Torsten Braun

IBM European Networking Center

Heidelberg

SPIE BNE SC 82

IP next generation

October 1996 - 1

The Next Generation IP (IPv6) and Resource Reservation Protocols

Torsten Braun

Short Course 82

SPIE Broadband Network Engineering

November 22, 1996

Hynes Convention Center, Boston, MA, USA

Torsten Braun

IBM European Networking Center

Heidelberg

SPIE BNE SC 82

IP next generation

October 1996 - 2

Instructor

Dr. Torsten Braun

IBM European Networking Center

Vangerowstr. 18, D-69115 Heidelberg, Germany
phone: +49 6221 59 4352
fax: +49 6221 59 3300
e-mail: t.braun@ieee.org or braun@heidelbg.ibm.com

- Study of Computer Science at University of Karlsruhe (Germany), Ph.D. in 1993
- 1994-1995 visiting scientist at the Institut National de Recherche en Informatique et en Automatique (INRIA) at Sophia-Antipolis (France), Rodeo project
- since 1995 guest scientist at IBM European Networking Center (ENC), Broadband Multimedia Communication department

Outline

- Internet Protocol version 6 (IPv6)
  - Addressing and header Formats
  - Routing
  - Segmentation and Reassembly
  - Quality-of-Service support
  - ICMP
  - Security
  - Transition mechanisms
  - IPv6 over ATM
- Resource Reservation Protocol (RSVP)
  - RSVP definitions
  - Integrated services
  - RSVP protocol mechanisms
  - RSVP over ATM and shared media

Addressing Problems

- Class B addresses are nearly exhausted.
  - CIDR (classless interdomain routing) provides a short-term solution
- Routing tables are growing extremely fast.
  - More hierarchical addressing levels are required.
- IP unicast address space will run out between the years 2000 and 2010 for several reasons:
  - Internet euphoria (on-line services, corporate networks over the Internet)
  - new low-end Internet devices (in cars, households etc.)
  - 10 billion people in 2020 and 100 IP addresses per person are not unrealistic.

Solution:

- Larger address space necessary

  -> new IP protocol required to substitute IPv4

Torsten Braun
IBM European Networking Center
Heidelberg

CIDR

- Class C address: 256 hosts, class B address: 65536 hosts
- Observation: Many organizations have more than 256 hosts but only a few have more than a few thousand hosts.

-> A few class C addresses are sufficient for many organizations, e.g. 1024 hosts can be addressed with 4 class C addresses.

- Non-contiguous class C addresses for a single organization lead to routing table explosion, because each class C address requires a routing table entry within a router.
- Solution:
  - divide available address prefixes into aggregates based on regions (e.g., 194.0.0.0 - 195.255.255.255 for Europe)
  - take contiguous class C addresses for an organization
  - add prefix length to routing entry

Torsten Braun
IBM European Networking Center
Heidelberg

CIDR Example

- Example: A company needs addresses for 800 IP hosts.
  -> 4 contiguous class C addresses required
- CIDR address format: prefix/length

110	0 0011 0000 0101 0000 00	00	host	195.5.0.0
110	0 0011 0000 0101 0000 00	01	host	195.5.1.0
110	0 0011 0000 0101 0000 00	10	host	195.5.2.0
110	0 0011 0000 0101 0000 00	11	host	195.5.3.0

CIDR:

110	0 0011 0000 0101 0000 00	195.5.0.0/22

- Routing protocol modifications required to allow CIDR format exchange (e.g., BGP-4 supports prefix lengths)

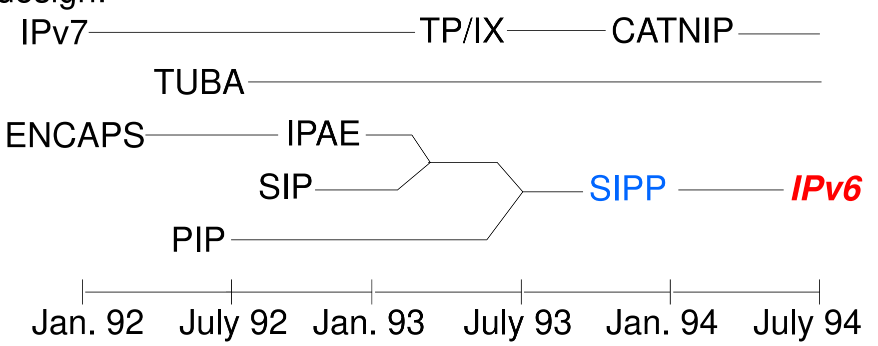
Torsten Braun

IBM European Networking Center

Heidelberg

IPng Development

- 1992: IETF issued a call for proposals for a next-generation IP (IPng) to replace the current IP version 4
- Several proposals have been made and evolved to a final design:



- 1994: SIPP (Simple Internet Protocol Plus) was suggested as the new basis of IPng
- 1995: Internet Draft “Internet Protocol, Version 6 (IPv6)” becomes “proposed standard” (9/95), publication as RFC1883 (12/95)

Torsten Braun

IBM European Networking Center

Heidelberg

IPng/IPv6 Characteristics

- Extended addressing capabilities
  - 128 bit addresses
    ($6.65 \cdot 10^{23}$ addresses per m^2 of earth surface)
  - address hierarchy levels
    (registry, provider, subscriber, subnet, interface)
  - anycast addresses to address one node of a group of nodes (the “nearest” node in terms of the routing metric)
  - address autoconfiguration
  - mobile system support
- New IP header format
  - simplified, minimal IP header
  - improved support for new options and extensions
  - segmentation and reassembly in end systems only

Torsten Braun

IBM European
Networking Center

Heidelberg

More IPng/IPv6 Characterisitcs

- Quality-of-Service support
  - flow labels
  - priorities
- Multicast integration
  - pre-defined groups for control functions
  - IGMP (Internet Group Management Protocol) integrated into ICMP (Internet Control Message Protocol)
  - special address format for multicast addresses
  - all routers and end systems support multicast communication, broadcast mechanisms may be avoided
- Security functions
  - authentication
  - encryption

Torsten Braun

IBM European
Networking Center

Heidelberg

Address Types

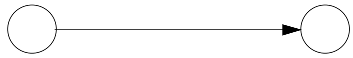
- Unicast addresses (1:1)
  - provider-based
    (variable length fields, more hierarchy levels are possible)

3 bit	n bit	m bit	o bit	p bit	125-m-n-o-p bit
010	registry ID	provider ID	subscriber ID	subnet ID	interface ID

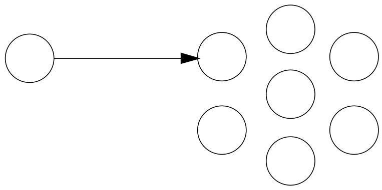
  - geographically-based
  - link local (for configuration or IP networks without router)
  - site-local (for IP networks not connected to the Internet, only the address prefix must be replaced when the site is connected to the Internet, the subnet and interface ID remain unchanged)
  - IPv4-, IPX-, NSAP- compatible addresses
- Anycast addresses (1: nearest node of a set of addresses, currently only used to address routers)
- Multicast addresses (1:n)

Differences of Address Types

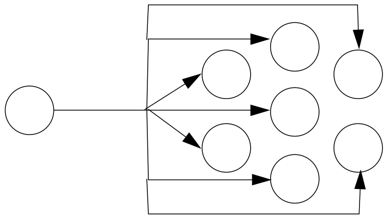
- unicast (1:1)



- anycast (1:nearest)



- multicast (1:n)



Unicast Addresses

address type	prefix	fraction of address space
reserved (e.g., for loop back addresses, IPv4 compatible addresses)	00000000	1/256
NSAP addresses	0000001	1/128
IPX addresses	0000010	1/128
provider-based unicast addresses	010	1/8
geographic-based unicast addresses	100	1/8
link-local unicast addresses	11111110 10	1/1024
site-local unicast addresses	11111110 11	1/1024

- link local addresses: prefix + interface ID (MAC address or interface number)
- IPv4-compatible address: prefix (96 0-bits) + IPv4 address (32 bits)
- site-local address: prefix + subnet ID + interface ID

Torsten Braun

IBM European Networking Center

Heidelberg

SPIE BNE SC 82

IP next generation

October 1996 - 13

Multicast Addresses

8bit	4bit	4 bit	112 bit
11111111	flags	scope	multicast group ID

- flags:
  transient/permanent address
- scope:
  node-local, link-local, site-local, organization-local
- reserved multicast addresses:
  - all nodes (node-, link-local)
  - all routers (node-, link-local)
  - DHCP server / relay agent (link-local)
  - solicited node (link-local)
    for each unicast and anycast address of a node there exists a multicast group address with a pre-defined 96 bit prefix and the least significant 32 bits of the unicast/ anycast address

Torsten Braun

IBM European Networking Center

Heidelberg

SPIE BNE SC 82

IP next generation

October 1996 - 14

Text Representation of Addresses

- 128 bit address consists of 8 16-bit integers,
  each integer is represented by four hexadecimal digits
  - FEDC:BA98:7654:3210:FEDC:BA98:7654:3210
- abbreviations for addresses with many 0's
  - 1080:0000:0000:0000:0008:0800:200C:4711
  - reduced form: 1080:0:0:0:8:0800:200C:4711
  - short form: 1080::8:0800:200C:4711
- double colon can be used once in an address representation
- IPv4-compatible addresses:
  - reduced form: 0:0:0:0:0:A00:1
  - short form: ::A00:1
  - decimal form: ::10.0.0.1
- Prefixes:
  - FEDC:BA98:7600::/40 (40 bit long prefix)

Torsten Braun

IBM European Networking Center

Heidelberg

IPv4 Header

version	hdr len	prec	type of service	total length	
identification				flags	fragment offset
time to live		protocol		header checksum	
source address					
destination address					

fields also available in IPv6
fields removed in IPv6
fields moved to IPv6 extension header
fields with new name in IPv6 but with similar meaning

- 20 bytes, 13 fields
- modifications
  - precedence -> priority
  - total length -> payload length
  - time to live -> hop limit
  - protocol -> next header

Torsten Braun

IBM European Networking Center

Heidelberg

SPIE BNE SC 82

IP next generation

IPv6 Header

version	priority	flow label	
payload length		next header	hop limit
source address			
destination address			

- 40 bytes, 8 fields
- less fields, no segmentation and reassembly
  -> simpler and faster router implementations

Torsten Braun
IBM European Networking Center
Heidelberg

Extension Headers

Daisy chain of extension headers

- small minimal header
- flexible headers dependent on application and network requirements
- simple introduction of new extensions and options without IPv4's 40-byte limit on options

IPv6 header next header = TCP	TCP header + data

IPv6 header next header = routing	routing header next header = TCP	TCP header + data

IPv6 header next header = routing	routing header next header = fragment	fragment header next header = TCP	TCP Header + data (fragment)

Torsten Braun
IBM European Networking Center
Heidelberg

Sequence of Headers

- IPv6 header (mandatory)
- Hop-by-hop options
- Destination options header
(processed by the first destination that appears in the IPv6 destination address field and destination listed in the routing header)
- Routing header
- Fragment header
- Authentication header
- Encapsulation security payload header
- Destination options header
(processed in the final destination of the packet)

Extension Headers following hop-by-hop options are only processed by nodes identified in IPv6 destination address field --> lower option processing overhead

Torsten Braun

IBM European Networking Center

Heidelberg

IPv6 Packet with Extension Headers

Header	Length	
IPv6 header	40	next header field
Hop-by-hop options	variable	
Destination options header	variable	
Routing header	variable	
Fragment header	8	
Authentication header	variable	
Encapsulation security payload header	variable	
Destination options header	variable	
TCP header	20 (optional part variable)	
Application data	variable	

Torsten Braun

IBM European Networking Center

Heidelberg

Routing

- based on “longest prefix match” routing like IPv4 CIDR
- unicast routing protocols (OSPF, RIP) and multicast routing protocols (DVMRP, MOSPF, PIM) must be modified to handle larger addresses
- Routing headers can be used (with anycast addresses) to route packets through particular (desired) regions
  - provider selection
  - policy
  - performance

Torsten Braun

IBM European Networking Center

Heidelberg

Routing Header

next header	routing type=0	hdr ext len	segments left
reserved	loose/strict bit mask		
address [0]			
address [1]			
...			

