L31 - IP Technology Details

IP Technology Details

IP Protocol, ICMP, ARP, RARP,
proxy ARP, HSRP, VRRP, PPP

Agenda

- IP Protocol
- ICMP
- Ping and Traceroute
- Address Resolution Protocol ARP
- RARP
- Proxy ARP
- VRRP, HSRP
- SLIP
- PPP

© 2007, D.I. Manfred Lindner IP Details, v4.7 2

© 2007, D.I. Manfred Lindner

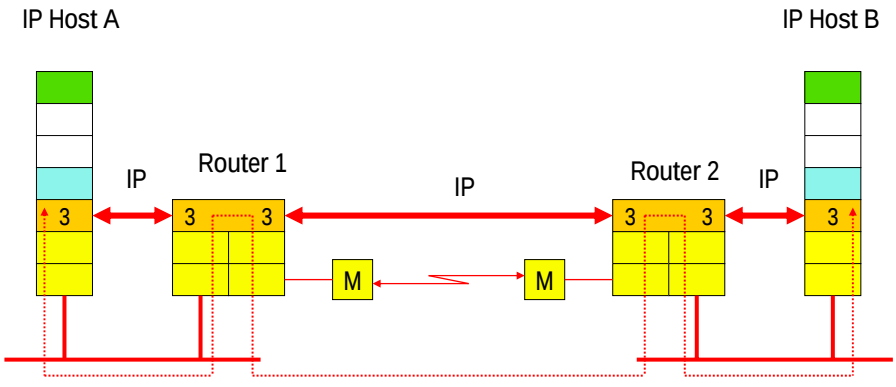
IP Internet Protocol (RFC 791)

- OSI layer 3 protocol with datagram service (unreliable connectionless service, "best effort service")
- Transports packets (datagrams) from a sender through one or more networks to a receiver
- Doesn't guarantee delivery or correct sequence of packets (task of higher layers)
- IP datagrams are encapsulated in layer 2 frames
- Encapsulation is a key feature of the TCP/IP suite, it provides versatility and independence from the physical network

© 2007, D.I. Manfred Lindner IP Details, v4.7 3

IP and OSI Network Layer 3

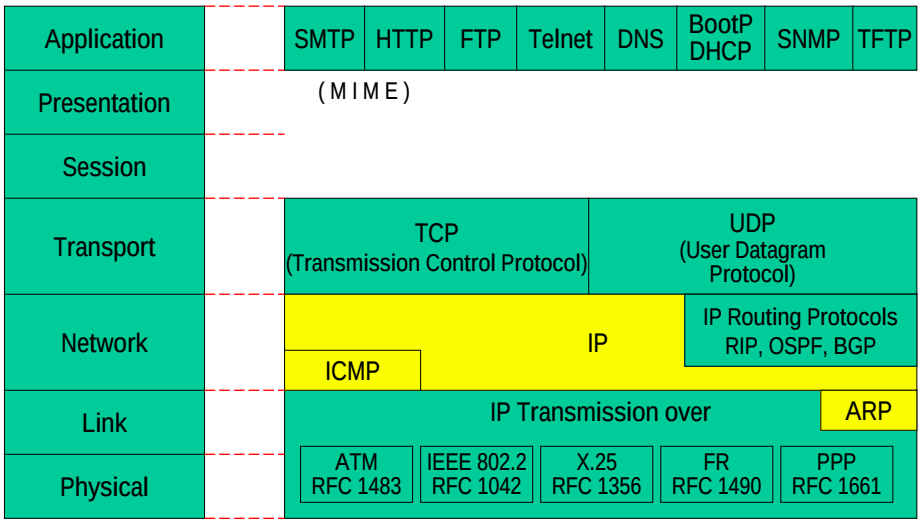
Layer 3 Protocol = IP
Layer 3 Routing Protocols = RIP, OSPF, EIGRP, BGP



© 2007, D.I. Manfred Lindner IP Details, v4.7 4

© 2007, D.I. Manfred Lindner

IP Related Protocols



© 2007, D.I. Manfred Lindner IP Details, v4.7 5

IP Protocol Functions

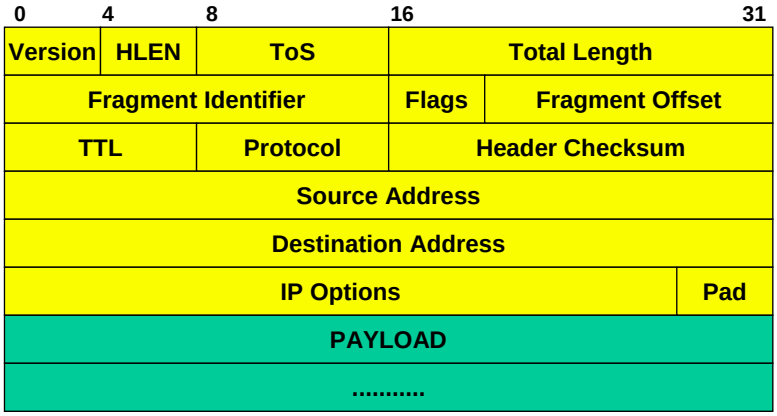
- **Mechanisms for packet forwarding, based on network addressing (Net-IDs)**
- **Error detection (only packet header)**
- **Fragmentation and reassembly of datagram's**
  - Necessary, if a datagram has to pass a network with a small max. frame size.
  - Reassembly by receiver
- **Mechanisms to limit the lifetime of a datagram**
  - To omit an endless circulating of datagrams if routing errors occur

© 2007, D.I. Manfred Lindner IP Details, v4.7 6

© 2007, D.I. Manfred Lindner

L31 - IP Technology Details

IP Header



IP Header Entries 1

- **Version**
  - Version of the IP protocol
  - Current version is 4
  - Useful for testing or for migration to a new version, e.g. "IP next generation“ (IPv6)
- **HLEN**
  - Length of the header in 32 bit words
  - Different header lengths result from IP options
    - HLEN 5 to 15 = 20 to 60 octets

© 2007, D.I. Manfred Lindner

IP Header Entries 2

- **Total Length**
  - Total length of the IP datagram (header + data) in octets
  - If fragmented: length of fragment
  - Datagram size max. = 65535 octets
  - Each host has to accept datagram's of at least 576 octets
    - either as a complete datagram or for reassembly

IP Header Entries 3

- **Protocol**
  - Indicates the higher layer protocols
    - Examples are: 1 (ICMP), 6 (TCP), 8 (EGP), 17 (UDP), 89 (OSPF) etc.
  - 100 different IP protocol types are registered so far
- **Source IP Address**
  - IP address of the source (sender) of a datagram
- **Destination IP Address**
  - IP address of the receiver (destination) of a datagram
- **Pad**
  - "0"-octets to fill the header to a 32 bit boundary

© 2007, D.I. Manfred Lindner

IP Header Entries 4

- **TTL Time To Live**
  - Limits the lifetime of a datagram in the network (Units are seconds, range 0-255)
  - Is set by the source to a starting value. 32 to 64 are common values, the current recommended value is 64 (RFC1700)
  - Every router decrements the TTL by the processing/waiting time. If the time is less than one second, TTL is decremented by one ("TTL = hop count").
  - If TTL reaches 0, the datagram (fragment) is discarded.
  - An end system can use the remaining TTL value of the first arriving fragment to set the reassembly timer.

© 2007, D.I. Manfred Lindner

IP Header Entries 5

- **Identification (for fragmentation)**
  - Unique identification of a datagram, used for fragmentation and reassembly
  - In praxis a hidden sequence number although not used because of connectionless behavior of IP
- **Flags (for fragmentation).**
  - DF (don't fragment)
    - If set: fragmentation is not allowed
    - Datagram's must be discarded by router if MTU (maximum transmission unit) size of next link is too small
  - MF (more fragments)
    - If set: more fragments of the same original datagram will follow

"0"	DF	MF	Fragment Offset

© 2007, D.I. Manfred Lindner

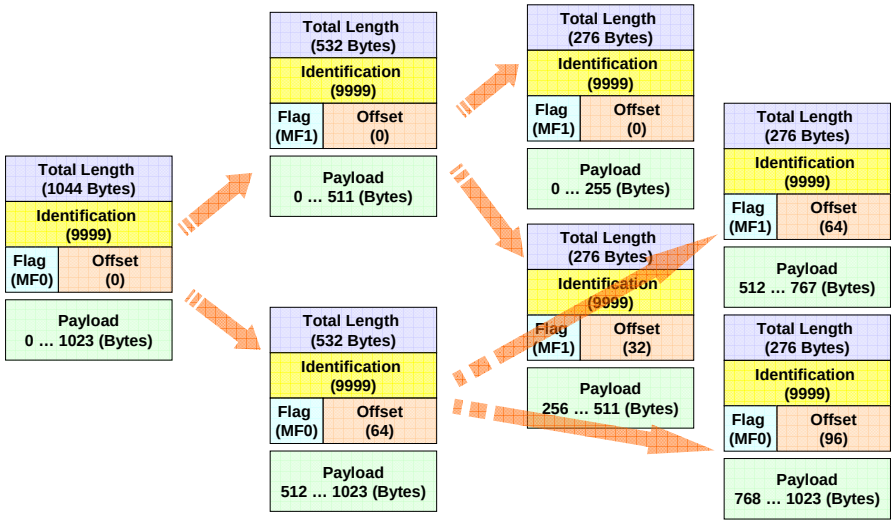
© 2007, D.I. Manfred Lindner

IP Header Entries 6

- **Fragment Offset**
  - Indicates the position of a fragment relative to the beginning of the original datagram
  - Offset is measured in multiples of 8 octets (64 bits)
  - The first fragment and unfragmented packets have an offset of 0
  - Fragments (except the last) must be a multiple of 8 octets
  - Fragments with the same combination of source address / destination address / protocol / identification will be reassembled to the original datagram

© 2007, D.I. Manfred Lindner

IP Fragmentation



© 2007, D.I. Manfred Lindner

© 2007, D.I. Manfred Lindner

Reassembly

- Reassembly is done at the destination, because fragments can take different paths
- Buffer space has to be provided at the receiver
- Some fragments may not arrive (unreliable nature of IP)
- Measures must be taken to free buffers if a packet can't be reconstructed in a timely manner
- The first arriving fragment of an IP packet (with MF=1 and/or Offset <> 0) starts a reassembly timer
- If the timer expires before the packet was reconstructed, all fragments will be discarded and the buffer is set free
- The reassembly timer limits the lifetime of an incomplete datagram and allows better use of buffer resources.

© 2007, D.I. Manfred Lindner

IP Header Entries 7

- **TOS field (Type Of Service)**
- **Old Meaning (RFC 1349)**
  - Tells the priority of a datagram (precedence bits) and the preferred network characteristics (low delay, high throughput, high reliability, low monetary cost.)
  - Precedence bits:
    - Define the handling of a datagram within the router
    - e.g. priority within the input / output queues
  - D, T, R and C bits:
    - Can be used to take a path decision for routing if multiple paths with different characteristics exist to the destination
      - needs one routing table per characteristic
    - TOS bits may be ignored by routers but may never lead to discarding a packet if the preferred service can't be provided

© 2007, D.I. Manfred Lindner

© 2007, D.I. Manfred Lindner

TOS Field Old Meaning (RFC 1349)

Precedence	D	T	R	C	"0"

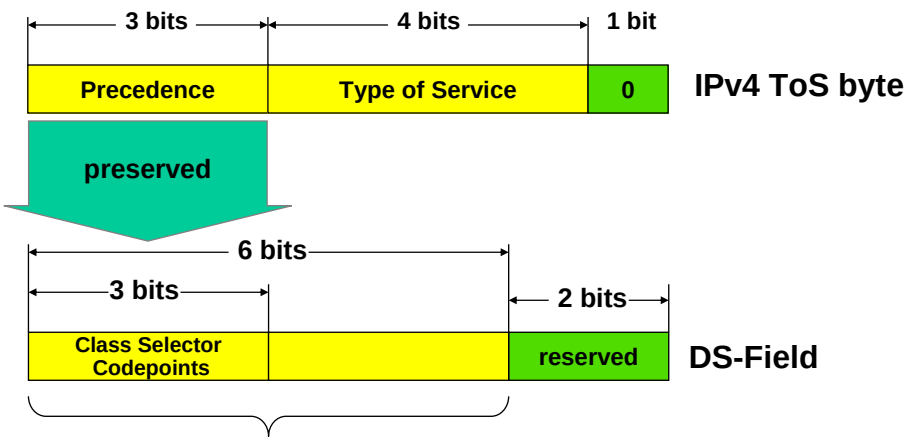
Precedence (Priority):	DTRC bits:		
111 Network Control	0 0 0 0 . .	. . . .	normal service
110 Internetwork Control	1 0 0 0 D	Delay	min. delay
101 Critic/ECP	0 1 0 0 T	Throughput	max. throughput
100 Flash Override	0 0 1 0 R	Reliability	max. reliability
011 Flash	0 0 0 1 C	Cost	min. cost
010 Immediate			
001 Priority	No other values are defined but have to be accepted (ignored) by a router or host.		
000 Routine			

IPv4 TOS Recycling

- **IPv4 TOS field was redefined by the IETF to become the "Differentiated Service CodePoint" (DSCP)**
- **Now the DSCP field is used to label the traffic class of a flow**
  - a flow is a given communication relationship (session) between two IP hosts
  - IP datagram's of a flow have the same
    - Source IP Address
    - Destination IP Address
    - Protocol Number
    - TCP/UDP Source Port
    - TCP/UDP Destination Port

© 2007, D.I. Manfred Lindner

IPv4 TOS Recycling



© 2007, D.I. Manfred Lindner

DSCP Usage

- **Important for IP QoS (Quality of Service)**
  - IP QoS Differentiated Services Model
    - RFC 2474: "Definition of the Differentiated Service Field in the IPv4 and IPv6 Headers"
    - RFC 2475: "An Architecture for Differentiated Services"
  - Remember
    - IP is basically a Best Effort Service, therefore not suited for interactive real-time traffic like voice and video
  - Using DSCP a IP datagram can be labelled at the border of IP QoS domain
    - with a certain traffic class
  - Traffic class will receive a defined handling within in IP QoS Domain
    - e.g. limited delay, guaranteed throughput

© 2007, D.I. Manfred Lindner

© 2007, D.I. Manfred Lindner

IP Header Entries 8

- **IP Options**
  - IP options have to be implemented by every IP station
  - The only thing optional is their presence in an IP header
  - Options include provisions for timestamps, security and special routing aspects
  - Some options may, others must be copied to all fragments
- **Today most IP Options are blocked by firewalls because of inherent security flaws**
  - e.g.source routing could divert an IP stream to a hacker´s network station

IP Options

- **Record Route**
  - Records the route of a packet through the network
  - Each router, which forwards the packet, enters its IP address into the provided space
- **Loose Source Route**
  - A datagramm or fragment has to pass the routers in the sequence provided in the list
  - Other intermediate routers not listed may also be passed
- **Strict Source Route**
  - A datagramm or fragment has to pass the routers in the sequence listed in the source route
  - No other router or network may be passed

© 2007, D.I. Manfred Lindner

IP Options

- **Security**
  - Four parameters ("security", "compartment", "handling" and "transmission control") tell a router about classification, restrictions and special treatments
  - The values and their meaning are defined by the US Defense Intelligence Agency
- **Timestamp**
  - Routers enter their current system time (or IP address and system time) into the provided space as they forward the packet
  - Format is millisecond since midnight UT
  - Prerequisite for analysis are synchronized clocks.

