-based Traffic

- Most of today's data traffic is IP-based. Therefore we have to find ways to carry IP network layer packets across an ATM network.

Native Mode

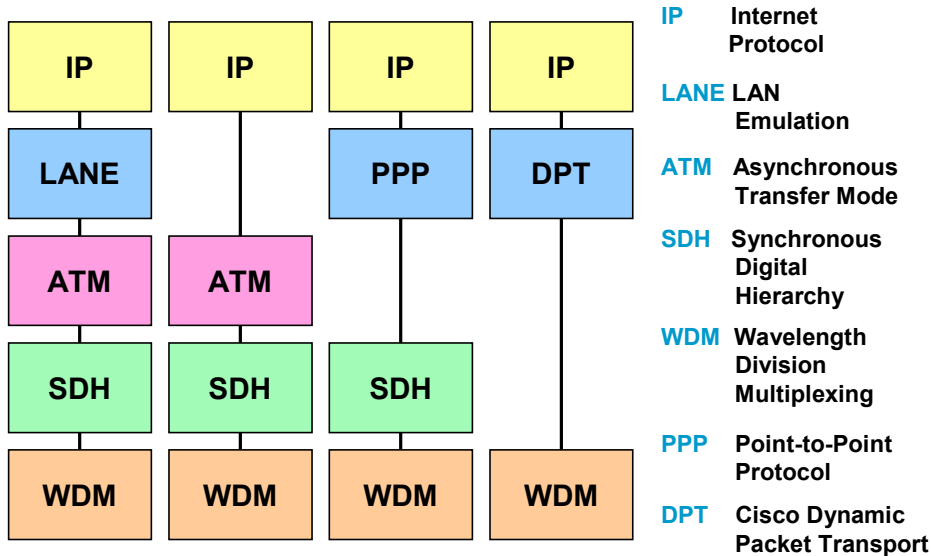
- IP addresses are mapped directly into ATM addresses and IP datagrams are carried in AAL 5 PDUs.
- **Pro:** Direct Mapping of IP onto ATM makes ATM's QoS features available for IP.
- **Contra:** Special drivers are required to support ATM network interfaces.

LAN Emulation

- By inserting an IEEE 802 MAC Layer between the ATM network interface and the IP stack, an Ethernet network interface is emulated. The ATM layer is hidden.
- **Pro:** No changes are required in the IP stack of an ATM enabled host.
- **Contra:** Since the intermediate MAC layer hides the ATM layer, ATM's QoS features cannot be used transparently. IP packets are transported using the ABR or UBR service classes.

Source: Anthony Alles, „ATM Internetworking“, Chapter 5.0 LAN Emulation, p. 24, May 1995, Cisco Systems, Inc.

IP over ATM over SDH over WDM



A. Steffen, 18.11.2001, KSy_ATM_2.ppt 17

IP over LANE over ATM over SDH over WDM

- Used for LAN emulation, not very efficient, no QoS support, no changes required.
- SDH layer monitors quality of physical link and allows fast protection switching.

IP over ATM over SDH over WDM

- Used by carriers in their backbone networks. Presently the only solution that offers true QoS support.
- SDH layer monitors quality of physical link and allows fast protection switching.

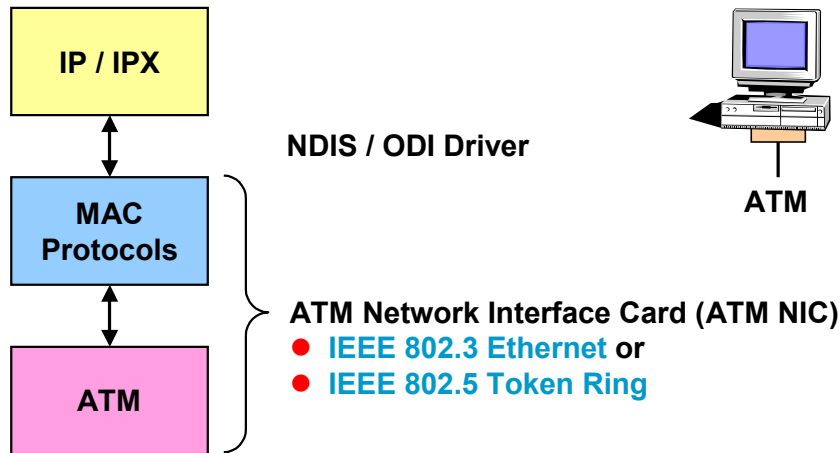
IP over PPP over SDH over WDM

- Used by many ISPs to connect POPs to the IP backbone. Eliminates ATM layer, therefore efficient use of available bandwidth. Presently no true QoS support.
- SDH layer monitors quality of physical link and allows fast protection switching.

IP over DPT over WDM

- Proprietary Cisco protocol. Eliminates ATM layer, therefore efficient use of available bandwidth. Uses SDH framing, but no SDH Automatic Protection switching (APS). Cisco implements proprietary Intelligent Protection Switching (IPS) instead.
- Classification into high and low priority IP packets possible, no true QoS support.