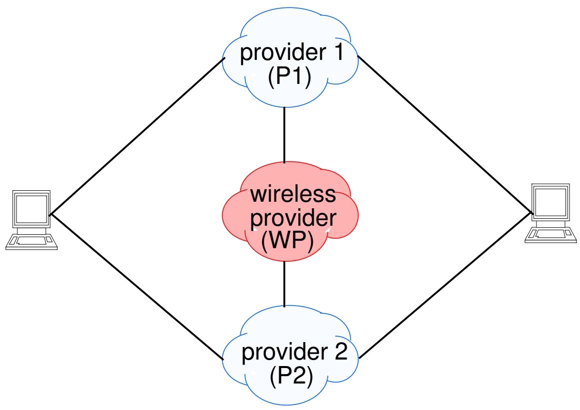
- List of unicast/anycast addresses to be visited (address[i]) (similar to source routing option in IPv4)
- segments left: number of nodes still to be visited
- loose/strict bit mask (1 bit per node):
  1 = strict routing, next address must be a neighbor

Torsten Braun

IBM European Networking Center

Heidelberg

Routing Header Example



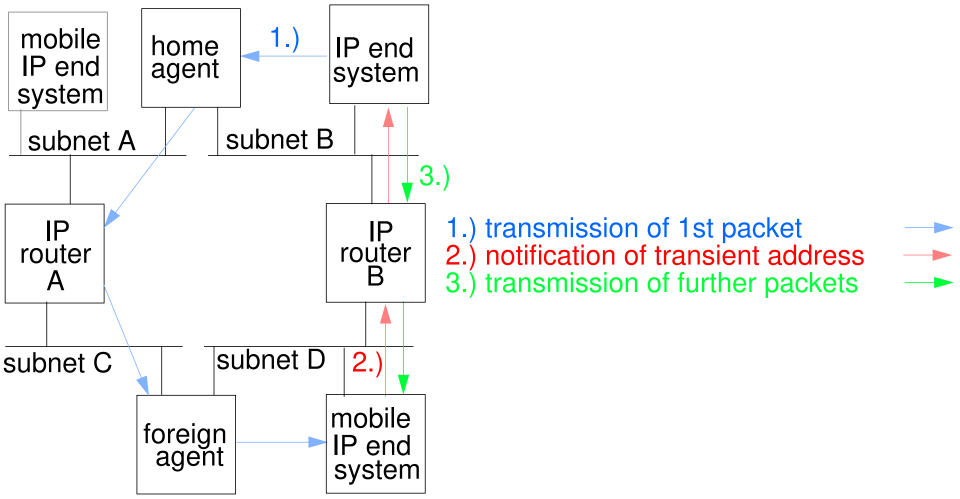
- Case 1: no routing header (H1, H2)
  -> either provider can be used
- Case 2: H1enforces traffic to H2 via provider 1 (H1, P1, H2) with strict routing
- Case 3: H1 becomes mobile and moves to the wireless provider -> (H1, WP, P1, H2)

Torsten Braun

IBM European Networking Center

Heidelberg

Mobility Support



- Mobile host gets temporary (care-of) address from foreign agent and notifies home agent
- Home agent forwards packets sent to permanent (home) address
- Notification of temporary address by options (2.)
- Use of transient and permanent address within routing header (3.) => no IP-in-IP encapsulation like in IPv4

Torsten Braun

IBM European Networking Center

Heidelberg

Hop-by-Hop Options Header

next header	Hdr Ext Len	
Options		

- list of variable options
- option

option type	option data len	option data

- option type

AIU	C	Option ID

  - AIU: action if unrecognized (00: skip, 01: discard packet, do not send ICMP report, 10: discard packet, send ICMP report even for multicast destination, 11: discard packet, send ICMP report only for unicast destination)
  - C (change en route): set if options may be changed by intermediate systems

Jumbo Payload Option

- Packet length is limited to 65535 bytes by 2 byte long “total length” field in IPv6 header.
- Jumbo payload length contains 4 byte long packet length field.
- Jumbo payload option is example for hop-by-hop option.

next header	header ext. len.	option type=194	opt data len=4
jumbo payload length			

- Fragment header is not allowed with jumbo datagrams.
- TCP and UDP can not currently use “jumbograms”

Torsten Braun
IBM European Networking Center
Heidelberg

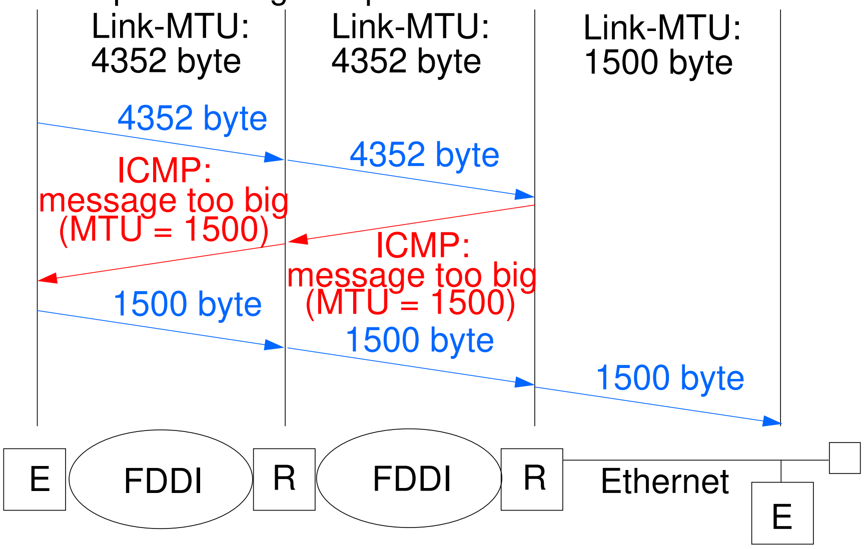
Segmentation and Reassembly

- link MTU = maximum IP packet size that can be transmitted over the link (MTU: maximum transmission unit)
- path MTU = minimum of all link MTUs in a path between source and destination
- no segmentation and reassembly (SAR) by IPv6 routers
- link MTU > 576 octets (IPv4: 68 bytes)
- links with smaller MTUs require SAR functions on link-level (e.g., ATM: AAL)
- Absence of SAR functions in routers requires SAR functions in end systems (-> fragment header):
    - sending of minimal link MTUs
    - path MTU discovery

Torsten Braun
IBM European Networking Center
Heidelberg

Path MTU Discovery

1. Send packets with link MTU size
2. “packet too big”-ICMP messages for packets exceeding path MTU size
3. Reduce packet length to path MTU size



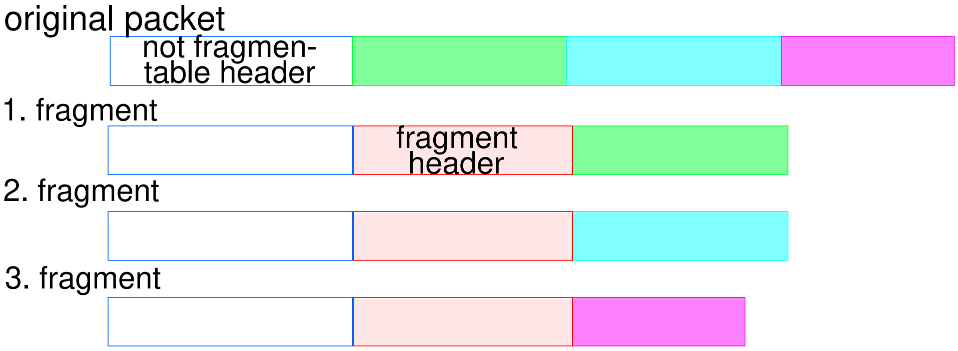
Torsten Braun
IBM European Networking Center
Heidelberg

Fragment Header

- Fragment header allows transmission of packets larger than path MTUs.

next header	reserved	fragment offset	00M
identification			

- fragment offset: relative position of fragment data within the original packet
- M flag: more fragments (M=1)
- identification: unique sender identification of a packet



Torsten Braun
IBM European Networking Center
Heidelberg

Priorities

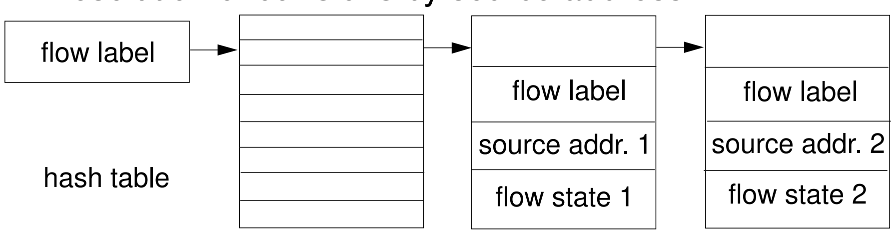
- Priority values allow source to identify packet priority relative to other packets from the same source.
- 2 independent priority classes:
  - congestion-controlled traffic (e.g., TCP traffic): values 0-7
  - non-congestion-controlled traffic (e.g., real-time data): values 8-15 to indicate discard priority (8: most willing to discard)

value	traffic type
0	uncharacterized traffic
1	filler traffic (e.g., news)
2	unattended data transfer (e.g., electronic mail)
3	reserved
4	attended bulk transfer (e.g., FTP, HTTP, NFS)
5	reserved
6	interactive traffic (e.g., telnet, X)
7	Internet control traffic (e.g., routing protocols, ICMP, SNMP)

Torsten Braun
IBM European Networking Center
Heidelberg

IPv6 Flows

- Sequence of packets for which the source desires non-default or real-time service
- Identification of data flows between end systems
- Identification by source address and random flow label value
  - hash table based on any flow label bits --> fast flow state look-up
  - resolution of collisions by source address



- Examples for flow state
  - next hop
  - information about quality-of-service to provide for this flow

Torsten Braun

IBM European Networking Center

Heidelberg

Flow Processing

- Special packet processing
  - quality-of-service support
  - processing information exchanged by
    - packet options (e.g., hop-by-hop options)
    - resource reservation protocols (e.g., RSVP)
- Simplified packet processing
  - Packet processing results are calculated for the first packet of a flow and stored in the hash table
  - Fast table look-up is performed for the following packets instead of recalculating the results
  - All packets of a flow must use identical source/destination addresses, priority, hop-by-hop options, and routing header
  - Flow labels are valid for 6 seconds, packet processing results (flow state) have to be recalculated

Torsten Braun

IBM European Networking Center

Heidelberg

ICMP

- Message types
  - error messages
  - informational messages
    - echo request and reply
    - group management
      (group membership query, report, termination)
  - neighbor discovery
    - router solicitation/advertisement
    - neighbor solicitation/advertisement
    - redirect
- Format

type	code	checksum
message body		

Torsten Braun

IBM European Networking Center

Heidelberg

ICMP Error Messages

- Destination unreachable
  - no route to destination
  - communication with destination administratively prohibited
  - address unreachable
  - port unreachable
- packet too big
- time exceeded
  - hop limit exceeded in transit
  - fragment reassembly time exceeded
- parameter problem
  - erroneous field
  - unrecognized next header type
  - unrecognized IPv6 option

