Системное программное обеспечение локальных компьютерных сетей

Стек протоколов TCP/IP

Денис Пынькин

2013 – 2014

e-mail: denis.pynkin@bsuir.by

http://goo.gl/32cTB

СЧАСТЬЕ ДЛЯ ВСЕХ, ДАРОМ, И ПУСТЬ НИКТО НЕ УЙДЕТ ОБИЖЕННЫЙ!

(c)Стругацкие, Пикник на обочине

Даты

1975

Первый тест между двумя системами

Даты

1975

Первый тест между двумя системами

1977

Тестовая сеть между тремя системами США, Великобритании и Норвегии

Даты

1975

Первый тест между двумя системами

1977

Тестовая сеть между тремя системами США, Великобритании и Норвегии

1978-1983

Тестируется 4 версии стека TCP/IP. И только в 3-й происходит разделение на протоколы TCP и IP.

Даты

1975

Первый тест между двумя системами

1977

Тестовая сеть между тремя системами США, Великобритании и Норвегии

1978-1983

Тестируется 4 версии стека TCP/IP. И только в 3-й происходит разделение на протоколы TCP и IP.

Март 1982

US DoD объявляет стек TCP/IP стандартом для военных сетей.

Даты

1975

Первый тест между двумя системами

1977

Тестовая сеть между тремя системами США, Великобритании и Норвегии

1978-1983

Тестируется 4 версии стека TCP/IP. И только в 3-й происходит разделение на протоколы TCP и IP.

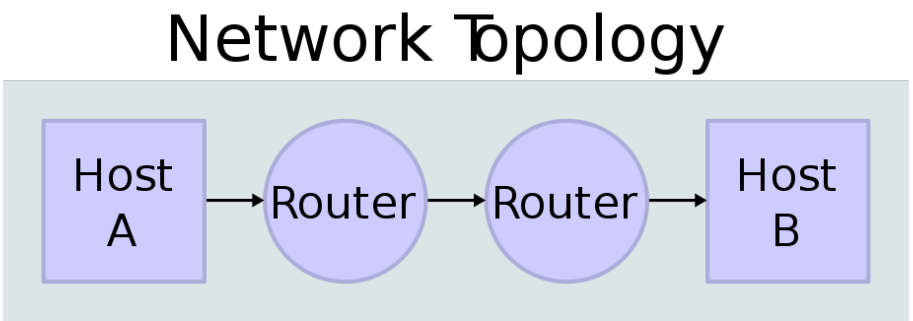
Март 1982

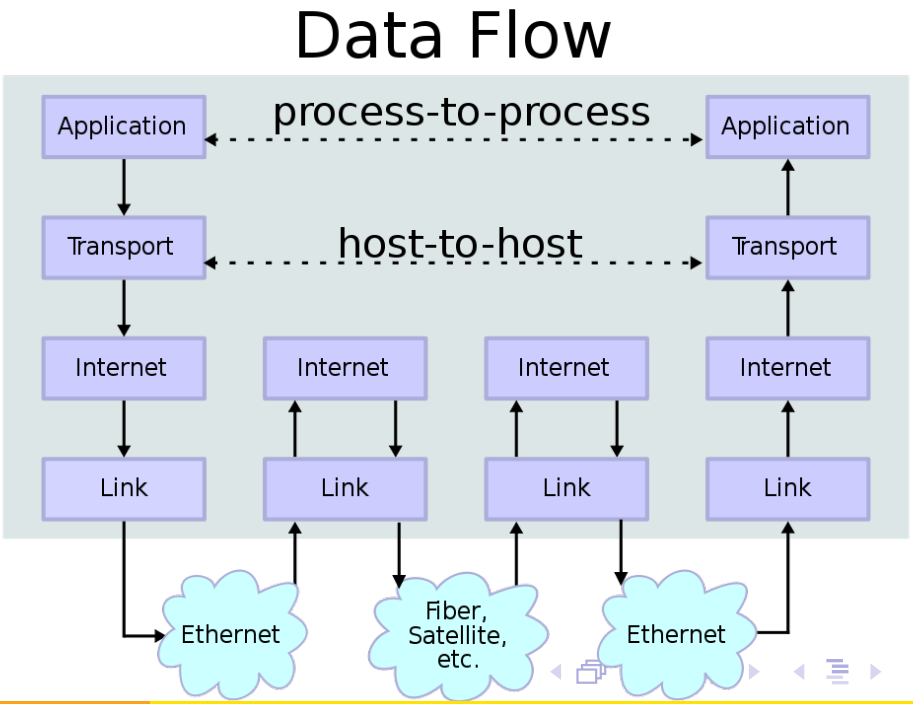
US DoD объявляет стек TCP/IP стандартом для военных сетей.

