# Chapter4. Network Layer - Data Plane

2019年12月17日 10:07

## NOTE TAKING AREA

### Overview

Network layer protocols in every host, router. Router examines header fields in all IP datagrams passing through it.

Network-layer functions: forwarding and routing.

- Forwarding: router's input to output.
- Routing: determine route taken packets from source to destination.

Data plane: local, per-router function, forwarding function.

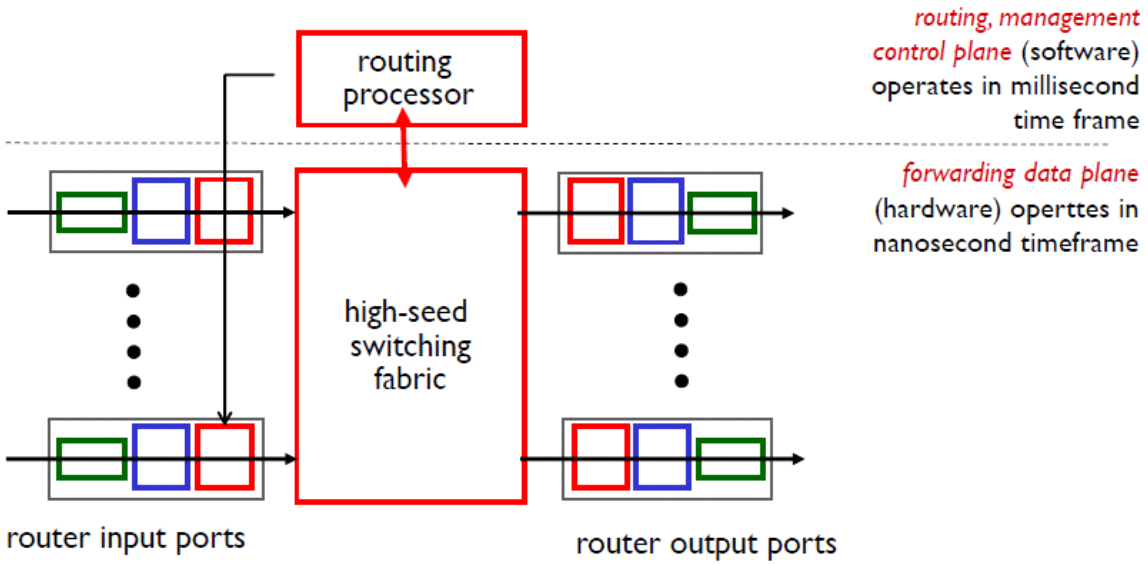
Control plane: network-wide logic (forwarding function).

- Traditional routing algorithms: in routers (per-router).
- Software-defined networking (SDN): in remote servers (logically centralized).

Service model: individual datagram guaranteed delivery, flow of datagrams in-order delivery.

Internet service model provide "best effort" service, no guarantee on **bandwidth**, **loss**, **order** or **timing**.

### Router architecture



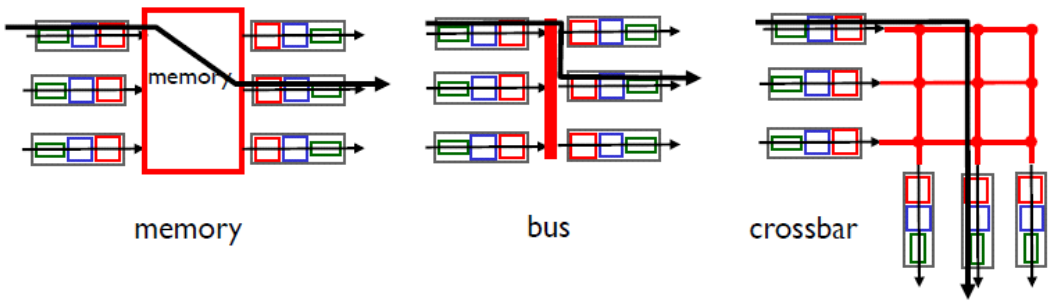

Input port functions: -> line termination -> data link layer protocol (receive) -> lookup, forwarding, queueing (decentralized switching) ->.

Forwarding: destination-based forwarding, generalized forwarding.

- Destination-based forwarding: longest prefix matching.

Switching fabrics: from input to output, has a switching rate.

Three types: memory, bus, crossbar.