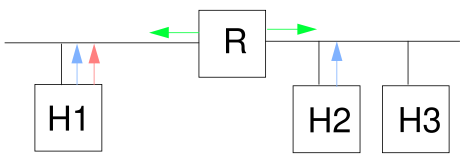
Torsten Braun

IBM European Networking Center

Heidelberg

Group Management

- Routers must know whether there are group members on a link
- Group management messages
  - Membership query messages
    are sent periodically (e.g., 1/min) from router to “all hosts” group (general) or to particular groups (group-specific)
  - Membership report messages
    are sent by a host to indicate group membership
  - Membership termination messages are sent by hosts leaving a group if they responded to the last membership query message -> router sends group-specific membership query messages to update routing tables



H3 detects membership report from H2 and suppresses own response

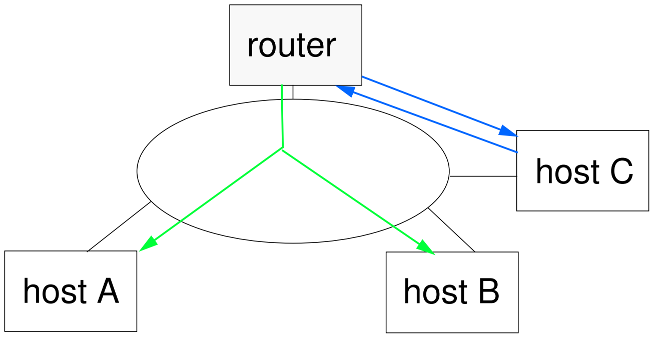
- Primitives: JoinHostGroup (address, interface), LeaveHostGroup (address, interface)

Neighbor Discovery Functions I

- **Router Solicitation**
  - request to all routers multicast address for router advertisement messages
- **Router Advertisement**
  - are sent periodically to all nodes multicast address or to the source of a router solicitation message
  - contains prefix value and length, router address and lifetime, link, Internet parameters, and stateful/stateless address autoconfiguration flag
- Functions
  - Router discovery
  - Prefix discovery
  - Next-hop determination
  - Parameter discovery (link and Internet parameters such as MTU size, maximum hop limit)
  - Stateless address autoconfiguration (IP address = combination of prefix and link-local unicast address)

Plug-and-Play

- IPv4: manual configuration of end system with several IP parameters (IP address, subnet mask, default gateway, ...)
- IPv6: Router advertises parameters periodically (router advertisement to all nodes multicast group) or host solicits advertisements from router (router solicitation to all routers multicast group).
- Hosts use link-local address (pre-defined link-local address prefix + interface ID (MAC address)) during initialization
- IP address = advertised prefix + interface ID



Torsten Braun

IBM European Networking Center

Heidelberg

Neighbor Discovery Functions II

- **Neighbor Solicitation**
  sent to solicited-nodes multicast address or to unicast address
- **Neighbor Advertisement**
  response to neighbor solicitation message or propagation of new information (e.g., modification of link address) to all-nodes multicast address
- Functions
  - duplicate address detection
    (by analysis of responses to neighbor solicitation messages to solicited nodes multicast address)
  - neighbor unreachability detection (neighbor solicitation message to unicast address)
  - link-layer address resolution (cf. IPv4 ARP, by sending a neighbor solicitation message to solicited nodes multicast address)
- **Redirects**
  to indicate a better next hop for a specific destination

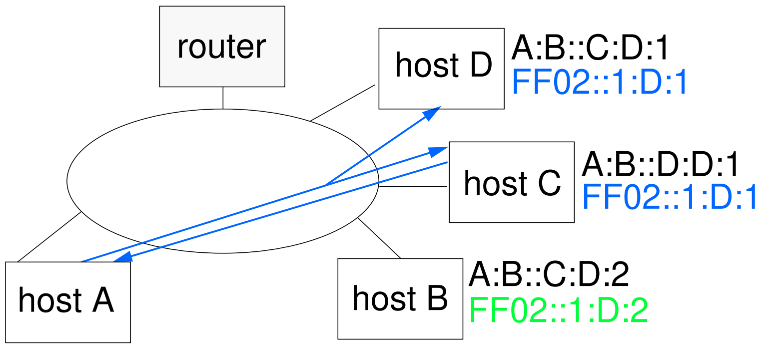
Torsten Braun

IBM European Networking Center

Heidelberg

Multicast for Address Resolution

- IPv4: broadcast of ARP_Request, target node answers by ARP_Reply
- IPv6: multicast of a neighbor solicitation to a small local group (solicited nodes multicast address)
- Solicited nodes multicast group: members with identical last 32 bits (e.g., 32 bits of MAC address)
- Target node responds with neighbor advertisement message.



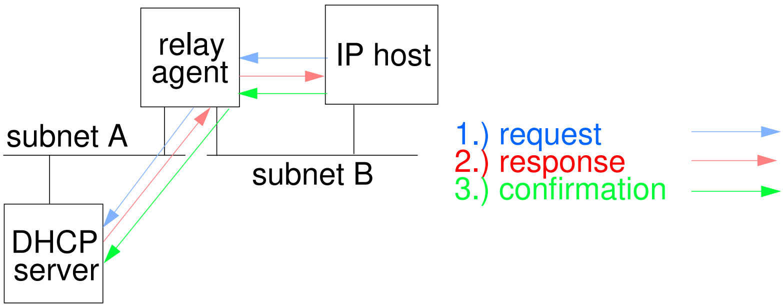
Torsten Braun

IBM European Networking Center

Heidelberg

Dynamic Host Configuration Protocol (DHCPv6)

- Stateful address autoconfiguration (client/server concept)
- Relay agents relay messages between client and DHCP server, if there is no DHCP server at the client's link. Client has initially only a link-local address !



- Client requests IP address or host name by a request message to the well-known link-local DHCP-server/relay-agent multicast address -> response from DHCP server
- Addresses and names with limited validity

Torsten Braun

IBM European Networking Center

Heidelberg

Security

- Authentication
- Privacy (encryption)
- Sender and receiver must agree on a set of parameters
- Security association
  - identified by destination address and security parameters index (SPI)
  - parameters
    - keys for authentication/encryption
    - authentication/encryption algorithm
    - key lifetime
    - association lifetime
    - source address
    - sensitivity level (secret, unclassified)
- Key exchange procedures required

Authentication

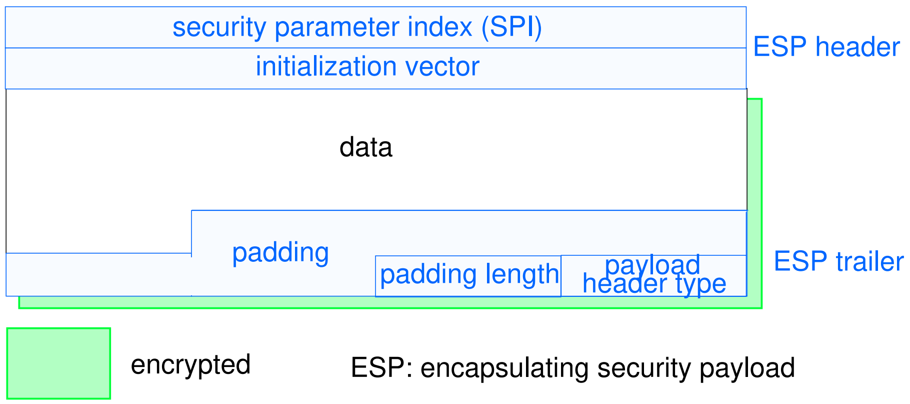
- Check whether the received packet was in fact transmitted by the sender identified as the source of the packet (sender authentication)
- Check whether the data has been manipulated in transit (data integrity)
- Default algorithm: Message Digest 5 (MD5) calculates a 128 bit value over the packet plus a secret key (-> authentication data)

next header	length	reserved
security parameter index (SPI)		
authentication data (n * 32 bit)		

- SPI identifies the pre-established security association and the corresponding parameters.
- Authentication header provides authentication and data integrity for all IPv6 packet fields that do not change en route.

Encryption

- Encryption of the packet for confidential and secure data exchange (privacy)
- DES-CBC as default algorithm
- Initialization vector (n*32 bit) as input for encryption algorithm
- Payload header type: protocol type of the data field (e.g., IP,TCP,UDP)



ESP: encapsulating security payload

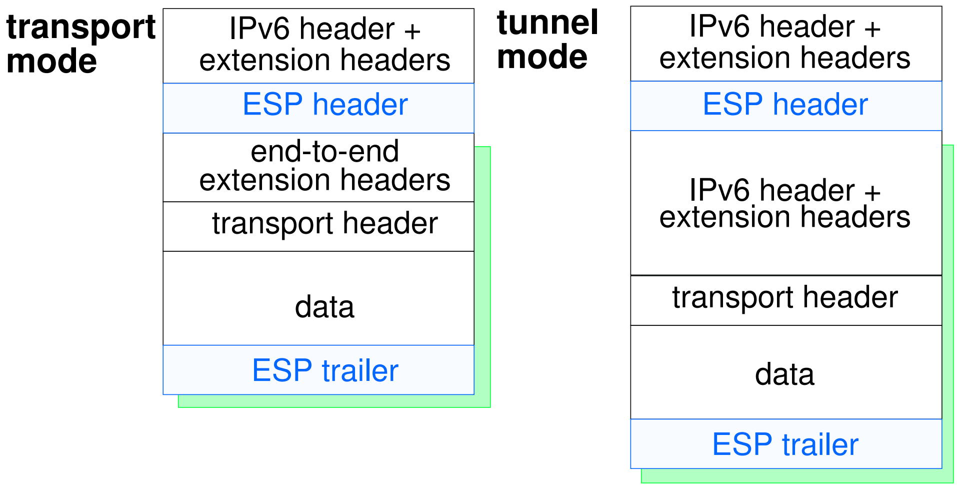
Torsten Braun

IBM European
Networking Center

Heidelberg

Encryption Modes

- Tunnel mode: encryption of the complete IP packet and encapsulation within a new IP packet
- Transport mode: end-to-end extension header and transport protocol packet are encrypted



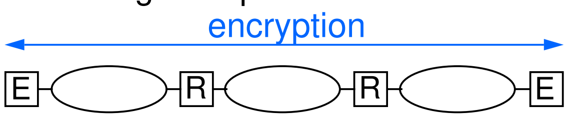
Torsten Braun

IBM European
Networking Center

Heidelberg

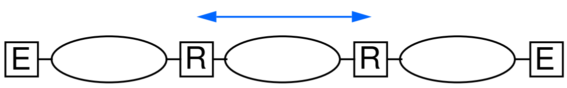
Encryption Scenarios

- End to end using transport mode



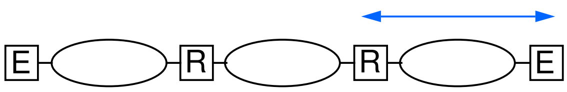
e.g., for confidential data exchange among 2 users

- Security-gateway to security-gateway using tunnel mode



e.g., for data exchange over insecure networks (corporate networks over the Internet)

- End to security-gateway using tunnel mode



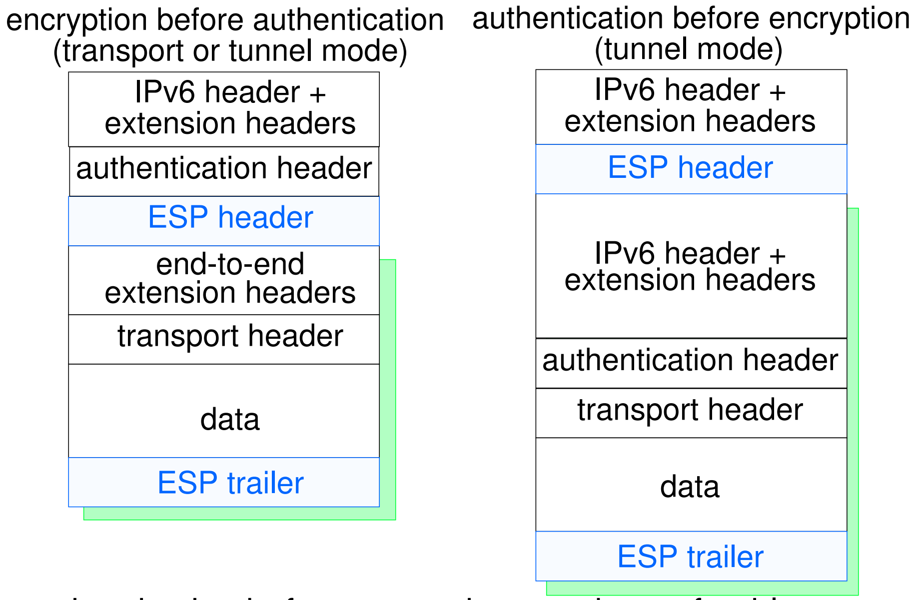
e.g., for corporate network access over insecure access networks (wireless networks)

Torsten Braun

IBM European Networking Center

Heidelberg

Authentication and Encryption



- authentication before encryption may be preferable
  - authentication header protected by ESP
  - encryption before authentication could require re-encryption at the receiver

Torsten Braun

IBM European Networking Center

Heidelberg

Transition Approach

- IPv4 addresses do not run out before the year 2000, transition must then be complete.
- Operating systems are replaced in cycles of ca. 3 years
- The first IPv6 implementations are already commercially available, others will be available in 1997.