01 января, 1983

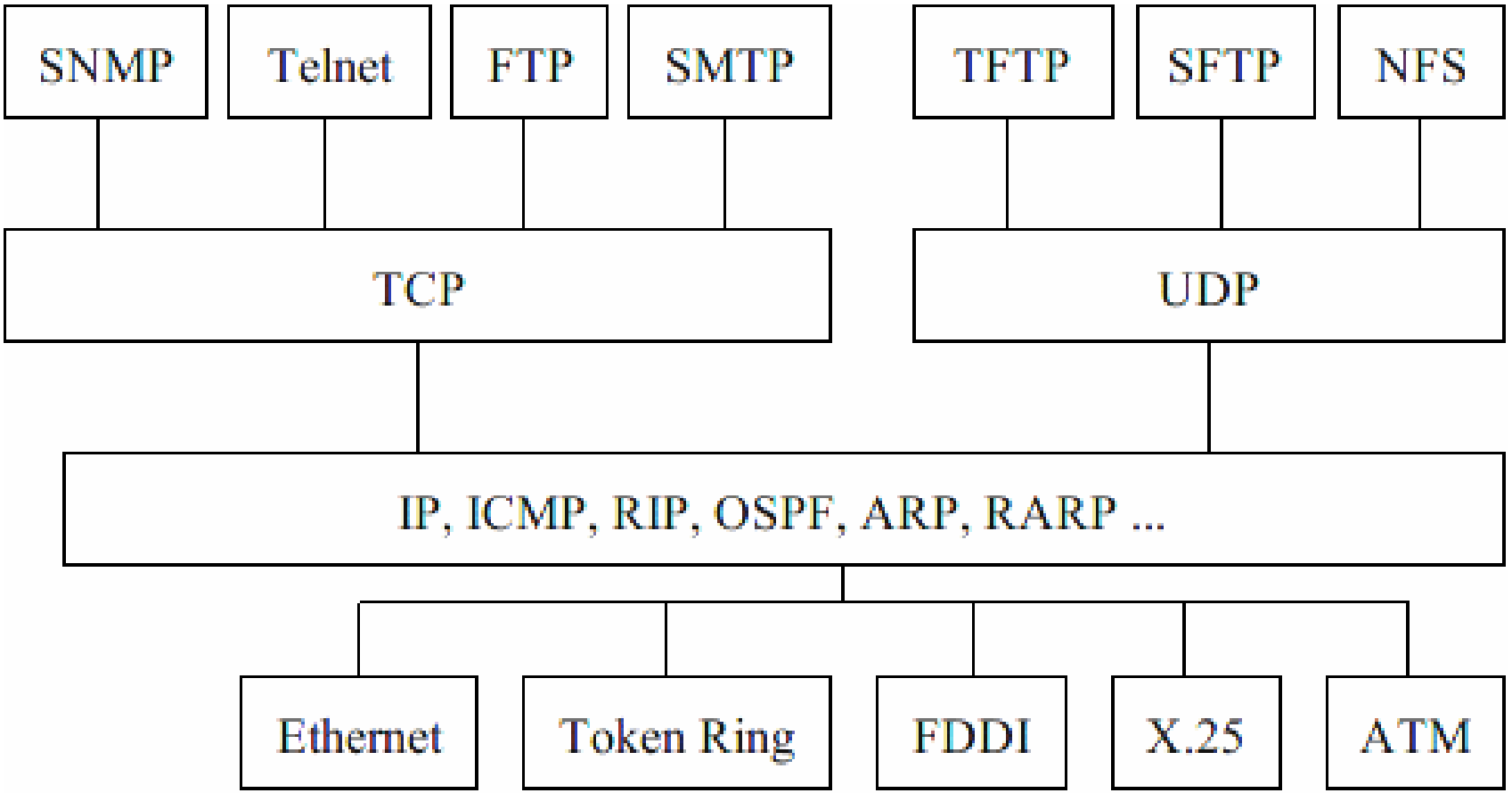
Топология сети

- Прикладной
- Транспортный
- Межсетовой
- От хоста к сети





Стек TCP/IP


SNMP
Telnet
FTP
SMTP
TFTP
SFTP
NFS
TCP
UDP
IP, ICMP, RIP, OSPF, ARP, RARP ...
Ethernet
Token Ring
FDDI
X.25
ATM

Межсетевой протокол IP (RFC 791)

- реализует обмен информации пакетами (IP-сегментами) (максимальный размер – 65535 байт);
- является протоколом взаимодействия без установления логического соединения;
- для адресации узлов сети используется адрес длиной 4 байта;
- обеспечивает в случае необходимости фрагментацию IP-сегментов;
- IP-сегменты имеют конечное время жизни в сети;
- не гарантирует надежность доставки IP-сегментов адресату;
- не имеет средств управления интенсивностью передачи IP-сегментов посылающей стороной (flow control);
- не гарантирует правильную последовательность IP-сегментов на принимающей стороне.

Заголовок IP-сегмента

Заголовок IP-сегмента: тип сервиса

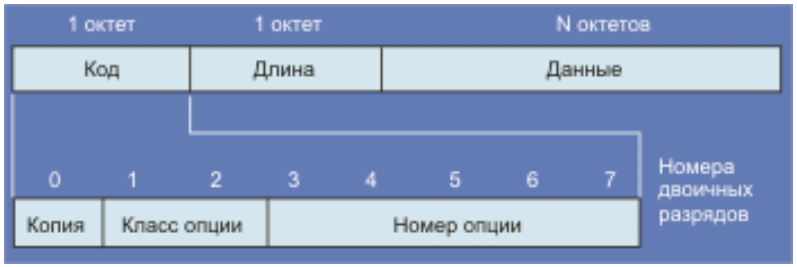
Приоритет:

- 0 Обычный уровень
- 1 Приоритетный
- 2 Немедленный
- 3 Срочный
- 4 Экстренный
- 5 ceitic/ecp
- 6 Межсетевое управление
- 7 Сетевое управление

Формат поля TOS определен в документе RFC-1349.
D=1 требует минимальной задержки,
T=1 - высокую пропускную способность,
R=1 - высокую надежность,
C=1 - низкую стоимость.

в настоящее время определяется RFC2474 как «Differentiated Services»

Заголовок IP-сегмента: опции



- Конец списка опций. Используется, если опции не укладываются в поле заголовка (смотри также поле "заполнитель")
- Никаких операций (используется для выравнивания октетов в списке опций)
- Ограничения, связанные с секретностью (для военных приложений)
- Свободная маршрутизация. Используется для того, чтобы направить дейтограмму по заданному маршруту
- Запись маршрута. Используется для трассировки
- Идентификатор потока. Устарело.
- Жесткая маршрутизация. Используется, чтобы направить дейтограмму по заданному маршруту
- Временная метка Интернет

