Функции протокола IP

Стандарт RFC-791

Функции протокола IP определены в стандарте RFC-791 следующим образом:

“Протокол IP обеспечивает передачу блоков данных, называемых дейтаграммами, от отправителя к получателям, где отправители и получатели являются компьютерами, идентифицируемыми адресами фиксированной длины (IP-адресами). Протокол IP обеспечивает при необходимости также фрагментацию и сборку дейтаграмм для передачи данных через сети с малым размером пакетов”.

Надежность протокола IP

- Протокол IP является ненадежным протоколом без установления соединения.
- Протокол IP не подтверждает доставку данных, не контролирует целостность полученных данных и не производит операцию квитирования (handshaking) - обмена служебными сообщениями, подтверждающими установку соединения с узлом назначения и его готовность к приему данных.
- Протокол IP обрабатывает каждую дейтаграмму как независимую единицу, не имеющую связи ни с какими другими дейтаграммами в Интернет.
- После того, как дейтаграмма отправляется в сеть, ее дальнейшее продвижение никак не контролируется отправителем (на уровне протокола IP).
- Если дейтаграмма не может быть доставлена, она уничтожается. Узел, уничтоживший дейтаграмму, может оправить по обратному адресу ICMP-сообщение о причине сбоя.
- Гарантию правильной передачи данных предоставляют протоколы вышестоящего уровня (например, протокол TCP), которые имеют для этого необходимые механизмы.

Работа протокола IP (прием)

- Одна из основных задач, решаемых протоколом IP, - маршрутизация дейтаграмм на основании адреса получателя.

Работа протокола IP на каком-либо узле сети при приеме дейтаграммы из сети:

- с одного из интерфейсов уровня доступа к среде передачи в модуль IP поступает дейтаграмма;
- Сетевой уровень анализирует заголовок дейтаграммы;

  если пунктом назначения дейтаграммы является данный компьютер:

  если дейтаграмма является фрагментом большей дейтаграммы, ожидаются остальные фрагменты для сбора исходной большой дейтаграммы;
- из дейтаграммы извлекаются данные и направляются на обработку одному из протоколов вышележащего уровня (указывается в заголовке дейтаграммы);
- если дейтаграмма не направлена на IP-адрес данного узла, то дальнейшие действия зависят от того, разрешена или запрещена ретрансляция “чужих” дейтаграмм;

  если ретрансляция разрешена, то определяются следующий узел сети, на который должна быть переправлена дейтаграмма для доставки ее по назначению, и интерфейс нижнего уровня; при необходимости может быть произведена фрагментация дейтаграммы;

  если же дейтаграмма ошибочна или по каким-либо причинам не может быть доставлена, она уничтожается;

  при этом, как правило, отправителю дейтаграммы отсылается ICMP-сообщение об ошибке.

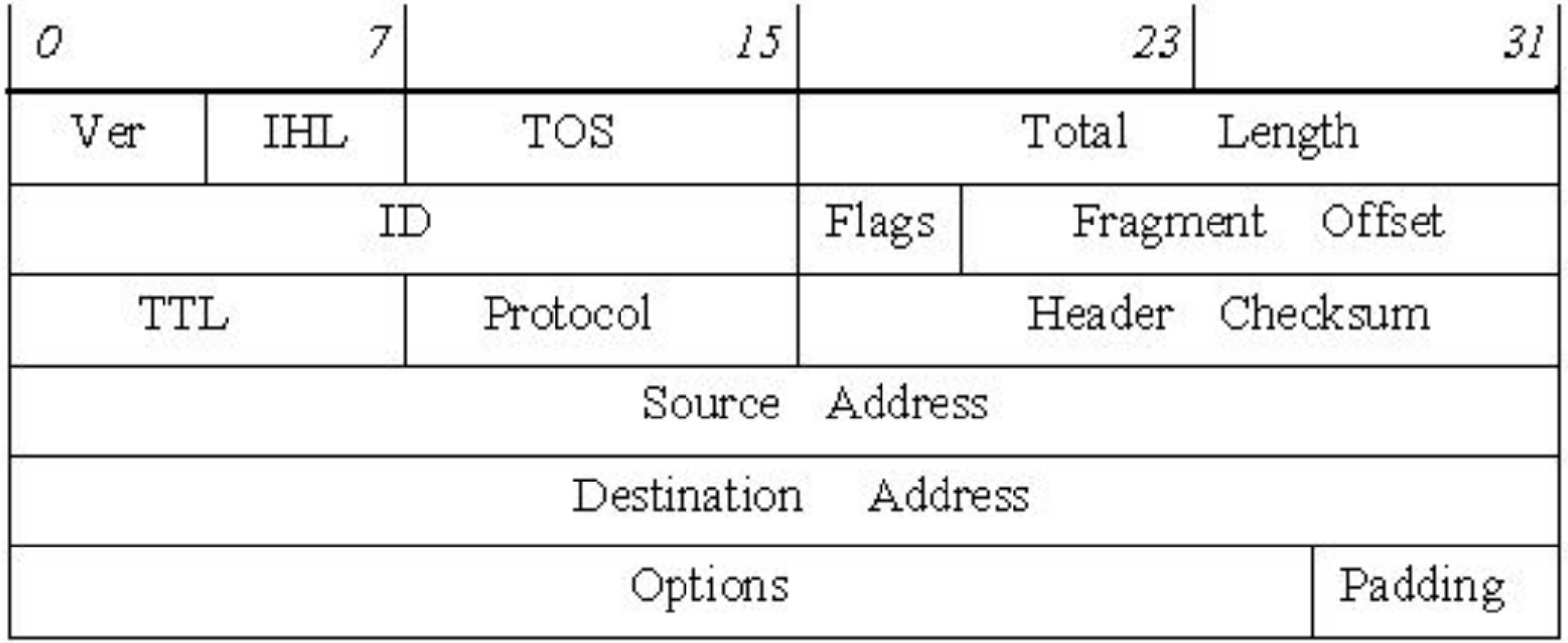
Работа протокола IP (передача)

При получении данных от вышестоящего уровня для передаче по сети IP-модуль формирует дейтаграмму с этими данными, в заголовок которой заносятся адреса отправителя и получателя (также полученные от транспортного уровня) и другая информация; после чего выполняются следующие шаги:

- Определяются узел сети, на который должна быть направлена дейтаграмма для доставки ее по назначению, и интерфейс нижнего уровня, после чего дейтаграмма передается на нижний уровень этому интерфейсу для отправки; при необходимости может быть произведена фрагментация дейтаграммы;
- если же дейтаграмма ошибочна или по каким-либо причинам не может быть доставлена, она уничтожается.

