BIBA Bremen Institute of Industrial Technology and Applied Work Science

ATM in TCP/IP environment: Adaptations and Effectiveness

Dipl.-Ing. Kai-Oliver Detken, BIBA

ATM Traffic Symposium,

Mykonos, Greece, September the 17th, 1997

BIBA

Content

- Introduction of the European ACTS project EIES (European Information Exchange Service)
- IP-over-ATM difficulty
- IP-over-ATM possibilities: Classical IP, LAN Emulation (LANE), and MPOA
- IP effectiveness: protocol overhead, TCP throughput obstacles, sender and receiver buffers, RTT and MTU deadlocks, UDP and Client performance
- Performance Improvements and further developments

BIBA

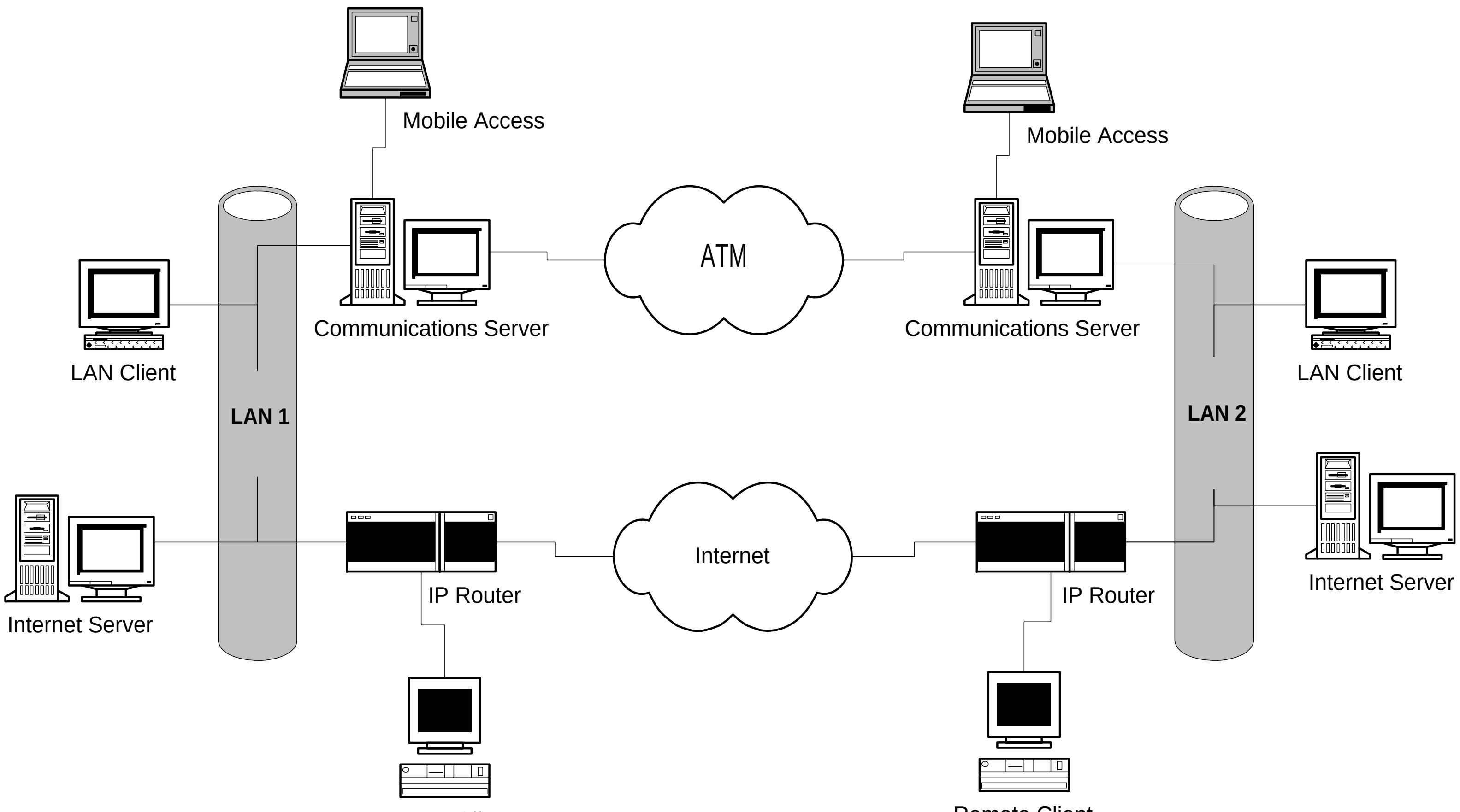
The ACTS project EIES

- Definition, implementation, and experimentation of advanced telecommunication services
- Support routine and non-routine communication between different maritime players
- Define and transpose the user requirements for the service development of EIES
- Build-up a pre-commercial global EIES service
- Establish an ATM link between the different harbour sites (Bordeaux, Bremen, Brest, Santander) to use the prototypes in real life

