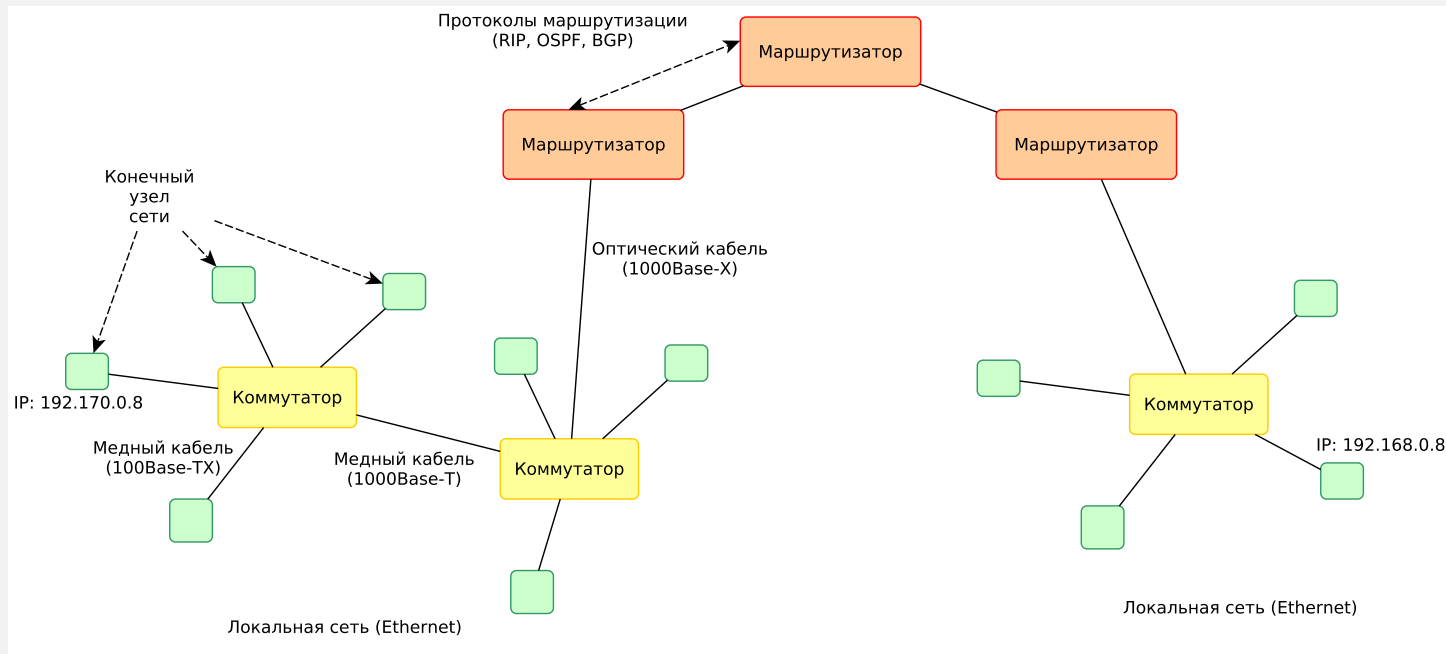


ПО сетевых устройств

Трещановский Павел Александрович, к.т.н.

23.05.19

Сетевой, канальный и физический уровни

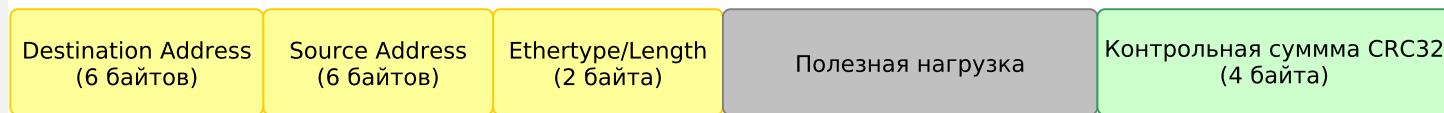


Локальные сети

- В локальных сетях поддерживается широковещательная (т.е. всем узлам сети) рассылка.
- Зачем нужна широковещательная рассылка? Для разрешения IP-адресов в MAC-адреса. Т.е. спрашиваем всех, где находится IP-адрес X.X.X.X, обладатель этого адреса отвечает и сообщает свой MAC-адрес.
- Если сеть с общей средой передачи, то все кадры широковещательные (например, WiFi).
- Почти все локальные сети - Ethernet или родственные (WiFi).
- Канальный уровень применительно к Ethernet также называется MAC - Media Access Control (управление доступом к среде передачи).
- Физические уровни (PHY) для Ethernet: 100Base-TX, 1000Base-T (медь), 100Base-FX, 1000Base-X (оптика).