RFC 1071 – вычисление контрольной суммы

```
1  register long sum = 0;
2  while( count > 1 ){
3  /*  This is the inner loop */
4    sum += * (unsigned short) addr++;
5    count -= 2;
6  }
7
8  /*  Add left-over byte,  if any */
9  if( count > 0 )
10   sum += * (unsigned char *) addr;
11
12 /*  Fold 32-bit sum to 16 bits */
13 while (sum>>16)
14   sum = (sum & 0xffff) + (sum >> 16);
15
16 checksum = ~sum;
```

Адреса IPv4

IP-адрес представляет собой четырехбайтовое число, старшие (крайние левые) биты которого определяют класс IP-адреса.

IP address Classes

Class	# Network Bits	# Hosts Bits	Decimal Address Range	Subnet mask
Class A	8 bits	24 bits	1-126	255.0.0.0
Class B	16 bits	16 bits	128-191	255.255.0.0
Class C	24 bits	8 bits	192-223	255.255.255.0
Class D	Reserved for Multicasting		224-239	N/A
Class E	Reserved for R & D		240-255	N/A

Специальные IP адреса

CIDR address block	Description	Reference
0.0.0.0/8	Current network (only valid as source address)	RFC 1700
10.0.0.0/8	Private network	RFC 1918
127.0.0.0/8	Loopback	RFC 5735
169.254.0.0/16	Link-Local	RFC 3927
172.16.0.0/12	Private network	RFC 1918
192.0.0.0/24	Reserved (IANA)	RFC 5735
192.0.2.0/24	TEST-NET-1, Documentation and example code	RFC 5735
192.88.99.0/24	IPv6 to IPv4 relay	RFC 3068
192.168.0.0/16	Private network	RFC 1918
198.18.0.0/15	Network benchmark tests	RFC 2544
198.51.100.0/24	TEST-NET-2, Documentation and examples	RFC 5737
203.0.113.0/24	TEST-NET-3, Documentation and examples	RFC 5737
224.0.0.0/4	Multicasts (former Class D network)	RFC 3171
240.0.0.0/4	Reserved (former Class E network)	RFC 1700
255.255.255.255	Broadcast	RFC 919

