Introduction to IP Routing

Static-, Default-, Dynamic-Routing, RIP
Classless/Classful Routing, Internet Routing Overview

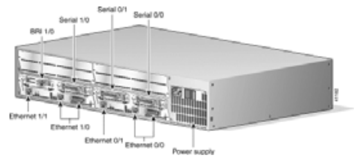
Agenda

- **Introduction to IP Routing**
  - Static Routing
  - Default Route
  - Dynamic Routing
- **RIP**
- **Classful versus Classless Routing**
- **Private Addresses and NAT**
- **Internet Routing**

IP Router

- **Initially Unix workstations with several network interface cards**
- **Today specialized hardware**

Cisco 3600 Router



IP Routing 1

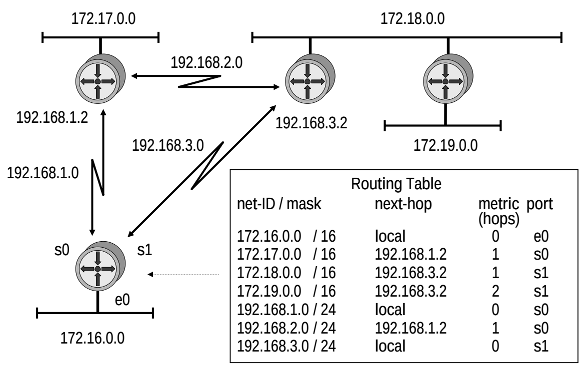
- **routing**
  - process of choosing a path over which to send IP datagram's
  - direct versus indirect delivery
    - depends on destination net-ID
      - net-ID equal to source net-ID -> direct delivery
      - net-ID unequal to source net-ID -> indirect delivery
  - IP hosts and routers take part in this process
    - IP hosts responsible for direct delivery of IP datagram's
    - IP hosts responsible for choosing a default router (“default gateway”) as next hop in case of indirect delivery of IP datagram's
    - routers responsible for selecting the best path in a meshed network in case of indirect delivery of IP datagram's

IP Routing 2

- **indirect routing of IP datagram's**
  - is done by routers based on routing tables
  - routing table
    - database of known destinations
  - database contains
    - next hop router (and next hop MAC address in case of LAN)
    - outgoing port
    - metric (information how far away is a certain destination network)
    - time reference (information about the age of the table entry)

    for every known (or specified) destination network
    - net-ID / subnet-mask

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

Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

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
    - note:some methods allow load balancing if paths are equal

IP Routing Types

- **routing can be either**
  - static
    - routing tables are preconfigured by network administrator
    - non-responsive to topology changes
    - can be labor intensive to set up and modify in complex networks
    - no overhead concerning CPU time and traffic
  - or dynamic
    - routing tables are dynamically updated with information received from other routers
    - responsive to topology changes
    - low maintenance labor cost
    - communication between routers is done by routing protocols using routing messages for their communication
    - routing messages need a certain percentage of bandwidth
    - dynamic routing need a certain percentage of CPU time of the router
    - that means overhead

© 2008, D.I. Manfred Lindner

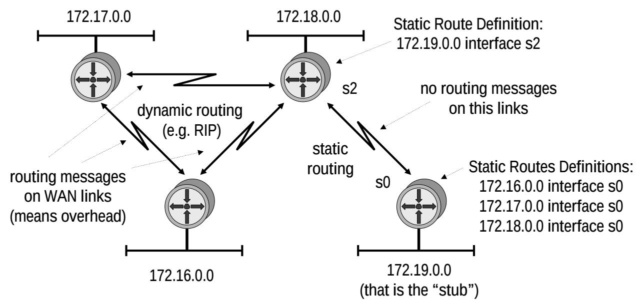
Agenda

- **Introduction to IP Routing**
  - Static Routing
  - Default Route
  - Dynamic Routing
- **RIP**
- **Classful versus Classless Routing**
- **Private Addresses and NAT**
- **Internet Routing**

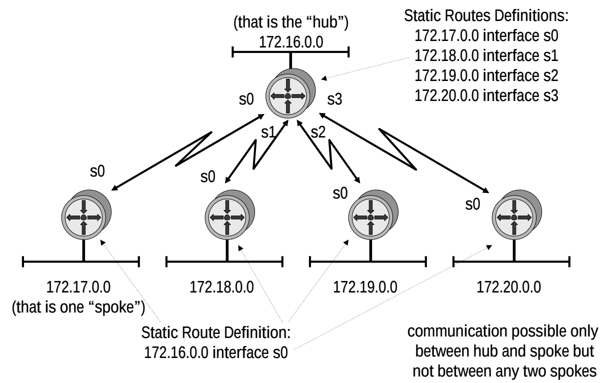
Static Routing

- **static routing**
  - preconfigured static routing tables
  - no overhead traffic
  - often sufficient in case of lack of any network redundancy
    - e.g. reaching stub networks
    - e.g. hub and spoke topology
  - but can be labor intensive to set up and modify in complex networks
    - overhead can be reduced by default route
  - sometimes only or preferred way in certain technologies
    - Dial on Demand Networks (e.g. X.25, ISDN, Frame Relay, ATM)
  - sometimes used for security reasons

Static Routes - Stub Network



Static Routes - Hub and Spoke



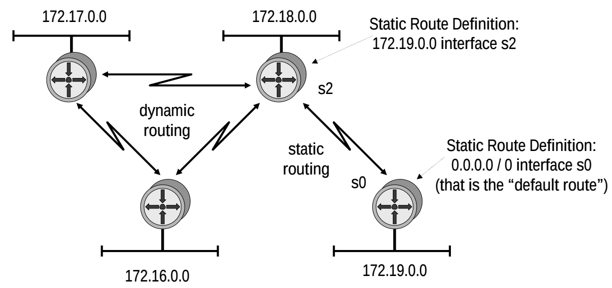
Agenda

- **<u>Introduction to IP Routing</u>**
  - Static Routing
  - <u>Default Route</u>
  - Dynamic Routing
- **RIP**
- **Classful versus Classless Routing**
- **Private Addresses and NAT**
- **Internet Routing**

Default Route

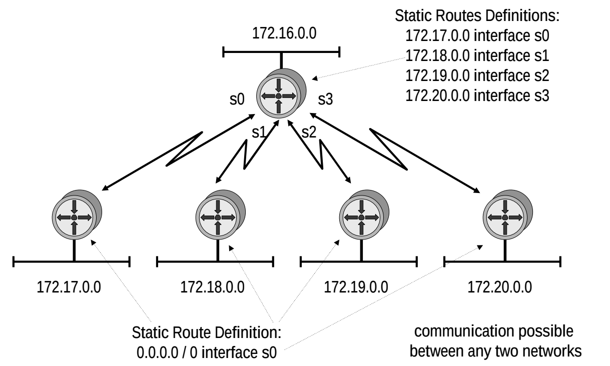
- **general routing principle**
  - traffic to destinations that are unknown to the router will be discarded by the router (ICMP message !!!)
  - behavior can be changed by default route
- **default routing principle**
  - traffic to destinations that are unknown to the router will be sent to a <u>default route</u> (default network)
  - implies that another router might know more networks
  - permits routers to carry less than full routing tables
  - default network marked with net-ID equal 0.0.0.0
    - in routing tables
    - in routing updates used by dynamic routing

Default Route - Stub Network

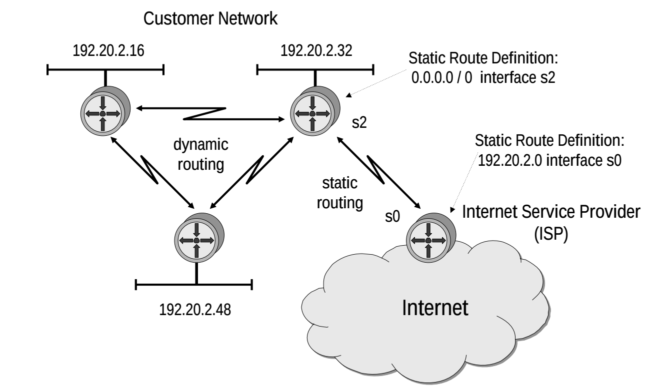


Questions: What happens to traffic generated in network 172.19.0.0 with an unknown destination address? What does this mean for the WAN link?