EIES services

- Computer Supported Cooperative Work (CSCW)
- Distribution of multimedia information about different databases (BluePages and Port Entry Guide)
- Integration of Electronic Data Interchange (EDI) and Electronic Mail (Email)
- Mobile communication (DSRR, DECT, and Inmarsat) for the mobile access to the services.
- Fixed networks: ATM, ISDN, and Internet (IP)

EIES platform



BIBA

IP features

- Internet Protocol works connectionless (hop-by-hop-transmission)
- No acknowledgements, error- and duplicate detection with IP
- Data packets are variable: 20 byte - 65 kbyte; destination address identified the other client
- Lost data packets have to repeat requested by higher layer protocols (TCP)
- Multicast/broadcast functions
- Type-of-Service (TOS)

ATM features

- Connection-oriented technologie: virtual connections will establish for the cell transport (PVC/SVC)
- ATM has its own address structur and routing funktions
- 53 byte cells with fixed size contain payload and control data
- Point-to-point connections
- Quality-of-Service (QoS)

BIBA

IP-over-ATM Adaptations

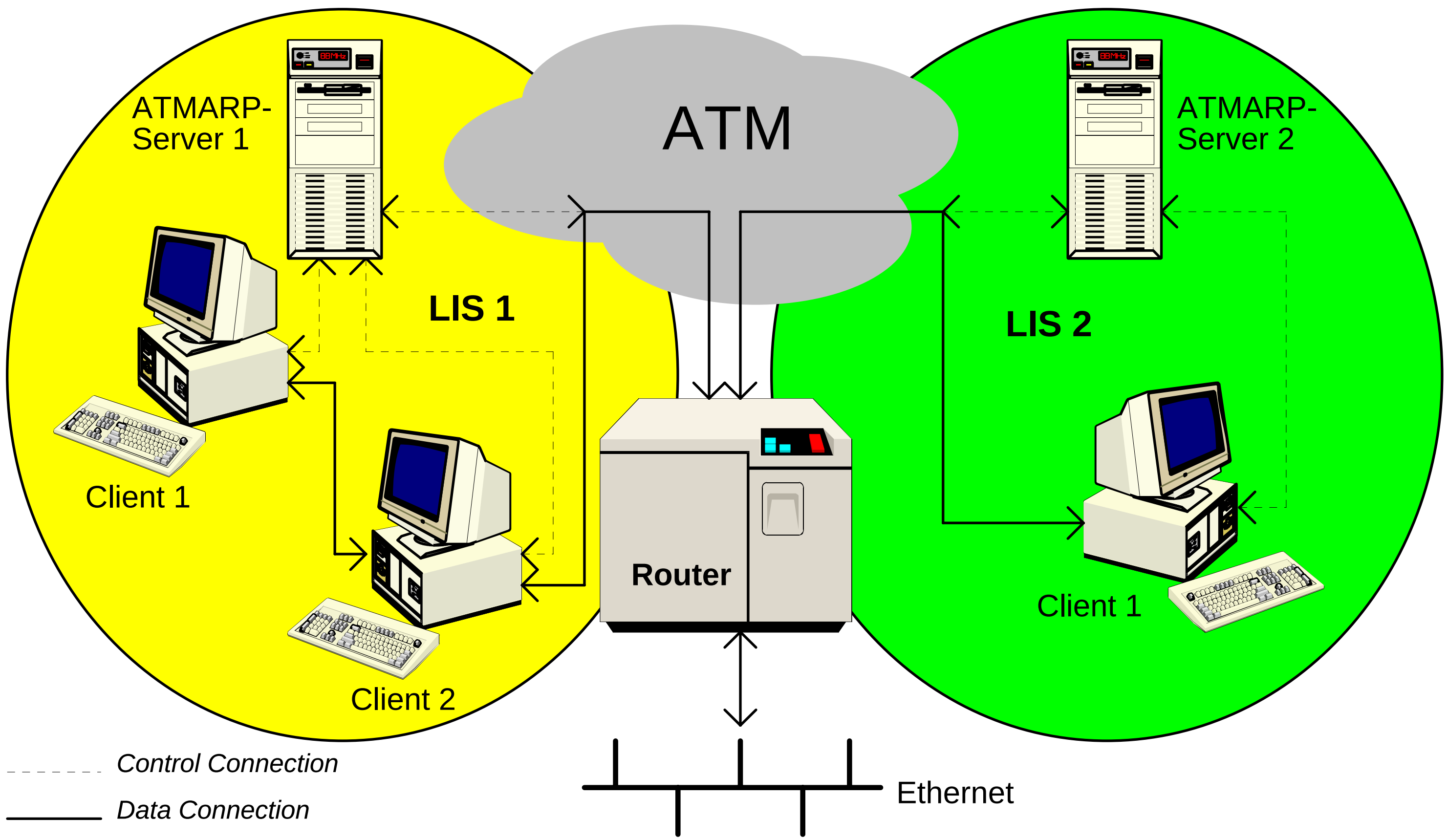
- Classical-IP
  - IETF: RFC-1577
  - ATM clients use IP
- LAN-Emulation (LANE)
  - ATM-Forum
  - Traditional networks use ATM
- Multiprotocol over ATM (MPOA)
  - ATM-Forum and IETF protocols
  - Routing mechanisms will be implemented