Структура Ethernet-кадра

Ethernet-кадр

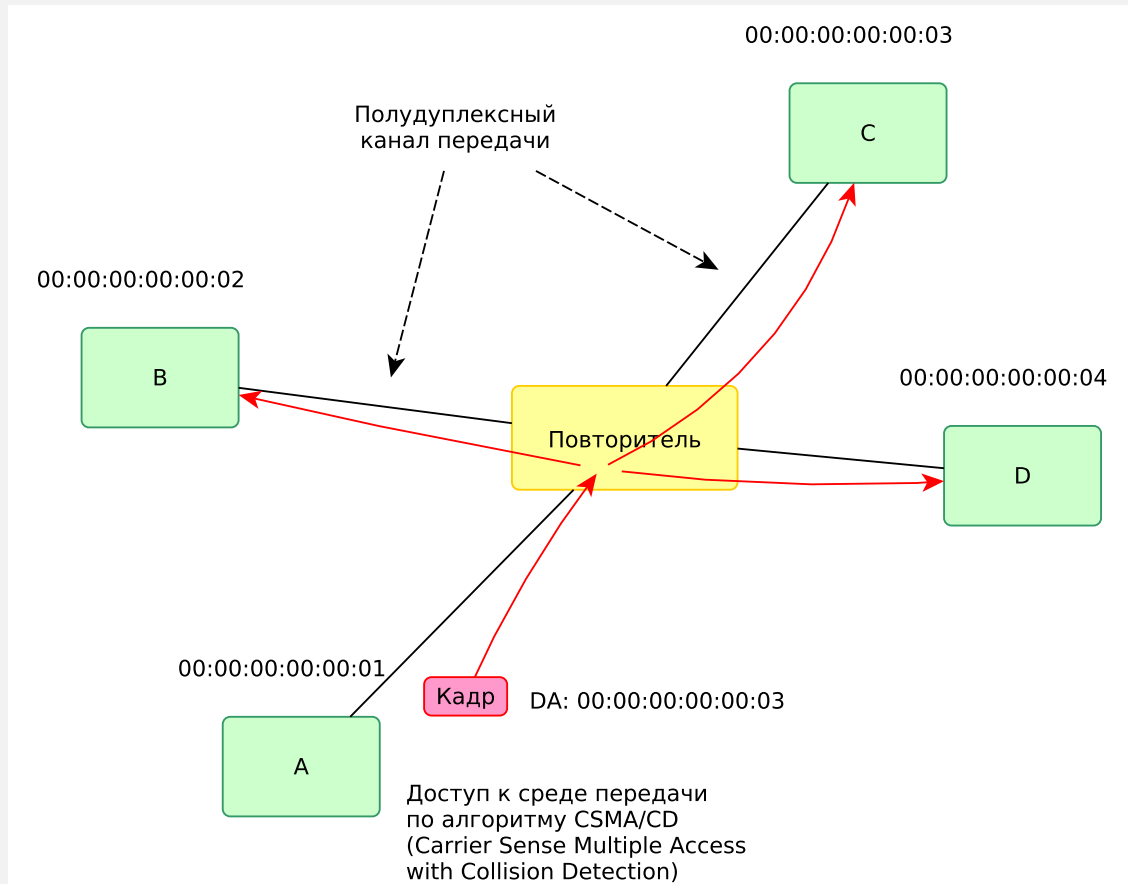


- Если значение Ethertype/Length меньше 0x800, то поле содержит длину кадра. В противном случае - тип кадра (IP, ARP и т.д.).
- Первые 3 байта MAC-адреса - OUI (Organizationally Unique Identifier), код производителя.
- CRC32 обычно устанавливается и проверяется на аппаратном уровне.
- Размер кадра - от 64 до 1522 байтов.