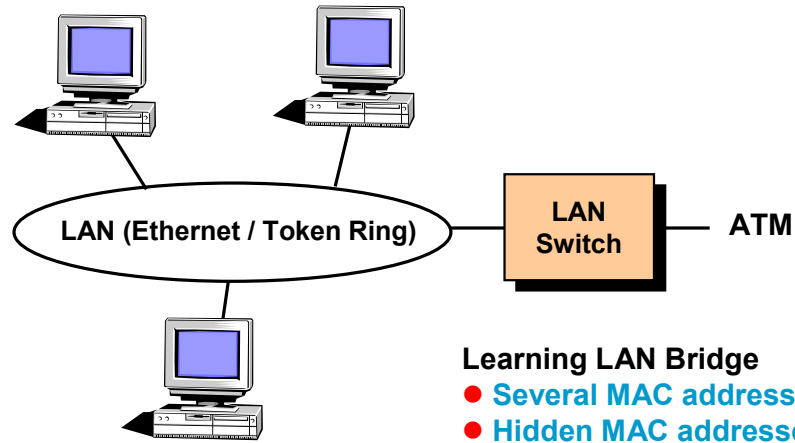
LAN Emulation Client (LEC) I ATM Network Interface Card



ATM Network Interface Card

- Under LAN emulation an ATM network interface card behaves like an IEEE 802.3 Ethernet or IEEE 802.5 Token Ring interface.
- Normal NDIS / ODI drivers can be used to interface the standard IP / IPX stack of a host to an ATM network interface card.

LAN Emulation Client (LEC) II Layer 2 LAN Switch

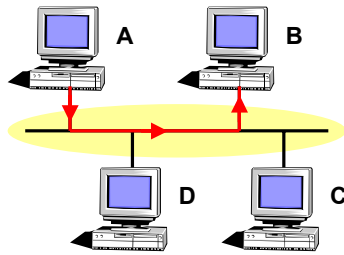


Layer 2 LAN Switch

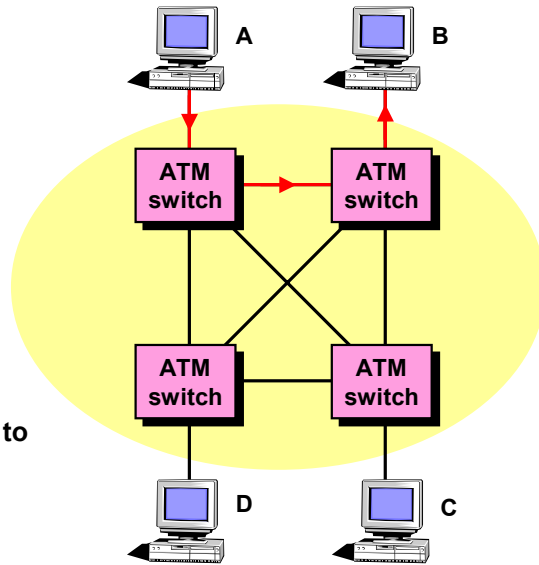
- In this constellation the LAN Emulation Client (LEC) is a Layer 2 LAN Switch with an Ethernet or Token Ring interface on one side and an ATM interface on the other side.
- The LAN switch works as a LAN bridge that learns the MAC addresses on both sides of the bridge.

LAN Emulation

Unicast from client A to known client B

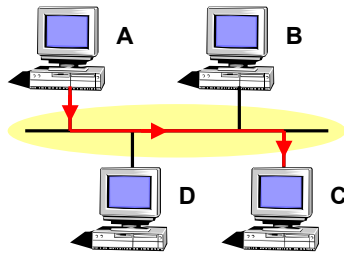


- Client A must know ATM address of client B
- With SVCs, client A must wait till direct connection to client B is set up
- Direct connection has a **high throughput !!**

