IP Technology

Introduction, IP Protocol Details
IP Addressing and IP Forwarding
ARP, ICMP, PPP, HSRP, VRRP

Agenda

- **<u>Introduction</u>**
- **IP**
  - IP Protocol
  - Addressing
- **IP Forwarding**
  - Principles
  - ARP
  - ICMP
  - PPP
- **First Hop Redundancy**
  - Proxy ARP, IDRP, HSRP, VRRP

Datenkommunikation 384.081 - SS 2008

L08 - IP Technology

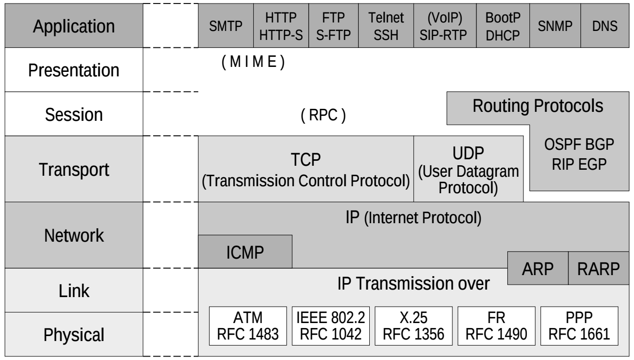
IP Technology

- **packet switching technology**
  - packet switch is called router or gateway (IETF terminology)
  - end system is called IP host
  - structured layer 3 address (IP address)
- **datagram service**
  - connectionless
    - datagrams are sent without establishing a connection in advance
  - best effort delivery
    - datagrams may be discarded due to transmission errors or network congestion

TCP Technology

- **shared responsibility between network and end systems**
  - routers responsible for delivering datagrams to remote networks based on structured IP address
  - IP hosts responsible for end-to-end control
- **end to end control**
  - is implemented in upper layers of IP hosts
  - TCP (Transmission Control Protocol)
    - connection oriented
    - sequencing, windowing
    - error recovery by retransmission
    - flow control

TCP/IP Protocol Suite



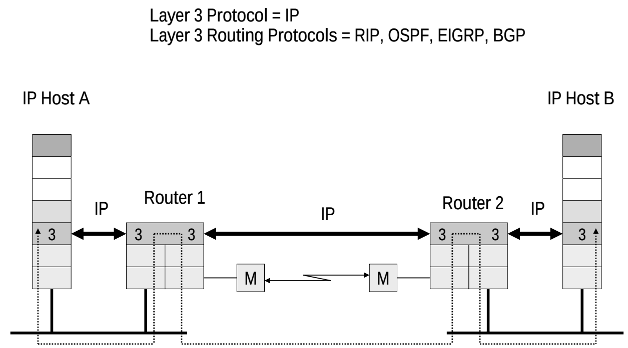
TCP/IP Story of Success

- **IP over everything**
  - technology independent
  - internetwork is built by layering a unique IP protocol on top of various network technologies
    - overlay technique
  - it is easy to adopt new network technologies
    - define how to transfer IP datagrams and how to use the possible switching capability of the new network
- **end-to-end principle**
  - avoids sophisticated tasks to be performed by network infrastructure (routers)
  - TCP takes care of reliability

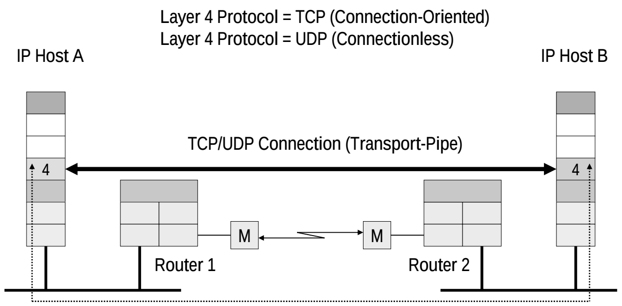
TCP/IP Story of Success

- **TCP**
  - tolerant and adaptive to network operational conditions
    - robust against network failures
    - adapts to varying network delays
    - adapts to varying network load
- **right functionality partition between**
  - IP
    - knows nothing about end systems applications
    - makes best effort to route packets through the network
  - and TCP
    - takes care of end-to-end issues
    - end users know nothing about network internals
- **WWW**
  - invented 1991, world take first notice in 1993

IP and OSI Network Layer 3



TCP/UDP and OSI Transport Layer 4



Key Players of Internet Technology

- **IAB (Internet Architecture Board)**
  - responsible for technical directions, coordination and standardization of the TCP/IP technology
  - the "Board" is highest authority and controls IETF, IRTF
- **IETF (Internet Engineering Task Force)**
  - provides solutions and extensions for TCP/IP
    - working groups organized in areas
    - area manager and IETF chairman form the IESG (Internet Engineering Steering Group)
- **IRTF (Internet Research Task Force)**
  - coordinates and prioritize research
    - research groups controlled by the IRSG (Internet Research Steering Group)

Datenkommunikation 384.081 - SS 2008

L08 - IP Technology

Internet in Europe

- **RIPE NCC (Reséaux IP Européens Network Coordination Center)**
  - Internet Registry
    - assigning IP addresses
    - assigning AS numbers
  - Routing Registry
    - coordinating policies between Internet Service Providers (ISP)
  - how to contact?
    - RIPE NCC
    - Singel 258
    - 1016 AB Amsterdam
    - The Netherlands
    - Phone: +31 20 535 4444 , Fax: +31 20 535 4445
    - E-Mail: <ncc@ripe.net>, WWW: <http://www.ripe.net>

Standardization by RFCs

- **all documentation, standards, proposals for new protocols and enhancements for the Internet**
  - are published as "Requests for Comments" (RFC)
  - RFCs were the initial approach of engineers to discuss questions, suggestions via e-mail to speed up development
    - part of the success story of TCP/IP
  - IETF (Internet Engineering Task Force) decides, which RFCs will be adopted as a standard after rigorous review (e.g. two different implementations have to exist)
  - RFCs are numbered in sequence of publishing
  - adopted enhancements or changes to a protocol will result in a new RFC number

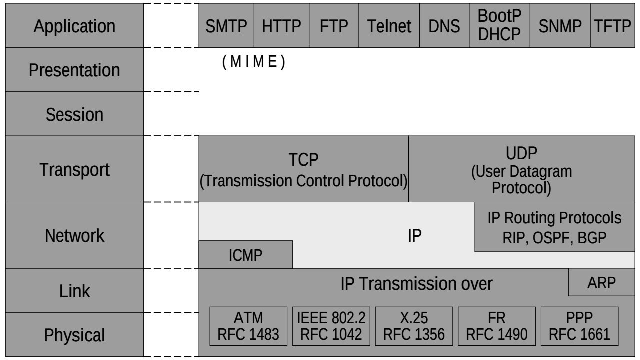
Standardization by RFCs

- **today's standardization process is best described**
  – in RFC-2026
    - The Internet Standards Process Revision3
- **not every RFC is an Internet Standard**
  – categories
    - Informational, Experimental, Historic
    - Proposed Standard
    - Draft Standard
    - Standard
- **IAB (Internet Architecture Board) publishes periodically a status list of all protocols:**
  – Official Protocol Standard RFC (currently RFC 3300).

Agenda

- **Introduction**
- **IP**
  – IP Protocol
  – Addressing
- **IP Forwarding**
  – Principles
  – ARP
  – ICMP
  – PPP
- **First Hop Redundancy**
  – Proxy ARP, IDRP, HSRP, VRRP

IP Related Protocols



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**

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

IP Header

0	4	8	16	31
Version	HLEN	ToS	Total Length	
Fragment Identifier			Flags	Fragment Offset
TTL		Protocol	Header Checksum	
Source Address				
Destination Address				
IP Options				Pad
PAYLOAD				
....				

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

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

© 2008, D.I. Manfred Lindner

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

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.

IP Header Entries 5

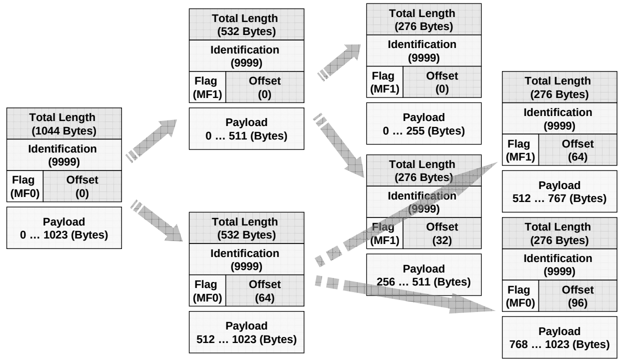
- **Identification (for fragmentation)**
  - Unique identification of a datagram, used for fragmentation and reassembly
  - In praxis a hidden sequence number although not used because of connectionless behavior of IP
- **Flags (for fragmentation).**
  - DF (don't fragment)
    - If set: fragmentation 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

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

IP Fragmentation



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.

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

TOS Field Old Meaning (RFC 1349)

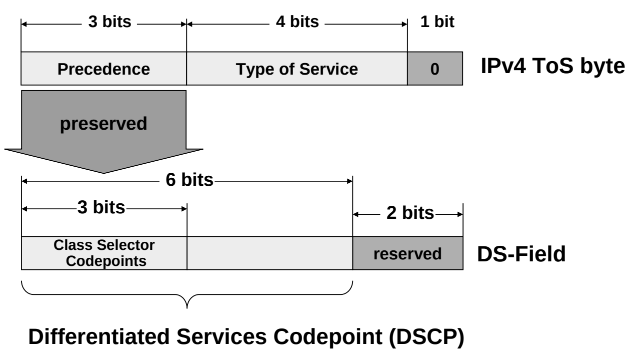
Precedence	D	T	R	C	"0"

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
000 Routine	accepted (ignored) by a router or host.

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

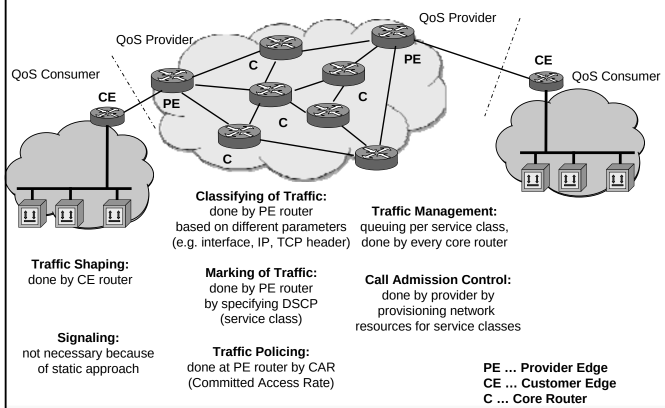
IPv4 TOS Recycling



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

IP QoS Scenario: Differentiated Services



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

Agenda

- **Introduction**
- **IP**
  - IP Protocol
  - Addressing
- **IP Forwarding**
  - Principles
  - ARP
  - ICMP
  - PPP
- **First Hop Redundancy**
  - Proxy ARP, IDRP, HSRP, VRRP

IP Address

- **IP address**
  - 32 bit , dotted decimal notation
  - identifies access to a network (network interface)
  - basic structure
    - network number (net-id)
    - host number (host-id)
  - two level hierarchy
  - net-id must be worldwide unique when a physical network with IP hosts is connected to the Internet
    - assignment controlled by Internet Registry
  - host-id is assigned by each local network manager

Address notation

IP address (example):

1 0 0 0 0 0 0 0	1 1 1 1 0 0 0 0	0 0 0 0 0 0 0 1	0 1 1 0 1 1 0 1
0x80	0xF0	0x01	0x6D

each octet of an IP address is written as the decimal equivalent:

128	240	1	109

The resulting four numbers are delimited with dots (dotted decimal notation):

128.240.1.109

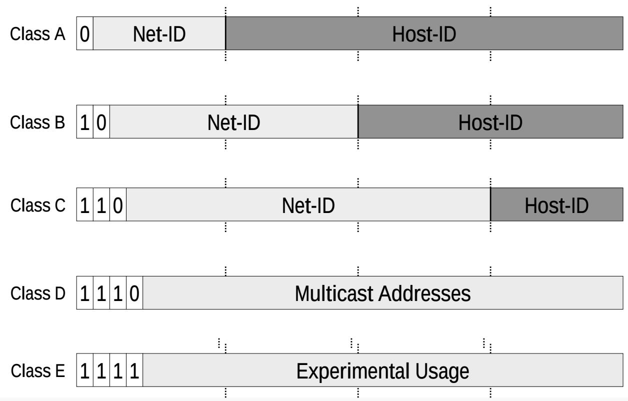
Binary vs Decimal Notation

2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0	
1	0	0	0	0	0	0	0	128
0	1	0	0	0	0	0	0	64
0	0	1	0	0	0	0	0	32
0	0	0	1	0	0	0	0	16
0	0	0	0	1	0	0	0	8
0	0	0	0	0	1	0	0	4
0	0	0	0	0	0	1	0	2
0	0	0	0	0	0	0	1	1
1	1	1	1	1	1	1	1	255

Classes

- **several classes of IP addresses**
  - A, B, C (unicast), D (multicast), E (experimental)
  - class defines numbers of address-bits to be used for net-id
    - class A: 7 bits of net-id, 24 bits of host-id; 126 nets / 16.777.214 hosts
    - class B: 14 bits of net-id, 16 bits of host-id; 16.384 nets / 65.534 hosts
    - class C: 21 bits of net-id, 8 bits of host-id; 2.097.512 nets / 254 hosts
    - class D: 28 bits multicast group number
  - first octet rule
    - class A range: 1 - 126
    - class B range: 128 - 191
    - class C range: 192 - 223
    - class D range: 224 - 239