BIBA

Classical-IP (RFC-1577)

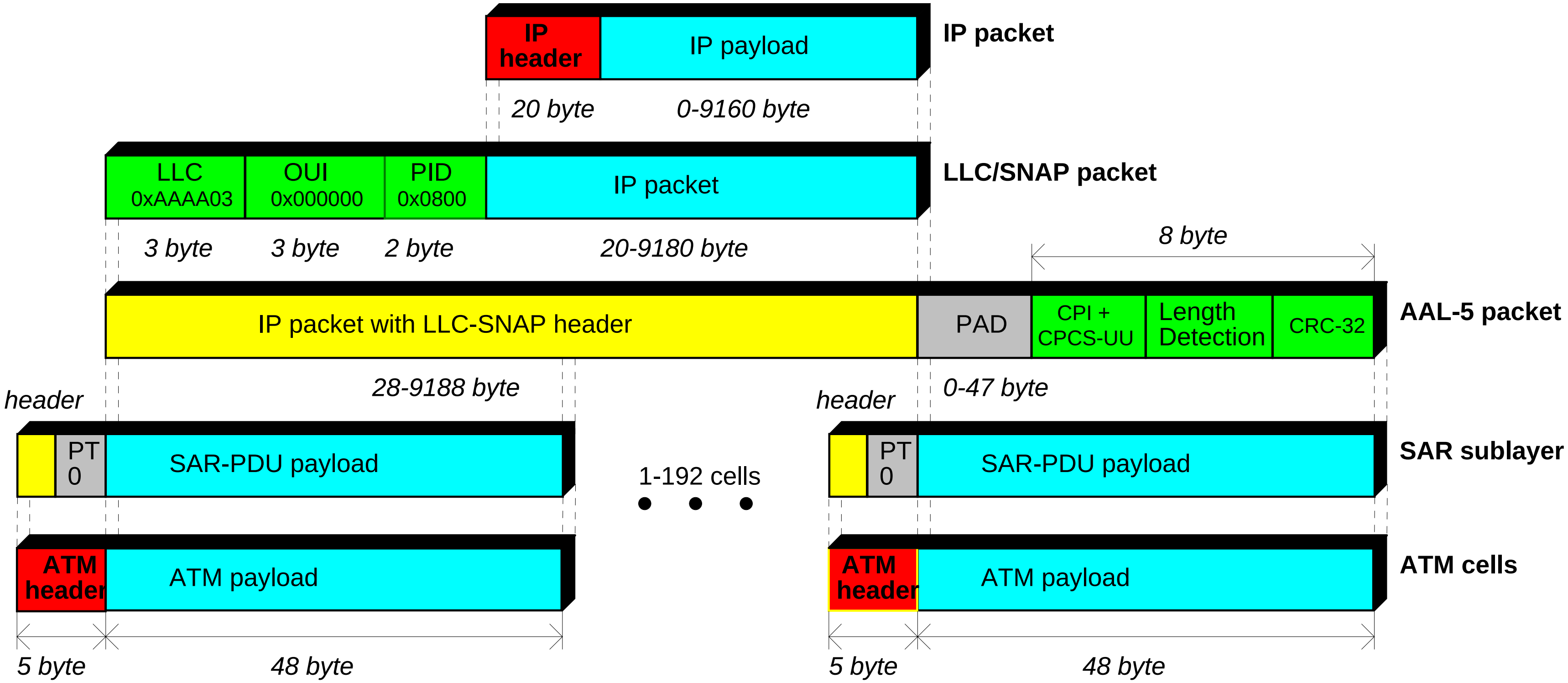
- RFC-1577: Classical-IP and ARP over ATM
- Logical Link Control Encapsulation: Multiple protocols can be transported over a single connection (RFC-1483)
- VC-Based Multiplexing: Single protocol is transported over an ATM connection (RFC-1483)
- D-MTU: 9180 Byte
- PVC/SVC connections
- Transport over the AAL-5 layer
- Point-to-point connections
- Address resolution by central ATMARP server