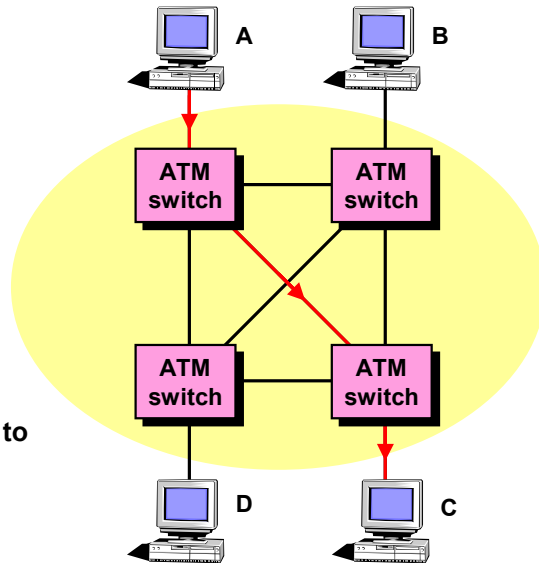


LAN Emulation

Unicast from client A to known client C

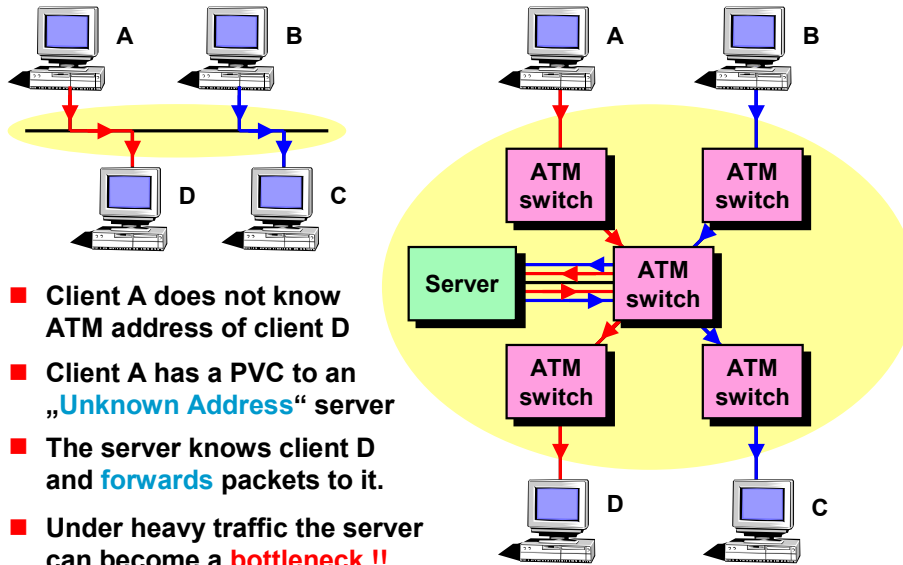


- Client A must know ATM address of client C
- With SVCs, client A must wait till direct connection to client C is set up
- Direct connection has a **high throughput !!**



LAN Emulation

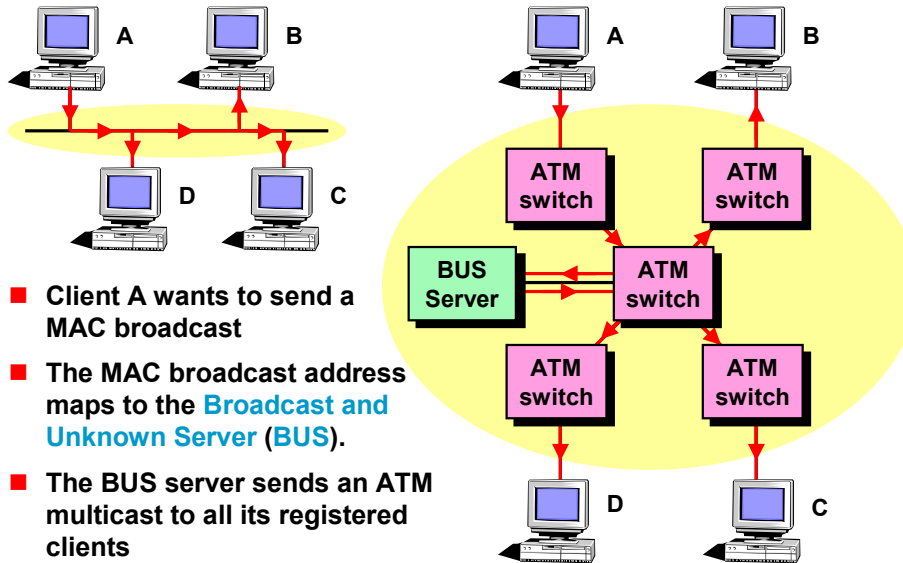
Unicast from client A to unknown client D



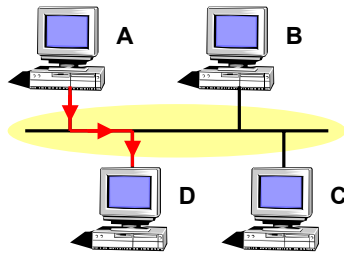
- Client A does not know ATM address of client D
- Client A has a PVC to an „Unknown Address“ server
- The server knows client D and forwards packets to it.
- Under heavy traffic the server can become a **bottleneck !!**

