Previous UP Next INDEX Поиск

Previous: 4.4.3 Протокол TCP **UP:** 4.4 Интернет
Next: 4.4.5 Протокол XTP

4.4.4 Протокол передачи команд и сообщений об ошибках (ICMP)

Семенов Ю.А. (ИТЭФ-МФТИ)
Semenov Yu (ITEP-MIPT)

Протокол передачи команд и сообщений об ошибках (ICMP - internet control message protocol, RFC-792, -1256) выполняет многие и не только диагностические функции, хотя у рядового пользователя именно этот протокол вызывает раздражение, сообщая об его ошибках или сбоях в сети. Именно этот протокол используется программным обеспечением ЭВМ при взаимодействии друг с другом в рамках идеологии TCP/IP. Осуществление повторной передачи пакета, если предшествующая попытка была неудачной, лежит на TCP или прикладной программе. При пересылке пакетов промежуточные узлы не информируются о возникших проблемах, поэтому ошибка в маршрутной таблице будет восприниматься как неисправность в узле адресата и достоверно диагностироваться не будет. 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, решаемые ICMP

Подводя итоги, можно сказать, что 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.

Следует только иметь в виду, что получив отклик на посланный запрос, мы узнаем состояние объекта не в данный момент, а RTT/2 тому назад.

ICMP-сообщения об ошибках никогда не выдаются в ответ 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.

Эти правила призваны блокировать потоки дейтограмм, посылаемым в отклик на мультикастинг или широковещательные ICMP-сообщения.

ICMP-сообщения имеют свой формат, а схема их вложения аналогична UDP или TCP и представлена на рис. 4.4.4.1.

Схема вложения ICMP-пакетов в Ethernet-кадр

Рис. 4.4.4.1. Схема вложения ICMP-пакетов в Ethernet-кадр

Все ICMP пакеты начинаются с 8-битного поля типа ICMP и его кода (15 значений).

По существу, значения полей типа и кода выполняют почти ту же функцию, что и порт в TCP и UDP-протоколах.

Код уточняет функцию ICMP-сообщения. Таблица этих кодов приведена ниже (символом * помечены сообщения об ошибках, остальные - являются запросами):

Типы и коды ICMP-сообщений

Таблица 4.4.4.1. Типы и коды ICMP-сообщений.

ICMP-сообщение		**Описание сообщения**
Тип	**Код**	
0		Эхо-ответ (ping-отклик)
3		Адресат недостижим

	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
4	0	* Отключение источника при переполнении очереди (**quench**)
5		Переадресовать (изменить маршрут)

	0	Переадресовать дейтограмму в сеть (устарело)
	1	Переадресовать дейтограмму на ЭВМ
	2	Переадресовать дейтограмму для типа сервиса (tos) и сети
	3	Переадресовать дейтограмму для типа сервиса и ЭВМ
8	0	Эхо запроса (ping-запрос).
9	0	Объявление маршрутизатора
10	0	Запрос маршрутизатора
11		Для дейтограммы время жизни истекло (ttl=0):
	0	*при передаче
	1	* при сборке (случай фрагментации).
12		* Проблема с параметрами дейтограммы
	0	* Ошибка в ip-заголовке
	1	* Отсутствует необходимая опция
13		Запрос временной метки
14		Временная метка-отклик
15		Запрос информации (устарел)
16		Информационный отклик (устарел)
17		Запрос адресной маски

18		Отклик на запрос адресной маски

Процедура "отключение источника" (quench, поле тип ICMP=4) позволяет управлять потоками данных в каналах Интернет. Не справляясь с обработкой дейтограмм, ЭВМ-адресат может послать запрос "отключить источник" отправителю, который может сократить темп посылки пакетов или вовсе прервать их посылку. Специальной команды, отменяющей прежние запреты, не существует. Если сообщения об отключении прекращаются, источник может возобновить посылку пакетов или увеличить частоту их отправки. Ниже (на рис. 4.4.4.2) представлен формат эхо-запроса (ping) и отклика для протокола ICMP.