Classical-IP architecture



BIBA

Encapsulation (RFC-1483)



RFC-1577 conclusions

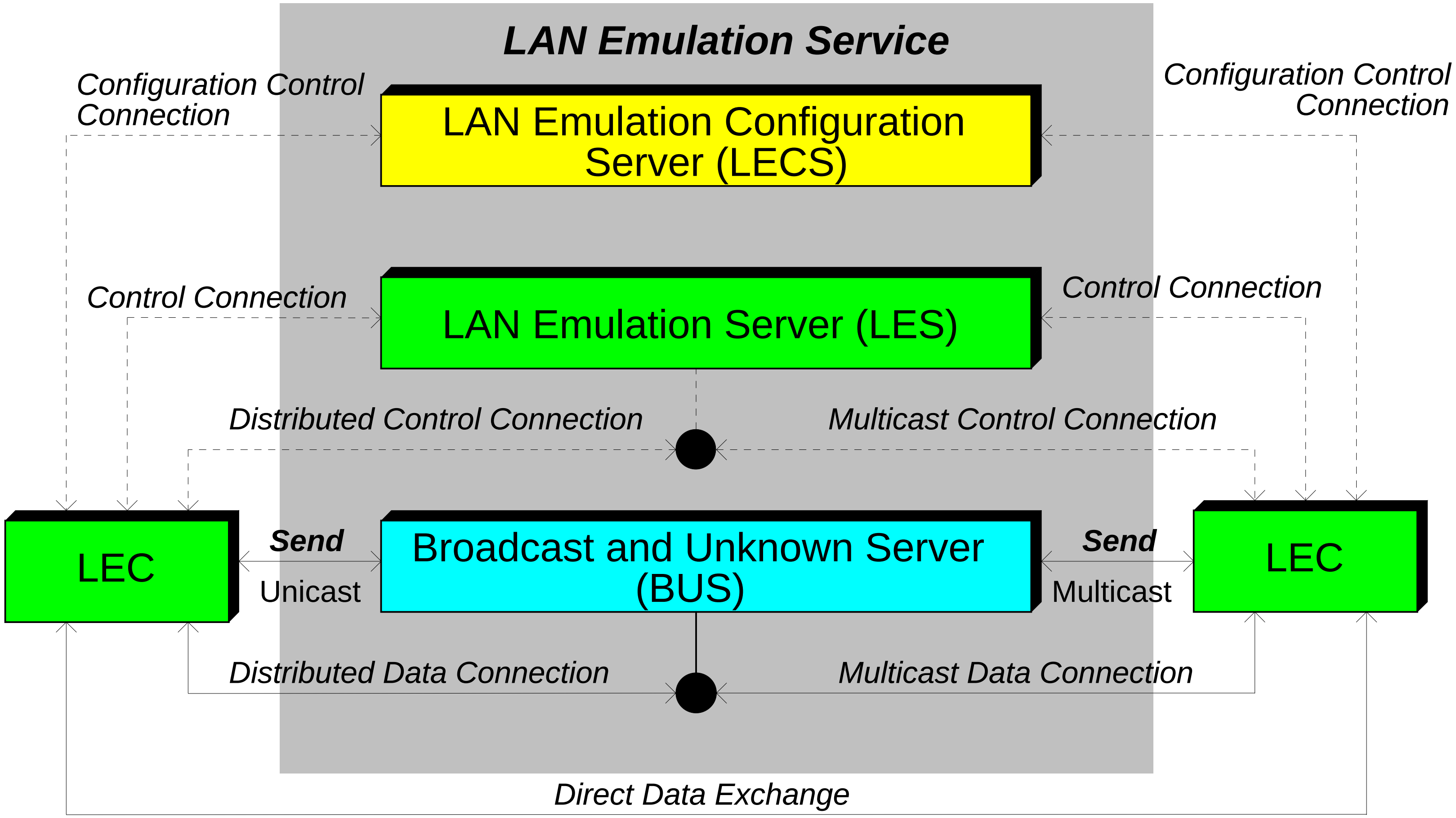
- Advantages:
  - high data throughput by high MTU (9180 byte)
  - simple and stabile standard
- Disadvantages:
  - only IP unicast is supported
  - high manual configuration effort
  - no multicast/broadcast support
  - no redundant ATMARP server
  - direct connections to other LIS subnets is not possible
  - no QoS