LAN Emulation

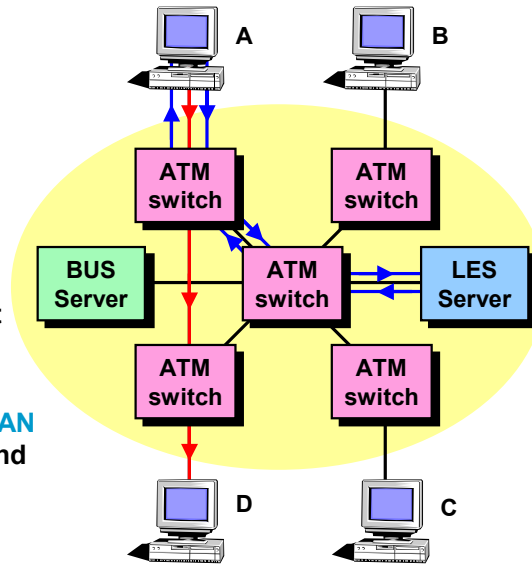
MAC Broadcast from client A

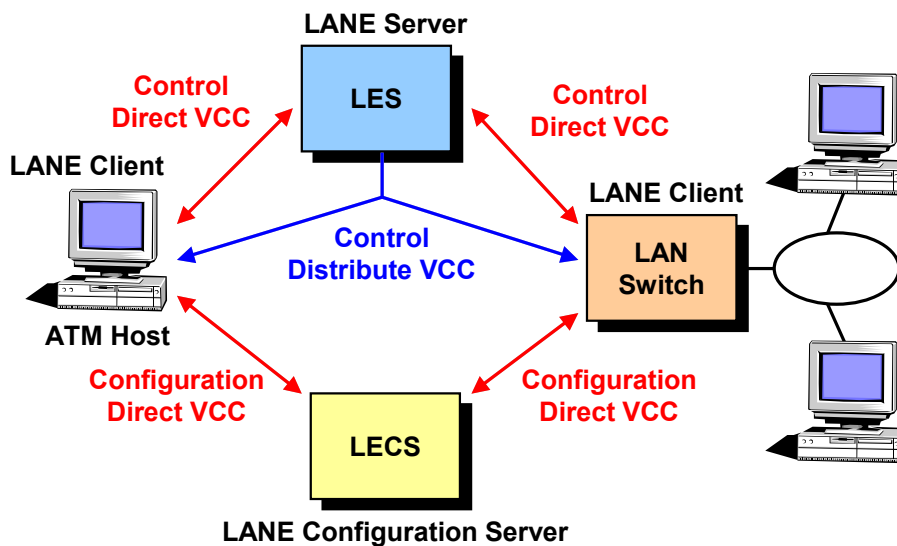


LAN Emulation ATM Address Resolution (LE_ARP)



- Client A wants to connect **directly** to client D
- Client A sends a **LE_ARP** request to his assigned **LAN Emulation Server (LES)** and gets ATM address of D
- Client A opens a direct connection to client D





A. Steffen, 18.11.2001, KSy_ATM_2.ppt 25

Configuration Direct Virtual Channel Connection to the LECS

- Upon initialization (such as power up) the LANE Client (LEC) is looking for the LANE Configuration Server (LECS) by either using a well-known ATM address or the well-known permanent connection (VPI=0, VCI=17).
- After finding the location of the LECS, the LEC sets up a bidirectional unicast **Configuration Direct VCC** to the LECS. The LEC then gets the information it requires to connect into the Emulated LAN (ELAN) it belongs to. This includes the ATM address of the LAN Emulation Server (LES) responsible for this particular ELAN.

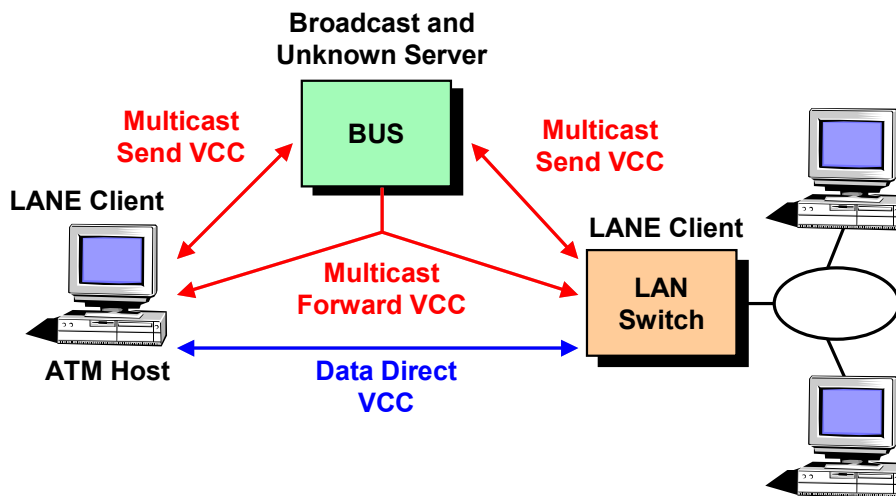
Control Direct Virtual Channel Connection to the LES

- Now the LEC sets up a bidirectional unicast **Control Direct VCC** to the LES and registers its own MAC and ATM addresses with the LES. In the case of a LAN switch operating as a bridge it may optionally also register any other learned MAC addresses.

Control Distribute Virtual Channel Connection from the LES

- The LES then sets up back to the LEC a branch of a unidirectional multicast **Control Distribute VCC** that can be used together with the Control Direct VCC for the LAN Emulation ARP (LE_ARP) procedure for requesting the ATM address that corresponds to a certain MAC address.
- When a LEC wants to resolve a MAC address, it sends an LE_ARP request to the LES over the Control Direct VCC. If the LES knows the ATM address, it immediately responds over the Control Direct VCC. Otherwise the LES sends a multicast over the Control Distribute VCC first, asking all connected LECs for the requested MAC address.

