

Chair for Network Architectures and Services – Prof. Carle
 Department for Computer Science
 TU München

**Master Course
 Computer Networks
 IN2097**

Prof. Dr.-Ing. Georg Carle
 Christian Grothoff, Ph.D.

Chair for Network Architectures and Services
 Institut für Informatik
 Technische Universität München
<http://www.net.in.tum.de>

TUM
 Technische Universität München

Chapter 4: Network Layer

Part 1

- **Introduction**
- IP: Internet Protocol
 - Datagram format
 - IPv4 addressing
 - ICMP

Part 2

- IPv6
- NAT and NAT Traversal
- Virtual circuit and datagram networks
- What's inside a router

Part 3

- Routing algorithms
 - Link state
 - Distance Vector
 - Hierarchical routing
- Routing in the Internet
 - RIP
 - OSPF
 - BGP
- Broadcast and multicast routing

IN2097 - Master Course Computer Networks, WS 2009/2010 3

Chapter 4: Network Layer

Chapter goals:

- understand principles behind network layer services:
 - network layer service models
 - forwarding versus routing
 - how a router works
 - routing (path selection)
 - dealing with scale
 - advanced topics: IPv6, mobility
- instantiation, implementation in the Internet

IN2097 - Master Course Computer Networks, WS 2009/2010 2

Network layer

- transport segment from sending to receiving host
- on sending side encapsulates segments into datagrams
- on rcving side, delivers segments to transport layer
- network layer protocols in every host, router
- router examines header fields in all IP datagrams passing through it

IN2097 - Master Course Computer Networks, WS 2009/2010 4

Two Key Network-Layer Functions

- **routing**: determine route taken by packets from source to dest.
 - **routing algorithms**
- **forwarding**: move packets from router's input to appropriate router output

analogy:

- **routing**: process of planning trip from source to dest
- **forwarding**: process of getting through single interchange

IN2097 - Master Course Computer Networks, WS 2009/2010 5

Datagram networks

- no call setup at network layer
- routers: no state about end-to-end connections
 - no network-level concept of "connection"
- packets forwarded using destination host address
 - packets between same source-dest pair may take different paths

IN2097 - Master Course Computer Networks, WS 2009/2010 7

Interplay between routing and forwarding

header value	output link
0100	3
0101	2
0111	2
1001	1

value in arriving packet's header: 0111

IN2097 - Master Course Computer Networks, WS 2009/2010 6

Forwarding table

4 billion possible entries

Destination Address Range	Link Interface
11001000 00010111 00010000 00000000 through 11001000 00010111 00010111 11111111	0
11001000 00010111 00011000 00000000 through 11001000 00010111 00011000 11111111	1
11001000 00010111 00011001 00000000 through 11001000 00010111 00011111 11111111	2
otherwise	3

IN2097 - Master Course Computer Networks, WS 2009/2010 8

Longest prefix matching

Prefix Match	Link Interface
11001000 00010111 00010	0
11001000 00010111 00011000	1
11001000 00010111 00011	2
otherwise	3

Examples

DA: 11001000 00010111 0001**0110 10100001** Which interface?

DA: 11001000 00010111 0001**1000 10101010** Which interface?

IN2097 - Master Course Computer Networks, WS 2009/2010 9

The Internet Network layer

Host, router network layer functions:

The diagram illustrates the network layer functions within a host or router. It shows a stack of layers: Transport layer (TCP, UDP), Network layer, Link layer, and physical layer. The Network layer is highlighted with a red double-headed arrow and contains three main components: Routing protocols (path selection, RIP, OSPF, BGP), IP protocol (addressing conventions, datagram format, packet handling conventions), and ICMP protocol (error reporting, router "signaling"). A forwarding table is also shown, connected to the Routing protocols and IP protocol components.