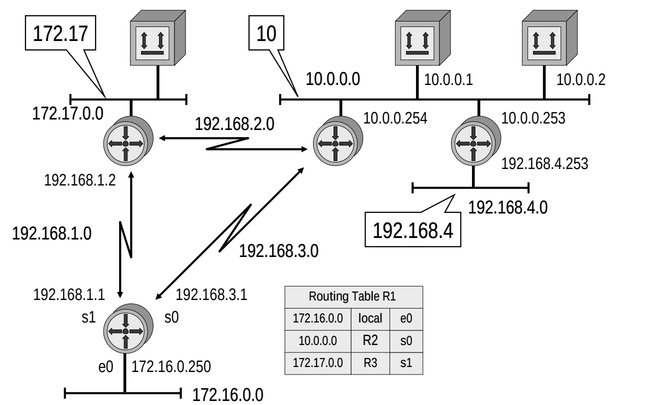
IP Address Classes



IP Address Classes First Octet Rule

Class A	0	1-127
Class B	1 0	128-191
Class C	1 1 0	192-223
Class D	1 1 1 0	224-239
Class E	1 1 1 1	240-255

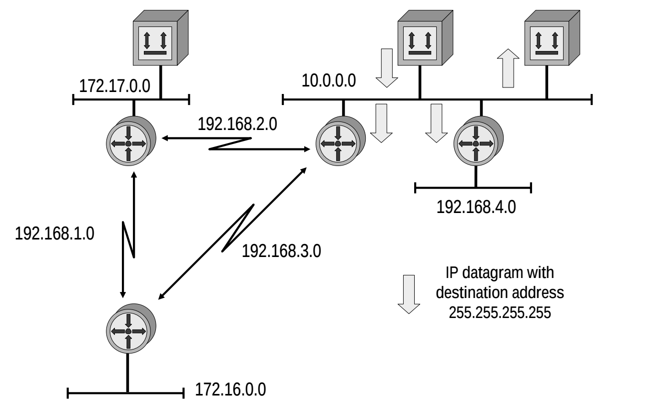
IP Address (Net-ID) Example



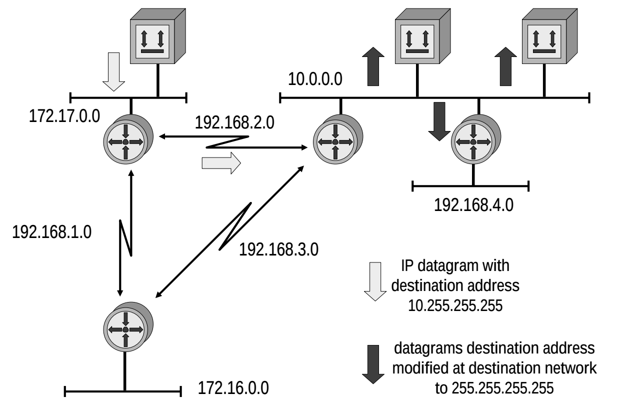
Special Addresses

- **basic IP address format**
  - { net-id, host-id }
- **special purpose addresses and rules**
  - { 0, 0 } this host on this network (0.0.0.0)
  - { 0, <host-id> } specified host on this network
  - { <net-id>, -1 } directed broadcast to specified network
  - { -1, -1 } limited broadcast on this network (255.255.255.255)
  - { 127, <any> } loopback address
  - { <net-id>, 0 } never used for a host number, identifies network itself
  - note:
    - 0 … means all corresponding bits = 0
    - -1 … means all corresponding bits = 1

IP Limited Broadcast



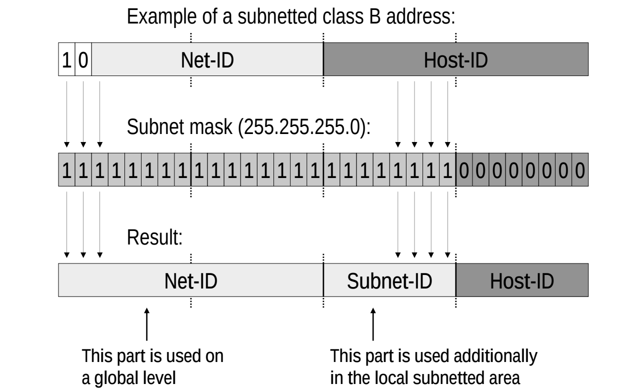
IP Directed Broadcast



Subnetting

- **two level hierarchy was sufficient in the early days of the Internet**
- **with local area networks a third hierarchical level was introduced by subnetting**
- **subnetting**
  - some bits of the host-id can be used as subnet-id
  - subnet-id extends classful net-id meaning
    - subnet-id bits are only locally interpreted inside subnetted area
    - net-id bits are still globally seen outside the subnetted area
  - number of bits to be used for network identification are specified by subnet mask (written in dotted decimal notation)
    - ones portion represents network part (must be contiguous)
    - zeros portion represent the host part

Subnet addressing



Possible Subnet Mask Values

2^7	2^6	2^5	2^4	2^3	2^2	2^1	2^0	
1	0	0	0	0	0	0	0	128
1	1	0	0	0	0	0	0	192
1	1	1	0	0	0	0	0	224
1	1	1	1	0	0	0	0	240
1	1	1	1	1	0	0	0	248
1	1	1	1	1	1	0	0	252
1	1	1	1	1	1	1	0	254
1	1	1	1	1	1	1	1	255

Subnet Mask

- **natural subnet mask**
  - address classes without subnetting
    - class A … 255.0.0.0
    - class B … 255.255.0.0
    - class C … 255.255.255.0
- **old notation of IP addresses**
  - with subnetmask
    - 10.0.0.0 255.0.0.0 (Class A)
    - 176.16.0.0 255.255.0.0 (Class B)
- **new notation of IP addresses**
  - with prefix/length
    - 10.0.0.0 / 8 (Class A)
    - 176.16.0.0 / 16 (Class B)

Rules with Subnetting

- **IP address format with subnetting**
  - { net-id, subnet-id, host-id }
- **additional special purpose addresses and rules**
  - { <net-id>, <subnet-id>, -1 }
    - directed broadcast to specified subnet
  - { <net-id>, -1, -1 }
    - directed broadcast to all subnets of specified subnetted network
  - { <net-id>, 0, <host-id> }
    - subnet zero never used for a subnet number for classful routing (see RFC 950)
  - { <net-id>, -1, <host-id> }
    - subnet broadcast never used for a subnet number for classful routing (see RFC 950)

Subnet Mask Examples 1

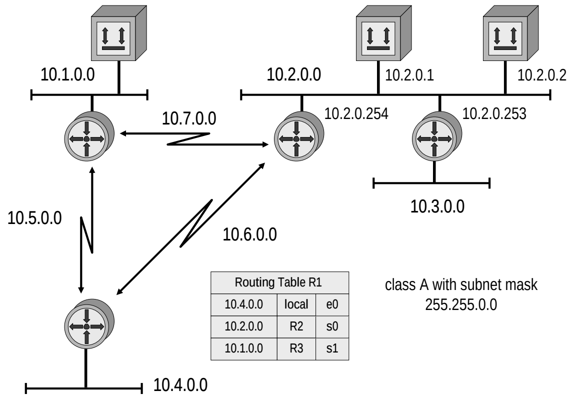
- **class A ⇨ pseudo class B (8 bit subnetting)**
  - 10.0.0.0 with 255.255.0.0 (10.0.0.0 / 16)
  - subnetworks:
    - 10.0.0.0 subnet zero
    - 10.1.0.0
      - 10.1.0.1 first IP host in net 10.1.0.0
      - 10.1.255.254 last IP host in net 10.1.0.0
      - 10.1.255.255 directed broadcast in net 10.1.0.0
    - 10.2.0.0
    - 10.3.0.0

      ……….
    - 10.254.0.0
    - 10.255.0.0 subnet broadcast
  - 254 subnets / 65534 hosts

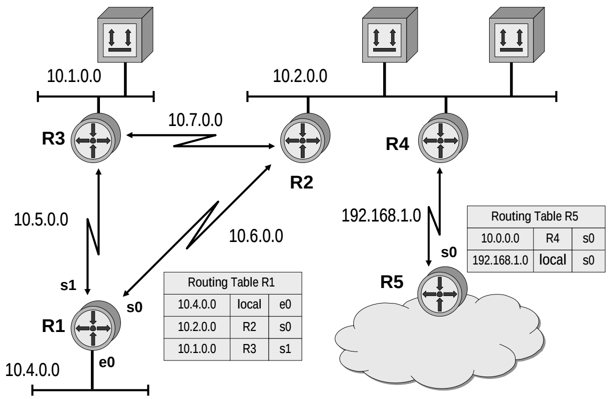
Subnet Zero / Subnet Broadcast

- **What is the problem?**
  - Does 10.0.0.0 mean net-ID
    of net 10
    or
    of subnet 10.0 ?
  - Does 10.255.255.255 mean directed broadcast
    for the whole net 10
    or
    for the subnet 10.255 ?
  - subnet zero and subnet broadcast are ambiguous

IP Address Example with Subnetting



Classful Routing



Datenkommunikation 384.081 - SS 2008

L08 - IP Technology

Subnet Mask Examples 2

- **class A ⇨ pseudo class C (16 bit subnetting)**
  - 10.0.0.0 with 255.255.255.0 (10.0.0.0 / 24)
  - subnetworks:
    - 10.0.0.0 subnet zero
    - 10.0.1.0
    - 10.0.2.0

      ……….
    - 10.0.255.0
    - 10.1.0.0
    - 10.1.2.0

      ………..
    - 10.255.254.0
    - 10.255.255.0 subnet broadcast
  - 65534 subnets / 254 hosts

Subnet Mask Examples 3

- **class B ⇨ pseudo class C (8 bit subnetting)**
  - 172.16.0.0 with 255.255.255.0 (172.16.0.0 / 24)
  - subnetworks:
    - 172.16.0.0 subnet zero
    - 172.16.1.0
    - 172.16.2.0

      ……….

      ………..
    - 172.16.254.0
    - 172.16.255.0 subnet broadcast
  - 254 subnets / 254 hosts

Subnet Mask -> Net-ID, Host-ID

- **class A address**

  subnet mask 255.255.0.0
  IP- Address 10.3.49.45
  ? net-id, ? host-id

  net-id = 10.3.0.0
  host-id = 0.0.49.45

  65534 IP hosts
  range: 10.3.0.1 -> 10.3.255.254
  10.3.0.0 -> network itself
  10.3.255.255 -> directed broadcast for this network

Subnet Mask Examples 4

- **class B address**

  subnet mask 255.255.255.192
  IP- Address 172.16.3.144
  ? net-id, ? host-id

  address binary 1010 1100 . 0001 0000 . 0000 0011 . 1001 0000
  mask (binary) 1111 1111 . 1111 1111 . 1111 1111 . 1100 0000
  ----
  logical AND (bit by bit)
  net-id 1010 1100 . 0001 0000 . 0000 0011 . 1000 0000

  net-id = 172.16.3.128
  host-id = 0.0.0.16

Subnet Mask Examples 5

- **class B ⇨ 10 bit subnetting**
  - 172.16.0.0 with 255.255.255.192 (172.16.0.0 / 26)
  - subnetworks: net-ID host-ID
    - 172.16.0.0 subnet zero 172.16.0. 00 | xx xxxx
    - 172.16.0.64 172.16.0. 01 | xx xxxx
      - 172.16.0.65 first IP host 172.16.0. 01 | 00 0001
      - 172.16.0.66 second IP host 172.16.0. 01 | 00 0010

        ………..
      - 172.16.0.126 last IP host 172.16.0. 01 | 11 1110
      - 172.16.0.127 directed broadcast 172.16.0. 01 | 11 1111
    - 172.16.0.128 172.16.0. 10 | xx xxxx
    - 172.16.0.192 172.16.0. 11 | xx xxxx

Subnet Mask Examples 5

- subnetworks (cont.):
  - 172.16.1.0 172.16.1. 00 | xx xxxx
  - 172.16.1.64 172.16.1. 01 | xx xxxx
  - 172.16.1.128 172.16.1. 10 | xx xxxx
  - 172.16.1.192 172.16.1. 11 | xx xxxx
  - 172.16.2.0 172.16.2. 00 | xx xxxx
  - 172.16.2.64 172.16.2. 01 | xx xxxx

    ………..
  - 172.16.255.0 172.16.255. 00 | xx xxxx
  - 172.16.255.64 172.16.255. 01 | xx xxxx
  - 172.16.255.128 172.16.255. 10 | xx xxxx
  - 172.16.255.192 subnet broadcast 172.16.255. 11 | xx xxxx
- 1022 subnets / 62 hosts

Subnet Mask Examples 6

- **class C ⇨ 2 bit subnetting**
  - 192.168.16.0 with 255.255.255.192 (192.168.16.0 / 26)
  - subnetworks:

	net-ID	host-ID
192.168.16.0 subnet zero	192.168.16. 00	xxxxxx
192.168.16.64	192.168.16. 01	xxxxxx
192.168.16.128	192.168.16. 10	xxxxxx
192.168.16.192 subnet broadcast	192.168.16. 11	xxxxxx

  - 2 subnets / 62 hosts

Subnet Mask Examples 7

- **class C ⇨ 6 bit subnetting**
  - 192.168.16.0 with 255.255.255.252 (192.168.16.0 / 30)
  - subnetworks:

	net-ID	host-ID
192.168.16.0 subnet zero	192.168.16. 000000	xx
192.168.16.4	192.168.16. 000001	xx
– 192.168.16.5 1st IP host	192.168.16. 000001	01
– 192.168.16.6 2nd IP host	192.168.16. 000001	10
– 192.168.16.7 directed broadcast	192.168.16. 000001	11
192.168.16.8	192.168.16. 000010	xx
………..		
192.168.16.248	192.168.16. 111110	xx
192.168.16.252 subnet broadcast	192.168.16. 111111	xx