Agenda

- **IP Protocol**
- **ICMP**
- **Ping and Traceroute**
- **Address Resolution Protocol ARP**
- **RARP**
- **Proxy ARP**
- **VRRP, HSRP**
- **SLIP**
- **PPP**

© 2007, D.I. Manfred Lindner

ICMP (Internet Control Message Protocol)

- The datagram service of IP doesn't guarantee or acknowledge the delivery of a datagram
- ICMP generates error messages to enhance the reliability and to provide information about errors and packet loss in the network
- ICMP allows to request information for debugging and diagnosis
- ICMP has to be supported by every IP station but different implementation may vary in the way how ICMP messages are responded to

ICMP (RFC 792)

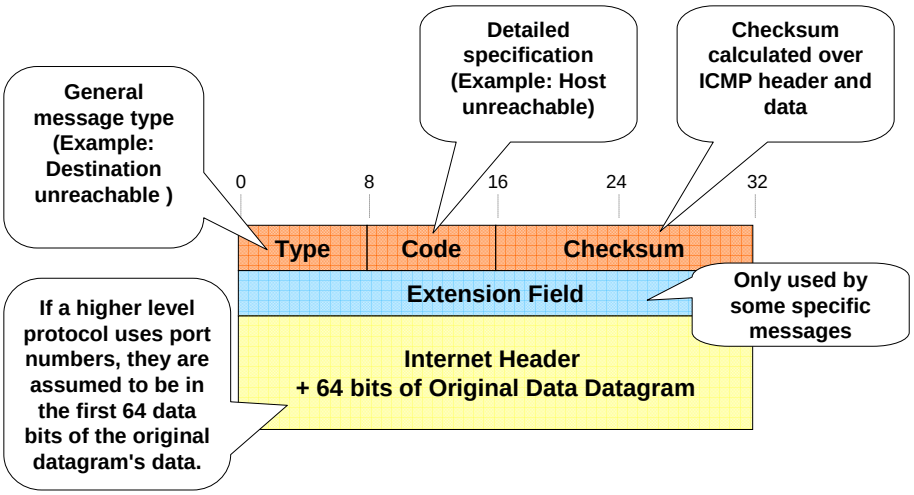
- **Principles of operation:**
  - The IP station (router or destination), which detects any transmission problems, generates an ICMP message
  - The ICMP message is addressed to the originating station (sender of the original IP packet)
- **ICMP messages are sent as IP packets**
  - protocol field = 1, ICMP header and code in the IP data area
- **Analysis of ICMP messages**
  - through network management systems or statistic programs can give valuable hints for network administrators

© 2007, D.I. Manfred Lindner

Important Rule

- **If a IP datagram carrying an ICMP message cannot be delivered**
  - No additional ICMP error message is generated to avoid an ICMP avalanche
  - "ICMP must not invoke ICMP"
- **Exception: PING command**
  - Echo request and echo response

ICMP Message Format



© 2007, D.I. Manfred Lindner

Type Field

0	Echo reply ("Ping")
3	Destination Unreachable
	Reason specified in Code
4	Source Quench (decrease data rate of sender)
	Theoretical Flow Control Possibility of IP
5	Redirect (use different router)
	More information in Code
8	Echo Request ("PING")
11	Time Exceeded (code = 0 time to live exceeded in transit code = 1 reassembly timer expired)
12	Parameter Problem (IP header)
13/14	Time Stamp Request / Time Stamp Reply
15/16	Information Request/ Reply (finding the Net-ID of the network; e.g. SLIP)
17/18	Address Mask Request / Reply

© 2007, D.I. Manfred Lindner

Using ICMP Types

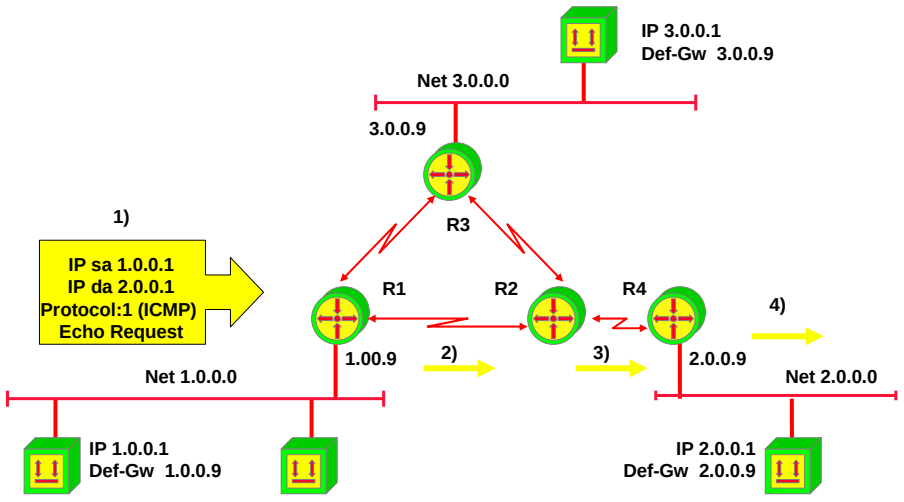
0, 8	"PING" testing whether an IP station (router or end system) can be reached and is operational
3, 11, 12	Signaling errors concerning reachability, TTL / reassambly timeouts and errors in the IP header
4	Flow control (only possibility to signal a possible buffer overflow)
5	Signaling of alternative (shorter) routes to a target
13 - 18	Diagnosis or management

© 2007, D.I. Manfred Lindner

© 2007, D.I. Manfred Lindner

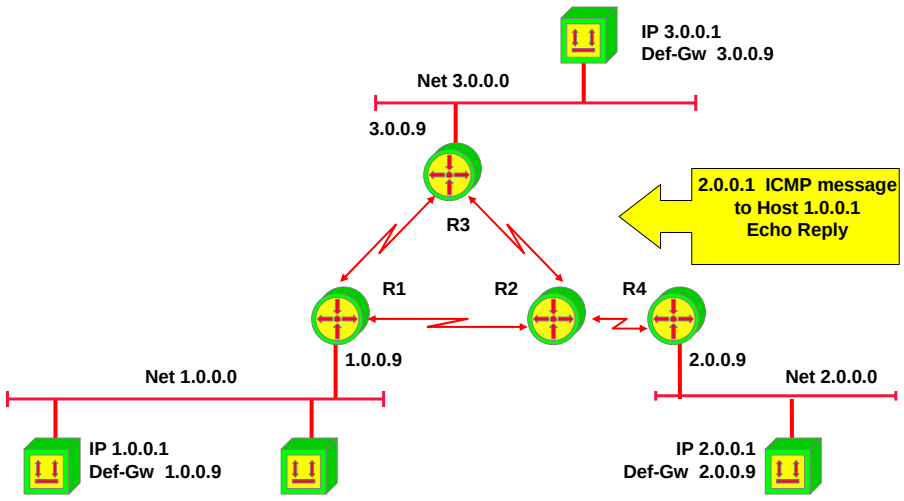
L31 - IP Technology Details

Ping 1.0.0.1 - > 2.0.0.1



© 2007, D.I. Manfred Lindner IP Details, v4.7 31

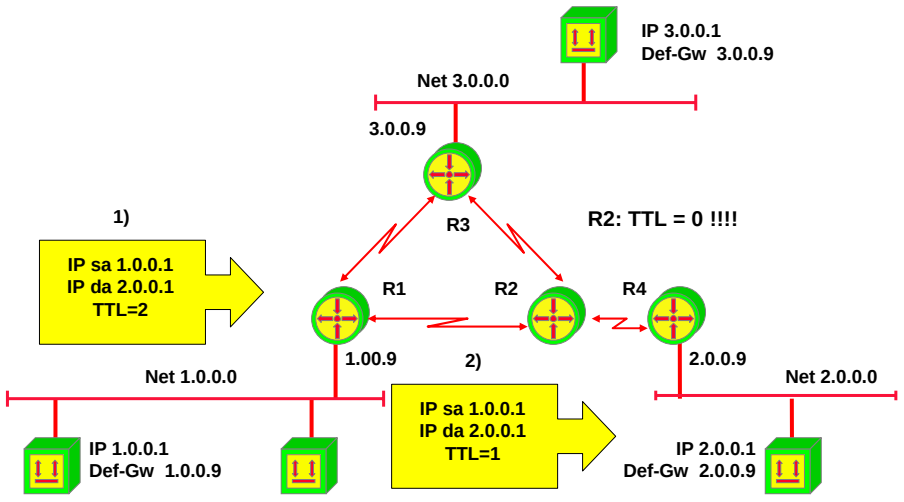
Ping Echo 2.0.0.1 - > 1.0.0.1



© 2007, D.I. Manfred Lindner IP Details, v4.7 32

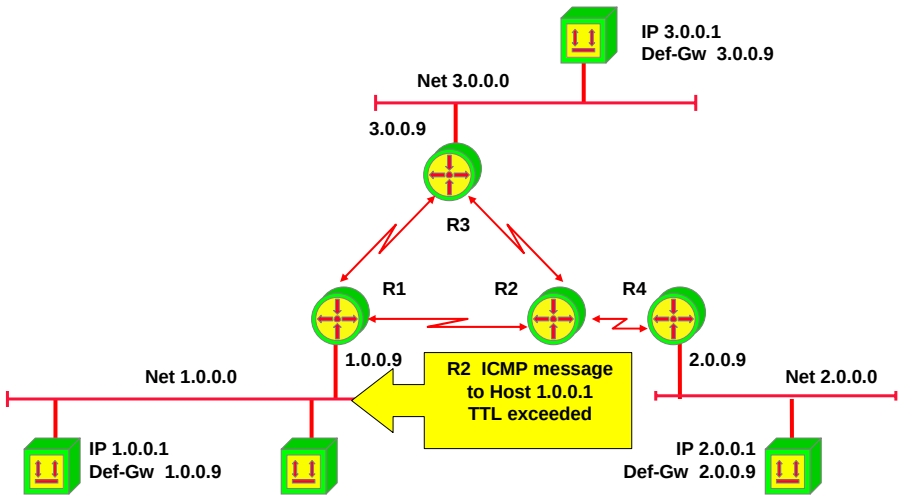
© 2007, D.I. Manfred Lindner

Delivery 1.0.0.1 - > 2.0.0.1 (TTL=2)



© 2007, D.I. Manfred Lindner IP Details, v4.7 33

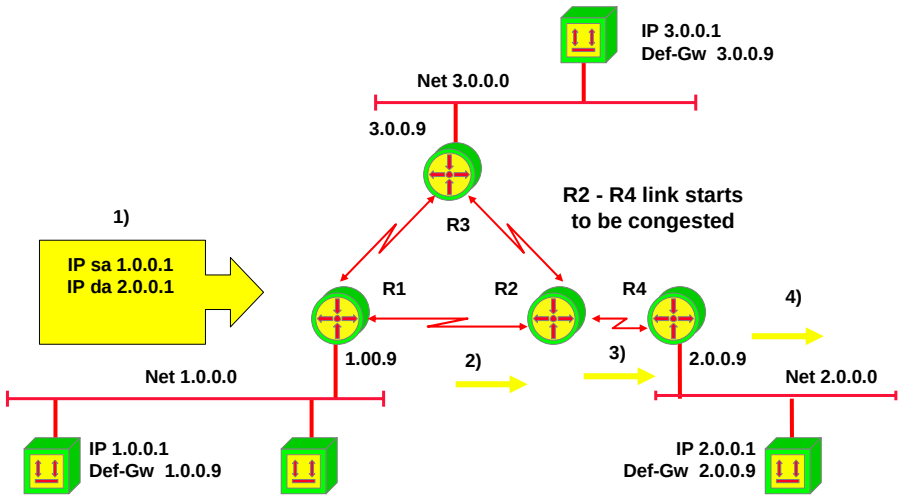
ICMP TTL exceeded



© 2007, D.I. Manfred Lindner IP Details, v4.7 34

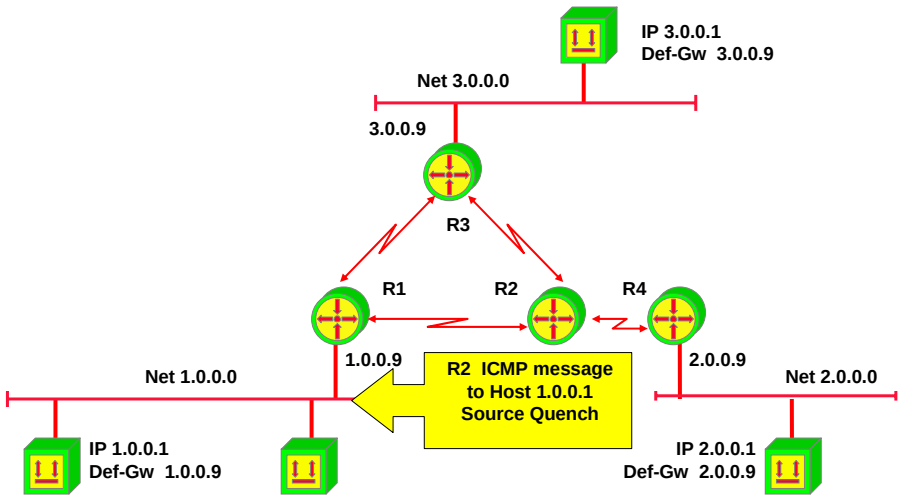
© 2007, D.I. Manfred Lindner

R2 -> R4 Link Congested



© 2007, D.I. Manfred Lindner IP Details, v4.7 35

ICMP Source Quench (Flow Control STOP?)



© 2007, D.I. Manfred Lindner IP Details, v4.7 36

© 2007, D.I. Manfred Lindner

Code Field for Type 3 (destination unreachable)

0 ... Network unreachable: no path to network known or network down; generated by intermediate or far-end router
1 ... Host unreachable: Host-ID can't be resolved or host not responding; generated by far-end router
2 ... Protocol unreachable: protocol specified in IP header not available; generated by end system
3 ... Port unreachable: port (service) specified in layer 4 not available; generated by end system
4 ... Fragmentation needed and do not fragment bit set: DF bit =1 but the packet is too big for the network (MTU); generated by router
5 ... Source route failed: Path in IP Options couldn't be followed; generated by intermediate or far-end router

Code Field for Type 3 (destination unreachable)

See RFC1122 (Host Requirements) page 38:

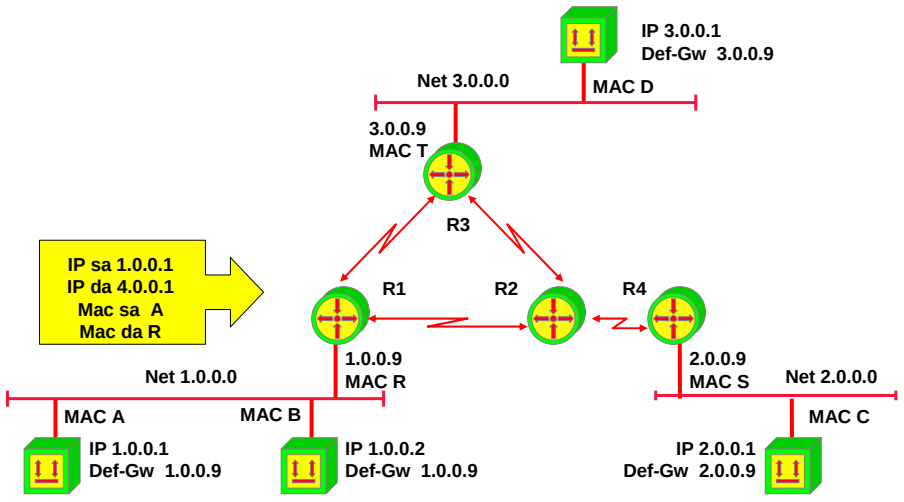
The following additional codes are hereby defined:

6 ... destination network unknown
7 ... destination host unknown
8 ... source host isolated
9 ... communication with destination network administratively prohibited
10 ... communication with destination host administratively prohibited
11 ... network unreachable for type of service
12 ... host unreachable for type of service