IN2097 - Master Course Computer Networks, WS 2009/2010 11

Chapter 4: Network Layer

Part 1

- Introduction
- **IP: Internet Protocol**
 - Datagram format
 - IPv4 addressing
 - ICMP

Part 2

- IPv6
- NAT and NAT Traversal
- Virtual circuit and datagram networks
- What's inside a router

Part 3

- Routing algorithms
 - Link state
 - Distance Vector
 - Hierarchical routing
- Routing in the Internet
 - RIP
 - OSPF
 - BGP
- Broadcast and multicast routing

IN2097 - Master Course Computer Networks, WS 2009/2010 10

IP datagram format

The diagram shows the structure of an IP datagram, which is 32 bits long. The fields are: IP protocol version number, header length (bytes), "type" of data, max number remaining hops (decremented at each router), upper layer protocol to deliver payload to, 32-bit source IP address, 32-bit destination IP address, Options (if any), and data (variable length, typically a TCP or UDP segment). The header is further divided into: version, header length, type of service, 16-bit identifier, flags, fragment offset, time to live, upper layer header, and checksum. Annotations include: "total datagram length (bytes)", "for fragmentation/reassembly", "E.g. timestamp, record route taken, specify list of routers to visit.", and "how much overhead with TCP?" with a list: □20 bytes of TCP, □20 bytes of IP, □= 40 bytes + app layer overhead.

IN2097 - Master Course Computer Networks, WS 2009/2010 12

IP Fragmentation & Reassembly

- network links have MTU (max. transfer size) - largest possible link-level frame.
 - different link types, different MTUs
- large IP datagram divided ("fragmented") within net
 - one datagram becomes several datagrams
 - "reassembled" only at final destination
 - IP header bits used to identify, order related fragments

fragmentation:
in: one large datagram
out: 3 smaller datagrams

reassemble

IN2097 - Master Course Computer Networks, WS 2009/2010 13

Chapter 4: Network Layer

Part 1

- Introduction
- IP: Internet Protocol**
 - Datagram format
 - IPv4 addressing**
 - ICMP

Part 2

- IPv6
- NAT and NAT Traversal
- Virtual circuit and datagram networks
- What's inside a router

Part 3

- Routing algorithms
 - Link state
 - Distance Vector
 - Hierarchical routing
- Routing in the Internet
 - RIP
 - OSPF
 - BGP
- Broadcast and multicast routing

IN2097 - Master Course Computer Networks, WS 2009/2010 15

IP Fragmentation and Reassembly

Example

- 4000 byte datagram
- MTU = 1500 bytes

length	ID	fragflag	offset
=4000	=x	=0	=0

One large datagram becomes several smaller datagrams

length	ID	fragflag	offset
=1500	=x	=1	=0

length	ID	fragflag	offset
=1500	=x	=1	=185

length	ID	fragflag	offset
=1040	=x	=0	=370

1480 bytes in data field

offset = 1480/8

IN2097 - Master Course Computer Networks, WS 2009/2010 14

IP Addressing: introduction

- IP address**: 32-bit identifier for host, router *interface*
- interface**: connection between host/router and physical link
 - router's typically have multiple interfaces
 - host typically has one interface
 - IP addresses associated with each interface

223.1.1.1, 223.1.1.2, 223.1.1.3, 223.1.1.4, 223.1.2.1, 223.1.2.2, 223.1.3.1, 223.1.3.2

223.1.1.1 = 11011111 00000001 00000001 00000001

223 1 1 1

IN2097 - Master Course Computer Networks, WS 2009/2010 16

Subnets

- IP address:
 - subnet part (high order bits)
 - host part (low order bits)
- What's a subnet?
 - device interfaces with same subnet part of IP address
 - can physically reach each other without intervening router

network consisting of 3 subnets

IN2097 - Master Course Computer Networks, WS 2009/2010 17

Subnets

How many?

IN2097 - Master Course Computer Networks, WS 2009/2010 19

Subnets

Recipe

- To determine the subnets, detach each interface from its host or router, creating islands of isolated networks. Each isolated network is called a **subnet**.

Subnet mask: /24

IN2097 - Master Course Computer Networks, WS 2009/2010 18

IP addressing: CIDR

CIDR: Classless InterDomain Routing

- subnet portion of address of arbitrary length
- address format: **a.b.c.d/x**, where x is # bits in subnet portion of address

200.23.16.0/23

IN2097 - Master Course Computer Networks, WS 2009/2010 20

IP addresses: how to get one?

Q: How does a *host* get IP address?