Source: Anthony Alles, „ATM Internetworking“, Chapter 5.2 LANE Operation, pp. 27-29, May 1995, Cisco Systems, Inc.

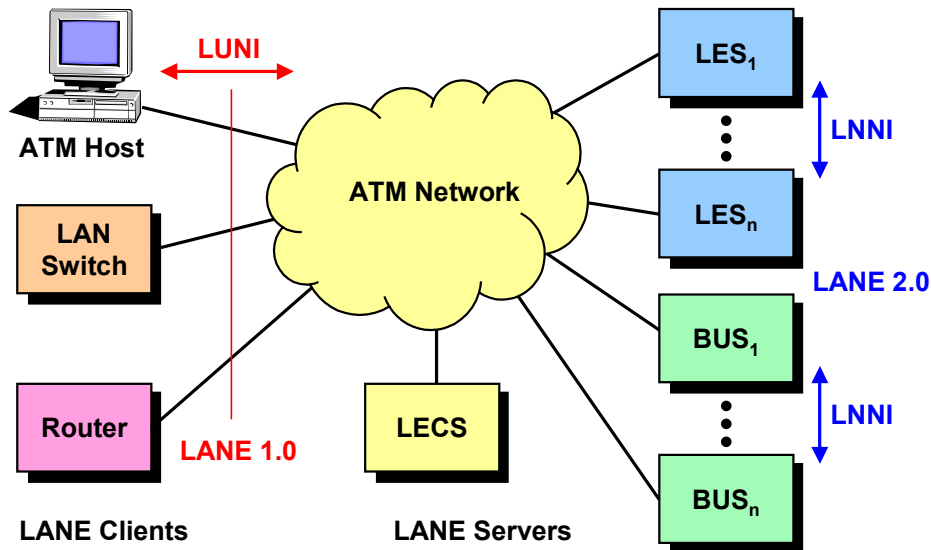


Broadcast and Unknown Server (BUS)

- The MAC broadcast address in an ELAN resolves to the ATM address of the Broadcast and Unknown Server (BUS). After registration with the LES, a LEC sets up a bidirectional point-to-point **Multicast Send VCC** to the BUS, which in turn adds the LEC as an additional leaf in its point-to-multipoint unidirectional **Multicast Forward VCC**.

Data Transfer

- When a LEC wants to send one or several layer 2 MAC packets but does not have the ATM address of the destination LEC, then it first sends an **LE_ARP** request to the LES.
- While waiting for a response from the **LE_ARP**, the LEC also forwards the packets to the BUS, which will in turn flood the packets to all LECs connected to the Multicast Forward VCC.
- When at last an **LE_ARP** response arrives from the LES, the LEC then sets up a **Data Direct VCC** to the destination node and uses this for data transfer rather than the BUS path.



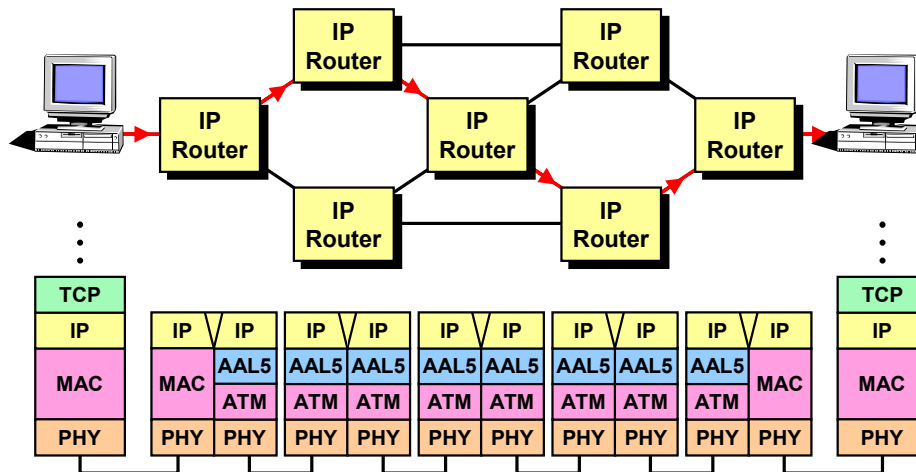
LAN Emulation User-Network Interface (LUNI)

- The LAN Emulation Specification version 1.0 (LANE 1.0) defined by the ATM Forum describes the communication between the LECs and the service components LECS, LES and BUS. This interface is called the **LAN Emulation User-Network Interface (LUNI)**.

LAN Emulation Network-Network Interface (LNNI)

- Version 2.0 of the LAN Emulation Specification (LANE 2.0) defines the communication (LECS-LES, LECS-BUS, LES-LES, BUS-BUS) among distributed service components that are replicated in order to support failure recovery and load balancing.