© 2007, D.I. Manfred Lindner

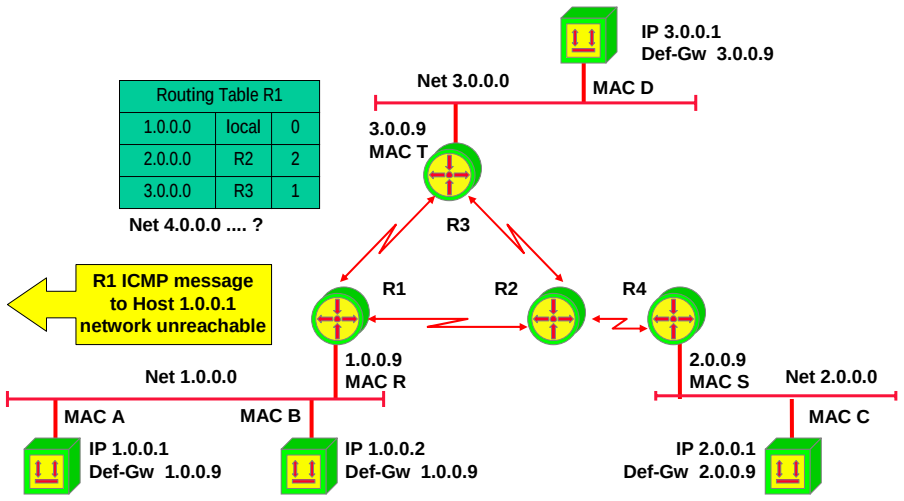
L31 - IP Technology Details

Delivery 1.0.0.1 - > 4.0.0.1



© 2007, D.I. Manfred Lindner IP Details, v4.7 39

ICMP network unreachable

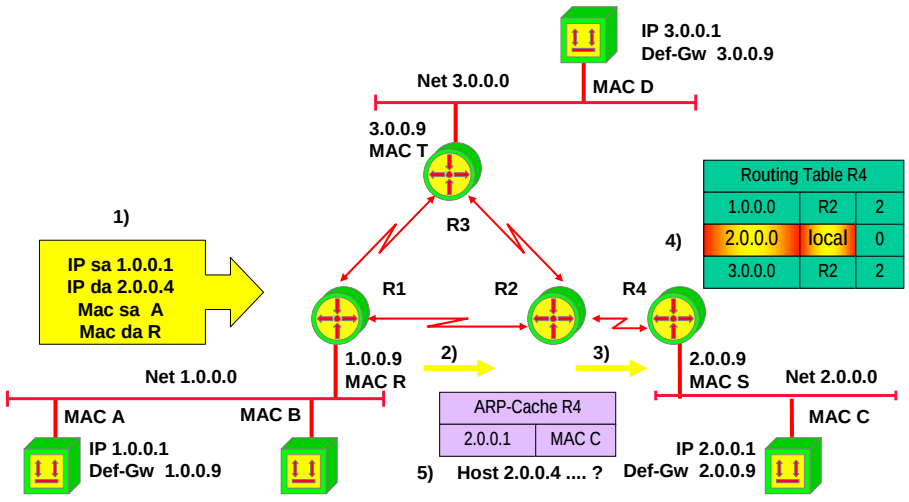


© 2007, D.I. Manfred Lindner IP Details, v4.7 40

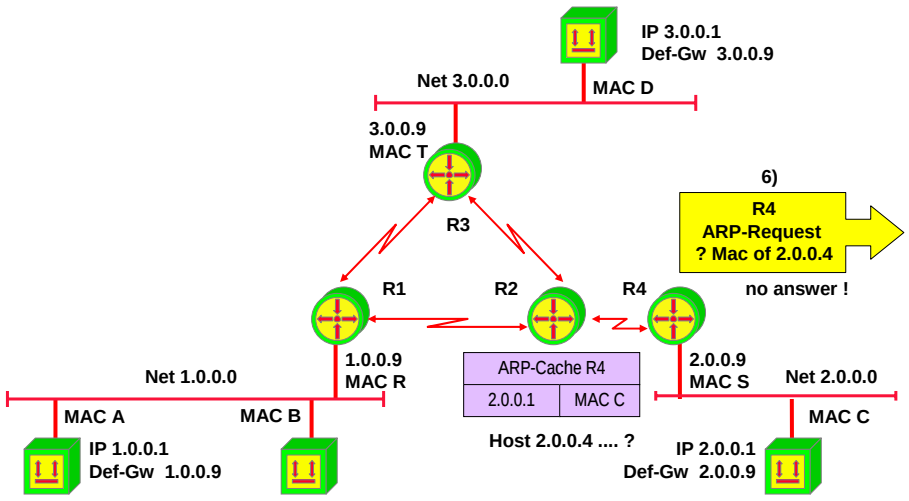
© 2007, D.I. Manfred Lindner

L31 - IP Technology Details

Delivery 1.0.0.1 - > 2.0.0.4

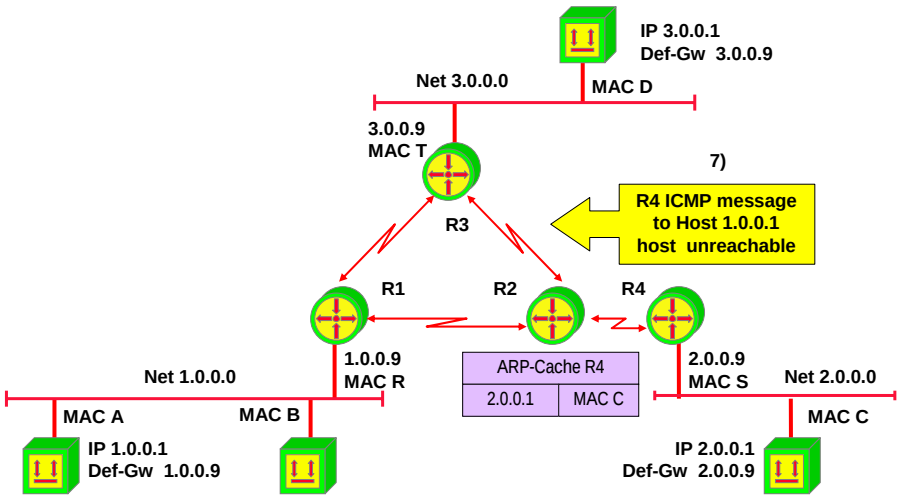


Delivery 1.0.0.1 - > 2.0.0.4

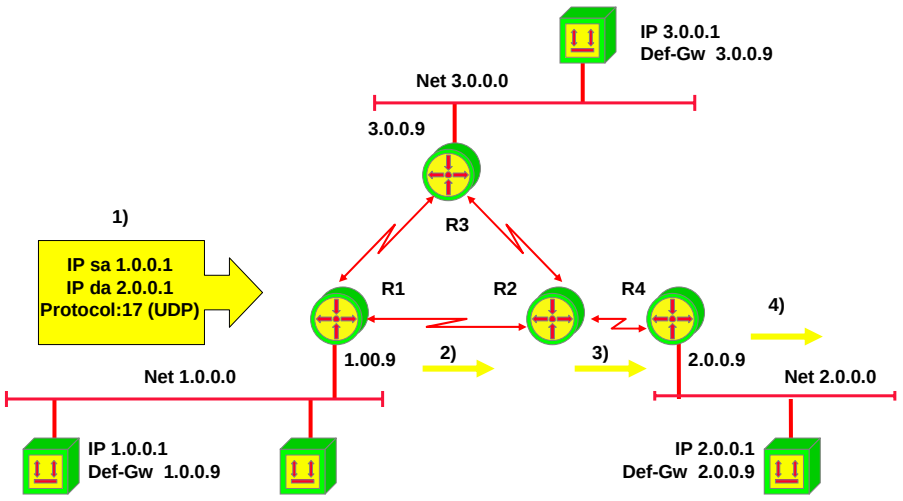


© 2007, D.I. Manfred Lindner

ICMP host unreachable

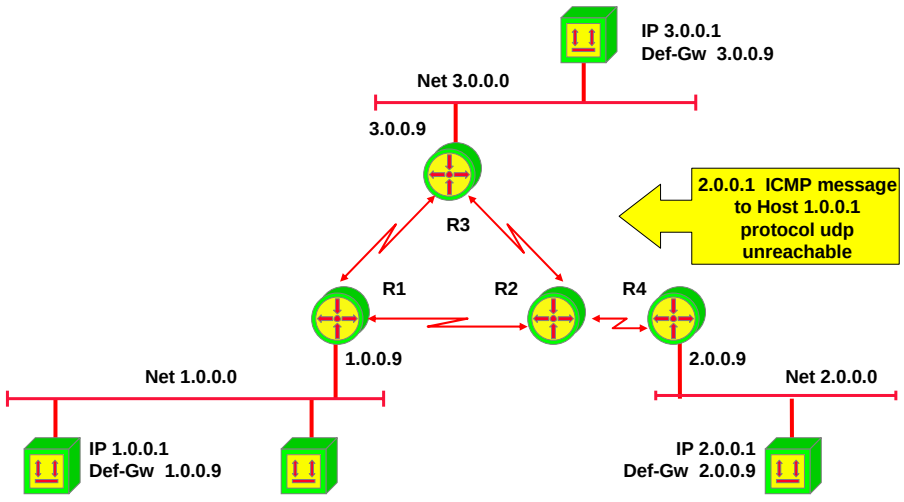


Delivery 1.0.0.1 - > 2.0.0.1 (protocol udp)



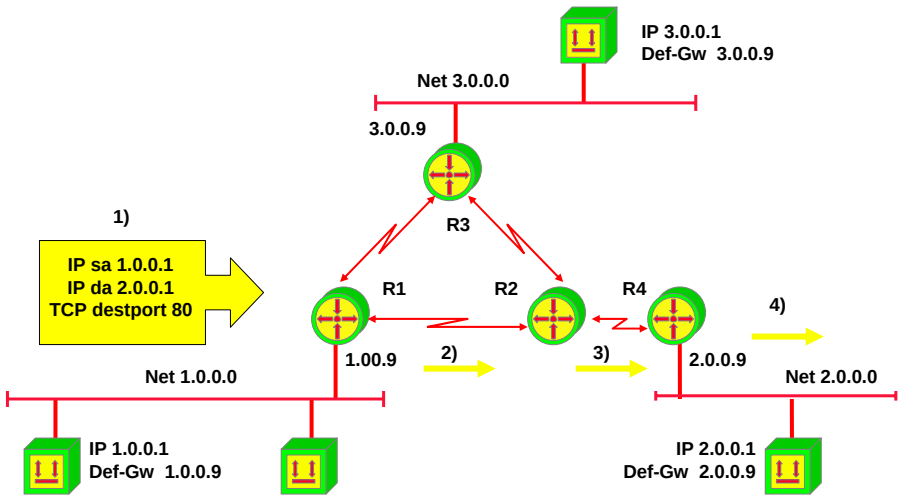
© 2007, D.I. Manfred Lindner

ICMP protocol unreachable



© 2007, D.I. Manfred Lindner IP Details, v4.7 45

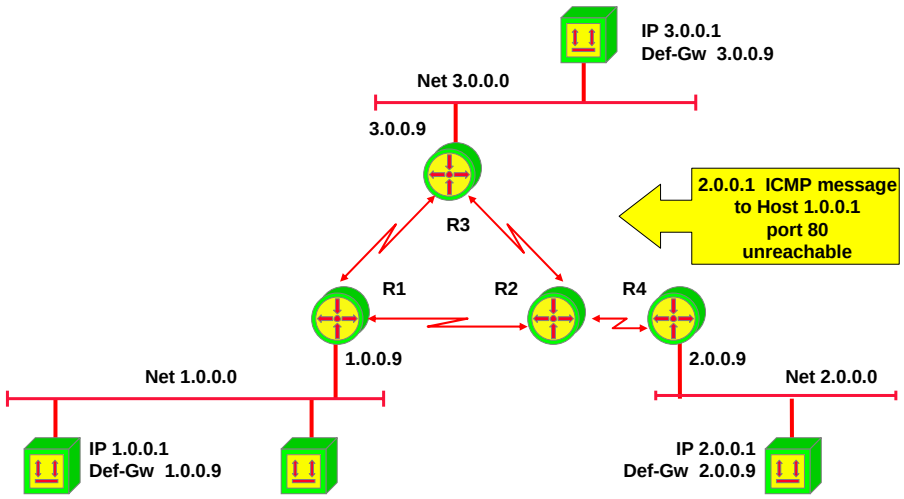
Delivery 1.0.0.1 - > 2.0.0.1 (http_server_proc)



© 2007, D.I. Manfred Lindner IP Details, v4.7 46

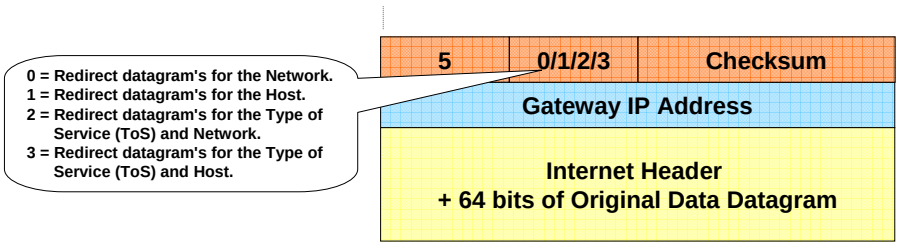
© 2007, D.I. Manfred Lindner

ICMP port unreachable (no http_server_proc)



Code Field for Type 5 (Redirect)

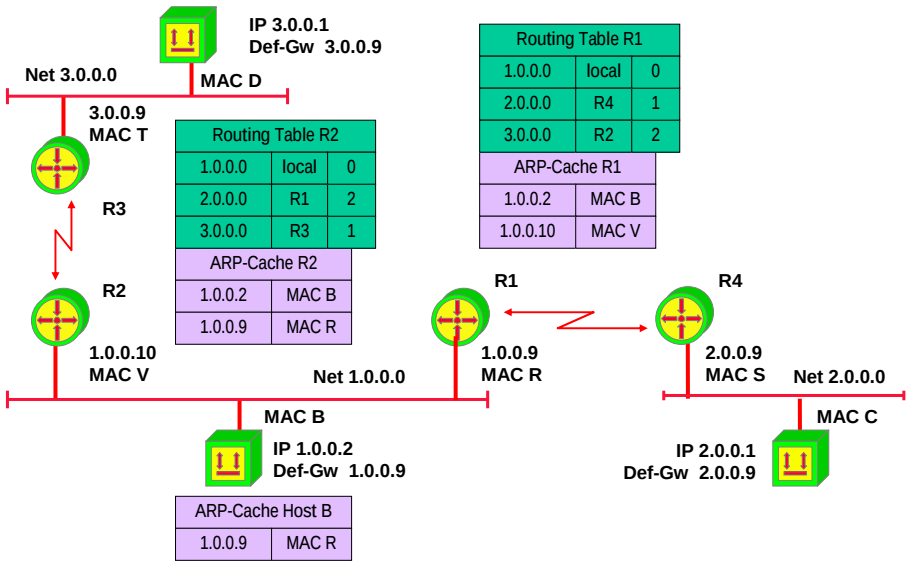
- **If a router knows of a better (faster, shorter) path to a target then it will notify the sender through ICMP redirect**
  - In any case the router will still forward the packets on the inefficient path
  - Datagram's will be sent twice through a LAN, if the sender ignores the redirect message