- hard-coded by system admin in a file
 - Windows: control-panel->network->configuration->tcp/ip->properties
 - UNIX: /etc/rc.config
- **DHCP:** Dynamic Host Configuration Protocol: dynamically get address from as server
 - “plug-and-play”

IN2097 - Master Course Computer Networks, WS 2009/2010 21

DHCP client-server scenario

IN2097 - Master Course Computer Networks, WS 2009/2010 23

DHCP: Dynamic Host Configuration Protocol

- **Goal:** allow host to dynamically obtain its IP address from network server when it joins network
 - Can renew its lease on address in use
 - Allows reuse of addresses (only hold address while connected an “on”)
 - Support for mobile users who want to join network (more shortly)
- DHCP overview:
 - host broadcasts “DHCP discover” msg
 - DHCP server responds with “DHCP offer” msg
 - host requests IP address: “DHCP request” msg
 - DHCP server sends address: “DHCP ack” msg

IN2097 - Master Course Computer Networks, WS 2009/2010 22

DHCP client-server scenario

IN2097 - Master Course Computer Networks, WS 2009/2010 24

IP addresses: how to get one?

Q: How does *network* get subnet part of IP addr?
A: gets allocated portion of its provider ISP's address space

ISP's block	11001000	00010111	00010000	00000000	200.23.16.0/20
Organization 0	11001000	00010111	00010000	00000000	200.23.16.0/23
Organization 1	11001000	00010111	00010010	00000000	200.23.18.0/23
Organization 2	11001000	00010111	00010100	00000000	200.23.20.0/23
...
Organization 7	11001000	00010111	00011110	00000000	200.23.30.0/23

IN2097 - Master Course Computer Networks, WS 2009/2010 25

Hierarchical addressing: more specific routes

ISPs-R-Us has a more specific route to Organization 1

Organization 0: 200.23.16.0/23

Organization 2: 200.23.20.0/23

Organization 7: 200.23.30.0/23

Organization 1: 200.23.18.0/23

Fly-By-Night-ISP: "Send me anything with addresses beginning 200.23.16.0/20"

ISPs-R-Us: "Send me anything with addresses beginning 199.31.0.0/16 or 200.23.18.0/23"

Internet

IN2097 - Master Course Computer Networks, WS 2009/2010 27

Hierarchical addressing: route aggregation

Hierarchical addressing allows efficient advertisement of routing information:

Organization 0: 200.23.16.0/23

Organization 1: 200.23.18.0/23

Organization 2: 200.23.20.0/23

Organization 7: 200.23.30.0/23

Fly-By-Night-ISP: "Send me anything with addresses beginning 200.23.16.0/20"

ISPs-R-Us: "Send me anything with addresses beginning 199.31.0.0/16"

Internet

IN2097 - Master Course Computer Networks, WS 2009/2010 26

IP addressing: the last word...

Q: How does an ISP get block of addresses?
A: ICANN: Internet Corporation for Assigned Names and Numbers

- allocates addresses
- manages DNS
- assigns domain names, resolves disputes

IN2097 - Master Course Computer Networks, WS 2009/2010 28

Chapter 4: Network Layer

Part 1

- Introduction
- **IP: Internet Protocol**
 - Datagram format
 - IPv4 addressing
 - **ICMP**

Part 2

- IPv6
- NAT and NAT Traversal
- Virtual circuit and datagram networks
- What's inside a router

Part 3

- Routing algorithms
 - Link state
 - Distance Vector
 - Hierarchical routing
- Routing in the Internet
 - RIP
 - OSPF
 - BGP
- Broadcast and multicast routing

IN2097 - Master Course Computer Networks, WS 2009/2010 29

Traceroute and ICMP

- Source sends series of UDP segments to dest
 - First has TTL =1
 - Second has TTL=2, etc.
 - Unlikely port number
- When nth datagram arrives to nth router:
 - Router discards datagram
 - And sends to source an ICMP message (type 11, code 0)
 - Message includes name of router & IP address
- When ICMP message arrives, source calculates RTT
- Traceroute does this 3 times
- **Stopping criterion**
- UDP segment eventually arrives at destination host
- Destination returns ICMP "host unreachable" packet (type 3, code 3)
- When source gets this ICMP, stops.

IN2097 - Master Course Computer Networks, WS 2009/2010 31

ICMP: Internet Control Message Protocol

- used by hosts & routers to communicate network-level information
 - error reporting:

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header
 - echo request/reply (used by ping)
- network-layer "above" IP:
 - ICMP msgs carried in IP datagrams
- **ICMP message:** type, code plus first 8 bytes of IP datagram causing error

IN2097 - Master Course Computer Networks, WS 2009/2010 30