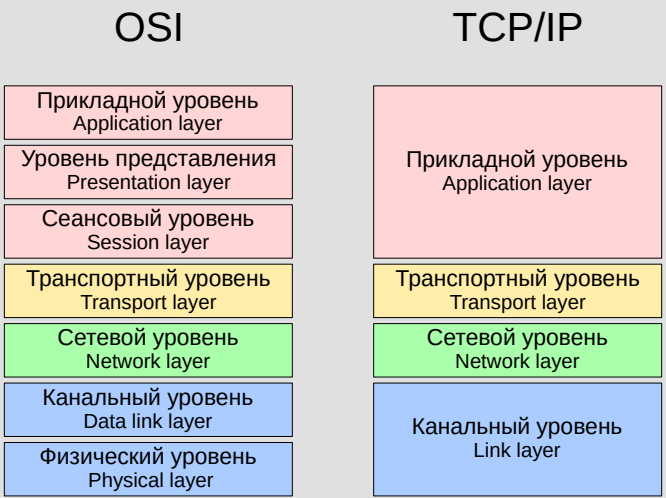
Факультет
биоинженерии и биоинформатики
Московский государственный университет имени М.В.Ломоносова

Сетевые средства GNU/Linux

Устройство компьютерных сетей

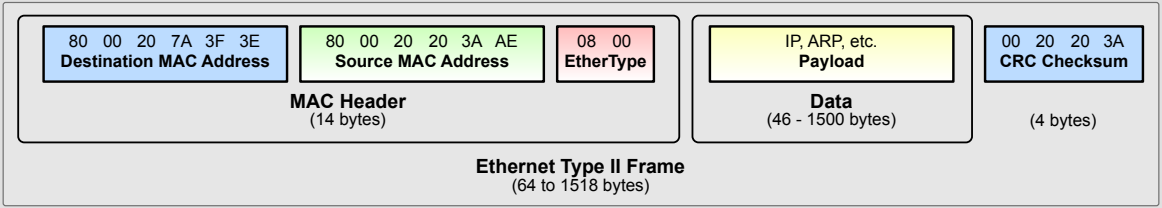
Стек сетевых протоколов



Канальный уровень

- Обмен данными в пределах сегмента сети.
- Единица передачи – кадр (фрейм, frame).
- Адрес – MAC (Media Access Control).

Ethernet

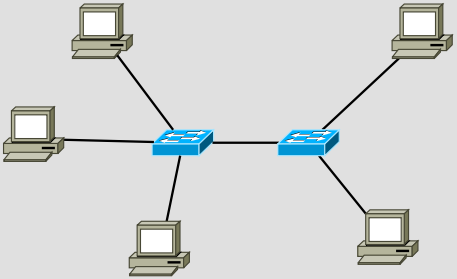

80 00 20 7A 3F 3E
Destination MAC Address
80 00 20 20 3A AE
Source MAC Address
08 00
EtherType
MAC Header
(14 bytes)
IP, ARP, etc.
Payload
Data
(46 - 1500 bytes)
00 20 20 3A
CRC Checksum
(4 bytes)
Ethernet Type II Frame
(64 to 1518 bytes)

Концентратор и коммутатор

Объединяют сегменты сети.

Концентратор (хаб) – передает полученные фреймы во все остальные порты (уже практически не используются).

Коммутатор (свитч) – передает фреймы только в нужный сегмент сети на основе динамической таблицы коммутации.



Сетевой уровень

- Обмен данными между подсетями с помощью маршрутизаторов.
- Определяет путь передачи данных.
- Единица передачи – пакет.
- Адрес – IPv4 или IPv6.

Internet Protocol (IP)

Актуальны две несовместимые версии протокола: IPv4 и IPv6, отличаются далеко не только адреса.

IP-адрес:

IPv4 – 4 октета в десятичном виде, разделенные точкой (127.0.1.255);

IPv6 – 8 групп шестнадцатеричных чисел 0-FFFF через двоеточие (2a02:06b8:0000:0000:0000:0000:0002:0242 -> 2a02:6b8::2:242).

Основные отличия IPv6:

- адресов хватит всем, NAT не необходим;
- маршрутизаторы не могут фрагментировать пакет;
- оптимизация заголовка;
- поддержка огромных пакетов;
- метки потоков;
- многоадресное вещание.