Форматы пакетов ICMP

Рис. 4.4.4.2. Формат эхо-запроса и отклика ICMP

Поля ***идентификатор*** (обычно это идентификатор процесса) и ***номер по порядку*** (увеличивается на 1 при посылке каждого пакета) служат для того, чтобы отправитель мог связать в пары запросы и отклики. Поле *тип* определяет, является ли этот пакет запросом (8) или откликом (0). Поле *контрольная сумма* представляет собой 16-разрядное дополнение по модулю 1 контрольной суммы всего ICMP-сообщения, начиная с поля *тип*. Поле *данные* служит для записи информации, возвращаемой

отправителю. При выполнении процедуры ping эхо-запрос с временной меткой в поле данные посылается адресату. Если адресат активен, он принимает IP-пакет, меняет адрес отправителя и получателя местами и посылает его обратно. ЭВМ-отправитель, восприняв этот отклик, может сравнить временную метку, записанную им в пакет, с текущим показанием внутренних часов и определить время распространения пакета (RTT - round trip time). Размер поля *данные* не регламентирован и определяется предельным размером IP-пакета.

Так как в пакете ICMP нет поля порт, то при запуске нескольких процессов PING одновременно может возникнуть проблема с тем какому из процессов следует передать тот или иной отклик. Для преодоления этой неопределенности следует использовать уникальные значения полей идентификатор.

Поле ***идентификатор*** бывает важно, когда ЭВМ используется как программируемый генератор трафика. В этом случае очередной ICMP-пакет посылается, не дожидаясь прихода отклика. Более того, такие пакеты могут генерироваться несколькими процессами одновременно. В этом случае поле ***идентификатор*** становится необходимым, чтобы определить, какому процессу ОС передать очередной отклик.

Время распространения ICMP-запроса, вообще говоря, не равно времени распространения отклика. Это связано не только с возможными изменениями в канале. В общем случае маршруты их движения могут быть различными.

Когда маршрутизатор не может доставить дейтограмму по месту назначения, он посылает отправителю сообщение "адресат не достижим" (destination unreachable). Формат такого сообщения показан ниже (на рис. 4.4.4.3).

Рис. 4.4.4.3. Формат ICMP-сообщения "адресат не достижим"

Поле *код* содержит целое число, проясняющее положение дел. Перечень кодов таких сообщений помещен в таблице 4.4.4.1. 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.

Если буфер приема сообщения переполнен, ЭВМ посылает сообщение типа 4 для каждого из не записанных в буфер сообщений.

Так как собственные часы различных ЭВМ имеют свое представление о времени, протокол ICMP, среди прочего, служит для синхронизации работы различных узлов, если это требуется (запросы временных меток). Протокол синхронизации NTP (network time protocol) описан в RFC-1119.

Когда дейтограммы поступают слишком часто и принимающая сторона не справляется с этим потоком, отправителю может быть послано сообщение с требованием резко сократить информационный поток (quench-запрос), снижение потока должно продолжаться до тех пор, пока не прекратятся quench-запросы. Такое сообщение имеет формат:

Рис. 4.4.4.4. Формат icmp-запроса снижения загрузки

В Internet таблицы маршрутизации остаются без изменений достаточно долгое время, но иногда таблицы все же меняются. Если маршрутизатор обнаружит, что ЭВМ использует неоптимальный маршрут, он может послать ей ICMP-запрос переадресации. Формат такого сообщения отображен на

рисунке 4.4.4.5.

Рис. 4.4.4.5. Формат ICMP-запроса переадресации

Поле *Internet-адрес маршрутизатора* содержит адрес маршрутизатора, который ЭВМ должна использовать, чтобы посланная дейтограмма достигла места назначения, указанного в ее заголовке. В поле *internet-заголовок* кроме самого заголовка лежит 64 первых бита дейтограммы, вызвавшей это сообщение. Поле *код* специфицирует то, как нужно интерпретировать адрес места назначения (см. табл. 4.4.4.1).

Команды переадресации маршрутизатор посылает только ЭВМ и никогда другим маршрутизаторам. Рассмотрим конкретный пример. Пусть некоторая ЭВМ на основе своей маршрутной таблицы посылает пакет маршрутизатору М1. Он, просмотрев свою маршрутную таблицу, находит, что пакет следует переслать маршрутизатору М2. Причем выясняется, что пакет из М1 в М2 следует послать через тот же интерфейс, через который он попал в М1. Это означает, что М2 доступен и непосредственно для ЭВМ-отправителя. В такой ситуации М1 посылает ICMP-запрос переадресации ЭВМ-отправителю указанного пакета с тем, чтобы она внесла соответствующие коррективы в свою маршрутную таблицу.