=> 2 cycles are required for transition

- During the transition phase both IPv4- and IPv6-nodes will be used.
- Both types of nodes must be able to communicate with each other.
- After the transition phase there will be only very few pure IPv4 nodes.

=> 2 strategies will be necessary:
dual protocol stacks and IPv6-in-IPv4 tunneling

- IPv6/IPv4 header translation is only required, if pure IPv4 nodes want to communicate with pure IPv6 nodes.

Torsten Braun

IBM European Networking Center

Heidelberg

Dual Protocol Stacks

- Dual protocol stacks
  - UDP/IPv4 and UDP/IPv6
  - TCP/IPv4 and TCP/IPv6
- IPv6 as an additional protocol family
- All IPv6-nodes will also have an IPv4 protocol stack during the transition phase.

Application			
Sockets			
UDP for IPv4	TCP for IPv4	UDP for IPv6	TCP for IPv6
IPv4		IPv6	
Network			

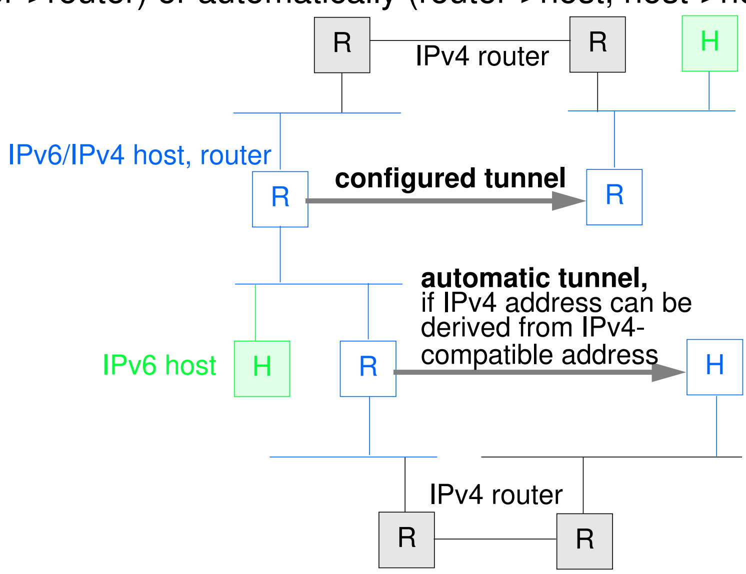
Torsten Braun

IBM European Networking Center

Heidelberg

Tunneling

- Tunneling to connect two IPv6 domains over IPv4 islands -> encapsulation of IPv6 packet in IPv4 packet
- IPv4 address derived from configuration (host->router, router->router) or automatically (router->host, host->host)



IPv4-Compatible Addresses

96 bit	32 bit
0000...0000	IPv4 address

- used by nodes to employ automatic tunneling
- Nodes can use their IPv4 address to build the IPv6 address
- useful in early transition phase only
- no benefits of “native” IPv6 addressing structure

Transition Plan

- maintain complete IPv4 routing system until addresses run out
- upgrade IPv4 routers to dual stack (IPv4 and IPv6) routers
- incrementally build IPv6 backbone (6Bone) routing system
  - use IPv6-over-IPv4 tunnels to construct IPv6 backbone
  - grow like multicast backbone (Mbone)
- de-activate tunnels as soon as underlying path upgraded to IPv6

Torsten Braun

IBM European Networking Center

Heidelberg

Transition Options for Providers

- do nothing
  - IPv6 users cross IPv4 providers
  - no IPv6 services can be provided
- build a tunneled infrastructure over the existing IPv4 system
  - no change to internal IPv4 routing
  - IPv6/IPv4 tunneling by routers at edges
  - IPv6 services can be provided
- build up dual IPv4/IPv6 routing system incrementally up to 100%

Torsten Braun

IBM European Networking Center

Heidelberg

Transition Options for End User Sites

- Upgrade IPv4 hosts to dual IPv4/IPv6 stack hosts incrementally
  - use IPv4-compatible addresses with existing IPv4 address assignments
  - Host-to-host automatic tunneling over IPv4
- deploy IPv6 routers
  - hosts acquire automatically native addresses by address autoconfiguration when on-link IPv6 routers are deployed
- connect IPv6 routers to IPv6 backbone

Torsten Braun

IBM European Networking Center

Heidelberg

Address Translation

- basic concept
  - IP addresses are only unique within a given region
  - DNS host names are globally-unique
  - addresses in packets leaving a region are replaced by addresses unique in the next region
  - similar approach to proxy servers or application-level gateways
- advantage
  - no changes in hosts or internal routers
- disadvantages
  - complexity
  - worse performance in address translation router
  - multicast support

Torsten Braun

IBM European Networking Center

Heidelberg

IPv6 over ATM

- Problem:
  - Control mechanisms are based on multicasting capability of the subnetwork
    (e.g., ND messages to solicited nodes multicast address)
  - IP multicast model (IGMP group management) has been designed for shared medium subnetworks
- Several solutions have been discussed
  - Multicast Address Resolution Server (MARS) as extension of the ATM ARP server (RFC1577) + NHRP in routers
  - hierarchical architecture of neighbor discovery servers (NDS); NDSs exchange and distribute ND messages, hosts send messages to corresponding NDSs
  - NHRP (NBMA Next Hop Resolution Protocol) replace certain ND functions (e.g., address resolution)
- Solutions are under study and will converge

Torsten Braun

IBM European
Networking Center

Heidelberg

Multicast Address Resolution Server

- MARS: Multicast Address Resolution Server
- ATM/IP end systems send special MARS PDUs in order to join or to leave a group
- MARS server maintains lists of the form
{IP multicast address, ATM address 1, ..., ATM address N}

IP multicast address	ATM addresses
224.5.6.7	ATMw, ATMx, ATMy, ATMz
230.6.7.8	ATMv
231.8.7.6	ATMu, ATMx

- ATM addresses are either
  - end system addresses (VC Mesh) or
  - multicast server addresses (MCS).
- The MARS server returns the set of addresses associated with requested IP multicast address.
- Requesting end systems establish point-to-point or point-to-multipoint VCs to the returned ATM end systems for sending multicast data

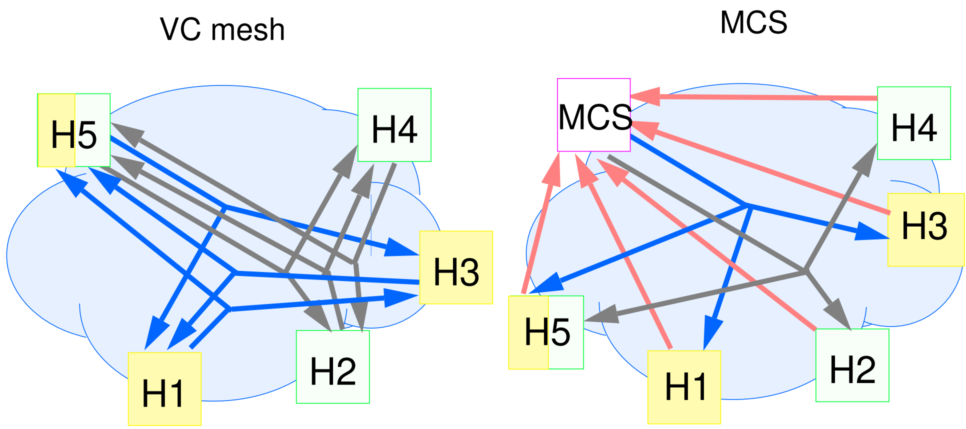
Torsten Braun

IBM European
Networking Center

Heidelberg

MARS: VC Mesh vs. MCS

- VC Mesh: direct data point-to-multipoint VCs from sender to receiver
- MCS: additional hop by multicast server
  - point-to-point VC from sender to MCS
  - point-to-multipoint-VC from MCS to receiver



Torsten Braun

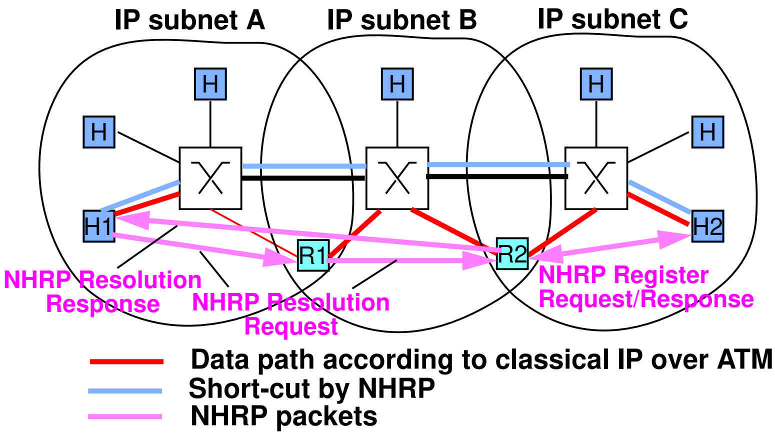
IBM European Networking Center

Heidelberg

SPIE BNE SC 82

IP next generation

NBMA Next Hop Resolution Protocol



1. Registration of hosts (H1, H2) at routers (NHRP servers) (R1, R2) via NHRP Registration Request / Reply
2. H1 sends NHRP Resolution Request to resolve NBMA address of NBMA next hop (destination or nearest NBMA router)
3. NHRP Resolution Request is exchanged among NHRP servers
4. R2 sends H2's NBMA address by NHRP Resolution Reply to H1
5. H1 sends the following data packets directly to H2

Torsten Braun

IBM European Networking Center

Heidelberg

SPIE BNE SC 82

IP next generation

MARS/NHRP for IPv6 over ATM

- MARS for sending multicast ND messages within a logical link group
  - simple approach: use MARS as "black-box" -> high VC resource consumption because of many solicited nodes multicast groups
  - VC resource optimization: send ND multicast messages point-to-point VC to MARS, MARS broadcasts the messages over point-to-multipoint VC to all registered nodes
- NHRP between routers to establish short-cuts between ATM hosts of different logical IP subnets
  - Router forwards packets but tries to find a better next hop for a destination using a NHRP Resolution Request (only for packets originating from a directly attached host)
  - Router sends Redirect messages to the sender of the forwarded packet after receiving the NHRP Resolution Response from another NHRP server
- IPv6 stack independent of link layer

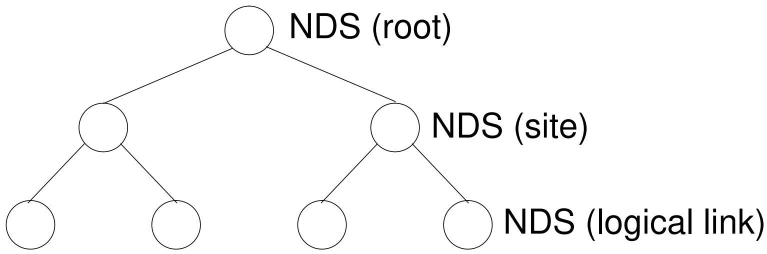
Torsten Braun

IBM European Networking Center

Heidelberg

Neighbor Discovery Servers

- Hierarchical tree of neighbor discovery servers (NDSs).