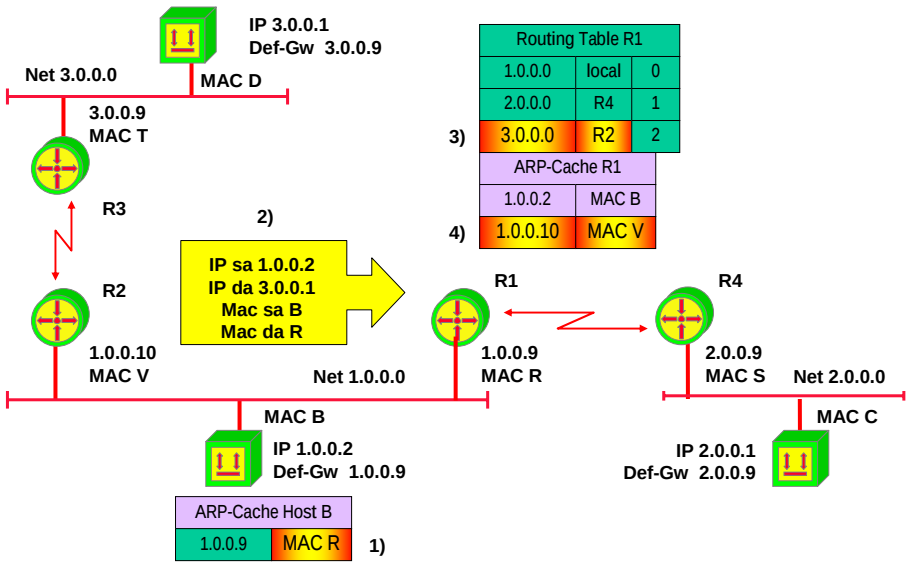
© 2007, D.I. Manfred Lindner

L31 - IP Technology Details

Delivery 1.0.0.2 -> 3.0.0.1

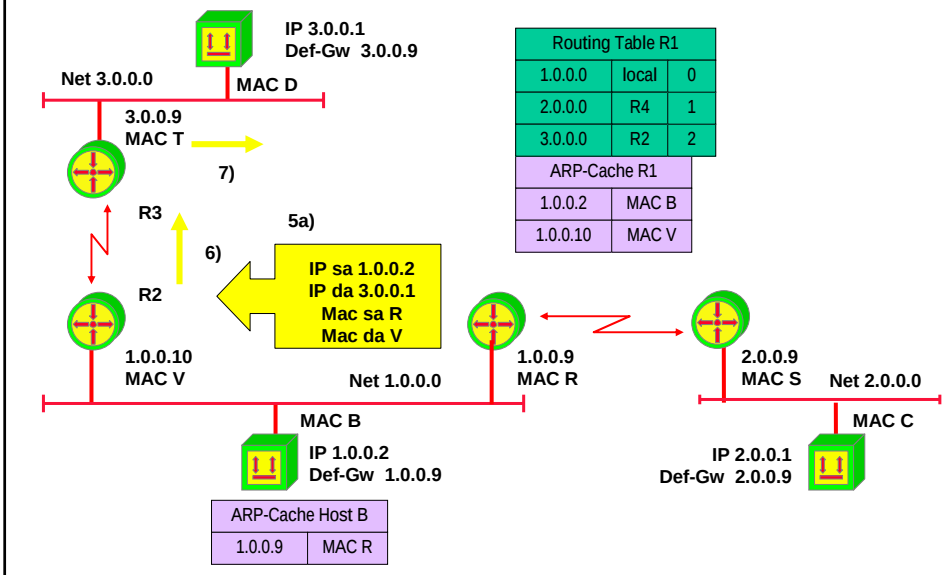


Delivery 1.0.0.2 -> 3.0.0.1

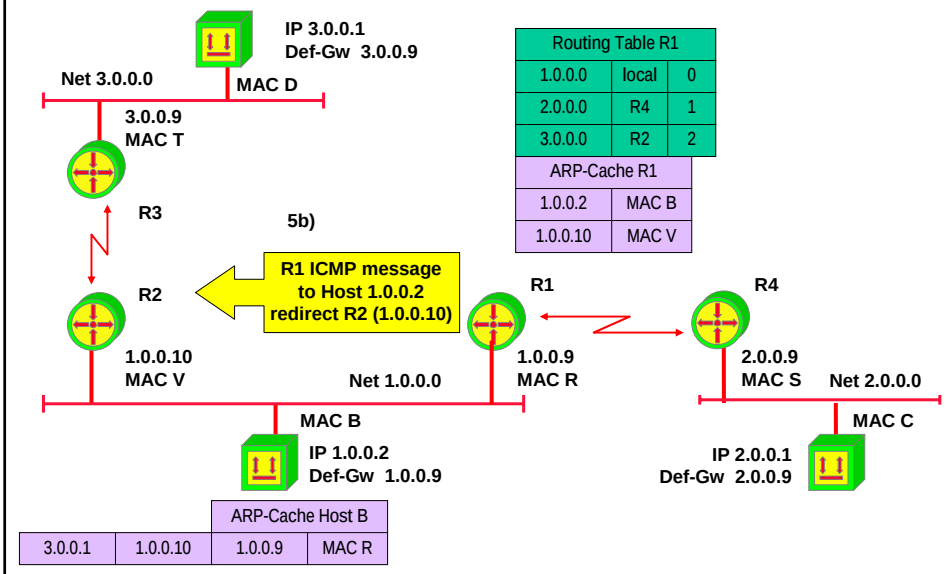


© 2007, D.I. Manfred Lindner

Delivery 1.0.0.2 -> 3.0.0.1

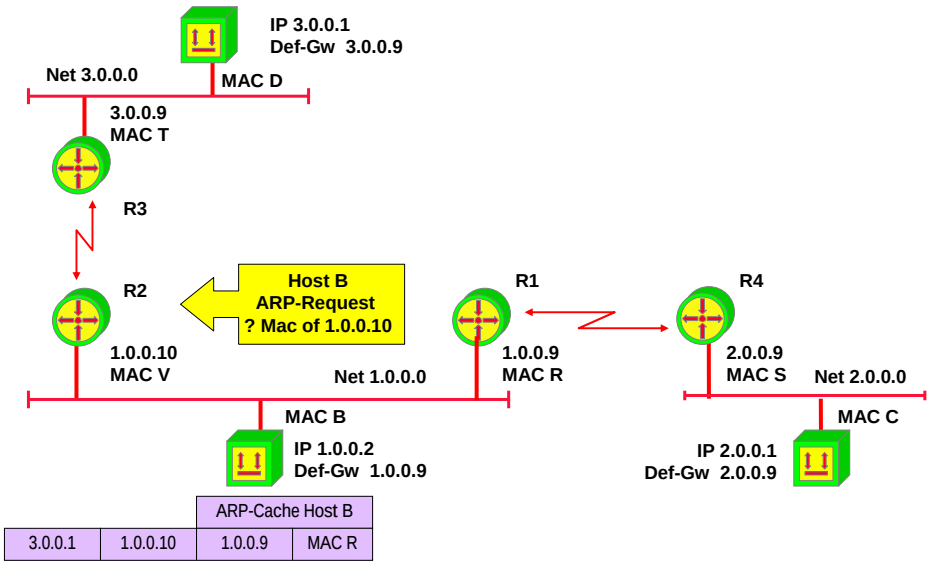


ICMP redirect

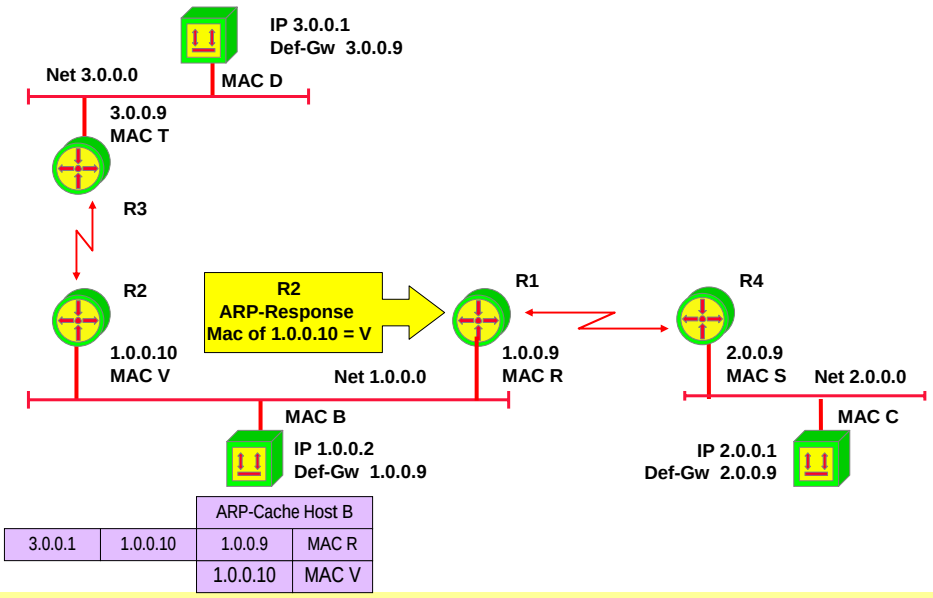


© 2007, D.I. Manfred Lindner

Delivery 1.0.0.2 -> 3.0.0.1

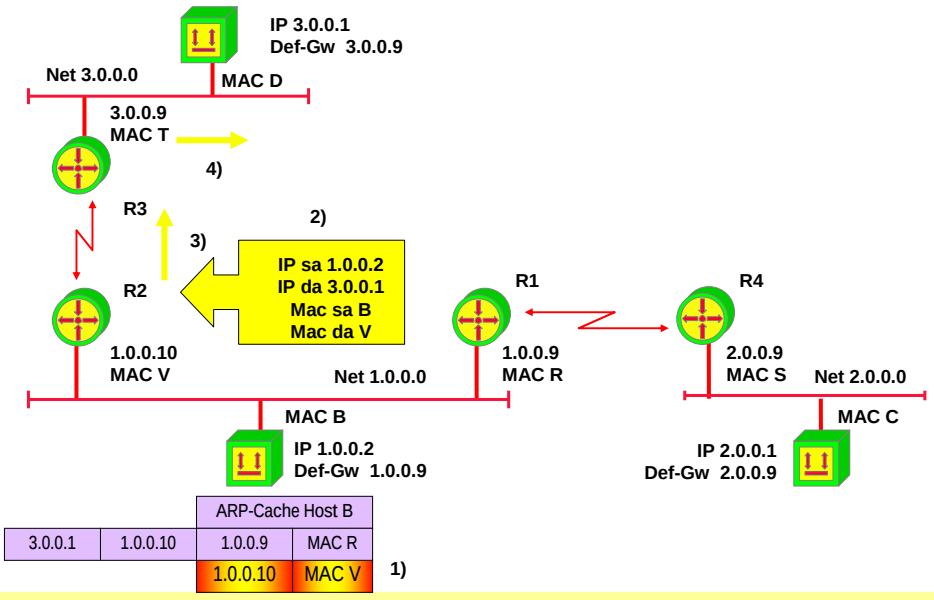


Delivery 1.0.0.2 -> 3.0.0.1



© 2007, D.I. Manfred Lindner

Next Packet 1.0.0.2 -> 3.0.0.1



Agenda

- IP Protocol
- ICMP
- Ping and Traceroute
- Address Resolution Protocol ARP
- RARP
- Proxy ARP
- VRRP, HSRP
- SLIP
- PPP

© 2007, D.I. Manfred Lindner

PING - Packet Internet Groper

- **Checks the reachability of an IP station.**
- **Measures time (round-trip-delay).**
- **Example:**
  - ping 132.105.56.3 (with IP address)
  - ping www.something.at (with a symbolic name, DNS)
- **If the station can be reached:**
  - 132.105.56.3 is alive
- **If no reply arrives within the timeout:**
  - no answer from 132.105.56.3

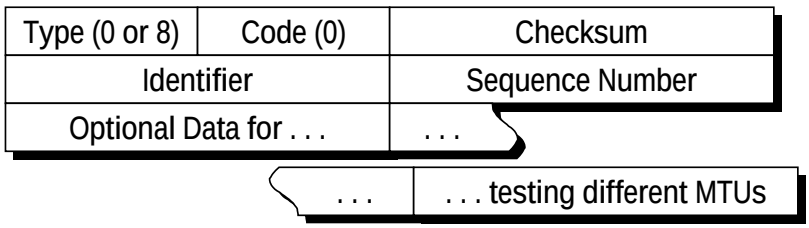
Ping Implementations

- **On most UNIX machines**
  - ping -s 132.105.56.3
- **sends continuous ICMP Echo Requests**
- **Ping operates in several modes (depending on the implementation), sending:**
  - continuous ECHO Requests every second; a final statistic is printed with Ctrl-C.
  - a single ECHO Request with 20 seconds timeout.
  - a fixed number of ECHO Request.

© 2007, D.I. Manfred Lindner

Ping - Message Format

- **ICMP type 8 (request) and 0 (reply)**
  - A replying station changes the type from 8 to 0, swaps IP source and destination and sends the packet back
  - Unix uses the process number as an Identifier
- **Ping of death**
  - giant ping causes some stations to crash



Typical Ping Options

-c	count:	Number of ECHO requests to send.
-i	wait:	Time between requests (1s).
-n	numeric:	IP addresses are shown numerically.
-q	quiet:	Only starting- and ending-line are shown
-s	packet size:	Size of IP packet (default 56 byte -> 32 Byte of data)
-v	verbose:	Other ICMP packets are shown.

© 2007, D.I. Manfred Lindner

Traceroute

- **Lists the exact route, a packet will take through the network**
- **UDP segment and manipulation of the TTL field (time to live) of the corresponding IP header is used**
  - to generate ICMP error messages
    - TTL exceeded
    - port not reachable
- **UDP segments with undefined port number (> 30000)**
  - Echo requests with TTL manipulation only can't be used because after reaching the final IP host no TTL exceeded message will be generated (done by routers only)

Traceroute - Operation

- **UDP datagram with TTL=1 is sent**
- **UDP datagram with TTL=2 is sent**
  .....
- **The routers in the path generate ICMP time exceeded messages because TTL reaches 0**
- **If the UDP datagram arrives at the destination, an ICMP port unreachable message is generated**
- **From the source addresses in the ICMP messages the path can be reconstructed**
- **The IP addresses are resolved to names through DNS**

© 2007, D.I. Manfred Lindner

Traceroute - Sample Output

```
tracert 140.252.13.65

1 bsdi (140.252.13.35)     20ms    10ms    10ms
2 slip  (140.252.13.65)       *   120ms   120ms
```

3 Packets are sent for each TTL value.
Output of "*", if no answer arrives within 5 seconds.

© 2007, D.I. Manfred Lindner

Agenda

- IP Protocol
- ICMP
- Ping and Traceroute
- Address Resolution Protocol ARP
- RARP
- Proxy ARP
- VRRP, HSRP
- SLIP
- PPP

© 2007, D.I. Manfred Lindner

© 2007, D.I. Manfred Lindner

ARP (Address Resolution Protocol)

- An IP address identifies the logical access to an IP network
- The station can be reached without any further addressing, if the physical network consists only of a point-to-point connection
- On a shared media LAN MAC addresses are used to deliver packets to a specific station
- A mapping between IP address and MAC address is needed
- RFC 826

ARP Operation 1

- The mapping between MAC- and protocol-address on a LAN can be static (table entries) or dynamic (ARP protocol and ARP cache)
- Operation of ARP:
  - Station A wants to send to station B and doesn't know the MAC address (both are connected to the same LAN)
  - A sends an ARP request in form of a MAC broadcast (dest. = FF, source = Mac_A), ARP request holds IP address of B
  - Station B sees the ARP request with its IP address and sends an ARP reply as a MAC frame (SA=Mac_B, DA=Mac_A), B puts the newly learned mapping (source MAC- and IP-address of A) into its ARP cache

© 2007, D.I. Manfred Lindner

ARP Operation 2

- The ARP reply holds MAC address of station B
- A stores the MAC- / IP-address mapping for station B in its ARP cache
- For subsequent packets from A to B or from B to A the MAC addresses are taken from the ARP cache (no further ARP request / reply)
- Entries in the ARP cache are deleted if they aren't used for a defined period (usually 5 min), this aging mechanism allows for changes in the network and saves table space
- ARP requests / reply are sent in Ethernet II (Type field 0x0806) or SNAP frames
- ARP has been designed to support different layer 3 protocols

© 2007, D.I. Manfred Lindner

ARP Request/Reply Format

Hardware		Protocol (IP = 0x0800)
hln	pln	Operation
Source Hardware Address (byte 0 - 3)		
Source HW Addr. (byte 4 - 5)		Source IP Addr. (byte 0 - 1)
Source IP Addr. (byte 2 - 3)		Dest. HW Addr. (byte 0 - 1)*
Destination Hardware Address (byte 2 - 5)*		
Destination IP Address (byte 0 - 3)		

*) Destination hardware address is left empty (hex FF FF FF FF FF FF) for ARP request.