  - 62 subnets / 2 hosts

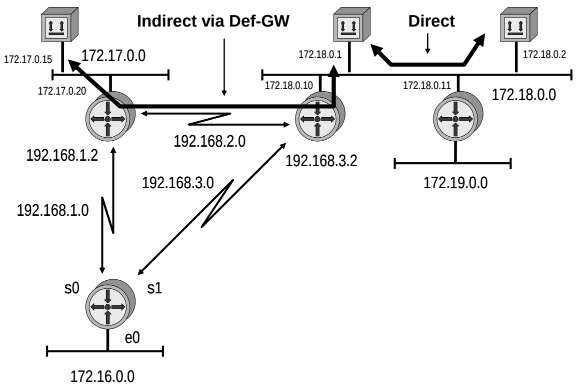
Agenda

- **Introduction**
- **IP**
  - IP Protocol
  - Addressing
- **IP Forwarding**
  - Principles
  - ARP
  - ICMP
  - PPP
- **First Hop Redundancy**
  - Proxy ARP, IDRP, HSRP, VRRP

IP Forwarding Responsibilities

- **IP hosts and IP routers take part in this process**
  - IP hosts responsible for direct delivery of IP datagram's
  - IP routers responsible for selecting the best path in a meshed network in case of indirect delivery of IP datagram's
    - decision based on current state of routing table
- **direct versus indirect delivery**
  - depends on destination net-ID
    - net-ID equal source net-ID -> direct delivery
    - net-ID unequal source net-ID -> indirect delivery
- **IP hosts choose a “default” router aka “Default Gateway”**
  - as next hop in case of indirect delivery of IP datagrams

Direct versus Indirect Delivery



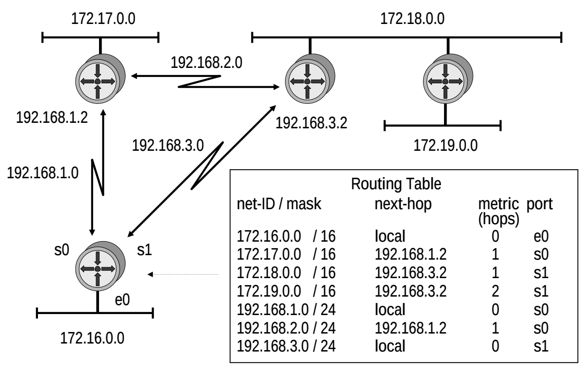
Principle

- **IP Forwarding is done by routers in case of indirect routing**
  - based on the destination address of a given IP datagram
  - following the path to the destination hop by hop
- **routing tables**
  - have information about which next hop router a given destination network can be reached
- **L2 header must be changed hop by hop**
  - if LAN then physical L2 address (MAC addresses) must be adapted for direct communication on LAN
- **mapping between IP and L2 address on LAN**
  - is done by Address Resolution Protocol (ARP)

IP Routing Paradigm

- **Destination Based Routing**
  - source address is not taken into account for the forward decision
- **Hop by Hop Routing**
  - IP datagram's follow the path, which is pointed by the current state of the routing tables
- **Least Cost Routing**
  - normally only the best path is considered for forwarding of IP datagram's
  - alternate paths will not be used in order to reach a given destination

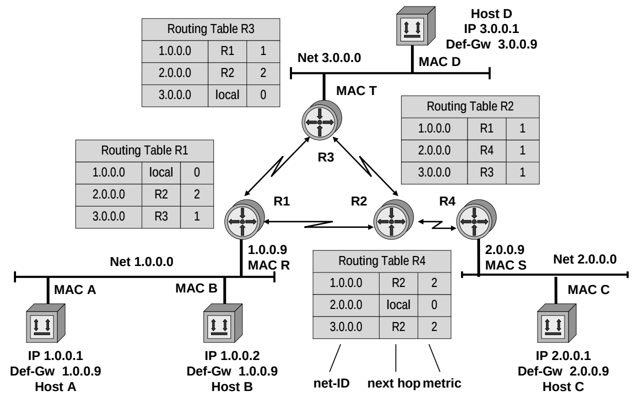
Routing Table Example



Routing Table

net-ID / mask	next-hop	metric (hops)	port
172.16.0.0 / 16	local	0	e0
172.17.0.0 / 16	192.168.1.2	1	s0
172.18.0.0 / 16	192.168.3.2	1	s1
172.19.0.0 / 16	192.168.3.2	2	s1
192.168.1.0 / 24	local	0	s0
192.168.2.0 / 24	192.168.1.2	1	s0
192.168.3.0 / 24	local	0	s1

Example Topology



Routing Table R3		
1.0.0.0	R1	1
2.0.0.0	R2	2
3.0.0.0	local	0

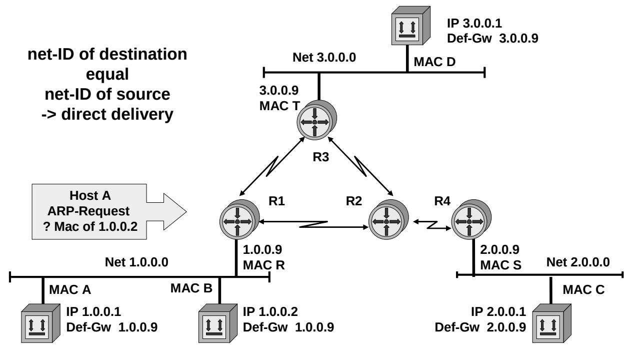
Routing Table R2		
1.0.0.0	R1	1
2.0.0.0	R4	1
3.0.0.0	R3	1