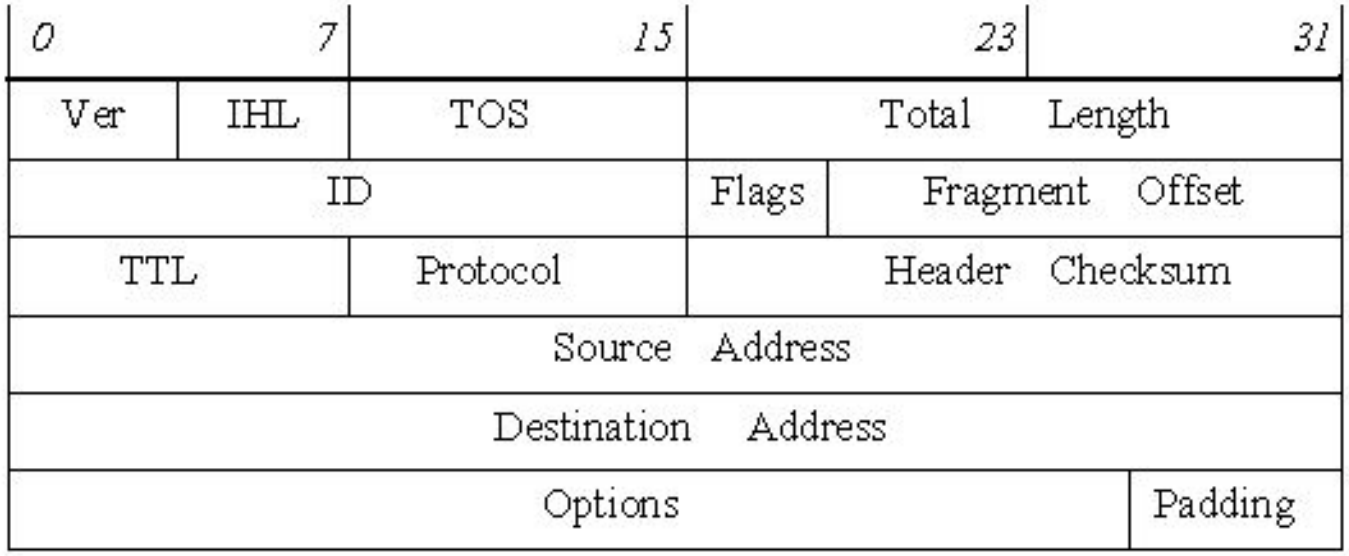
Формат заголовка IP-дейтаграммы

- IP-дейтаграмма состоит из заголовка и данных.
- Заголовок дейтаграммы состоит из 32-разрядных слов и имеет переменную длину, зависящую от размера поля “Options”, но всегда кратную 32 битам. За заголовком непосредственно следуют данные, передаваемые в дейтаграмме.



Поля дейтаграммы

- **Ver** (4 бита) - версия протокола IP, в настоящий момент используется версия 4, новые разработки имеют номера версий 6-8.
- **IHL (Internet Header Length)** (4 бита) - длина заголовка в 32-битных словах; диапазон допустимых значений от 5 (минимальная длина заголовка, поле “Options” отсутствует) до 15 (т.е. может быть максимум 40 байт опций).

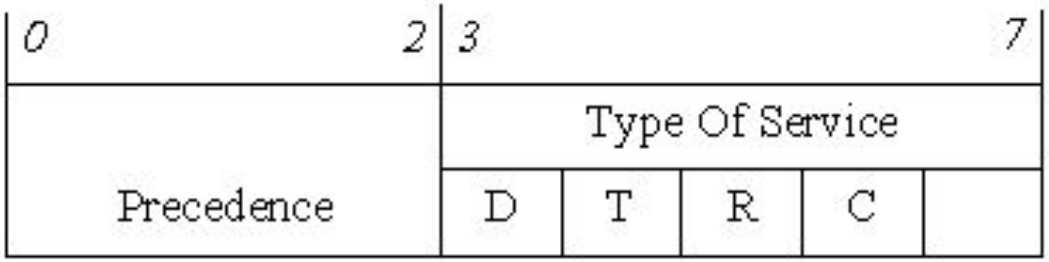


Поля дейтаграммы

TOS (Type Of Service) (8 бит) - значение поля определяет приоритет дейтаграммы и желаемый тип маршрутизации.

Три младших бита (“Precedence”) определяют приоритет дейтаграммы:

- 111 - управление сетью
- 110 - межсетевое управление
- 101 - CRITIC-ECP
- 100 - более чем мгновенно
- 011 - мгновенно
- 010 - немедленно
- 001 - срочно
- 000 - обычно