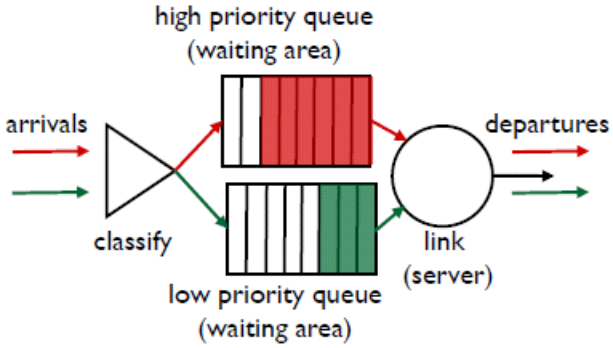



Output ports: -> datagram buffer, queueing -> link layer protocol (send) -> line termination ->.

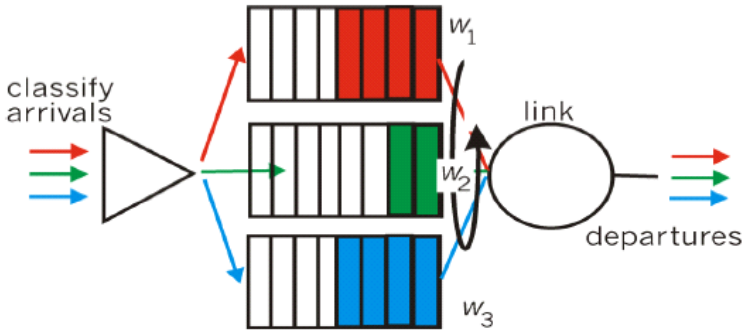
Scheduling datagrams: priority scheduling, network neutrality.

Scheduling mechanisms: FIFO scheduling, tail drop / priority / random.

- Priority: multiple classes, with different priorities.




- Round Robin (RR) scheduling: multiple classes, send complete packet.
- Weighted Fair Queueing (WFQ): generalized RR, with weight.




IP: Internet Protocol

Network layer components: routing protocols (path selection: RIP, OSPF, BGP), forwarding table, IP protocol (addressing conversion, datagram format, packet handling conventions), ICMP protocol (error reporting).

IP fragmentation, reassembly: divide according to MTU, and reassemble.

IPv4 address: 32-bit for interface. Each interface has an IP address.

Subnets: IP address subnet part (high, subnet mask) and host part (low).

Classless Inter Domain Routing (CIDR): $a.b.c.d/x$.

Get an IP address: static, Dynamic Host Configuration Protocol (DHCP).

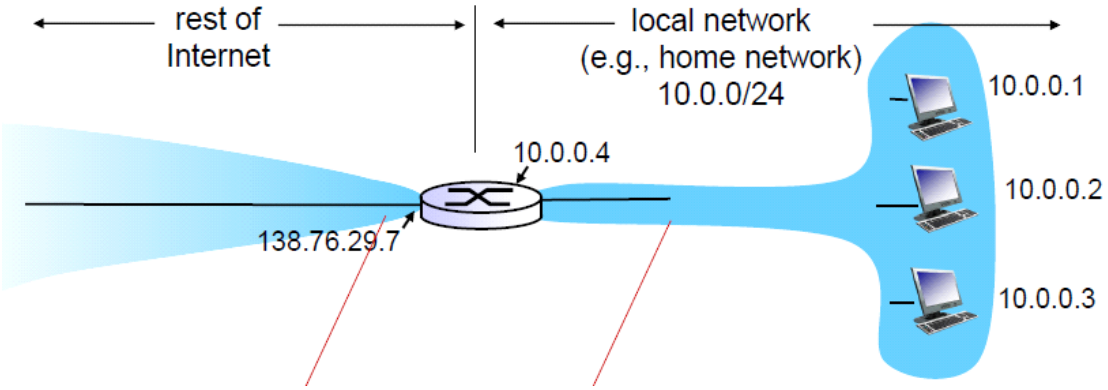
DHCP: discover, offer, request, ack.

DHCP -> UDP -> IP -> Ethernet -> Physics.

Hierarchical addressing: route aggregation.

ISP get block of addresses from ICANN (Internet Corporation for Assigned Names and Numbers).

NAT (Network Address Translation): intermediate between Internet and subnet.




Implementation NAT router:

- Outgoing datagrams replace source IP to NAT IP, source port to NAT new port.
- Remember (in NAT translation table) source pair to NAT pair.
- Incoming datagrams replace pair to corresponding subnet pair.

NAT is controversial.

IPv6: header format helps speed processing / forwarding, header changes to facilitate QoS, fixed-length 40 byte header, no fragmentation allowed.

Remove checksum, options indicated by "next header", ICMPv6.