BIBA

LAN Emulation (LANE)

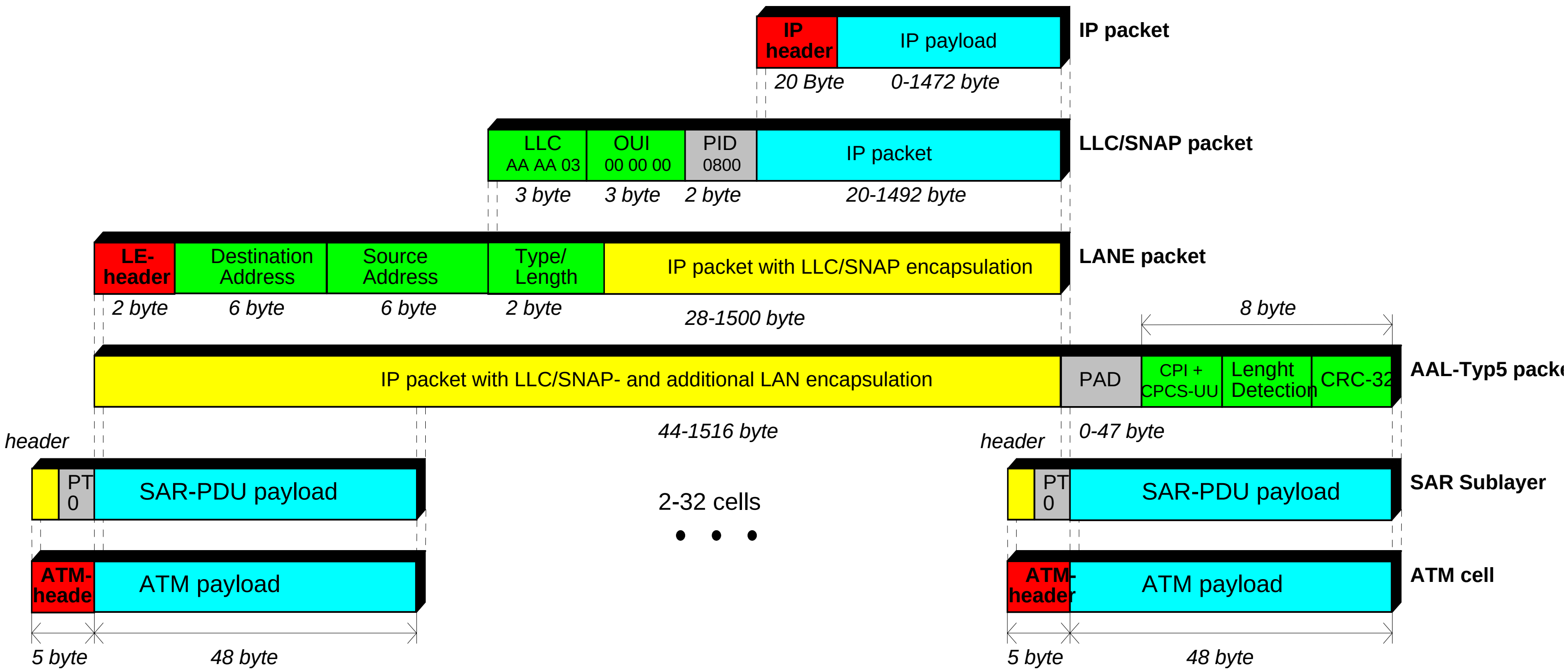
- Migration of traditional LANs to ATM: Ethernet switch (layer 2) und router (layer 3) coupling
- Universal implementation (MAC layer is emulated): arbitrary LAN protocols
- Using of the application layer without ATM configuration
- Support of PVC/SVC connections
- IP multicasting
- D-MTU: 1500 byte