- Packets to the following multicast addresses are sent via the NDS tree: all-nodes, all-routers, solicited node, DHCP server
- Routers and hosts send ND messages to the next NDS, NDS relays neighbor solicitations but generates new router advertisement
- Multicast messages may travel up the tree depending on the scope via point-to-point VCs
- Messages are then distributed down the tree via point-to-point or point-to-multipoint VCs, NDSs filter messages
- IPv6 stack independent of link layer

Torsten Braun

IBM European Networking Center

Heidelberg

NHRP for ND Functions I

- Router discovery
  - Hosts sends NHRP resolution request to NHS, which responds with the ATM addresses of the routers.
  - Host sends periodically Router Solicitation messages to the routers using link-local addresses
- Address autoconfiguration
  - use 6 right-most octets of ATM address as interface ID
  - hosts register IP addresses at NHSs
- Duplicate address detection
  - add a bit to NHRP register messages to indicate that upper layer / NBMA address binding (in this case: IPv6/ATM) must be unique
  - registration attempts with registered addresses are rejected

Torsten Braun

IBM European Networking Center

Heidelberg

NHRP for ND Functions II

- Neighbor unreachability detection
  - Neighbor solicitation/advertisement indicate unreachability
  - hosts send NHRP resolution requests and receive NHRP purge requests in case of an unreachable neighbor
- Link-layer address resolution
  - Hosts use NHRP to resolve ATM address for a IPv6 address
- Redirects as defined for “traditional” IPv6 ND
  - ND functions are mapped to NHRP or ND dependent on link-level type

Torsten Braun

IBM European Networking Center

Heidelberg

Effects of IPv6 on Higher Layers

- Application modifications are required due to modified socket API (new address family, flow label).
- Maximum payload size changed because of larger IP header.
- Domain name system (DNS)
  - New DNS record type (AAAA)
  - Dynamic DNS updates
- TCP
  - New IP pseudo header requires modifications of the checksum calculation in UDP and TCP implementations
  - TCP does not currently allow that the IP address changes during a TCP connection (IP address may change due to stateful address autoconfiguration)

Torsten Braun

IBM European Networking Center

Heidelberg

Implementation Status

- Several research organizations released IPv6 source code.
- All main router and end system vendors are currently implementing IPv6. ->http://playground.sun.com/ipng/ipng-implementations.html
- Beta releases of several vendors are freely available.
- First commercial implementations are already available.
- Multi-vendor demos and interoperability tests have already been done and more tests are planned (at University of New Hampshire).
- IPv6-Backbone (6Bone) has been established over the Internet.

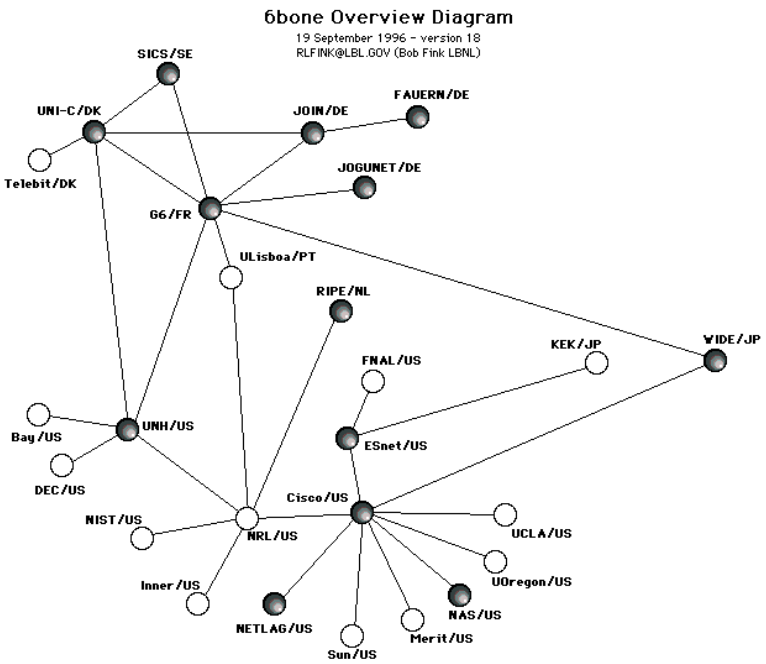
Torsten Braun

IBM European Networking Center

Heidelberg

6Bone

from: http://www-cnr.lbl.gov/6bone/



Specifications I

- IPv6 Specification
  - S. Deering, R. Hinden: Internet Protocol, Version 6 (IPv6) Specification, RFC 1883, December 1995
  - S. Bradner, A. Mankin: The Recommendation for the IP Next Generation Protocol, RFC 1752, January 1995.
- Addressing Architecture
  - R. Hinden, S. Deering: IP Version 6 Addressing Architecture, RFC 1884, December 1995.
  - IAB, IESG: IPv6 Address Allocation Management, RFC 1881, December 1995.
  - Y. Rekhter, T. Li: An Architecture for IPv6 Unicast Address Allocation, RFC 1887, December 1995.
  - Y. Rekhter, P. Lothberg, R. Hinden, S. Deering, J. Postel: An IPv6 Provider-Based Unicast Address Format, Internet Draft, draft-ietf-ipngwg-unicast-addr-fmt-04.txt, April 1996.
  - R. Hinden: IPv6 Testing Address Allocation, RFC 1897, January 1996.
  - R. Elz: Identifying Interfaces in IPv6 link-local addresses, Internet Draft, draft-ietf-ipngwg-iid-02.txt, May 1996.
  - D. Harrington: Identifying Interfaces in IPv6 link-local addresses, Internet Draft, draft-ietf-ipngwg-linkname-00.txt, June 1996.
  - J. Bound, P. Roque: IPv6 Anycasting Service: Minimum requirements for end nodes, Internet Draft, draft-bound-anycast-00.txt, June 1996.
  - J. Bound, B. Carpenter, D. Harrington, J. Houldsworth, A. Lloyd: OSI NSAPs and IPv6, Internet Draft, draft-ietf-ipngwg-nsap-ipv6-01.txt, May 1996.

Torsten Braun
IBM European Networking Center
Heidelberg

Specifications II

- Internet Control Message Protocol
  - A. Conta, S. Deering: Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6), RFC 1885, December 1995.
- Path MTU Discovery
  - J. McCann, S. Deering, J. Mogul: Path MTU Discovery for IP version 6, RFC 1981, August 1996.
- Header Compression
  - M. Degermark, B. Nordgren, S. Pink: Header Compression for IPv6, Internet Draft, draft-degermark-ipv6-hc-01.txt, June 1996.
- Packet Tunneling
  - A. Conta, S. Deering: Generic Packet Tunneling in IPv6 Specification, Internet Draft, draft-ietf-ipngwg-ipv6-tunnel-02.txt, June 1996.
- Domain Name System
  - S. Thomson, C. Huitema: DNS Extensions to support IP version 6, RFC 1886, December 1995.
  - P. Vixie: Reverse Name Lookups of Encapsulated IPv4 Addresses in IPv6, Internet Draft, draft-vixie-ipng-ipv4ptr-00.txt, May 1996.

Specifications III

- Transition Mechanisms
  - R. Gilligan, E. Nordmark: Transition Mechanisms for IPv6 Hosts and Routers, RFC 1933, April 1996.
  - D. Haskin, R. Callon: Routing Aspects Of IPv6 Transition, Internet Draft, draft-ietf-ngtrans-routing-aspects-01.txt, April 1996.
- Routing
  - G.Malkin,R.Minnear:RIPng for IPv6,Internet Draft,draft-ietf rip-riping-03.txt,June1996.
  - Y. Rekhter, P. Traina: IDRP for IP v4 and v6, Internet Draft, draft-ietf-idr-idrp-v4v6-02.txt, January 1996.
  - R. Coltun, D. Ferguson, J. Moy: OSPF for IPv6, Internet Draft, draft-ietf-ospf-ospfv6-02.txt, June 1996.
  - M. Sand, M. Thomas: Multi-homing Support in IPv6, Internet Draft, draft-shand-ipv6-multi-homing-01.txt, February 1996.
  - R. Atkinson, D. Katz: IPv6 Router Alert Option, Internet Draft, draft-cisco-ipv6-router-alert-00.txt, June 1996.
- Mobility
  - W. Simpson: IPv6 Mobility Support, Internet Draft, draft-simpson-ipv6-mobility-00.txt, November 1994.
  - C. Perkins, D. Johnson: Mobility Support in IPv6, Internet Draft, draft-perkins-ipv6-mobility-sup-02.txt, July 1995.
  - F. Teraoka: Mobility Support in IPv6, Internet Draft, draft-teraoka-ipv6-mobility-sup-03.txt April 1996.

Torsten Braun

IBM European
Networking Center

Heidelberg

Specifications IV

- Security
  - R. Atkinson: Security Architecture for the Internet Protocol, RFC 1825, August 1995.
  - R. Atkinson: IP Authentication Header, RFC 1826, August 1995.
  - R. Atkinson: IP Encapsulating Security Payload (ESP), RFC 1827, August 1995.
  - P. Metzger, W. Simpson: IP Authentication using Keyed MD5, RFC 1828, August 1995.
  - P. Karn, P. Metzger, W. Simpson: The ESP DES-CBC Transform, RFC 1829, August 1995.
- Neighbor Discovery and Address Autoconfiguration
  - T. Narten, E. Nordmark, W. Simpson: Neighbor Discovery for IP Version 6 (IPv6), RFC 1970, August 1996.
  - S. Thompson, T. Narten: IPv6 Stateless Address Autoconfiguration, RFC 1971, August 1996.
  - J. Bound: Dynamic Host Configuration Protocol for IPv6 (DHCPv6), Internet Draft, draft-ietf-dhc-dhcpv6-04.txt, February 1996.
  - C. Perkins, Extensions for DHCPv6: Internet Draft, draft-ietf-dhc-v6exts-01.txt, June 1996.
- Program Interfaces
  - R. Gilligan, S. Thomson, J. Bound: IPv6 Program Interfaces for BSD Systems, Internet Draft, draft-ietf-ipngwg-bsd-api-05.txt, April 1996.