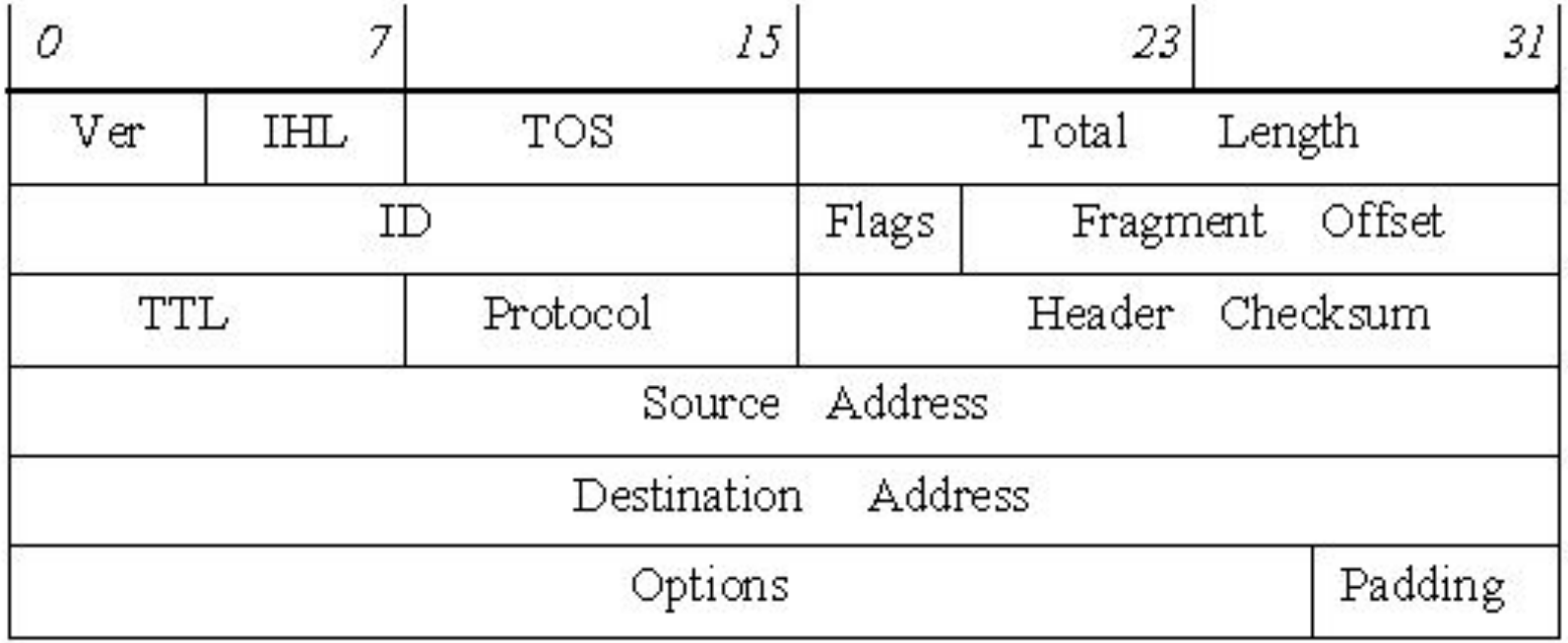
Биты D,T,R,C определяют желаемый тип маршрутизации:

- D (Delay) - выбор маршрута с минимальной задержкой,
- T (Throughput) - выбор маршрута с максимальной пропускной способностью,
- R (Reliability) - выбор маршрута с максимальной надежностью,
- C (Cost) - выбор маршрута с минимальной стоимостью.

В дейтаграмме может быть установлен только один из битов D,T,R,C. Старший бит байта не используется.

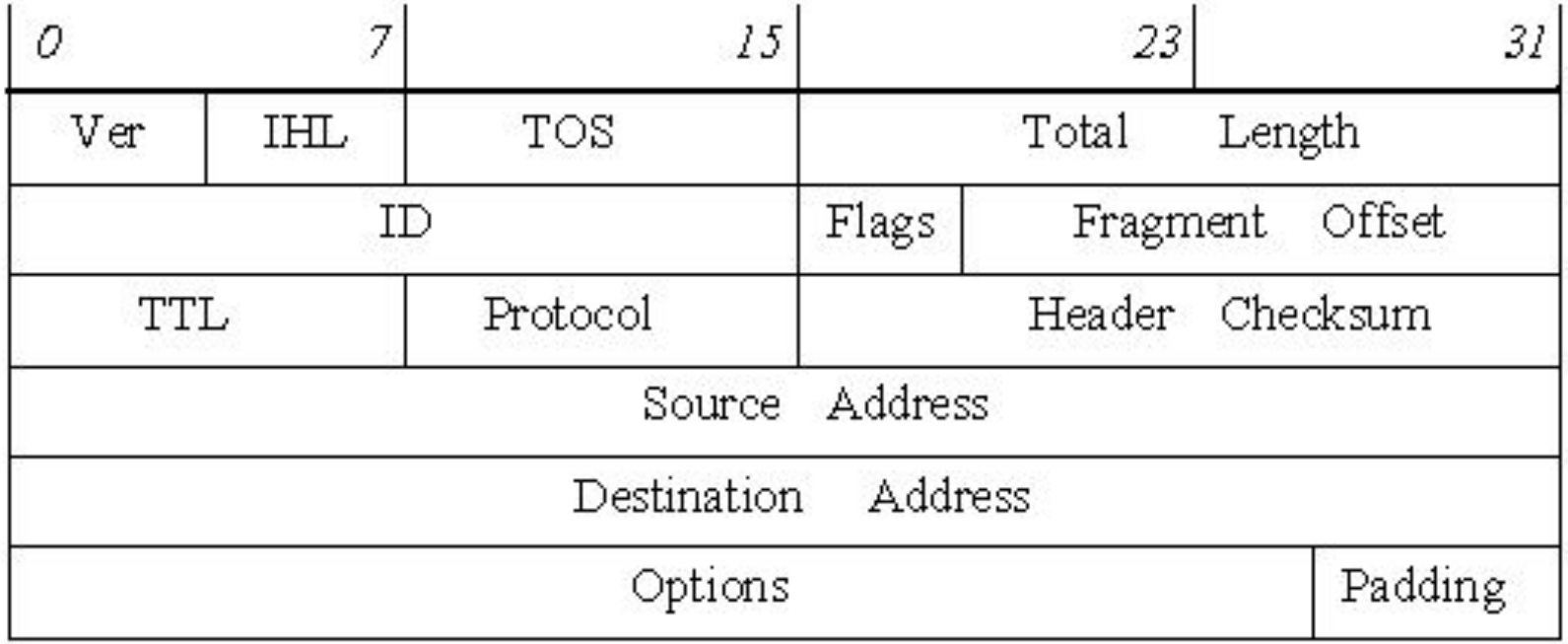
Поля дейтаграммы

- **Total Length** (16 бит) - длина всей дейтаграммы в октетах, включая заголовок и данные, максимальное значение 65535, минимальное - 21 (заголовок без опций и один октет в поле данных).
- **ID (Identification)** (16 бит), Flags (3 бита), Fragment Offset (13 бит) используются для фрагментации и сборки дейтаграмм