Routing Table R1		
1.0.0.0	local	0
2.0.0.0	R2	2
3.0.0.0	R3	1

Routing Table R4		
1.0.0.0	R2	2
2.0.0.0	local	0
3.0.0.0	R2	2

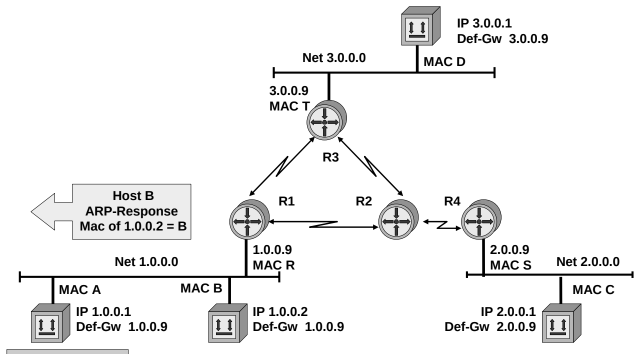
Direct Delivery 1.0.0.1 - > 1.0.0.2

net-ID of destination
equal
net-ID of source
-> direct delivery



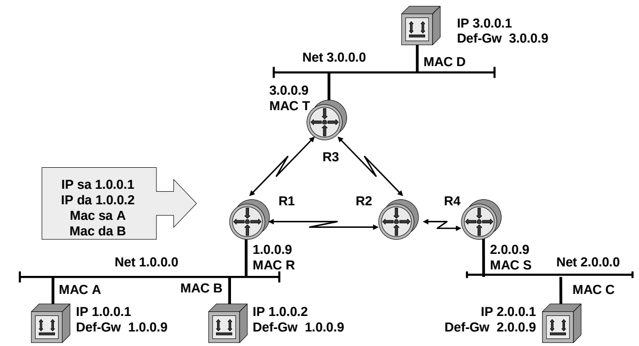
ARP ... Address Resolution Protocol

Direct Delivery 1.0.0.1 - > 1.0.0.2



ARP-Cache Host A	
1.0.0.2	MAC B

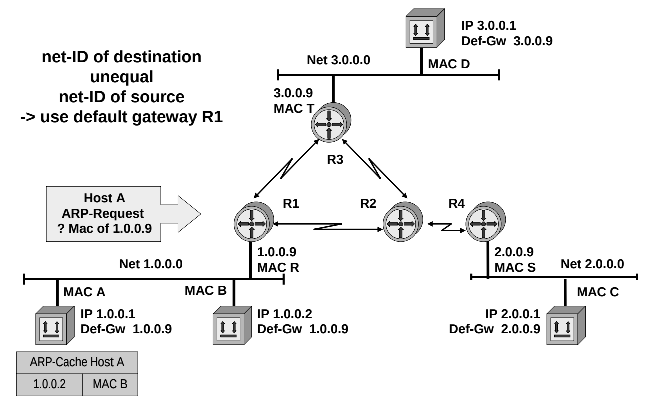
Direct Delivery 1.0.0.1 - > 1.0.0.2



ARP-Cache Host A	
1.0.0.2	MAC B

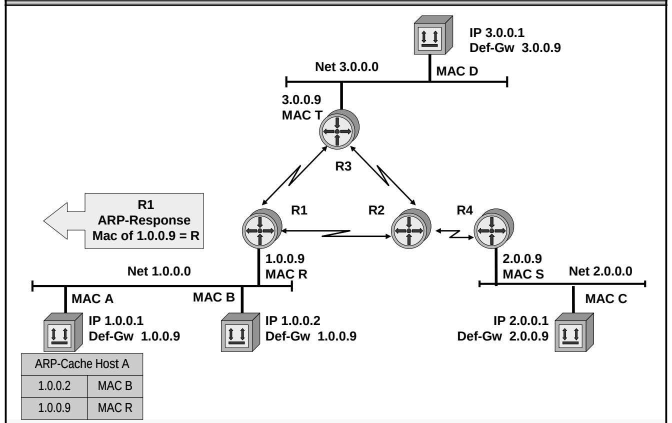
L08 - IP Technology

Indirect Delivery 1.0.0.1 - > 2.0.0.1



 IP Technology, v4.8 73

Indirect Delivery 1.0.0.1 - > 2.0.0.1

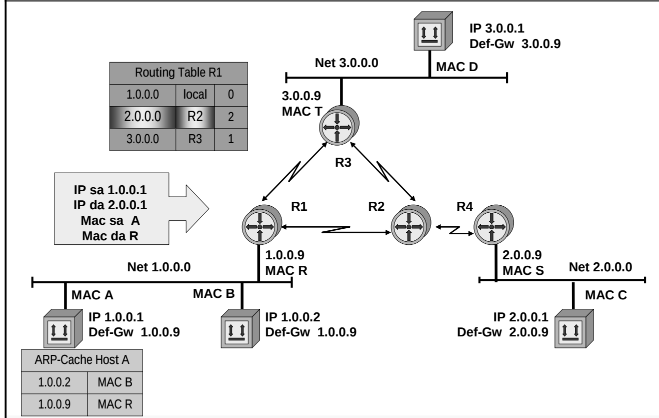


 IP Technology, v4.8 74

Datenkommunikation 384.081 - SS 2008

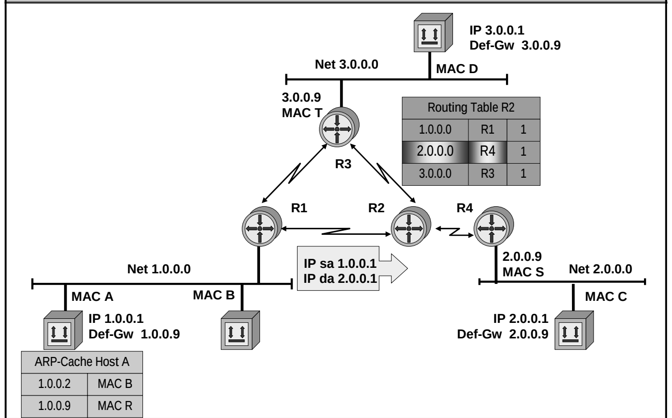
L08 - IP Technology

Indirect Delivery 1.0.0.1 - > 2.0.0.1



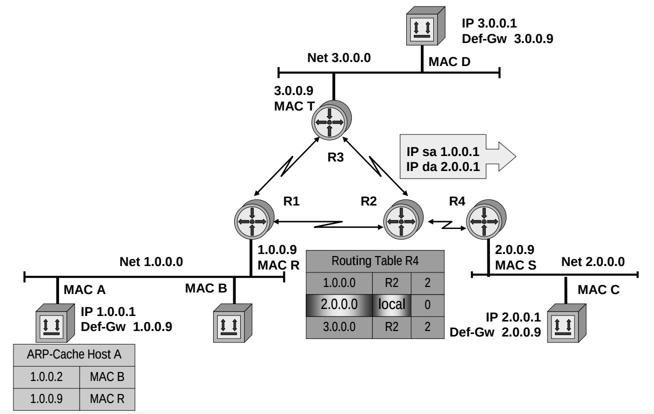
 IP Technology, v4.8 75

Indirect Delivery 1.0.0.1 - > 2.0.0.1



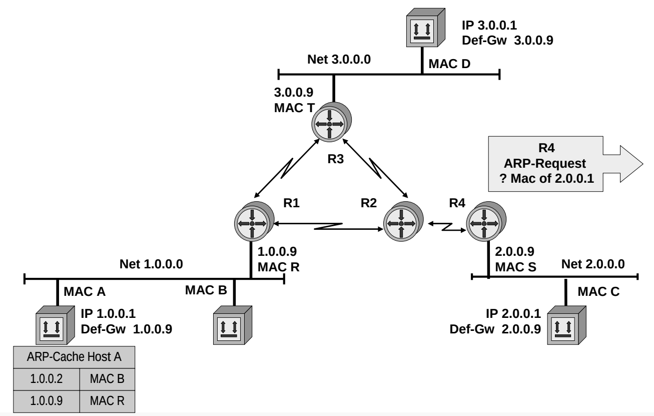
 IP Technology, v4.8 76

Indirect Delivery 1.0.0.1 - > 2.0.0.1



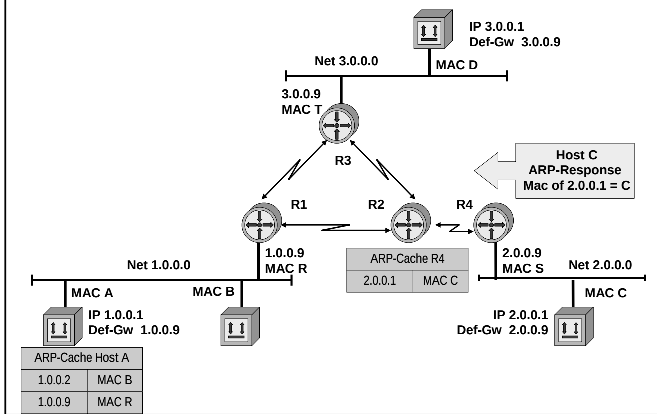
 IP Technology, v4.8 77

Indirect Delivery 1.0.0.1 - > 2.0.0.1



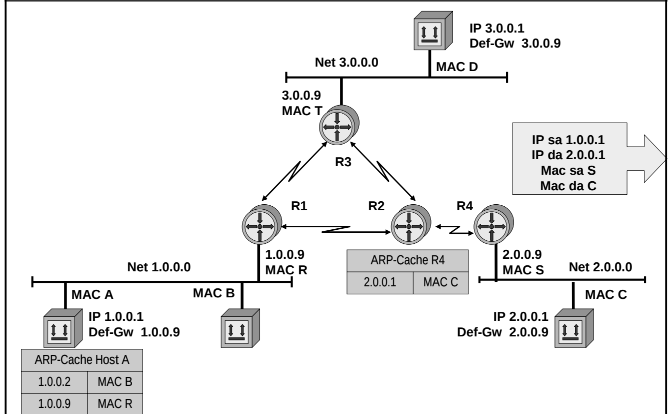
 IP Technology, v4.8 78

Indirect Delivery 1.0.0.1 - > 2.0.0.1



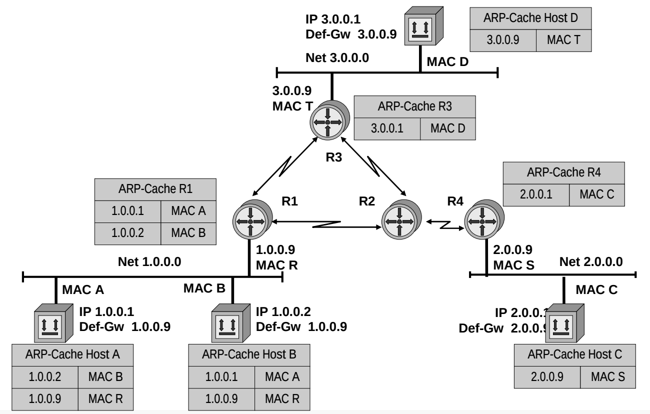
 IP Technology, v4.8 79

Indirect Delivery 1.0.0.1 - > 2.0.0.1



 IP Technology, v4.8 80

ARP Cache - Final Picture



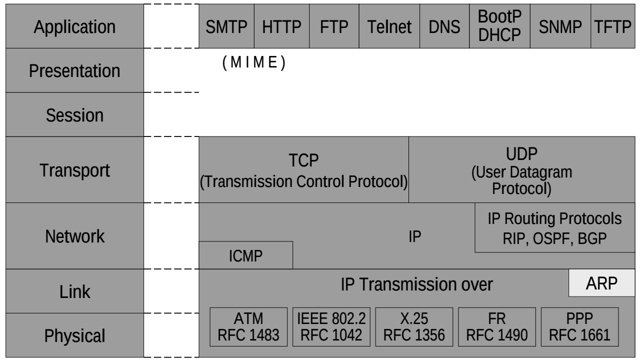
Agenda

- **Introduction**
- **IP**
  - IP Protocol
  - Addressing
- **IP Forwarding**
  - Principles
  - ARP
  - ICMP
  - PPP
- **First Hop Redundancy**
  - Proxy ARP, IDRP, HSRP, VRRP

Datenkommunikation 384.081 - SS 2008

L08 - IP Technology

IP Related Protocols



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 response 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

ARP Operation 2

- The ARP response 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 / response)
- 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 / responses are sent in Ethernet II or SNAP frames (Type field 0x0806)
- ARP has been designed to support different layer 3 protocols

Datenkommunikation 384.081 - SS 2008

L08 - IP Technology

ARP Request/Response Format

<table>
<tr><td colspan="2">Hardware</td><td colspan="2">Protocol (IP = 0x0800)</td></tr>
<tr><td>hln</td><td>pln</td><td colspan="2">Operation</td></tr>
<tr><td colspan="4">Source Hardware Address (byte 0 - 3)</td></tr>
<tr><td colspan="2">Source HW Addr. (byte 4 - 5)</td><td colspan="2">Source IP Addr. (byte 0 - 1)</td></tr>
<tr><td colspan="2">Source IP Addr. (byte 2 - 3)</td><td colspan="2">Dest. HW Addr. (byte 0 - 1)*</td></tr>
<tr><td colspan="4">Destination Hardware Address (byte 2 - 5)*</td></tr>
<tr><td colspan="4">Destination IP Address (byte 0 - 3)</td></tr>
</table>

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

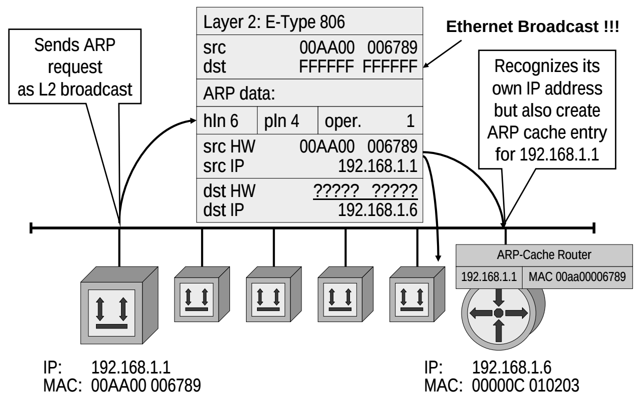
ARP Request/Response Fields

- **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/Response Fields

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

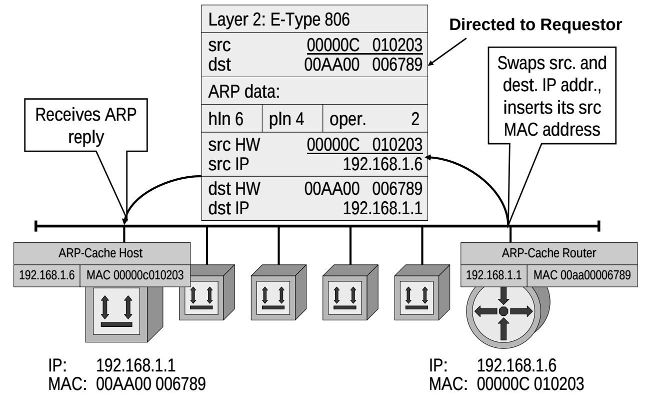
ARP Request



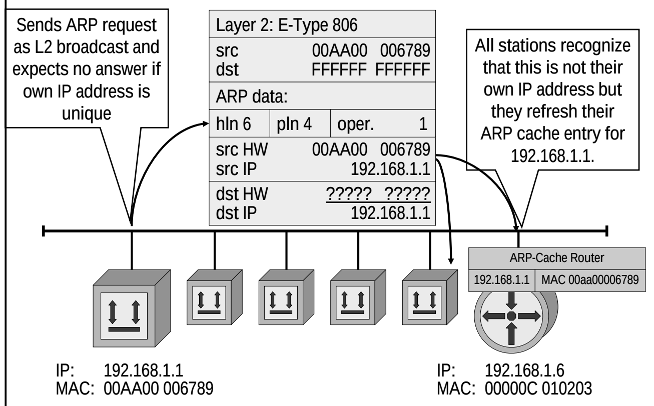
Datenkommunikation 384.081 - SS 2008

L08 - IP Technology

ARP Reply



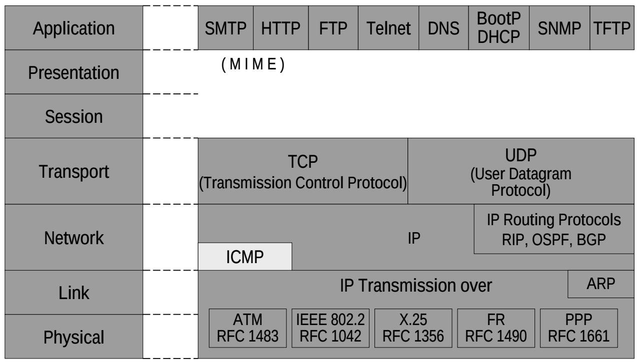
Gratuitous ARP for Duplicate Address Check and ARP Cache Refresh



Agenda

- **Introduction**
- **IP**
  - IP Protocol
  - Addressing
- **<u>IP Forwarding</u>**
  - Principles
  - ARP
  - <u>ICMP</u>
  - PPP
- **First Hop Redundancy**
  - Proxy ARP, IDRP, HSRP, VRRP

IP Related Protocols



Datenkommunikation 384.081 - SS 2008

L08 - IP Technology

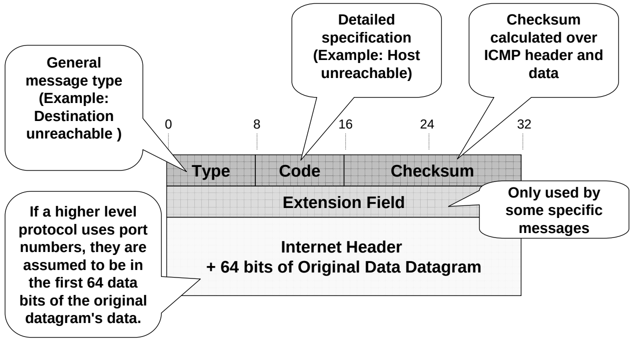
ICMP (RFC 792)

- **datagram service of IP**
  - best effort -> IP datagram's can be lost
- **ICMP (<u>I</u>nternet <u>C</u>ontrol <u>M</u>essage <u>P</u>rotocol)**
  - generates error messages to enhance the reliability and to provide information about errors and packet loss in the network
  - allows to request information for debugging and diagnosis
- **principle of ICMP operation**
  - IP station (router or destination), which detects any transmission problems, generates an ICMP message
  - ICMP message is addressed to the originating station (sender of the original IP packet)

ICMP

- **ICMP messages are sent as IP packets**
  - protocol field = 1, ICMP header and code in the IP data area
- **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)
- **Analysis of ICMP messages**
  - through network management systems or statistic programs can give valuable hints for network administrators

ICMP Message Format



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

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

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

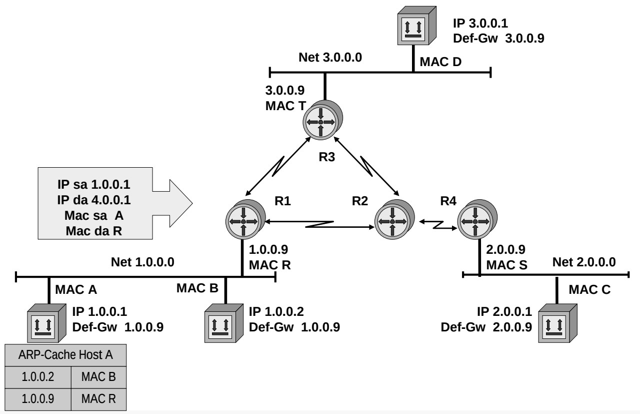
Code Field for Type 3 (destination unreachable)

See RFC1122 (Host Requirements) page 38:

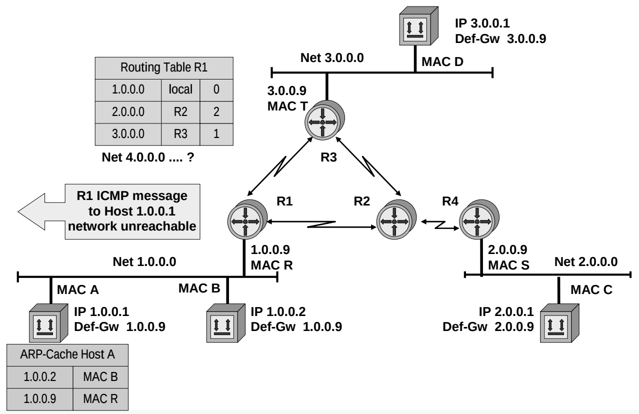
The following additional codes are hereby defined:

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

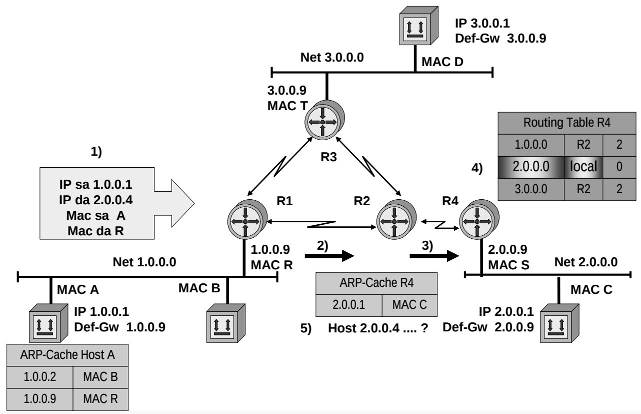
Delivery 1.0.0.1 - > 4.0.0.1



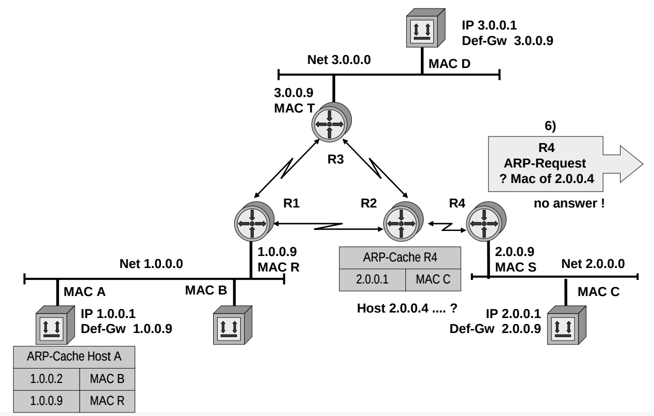
ICMP network unreachable



Delivery 1.0.0.1 - > 2.0.0.4

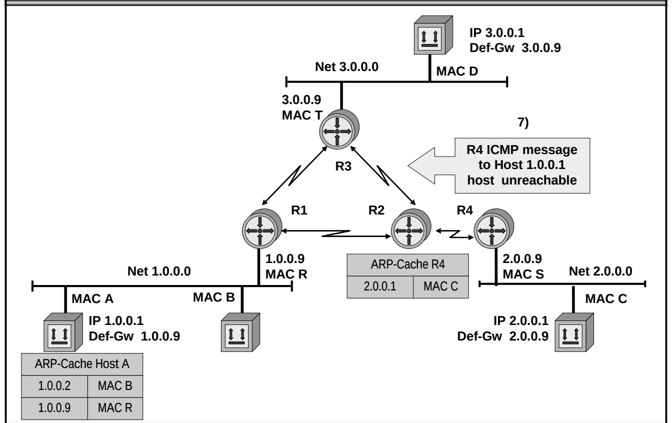


Delivery 1.0.0.1 - > 2.0.0.4



 IP Technology, v4.8 105

ICMP host unreachable

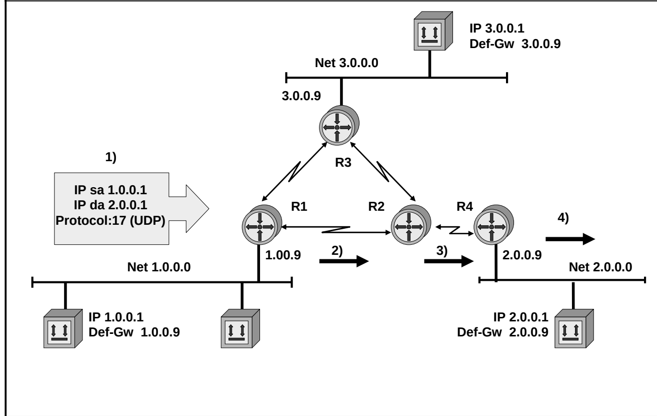


 IP Technology, v4.8 106

Datenkommunikation 384.081 - SS 2008

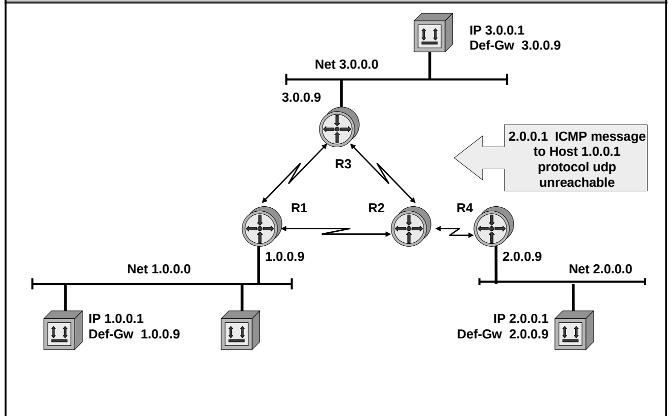
L08 - IP Technology

Delivery 1.0.0.1 - > 2.0.0.1 (protocol udp)



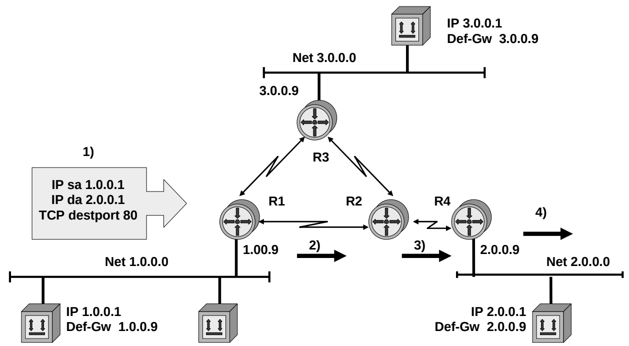
 IP Technology, v4.8 107

ICMP protocol unreachable

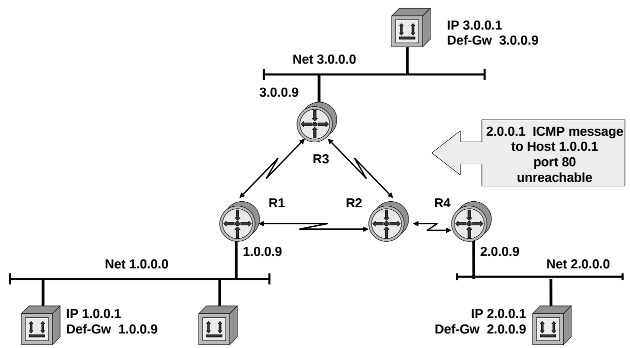


 IP Technology, v4.8 108

Delivery 1.0.0.1 - > 2.0.0.1 (http_server_proc)



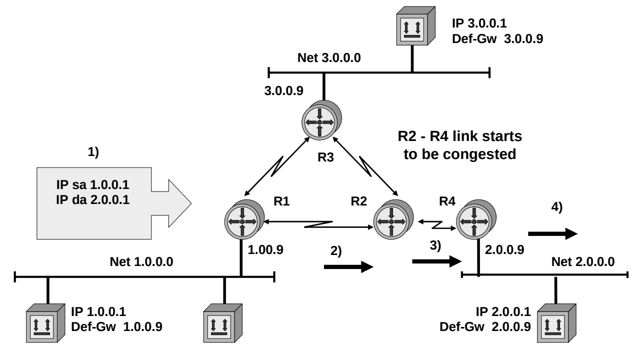
ICMP port unreachable (no http_server_proc)



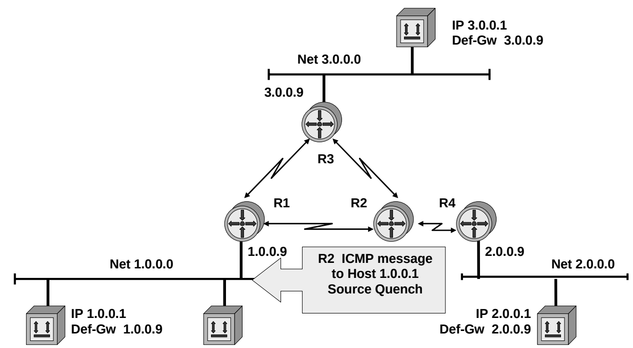
Datenkommunikation 384.081 - SS 2008

L08 - IP Technology

R2 -> R4 Link Congested



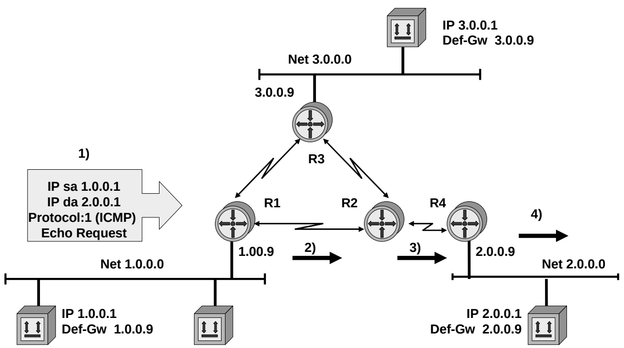
ICMP Source Quench (Flow Control STOP?)



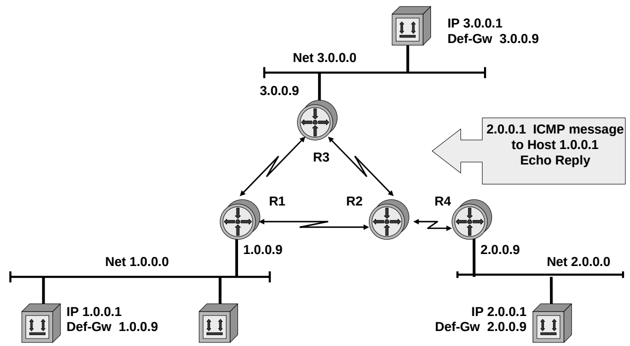
PING - Packet Internet Groper

- **Checks the reachability of an IP station.**
- **Measures time (round-trip-delay).**
- **Example:**
  - ping 132.105.56.3 (with IP address)
  - ping www.proin.via.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

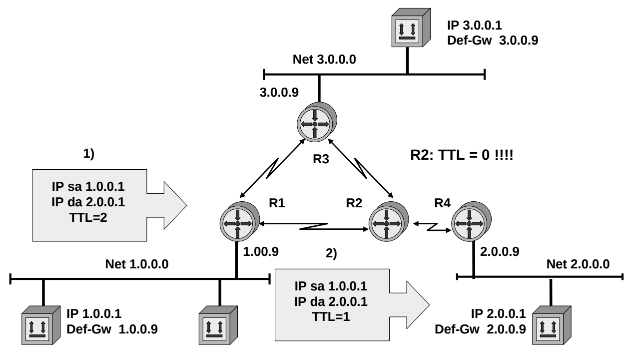
Ping 1.0.0.1 - > 2.0.0.1



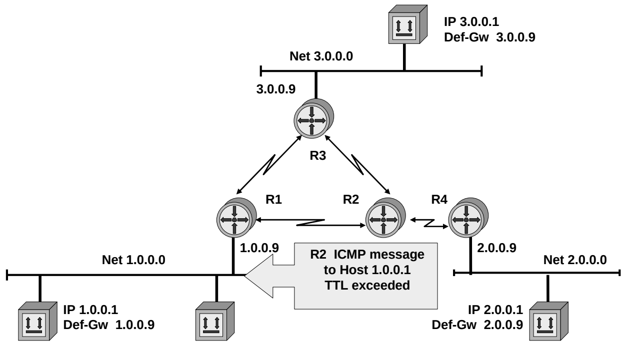
Ping Echo 2.0.0.1 - > 1.0.0.1



Delivery 1.0.0.1 - > 2.0.0.1 (TTL=2)



ICMP TTL exceeded



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**

Traceroute - Sample Output

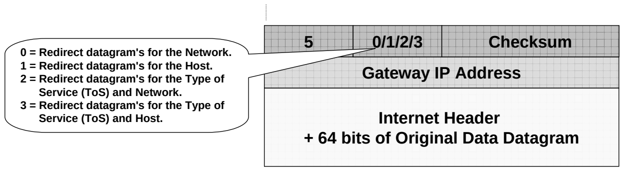
```
tracert 140.252.13.65

1 ny-providerx-int-99 (140.252.13.35)   20ms    10ms   10ms
2 www.example.com  (140.252.13.65)   *         120ms  120ms
```

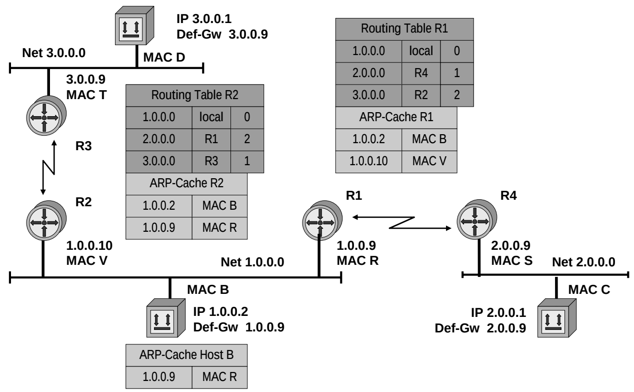
3 Packets are sent for each TTL value.
Output of "*", if no answer arrives within 5 seconds.

Code Field for Type 5 (Redirect)

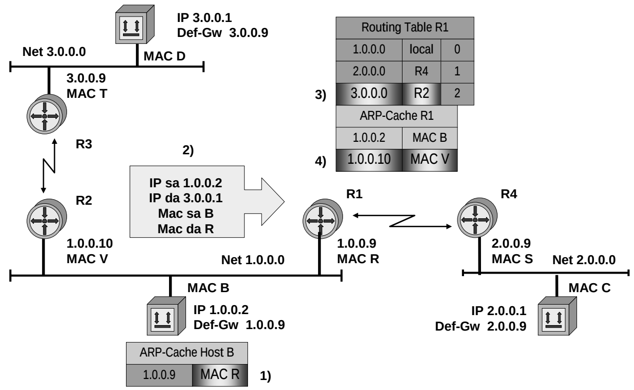
- **If a router knows of a better (faster, shorter) path to a target then it will notify the sender through ICMP redirect**
  - In any case the router will still forward the packets on the inefficient path
  - Datagram's will be sent twice through a LAN, if the sender ignores the redirect message