Таблица: Специальные адреса

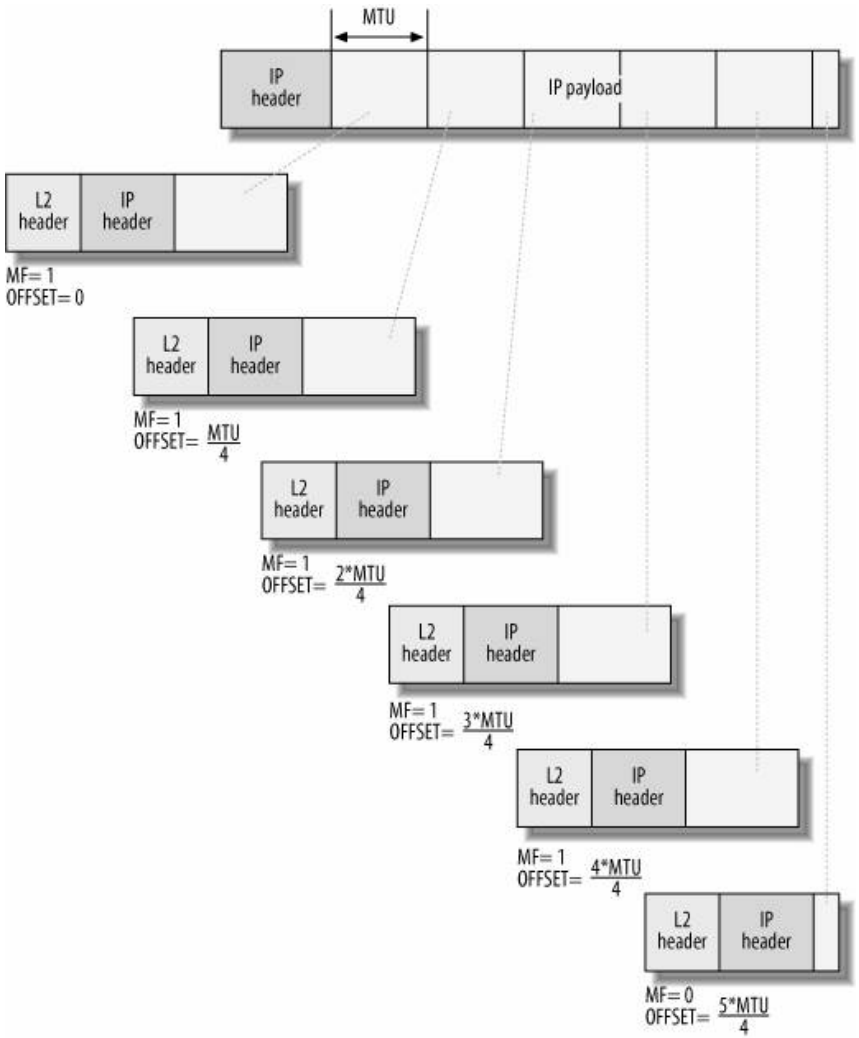
Фрагментация IP

Для того, чтобы существовала возможность передачи IP-сегментов через сети различного типа, межсетевой протокол обеспечивает адаптацию их размера к требованиям каждой сети. Это дает возможность, например, IP-сегментам, порожденным в сети на базе Ethernet (максимальный размер кадра – 1526 байт), беспрепятственно перемещаться до адресата по сети на базе X.25 (максимальный размер кадра - 128 байт).
Изменение размера IP-сегмента в процессе перемещения по сети может быть связано и с соображениями эффективности передачи.

Фрагментация IP

Каждый IP-фрагмент представляет собой полноценный IP-сегмент со своим собственным IP-заголовком. Однако заголовки всех IP-фрагментов содержат одинаковый идентификатор, совпадающий с идентификатором исходного IP-сегмента. Это позволяет распознавать все IP-фрагменты, относящиеся к одному исходному IP-сегменту.

Фрагментация IP



Фрагментация IP

IP-модуль на принимающем IP-фрагменты узле в ситуации, когда он должен транслировать IP-сегмент далее по сети, имеет три варианта действий с фрагментами:

- переслать IP-фрагменты далее неизменными;
- разбить (если в этом есть необходимость) полученные IP-фрагменты на более короткие IP-фрагменты;
- восстановить исходный IP-сегмент из фрагментов.

В работе с IP-фрагментами на принимающей стороне используется специальный таймер, который с приходом первого фрагмента IP-сегмента устанавливается в исходное состояние (для UNIX-реализаций это, обычно, 30 сек) и начинает обратный счет.

Протокол ICMP

Протокол передачи команд и сообщений об ошибках (ICMP – internet control message protocol, RFC-792)

Протокол ICMP

Протокол передачи команд и сообщений об ошибках (ICMP – internet control message protocol, RFC-792)

ICMP позволяет маршрутизатору либо конечному узлу сообщить узлу-отправителю об ошибках, с которыми машрутизатор столкнулся при передаче какого-либо IP-пакета от данного конечного узла.

ICMP только для конечных узлов

Управляющие сообщения ICMP не могут направляться промежуточному маршрутизатору, который участвовал в передаче пакета, с которым возникли проблемы, так как для такой посылки нет адресной информации - пакет несет в себе только адрес источника и адрес назначения, не фиксируя адреса промежуточных маршрутизаторов.

Спасение утопающих – дело рук самих утопающих!

Протокол ICMP - это протокол сообщения об ошибках, а не протокол коррекции ошибок.

Конечный узел может предпринять некоторые действия для того, чтобы ошибка больше не возникала, но эти действия протоколом ICMP не регламентируются.