© 2007, D.I. Manfred Lindner

© 2007, D.I. Manfred Lindner

ARP Request/Reply Fields

- **Hardware**
  - Defines the type of network hardware, e.g.:
    1 Ethernet DIX
    6 802.x-LAN
    7 ARCNET
    11 LocalTalk
- **Protocol**
  - Selects the layer 3 protocol (uses the values which are defined for the Ethernet type field, e.g. 0x800 for IP)
- **hln**
  - Length of hardware address in bytes

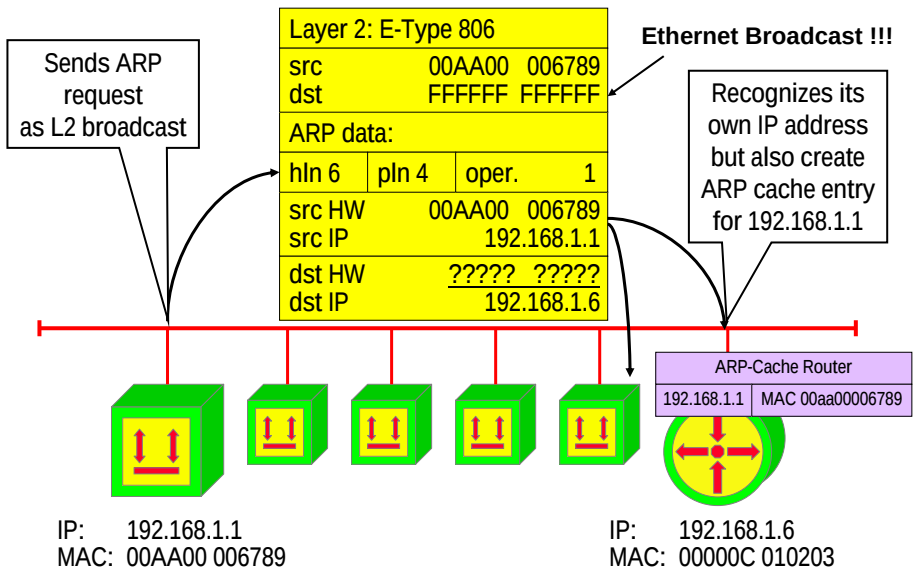
ARP Request/Reply Fields

- **pln**
  - Length of layer 3 address in bytes
- **Operation**
  - 1 .... ARP Request
  - 2 .... ARP Reply
  - 3 .... RARP Request
  - 4 .... RARP Reply
- **Addresses**
  - Hardware addresses: MAC addresses (src. and dest.)
  - IP addresses: layer 3 addresses (src. and dest.)
- **ARP request and replies are not forwarded by routers (only LAN broadcast used)**

© 2007, D.I. Manfred Lindner

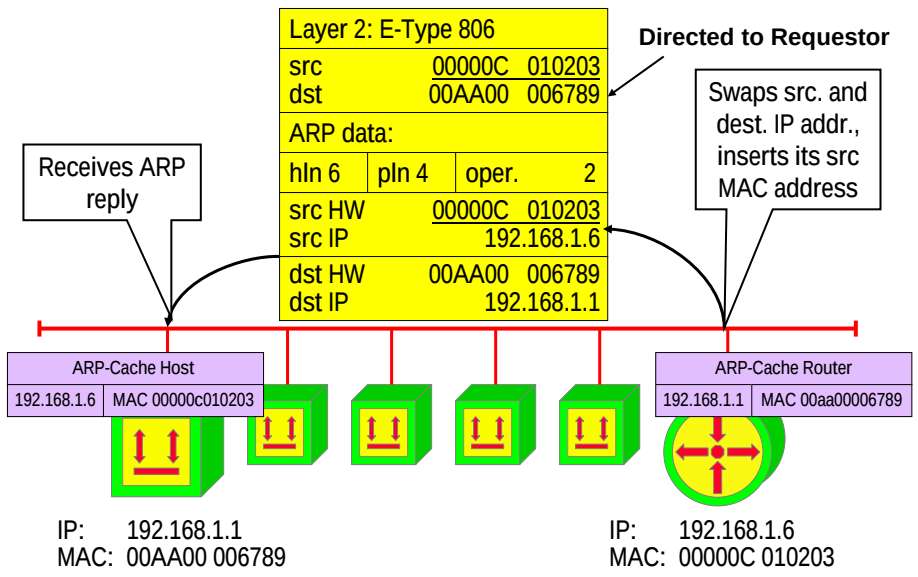
L31 - IP Technology Details

ARP Request



© 2007, D.I. Manfred Lindner IP Details, v4.7 71

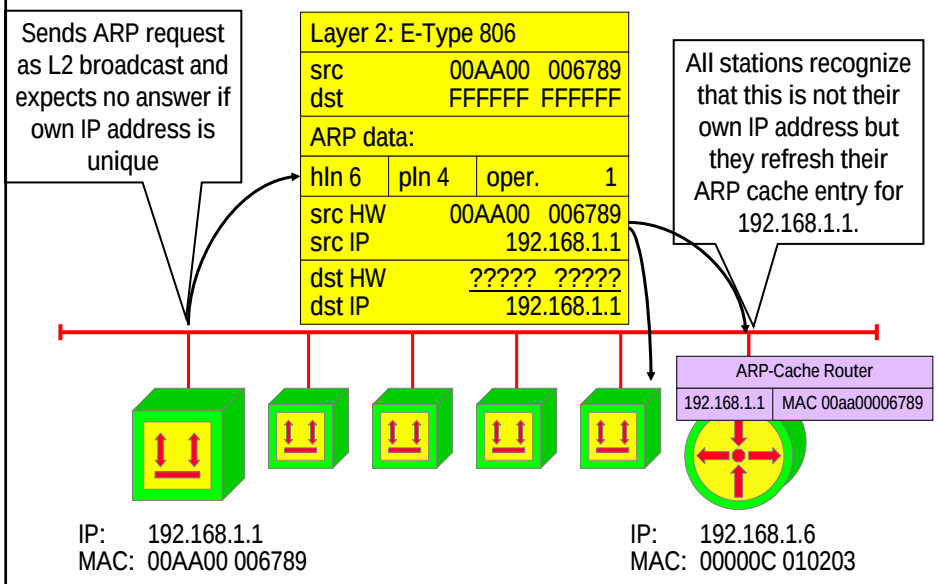
ARP Reply



© 2007, D.I. Manfred Lindner IP Details, v4.7 72

© 2007, D.I. Manfred Lindner

Gratuitous ARP for Duplicate Address Check and ARP Cache Refresh



Agenda

- IP Protocol
- ICMP
- Ping and Traceroute
- Address Resolution Protocol ARP
- RARP
- Proxy ARP
- VRRP, HSRP
- SLIP
- PPP

© 2007, D.I. Manfred Lindner

RARP (Reverse ARP, RFC 903)

- ARP assumes, that an IP station knows its IP address (stored in NVRAM, on hard disk, in config file etc.)
- Diskless Machines usually don't have such means so they must retrieve an IP address for network booting
- RARP (Reverse ARP) provides IP addresses for unconfigured stations

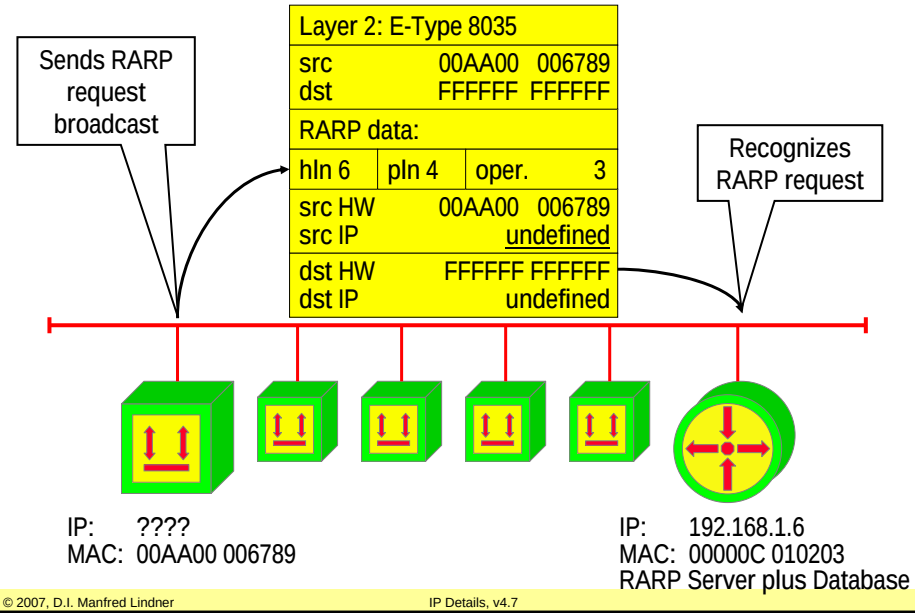
RARP Operation

- A station sends a RARP request broadcast
- One station, the RARP server, looks up the IP address for that MAC address in a database and replies
- Newer methods:
  - BOOTP
  - DHCP

© 2007, D.I. Manfred Lindner

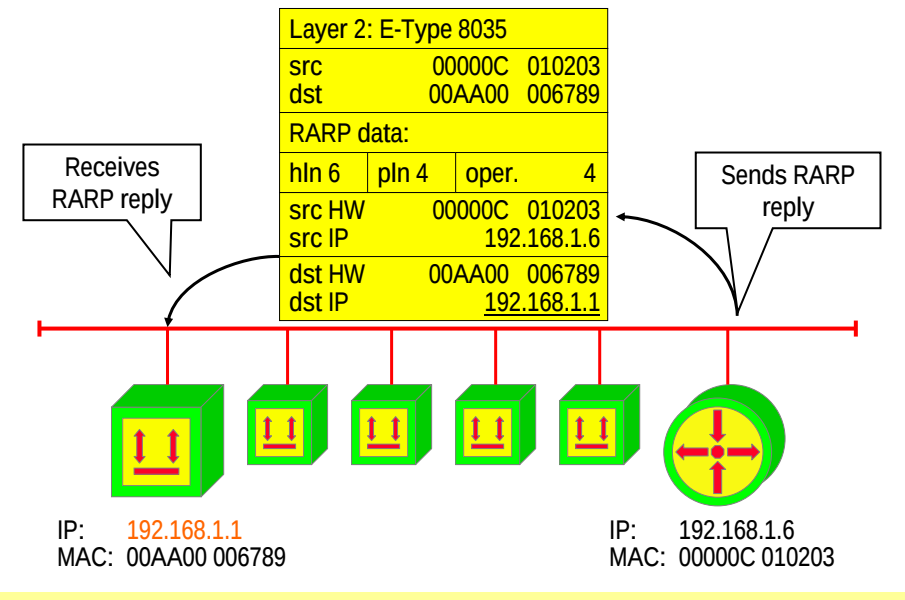
L31 - IP Technology Details

RARP Request



© 2007, D.I. Manfred Lindner IP Details, v4.7 77

RARP Reply



© 2007, D.I. Manfred Lindner IP Details, v4.7 78

© 2007, D.I. Manfred Lindner

Agenda

- IP Protocol
- ICMP
- Ping and Traceroute
- Address Resolution Protocol ARP
- RARP
- Proxy ARP
- VRRP, HSRP
- SLIP
- PPP

Proxy ARP

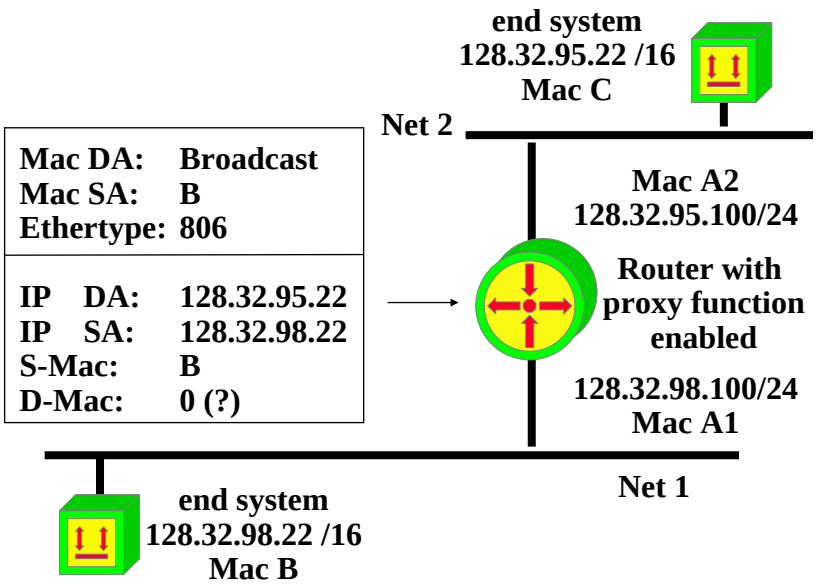
- Networks, which are connected through a router
  - need different Net-IDs
- If two networks coupled by a router need to have the same Net-ID
  - e.g. for the time a bridged network should be migrated to a routed network
  - a proxy ARP component must be installed
- Term “proxy” means “instead of”
  - some system is doing some function instead of the expected system

© 2007, D.I. Manfred Lindner

Proxy ARP

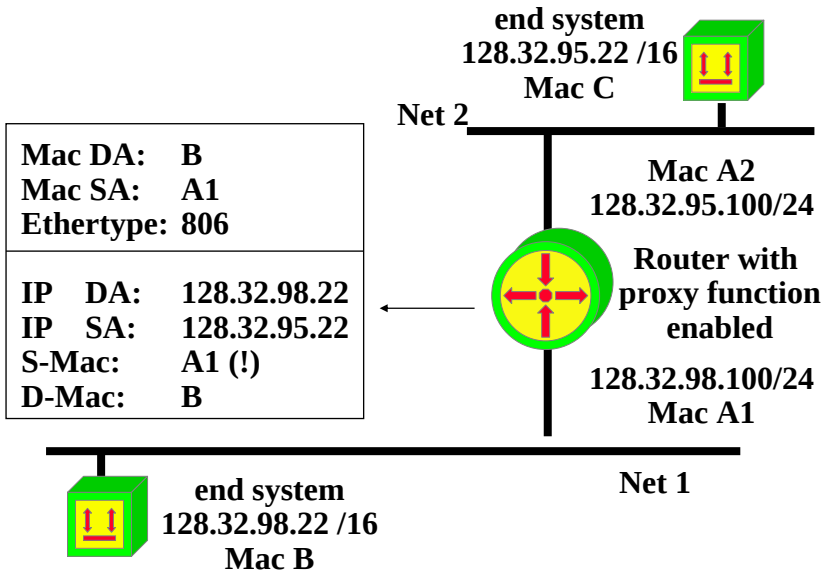
- **In such as migration phase**
  - Proxy ARP gives IP Hosts on both sides the impression of a homogenous network (single LAN)
- **ARP requests for a station on the "other side“**
  - are responded to on behalf of the destination with the MAC address of the Proxy system (the router in our case)
- **IP datagram's will be sent to the layer 2 address (MAC) of the Proxy system**
  - who forwards them to the destination network
- **Important ARP rule:**
  - Normal IP Protocol stack sends out an ARP Request for stations located on same subnet (same NET-ID) only !!!