Default Route - Any to Any



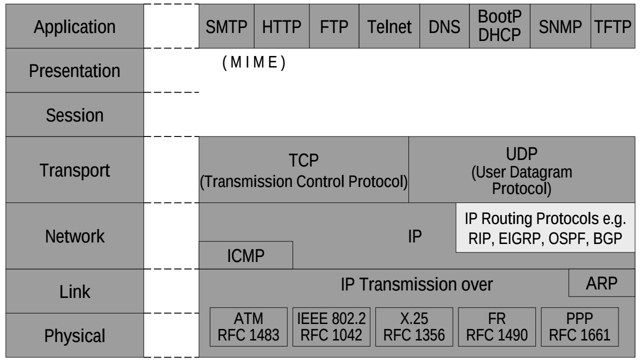
Default Route - Internet Access



Agenda

- **<u>Introduction to IP Routing</u>**
  - Static Routing
  - Default Route
  - <u>Dynamic Routing</u>
- **RIP**
- **Classful versus Classless Routing**
- **Private Addresses and NAT**
- **Internet Routing**

IP (Dynamic) Routing Protocols



Dynamic Routing

- **dynamic routing**
  - routing tables are dynamically updated with information from other routers done by routing protocols
  - routing protocol
    - discovers current network topology
    - determines the best path to every reachable network
    - stores information about best paths in the routing table
  - metric information is necessary for best path decision
    - in most cases summarization along the a given path of <u>static</u> preconfigured values
      - hops, interface cost, interface bandwidth, interface delay, etc.
  - two basic technologies
    - distance vector, link state

Distance Vector Protocols 1

- **routing table is sent periodically to all immediately-neighboring routers (IP limited broadcast)**
  - after power-up routing table contains information about local attached networks only
- **incoming updates are checked for changes**
  - new networks, metric change of already known networks, etc.
- **own routing table is adapted accordingly**
  - changes announced by next periodic routing update

Distance Vector Protocols 2

- **metric information based on hops (distance between hops)**
- **limited view of topology**
  - routers view is based on its routing table only
    - exact view how to reach local neighbors
    - but topology behind neighbors is hidden
  - based on signpost principle only
- **several procedures necessary**
  - to solve problems caused by limited view
    - e.g. count to infinity, routing loops
  - to reduce convergence time
    - time to reach consistent routing tables after topology change

Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

Distance Vector Protocols 3

- **some usual procedures for to solve inherent problems**
  - maximum hop count
  - split horizon, poison reverse
  - triggered update
  - hold down, route poisoning
- **distance vector protocols examples**
  - RIP, RIPv2 (Routing Information Protocol)
  - IGRP (Cisco, Interior Gateway Routing Protocol)
  - IPX RIP (Novell)
  - AppleTalk RTMP (Routing Table Maintenance Protocol)

Link State Protocols 1

- **routers have a global view of network topology**
  - exact knowledge about all routers, links and their costs (metric) of a network stored in topology database (“roadmap”)
  - roadmap principle
- **routing table entries are based**
  - on computation of own router-resident topology database
- **SPF computation**
  - **S**hortest-**P**ath-**F**irst (Dijkstra) algorithm to find lowest cost path to every destination network
  - lowest cost path is stored in routing table

Link State Protocols 2

- **topology changes (link up or down, link state)**
  - are recognized by routers responsible for supervising those links
  - are flooded by responsible routers to the whole network (**L**ink **S**tate **A**dvertisements, LSAs)
- **flooding**
  - is a controlled multicast procedure to guarantee that every router gets corresponding LSA information as fast as possible
  - is used to update network topology database and hence may lead to change of routing table

Link State Protocols 3

- **with the lack of topology changes**
  - local hello messages are used to supervise local links (to test reachability of immediate-neighboring routers)
  - therefore less routing overhead concerning link bandwidth than periodic updates of distance vector protocols
- **but more network load is caused by such a routing protocol**
  - during connection of former separate parts of the network
  - topology database synchronization

Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

Link State Protocols 4

- **link state protocols examples**
  - OSPF (Open Shortest Path First)
  - Integrated IS-IS (IP world)
    - note: Integrated IS-IS takes another approach to handle large networks (topic outside the scope of this course)
  - IS-IS (OSI world)
  - PNNI (in the ATM world)
  - APPN (IBM world),
  - NLSP (Novell world)

Routing Protocol Comparison

Routing Protocol	Complexity	Max. Size	Convergence Time	Reliability	Protocol Traffic
RIP	very simple	16 Hops	High (minutes)	Not absolutely loop-safe	High
RIPv2	very simple	16 Hops	High (minutes)	Not absolutely loop-safe	High
IGRP	simple	x	High (minutes)	Medium	High
EIGRP	complex	x	Fast (seconds)	High	Medium
OSPF	very complex	Thousands of Routers	Fast (seconds)	High	Low
IS-IS	complex	Thousands of Routers	Fast (seconds)	High	Low
BGP-4	very complex	more than 100,000 networks	Middle	Very High	Low

Agenda

- **Introduction to IP Routing**
  - Static Routing
  - Default Route
  - Dynamic Routing
- **<u>RIP</u>**
- **Classful versus Classless Routing**
- **Private Addresses and NAT**
- **Internet Routing**

RIP - Routing Information Protocol

- **RIP is an Interior Gateway Protocol (IGP)**
  - due to inherent administrative traffic, RIP suits best for smaller networks
  - routing decisions are based upon hop count measure
- **RIP was initially released as part of BSD 4.2 UNIX**
  - hence RIP got wide-spread availability
- **RIP is specified in RFC 1058**
  - RFC category „historic“

Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

RIP Principles 1

- **RIP is a Distance Vector Protocol**
- **after booting the non-volatile configuration-memory tells a RIP router to which networks it is <u>directly</u> connected**
- **this information is loaded into the routing table**
- **basically the routing table contains**
  - the net-ID of the directly connected networks
  - and the associated distance (in hops) to them
    - directly connected networks have hop-count = 0

RIP Principles 2

- **then, this routing table is distributed periodically (every 30 seconds) to all connected networks = <u>routing update</u>**
  - using a broadcast MAC-frame containing
  - an IP-broadcast datagram containing
  - an UDP-datagram with port number 520
  - <u>metric entries of the routing table will be risen by the distance</u> <u>of the interface where transmitted</u>-> in case of RIP -> distance is one -> hop count
- **directly reachable routers**
  - receive this message, update their own routing tables,
  - and hence generate their own routing updates reflecting any corresponding modifications

RIP Principles 3

- **after a specific time**
  - all routers know about all network addresses of the whole network
- **if different routing updates (from different routers) contain the same net-ID**
  - then there are redundant paths to this network
  - only the path with the lowest hop-count is stored in the routing-table
  - on receiving equal hop counts, the net-ID of the earlier one will be selected (and all other associated data)
- **hence, between each two networks exists exactly one active path**

RIP Principles 4

- **all routing table's net-IDs are periodically refreshed by routing-update messages**
- **if a routing update tells a <u>better</u> metric than that one currently stored in the table**
  - the routing table must be updated with this new information
  - this update does not take care about if the sender of this routing-update is also the router which is currently selected as next hop
  - "good news" are quickly adapted
  - RIP trusts good news from <u>any</u> source ("trusted news")

Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

RIP Principles 5

- **if a routing update tells a <u>worse</u> metric than that one currently stored in the table**
  - the routing table must be updated with this new information if the sender of this routing-update is the next-hop router for this network
    - that is: the actual VECTOR in the table is identical with the source address of the routing-update
  - routing-updates from other routers than that one currently registered in the table are ignored
- **summary: routing-updates with worse metric is only relevant if it comes from that router mentioned in the actual table entry**

RIP Principles 6

- **when a routing table entry is not refreshed within 180sec**
  - this entry is considered to be obsolete
  - possible reasons: router-failure, network not reachable
  - without special mechanism
    - we have to wait for 180sec at least in order that all routers have consistent routing tables again
  - improvement by a special network-unreachable message
    - which is distributed to all other routers
    - it takes 180sec in the worst case
  - slow adaptation of "bad news"
  - during these 180sec, forwarding of IP datagram's is done according to the routing table !!

RIP Header plus RIP Data

0	8	16	31
command	version	must be zero	
family of net 1		must be zero	
IP address of net 1			
must be zero			
must be zero			
distance to net 1			
family of net 2		must be zero	
IP address of net 2			
must be zero			
must be zero			
distance to net 2			
family of net 3		must be zero	
. . . .			