Поля дейтаграммы

- **TTL (Time To Live)** (8 бит) - “время жизни” дейтаграммы. Устанавливается отправителем, измеряется в секундах.
- Каждый маршрутизатор, через который проходит дейтаграмма, переписывает значение TTL, предварительно вычтя из него единицу.
- При достижении значения TTL=0 дейтаграмма уничтожается, при этом отправителю может быть послано соответствующее ICMP-сообщение.
- Контроль TTL предотвращает зацикливание дейтаграммы в сети.



Поля дейтаграммы

- **Protocol** (8 бит) - определяет программу (вышестоящий протокол стека), которой должны быть переданы данные дейтаграммы для дальнейшей обработки.

Код	Протокол	Описание
1	ICMP	Протокол контрольных сообщений
2	IGMP	Протокол управления группой хостов
3	IP	IP поверх IP (инкапсуляция)
4	TCP	TCP
5	EGP	Протокол внешней маршрутизации (устарел)
6	IGP	Протокол внутренней маршрутизации (устарел)
7	UDP	UDP
8	RSVP	Протокол резервирования ресурсов при мультикастинге
9	IGRP	Протокол внутренней маршрутизации от фирмы cisco
10	OSPF	Протокол внутренней маршрутизации

Поля дейтаграммы

- **Header Checksum** (16 бит) - контрольная сумма заголовка, представляет из себя 16 бит, дополняющие биты в сумме всех 16-битовых слов заголовка. Перед вычислением контрольной суммы значение поля “Header Checksum” обнуляется. Поскольку маршрутизаторы изменяют значения некоторых полей заголовка при обработке дейтаграммы (как минимум, поля “TTL”), контрольная сумма каждым маршрутизатором пересчитывается заново. Если при проверке контрольной суммы обнаруживается ошибка, дейтаграмма уничтожается.

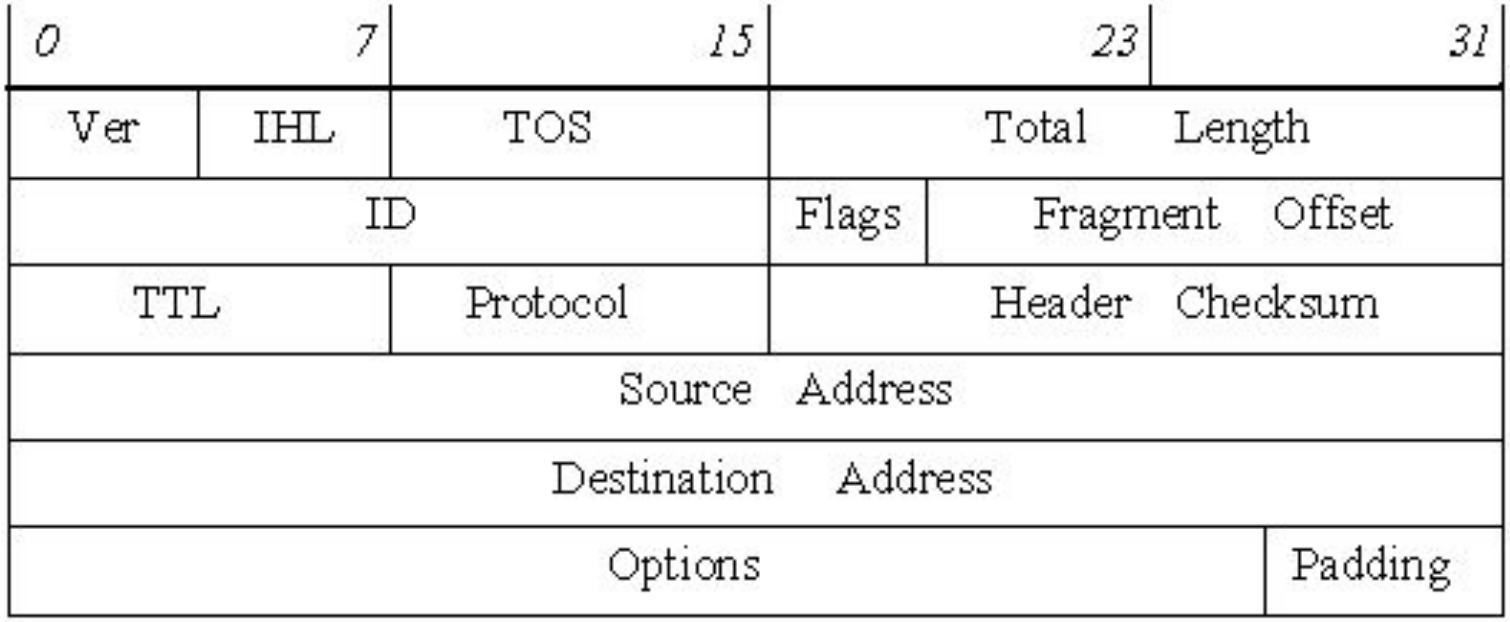
- **Source Address** (32 бита) - IP-адрес отправителя.
- **Destination Address** (32 бита) - IP-адрес получателя.

- **Options** - опции, поле переменной длины. Опций может быть одна, несколько или ни одной. Опции определяют дополнительные услуги модуля IP по обработке дейтаграммы, в заголовок которой они включены. Подробнее опции рассматриваются в пп. 2.4.3, 2.4.4.