ARP Request Net 1



© 2007, D.I. Manfred Lindner

ARP Response Net 1



Proxy ARP Usage

- **Old method for efficient use of address space**
- **Replaced by subnetting**
- **Differences to subnetting:**
  - Only Proxy ARP needs to know about the network structure, e.g. static tables or through interpretation of the Host-ID with rules similar to subnetting.
  - End systems have no knowledge of the network / address structure -> transparent network
  - End systems don't have to care about subnet masks.
- **Still useful for security ("demilitarized areas"), especially for Internet access and employment of firewalls**

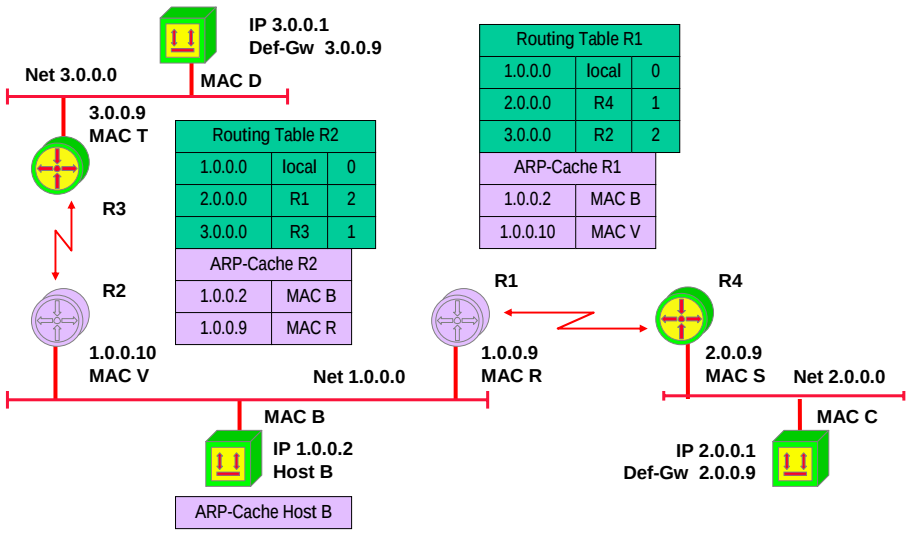
© 2007, D.I. Manfred Lindner

Proxy ARP Usage Nowadays

- **Proxy ARP is also be used if an IP host didn't know the address of the default gateway:**
  - In an IP host normally a static entry will tell the IP address of the router
    - if an IP datagram has to be sent to a non-local Net-ID, an ARP request will find the MAC address of the default gateway
  - With Proxy ARP extensions in the IP host and in the router
    - the MAC address of the router can be found without knowing the routers IP address
    - An ARP request will be sent for IP hosts with NET-IDs different from the local Net-ID and the router will respond
  - With Unix stations or Windows NT/XP:
    - proxy ARP extensions are triggered by setting the default gateway to the systems IP address itself

© 2007, D.I. Manfred Lindner

1.0.0.2 -> 3.0.0.1 / 2.0.0.1 with proxy ARP 1



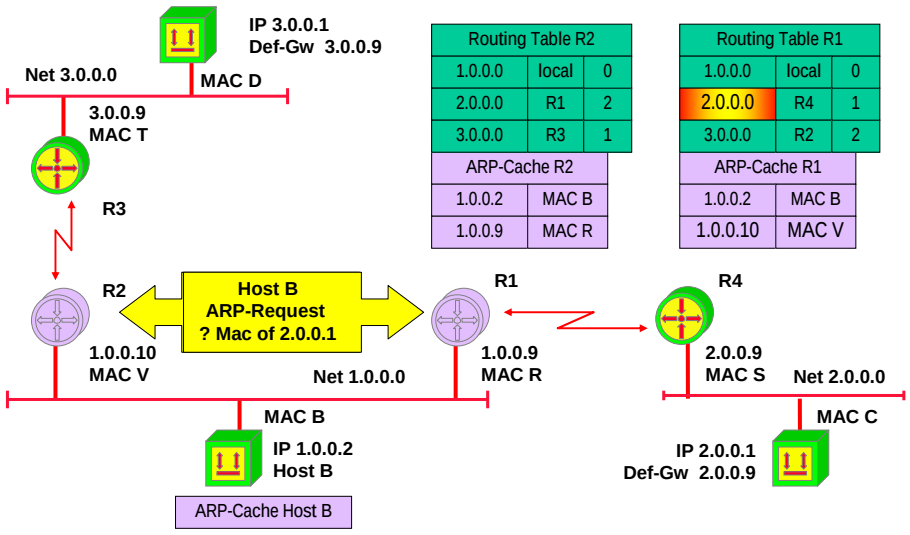
R1 and R2 proxy ARP enabled; Host B sends ARP also for net-ID unequal own net-ID

© 2007, D.I. Manfred Lindner

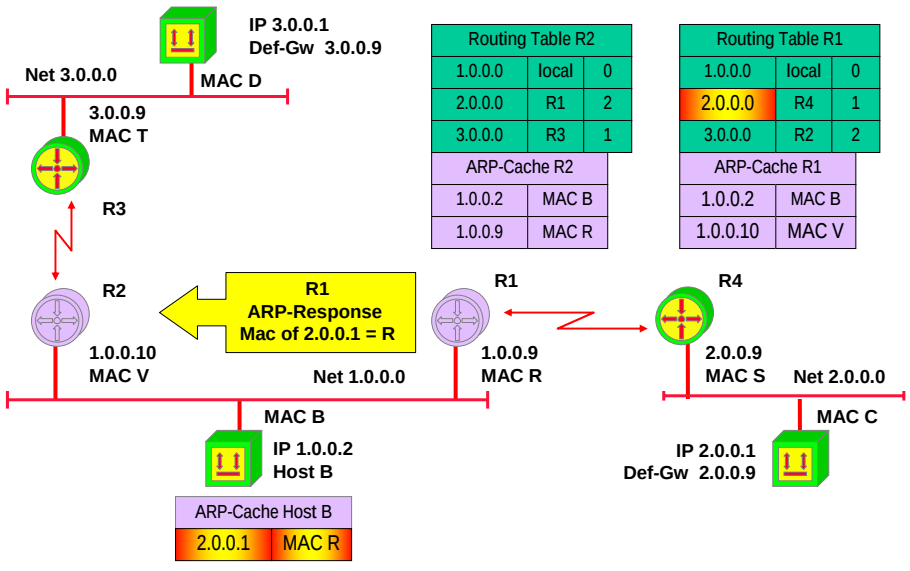
© 2007, D.I. Manfred Lindner

L31 - IP Technology Details

1.0.0.2 -> 3.0.0.1 / 2.0.0.1 with proxy ARP 2



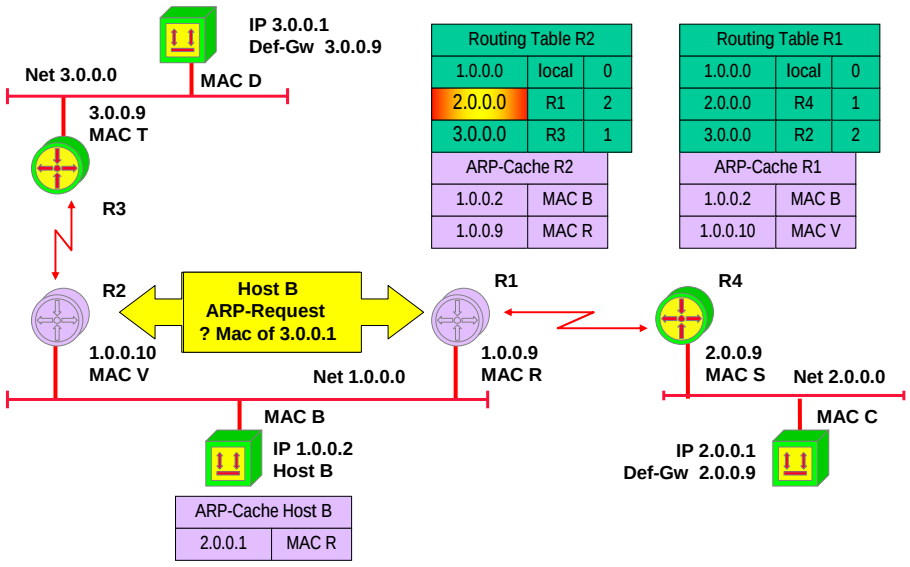
1.0.0.2 -> 3.0.0.1 / 2.0.0.1 with proxy ARP 3



© 2007, D.I. Manfred Lindner

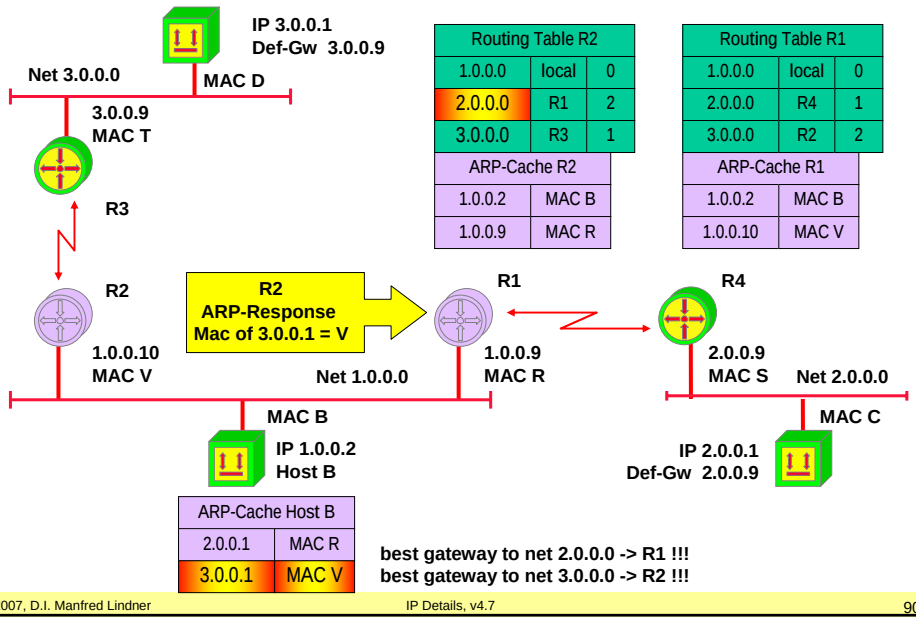
L31 - IP Technology Details

1.0.0.2 -> 3.0.0.1 / 2.0.0.1 with proxy ARP 4



© 2007, D.I. Manfred Lindner IP Details, v4.7 89

1.0.0.2 -> 3.0.0.1 / 2.0.0.1 with proxy ARP 5



© 2007, D.I. Manfred Lindner IP Details, v4.7 90

© 2007, D.I. Manfred Lindner

Agenda

- IP Protocol
- ICMP
- Ping and Traceroute
- Address Resolution Protocol ARP
- RARP
- Proxy ARP
- VRRP, HSRP
- SLIP
- PPP

Other Techniques to Solve the Problem 1

- IDRP
  - ICMP Router Discovery Messages (RFC 1256)
  - Routers periodically advertise their IP address on a shared media together with an preference value and a lifetime
    - ICMP Router Advertisement Message
  - Hosts may listen to these messages to find out all possible Default Gateways
    - or may ask by sending an ICMP Router Solicitation Message
- DHCP
  - Dynamic Host Configuration Protocol (RFC 2131)
  - More than one Default Gateway can be specified
  - Every Default Gateway has a preference value

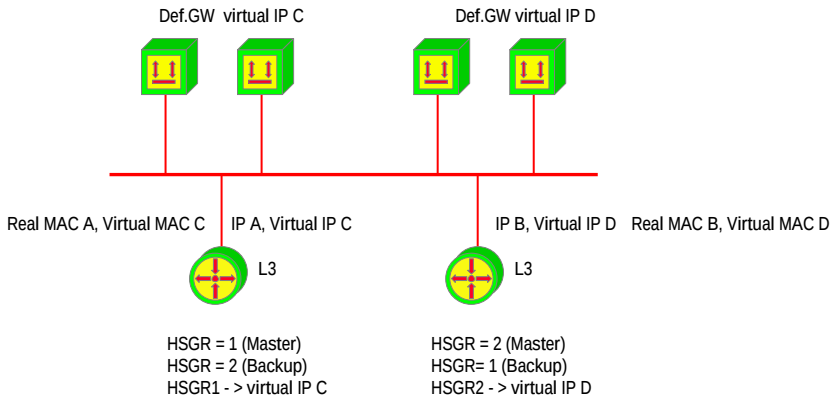
© 2007, D.I. Manfred Lindner

Other Techniques to Solve the Problem 2

- **HSRP**
  - Hot Standby Routing Protocol (Cisco, RFC 2281)
  - Default Gateway in IP host configured with IP address of a Virtual Router
    - Virtual Router listens to virtual MAC Address
  - Real routers work together and spoof Virtual Router
- **VRRP**
  - Virtual Router Redundancy Protocol (RFC 2338)
- **Running a routing protocol on IP host**
  - RIP
  - OSPF

HSRP (Cisco, informational RFC 2281)

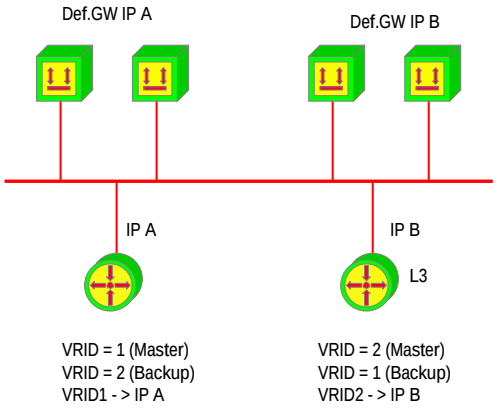
Two Hot Standby Groups (HSGR) defined to allow different Default Gateways (Exit Points of LAN) for IP end systems. End system configured with virtual IP Address. Master router of a given group listens to his virtual IP address and replies with his virtual MAC address on ARP requests. Backup router of a given group supervises the master and can take over if master fails.



© 2007, D.I. Manfred Lindner

VRRP (proposed standard RFC 2338)

With VRRP, the real IP Address/the real MAC address of one router (master) could be supported as virtual IP Address/virtual MAC address of the other router (backup) in case the master fails.



Agenda

- IP Protocol
- ICMP
- Ping and Traceroute
- Address Resolution Protocol ARP
- RARP
- Proxy ARP
- VRRP, HSRP
- SLIP
- PPP

© 2007, D.I. Manfred Lindner

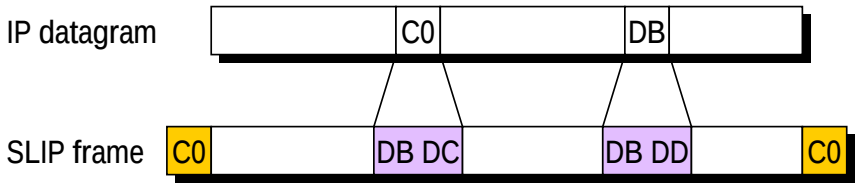
SLIP

- SLIP (Serial Line IP) defines layer 2 frame format for serial lines with two escape characters
  - byte stuffing to avoid SD, ED !!!
- Simple transmission of IP datagrams on serial lines
- Async, character orientated, 8 bit, no parity
- No flow control with XON/XOFF possible
- For the connection of two IP stations
- TCP/IP protocol stack must be implemented
- Suitable for "dial In" modem connections

© 2007, D.I. Manfred Lindner

SLIP Frame Format

- Start and end are delimited with a special character, SLIP END (0xC0)
- SLIP END within the datagram is substituted with SLIP ESC (0xDB) and 0xDC
- SLIP ESC within the datagram is substituted with SLIP ESC (0xDB) and 0xDD



© 2007, D.I. Manfred Lindner

© 2007, D.I. Manfred Lindner

SLIP Disadvantages, SLIP versus PPP

- Both stations have to know the IP address of the other side
- Addresses can't be discovered dynamically
- No Protocol Type Field (like PPP):
- only IP protocols can be transmitted over the line.
- No Checksum
  - error detection and correction has to be done by higher layer protocols

- **SLIP replaced by PPP**
  - Multiple network layer protocols on a serial line
  - CRC (error detection) in every frame
  - Dynamic negotiation of IP addresses (IPCP)
  - Security through password authentication (PAP, CHAP)

Agenda

- **IP Protocol**
- **ICMP**
- **Ping and Traceroute**
- **Address Resolution Protocol ARP**
- **RARP**
- **Proxy ARP**
- **VRRP, HSRP**
- **SLIP**
- **PPP**

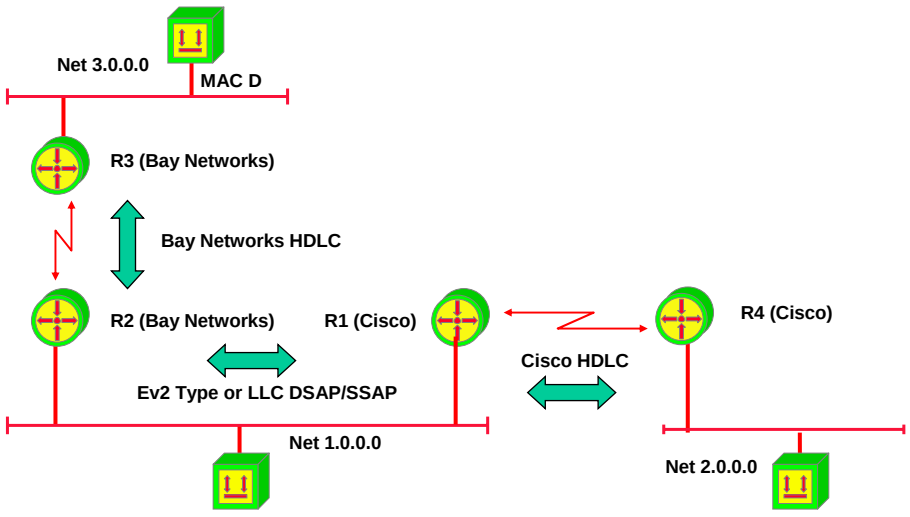
© 2007, D.I. Manfred Lindner

Reasons for Point-to-Point Protocol (PPP)