RIP Header and Data Fields

- **command**
  - Request (1): Router or Host requests for a routing update
  - Response (2): response to a request but also used for periodic routing-updates
- **version**
  - version number of the RIP protocol (= 1 for RIP)
- **family of net x**
  - because RIP is not only build for IP (in the case of IP: 2)
- **IP address of net x**
  - IP-address of the announced network x
- **distance to net x**
  - = hop-count to net x

Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

RIP Summary

to give a clearer idea of the RIP method:

- **every router holds a directory, pointing to other networks (VECTOR entries) without knowing the exact location of them**
- **datagram's follow these "signposts" and finally reach their targets**
- **information about these signposts is based upon rumours (routing-updates)**

Creating Routing Tables

Routing table A

Net	Hops	Hop-ID
1	0	direct
2	0	direct

Routing table B

Net	Hops	Hop-ID
2	0	direct
3	0	direct

Situation after booting

Creating Routing Tables

Net	Hops
1	1
2	1

Router A: routing update to net 2

values used by router B to actualise its routing table

Routing table A

Net	Hops	Hop-ID
1	0	direct
2	0	direct

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	0	direct

Situation after distribution of A's routing update

Creating Routing Tables

Net	Hops
1	2
2	1
3	1

Router B: routing update to net 2

Routing table A

Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	1	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	0	direct

Situation after distribution of B's routing update

Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

Routing Updates and Stable States

Router A: routing update to net 2

Net	Hops
1	1
2	1
3	2

Net	Hops
1	2
2	1
3	1

Router B: routing update to net 2

Routing table A

Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	1	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	0	direct

Routing tables keep stable state

RIP-Problems

- **slow convergence**
  - due to 30s routing-update interval and 180s aging timeout
- **routing-loops cannot be safely detected**
  - because routing updates are obligatory ("Trusted Information Principle" of RIP)
- **so, failures and rooting-loops keep large networks inconsistent for a long period of time**
  - datagram's circle around along redundant paths
  - "Count to Infinity" -problem

Count to Infinity 1

Router A: routing update to net 2

Net	Hops
1	1
2	1
3	2

Router B: routing update to net 2

Net	Hops
1	2
2	1
3	1

Routing table A

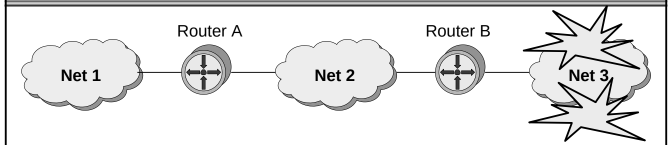
Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	1	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	0	direct

Initial situation

Count to Infinity 2



Routing table A

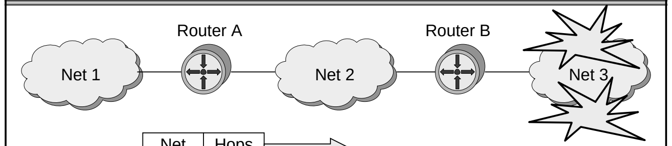
Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	1	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	?	?

Failure: net 3 unreachable

Count to Infinity 3



Router A: routing update to net 2

Net	Hops
1	1
2	1
3	2

Routing table A

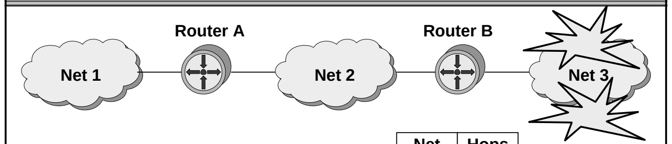
Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	1	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	2	2.A

Situation after sending A's routing update

Count to Infinity 4



Router B: routing update to net 2

Net	Hops
1	2
2	1
3	3

Routing table A

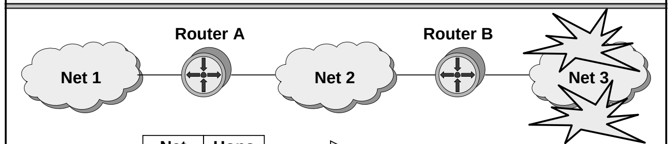
Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	3	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	2	2.A

Situation after sending B's routing update

Count to Infinity 5



Router A: routing update to net 2

Net	Hops
1	1
2	1
3	4

problem: routing updates do not contain explicit VECTOR-information !!

(Vector is given only implicitly through source address)

Routing table A

Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	3	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	4	2.A

Situation after sending A's next routing update

Solutions against Count to Infinity

- **<u>problem:</u> good news are distributed faster than bad news**
  - the information of a path with lower hop-count is distributed every 30 seconds
  - the information of a network-failure is distributed at least after 180 (or later)
- **<u>remedy:</u>**
  - Maximum Hop Count
  - Split Horizon
  - Poison Reverse
  - Triggered Update
  - Hold Down

Maximum Hop Count

- **maximal distance between each two subnets is limited to 16**
  - hop count between two end-systems cannot exceed 15
- **a DISTANCE-value of 16 in the routing-table means that the corresponding network is not reachable**
  - using hop count = 16 in a routing update allows a router to indicate the failure of a network
  - we have not to age out this entry in all routing tables hence waiting at least for 180 s
  - IP-datagram's with a net-ID pointing to such an entry are discarded by the router and additionally a ICMP message "network unreachable" is generated

Split Horizon 1

- **Maximum Hop Count technique alone (counting to 16) does not avoid routing-loops !**
- **to overcome routing-loops and also slow-convergence, <u>Split-Horizon</u> has been introduced**
  - prevents router from sending information about reachability of networks in that direction from where the information originally came
- **exception of this rule:**
  - if the router knows a better path

Split Horizon 2

Router A: routing update to net 2

Net	Hops
1	1

Router B: routing update to net 2

Net	Hops
3	1

Routing table A

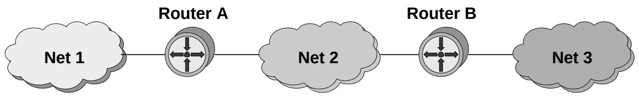
Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	1	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	0	direct

Routing update to net 2

Split Horizon 3



Router A: routing update to net 1

Net	Hops
2	1
3	2

Router B: routing update to net 3

Net	Hops
1	2
2	1

Routing table A

Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	1	2.B

Routing table B

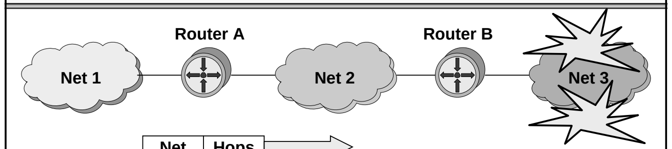
Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	0	direct

Routing update to net1 and net 3

Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

Net 3 unreachable



Router A: routing update to net 2

Net	Hops
1	1

Routing table A

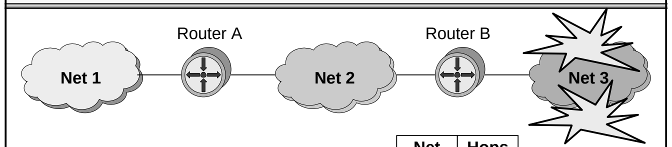
Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	1	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	16	-

Situation after sending A's routing update

Net 3 unreachable



Router B: routing update to net 2

Net	Hops
3	16

Routing table A

Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	16	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	16	-

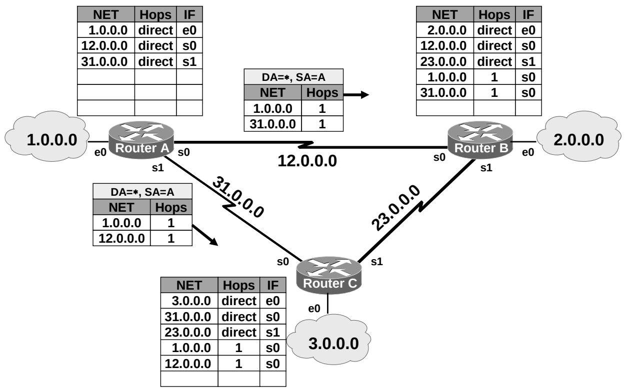
Situation after sending B's routing update

Split Horizon Facts

- **Don't tell neighbour of routes that you learned from this neighbour**
- **That's what humans (almost) always do:**

  <u>Don't tell me what I have told you!</u>
- **Using split horizon a router will not send information about routes he isn't really aware of**
- **the convergence time**
  - is reduced to the time of failure-detection (180 s)
  - note: method of Maximum Hop Count alone would take 16 x 30 s = 480 s in the worst case

RIP At Work (Update Router A)



NET	Hops	IF
1.0.0.0	direct	e0
12.0.0.0	direct	s0
31.0.0.0	direct	s1

DA=∗, SA=A	
NET	Hops
1.0.0.0	1
31.0.0.0	1

NET	Hops	IF
2.0.0.0	direct	e0
12.0.0.0	direct	s0
23.0.0.0	direct	s1
1.0.0.0	1	s0
31.0.0.0	1	s0

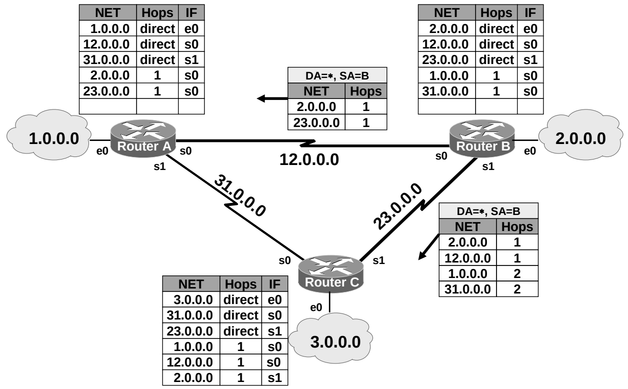
DA=∗, SA=A	
NET	Hops
1.0.0.0	1
12.0.0.0	1

NET	Hops	IF
3.0.0.0	direct	e0
31.0.0.0	direct	s0
23.0.0.0	direct	s1
1.0.0.0	1	s0
12.0.0.0	1	s0

Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

RIP At Work (Update Router B)



NET	Hops	IF
1.0.0.0	direct	e0
12.0.0.0	direct	s0
31.0.0.0	direct	s1
2.0.0.0	1	s0
23.0.0.0	1	s0