ICMP и IP

ICMP-протокол сообщает об ошибках в IP-дейтограммах, но не дает информации об ошибках в самих ICMP-сообщениях.

ICMP и IP

ICMP-протокол сообщает об ошибках в IP-дейтограммах, но не дает информации об ошибках в самих ICMP-сообщениях.

ICMP использует IP, а IP-протокол должен использовать ICMP.

ICMP и IP

ICMP-протокол сообщает об ошибках в IP-дейтограммах, но не дает информации об ошибках в самих ICMP-сообщениях.

ICMP использует IP, а IP-протокол должен использовать ICMP.

В случае ICMP-фрагментации сообщение об ошибке будет выдано только один раз на дейтограмму, даже если ошибки были в нескольких фрагментах.

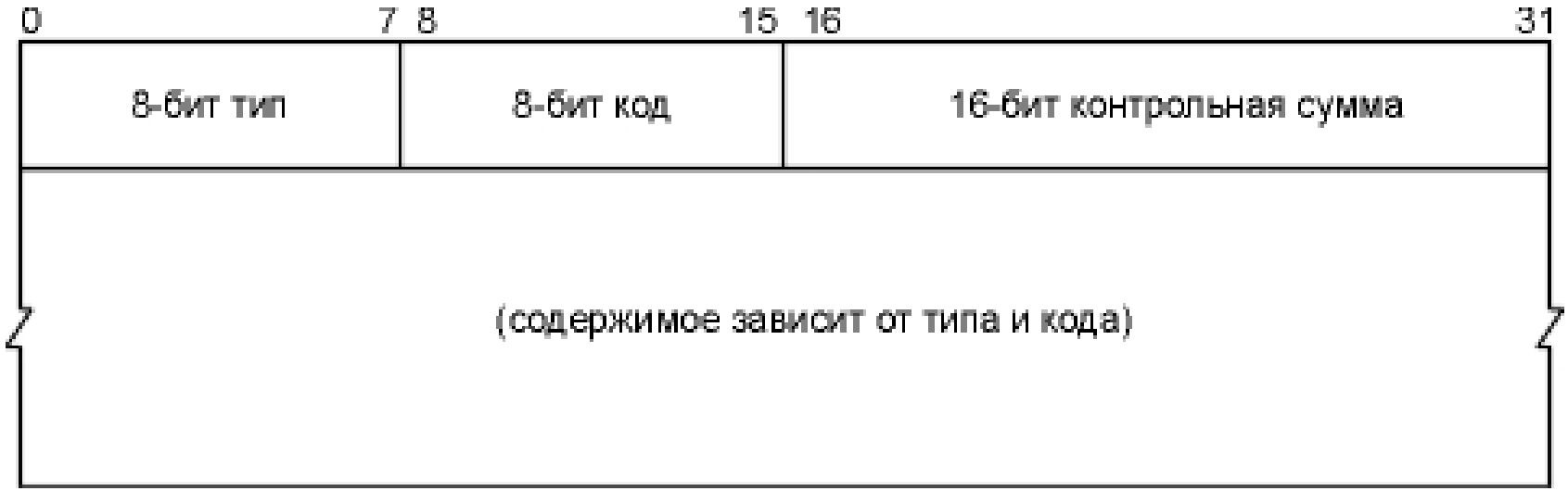
Задачи ICMP

- передача отклика на пакет или эхо на отклик;
- контроль времени жизни дейтограмм в системе;
- реализует переадресацию пакета;
- выдает сообщения о недостижимости адресата или о некорректности параметров;
- формирует и пересылает временные метки;
- выдает запросы и отклики для адресных масок и другой информации.