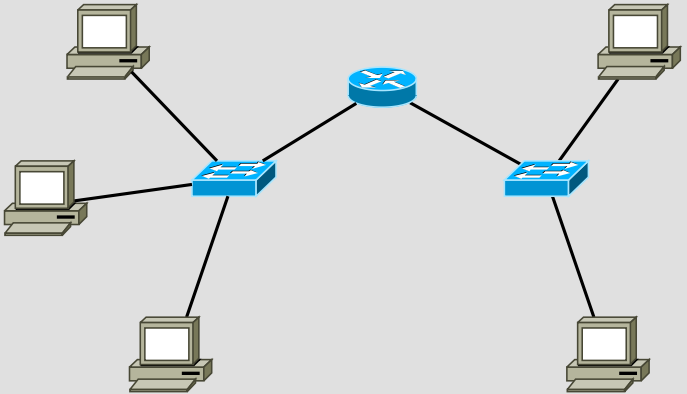
Address Resolution Protocol (ARP)

Протокол для определения MAC-адреса компьютера в локальной сети по IP-адресу.

- Широковещательный (по MAC) запрос Who-has с требуемым IP-адресом.
- Все устройства в сегменте сети получают запрос.
- Если IP запроса совпадает с IP устройства, то оно шлет ответный фрейм.
- Искомый адрес содержится в поле "MAC отправителя".
- Соответствие IP-MAC кэшируется на некоторое время.

Маршрутизатор

Устройство, перенаправляющее пакеты между подсетями согласно таблицам маршрутизации.



Network Address Translation (NAT)

Изменение роутером IP-адреса и, возможно, порта отправителя/получателя пакета при пересылке.

Виды NAT:

Статический – статическая взаимооднозначная трансляция IP.

Динамический – динамическая (временная) взаимооднозначная трансляция IP.

Перегруженный (PAT) – динамическая взаимооднозначная трансляция пар IP-порт – позволяет сократить количество зарегистрированных IP вплоть до одного.

Назначение NAT:

- решение проблемы недостатка адресов IP;
- предотвращение доступа к устройствам в локальной сети извне;
- сокрытие стандартных портов.

Транспортный уровень

- Определяет механизм передачи данных (байтов).
- Единица передачи – датаграмма (UDP) или сегмент (TCP).
- Адрес – номер порта.

User Datagram Protocol (UDP)

- Не гарантирует доставку всех датаграмм.
- Не гарантирует правильный порядок датаграмм.
- Может гарантировать целостность полученных датаграмм.

Заголовок датаграммы UDP

	октет	0								1								2								3							
октет	бит	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	0	Порт отправителя																Порт получателя															
4	32	Длина датаграммы																Контрольная сумма															

Transmission Control Protocol (TCP)

- Обе стороны должны подтвердить установку и завершение соединения.
- На каждый переданный сегмент требуется подтверждение получения;
- иначе сегмент посылается заново.
- Сегменты нумеруются, поэтому гарантируется сохранение их порядка.
- Существует механизм динамического контроля скорости передачи (не очень надежный).

Прикладной уровень

- Определяет формат данных, которыми программы обмениваются по сети.
- Существует множество протоколов прикладного уровня (HTTP, FTP, SMTP, SSH, ...).
- Передаваемые данные могут быть текстовыми или бинарными.
- Данные часто шифруются, например с помощью SSL/TLS.

Сетевые средства Linux

Сетевой интерфейс

Представление сетевого устройства (в том числе виртуального) в ядре Linux.

```
$ ip address
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
       valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
       valid_lft forever preferred_lft forever
2: enp2s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP group default qlen 1000
    link/ether xx:xx:xx:xx:xx:xx brd ff:ff:ff:ff:ff:ff
    inet 192.168.0.99/24 brd 192.168.0.255 scope global dynamic enp2s0
       valid_lft 646sec preferred_lft 646sec
3: wlp3s0: <BROADCAST,MULTICAST> mtu 1500 qdisc noop state DOWN group default qlen 1000
    link/ether xx:xx:xx:xx:xx:xx brd ff:ff:ff:ff:ff:ff
```

IP-сокет

Файлоподобный объект для передачи данных по протоколу IP.