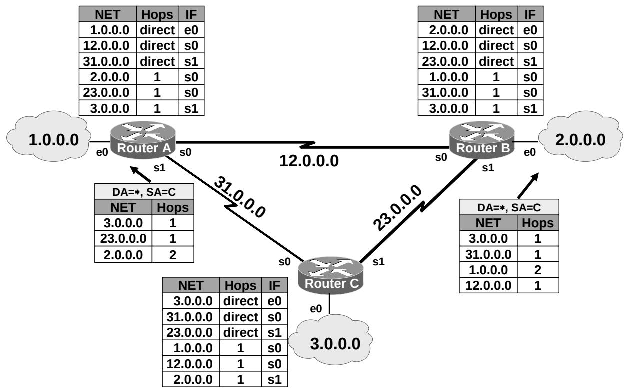
DA=∗, SA=B	
NET	Hops
2.0.0.0	1
23.0.0.0	1

NET	Hops	IF
2.0.0.0	direct	e0
12.0.0.0	direct	s0
23.0.0.0	direct	s1
1.0.0.0	1	s0
31.0.0.0	1	s0

DA=∗, SA=B	
NET	Hops
2.0.0.0	1
12.0.0.0	1
1.0.0.0	2
31.0.0.0	2

NET	Hops	IF
3.0.0.0	direct	e0
31.0.0.0	direct	s0
23.0.0.0	direct	s1
1.0.0.0	1	s0
12.0.0.0	1	s0
2.0.0.0	1	s1

RIP At Work (Update Router C)



NET	Hops	IF
1.0.0.0	direct	e0
12.0.0.0	direct	s0
31.0.0.0	direct	s1
2.0.0.0	1	s0
23.0.0.0	1	s0
3.0.0.0	1	s1

NET	Hops	IF
2.0.0.0	direct	e0
12.0.0.0	direct	s0
23.0.0.0	direct	s1
1.0.0.0	1	s0
31.0.0.0	1	s0
3.0.0.0	1	s1

DA=∗, SA=C	
NET	Hops
3.0.0.0	1
23.0.0.0	1
2.0.0.0	2

DA=∗, SA=C	
NET	Hops
3.0.0.0	1
31.0.0.0	1
1.0.0.0	2
12.0.0.0	1

NET	Hops	IF
3.0.0.0	direct	e0
31.0.0.0	direct	s0
23.0.0.0	direct	s1
1.0.0.0	1	s0
12.0.0.0	1	s0
2.0.0.0	1	s1

Poison Reverse 1

- **alternative method against routing loops and slow convergence: <u>Poison Reverse</u>**
- **router sends not-reachability messages (="poison") via routing-updates in the direction from which the information about this network originally came**
- **so the convergence time is reduced to the time of failure-detection (180 s)**
  - method of Maximum Hop Count would take 16 x 30 s = 480 s !

Poison Reverse 2

Router A: routing update to net 2

Net	Hops
1	1
2	16
3	16

Router B: routing update to net 2

Net	Hops
1	16
2	16
3	1

Routing table A

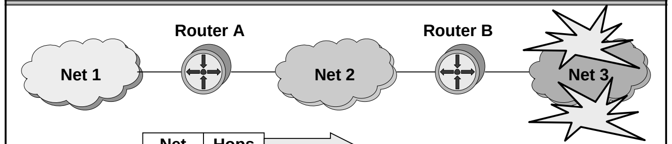
Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	1	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	0	direct

Routing update to net1 and net 2

Net 3 unreachable



Router A: routing update to net 2

Net	Hops
1	1
2	16
3	16

Routing table A

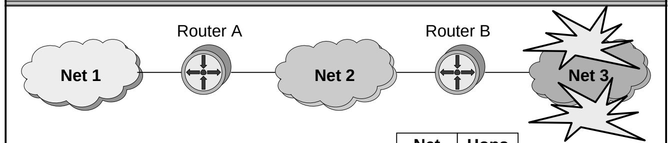
Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	1	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	16	-

Situation after sending A's routing update

Net 3 unreachable



Router B: routing update to net 2

Net	Hops
1	16
2	16
3	16

Routing table A

Net	Hops	Hop-ID
1	0	direct
2	0	direct
3	16	2.B

Routing table B

Net	Hops	Hop-ID
1	1	2.A
2	0	direct
3	16	-

Situation after sending B's routing update

Triggered Update

- **<u>Triggered Update</u> has been introduced in order to speed up the convergence time**
- **after notice of a network-failure the router generates immediately a routing-update to indicate this failure**
  - setting hop-count = 16
  - note: the router does not wait for the expiration of the 30 s
- **Triggered Update can also be used when other events occur (e.g. additional links)**
- **Triggered Update without employing additional methods (like Split Horizon) cannot avoid routing-loops**

Hold Down

- **Split Horizon is a good means to avoid temporary routing-loops and to improve the convergence time in simple network topologies**
- **complex network topologies require an additional tool to avoid temporary rooting-loops:**

  <u>Hold Down</u>
- **if a router gets information about a network failure, it ignores further information about that network from other routers for a specific duration of time**
  - typically 240 seconds

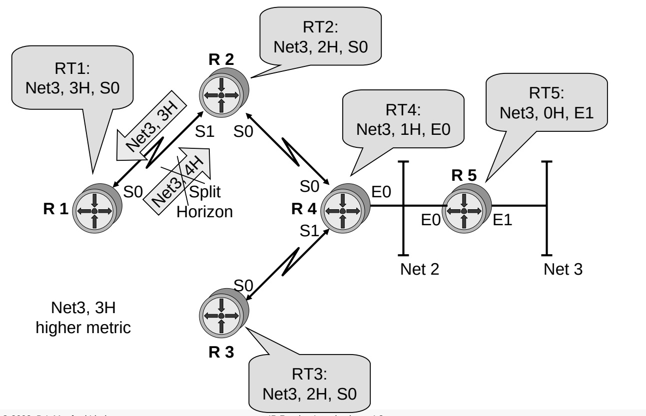
Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

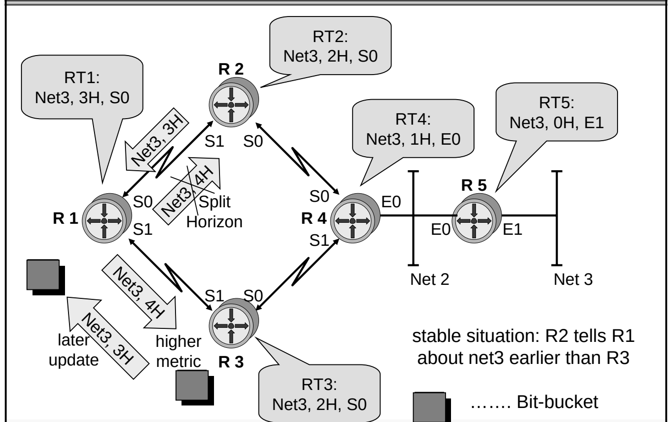
Hold Down

- **basic idea:**
  - network-failure message requires a specific amount of time to spread across the whole network (like a wave)
  - with Hold Down, all routers get the chance to receive the network-failure message
  - steady-state will be awaited to avoid inconsistent routing-tables and rooting-loops
- **disadvantages of Hold Down**
  - longer convergence time...
  - ...can be a drawback at all (in special cases)