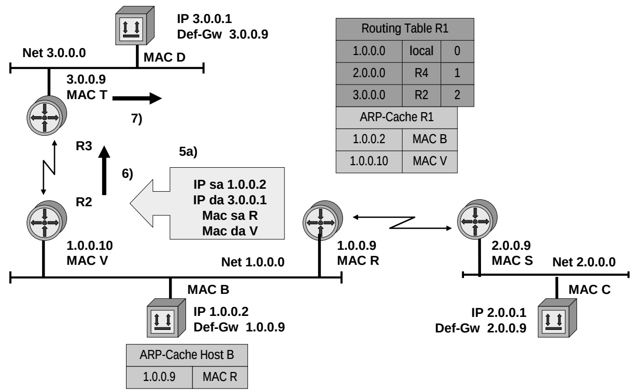
Delivery 1.0.0.2 -> 3.0.0.1



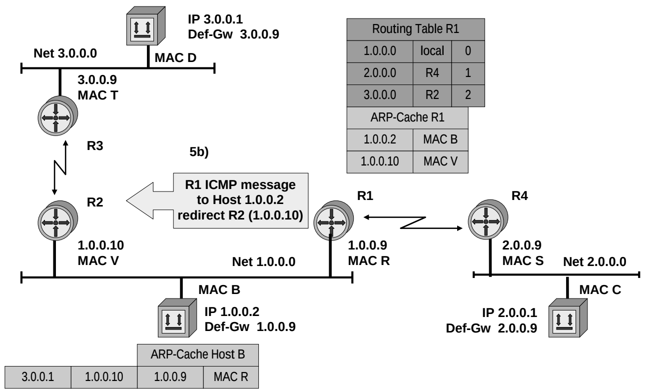
Delivery 1.0.0.2 -> 3.0.0.1



Delivery 1.0.0.2 -> 3.0.0.1



ICMP redirect

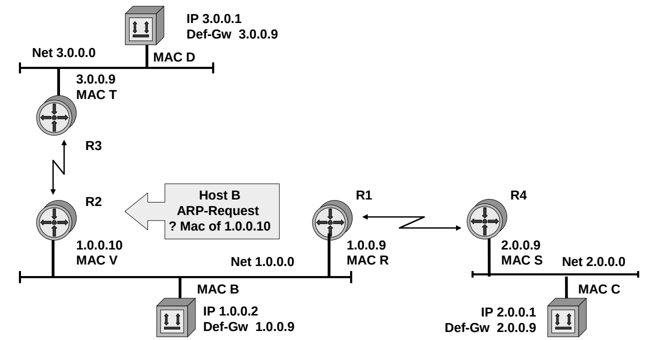


Routing Table R1		
1.0.0.0	local	0
2.0.0.0	R4	1
3.0.0.0	R2	2

ARP-Cache R1	
1.0.0.2	MAC B
1.0.0.10	MAC V

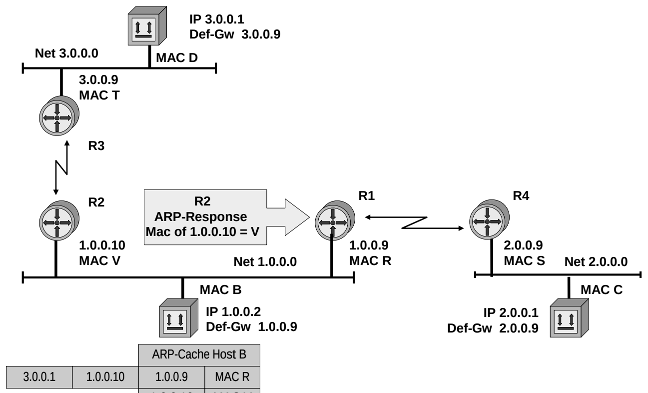
		ARP-Cache Host B	
3.0.0.1	1.0.0.10	1.0.0.9	MAC R

Delivery 1.0.0.2 -> 3.0.0.1



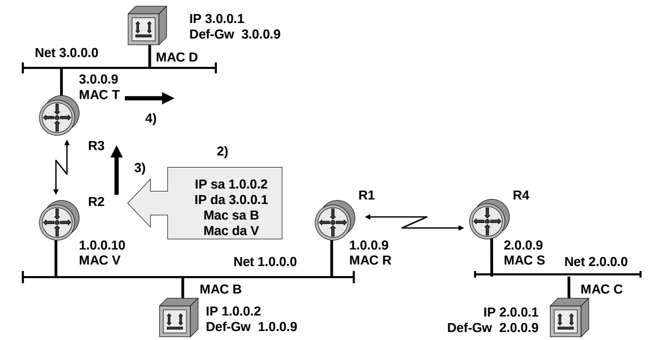
		ARP-Cache Host B	
3.0.0.1	1.0.0.10	1.0.0.9	MAC R

Delivery 1.0.0.2 -> 3.0.0.1



		ARP-Cache Host B	
3.0.0.1	1.0.0.10	1.0.0.9	MAC R
		1.0.0.10	MAC V

Next Packet 1.0.0.2 -> 3.0.0.1



		ARP-Cache Host B	
3.0.0.1	1.0.0.10	1.0.0.9	MAC R
		1.0.0.10	MAC V

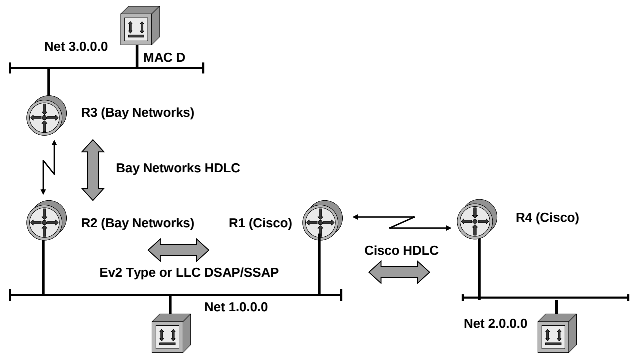
Agenda

- **Introduction**
- **IP**
  - IP Protocol
  - Addressing
- **IP Forwarding**
  - Principles
  - ARP
  - ICMP
  - PPP
- **First Hop Redundancy**
  - Proxy ARP, IDRP, HSRP, VRRP

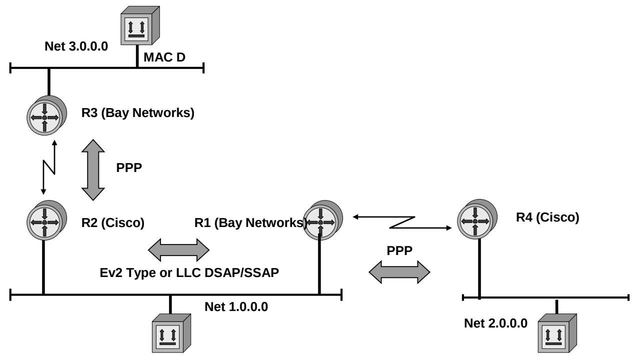
Reasons for Point-to-Point Protocol (PPP)

- **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

Interoperability without PPP



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), IETF-RFC

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
- **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

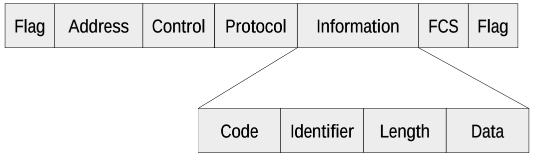
PPP Frame Format

Flag	Address	Control	Protocol	Information	FCS	Flag

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 Protocol

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)

Types of LCP Packets

- **There are three classes of LCP packets:**
  - <u>class 1</u>: Link Configuration packets used to establish and configure a PPP 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)
  - <u>class 2</u>: Link Termination packets used to terminate a link
    - Terminate-Request (code 5) and Terminate-Ack (code 6)
  - <u>class 3</u>: Link Maintenance packets used to manage and debug a PPP 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)

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**
  - <u>phase 1</u>: link establishment and configuration negotiation
    - done by LCP (note: deals only with link operations, does not negotiate the implementation of network layer protocols)
  - <u>phase 2</u>: optional procedures that were agreed during negotiation of phase 1 (e.g. CHAP authentication or compression)
  - <u>phase 3</u>: network layer protocol configuration negotiation done by corresponding NCP´s
    - e.g. IPCP, IPXCP, …
  - <u>phase 4</u>: link termination

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)

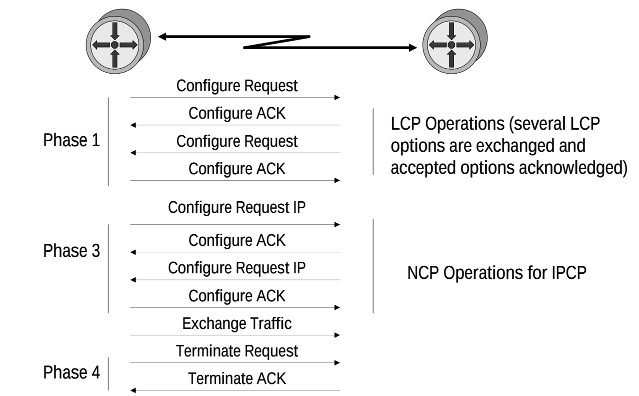
PPP Phases

- **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

PPP Phases

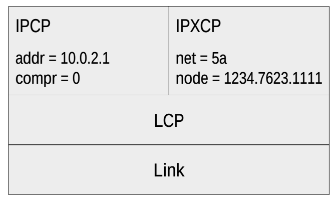
- **task of phase 2**
  - providing of optional facilities
    - authentication, compression initialization, multilink, etc.
- **task of phase 3**
  - network layer protocol configuration negotiation
    - after link establishment, stations negotiate/configure the protocols that will be used at the network 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

PPP Link Operation Example



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



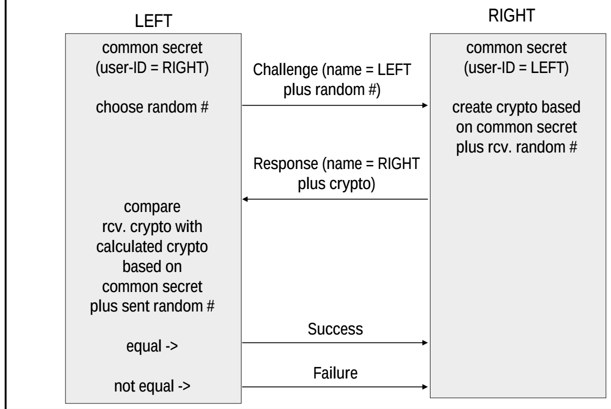
CHAP Authentication RFC 1994

- **Challenge Authentication Protocol**
- **follows establishment of LCP**
- **identifies user**
- **three way handshake**
- **one way authentication only**
  - station which starts the three way handshake proofs authentication of 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

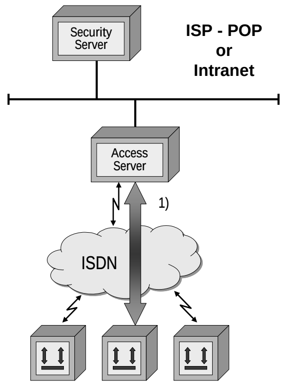
CHAP Authentication Procedure



PPP as Dial-In Technology

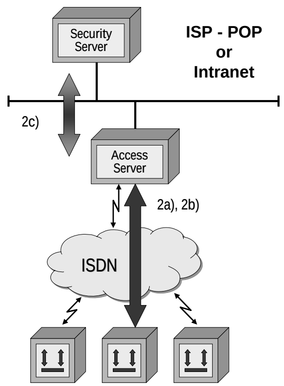
- **Dial-In:**
  - Into a corporate network (Intranet) of a company
    - Here the term RAS (remote access server) is commonly used to describe the point for accessing the dial-in service
  - Into the Internet by having an dial-in account with an Internet Service Provider (ISP)
    - Here the term POP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

RAS Operation 1



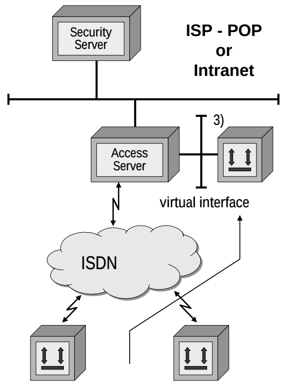
- **remote PC places ISDN call to access server, ISDN link is established (1)**

RAS Operation 2



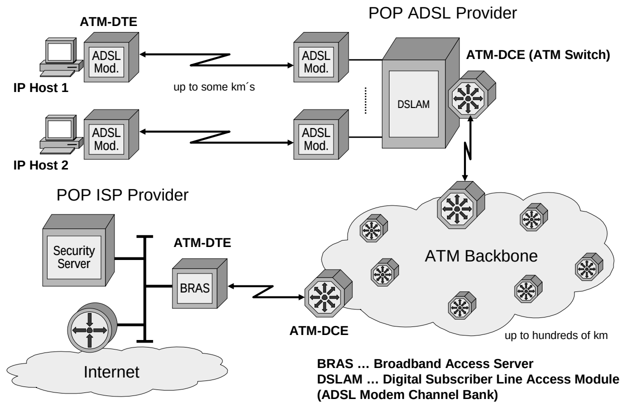
- **PPP link (multiprotocol over serial line) is established**
  - LCP Link Control Protocol (2a)
    - establishes PPP link plus negotiates parameters like authentication CHAP
  - authentication
    - CHAP Challenge Authentication Protocol to transport passwords (2b)
    - verification maybe done by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

RAS Operation 3

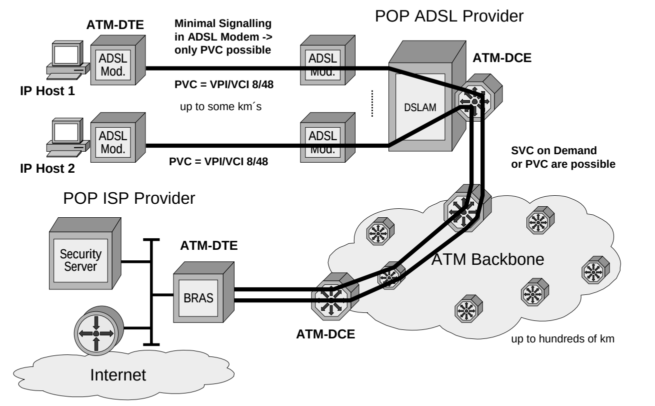


- **PPP NCP (Network Control Protocol) IPCP**
  - assigns IP address, Def. GW, DNS to remote PC
- **remote PC appears as**
  - device reachable via virtual interface (3), IP host Route
- **optionally**
  - filter could be established on that virtual interface
    - authorization
  - accounting can be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