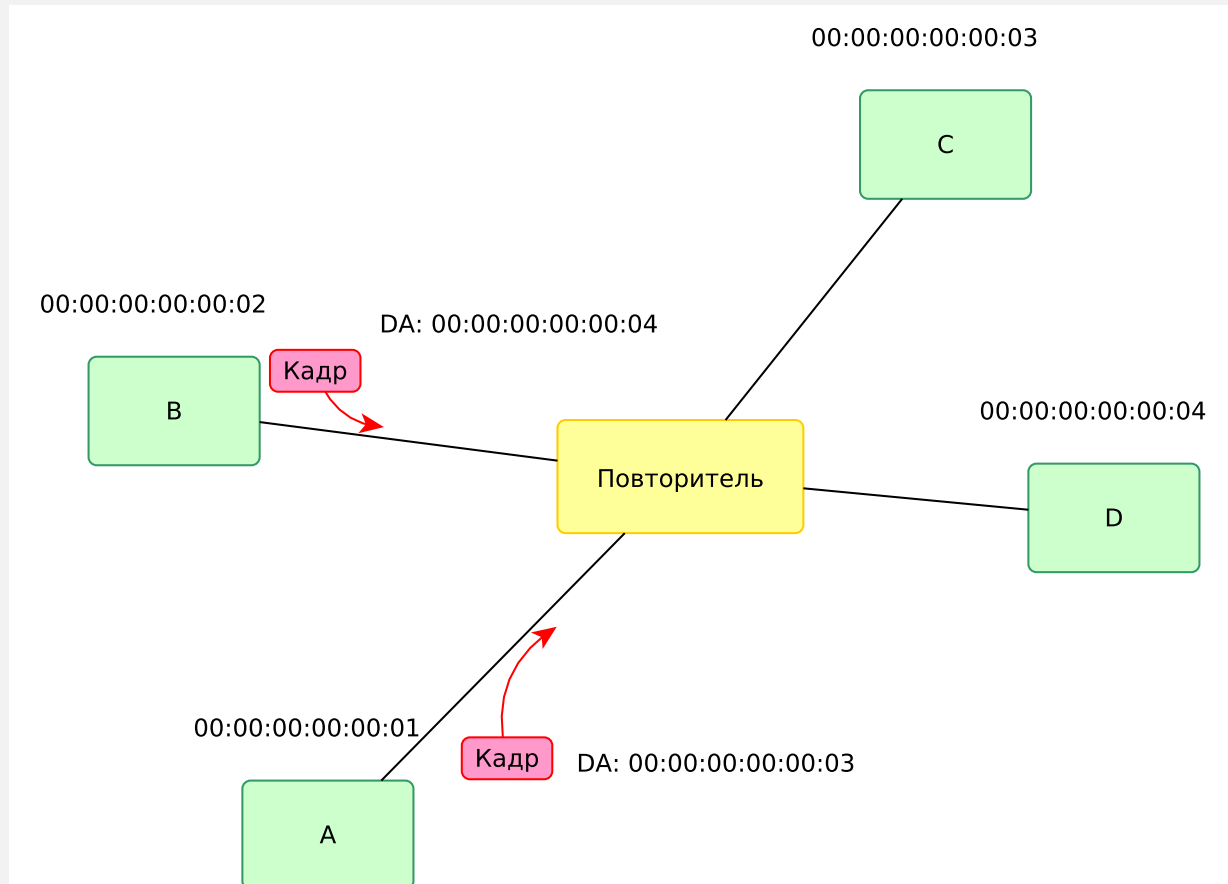
Передача Ethernet-кадра



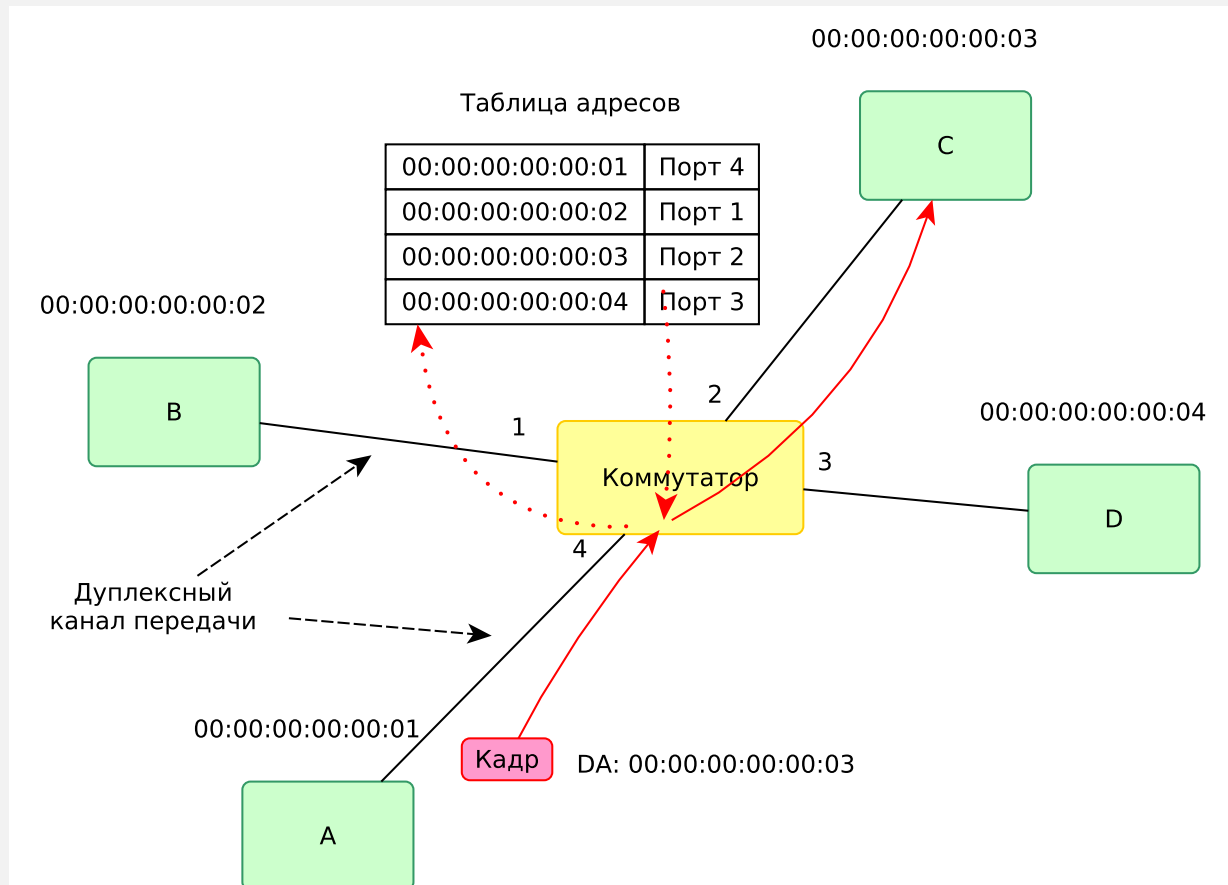
Сеть на основе повторителя



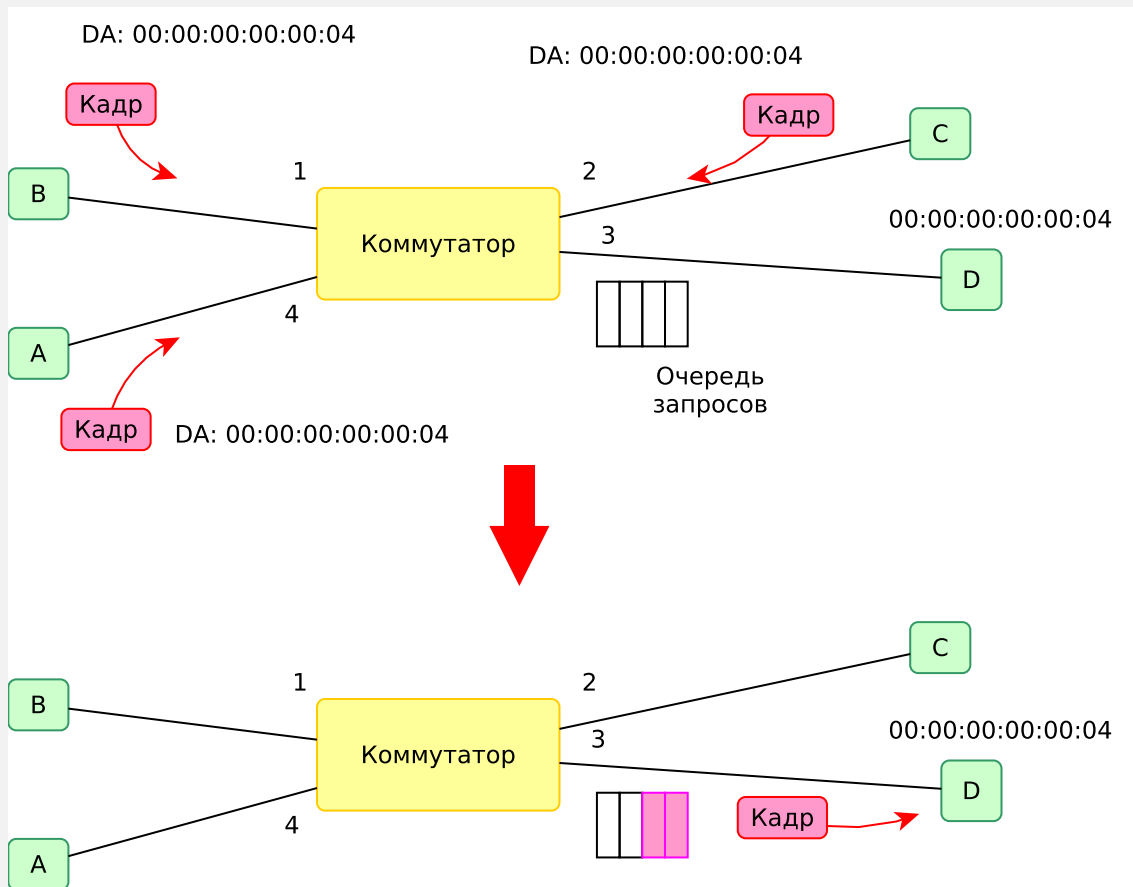
Коллизия в полудуплексных сетях



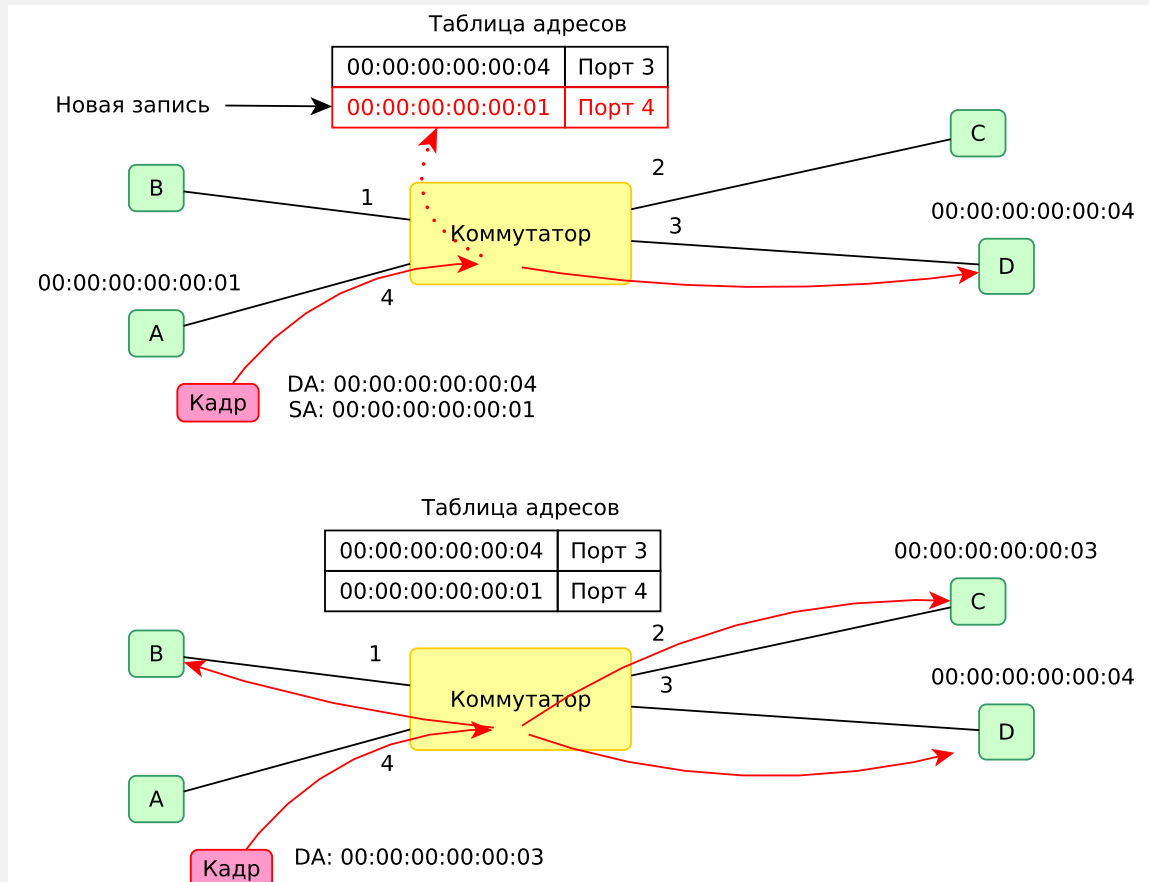
Сеть на основе коммутатора



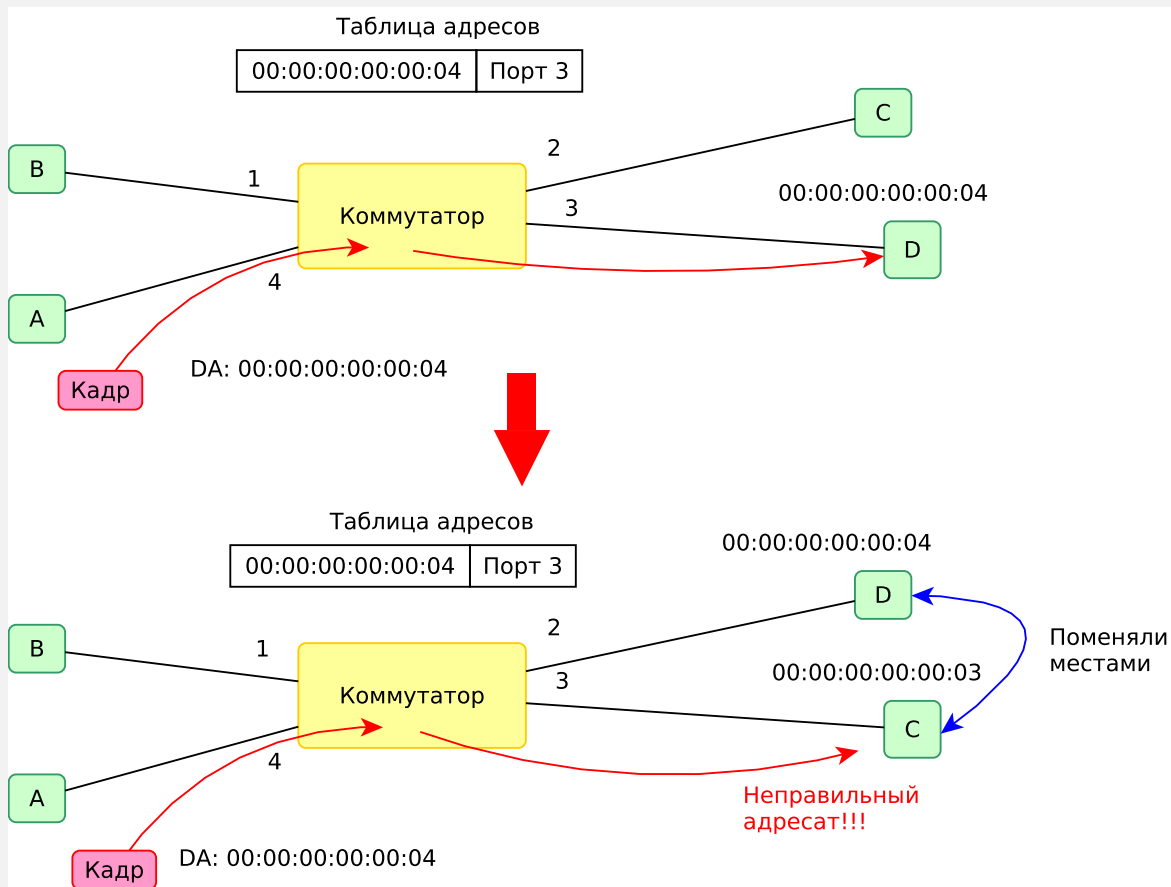
Буферизация кадров



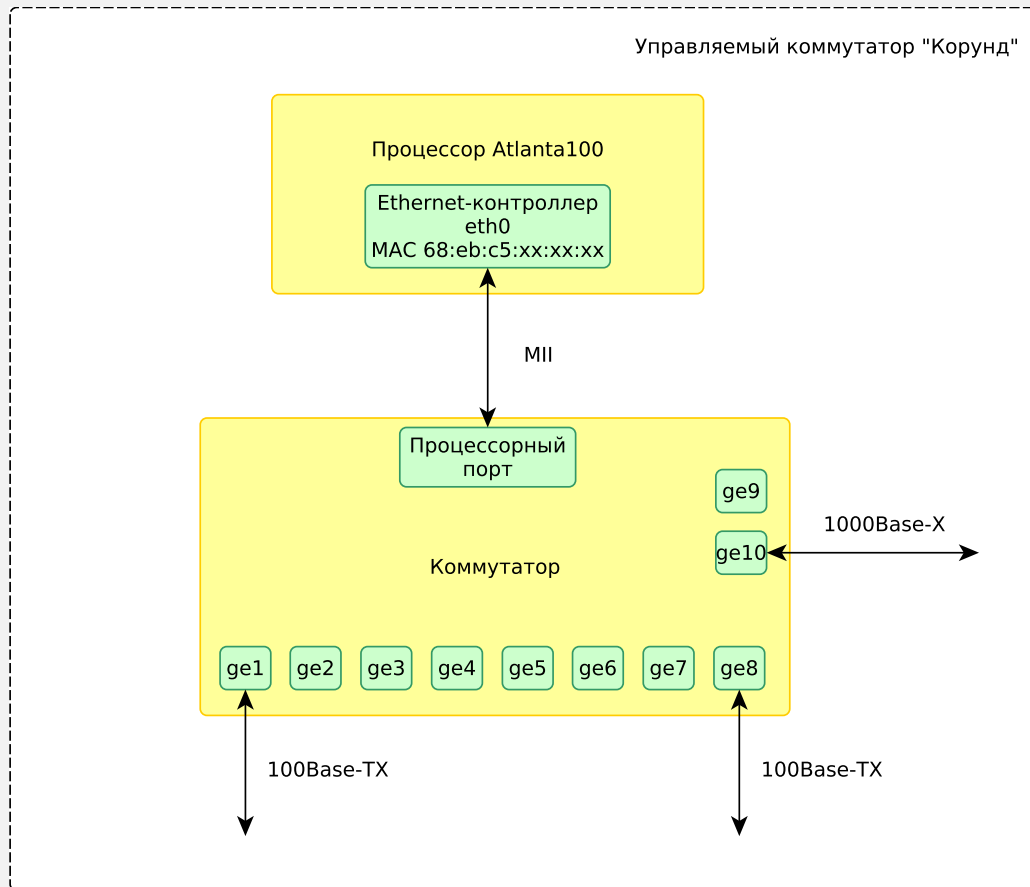