Multiprotocol over ATM (MPOA)



IP Routing

- IP is a connection-less layer 3 network protocol. Every IP packet arriving at a router is forwarded on an individual basis by consulting a routing table that is being constantly updated by various dynamic routing protocols.
- Looking up each IP header is a costly operation. Much more efficient would be to determine the optimal path through the network at the start of a new connection or **stream** and then to switch all packets along this same path without consulting the routing table on a packet-by-packet basis.

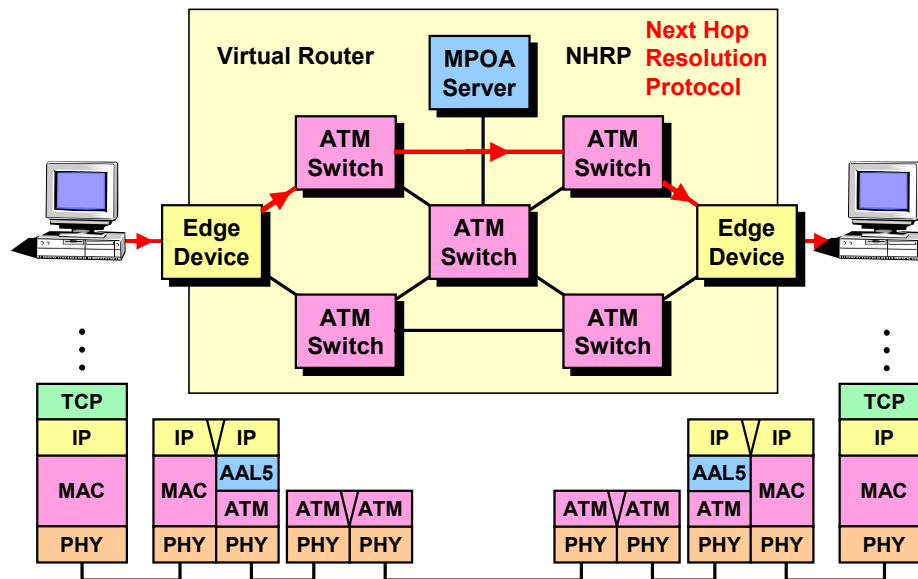
Traffic Engineering is difficult to implement !

- Many IP routing algorithms base their metrics on the smallest number of hops to the end-point. Therefore all packets take this route even if it becomes heavily congested. It is very difficult to apply traffic engineering by forcing IP packets onto alternative routes that still have spare capacity.

Quality of Service is still missing !

- Although a couple of QoS schemes have been proposed for IP-based traffic, neither **IntServ** based on **RSVP**, nor **DiffServ** are widely deployed in the internet world yet. IP-based communication is still mostly „best effort“!

Multiprotocol over ATM (MPOA)



A. Steffen, 18.11.2001, KSy_ATM_2.ppt 30

IP Streams based on Layer 2 Switching

- In many applications an IP connection is not constrained to the transmission of a single IP packet but quite a large number of packets follow the first one. This is called an „**IP stream**“. With MPOA it becomes possible to set up an optimal route based on ATM switched virtual connections, so that all packets take the same path based on fast and efficient layer 2 switching.

Edge Devices

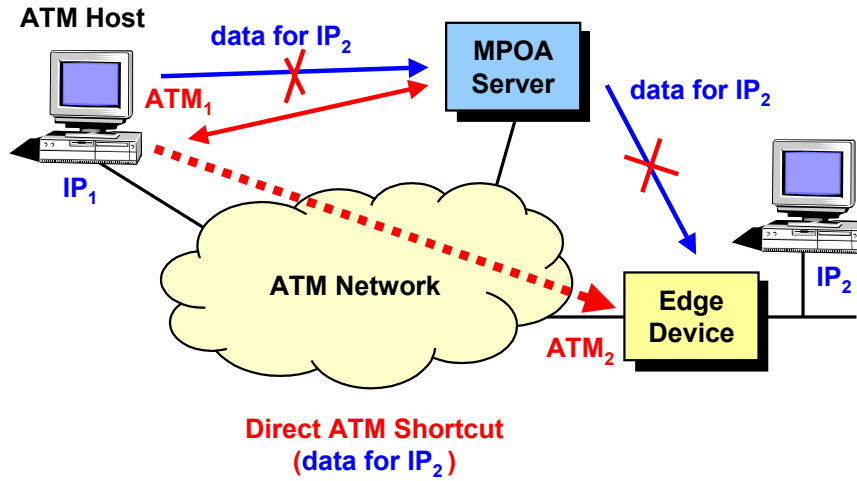
- In MPOA based networks, an IP packet enters the network at an **ingress edge router** and leaves it again at an **egress edge router**. When the first packet arrives, the ingress edge router contacts the MPOA server situated somewhere in the IP network and asks for an optimal route for the incoming packet stream, given the current traffic situation in the network.