LANE architecture



BIBA

LANE encapsulation



BIBA

LANE conclusions

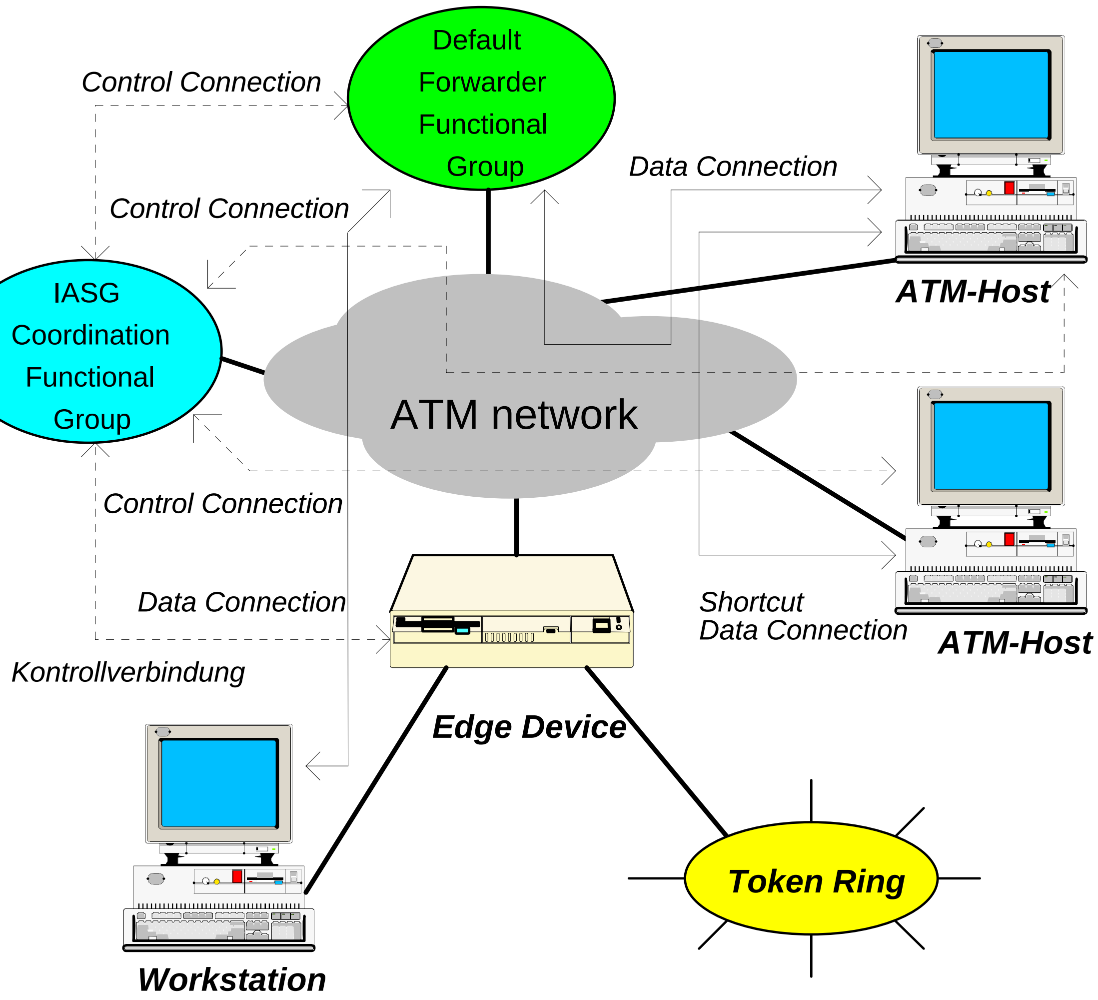
- Advantages:
  - integration of traditional networks
  - transparency for arbitrary network protocols
  - multicast/broadcast support
  - automatic configuration
- Disadvantages:
  - high functionality (layer 2 emulation)
  - MTU and older network boards limitated throughput
  - bad scalability
  - no redundant BUS (overload)
  - no QoS

MPOA capacity

- Support multiple protocols effective over ATM networks
- Distribute the routing functions between route servers
- Separate routing from switching functions
- Leverage performance and QoS capabilities of ATM network
- Direct connections between ELANs rather than passing through traditional routers
- Direct Virtual Channel Connections (VCC) between data forwarding devices
- Interworking with unified routers
- Enables subnet members to be distributed across the network
- Efficient scalability of the ATM network

BIBA

MPOA architecture



BIBA

Extensions for MPOA attempts

- Multicast Address Resolution Server (MARS)
  - support of multicasting by additional MARS server
- Next-Hop Resolution Protocol (NHRP)
  - direct connection between subnets possible
- Resource Reservation Protocol (RSVP)
  - QoS parameter at connectionsless packets
  - similar to the UNI signaling
- Private Network-to-Network-Interface (P-NNI)
  - control data exchange between ATM systems
  - transparency pass on the virtual connection

MPOA conclusions

- Advantages:
  - QoS is supported for the first time
  - router bottlenecks are eliminated
  - management of large networks
  - build up of VLANs
- Disadvantages:
  - MPOA is not a full defined spezification (1998)
  - IETF protocols must further developed
  - scalability is an extremely complex issue
  - MPOA does not provide a stable network yet