Specifications V

- IPv6 over Different Media
  - M. Crawford: A Method for the Transmission of IPv6 Packets over Ethernet Networks, Internet Draft, draft-ietf-ipngwg-ethernet-ntwrks-02.txt, February 1996.
  - M. Crawford: A Method for the Transmission of IPv6 Packets over Ethernet Networks, RFC 1972, August 1996.
  - M. Crawford: Transmission of IPv6 Packets Over FDDI, RFC 2019, October 1996.
  - S. Thomas: A Method for the Transmission of IPv6 Packets over Token Ring Networks, Internet Draft, draft-ietf-ipngwg-token-ring-01.txt, January 1996.
  - D. Haskin, E. Allen: IP Version 6 over PPP, Internet Draft, draft-ietf-ipngwg-pppext-ipv6cp-03.txt, May 1996.
  - R. Atkinson, D. Haskin, J. Luciani: IPv6 over NBMA Networks, Internet Draft, draft-ietf-ion-ipv6-nbma-00.txt, June 1996.
  - G. Armitage: Transient Neighbors for IPv6 ATM, Internet Draft, draft-armitage-ion-tn-00.txt, June 1996.
  - G. Armitage: IPv6 and Neighbor Discovery over ATM, Internet Draft, draft-ietf-ion-ipv6nd-00.txt, April 1996.

Torsten Braun
IBM European Networking Center
Heidelberg

Specifications VI

- IPv6 over ATM
  - G. Armitage: IPv6 and Neighbor Discovery over ATM, Internet Draft, draft-ietf-ion-ipv6nd-00.txt, June 1996
  - R. Atkinson, D. Haskin, J. Luciani: IPv6 over NBMA Networks, Internet Draft, draft-ietf-ion-ipv6-nbma-00.txt, June 1996
  - P. Schulter: A Framework for IPv6 Over ATM, Internet Draft, draft-schulter-ipv6atm-framework-01.txt, February 1996.
  - R. Elz: Identifying Interfaces in IPv6 link-local addresses, Internet Draft, draft-ietf-ipngwg-iid-01.txt, February 1996.
  - G. Armitage: Transient Neighbors for IPv6 over ATM, Internet Draft, draft-armitage-ion-tn-00.txt, June 1996

Torsten Braun
IBM European Networking Center
Heidelberg

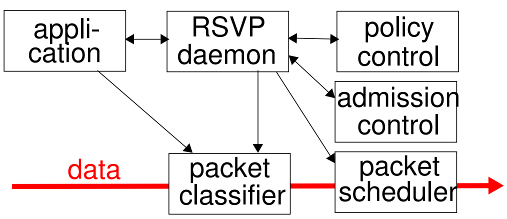
IPv6 References

- http://playground.sun.com/ipng
- http://www.computermethods.com/ipng/IPng.htm
- http://www.ietf.cnri.reston.va.us/html.charters/ipngwg-charter.html
- http://www.iol.unh.edu/interop/ip/htmls/info1.html
- W. Stallings: IPv6: The New Internet Protocol, IEEE Communications Magazine, July 1996, pp. 96-108
- R. Hinden: IP Next Generation Overview, Communications of the ACM, June 1996, pp. 61-71
- T. Braun: Die Internet-Protokollfamilie der nächsten Generation, PIK 2/96, pp. 94-102 (in German)
- C. Huitema: IPv6 The New Internet Protocol, Prentice Hall, 1996
- S. Bradner and A. Mankin: IPng Internet Protocol Next Generation, Addison-Wesley, 1996

Torsten Braun
IBM European Networking Center
Heidelberg

Resource Reservation

- Hop-by-hop reservation of resources to ensure quality-of-service of data transfer (link, node, end system)
- Reservation requirements
  - admission control (determines whether the node has sufficient available resources to supply the requested QoS)
  - policy control (determines whether the user has administrative permission to make the reservation)
  - scheduling of classified packets



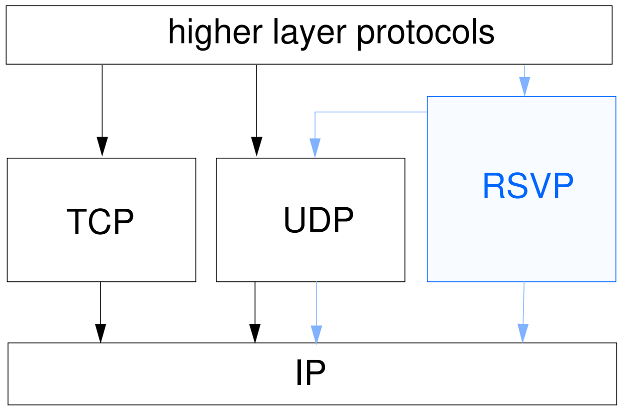
- Examples
  - Networks: ATM, ISDN
  - Protocols: ST2, RSVP

Torsten Braun

IBM European
Networking Center

Heidelberg

Resource Reservation Protocol (RSVP)



- RSVP runs over UDP or IP alternatively
- User data exchange over TCP/IP or UDP/IP
- RSVP is a reservation protocol and does not transfer user data, i.e. RSVP supports QoS in TCP/IP and UDP/IP stacks
- Dynamic receiver-oriented reservation based on soft states for simplex data streams, supports heterogeneous receivers

Torsten Braun

IBM European
Networking Center

Heidelberg

Flow and Session

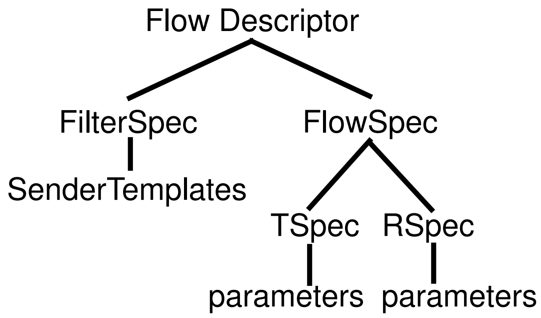
- Flow
  - Sequence of packets between 2 end points
  - Identification by filter specification and session ID
- Session
  - set of flows to a single destination (unicast or multicast)
  - identification by session ID consisting of
    - source IP address
    - IP protocol ID
    - destination IP address

Torsten Braun

IBM European Networking Center

Heidelberg

Flow Descriptor



- Filter specification (FilterSpec)
  - defines the subset of packets (of a session) to be served with the QoS described in the FlowSpec
  - parameter for packet classification
- Flow specification (FlowSpec)
  - defines the desired QoS of a flow
  - parameters for packet scheduling components

Torsten Braun

IBM European Networking Center

Heidelberg

Filter Specification

- IPv4
  - source IP address
  - source port
- IPv6
  - source IP address
  - source port
- IPv6
  - source IP address
  - flow label

- IPv6 flow labels allows to describe the FilterSpec by layer-3-information only!

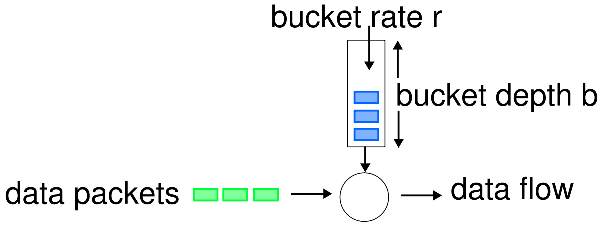
Torsten Braun

IBM European Networking Center

Heidelberg

Flow Specification

- Service class
- TSpec (traffic specification) is based on a token bucket description:
  - bucket depth b [bytes]
  - bucket rate r [bytes/s]; Token bucket is filled with rate r. 1 token is generated during the time interval 1/r. During a time interval T, no more than $r \cdot T + b$ data can be sent.



  - Further TSpec parameters
    - peak rate p [bytes/second]
    - minimum policed unit m [bytes]
    - maximum packet size M [bytes]
- RSpec (requested service specification)

Torsten Braun

IBM European Networking Center

Heidelberg

IP Integrated Service (Intserv) Classes

- Several service classes have been defined in addition to IP's best-effort service:
  - Controlled load service
  - Controlled delay service
  - Predictive delay service
  - Committed rate service
  - Guaranteed service

Torsten Braun

IBM European Networking Center

Heidelberg

Controlled Load Service

- Characterisitcs
  - provides quality-of-service (QoS) closely approximating the QoS one would get from an unloaded network
  - high percentage of successfully delivered packets
  - delay will not greatly exceed the minimum delay
- Applications
  - intended for applications that are highly sensitive to overload conditions such as adaptive real-time applications
- FlowSpec
  - TSpec: r, b, p, m, M
- Control Functions
  - admission control to assure the service in an overloaded network
  - packet scheduling

Torsten Braun

IBM European Networking Center

Heidelberg

Committed Rate Service

- Characteristics
  - provides applications with a firm commitment from the network, that the requested minimum transmission rate is available
  - not associated with a delay guarantee
  - delay of any packet should not significantly exceed the delay over a dedicated circuit
- Applications
  - intended for applications that require a given amount of bandwidth in order to work properly
- FlowSpec
  - TSpec: r, b, p, m, M
  - RSpec: desired service rate R [bytes/s]
- Control Functions
  - admission control for R, traffic policing, reshaping to R

Guaranteed Service

- Characteristics
  - guarantees delay and bandwidth but not does not try to minimize jitter
- Applications
  - intended for audio/video playback applications
- FlowSpec
  - TSpec: r, b, p, m, M
  - RSpec:
    - desired service rate R [bytes/s]
    - slack term s: difference between desired delay and delay obtained using reservation level R
      can be used by network elements to reduce the resource reservations
- Control Functions
  - admission control, traffic policing, reshaping

Reservation Styles

- FF style: the resources are reserved for a single sender
- WF style: all senders share the reserved resources
- SE style: the explicitly defined senders share the reserved resources

sender selection	distinct reservation	shared reservation
explicit	fixed filter (FF)	shared explicit (SE)
wildcard	undefined	wildcard filter (WF)

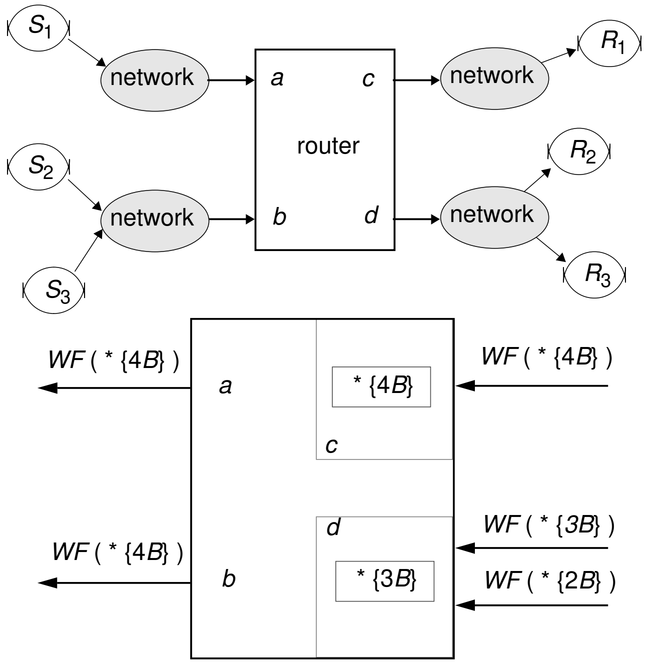
- distinct reservation is used for video conferencing: multiple senders are sending simultaneously
- shared reservation is used for audio conferencing: only one speaker should be active at a given time

Torsten Braun

IBM European Networking Center

Heidelberg

Wildcard Filter (WF)