MPOA Server

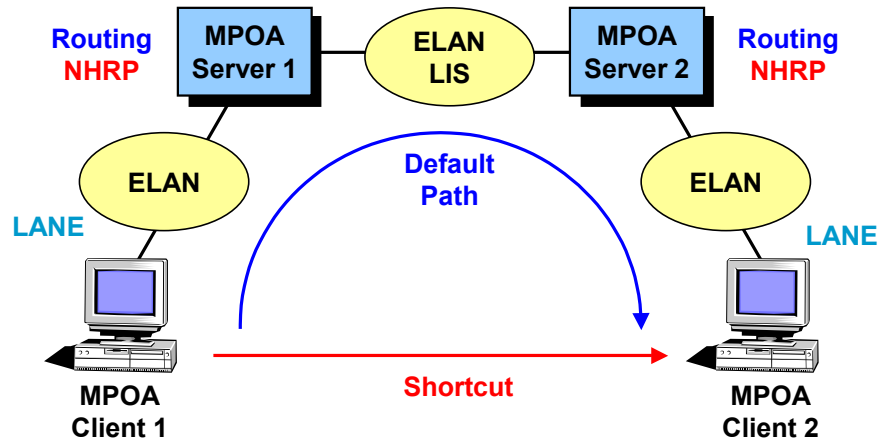
- The MPOA server continuously gathers routing information all over the network using the **Next Hop Routing Protocol (NHRP)** and determines the best route between two edge devices at any given moment.

Virtual Router

- From the outside an MPOA network looks like a huge **virtual router**, with the edge devices being the ports on the interface cards, the ATM switches being the backplane wiring and the MPOA server acting as the central routing processor.



- An ATM host starting up an IP stream sends the first packets to a **collocated** IP router / MPOA server at some node in the network. The MPOA server forwards the packets using IP routing to the egress edge device behind which the destination IP host is located.
- At the same time the MPOA server determines the ATM address of the egress edge device and communicates it to the ingress edge device which in turn sets up a switched virtual connection through the ATM network to the egress edge device. Thus a direct ATM shortcut is created over which the stream can be switched efficiently.



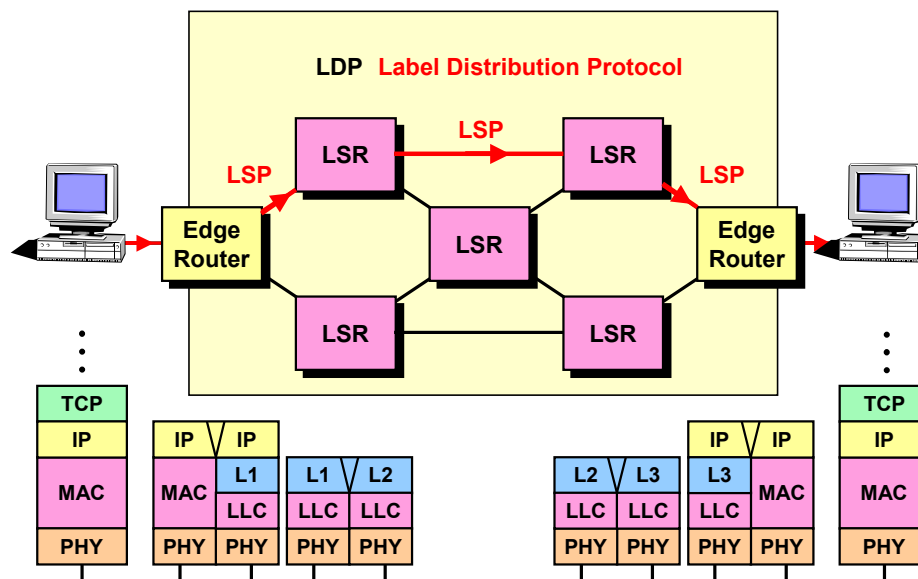
ELAN Emulated LAN **LIS** Logical IP Subnet

- Large ATM networks carrying IP-based data traffic are often subdivided into local Emulated LANs (ELAN) or IP-over-ATM based Logical IP Subnets (LIS).
- Within the subnets layer 2 based bridging based on ATM switching can be applied. The individual subnets must be connected using layer 3 IP routing.
- When the IP routers connected to the edges of the individual subnets act at the same time as MPOA servers, then using the NHRP protocol an optimal shortcut can be found through the network, based on end-to-end ATM switching.

**Multiprotocol Label Switching
(MPLS)**

Multiprotocol Label Switching (MPLS)

RFC 3031 - Architecture



A. Steffen, 18.11.2001, KSy_ATM_2.ppt 34

Labels (L1, L2, L3, ...)

- With Multiprotocol Label Switching (MPLS) labels are inserted between the layer 3 IP-datagrams and an arbitrary layer 2 link layer (e.g. ATM, PPP over SDH, etc.). Similar to an ATM VPI/VCI identifier the outermost label is valid only over the segment connecting two neighbouring Label Switched Routers (LSRs).

Label Switched Router (LSR)

- Similar to an ATM switch, a label switched router switches IP packets solely on the basis of looking at the label. The IP header information is not examined.

Edge Router

- The first time a new IP stream is fed into an IP-based network through an edge router, an optimal source originated route is constructed by the edge router, based on network information gathered through conventional, but enhanced IP routing protocols. Traffic engineering is achieved by applying additional constraints to the routing decisions.