- **Padding** - выравнивание заголовка по границе 32-битного слова, если список опций занимает нецелое число 32-битных слов. Поле “Padding” заполняется нулями

Фрагментация дейтаграмм

- Различные среды передачи имеют различный максимальный размер передаваемого блока данных (MTU - Media Transmission Unit), это число зависит от скоростных характеристик среды и вероятности возникновения ошибки при передаче.
- При передаче дейтаграммы из среды с большим MTU в среду с меньшим MTU может возникнуть необходимость во фрагментации дейтаграммы. Фрагментация и сборка дейтаграмм осуществляются модулем протокола IP. Для этого применяются поля “ID” (Identification), “Flags” и “Fragment Offset” .



Фрагментация дейтаграмм

Flags -поле состоит из 3 бит, младший из которых всегда 0.

Значения бита DF (Don’t Fragment):

0 - фрагментация разрешена,
1 - фрагментация запрещена (если дейтаграмму нельзя передать без фрагментации, она уничтожается).

Значения бита MF (More Fragments):

0 - данный фрагмент последний (единственный),
1 - данный фрагмент не последний.

0 DF MF

Фрагментация дейтаграмм

- **ID (Identification)** - идентификатор дейтаграммы, устанавливается отправителем; используется для сборки дейтаграммы из фрагментов для определения принадлежности фрагментов одной дейтаграмме.

- **Fragment Offset** - смещение фрагмента, значение поля указывает, на какой позиции в поле данных исходной дейтаграммы находится данный фрагмент. Смещение считается 64-битовыми порциями, т.е. минимальный размер фрагмента равен 8 октетам, а следующий фрагмент в этом случае будет иметь смещение 1. Первый фрагмент имеет смещение нуль.

Протокол ICMP

- Протокол ICMP (Internet Control Message Protocol, Протокол Управляющих Сообщений Интернет) является неотъемлемой частью IP-модуля. Он обеспечивает обратную связь в виде диагностических сообщений, посылаемых отправителю при невозможности доставки его дейтаграммы и в других случаях. ICMP стандартизован в RFC-792, дополнения — в RCF-950,1256.

Функции протокола ICMP

ICMP-протокол осуществляет:

- передачу 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 никогда не выдаются в ответ на:

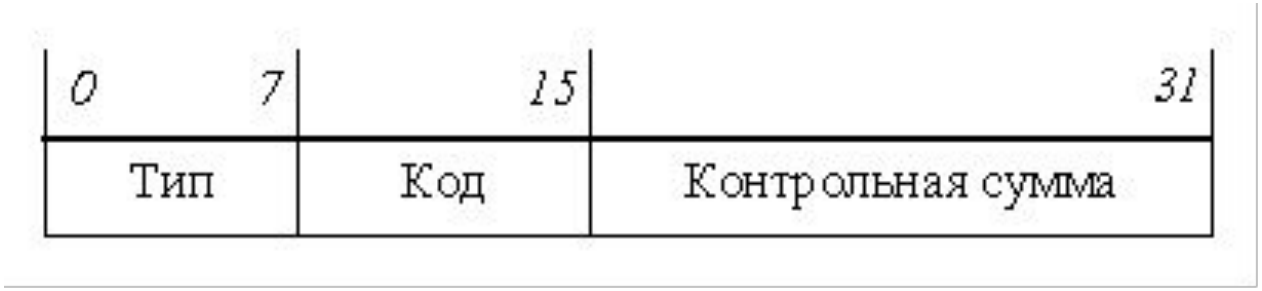
- Дейтаграммы, содержащие ICMP-сообщения.
- При мультикастинг или широковещательной адресации.
- Для фрагмента дейтограммы (кроме первого).
- Для дейтограмм, чей адрес отправителя является нулевым, широковещательным или мультикастинговым.

Формат ICMP-сообщения

После IP-заголовка следует 32-битное слово с полями “Тип”, “Код” и “Контрольная сумма”. Поля типа и кода определяют содержание ICMP-сообщения.

Формат остальной части дейтаграммы зависит от вида сообщения.

Контрольная сумма считается так же, как и в IP-заголовке, но в этом случае суммируется содержимое ICMP-сообщения, включая поля “Тип” и “Код”.



Виды ICMP сообщений

Тип	Код	Сообщение
0	0	Echo Reply (эхо-ответ)
3		Destination Unreachable (адресат недостижим по различным причинам):
	0	Net Unreachable (сеть недоступна)
	1	Host Unreachable (хост недоступен)
	2	Protocol Unreachable (протокол недоступен)
	3	Port Unreachable (порт недоступен)
	4	DF=1 (необходима фрагментация, но она запрещена)
	5	Source Route failed (невозможно выполнить опцию Source Route)
4	0	Source Quench (замедление источника)
5		Redirect (выбрать другой маршрутизатор для посылки дейтаграмм)
	0	в данную сеть
	1	на данный хост
	2	в данную сеть с данным TOS
	3	на данный хост с данным TOS

Виды ICMP сообщений

Тип	Код	Сообщение
8	0	Echo Request (эхо-запрос)
9	0	Router Advertisement (объявление маршрутизатора)
10	0	Router Solicitation (запрос объявления маршрутизатора)
11		Time Exceeded (время жизни дейтаграммы истекло)
	0	при передаче
	1	при сборке
12		Parameter problem (ошибка в параметрах)
	0	Ошибка в IP-заголовке
	1	Отсутствует необходимая опция
13	0	Timestamp (запрос временной метки)
14	0	Timestamp Reply (ответ на запрос временной метки)
17	0	Address Mask Request (запрос сетевой маски)
18	0	Address Mask Reply (ответ на запрос сетевой маски)