- **Communication between router of different vendors on a LAN was possible**
  - from the very beginning
    - Remember: Ethernet V2 Protocol Type field or LLC-DSAP/SSAP fields carry information about the protocol stack (e.g. IP or IPX or SAN or NetBEUI or AppleTalk)
- **Communication between router of different vendors on a serial line was not possible**
  - because of the proprietary “kind of HDLC” encapsulation method used by different vendors
- **PPP standardizes multiprotocol encapsulation on a serial line**
  - hence interoperability is the main focus

© 2007, D.I. Manfred Lindner IP Details, v4.7 101

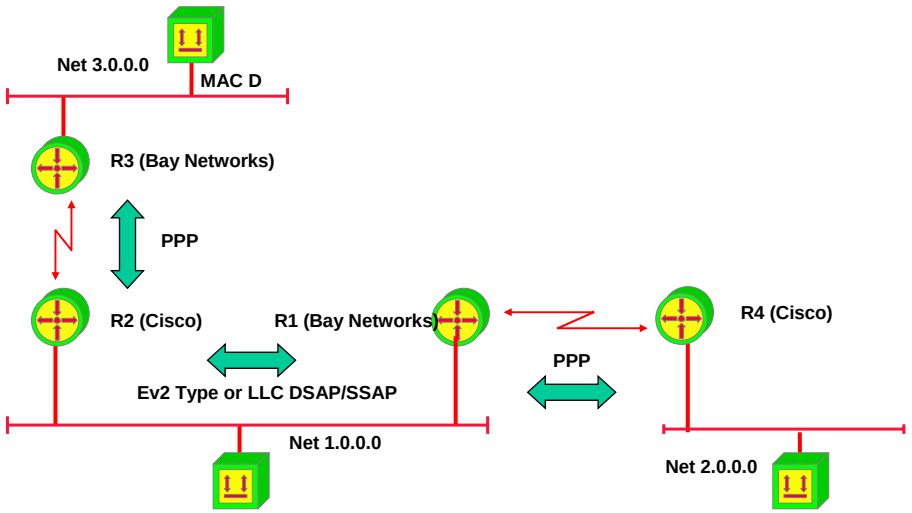
Interoperability without PPP



© 2007, D.I. Manfred Lindner IP Details, v4.7 102

© 2007, D.I. Manfred Lindner

Interoperability with PPP



Today's Main Focus of PPP

- **Providing Dial-In connectivity for IP systems**
  - using modems and Plain Old Telephone Network (POTS)
    - PPP
  - using ISDN
    - PPP over transparent B-channel
  - using ADSL (Asymmetric Digital Subscriber Line)
    - PPPoE (PPP over Ethernet)
    - PPPoA (PPP over ATM)
  - using Dial-In VPN technology
    - Microsoft PPTP (Point-to-Point Tunneling Protocol)
    - Cisco L2F (L2 Forwarding Protocol)
    - L2TP (Layer2 Tunneling Protocol), RFC

© 2007, D.I. Manfred Lindner

PPP Overview

- **data link protocol (L2)**
- **used to encapsulate network layer datagram's or bridged packets (multiprotocol traffic)**
  - over serial communication links in a well defined manner
- **connectionless service**
  - although we speak about a PPP connection, details are provided later
- **symmetric point-to-point protocol**
- **industry standard for dial-in service**
  - used for interoperability, even over leased lines
  - finally displacing SLIP (serial line IP) in the field
- **supports the simultaneous use of network protocols**

PPP Components

- **three major components**
  - HDLC framing and encapsulation (RFC 1662)
    - bitstuffing for synchronous serial lines
    - modified bytestuffing for asynchronous serial
    - only connectionless service used (UI frame)
  - Link Control Protocol (LCP, RFC 1661)
    - establishes and closes the PPP connection / PPP link
    - tests the link for quality of service features
    - negotiation of parameters
    - configures the PPP connection / PPP link
  - family of Network Control Protocols (NCP, div. RFCs)
    - configures and maintains network layer protocols
    - NCP´s exist for IP, OSI, DECnet, AppleTalk, Novell
    - NCP´s are started after PPP link establishment through LCP

© 2007, D.I. Manfred Lindner

PPP Frame Format

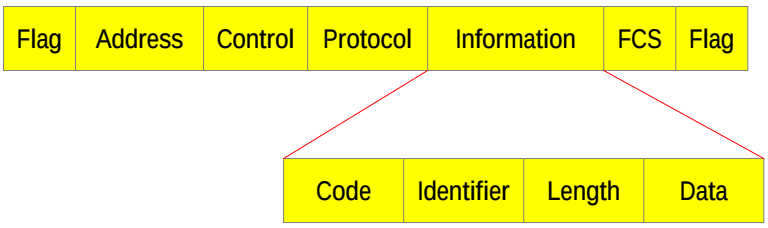
Flag = 01111110
Address = 11111111
Control = 00000011 (UI frame)

Protocol = see RFC 1700 (assigned numbers)
Information= Network Layer PDU
FCS = 16 bit

- **some protocol fields**
  - 0021 Internet Protocol
  - 0029 AppleTalk
  - 8021 IP Control Protocol
  - 8029 AppleTalk Control Prot.
  - C021 Link Control Protocol
  - C223 Authentication CHAP
  - 0027 DECnet Phase 4
  - 002B Novell IPX
  - 8027 DECnet Control Protocol
  - 802B IPX Control Protocol
  - C023 Authentication PAP

© 2007, D.I. Manfred Lindner

Link Control Protocol (LCP) Frame Format



- **carried in PPP information field**
  - protocol field has to be 0xC021
  - code field indicates type of LCP packet
  - identifier field is used to match requests and replies
  - data field values are determined by the code field (e.g. contains options to be negotiated)

© 2007, D.I. Manfred Lindner

© 2007, D.I. Manfred Lindner

Types of LCP Packets

- **There are three classes of LCP packets:**
  - class 1: Link Configuration packets used to establish and configure a link
    - Configure-Request (code 1, details in option field), Configure-Ack (code 2), Configure-Nak (code 3, not supported option) and Configure-Reject (code 4, not supported option)
  - class 2: Link Termination packets used to terminate a link
    - Terminate-Request (code 5) and Terminate-Ack (code 6)
  - class 3: Link Maintenance packets used to manage and debug a link
    - Code-Reject (code 7, unknown LCP code field), Protocol-Reject (code 8, unknown PPP protocol field), Echo-Request (code 9), Echo-Reply (code 10) and Discard-Request (code 11)

© 2007, D.I. Manfred Lindner

LCP and PPP Connection

- **LCP**
  - supports the establishment of the PPP connection and allows certain configuration options to be negotiated
- **PPP connection is established in four phases**
  - phase 1: link establishment and configuration negotiation
    - done by LCP (note: deals only with link operations, does not negotiate the implementation of network layer protocols)
  - phase 2: optional procedures that were agreed during negotiation of phase 1 (e.g. CHAP authentication or compression)
  - phase 3: network layer protocol configuration negotiation done by corresponding NCP´s
    - e.g. IPCP, IPXCP, …
  - phase 4: link termination

© 2007, D.I. Manfred Lindner

© 2007, D.I. Manfred Lindner

PPP Phases

- **task of phase 1**
  - LCP is used to automatically
    - agree upon the encapsulation format options
    - handle varying limits on sizes of packets
    - detect a looped-back link and other common configuration errors (magic number for loopback detection)
  - options which may be negotiated
    - maximum receive unit
    - authentication protocol
    - quality protocol
    - Protocol-Field-Compression
    - Address-and-Control-Field-Compression
    - these options are described in RFC 1661 (except authentication protocols)

PPP Phases

- **task of phase 1 (cont.)**
  - options which may be negotiated but implementations are specified in other RFCs
    - PPP link quality protocol (RFC 1989)
    - PPP compression control protocol (RFC 1962)
    - PPP compression STAC (RFC 1974)
    - PPP compression PREDICTOR (RFC 1978)
    - PPP multilink (RFC 1990)
    - PPP callback (draft-ietf-pppext-callback-ds-01.txt)
    - PPP authentication CHAP (RFC 1994)
    - PPP authentication PAP (RFC 1334)
    - PPP Extensible Authentication Protocol (EAP), RFC 2284

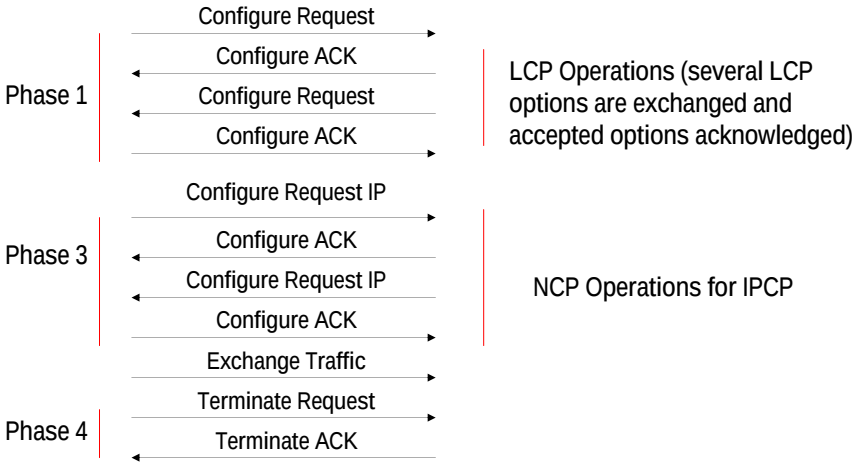
© 2007, D.I. Manfred Lindner

PPP Phases

- **task of phase 2**
  - providing of optional facilities
    - authentication, compression initialization, multilink, etc.
- **task of phase 3**
  - network layer protocol configuration negotiation
    - after link establishment, stations negotiate/configure the protocols that will be used at the network layer; performed by the appropriate network control protocol
    - particular protocol used depends on which family of NCPs is implemented
- **task of phase 4**
  - link termination
    - responsibility of LCP, usually triggered by an upper layer protocol of a specific event

© 2007, D.I. Manfred Lindner

PPP Link Operation Example

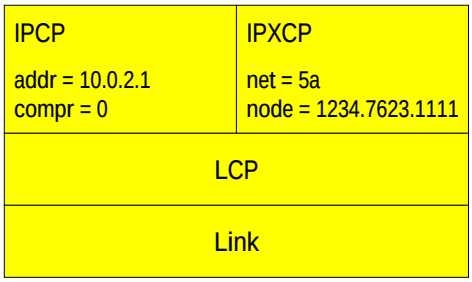


© 2007, D.I. Manfred Lindner

© 2007, D.I. Manfred Lindner

Network Control Protocol

- one per upper layer protocol (IP, IPX…)
- each NCP negotiates parameters appropriate for that protocol
- NCP for IP (IPCP)
  - IP address, Def. Gateway, DNS Server, TTL, TCP header compression can be negotiated
  - Similar functionality as DHCP for LAN



PAP Authentication (RFC 1334)

- **Password Authentication Protocol**
- **simple two way handshake**
- **passwords are sent in clear text**
- **snooping gets you the password**
- **not compatible with bi-directional authorization**

© 2007, D.I. Manfred Lindner

CHAP Authentication RFC 1994

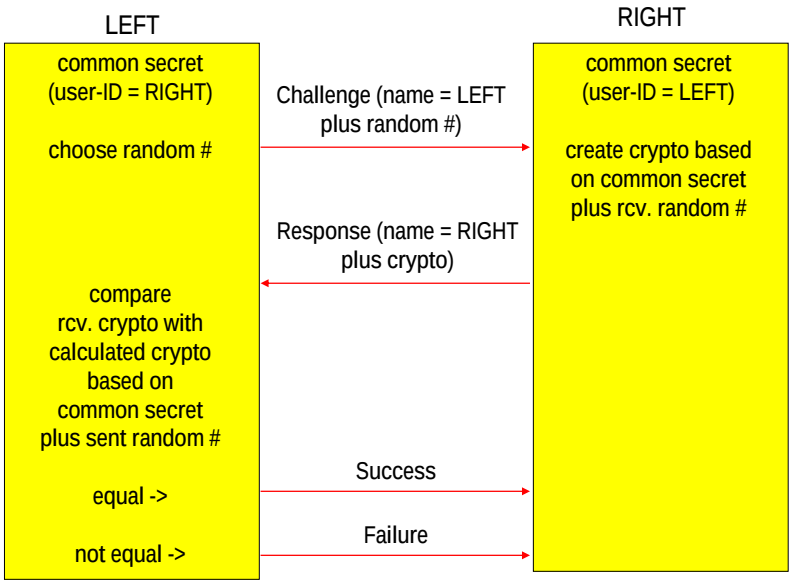
- **Challenge Authentication Protocol**
- **follows establishment of LCP**
- **identifies user**
- **three way handshake**
- **one way authentication only**
  - station which starts (authenticator) the three way handshake proofs authentication of the other station
  - must be configured on both sides if two way authentication is necessary
- **snooping does not discover password**

CHAP Operation

- **three way handshake**
  - PPP link successfully installed by LCP
  - local station sends a challenge message to remote station
  - challenge contain random number and own user-id
  - remote station replies with value using one way hash function based on crypto negotiated for this user-id
  - response is compared with stations own calculation of random number with same crypto
  - if equal success messages is sent to remote station
  - if unequal failure message is sent

© 2007, D.I. Manfred Lindner

CHAP Authentication Procedure

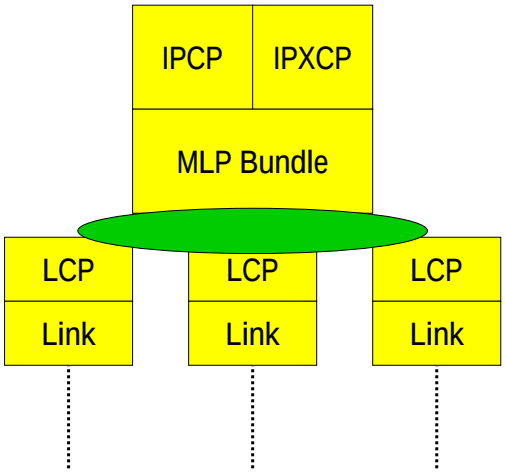


Multilink PPP (RFC 1990)

- **"bundles" bandwidth of multiple interfaces**
  - based on LAPB Multilink procedures (ISO 7776)
- **enabled by LCP negotiation**
- **inactive until authentication**
- **protocol-id 0x003D**
- **single NCP across all links in bundle**
- **large packets are fragmented and sent across links in balanced way**
  - fragments carry sequence number and are tagged with begin and end flags
- **better granularity than packet-level switching**

© 2007, D.I. Manfred Lindner

Multilink PPP

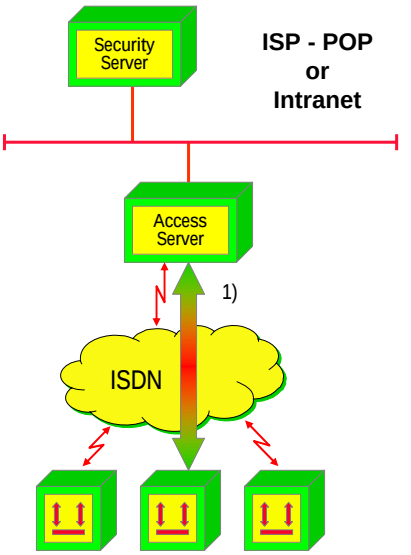


PPP as Dial-In Technology

- **Dial-In:**
  - Into a corporate network (Intranet) of a company
    - Here the term RAS (remote access server) is commonly used to describe the point for accessing the dial-in service
  - Into the Internet by having an dial-in account with an Internet Service Provider (ISP)
    - Here the term POP (point-of-presence) is used to describe the point for accessing the service