Label Distribution Protocol (LDP)

- The labels that are going to be used from hop to hop are distributed to the nodes on the route either using the reliable TCP/IP-based **Label Distribution Protocol** (LDP) or via extensions of the unreliable Resource Reservation Protocol (RSVP) with IP protocol number 46.

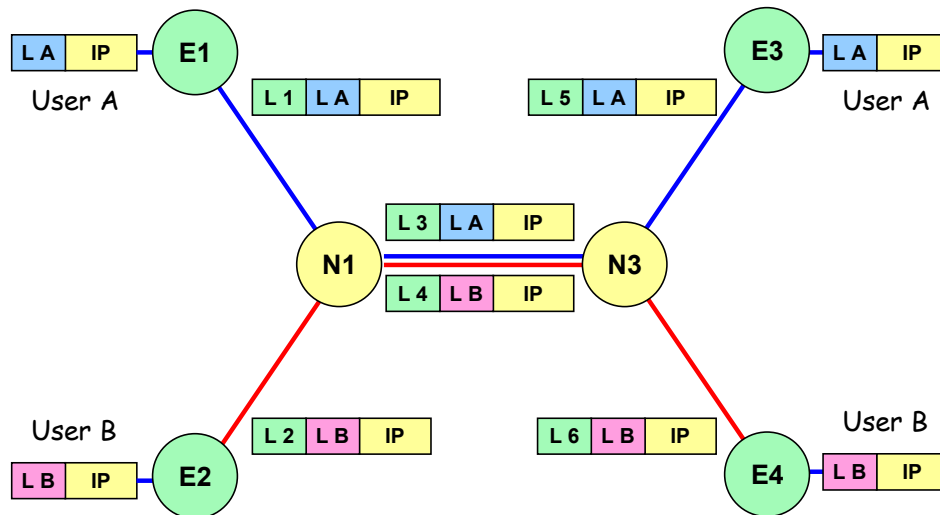
Label Switched Path (LSP)

- the path set up through the network by the label distribution protocol is called a Label Switched Path (LSP).

Why is MPLS the most attractive solution at the moment ?

- MPLS has all the advantages of a connection-oriented network while still retaining the underlying efficiency of a datagram based network.

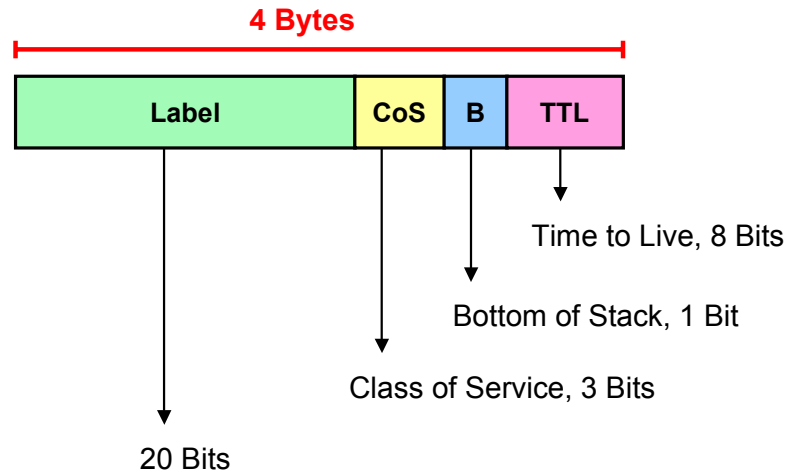
MPLS - Label Stack Virtual Private Networks (VPNs)



Virtual Private Network (VPN) based on MPLS

- By prepending a user label LA in front of each IP packet the whole IP-traffic from user A can be transported from ingress node E1 (e.g. home office) to egress node E3 (e.g. company headquarters) without examining the IP headers along the route. Even private network addresses (e.g. 10.x.x.x) could be transported.
- From hop to hop an outer switching label (L1, L3, L5) that defines the label switched path is pushed onto the label stack at the beginning of a hop and popped again after the switching decision at the end of a hop.
- Labelling client traffic also allows efficient billing based on the number of transmitted IP-packets.

Multiprotocol Label Switching (MPLS) RFC 3032 - Shim Header



Shim Header

- The shim header is carried after the data link layer header (layer 2) and before the IP header (layer 3).

Label

- Multiple encapsulations may exist, i.e. labels may be stacked to an arbitrary depth:
- Stacking maintains identity of several streams when they are aggregated into a single Label Switched Path (LSP)
- Labels are a generalization of ATM's concept of a dual hierarchy established by VPCs and VCCs.

Class of Service

- The Class of Service (CoS) field is similar to the seldom used Type of Service (TOS) field in the IP header. MPLS uses these experimental bits for QoS purposes.

Time to Live

- The Time to Live (TTL) field takes over the function of the TTL field in the IP header hidden by encapsulation.

Source: Hans Weibel, ZHW Course „Kommunikationsnetze“, May 2000