Формат эхо-запроса и отклика ICMP

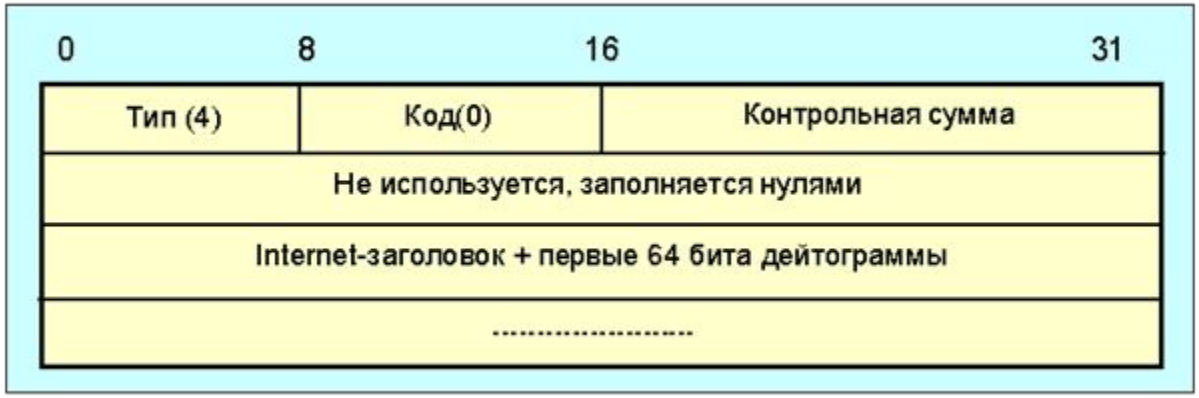
- Сообщения типов 0 и 8 используются для тестирования связи по протоколу IP между двумя узлами сети. Тестирующий узел генерирует сообщения типа 8 (“Эхо-запрос”), при этом “Идентификатор” определяет данный сеанс тестирования (номер последовательности отправляемых сообщений), поле “Номер по порядку” содержит номер данного сообщения внутри последовательности. В поле данных содержатся произвольные данные, размер этого поля определяется общей длиной дейтаграммы, указанной в поле “Total length” IP-заголовка.
- IP-модуль, получивший эхо-запрос, отправляет эхо-ответ. Для этого он меняет местами адреса отправителя и получателя, изменяет тип ICMP-сообщения на 0 и пересчитывает контрольную сумму.
- Анализируется времени оборота дейтаграмм, процент потерь и последовательность прибытия ответов.
- На основе посылки и приема эхо-сообщений работает программа ping.

ICMP-сообщение "адресат не достижим"

- Когда маршрутизатор не может доставить дейтограмму по месту назначения, он посылает отправителю сообщение "адресат не достижим" (destination unreachable).
- Поле «код» содержит целое число, проясняющее причину (в соответствии с таблицей).
- Поле «MTU на следующем этапе» характеризует максимальную длину пакетов на очередном шаге пересылки.
- Так как в сообщении содержится Интернет-заголовок и первые 64-бита дейтограммы, легко понять, какой адрес оказался недостижим.
- Этот тип ICMP-сообщения посылается и в случае, когда дейтограмма имеет флаг DF=1 ("не фрагментировать"), а фрагментация необходима. В поле код в этом случае будет записано число 4.

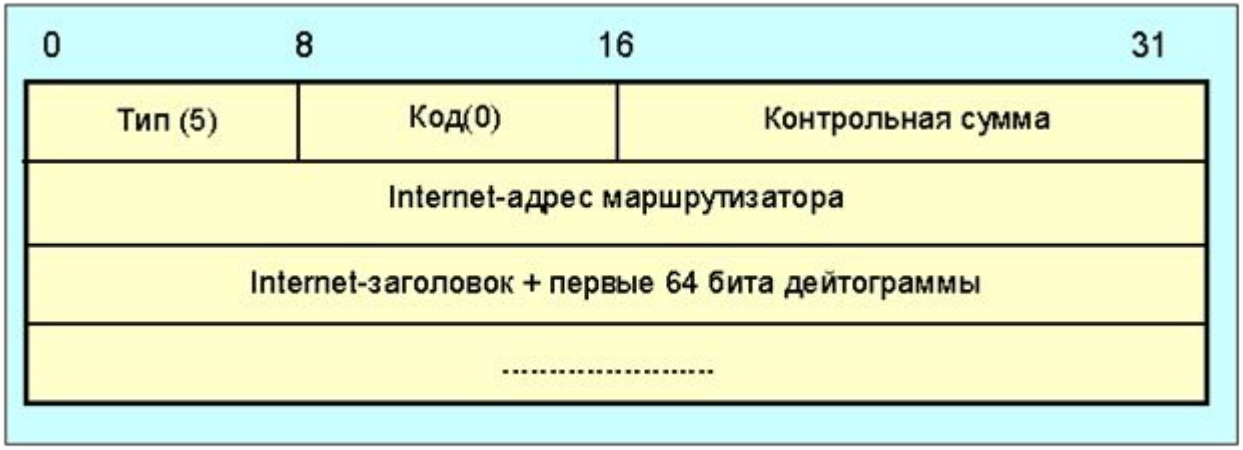
ICMP-сообщение «Замедление источника»

- Сообщения типа 4 (“Замедление источника”) генерируются в случае переполнения (или опасности переполнения) буферов обработки дейтаграмм адресата или промежуточного узла на маршруте. При получении такого сообщения отправитель должен уменьшить скорость или приостановить отправку дейтаграмм до тех пор, пока он не перестанет получать сообщения этого типа.
- IP-заголовок и начальные слова оригинальной дейтаграммы приводятся для опознания ее отправителем и, возможно, анализа причины сбоя.



ICMP-сообщение Redirect

- Сообщения типа 5 направляются маршрутизатором отправителю дейтаграммы в случае, когда маршрутизатор считает, что дейтаграммы в данное место назначения следует направлять через другой маршрутизатор. Адрес нового маршрутизатора приведен во втором слове сообщения.
- Понятие “место назначения” конкретизируется значением поля “Код”. Информация о том, куда была направлена дейтаграмма, породившая ICMP-сообщения, извлекается из ее заголовка, присоединенного к сообщению. Отсутствие передачи сетевой маски ограничивает область применения сообщений типа 5.