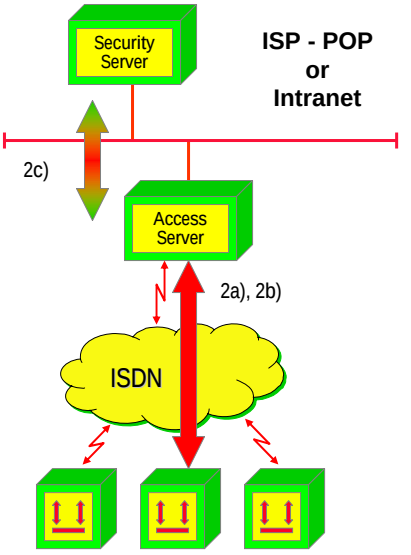
© 2007, D.I. Manfred Lindner

RAS Operation 1



- **remote PC places ISDN call to access server, ISDN link is established (1)**

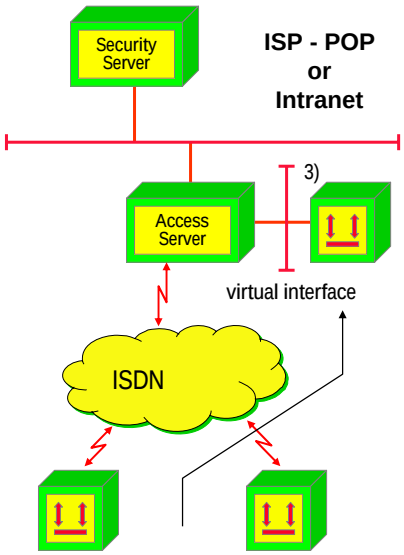
RAS Operation 2



- **PPP link (multiprotocol over serial line) is established**
  - LCP Link Control Protocol (2a)
    - establishes PPP link plus negotiates parameters like authentication CHAP
  - authentication
    - CHAP Challenge Authentication Protocol to transport passwords (2b)
    - verification maybe done by central security server (2c) -> Radius, TACACS, TACACS+

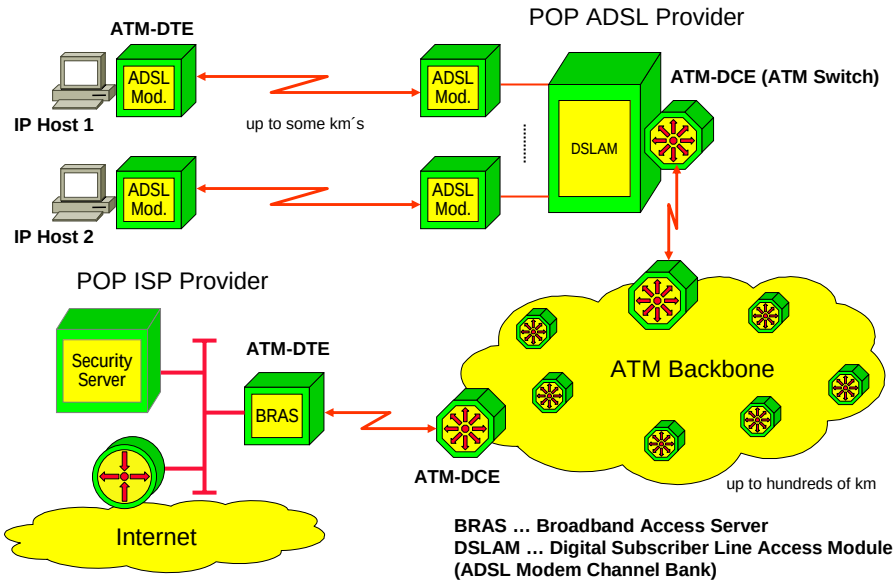
© 2007, D.I. Manfred Lindner

RAS Operation 3



- **PPP NCP (Network Control Protocol) IPCP**
  - assigns IP address, Def. GW, DNS to remote PC
- **remote PC appears as**
  - device reachable via virtual interface (3), IP host Route
- **optionally**
  - filter could be established on that virtual interface
    - authorization
  - accounting can be performed
    - actually done by security server (AAA server)
    - TACACS, Radius

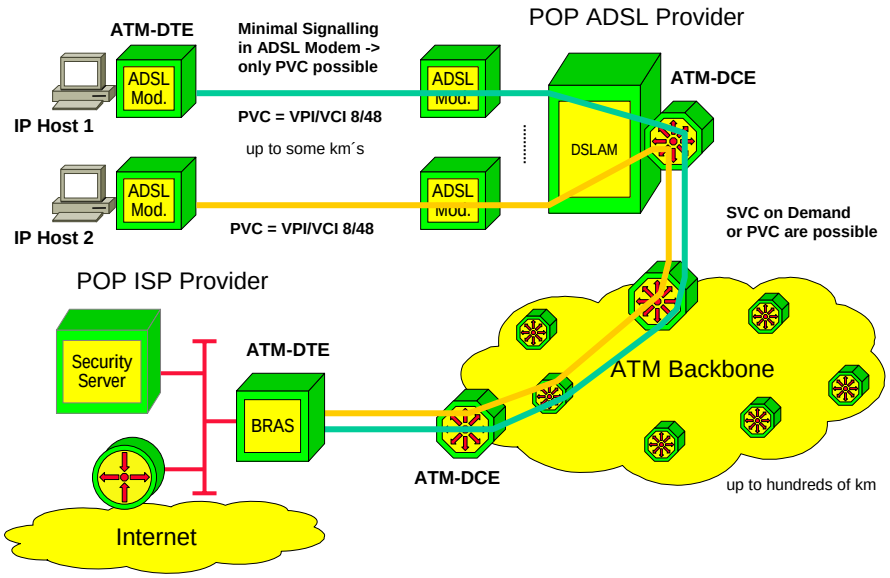
ADSL: Physical Topology



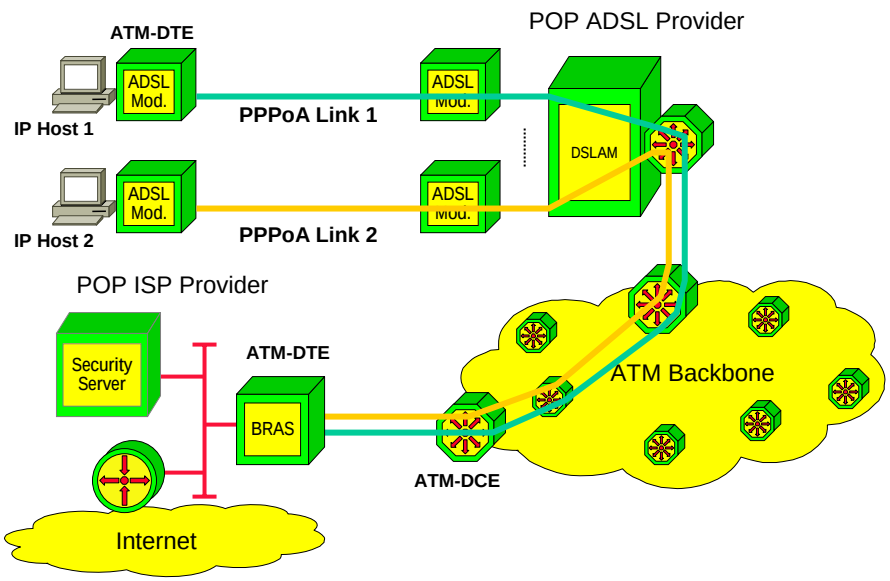
BRAS … Broadband Access Server
DSLAM … Digital Subscriber Line Access Module (ADSL Modem Channel Bank)

© 2007, D.I. Manfred Lindner

ADSL: ATM Virtual Circuits



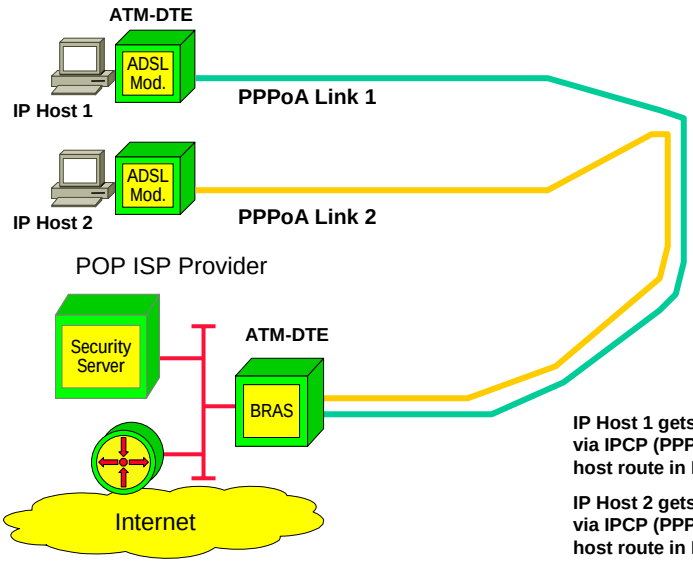
ADSL: PPP over ATM (PPPoA)



© 2007, D.I. Manfred Lindner

L31 - IP Technology Details

ADSL: PPP over ATM (PPPoA), IPCP

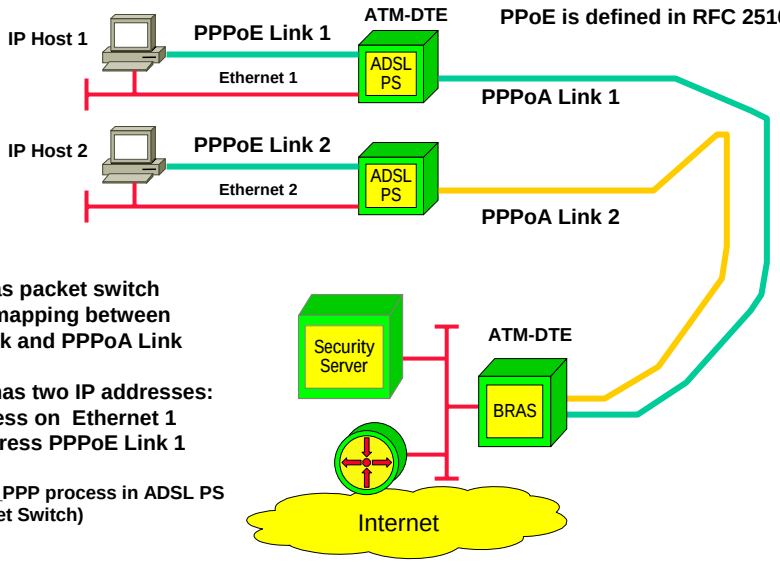


IP Host 1 gets global IP address via IPCP (PPP-NCP), appears as host route in BRAS

IP Host 2 gets global IP address via IPCP (PPP-NCP), appears as host route in BRAS

© 2007, D.I. Manfred Lindner IP Details, v4.7 129

ADSL: PPP over Ethernet (PPPoE)



ADSL PS as packet switch performs mapping between PPPoE Link and PPPoA Link

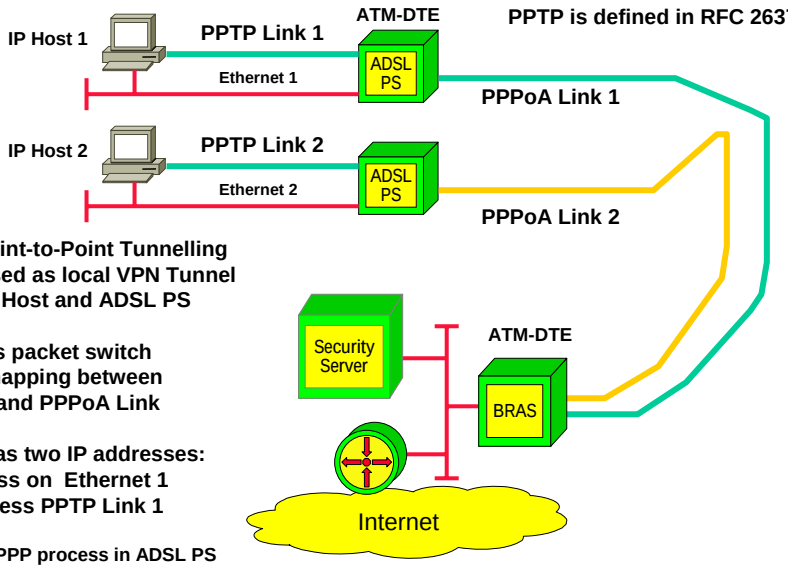
IP Host 1 has two IP addresses:
local address on Ethernet 1
global address PPPoE Link 1

note: Relay_PPP process in ADSL PS (PS ... Packet Switch)

© 2007, D.I. Manfred Lindner IP Details, v4.7 130

© 2007, D.I. Manfred Lindner

ADSL: PPTP over Ethernet (Microsoft VPN)



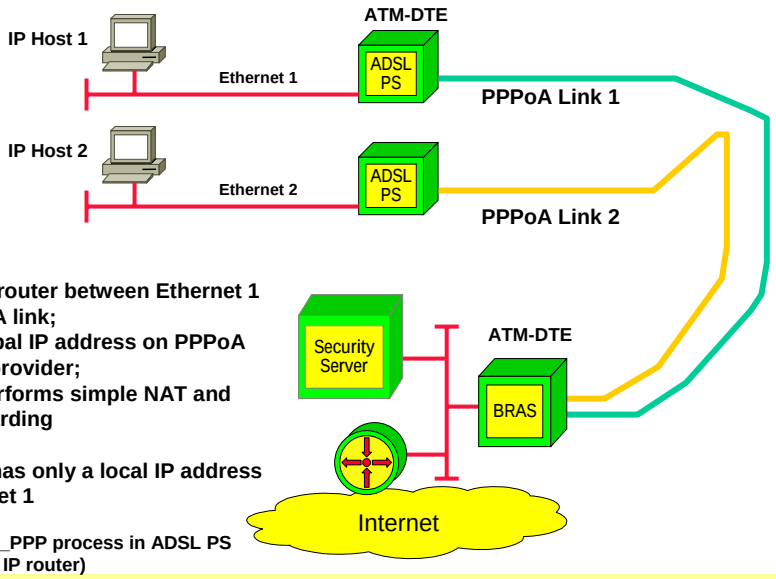
PPTP ... Point-to-Point Tunnelling Protocol used as local VPN Tunnel between IP Host and ADSL PS

ADSL PS as packet switch performs mapping between PPTP Link and PPPoA Link

IP Host 1 has two IP addresses:
local address on Ethernet 1
global address PPTP Link 1

note: Relay_PPP process in ADSL PS

ADSL: Routed PPPoA



ADSL PS :
acts as IP router between Ethernet 1 and PPPoA link;
gets a global IP address on PPPoA link from provider;
usually performs simple NAT and DNS forwarding

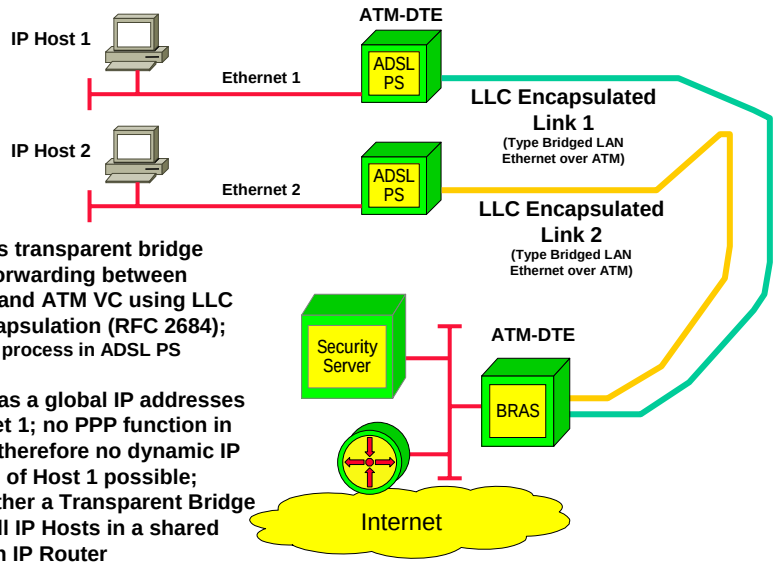
IP Host 1 has only a local IP address on Ethernet 1

note: Dialup_PPP process in ADSL PS (PS is a real IP router)

© 2007, D.I. Manfred Lindner

L31 - IP Technology Details

ADSL: Ethernet Approach



ADSL PS as transparent bridge performs forwarding between Ethernet 1 and ATM VC using LLC bridge encapsulation (RFC 2684);
note: Bridge process in ADSL PS

IP Host 1 has a global IP addresses on Ethernet 1; no PPP function in BRAS and therefore no dynamic IP addressing of Host 1 possible;
BRAS is either a Transparent Bridge grouping all IP Hosts in a shared media or an IP Router

© 2007, D.I. Manfred Lindner IP Details, v4.7 133

© 2007, D.I. Manfred Lindner