Notation: filter type(sender{resource}), WF -> max. reservation

SPIE BNE SC 82

IP next generation

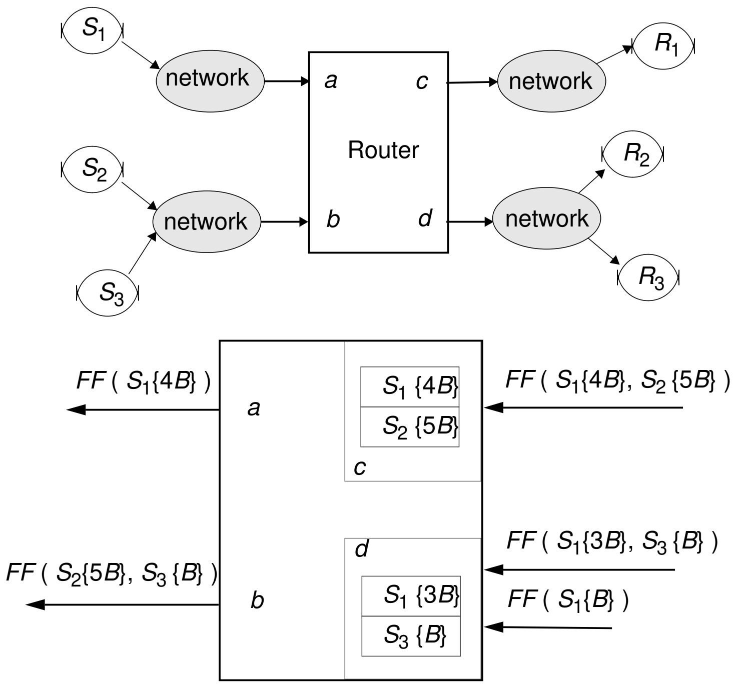
October 1996 - 83

Torsten Braun

IBM European Networking Center

Heidelberg

Fixed Filter (FF)



FF -> reservation are added

SPIE BNE SC 82

IP next generation

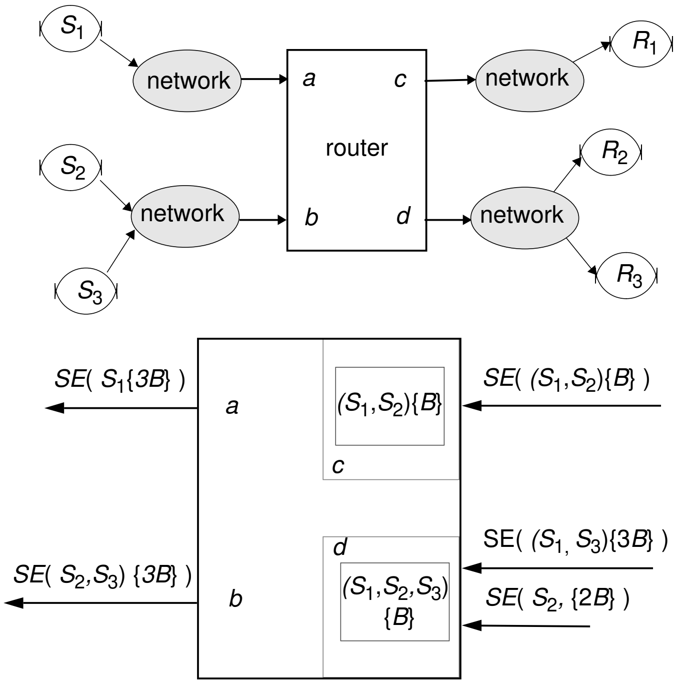
October 1996 - 84

Torsten Braun

IBM European Networking Center

Heidelberg

Shared Explicit (SE) Filter



Torsten Braun

IBM European Networking Center

Heidelberg

RSVP Messages

- PATH
  - transmission of path messages from the sender to the destination address, e.g. messages about data flow characteristics
  - establishment and maintenance of the path by sending PATH messages periodically
- RESV
  - resource reservation on the return path from the receiver to the sender (routers store the previous hop)
  - periodic reservation refreshments
  - soft reservation states
  - dynamic reservations
  - heterogeneous receivers
- RESV_CONFIRM to confirm a reservation
- error messages: PATH_ERR, RESV_ERR
- termination of a reservation: PATH_TEAR, RESV_TEAR

Torsten Braun

IBM European Networking Center

Heidelberg

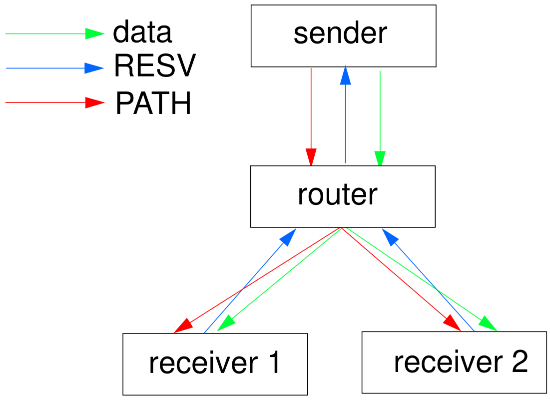
SPIE BNE SC 82

IP next generation

October 1996 - 87

PATH/RESV Message Exchange

- PATH messages are sent with the same source and destination address as the data using the same route
- PATH messages indicate session parameters
- Reservation messages (RESV) are sent periodically on the reverse path from each receiver to the sender using unicast packets
- Intermediate routers reserve appropriate resources based on RESV messages



Torsten Braun

IBM European Networking Center

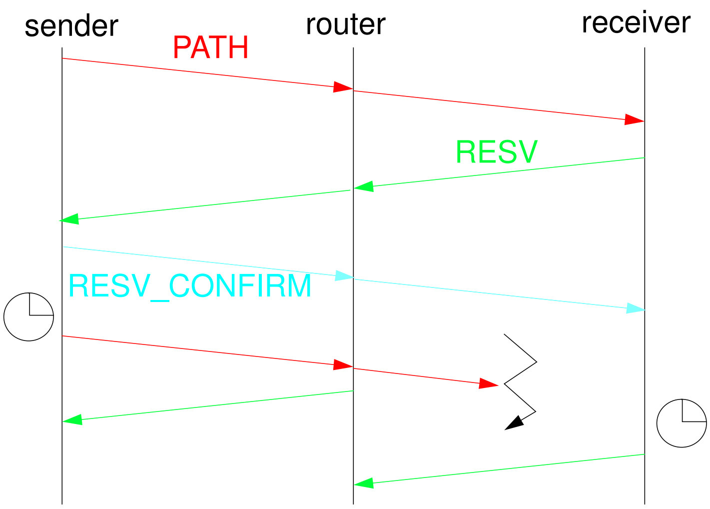
Heidelberg

SPIE BNE SC 82

IP next generation

October 1996 - 88

Establishment of a Reservation



- PATH and RESV messages are sent periodically (-> error control) or for reservation changes
  -> soft-state
- RESV messages must be acknowledged if confirmation flag is set

Torsten Braun

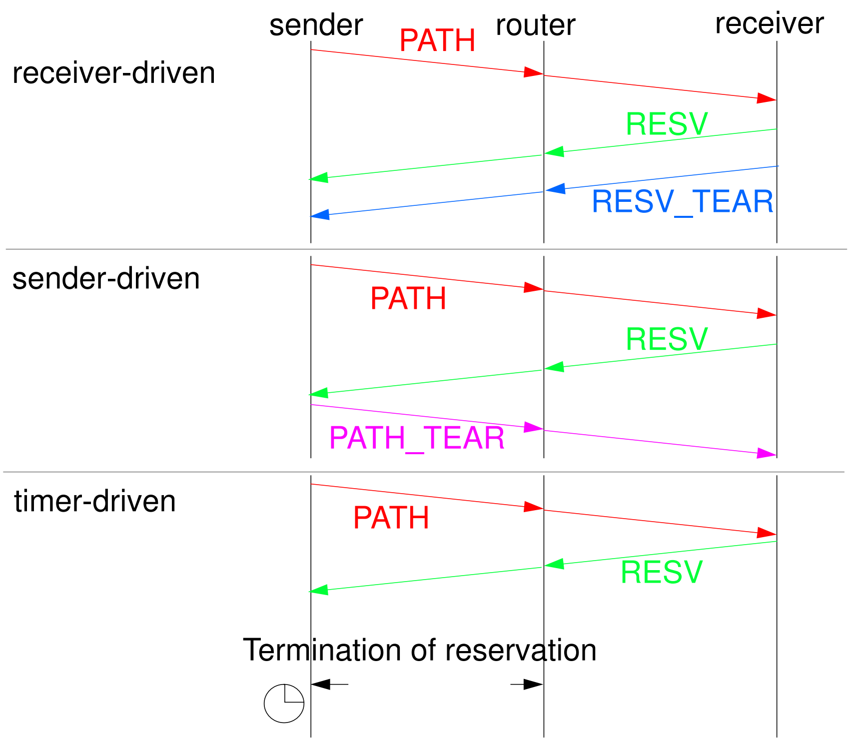
IBM European Networking Center

Heidelberg

SPIE BNE SC 82

IP next generation

Termination of a Reservation



Torsten Braun

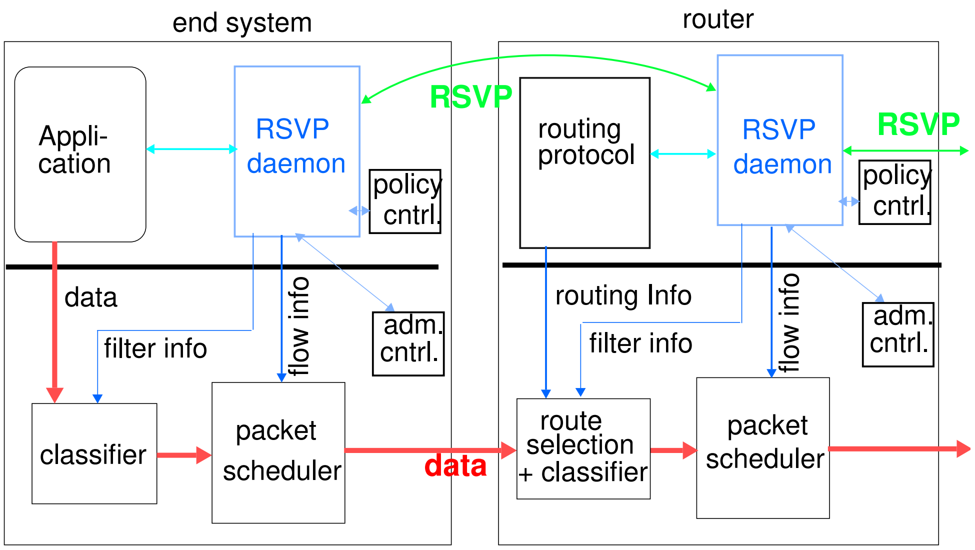
IBM European Networking Center

Heidelberg

SPIE BNE SC 82

IP next generation

Implementation Model

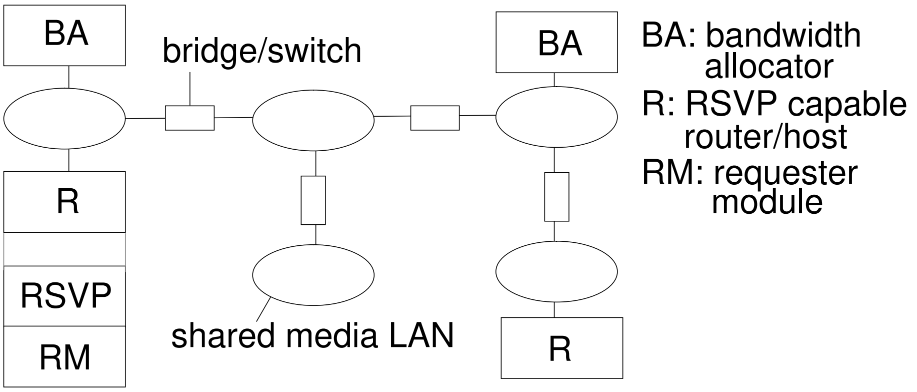


Torsten Braun

IBM European Networking Center

Heidelberg

RSVP over Shared Media



- protocol (e.g., LLRMP, SBM) between RM and BA for
  - admission control
  - bandwidth reservation
  - policing
    - traffic shaping
    - packet scheduling at sources and switches
    - flow mapping to LAN priorities