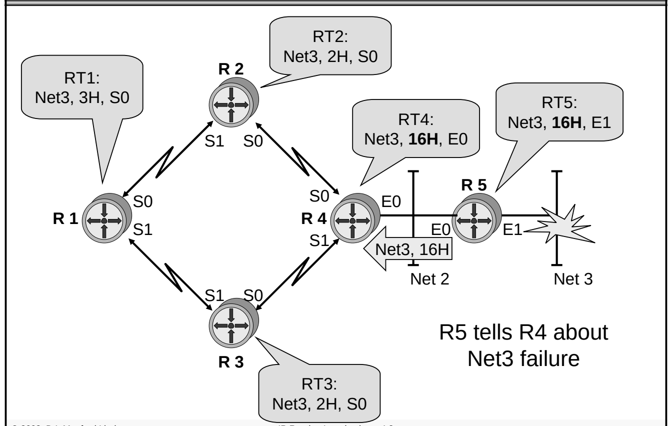
Split Horizon 1



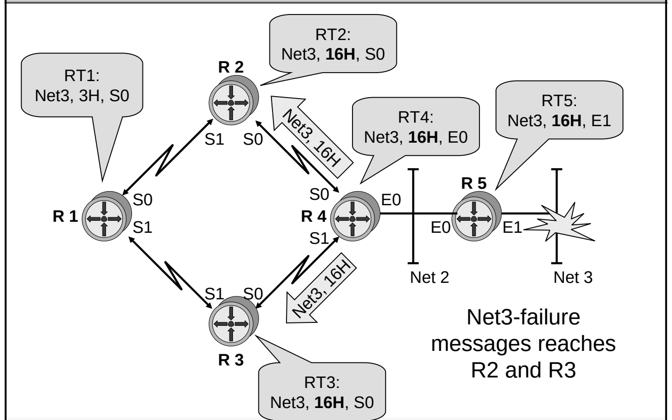
Split Horizon 2



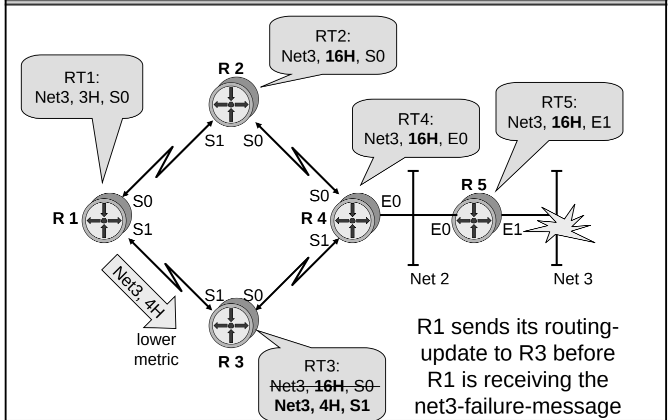
Split Horizon 3



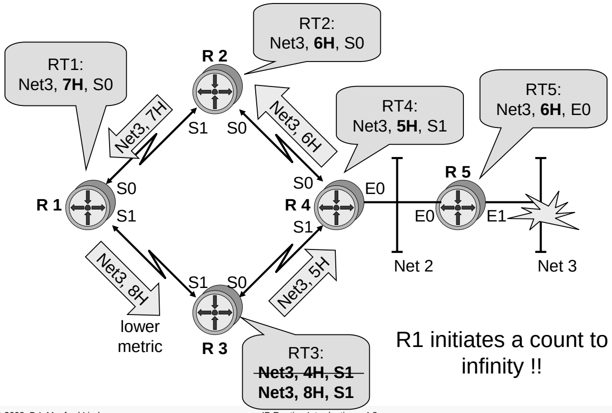
Split Horizon 4



Split Horizon 5

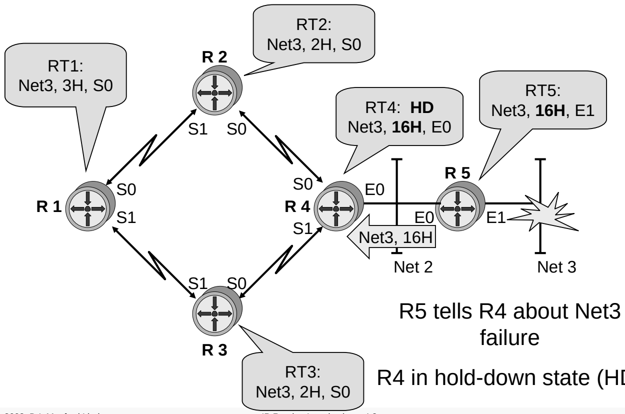


Split Horizon 6



R1 initiates a count to infinity !!

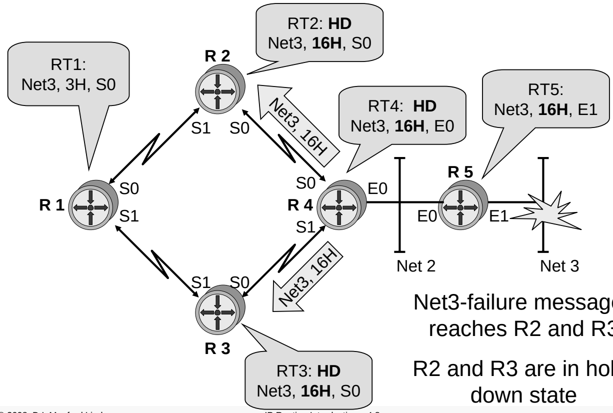
Hold Down 1



R5 tells R4 about Net3 failure

R4 in hold-down state (HD)

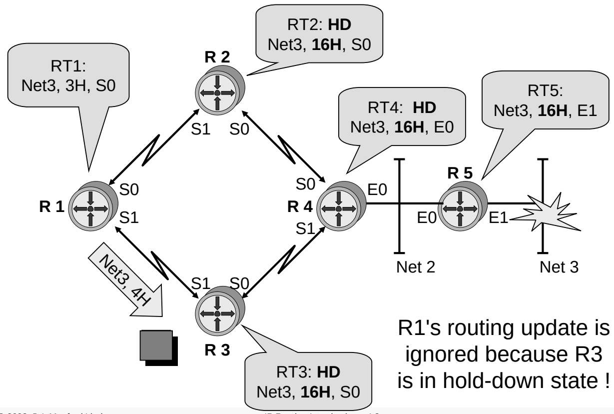
Hold Down 2



Net3-failure messages reaches R2 and R3

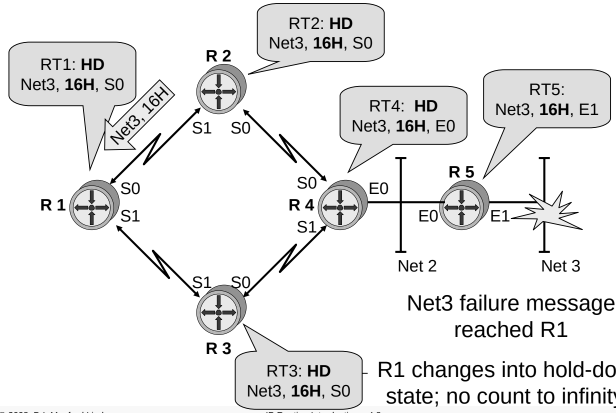
R2 and R3 are in hold-down state

Hold Down 3



R1's routing update is ignored because R3 is in hold-down state !

Hold Down 4



Net3 failure message reached R1

R1 changes into hold-down state; no count to infinity !

RIP Version 2

- **RFC 2453 specifies a new, extended RIP version:**
  - RIPv2 is RFC category "Standard"
  - RIPv1 is RFC category "Historic"
- **RIPv2 is an alternative choice to OSPF**
- **RIPv2 utilizes the unused fields of the RIPv1 message-format**
- **several new features are supported:**
  - routing domains and route tags
  - transmission of subnet-masks
  - transmission of next hop redirect information
  - authentication

RIPv2 Header plus RIPv2 Data

0	8	16	31
command	version	routing domain	
family of net 1		route tag	
IP address of net 1			
subnet mask			
next hop			
distance to net 1			
family of net 2		route tag	
IP address of net 2			
subnet mask			
next hop			
distance to net 2			
family of net 3		route tag	
. . . .			

Message Fields of RIPv2 1

- **version = 2 (RIPv2)**
- **IP address of net x, distance to net x, command fields**
  - have the same meaning as for RIPv1
- **subnet mask**
  - contains the subnet-mask to the "IP address"-field
  - now discontiguos subnetting and variable length subnet masks (VLSM) are supported
- **RIPv2 is a classless routing protocol**

Message Fields of RIPv2 2

- **routing domain**
  - indicates the routing-process for which the routing-update is destined
  - now routers can support several domains within the same subnet
- **route tag**
  - contains the autonomous system number for EGP and BGP
  - on receiving a routing-update with a routing tag unequal zero, the associated path must be distributed to other routers; so interior routers notice the existence of exterior networks (tagging exterior routes)

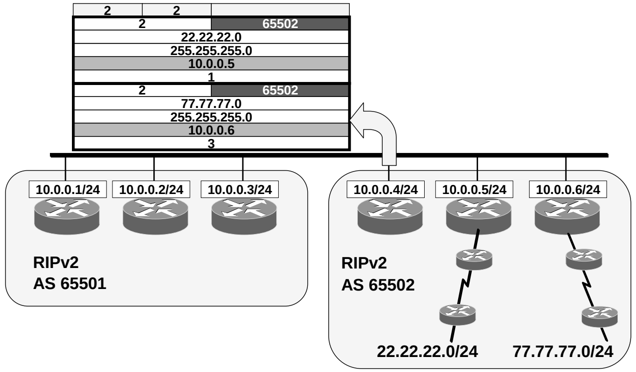
Message Fields of RIPv2 3

- **next hop**
  - datagram's for the network specified in the "IP address" -field have to be redirected to that router whose IP address is specified in the "next hop" field
  - this next-hop router must be located in the same subnet as the sender of the routing-update
  - a next hop value of 0.0.0.0 indicates, that the sender-router acts as next hop itself for the given network
  - so in cases when there are several routers in a subnet, just one router needs to send a routing-update
  - using RIPv2 this router announces which networks can be reached over other routers

Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

Next Hop and Route Tag



Authentication

- **family of net = hex FFFF**
  - this value registered in the first family of net entry announces 16 authentication octets to follow
  - currently only a single type (type = 2) defined with a simple (clear text) password protection
  - Cisco supports also type 3 based on Message Digest 5 (MD5)
  - routing updates without valid authentication are ignored by the receiving router (only trusted router are accepted)
- **family of net unequal hex FFFF**
  - has the same meaning as for RIP1

RIP2 Header plus Authentication

0	8	16 31
command	version	routing domain
hex FFFF		authentication type
authentication		
authentication		
authentication		
authentication		
family of net 1		route tag
IP address of net 1		
subnet mask		
next hop		
distance to net 1		
family of net 2		route tag
. . . .		