ICMP-сообщение об имеющихся маршрутах

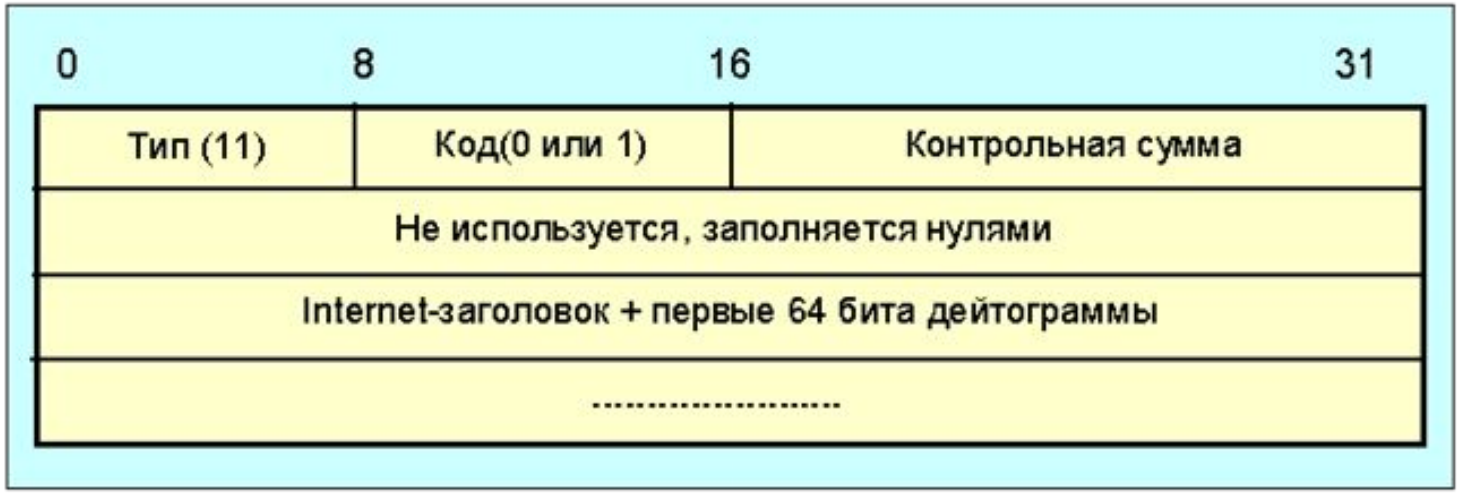
- Таблица маршрутизации создается в результате запросов и объявлений, посылаемых маршрутизаторами.
- Маршрутизаторы посылают в ответ сообщения об имеющейся маршрутной информации. В RFC-1256 описаны форматы ICMP-сообщений такого рода.
- Поле «число адресов» характеризует количество адресных записей в сообщении.
- Поле «длина адреса» - число 32-битных слов для описания адреса маршрутизатора.
- Поле «время жизни» предназначено для записи продолжительности жизни объявленных маршрутов (адресов) в секундах (по умолчанию 30 мин).

Поля «уровень приоритета» представляют собой меру приоритетности маршрута по отношению к другим маршрутам данной подсети. Чем больше этот код тем выше приоритет. Маршрут по умолчанию имеет уровень приоритета 0.

Формат запроса маршрутной информации

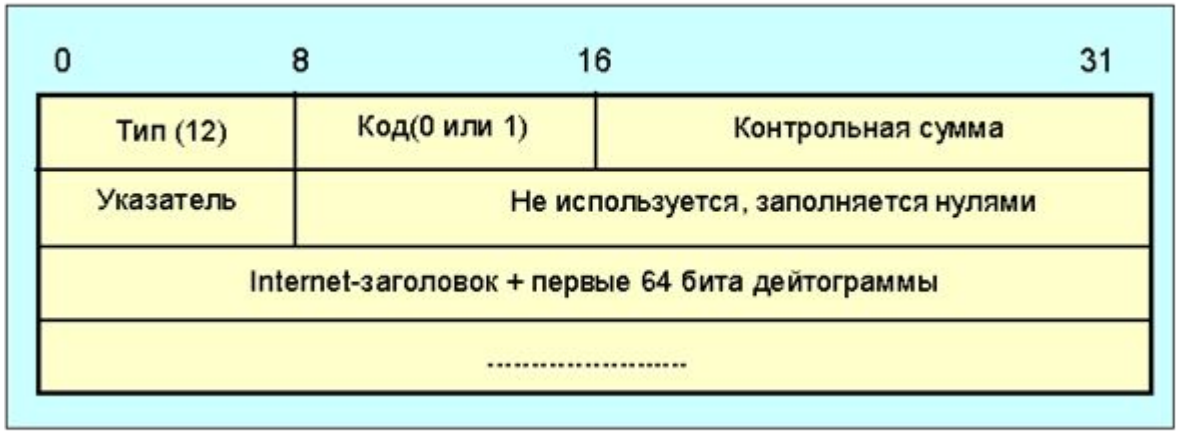
Формат сообщения "время (TTL) истекло"

- Следующий шаг (hop) дейтограммы определяется на основании локальной маршрутной таблицы, а ошибки в ней могут привести к зацикливанию пакетов. Для предотвращения таких зацикливаний используется контроль по времени жизни пакета (TTL). При ликвидации пакета по истечении TTL маршрутизатор посылает отправителю сообщение «время истекло».



Формат сообщения типа "конфликт параметров"

- При выявлении неправильного заголовка дейтограммы, посылается сообщение "конфликт параметров". Это может произойти при неверных параметрах опций.
- Поле «указатель» отмечает октет дейтограммы, который создал проблему.
- Код=1 используется для сообщения о том, что отсутствует требуемая опция (например, опция безопасности при конфиденциальных обменах), поле указатель при значении поля код=1 не используется.



Формат ICMP-запроса временной метки

- В процессе трассировки маршрутов возникает проблема синхронизации работы часов в различных узлов сети. Для запроса временной метки другого узла используется сообщение запрос временной метки.
- Поле тип=13 указывает на то, что это запрос, а тип=14 - на то, что это отклик.
- Поле «идентификатор» и номер по порядку используются отправителем для связи запроса и отклика.
- Поле «исходная временная метка» заполняется отправителем непосредственно перед отправкой пакета.
- Поле «временная метка на входе» заполняется маршрутизатором при получении этого пакета, а «Временная метка на выходе» - непосредственно перед его отправкой.
- Используется в процедурах ping и traceroute.