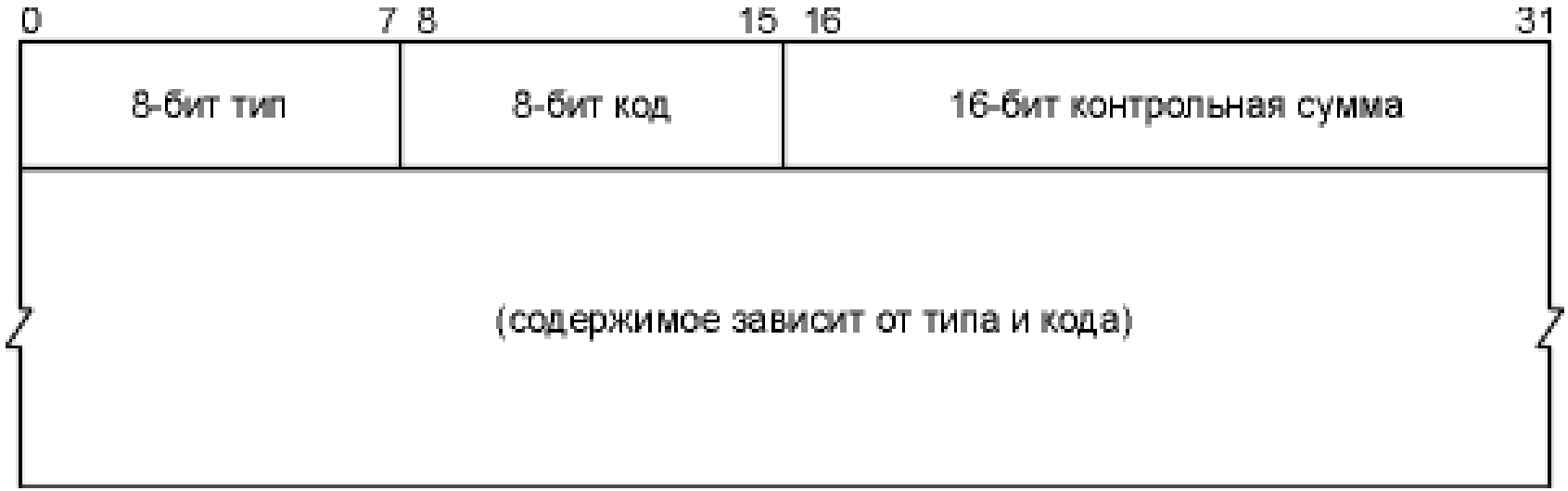
ICMP-сообщения об ошибках не выдаются:

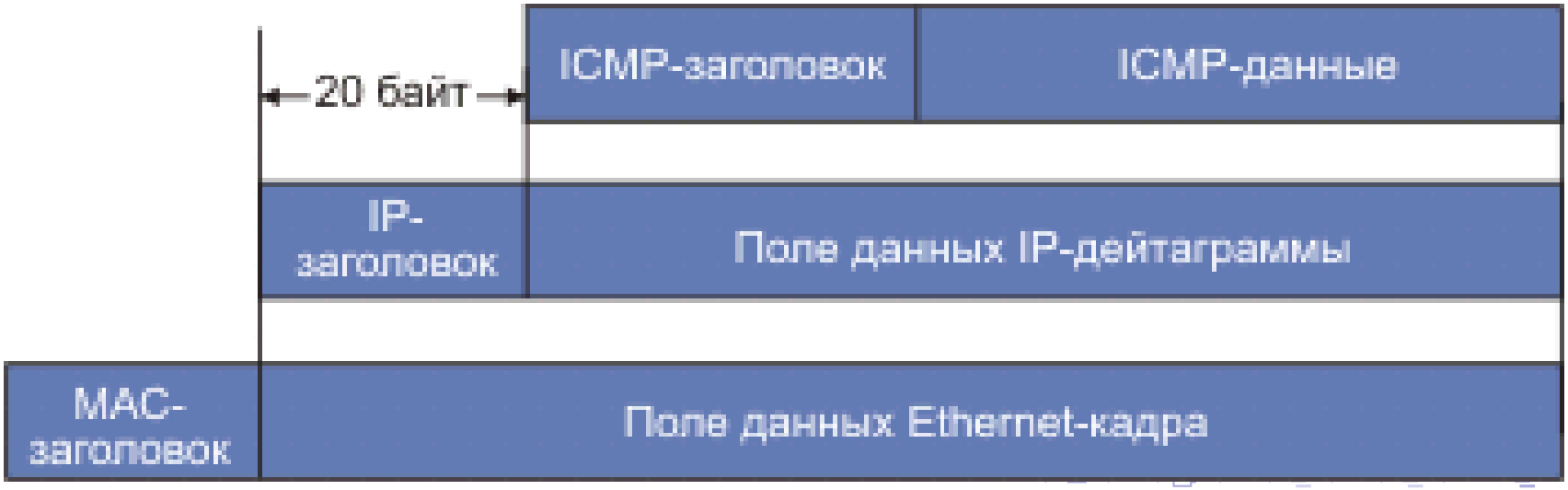
- на ICMP-сообщение об ошибке;
- При мультикастинг или широковещательной адресации;
- Для фрагмента дейтограммы (кроме первого);
- Для дейтограмм, чей адрес отправителя является нулевым, широковещательным или мультикастинговым.