Изучение MAC-адресов



Устаревание MAC-адресов



Коммутатор Корунд



Сокеты семейства AF_PACKET

```
#include <linux/if_packet.h>
#include <net/ethernet.h>

int sock;
struct sockaddr_ll bind_addr, rx_addr;
char pkt_buf[2048];
struct ether_header *eth;
struct ether_addr *da;

sock = socket(AF_PACKET, SOCK_RAW, 0);

memset(&bind_addr, 0, sizeof(bind_addr));
bind_addr.sll_family    = AF_PACKET;
bind_addr.sll_protocol  = htons(ETH_P_ALL);
bind_addr.sll_ifindex   = ifindex; /* Индекс можно найти в выводе ip addr */
bind(sock, (struct sockaddr*)&bind_addr, sizeof(bind_addr));

recvfrom(sock, pkt_buf, sizeof(pkt_buf), 0, &rx_addr, sizeof(rx_addr));
if (rx_addr.sll_pkttype == PACKET_OTHERHOST) {
    eth = pkt_buf;
    da = (struct ether_addr *)eth->ether_dhost;
    printf("Destination address: %s\n", ether_ntoa(da));
}
```

Виртуализация сетевого стека

- Если в Linux настроено два интерфейса eth0 и eth1 192.168.0.8 и 192.168.0.9, то команда ping 192.168.0.9 не будет передавать запрос через eth0.
- Почему - общая таблица маршрутизации. Linux знает, что оба адреса локальные.
- Нужна виртуализация. В Linux'е виртуализация осуществляется с помощью контейнеров (aka LXC-контейнеры).
- Контейнер виртуализирует среду исполнения процесса, а не всей ОС.
- Контейнер состоит из пространств имен (namespace).
- Есть пространства имен для процессов, пользователей, сетевого стека.
- Пространства имен можно использовать по отдельности, без полной виртуализации.