Формат запроса (отклика) маски подсети

- При работе с подсетью важно знать ее маску. Для получения маски подсети рабочая станция может послать "запрос маски" в маршрутизатор и получить отклик, содержащий эту маску. Рабочая станция может это сделать непосредственно, если ей известен адрес маршрутизатора, либо прибегнув к широковещательному запросу.
- Поле «тип» специфицирует модификацию сообщения, тип=17 - это запрос, а тип=18 - отклик.
- Поля «идентификатор» и «номер по порядку» обеспечивают взаимную привязку запроса и отклика, а поле адресная маска содержит 32-разрядную маску сети.

Протокол ARP

- Протокол ARP (Address Resolution Protocol, Протокол разрешения адреса) предназначен для преобразования IP-адресов в MAC-адреса, часто называемые также физическими адресами.

- MAC-адреса идентифицируют устройства, подключенные к физическому каналу.

- Для передачи IP-дейтаграммы по физическому каналу требуется инкапсулировать эту дейтаграмму в кадр Ethernet и в заголовке кадра указать MAC адрес Ethernet-карты, на которую будет доставлена эта дейтаграмма для ее последующей обработки вышестоящим по стеку протоколом IP.

Работа протокола ARP

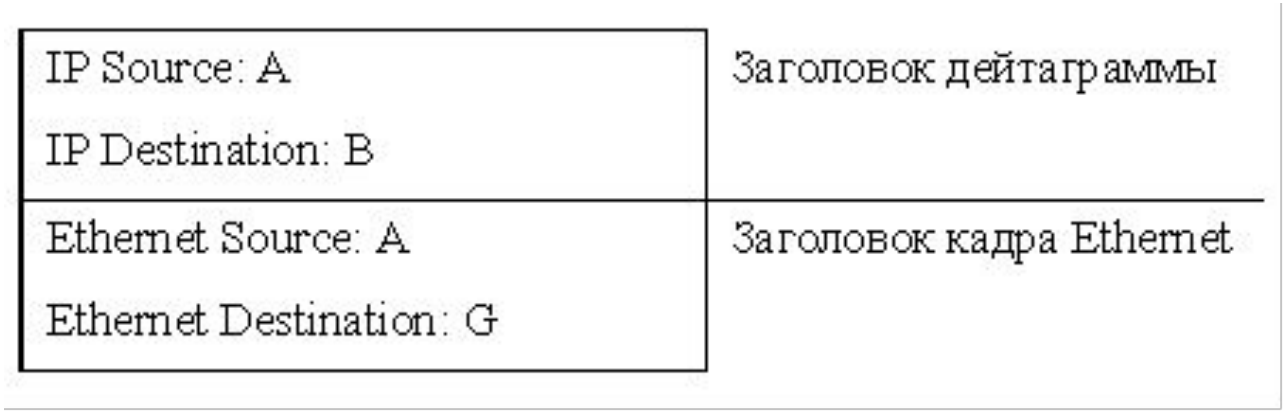
- С сетевого уровня поступает IP-дейтаграмма для передачи в физический канал (Ethernet), вместе с дейтаграммой передается, среди прочих параметров, IP-адрес узла назначения.
- Если в arp-таблице не содержится записи об Ethernet-адресе, соответствующем нужному IP-адресу, модуль arp ставит дейтаграмму в очередь и формирует широковещательный запрос.
- Запрос получают все узлы, подключенные к данной сети; узел, опознавший свой IP-адрес, отправляет arp-ответ (arp-response) со значением своего адреса Ethernet.
- Полученные данные заносятся в таблицу, ждущая дейтаграмма извлекается из очереди и передается на инкапсуляцию в кадр Ethernet для последующей отправки по физическому каналу.
- «Время жизни» записи в таблице 2 мин.
- При повторном обращении в течении этого времени «время жизни» продляется до 10 минут.

Форматы запроса и ответа

- Форматы запроса и ответа одинаковы и отличаются только кодом операции (Operation code, 1 и 2 соответственно).
- HA-Len - длина аппаратного адреса.
- PA-Len – длина адреса сетевого уровня (длина в байтах, например, для IP-адреса PA-Len=4).
- Тип оборудования - это тип интерфейса, для которого отправитель ищет адрес; код содержит 1 для Ethernet.
- Тип протокола сетевого уровня (IP=2048).

ARP для дейтаграмм, направленных в другую сеть

- Дейтаграмма, направленная во внешнюю (в другую) сеть, должна быть передана маршрутизатору.
- Предположим, хост A отправляет дейтаграмму хосту B через маршрутизатор G. Несмотря на то, что в заголовке дейтаграммы, отправляемой из A, в поле “Destination” указан IP-адрес B, кадр Ethernet, содержащий эту дейтаграмму, должен быть доставлен маршрутизатору. Для этого IP-модуль при вызове ARP-модуля передает тому вместе с дейтаграммой в качестве IP-адреса узла назначения адрес маршрутизатора, извлеченный из таблицы маршрутов. Таким образом, дейтаграмма с адресом B инкапсулируется в кадр с MAC-адресом G.



Proxy ARP

- ARP-ответ может отправляться не обязательно искомым узлом, вместо него это может сделать другой узел. Такой механизм называется proxy ARP.

Рассмотрим пример.

- Удаленный хост A подключается по коммутируемой линии к сети 194.84.124.0/24 через сервер доступа G.
- Сервер G выдает хосту A IP-адрес 194.84.124.30, принадлежащий сети 194.84.124.0. Следовательно, любой узел этой сети, например, хост B, полагает, что может непосредственно отправить дейтаграмму хосту A, поскольку они находятся в одной IP-сети.
- IP-модуль хоста B вызывает ARP-модуль для определения физического адреса A. Однако вместо A (который, разумеется, откликнуться не может, потому что физически не подключен к сети Ethernet) откликается сервер G, который и возвращает свой Ethernet-адрес как физический адрес хоста A.
- Вслед за этим B отправляет, а G получает кадр, содержащий дейтаграмму для A, которую G