Типы:

- STREAM – соединение по протоколу TCP;
- DGRAM – соединение по протоколу UDP;
- RAW – доступ на уровне IP.

Netfilter

nftables – подсистема ядра Linux для фильтрации, изменения и отслеживания пакетов.

- Правило фильтрации состоит из проверяемых критериев (expressions) и действий (statements), которые выполняются, если пакет удовлетворил всем критериям.
- Правила фильтрации пакетов объединены в цепочки.
- Цепочки организованы в таблицы.

Пользовательская программа для управления nftables – `nft` .

Таблицы маршрутизации

Набор правил, определяющих путь пакета (сетевой интерфейс, шлюз), в зависимости от IP-адреса получателя.

```
# local
broadcast 127.0.0.0 dev lo proto kernel scope link src 127.0.0.1
local 127.0.0.0/8 dev lo proto kernel scope host src 127.0.0.1
broadcast 127.255.255.255 dev lo proto kernel scope link src 127.0.0.1
broadcast 192.168.0.0 dev enp2s0 proto kernel scope link src 192.168.0.99
local 192.168.0.99 dev enp2s0 proto kernel scope host src 192.168.0.99
broadcast 192.168.0.255 dev enp2s0 proto kernel scope link src 192.168.0.99

# main
default via 192.168.0.1 dev enp2s0
192.168.0.0/24 dev enp2s0 proto kernel scope link src 192.168.0.101
```

Менеджер сетевых подключений

Системный демон, позволяющий настраивать сетевые интерфейсы автоматически при включении, или вручную с помощью пользовательского интерфейса.

- NetworkManager – наиболее популярный менеджер: GUI, CLI, апплет для системного трея.
- ConnMan – распространенная альтернатива, более минималистичный.
- Wicd – еще одна альтернатива.
- systemd-networkd – менеджер в составе systemd, без GUI.
- ...

Конфигурация подключений без менеджера

- Утилита ip – ip link, ip address, ip route, ip rule;
- dhclient – настройка интерфейса по DHCP;
- пакет ifupdown – программы ifup и ifdown и конфигурация в /etc/network/interfaces;
- systemd – юниты systemd.link.

Сетевые службы

Domain Name System (DNS)

Распределенная иерархическая система для получения информации о интернет-доменах и используемый протокол.

- Доменное имя – человекочитаемый идентификатор компьютера, который можно превратить в IP-адрес с помощью DNS.
- Доменные имена формируют иерархию – полное имя (FQDN) состоит из перечня имен всех вложенных доменов через точку, от листа к корню.
- Каждому домену могут соответствовать записи, его описывающие, в том числе IP-адрес.
- База DNS распределенная – она разделена на зоны (совокупность доменных имен определенного уровня, входящих в домен), за каждую зону отвечает один или несколько DNS-серверов.
- /etc/resolv.conf – параметры для DNS-клиентов (в основном, адрес сервера).
- BIND – популярная реализация полнофункционального DNS-сервера.

Dynamic Host Configuration Protocol (DHCP)

Протокол для автоматической настройки сетевых подключений.

- UDP порты 67 (сервер) и 68 (клиент).
- Позволяет выдавать клиентам динамические IP из заданного диапазона, или статические (по MAC или другому идентификатору).
- Кроме IP-адреса можно настраивать другие параметры – маску подсети, шлюз по умолчанию, широковещательный адрес, адреса DNS-серверов, имя машины, домен DNS и т.д.
- Наиболее распространена реализация DHCP-клиента и DHCP-сервера от Internet Software Consortium (ISC) – пакеты isc-dhcp-client и isc-dhcp-server.

Network Time Protocol (NTP)

Протокол для синхронизации системного времени.

- UDP порт 123.
- Иерархическая организация серверов в слои (stratum).
- Слой 0 – устройства с аппаратными часами высокой точности.
- Серверы слоя 1 – синхронизующиеся напрямую хотя бы с одним устройством слоя 0, серверы слоя 2 – хотя бы с одним сервером слоя 1, и т.д.
- Протокол предполагает периодические запросы для сравнения времени с сервером.
- systemd-timesyncd – встроенный в systemd клиент SNTP (упрощенная версия протокола NTP).
- ntpd, chronyd – полнофункциональные клиенты+серверы NTP.