Treatment of Routing Updates

- **RIPv2 uses a class D multicast-address (224.0.0.9)**
  - no use of broadcast messages (like RIPv1 does)
  - only a router who is member of this group will receive and must process this routing update
- **remember:**
  - RIPv1 used broadcast addresses
  - Seen by each IP host
  - Slows down stations

Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

Agenda

- **Introduction to IP Routing**
  - Static Routing
  - Default Route
  - Dynamic Routing
- **RIP**
- **Classful versus Classless Routing**
- **Private Addresses and NAT**
- **Internet Routing**

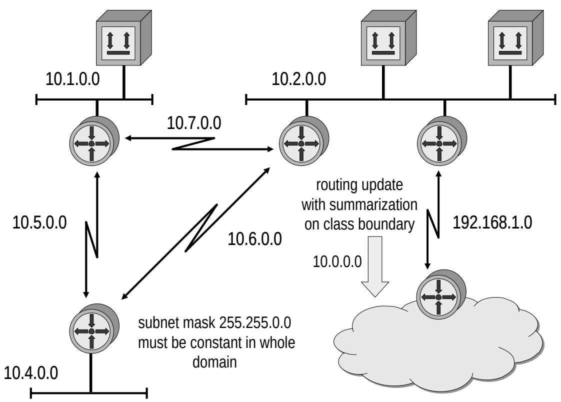
Classful Routing

- **routing protocols like RIP, IGRP cannot carry subnetmask information in routing updates**
- **this has several consequences**
  - if a given class A, B or C address is subnetted the subnetmask must be constant in the whole area
    - no variable length subnet mask (VLSM) can be used
  - if a routing update is sent to an interface with an network number different to the subnetted network
    - only the major class A, B or C network number will be announced
    - route summarization will be performed on class boundaries
    - hence a subnetted area must be contiguous
  - classful routing

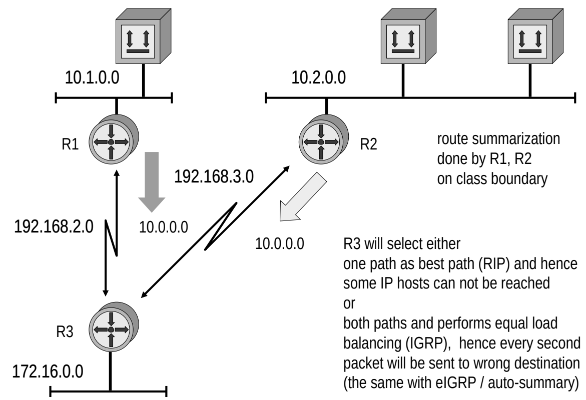
Routing Table Lookup (Classful)

- **assumption:**
  - IP datagram with a given IP address is received by a classful router
- **IP address is interpreted as class A, B or C**
  - the major net is determined
- **next the a lookup in the routing table for the major net is performed**
  - if there is no entry the IP datagram will be discarded
- **if there is a match the IP address is compared to every known subnet of this major network**
  - if there is no such subnet the IP datagram will be discarded

Classful Routing



Discontiguous Subnetting Classful



Classless Routing

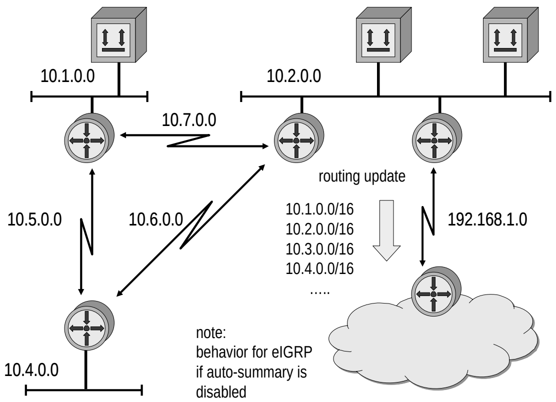
- **routing protocols like RIPv2, OSPF, eIGRP can carry subnet mask information in routing updates**
- **this has several advantages**
  - variable length subnet mask (VLSM) can be used
    - subnetting of a given address can be done according to the number of hosts required on a certain subnet
    - more efficient use of address space ⇨ sub-subnetting
  - route summarization can be performed on any address boundary and not only on class boundaries
    - a routing update contains prefix (relevant part of IP address) and length (number of ones used in subnetmask)
    - supernetting
      - actual subnetmask is smaller than natural subnetmask of given class

Routing Table Lookup (Classless)

- **assumption:**
  - IP datagram with a given IP address is received by a classful router
- **IP address is not interpreted as class A, B or C**
- **a lookup in the routing table for the best match for this IP address is performed**
  - IP prefixes of the routing table are compared with the given IP address bit by bit from left to right
  - IP datagram is passed on to the network which matches best
  - “Longest Match Routing Rule”
  - result: IP addresses with any kind of subnetting can be used independently of the underlying network topology

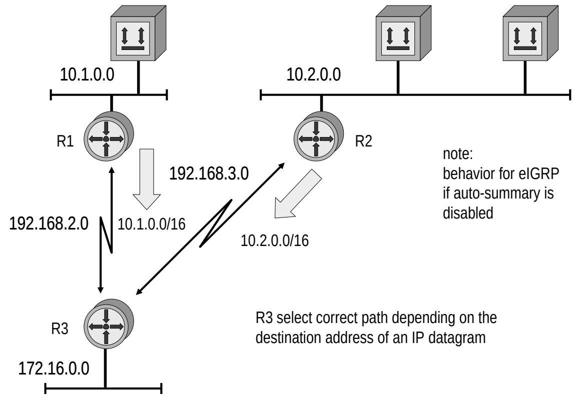
 IP Routing Introduction, v4.8 93

Classless Routing



 IP Routing Introduction, v4.8 94

Discontiguous Subnetting Classless



 IP Routing Introduction, v4.8 95

VLSM Example (1)

- **First step 6 bit subnetting of 172.16.0.0**
  - 172.16.0.0 with 255.255.252.0 (172.16.0.0 / 22)
  - subnetworks:
    - 172.16.0.0
    - 172.16.4.0
    - 172.16.8.0
    - 172.16.12.0
    - 172.16.16.0

      ……….
    - 172.16.248.0
    - 172.16.252.0
  - subnetworks capable of addressing 1022 IP systems

 IP Routing Introduction, v4.8 96

VLSM Example (2)

- **next step sub-subnetting**
  - basic subnet 172.16.4.0 255.255.252.0 (172.16.4.0 / 22)
  - sub-subnetworks with mask 255.255.255.252 (/ 30):
    - 172.16.4.0 / 30
    - 172.16.4.4 / 30
      - 172.16.4.4 net-ID
      - 172.16.4.5 first IP host of subnet 172.16.4.4
      - 172.16.4.6 last IP host of subnet 172.16.4.4
      - 172.16.4.7 directed broadcast of subnet 172.16.4.4
    - 172.16.4.8 / 30
    - 172.16.4.12 / 30
    - ……….
    - 172.16.4.252 / 30
  - sub-subnetworks capable of addressing 2 IP systems

VLSM Example (3)

- **next step sub-subnetting**
  - basic subnet 172.16.8.0 255.255.252.0 (172.16.8.0 / 22)
  - sub-subnetworks with mask 255.255.255.0 (/ 24):
    - 172.16.8.0 / 24
    - 172.16.9.0 / 24
      - 172.16.9.0 net-ID
      - 172.16.9.1 first IP host of subnet 172.16.9.0

      - 172.16.9.254 last IP host of subnet 172.16.9.0
      - 172.16.9.255 directed broadcast of subnet 172.16.9.0
    - 172.16.10.0 / 24
    - 172.16.11.0 / 24
  - sub-subnetworks capable of addressing 254 IP systems

Datenkommunikation 384.081 - SS 2008

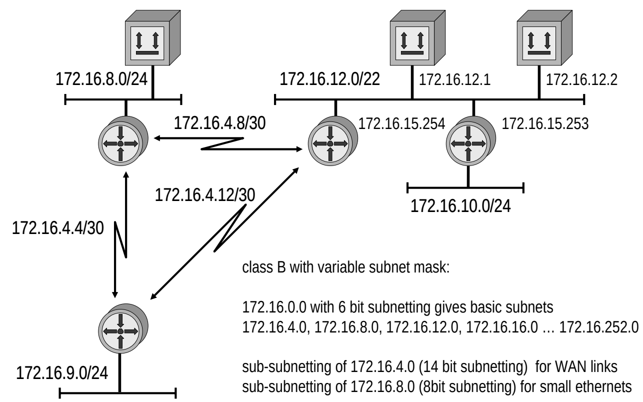
L09 - IP Routing Introduction

VLSM Example (4)