Handle both IPv4 and IPv6: tunneling, IPv6 datagram carried as payload in IPv4 datagram among IPv4 routers.

## Generalized forwarding and SDN

Routers containing a flow table being computed and distributed by a logically centralized routing controller.

Open flow data plane abstraction: flow defined by header fields.

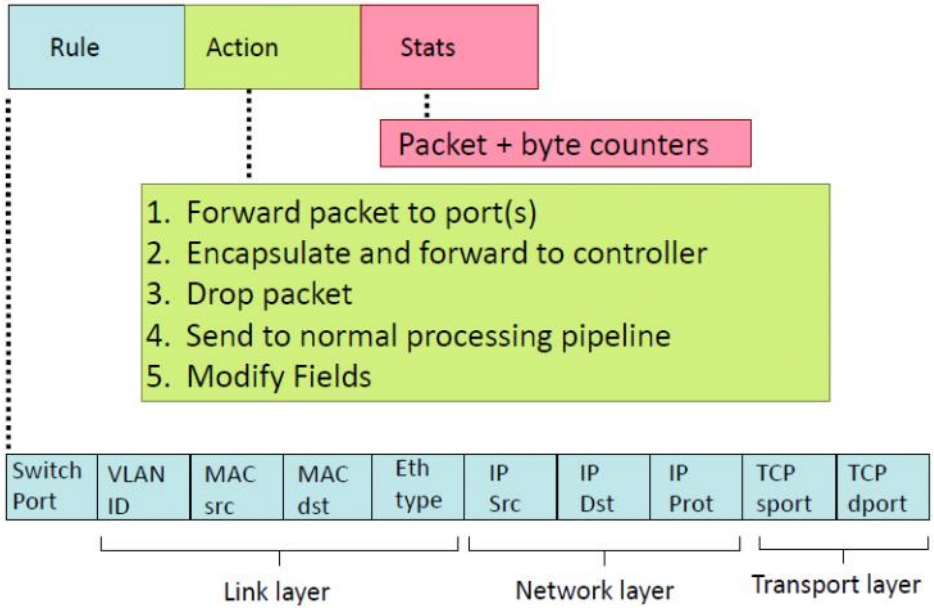
Generalized forwarding (simple packet-handling rules):

1. src=1.2.*.*, dest=3.4.5.* → drop
2. src = *.*.*.*, dest=3.4.*.* → forward(2)
3. src=10.1.2.3, dest=*.*.*.* → send to controller

- Pattern: match values in packet header fields.
- Actions: for matched packet, drop, forwarding, modify, match, send.
- Priority: disambiguate overlapping patterns.
- Counters: #bytes and #packets.

Flow table defines router's match and action rules.

Flow table entries: rule, action, and stats.




Match + action unifies different kinds of device: router, switch, firewall, NAT.

## CUE COLUMN

Example of longest prefix matching

| Destination Address Range | Link interface |
|---|---|
| 11001000 00010111 00010*** ********* | 0 |
| 11001000 00010111 00011000 ********* | 1 |
| 11001000 00010111 00011*** ********* | 2 |
| otherwise | 3 |

examples:

DA: 11001000 00010111 00010110 10100001 which interface?

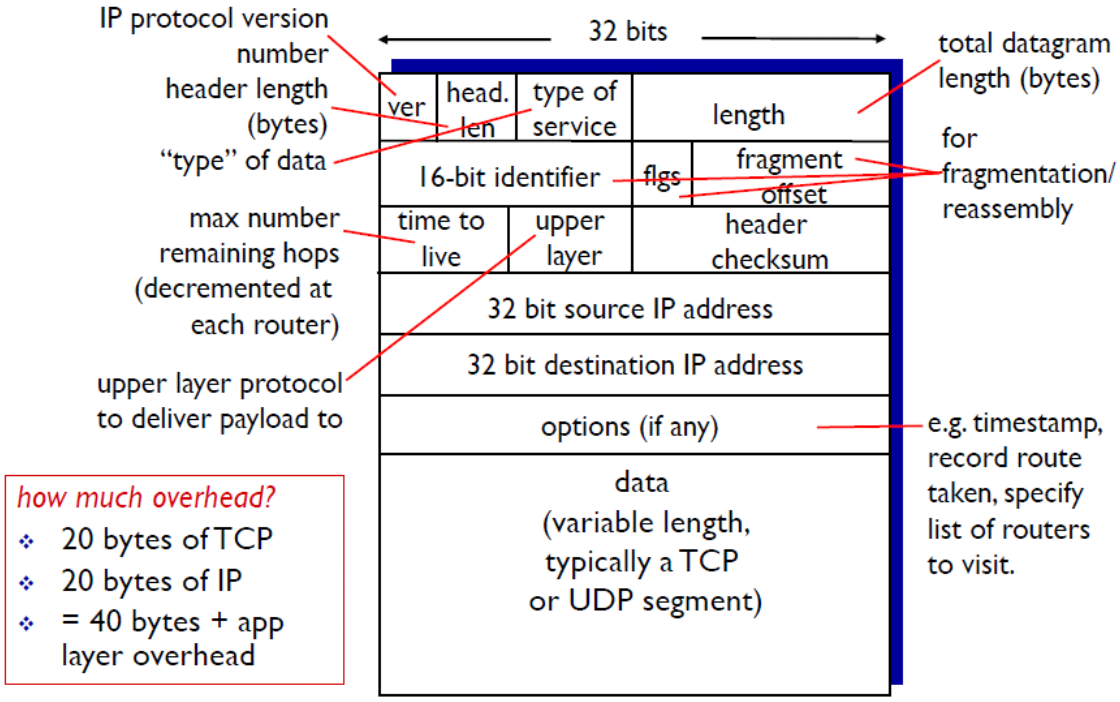
DA: 11001000 00010111 00011000 10101010 which interface?