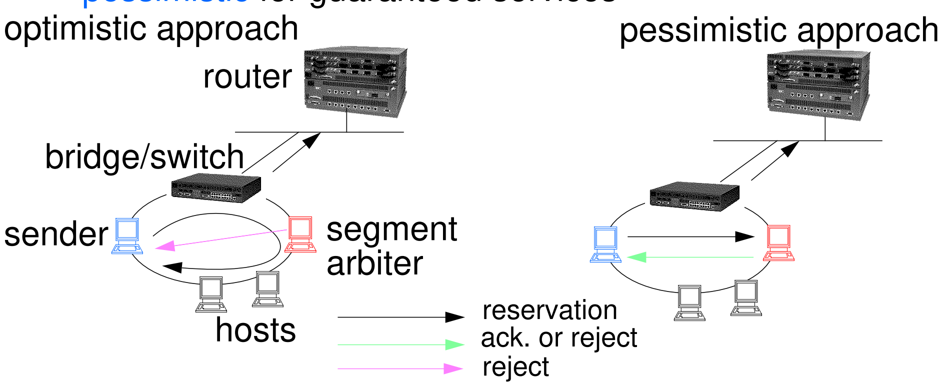
Torsten Braun

IBM European Networking Center

Heidelberg

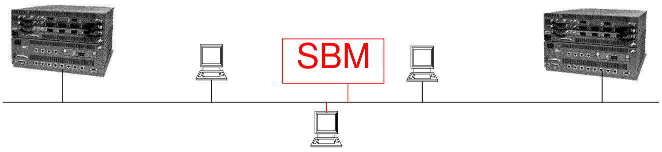
Link Level Resource Management Protocol

- soft-state
- independent of network layer resource reservation protocol
- 2 approaches
  - optimistic for services which do not provide hard guarantees, e.g. controlled load service
  - pessimistic for guaranteed services



Subnet Bandwidth Manager (SBM)

- 2 mechanisms to support admission control so that aggregate traffic does not exceed available bandwidth
  - RSVP-based admission control
    - receivers send RSVP RESV messages in response to PATH messages via the SBMs to the previous hop
    - RSVP extension to address SBMs and previous hop
  - admission control for non-RSVP flows
    SBM_RESV message by receiver ->SBM_ERROR or SBM_CONM by SBM



- networks with several switches require topology discovery
- mechanisms are performed between distributed SBMs that are responsible for bandwidth management on separate LAN portions

Comparison RSVP/ATM

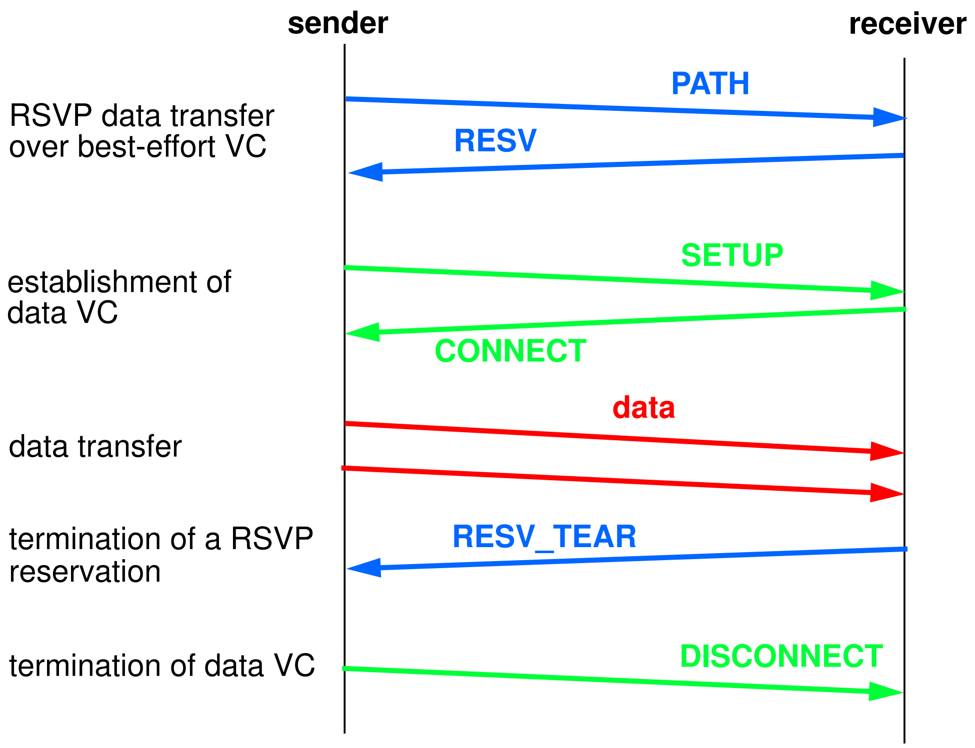
	RSVP	ATM (UNI 3.x)
initiator of the reservation	receiver	sender
connection	soft-state	explicit
connection types	m:n uni-directional	1:1 bi-directional 1:n uni-directional
relation routing/ reservation	independent	integrated
heterogeneity of QoS	heterogeneous	homogeneous
dynamic/static QoS	dynamic	static

Torsten Braun

IBM European
Networking Center

Heidelberg

RSVP over ATM



Torsten Braun

IBM European
Networking Center

Heidelberg

Multicast Issues for RSVP over ATM

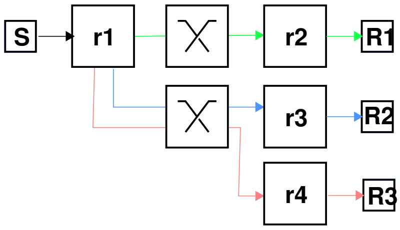
- Data VC Management due to heterogeneous receivers
  - many VCs per reservation
  - 2 VCs per reservation
  - 1 VC per reservation
  - VC aggregation
- Control VC Management
  - same control VC as for data
  - single control VC per reservation
  - multiplexed point-to-multipoint VCs
  - multiplexed point-to-point VCs
- Receiver Transitions (move receivers to another VC)
  - add before leave -> duplicated packets
  - leave before add -> lost packets
- Dynamic QoS (-> tear down old VC and establish new one)

Torsten Braun

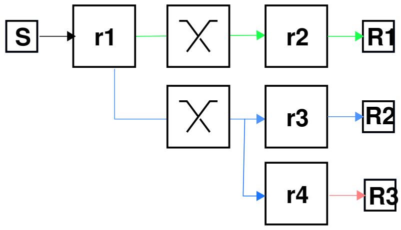
IBM European Networking Center

Heidelberg

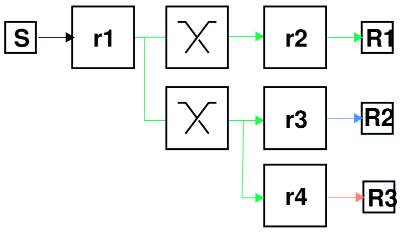
Data VC Management



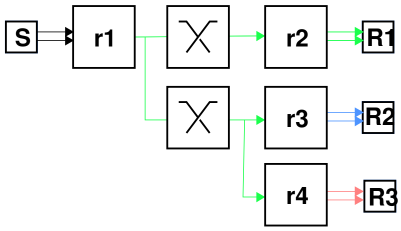
full heterogeneity with many VCs per reservation



limited heterogeneity with two VCs (QoS and best-effort) per reservation



single VC for each reservation



VC aggregation with multiple reservations per VC

Torsten Braun

IBM European Networking Center

Heidelberg

Control VC Management

- Mixed data and control traffic
  - [+] no additional RSVP VC
  - [–] RSVP messages may be dropped due to non-conforming data traffic
- Parallel RSVP VC for each reservation
  - [+] no dropping of RSVP packets due to non-conforming data traffic
  - [–] twice the number of VCs (resources, signaling)
- Single point-to-multipoint VC for each ingress router and each set of egress routers
  - [+] reduced number of VCs
  - [–] selection and management of point-to-multipoint VC is complex
- Multiplexed point-to-point VCs
  - [+] VCs can also be used in reverse direction
  - [–] data must be duplicated for different VCs

Intserv/ATM QoS Parameter Mappings

- ATM service classes

UBR	unspecified bit rate	best-effort service
ABR	available bit rate	feedback mechanism to control loss
nrt-VBR	non-real-time variable bit rate	for variable bandwidth applications
rt-VBR	real-time variable bit rate	for variable bandwidth applications with delay constraints
CBR	constant bit rate	emulates a leased line

- Mappings of integrated services to ATM service classes

IP integrated service	ATM service	Comment
best-effort	UBR	natural service mapping
	ABR	in particular for TCP
controlled-load	nrt-VBR	sustainable cell rate = bucket rate
	ABR	minimum cell rate = bucket rate
guaranteed	CBR	utilization loss
	rt-VBR	reasonable but not-controlled delay

RSVP Specifications I

- RSVP
  - R. Braden, L. Zhang, S. Berson, S. Herzog, S. Jamin: Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification, Internet Draft, draft-ietf-rsvp-spec-12.ps, May 1996
- Integrated Services
  - S. Shenker, C. Partridge, R. Guerin: Specification of Guaranteed Quality of Service, Internet Draft, draft-ietf-intserv-guaranteed-svc-06.txt, August 1996
  - J. Wroclawski: Specification of the Controlled-Load Network Element Service, Internet Draft, draft-ietf-intserv-ctrl-load-svc-03.txt, August 1996
  - F. Baker, R. Guerin, D. Kandlur: Specification of Committed Rate Quality of Service, Internet Draft, draft-ietf-intserv-commit-rate-svc-00.txt, June 1996
  - J. Wroclawski: The Use of RSVP with IETF Integrated Services, Internet Draft, draft-ietf-intserv-rsvp-use-00.txt, August 1996
  - B. Braden, D. Clark, S. Shenker: Integrated Services in the Internet Architecture: an Overview, RFC 1633, June 1994
- Integrated Services over ATM
  - S. Berson: IP Integrated Services with RSVP over ATM, Internet Draft, draft-ietf-issll-atm-support-01.ps, Septmber 1996
  - M. Borden: Interoperation of Controlled-Load and Guaranteed-Service with ATM, Internet Draft, draft-issll-intserv-atm-mapping-01.txt, September 1996
  - M. Borden, E. Crawley, B. Davie, S. Batsell: Integration of Real-Time Services in an IP-ATM Network Architecture, RFC 1831, August 1995

RSVP Specifications II

- RSVP over Shared Media
  - A. Ghanwani, J. Pace, V. Srinivasan: A Framework for Providing Integrated Services Over Shared and Switched LAN Technologies, Internet Draft, draft-ghanwani-framework-is-lan-0.txt, September 1996
  - R. Yavatkar, D. Hoffmann, Y. Bernet, E. Patki: SBM (Subnet Bandwidth Manager): A Proposal for Admission Control over Ethernet, Internet Draft, draft-yavatkar-sbm-ethernet-01.txt, September 1996
  - P. Kim: Link Level Resource Management Protocol (LLRMP), Internet Draft, draft-kim-llrmp-00.ps, June 1996

Torsten Braun

IBM European Networking Center

Heidelberg

RSVP/Intserv References

- http://www.isi.edu/div7/rsvp/rsvp.html
- http://www.ietf.cnri.reston.va.us/html.charters/rsvp-charter.html
- L. Zhang, S. Deering, D. Estrin, S. Shenker, D. Zappala: RSVP: A New Resource ReSerVation Protocol, IEEE Network, September 1993
- S. Shenker, L. Breslau: Two Issues in Reservation Establishment, ACM SIGCOMM ’95.
- S. Shenker: Fundamental Design Issues for the Future Internet, IEEE Journal on Selcted Areas in Communications, Vol. 13, No. 7
- S. Berson: Classical RSVP and IP over ATM, INET ’96, April 1996