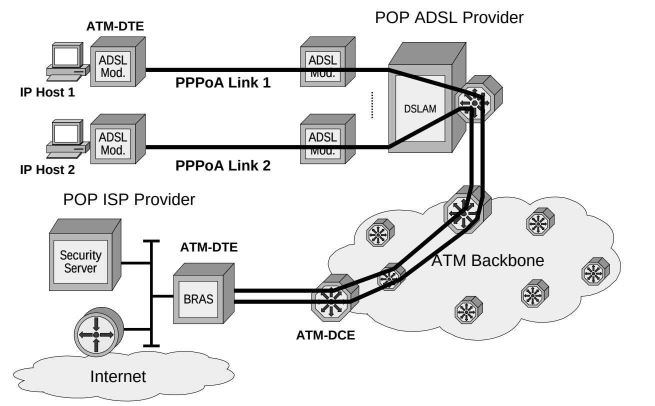
ADSL: Physical Topology



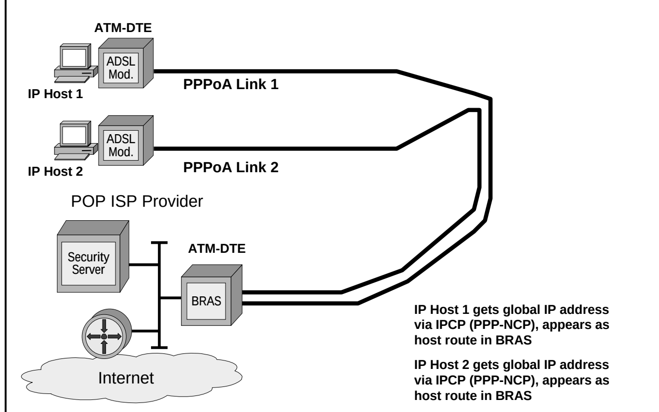
ADSL: ATM Virtual Circuits



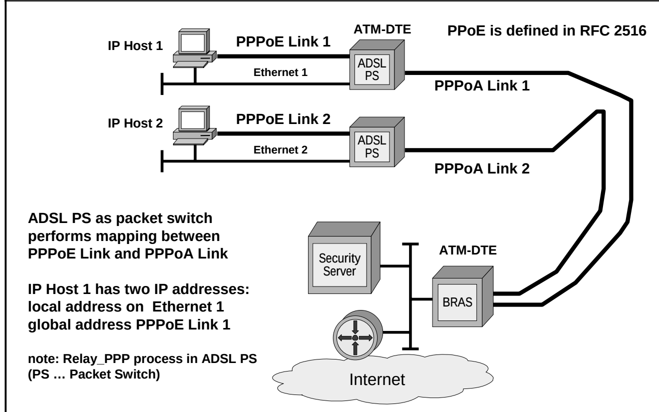
ADSL: PPP over ATM (PPPoA)



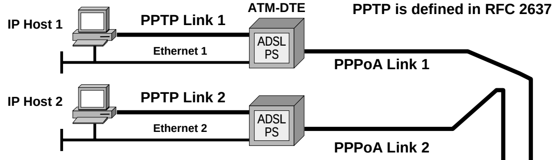
ADSL: PPP over ATM (PPPoA), IPCP



ADSL: PPP over Ethernet (PPPoE)



ADSL: PPTP over Ethernet (Microsoft VPN)



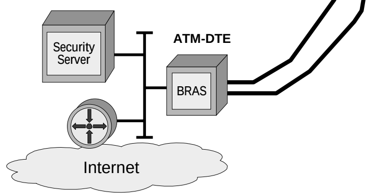
PPTP is defined in RFC 2637

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

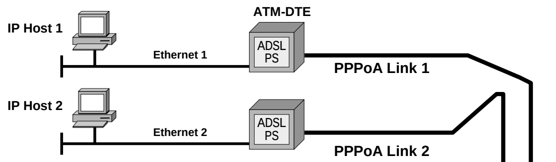
ADSL PS as packet switch performs mapping between PPTP Link and PPPoA Link

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

note: Relay_PPP process in ADSL PS



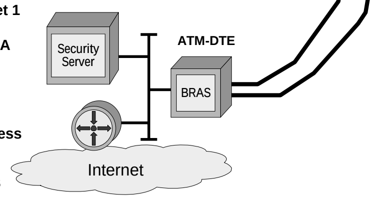
ADSL: Routed PPPoA



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)



Agenda

- **Introduction**
- **IP**
  - IP Protocol
  - Addressing
- **IP Forwarding**
- **<u>First Hop Redundancy</u>**
  - <u>Proxy ARP, IDRP</u>
  - HSRP
  - VRRP

First Hop Redundancy (Layer 3) 1

- **The problem:**
  - How can <u>local routers</u> be recognized by IP hosts?
  - Note: Normally IP host has limited view of topology
    - IP host knows to which IP subnet connected
    - IP host knows <u>one</u> “Default Gateway” to reach other IP networks
  - Static configuration of “Default Gateway”:
    - Loss of the default router results in a catastrophic event, isolating all end-hosts that are unable to detect any alternate path that may be available
- **Two design philosophies:**
  - Solve the problem at the IP host level
    - OS of the IP host need to support certain functionality in a appropriate way
  - Solve the problem at the IP router level
    - OS of the IP host need to support the basic functionality only
      - that is static configuration of one “Default Gateway”
    - Proprietary functionality may be needed at the router

First Hop Redundancy (Layer 3) 2

- **Methods for solving it at the IP host level:**
  – Proxy ARP
  – IDRP
  – DHCP
  – IP Routing (RIPv2, OSPF)
- **Methods for solving it at the IP router level:**
  – HSRP
  – VRRP
  – GLBP

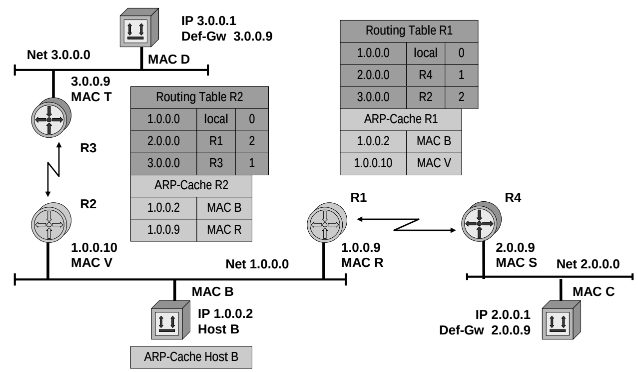
Old Proxy ARP Usage

- **Old method for efficient use of address space**
  – If two networks coupled by a router need to have the same IP Net-ID
    - e.g. for the time a bridged network should be migrated to a routed network a proxy ARP component must be installed in the network component to be migrated (bridge –>router)
  – Term "proxy" means "instead of"
    - some system is doing some function instead of the expected system
- **Replaced nowadays by IP subnetting**

Proxy ARP Usage Nowadays

- **Proxy ARP is can be used if an IP host didn't know the address of the default gateway or find it out dynamically:**
  – 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

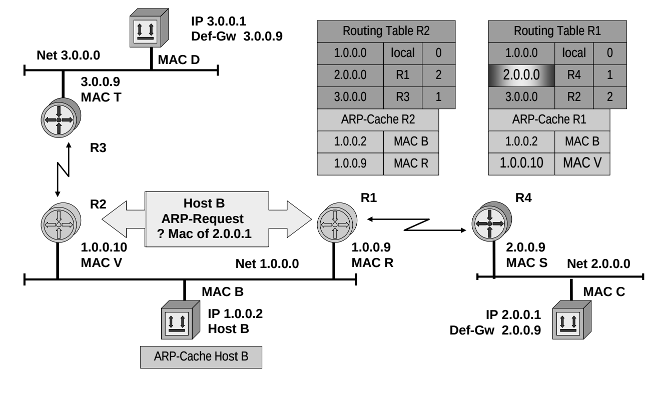
1.0.0.2 -> 3.0.0.1 / 2.0.0.1 with proxy ARP 1



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

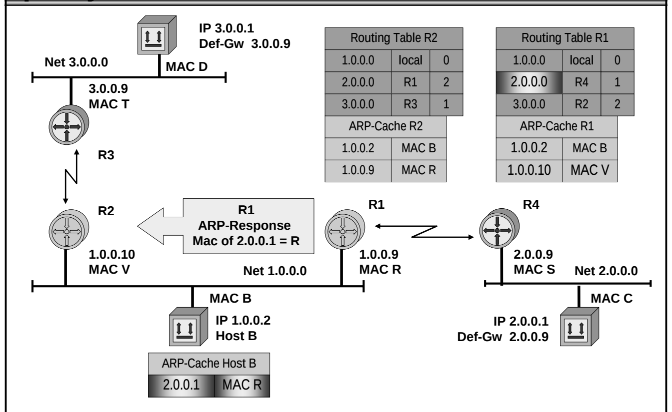
L08 - IP Technology

1.0.0.2 -> 3.0.0.1 / 2.0.0.1 with proxy ARP 2



 IP Technology, v4.8 165

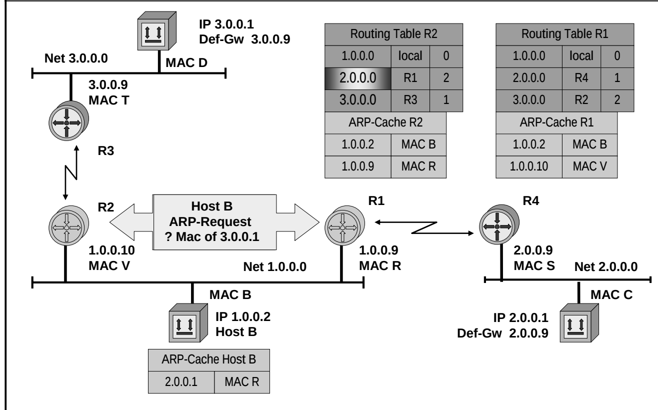
1.0.0.2 -> 3.0.0.1 / 2.0.0.1 with proxy ARP 3



 IP Technology, v4.8 166

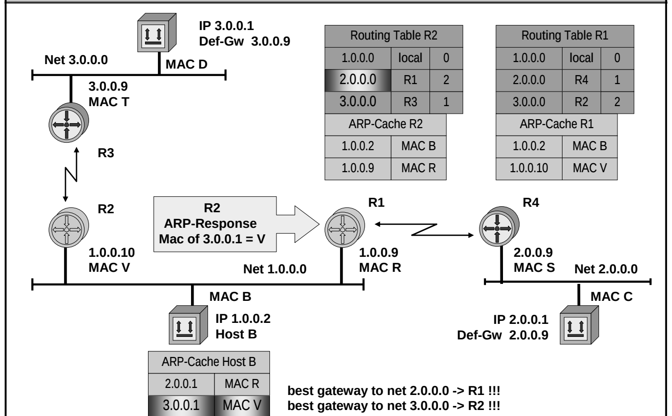
L08 - IP Technology

1.0.0.2 -> 3.0.0.1 / 2.0.0.1 with proxy ARP 4



 IP Technology, v4.8 167

1.0.0.2 -> 3.0.0.1 / 2.0.0.1 with proxy ARP 5



 IP Technology, v4.8 168

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

Other Techniques to Solve the Problem 2

- **With IDRP and DHCP**
  - You still depend on OS functionality in order to trigger switchover between redundant local routers
    - How often the currently selected router will be tested for reachability? What is if the currently selected router is reachable via LAN but networks behind are not reachable?
- **Therefore running a classical IP routing protocol on the IP host would be optimal**
  - RIPv2
    - But slow convergence if the currently selected router fails, no hello messages hence 180 seconds for recognizing that event
  - OSPF
    - Fast convergence because of hello messages, the best but the most complex solution

Datenkommunikation 384.081 - SS 2008

L08 - IP Technology

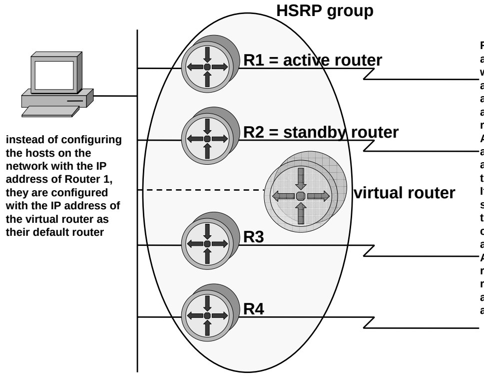
Agenda

- **Introduction**
- **IP**
  - IP Protocol
  - Addressing
- **IP Forwarding**
- **<u>First Hop Redundancy</u>**
  - Proxy ARP, IDRP
  - <u>HSRP</u>
  - VRRP

HSRP – Hot Standby Router Protocol

- **HSRP (<u>H</u>ot <u>S</u>tandby <u>R</u>outer <u>P</u>rotocol)**
  - Proprietary protocol invented by Cisco
  - RFC 2281 (Informational)
- **Basic idea: a set of routers present the illusion of a single virtual router to the hosts on the LAN**
  - Active router
    - one router is responsible for forwarding the packets that hosts send to the virtual router
  - Standby router
    - if active router fails, the standby assumes the packet forwarding duties of the active router
  - Conspiring routers form a HSRP group

Terminology



Router 1 is configured as the active router. It is configured with the IP address and MAC address of the virtual router and sends any packets addressed to the virtual router out interface.
As the standby router, R2 is also configured with the IP address and MAC address of the virtual router.
If for any reason Router 1 stops transferring packets, the routing protocol converges, and Router 2 assumes the duties of Router A and becomes the active router. That is, Router 2 now responds to the virtual IP address and the virtual MAC address.