## Router output buffering size

RFC 3439 rule of thumb: $\boldsymbol{buffer = RTT \times link\ capacity\ C}$.

Recent recommendation: $\boldsymbol{buffer} = \frac{\boldsymbol{RTT \times C}}{\sqrt{\boldsymbol{N}}}, N\ is\ number\ of\ flows.$

## IP datagram format



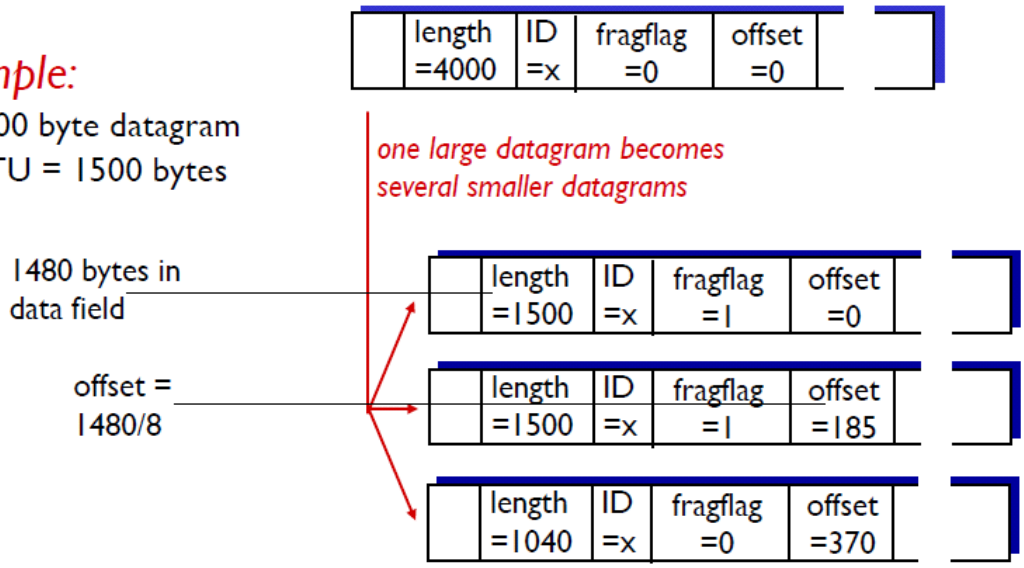

*how much overhead?*

- 20 bytes of TCP
- 20 bytes of IP
- = 40 bytes + app layer overhead

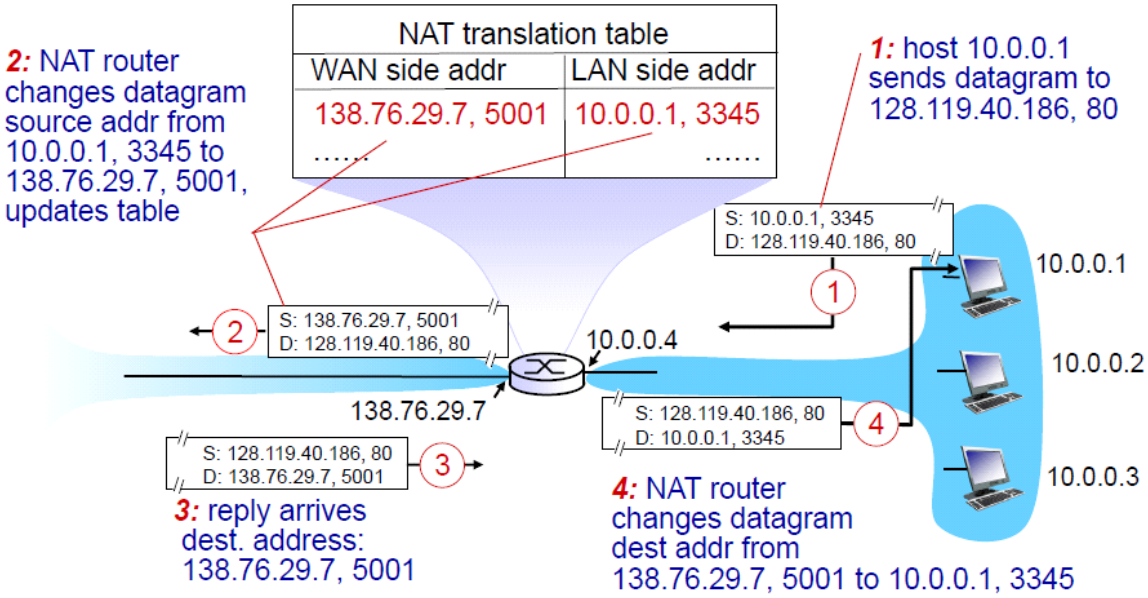
## IP fragmentation and reassembly example

## example:

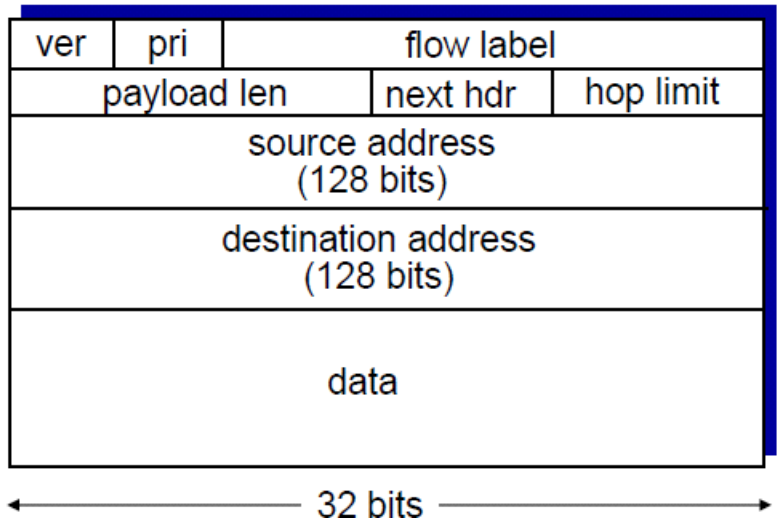
- 4000 byte datagram
- MTU = 1500 bytes




## NAT example




## IPv6 datagram format




Priority: among datagrams in flow.
Flow label: identify datagrams in same flow.
Next header: identify upper layer protocol for data.

## Example of flow table

Destination-based forwarding:

| Switch Port | MAC src | MAC dst | Eth type | VLAN ID | IP Src | IP Dst | IP Prot | TCP sport | TCP dport | Action |
|---|---|---|---|---|---|---|---|---|---|---|
| * | * | * | * | * | * | 51.6.0.8 | * | * | * | port6 |

IP destined to 51.6.0.8 should be forwarded to output port 6.

Firewall:

| Switch Port | MAC src | MAC dst | Eth type | VLAN ID | IP Src | IP Dst | IP Prot | TCP sport | TCP dport | Action |
|---|---|---|---|---|---|---|---|---|---|---|
| * | * | * | * | * | * | * | * | * | 22 | drop |

Do not forward TCP destination port 22.

Destination-based layer 2 (switch) forwarding:

| Switch Port | MAC src | MAC dst | Eth type | VLAN ID | IP Src | IP Dst | IP Prot | TCP sport | TCP dport | Action |
|---|---|---|---|---|---|---|---|---|---|---|
| * | 22:A7:23:11:E1:02 | * | * | * | * | * | * | * | * | port3 |

Frames from MAC address 22:A7:23:22:E1:02 should be forwarded to output port 6.

## SUMMARIES

1. Overview of network layer: data plane and control plane, forwarding and routing, service model.
2. Router architecture: input port, switching fabrics, output ports.
3. IP: subnet, DHCP, NAT, IPv6.
4. Generalized forwarding and SDN: flow table entries.