IPoverATM-Overhead

• STM-1-Frame (OC-3):	155,52 Mbps
• ATM Layer:	149,76 Mbps
• Adaptation Layer AAL-5:	135,632 Mbps
• LLC/SNAP encapsulation:	
- MTU: 576 byte	126,937 Mbps
- MTU: 9180 byte	135,22 Mbps
- MTU: 65527 byte	135,563 Mbps
• Internet Protocol (IP):	125,2/135,1/135,6 Mbps
• Transport Layer:	120,9/134,8/135,5 Mbps
• Application Layer via TCP:	116,5/134,5/135,5 Mbps
• Application Layer via UDP:	119,1/134,7/135,5 Mbps

BIBA

TCP-Throughput Obstacles

- Socket buffer size at sender/receiver
- Network: Maximum Transport Unit (MTU)
- Protocol: Maximum Segment Size (MSS)
- Sender: using of the Nagle’s algorithm
- Round-Trip-Time (RTT)
- Receiver: delayed acknowledegements
- Sender: Silly Window Syndrom (SWS)
- Copy strategy at the socket interface
- Network congestion and lost packet notice

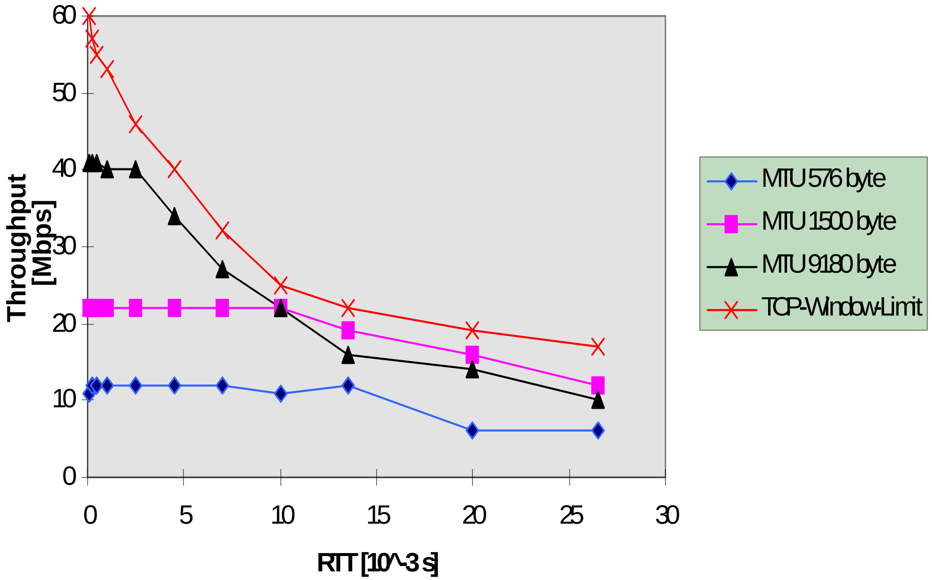
BIBA

Sender and Receiver buffer

S/E	16 kbyte	20 kbyte	24 kbyte	28 kbyte	32 kbyte	36 kbyte	40 kbyte	44 kbyte	48 kbyte	52 kbyte
16 kbyte	15,05	13,60	0,322	0,319	0,319	0,467	0,469	0,466	0,469	0,469
20 kbyte	15,99	14,60	15,07	14,87	15,40	14,24	1,095	1,095	0,548	0,549
24 kbyte	17,71	16,79	16,74	16,32	17,40	17,31	17,42	17,12	0,760	0,740
28 kbyte	16,57	17,69	17,93	18,13	18,36	19,20	19,74	19,78	18,38	18,20
32 kbyte	14,63	18,96	18,42	19,23	19,14	19,74	19,96	20,31	19,69	19,17
36 kbyte	14,33	19,22	18,12	19,82	19,77	19,92	20,56	20,49	20,13	20,20
40 kbyte	15,16	19,34	18,85	19,73	20,11	20,41	20,81	20,74	20,69	20,57
44 kbyte	14,80	19,40	18,27	20,39	20,16	20,74	20,99	20,87	20,89	20,70
48 kbyte	14,62	19,46	18,34	20,48	20,26	20,41	20,85	20,83	20,93	20,83
52 kbyte	13,92	19,41	18,26	20,50	20,06	20,21	20,88	20,91	21,21	21,06