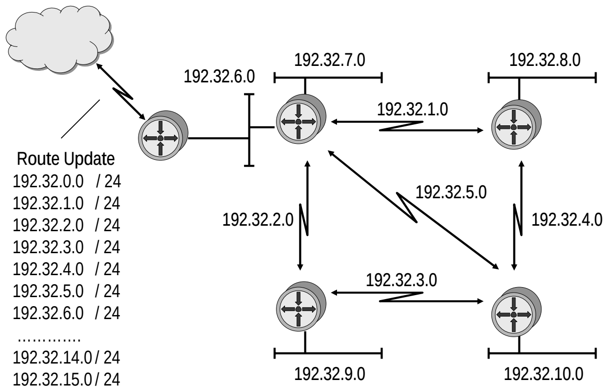
- **no sub-subnetting for basic subnet 172.16.12.0**
  - 172.16.12.0 with 255.255.252.0 (172.16.12.0 / 22)
      - 172.16.12.0 net-ID
      - 172.16.12.1 first IP host of subnet 172.16.12.0

      - 172.16.15.254 last IP host of subnet 172.16.12.0
      - 172.16.15.255 directed broadcast of subnet 172.16.12.0
  - subnetwork capable of addressing 1022 IP systems

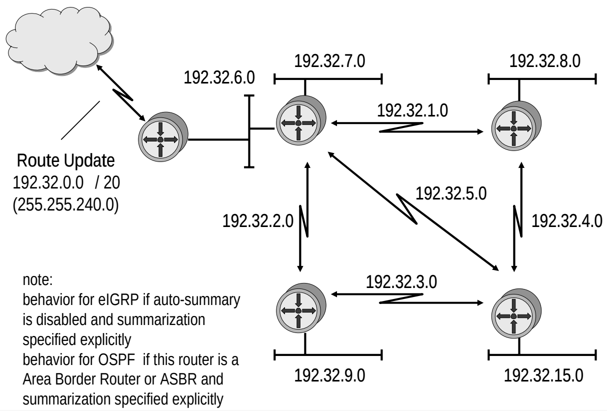
VLSM Classless



Routing Updates without Supernetting



Route Summarization with Supernetting



IP Address Space Depletion

- **the growing demand of IP addresses**
  - has put a strain on the classful model
  - class B exhaustion
  - class C are to small for most organization
  - many class C addresses given to a certain organization leads to explosion of routing table entries in the Internet core routers
- **measures to handle these problems**
  - CIDR (Classless Interdomain Routing)
  - private IP addresses and network address translation (NAT)
  - IPv6

CIDR

- **Classless Interdomain Routing (CIDR)**
  - address assignment and aggregation (route summarization) strategy
- **basic ideas**
  - classless routing (prefix, length)
  - supernetting
  - coordinated address allocation
    - until 1992 IP addresses had no relation at all to the networks topology
- **in order to implement CIDR**
  - classless routing protocols between routing domains must be used
    - BGP-4 as interdomain routing protocol
  - classless routing within an routing domain
    - RIPv2, OSPF, eIGRP

© 2008, D.I. Manfred Lindner

CIDR

- **CIDR address allocation**
  - addressing plan for class C addresses by continents
    - 192.0.0.0 - 193.255.255.255 … Multiregional
    - 194.0.0.0 - 195.255.255.255 … Europe
    - 198.0.0.0 - 199.255.255.255 … North America
    - 200.0.0.0 - 201.255.255.255 … Central/South America
  - provider addressing strategy
    - Internet Service Providers (ISP) are given contiguous blocks of class C addresses which in turn are granted to their customers
    - consequence: change of provider means renumbering
  - class C network numbers are allocated in such a way that route summarization (or sometimes called route aggregation) into supernets is possible

Agenda

- **Introduction to IP Routing**
  - Static Routing
  - Default Route
  - Dynamic Routing
- **RIP**
- **Classful versus Classless Routing**
- **<u>Private Addresses and NAT</u>**
- **Internet Routing**

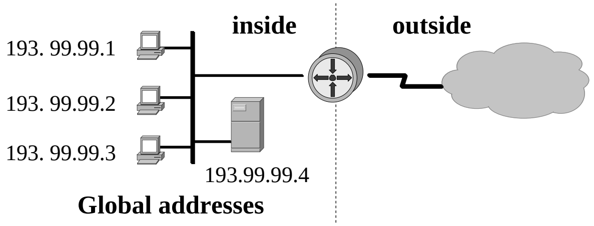
Private Address Range - RFC 1918

- **Three blocks of address ranges are reserved for addressing of private networks**
  - 10.0.0.0 - 10.255.255.255 (10/8 prefix)
  - 172.16.0.0 - 172.31.255.255 (172.16/12 prefix)
  - 192.168.0.0 - 192.168.255.255 (192.168/16 prefix)
  - Note:
    - In pre-CIDR notation the first block is nothing but a single class A network number, while the second block is a set of 16 contiguous class B network numbers, and third block is a set of 256 contiguous class C network numbers.
- **Translation between private addresses and globally unique addresses -> NAT**

Reasons for Network Address Translation

- **Mitigate Internet address depletion**
  - NAT was originally developed as an interim solution to combat IPv4 address depletion by allowing globally registered IP addresses to be reused or shared by several hosts (RFC 1631)
- **Save global addresses (and money)**
  - NAT is most often used to map IPs from the nonroutable private address spaces defined by RFC 1918
    - 10.0.0.0/8, 172.16.0.0/16, 192.168.0.0/16
- **Conserve internal address plan**
- **TCP load sharing**
- **Hide internal topology**

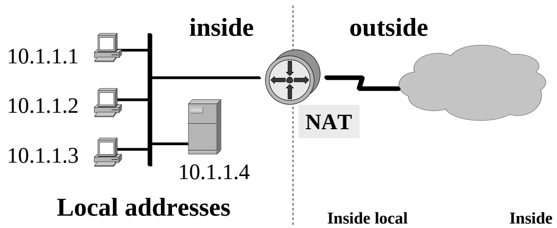
Terms (1)



(NAT not necessary in this case)

 IP Routing Introduction, v4.8 109

Terms (2)

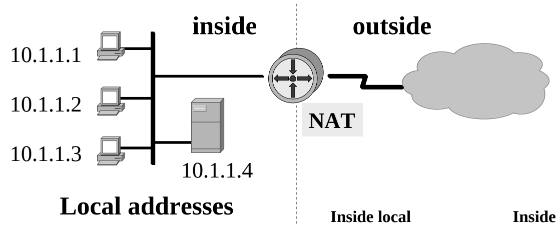


Static one-to-one mapping is maintained by router-internal NAT–Table

Inside local IP address		Inside global IP address
10.1.1.1	⟷	193.99.99.1
10.1.1.2	⟷	193.99.99.2
10.1.1.3	⟷	193.99.99.3
10.1.1.4	⟷	193.99.99.4

 IP Routing Introduction, v4.8 110

Terms (3)

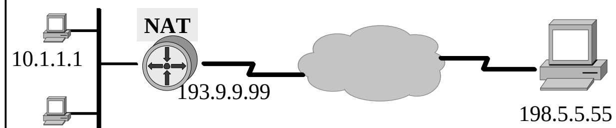


Dynamic mapping via pool is maintained by router-internal NAT–Table

Inside local IP address		Inside global IP address
10.1.1.1	⟷	193.99.99.5
10.1.1.2	⟷	193.99.99.6
10.1.1.3		
10.1.1.4		

 IP Routing Introduction, v4.8 111

Basic Principle (1a)

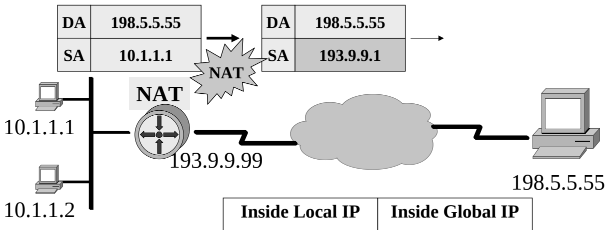


Simple NAT Table

Inside Local IP	Inside Global IP
10.1.1.1	193.9.9.1
10.1.1.2	193.9.9.2
....	....

 IP Routing Introduction, v4.8 112

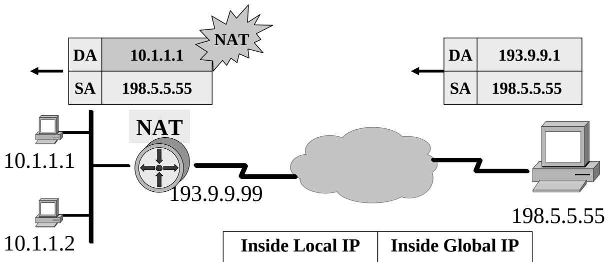
Basic Principle (1b)



Simple NAT Table

Inside Local IP	Inside Global IP
10.1.1.1	193.9.9.1
10.1.1.2	193.9.9.2
....	....

Basic Principle (1c)



Simple NAT Table

Inside Local IP	Inside Global IP
10.1.1.1	193.9.9.1
10.1.1.2	193.9.9.2
....	....