HSRP Operation 1

- **Principle:**
  - A group of routers forms a HSRP group
  - The group is represented by a virtual router
    - With a virtual IP address and virtual MAC address for that group
  - IP hosts are configured with the virtual IP address as default gateway
  - One router is elected as the active router, one router is elected as the standby router of that group
  - Active router responds to ARP request directed to the virtual IP address with the virtual MAC address
  - Standby router supervise if the active router is alive and can take over the role of the active router
    - HSRP protocol using UDP messages to port 1985, IP multicast 224.0.0.2, and Ethernet multicast as destination address
  - Router must be able to support more than one unicast MAC address on an Ethernet interface

HSRP Operation 2

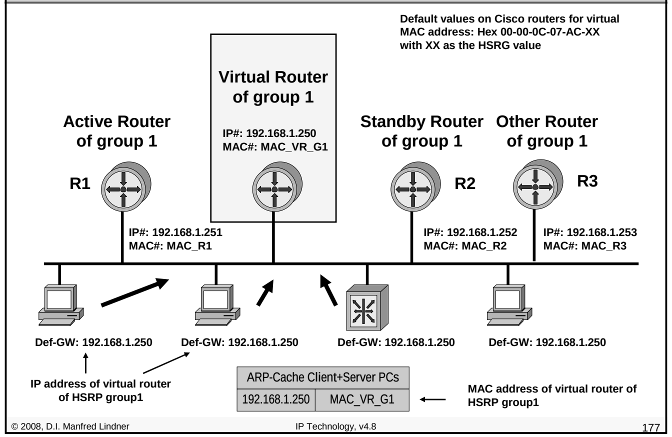
- **Roles of router:**
  - Active, Standby, Other defined by HSRP priority
  - Priority value can be configured
    - Default value is 100
  - The higher the better
    - Will become the active router after initialization
    - If priority is equal than the higher IP address decides
  - Preempt allows to give up the role of the active router when a router with higher priority is activated or reported
    - e.g. a failed router comes back or tracking has changed priority
- **Load Balancing:**
  - Specify at least two different HSRP groups with complementary roles
- **HSRP authentication:**
  - Based on keyed MD5
  - Against HSRP spoofing

HSRP Operation 3

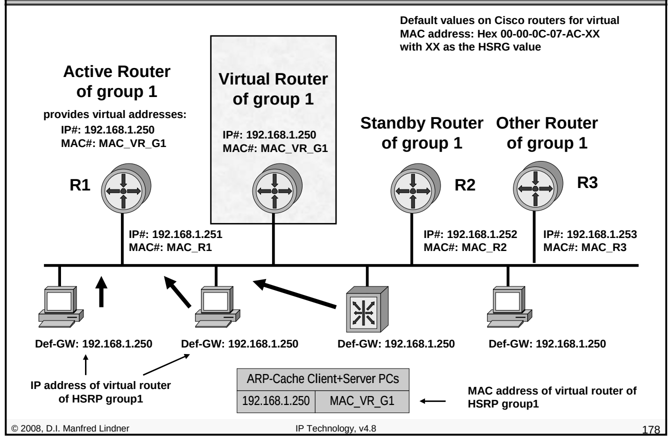
- **Failover scenarios:**
  - Active router not reachable via LAN
    - Standby router will take over active role
    - A new standby router is elected from the remaining routers of a HSRP group
    - Timing depends on HSRP hello message interval and hold-time
      - Default hello-time = 3 seconds, default hold-time = 10 seconds
  - Active router losses connectivity to a WAN interface (basic tracking options) or losses connectivity to an IP route (enhanced tracking options)
    - If tracking and preempt is configured standby router will take over
      - Tracking will lower the priority
      - Preempt allows another router to take over the role of the active router even if the current active router does not fail
  - Enhanced tracking options depend on IOS version

L08 - IP Technology

HSRP 1

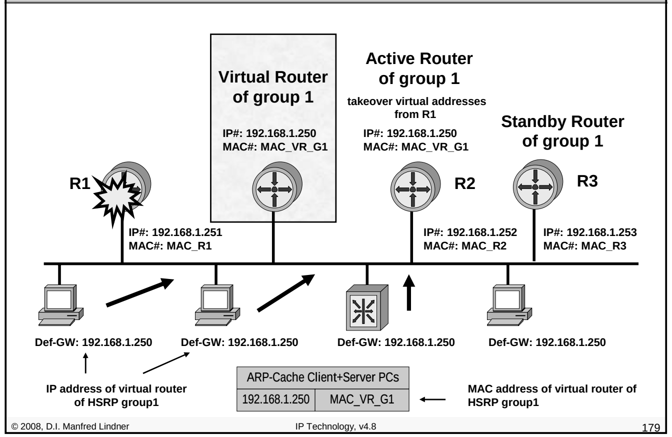


HSRP 2

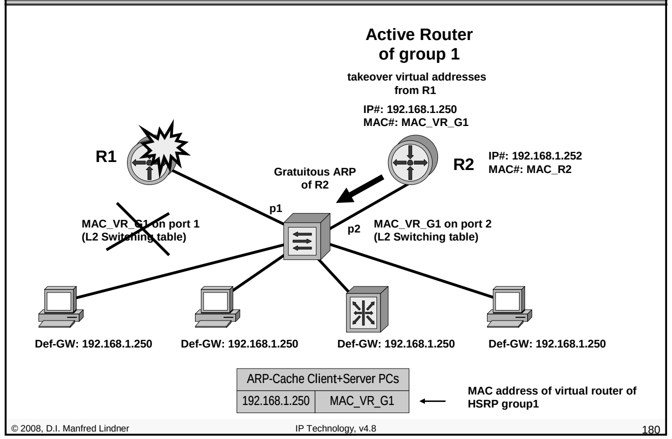


L08 - IP Technology

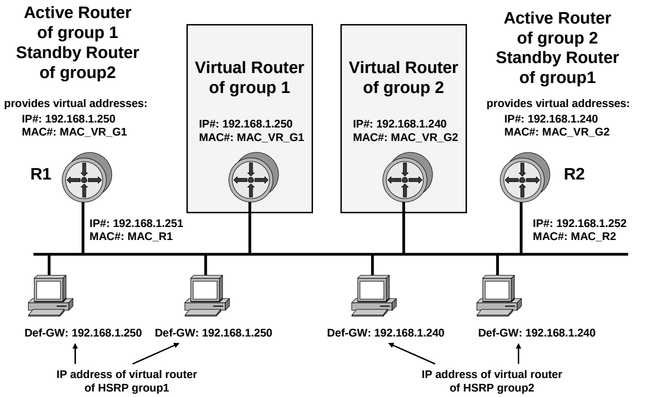
HSRP 3



HSRP – Gratuitous ARP



HSRP Load Balancing



 IP Technology, v4.8 181

HSRP Versions

- **HSRP version 1:**
  - Second timers
  - 256 groups (0 – 255)
  - Virtual Mac Address: 00-00-0C-07-AC-XX
    - XX value = group number
  - IP multicast 224.0.0.2
- **HSRP version 2:**
  - Millisecond timers
    - Hello-time 15 - 999 msec
    - Hold-time - 3000 msec
  - 4096 groups (0-4095)
    - Allow a group number to match the VLAN-ID
  - Virtual Mac Address: 00-00-0C-9F-FX-XX
    - X-XX value = group number
  - IP multicast 224.0.0.102
    - To avoid conflicts with CGMP (Cisco Group Management Protocol)

 IP Technology, v4.8 182

HSRP Protocol Fields

- **standby protocol runs on top of UDP (port 1985)**
  - IP packets are sent to multicast address 224.0.0.2 or 224.0.0.102 with a IP TTL = 1

0 4	8	16	31
Version	Op Code	State	Hellotime
Holdtime	Priority	Group	Reserved
Authentication Data			
Authentication Data			
Virtual IP Address			

- **version**: version of the HSRP messages
- **op code**: 3 types
  - hello: indicates that a router is running and is capable of becoming the active or standby router
  - coup: when a router wishes to become the active router
  - resign: when a router no longer wishes to be the active router
- **states**: initial, learn, listen, speak, standby, active
- **hellotime**: contains the period between the hello messages that the router sends
- **holdtime**: amount of time the current hello message is valid
- **priority**: compares priorities of 2 different routers
- **group**: identifies standby group (0...255)
- **authentication data**: cleartext 8 character reused password

 IP Technology, v4.8 183

Agenda

- **Introduction**
- **IP**
  - IP Protocol
  - Addressing
- **IP Forwarding**
- **<u>First Hop Redundancy</u>**
  - Proxy ARP, IDRP
  - HSRP
  - <u>VRRP</u>

 IP Technology, v4.8 184

VRRP Operation 1

- **VRRP (<u>V</u>irtual <u>R</u>outer <u>R</u>edundancy <u>P</u>rotocol)**
  - RFC 2338 (Standards Track)
- **Principle:**
  - A group of routers forms a VRRP group
  - The group is represented by a virtual router
    - With is identified by a VRID (<u>V</u>irtual <u>R</u>outer <u>ID</u>) and a virtual MAC address
  - One router is elected as the **<u>virtual router master</u>**, all other routers get the role of **<u>virtual router backup</u>** routers
  - The real IP address of the virtual router master become the IP address of the virtual router for a given VRRP group
    - <u>IP address owner</u>
  - Default Gateway of IP hosts is configured with the IP address of the virtual router for a given VRRP group
  - Virtual router master responds to ARP request directed to the IP address of the virtual router with the virtual MAC address
  - Backup routers supervise if master router is alive and take over the role of the master in case of failure
    - VRRP protocol using IP protocol number 112, IP multicast 224.0.0.18, and Ethernet multicast as destination address
  - Router must be able to support more than one unicast MAC address on an Ethernet interface

VRRP Operation 2

- **Roles of router:**
  - Virtual router master, virtual router backup defined by VRRP priority
  - Priority value can be configured
    - Default value is 100
  - The higher the better
    - Will become the master after initialization
    - If priority is equal than the higher IP address decides
  - Preempt allows to give up the role of the master router when a router with higher priority is activated or reported
    - e.g. a failed router comes back or tracking has changed priority
- **Load Balancing:**
  - Specify at least two different VRRP groups with complementary roles
- **VRRP authentication:**
  - Based on keyed MD5
  - Against VRRP spoofing

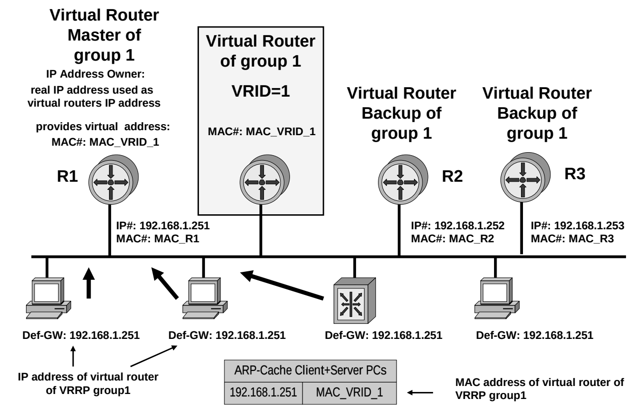
Datenkommunikation 384.081 - SS 2008

L08 - IP Technology

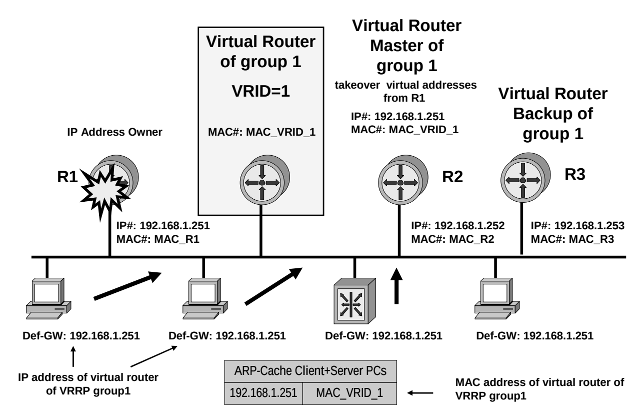
VRRP Operation 3

- **Failover scenarios:**
  - Master router not reachable via LAN
    - Backup router with highest priority will take over master role
    - Timing depends on VRRP advertisements interval and master down interval
      - Default advert-interval = 1 seconds
      - Default master-down-interval = 3 * advert-interval + skew-time
  - Master router losses connectivity to a WAN interface (basic tracking options) or losses connectivity to an IP route (enhanced tracking options)
    - If tracking and preempt is configured backup router will take over
      - Tracking will lower the priority
      - Preempt allows another router to take over the role of the master router even if the current master router does not fail
  - Enhanced tracking options depend on IOS version

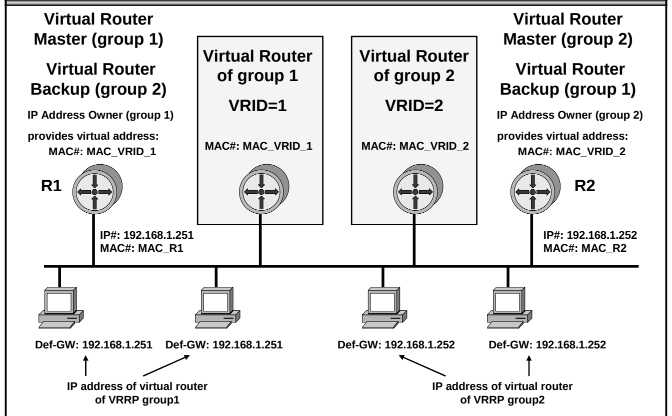
VRRP 1



VRRP 2



VRRP Load Balancing



Some VRRP Details

- **VRRP:**
  - Second or milliseconds timers
  - VRID range
    - 1 – 255
    - Maximum 255 groups
  - Virtual Mac Address: 00-00-5E-00-01-VRID
    - VRID value = group number
  - IP multicast 224.0.0.18

VRRP Protocol Fields

0	4	8	16	31
Version	Type	Virtual Rtr ID	Priority	Count IP Addrs
Auth Type		Advert Int	Checksum	
IP Address 1				
...				
IP Address n				
Authentication Data 1				
Authentication Data 2				

- Version - This version is version 2.
- Type - The only packet type defined in this version of the protocol is: 1 ADVERTISEMENT.
- Virtual Rtr ID - The Virtual Router Identifier (VRID) field identifies the virtual router this packet is reporting status for.
- Priority - VRRP routers backing up a virtual router MUST use priority values between 1-254 (decimal).
- Count IP Addresses -The number of IP addresses
- Auth Type - Identifies the authentication method being utilized.
- Advertisement Interval - Indicates the time interval (in seconds) between advertisements.
- Checksum - used to detect data corruption
- IP Address(es) - One or more IP addresses that are associated with the virtual router.
- Authentication Data - The authentication string is currently only utilized for simple text authentication