RTT and MTU deadlocks

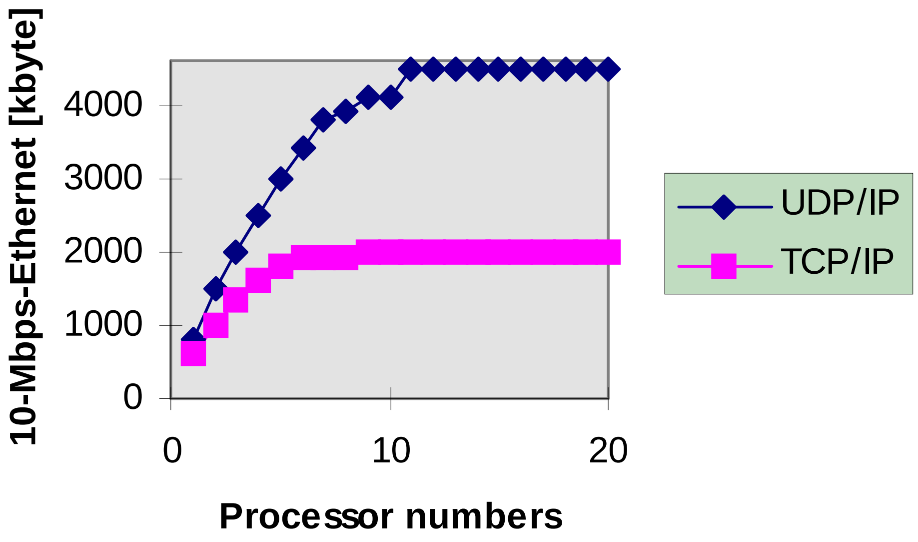
- MTU of 9180 byte has the best performance if the RTT is low
- MTU of 1500 byte has the best performance if the RTT is high
- If the MTU is smaller as the used packet size: fragmentation and reassembling is necessary!



BIBA

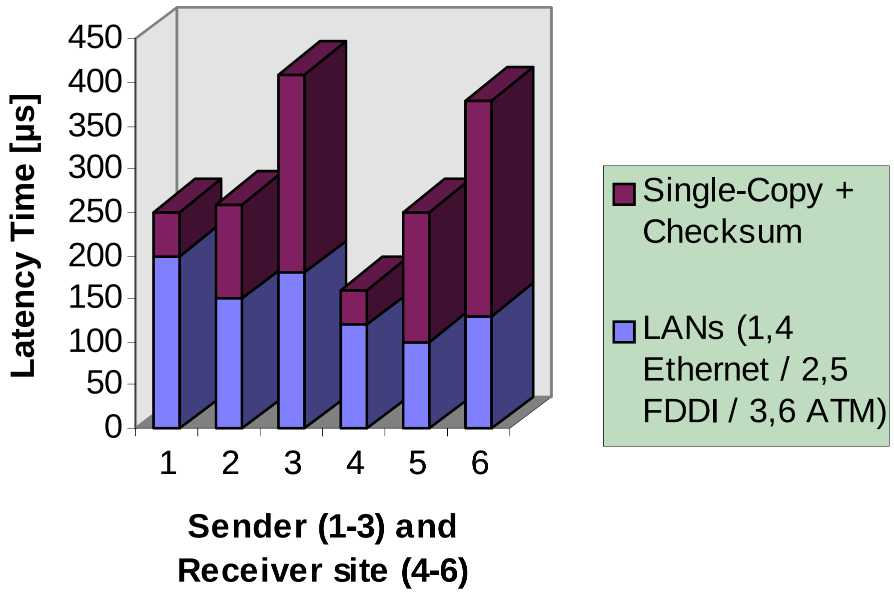
UDP performance

- Connectionless structure
- UDP has no acknowledgements, no end-to-end control and no duplicate detection
- Less locking effects for protect shared protocol state and data
- Locking Effects limits the data throughput.



Client performance

- Operation system (inter-process communication)
- Data processing
- CPU performance
- Memory speed
- I/O bus bandwidth
- ATM adapter implementation



BIBA

Client Tuning

- Using of the D-MTU value
- Path-MTU: max. possible packet size should be switch on (none fragmentation)
- MSS value: 1436 byte for high speed networks
- Sender and receiver buffer at least 2 times bigger than MTU value
- Window Scale Option (RFC-1323) higher Window-Size

Classical-IP:	Throughput of 134,01 Mbps (Sun20-Ultra)
LAN-Emulation, test 1:	Throughput of 76,30 Mbps (Sun20-Ultra)
LAN-Emulation, test 2:	Throughput of 117,63 Mbps (Ultra-Ultra)

BIBA

...further developments

- Classical-IP, Version 2
  - distributed ATMARP server
- LANE, Version 2
  - L-UNI: QoS, multicast, ELAN multiplexing
  - L-NNI: distributed LES/BUS
- IP-Switching (Ipsilon)
  - none standard
  - other proposal Tag-Switching from CISCO