Agenda

- **Introduction to IP Routing**
  - Static Routing
  - Default Route
  - Dynamic Routing
- **RIP**
- **Classful versus Classless Routing**
- **Private Addresses and NAT**
- **Internet Routing**

Routing in Small Networks

- **in small networks**
  - distance vector or link state protocols like RIP or OSPF can be used for dynamic routing
  - it is possible that every router of the network knows about all destinations
    - all destination networks will appear in the routing tables
  - routing decisions are based on technical parameters
    - e.g. hop count, link bandwidth, link delay, interface costs
  - it is sufficient that routing relies only on technical parameters
    - small networks will be administered by a single authority
    - non-technical parameter like traffic contracts have no importance

Routing in Large Networks

- **with increasing network size limitations of these protocols can be recognized**
  - some limitations for example
    - maximum hop count (RIP)
    - time to transmit routing tables (RIP) on low speed links
    - CPU time for SPF calculation (OSPF)
    - memory used for storing routing table (RIP, OSPF)
    - memory used for storing topology database (OSPF)
    - two level hierarchy centered around a core network (OSPF)
    - route fluctuation caused by link instabilities (OSPF)
    - routing based on non-technical criteria like financial contracts or legal rules is not possible

Routing in the Internet

- **limitations prevent using routing protocols like RIP or OSPF for routing in the Internet**
  - note: routing tables of Internet-core routers have about 220.000 net-ID entries
- **routing in the Internet**
  - is based on non-technical criteria like financial contracts or legal rules
  - policy routing
    - e.g. Acceptable Use Policy (AUP) in parts of the Internet
    - e.g. contracts between Internet Service Providers (ISP)

Datenkommunikation 384.081 - SS 2008

L09 - IP Routing Introduction

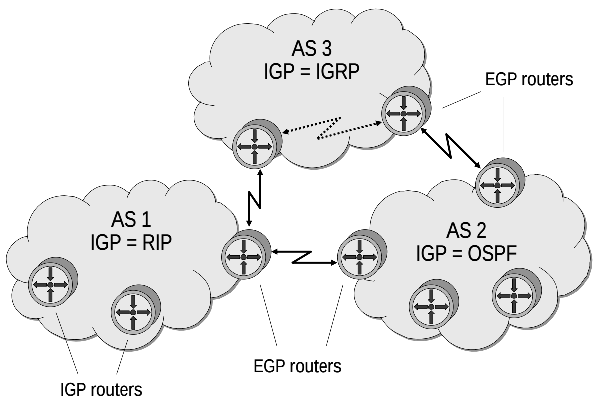
Routing Hierarchy, Autonomous Systems

- **routing hierarchy is necessary for large networks**
  - to control expansion of routing tables
  - to provide a more structured view of the Internet
- **routing hierarchy used in the Internet**
  - based on concept of autonomous system (AS)
- **AS concept allows**
  - segregation of routing domains into separate administrations
  - note:
    - routing domain is a set of networks and routers having a single routing policy running under a single administration

IGP, EGP

- **within an AS one or more IGP protocols provide interior routing**
  - IGP - Interior Gateway Protocol
  - IGP examples
    - RIP, RIPv2, OSPF, IGRP, eIGRP, Integrated IS-IS
  - IGP router responsible for routing to internal destinations
- **routing information between ASs is exchanged via EGP protocols**
  - EGP - Exterior Gateway Protocols)
  - EGP examples
    - EGP-2, BGP-3, BGP-4
    - EGP router knows how to reach destination networks of other ASs

AS, IGP, EGP



AS Numbers

- **hierarchy based on ASs allows forming of a large network**
  - by dividing it into smaller and more manageable units
  - every unit may have its own set of rules and policies
- **AS are identified by a unique number**
  - can be obtained like IP address from an Internet Registry
    - e.g. RIPE NCC (reséaux IP Européens Network Coordination Center)

BGP-4

- **Border Gateway Protocol (BGP)**
  - is the Exterior Gateway Protocol used in the Internet nowadays
  - was developed to overcome limitations of EGP-2
  - RFC 1267 (BGP-3) older version
    - classful routing only
  - RFC 1771 (BGP-4) current version, DS
    - classless routing
  - is based on relationship between neighboring BGP-routers
    - peer to peer
    - called BGP session or BGP connection

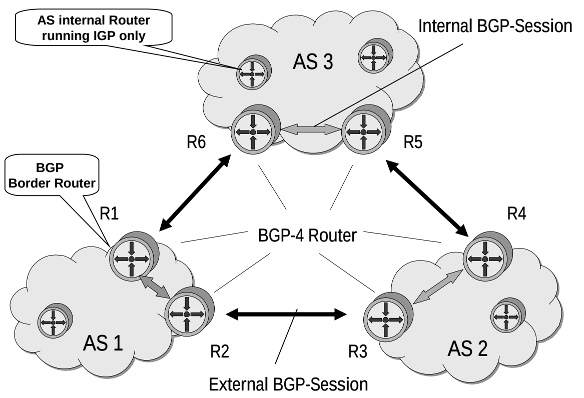
BGP-4 Concepts 1

- **Border Gateway Protocol (cont.)**
  - primary function
    - exchange of network reachability information with other autonomous systems via external BGP sessions
    - but also within an autonomous system between BGP border routers via internal BGP sessions
  - BGP session runs on top of TCP
    - reliable transport connection
    - well known port 179
    - TCP takes care of fragmentation, sequencing, acknowledgement and retransmission
    - hence these procedures need not be done by the BGP protocol itself

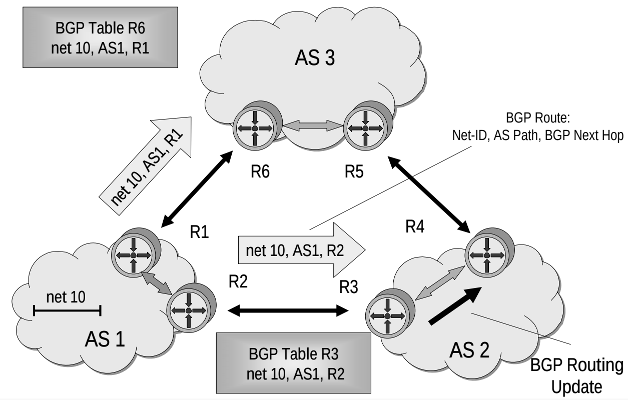
BGP-4 Concepts 2

- **basic ideas**
  - reachability information exchanged between BGP routers carries a <u>sequence of AS</u> numbers
    - indicates the path of ASs a route has traversed
  - path vector protocol
  - this allows BGP to construct a graph of autonomous systems
    - loop prevention
    - no restriction on the underlying topology
  - the best path
    - minimum number of AS hops
  - incremental update
    - after first full exchange of reachability information between BGP routers only changes are reported

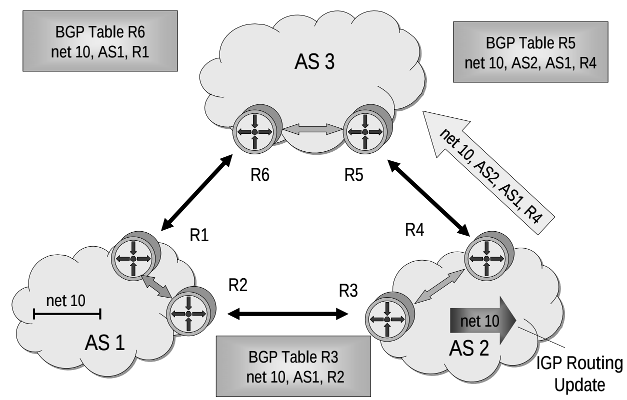
BGP-4 Basic Example (1)



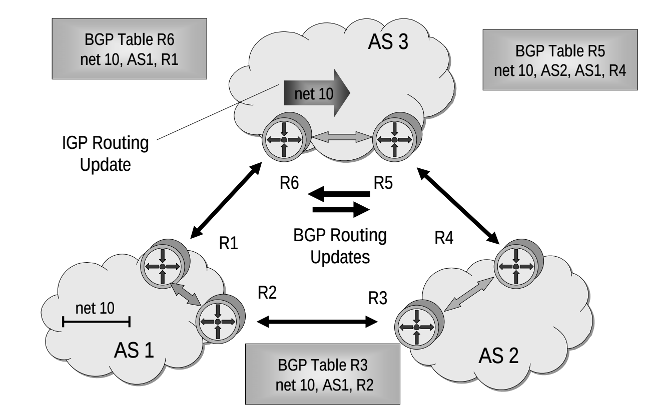
Basic Example (2)



Basic Example (3)



Basic Example (4)



BGP-4 Concepts 3

- **basic ideas (cont.)**
  - description of reachability information by BGP attributes
    - for BGP routing
    - for establishing of routing policy between ASs
  - BGP-4 advertises so called BGP routes
    - a BGP route is unit of information that pairs a destination with the path attributes to that destination
    - AS Path is one among many BGP attributes
  - IP prefix and mask notation
    - supports VLSM
    - supports aggregation (CIDR) and supernetting
  - routes can be filtered using attributes, attributes can be manipulated

    --> routing policy can be established