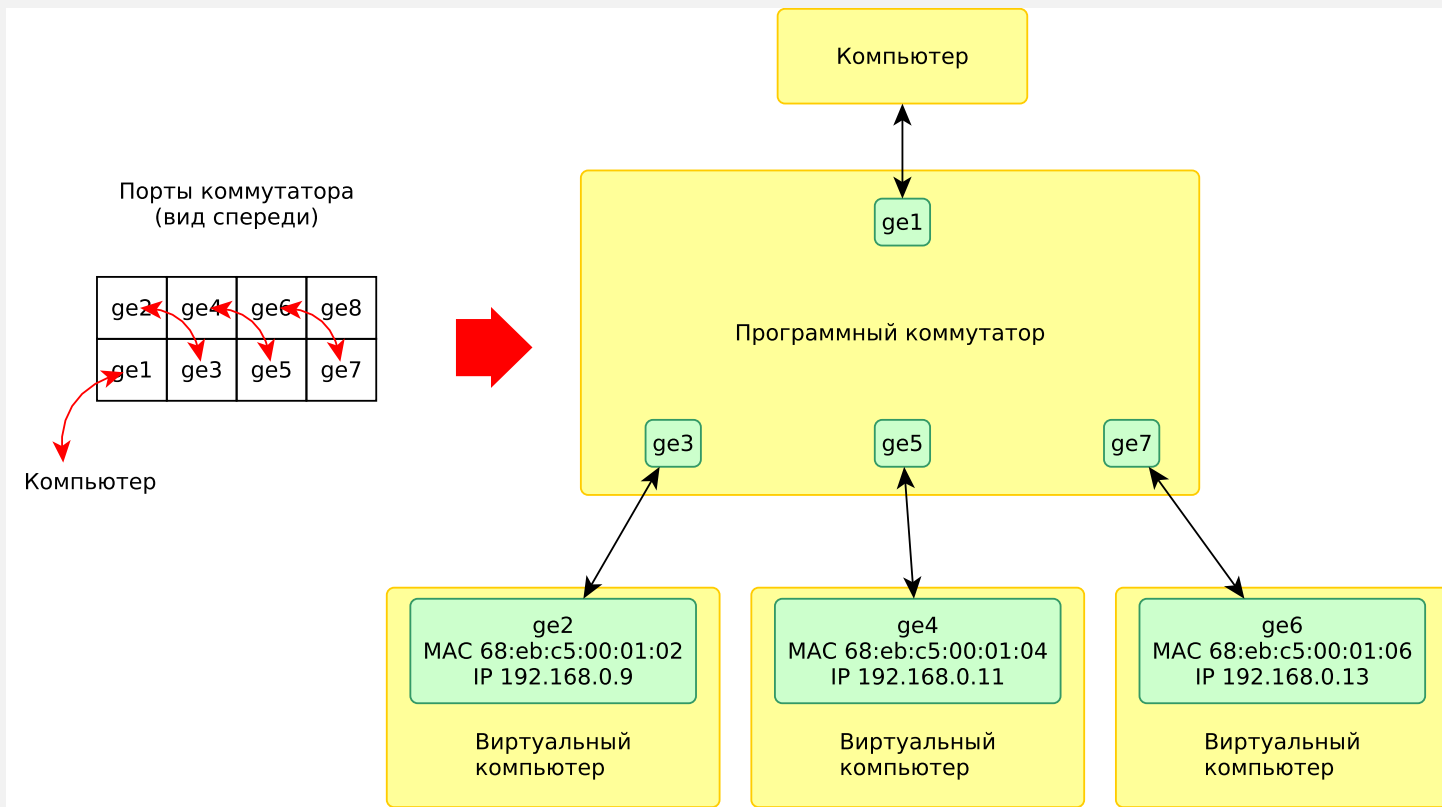
Запуск приложений в определенном пространстве имен

Вывод списка интерфейсов в сетевом пространстве имен /etc/ns/ns1:

```
# nsenter --net=/etc/ns/ns1 ip addr
1: lo: <LOOPBACK> mtu 65536 qdisc noop state DOWN qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
4: ge2@if2: <NO-CARRIER,BROADCAST,MULTICAST,UP,M-DOWN> ...
    link/ether 68:eb:c5:00:01:02 brd ff:ff:ff:ff:ff:ff
    inet 192.168.0.10/24 scope global ge2
        valid_lft forever preferred_lft forever
```

Замечание. Имена /etc/ns/nsX специфичны для Корунда. На других системах именование пространств имен будет отличаться.

Схема виртуальной локальной сети



Замечания

- Тестирование осуществляется с помощью утилиты ping.
- Аргументом ping является IP-адрес. Чтобы узнать MAC-адрес, Linux отправляет широковещательный ARP-запрос. Владелец указанного адреса возвращает ARP-ответ. Только после этого отправляется ICMP-запрос.
- Что должно быть обязательно проверено.
 - Если адрес не изучен, кадр отправляется на все порты.
 - Если изучен, то только на один.
 - Если поменять порт подключения виртуального компьютера, к которому периодически отправляются ICMP-запросы, то должна произойти временная потеря связи и последующим автоматическим восстановлением.