Заголовок ICMP



Заголовок ICMP





Тип 0 – Эхо-ответ (ping-отклик)

Код	Описание
0	Эхо-ответ (ping-отклик)

Тип 3 – Адресат недоступен

Код	Описание
0	Сеть недостижима
1	ЭВМ не достижима
2	Протокол не доступен
3	Порт не доступен
4	Необходима фрагментация сообщения
5	Исходный маршрут вышел из строя
6	Сеть места назначения не известна
7	ЭВМ места назначения не известна
8	Исходная ЭВМ изолирована
9	Связь с сетью места назначения административно запрещена
10	Связь с ЭВМ места назначения административно запрещена
11	Сеть не доступна для данного вида сервиса
12	ЭВМ не доступна для данного вида сервиса
13	Связь административно запрещена с помощью фильтра.
14	Нарушение старшинства ЭВМ
15	Дискриминация по старшинству

Тип 5 - Переадресовать (изменить маршрут)

Код	Описание
0	Переадресовать дейтаграмму в сеть (устарело)
1	Переадресовать дейтаграмму на ЭВМ
2	Переадресовать дейтаграмму для типа сервиса (tos) и сети
3	Переадресовать дейтаграмму для типа сервиса и ЭВМ

Тип 8 – Эхо-запрос

Код	Описание
0	Эхо запрос (ping-запрос)

Тип 11 – Превышение временного интервала (для дейтаграммы время жизни истекло) (ttl=0)

Код	Описание
0	при передаче
1	при сборке (случай фрагментации)

Тип 12 – Проблема с параметрами дейтаграммы

Код	Описание
0	Ошибка в ip-заголовке
1	Отсутствует необходимая опция

Там еще много подозрительных типов!

Тип	Описание
4	Сдерживание источника (отключение источника при переполнении очереди)
9	Объявление маршрутизатора
10	Запрос маршрутизатора
13	Запрос метки времени
14	Ответ с меткой времени
15	Информационный запрос
16	Информационный ответ
17	Запрос адресной маски (RFC-950)
18	Отклик на запрос адресной маски (RFC-950)
30	Трассировка маршрута (RFC-1393)
31	Ошибка преобразования датаграммы (RFC-1475)
32	Перенаправление для мобильного узла
33	IPv6 Where-Are-You (где вы находитесь)
34	IPv6 I-Am-Here (я здесь)
35	Запрос перенаправления для мобильного узла
36	Отклик на запрос перенаправления для мобильного узла
37	Запрос доменного имени (Domain Name Request)
38	Ответ на запрос доменного имени (Domain Name Reply)
39	SKIP

Спасибо за внимание!
Вопросы?