Маршрутная таблица может загружаться из файла, формироваться менеджером сети, но может создаваться и в результате запросов и объявлений, посылаемых маршрутизаторами. После загрузки системы маршрутизатор посылает широковещательный запрос. Один или более маршрутизаторов посылают в ответ сообщения об имеющейся маршрутной информации. Кроме того, маршрутизаторы периодически (раз в 450-600 сек.) широковещательно объявляют о своих маршрутах, что позволяет другим маршрутизаторам скорректировать свои маршрутные таблицы. В RFC-1256 описаны форматы ICMP-сообщений такого рода (см. рис. 4.4.4.6).

Рис. 4.4.4.6. Формат ICMP-сообщений об имеющихся маршрутах

Поле *число адресов* характеризует количество адресных записей в сообщении. Поле *длина адреса* содержит число 32-битных слов, необходимых для описания адреса маршрутизатора. Поле *время жизни* предназначено для записи продолжительности жизни объявленных маршрутов (адресов) в секундах. По умолчанию время жизни равно 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 (8 октетов, рис. 4.4.4.7).

Рис. 4.4.4.7 Формат запроса маршрутной информации

Так как следующий прогон (hop) дейтограммы определяется на основании локальной маршрутной таблицы, ошибки в последней могут привести к зацикливанию пакетов. Для подавления таких циркуляций используется контроль по времени жизни пакета (TTL). При ликвидации пакета по истечении TTL маршрутизатор посылает отправителю сообщение "время истекло", которое имеет формат (рис. 4.4.4.8):

Рис. 4.4.4.8. Формат сообщения "время (ttl) истекло"

Значение поля *код* определяет природу тайм-аута (см. табл. 4.4.4.1).

Когда маршрутизатор или ЭВМ выявили какую-либо ошибку, не из числа описанных выше (например, нелегальный заголовок дейтограммы), посылается сообщение "конфликт параметров". Это может произойти при неверных параметрах опций. При этом посылается сообщение вида (рис. 4.4.4.9):

Рис. 4.4.4.9. Формат сообщения типа "конфликт параметров"

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.

В процессе трассировки маршрутов возникает проблема синхронизации работы часов в различных ЭВМ. К счастью синхронизация внутренних часов ЭВМ требуется не так часто (например, при выполнении синхронных измерений), негативную роль здесь могут играть задержки в каналах связи. Для запроса временной метки другой ЭВМ используется сообщение запрос временной метки, которое вызывает отклик с форматом (рис. 4.4.4.10):

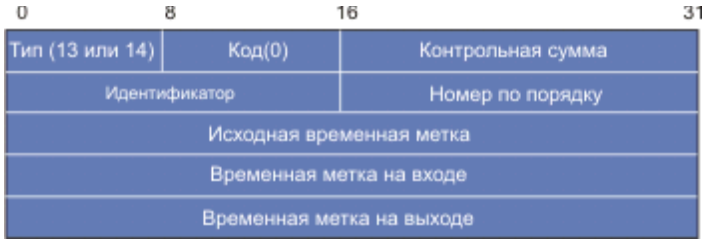

Рис. 4.4.4.10. Формат ICMP-запроса временной метки

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. Именно этот формат используется в процедурах ping и traceroute. Эти процедуры позволяют не только диагностировать, но и оптимизировать маршруты. Например, команда traceroute cernvm.cern.ch, выданная в ЭВМ SUN (ИТЭФ), может отобразить на экране (в скобках указаны IP-адреса узлов и значения времени жизни дейтограмм, значения RTT приводится в миллисекундах):

	traceroute to crnvma.cern.ch	*(128.141.2.4) 30 hops max, 40 byte packets*
1	itep-fddi-bbone	(193.124.224.50) 3 ms 2 ms 3 ms
2	msu-tower.moscow.ru.radio-msu.net	(193.124.137.13) 3 ms 3 ms 3 ms
3	npi-msu.moscow.ru.radio-msu.net	(193.124.137.9) 27 ms 3 ms 9 ms
4	desy.hamburg.de.radio-msu.net	(193.124.137.6) 556 ms 535 ms 535 ms
5	* 188.1.133.56	(188.1.133.56) 637ms 670ms
6	duesseldorf2.empb.net	(193.172.4.12) 740ms(ttl=59!) 839ms(ttl=59!) 2066ms(ttl=59!)

7	bern3.empb.net	(193.172.4.30) 2135ms (ttl=58!) 1644ms (ttl=58!) 1409ms (ttl=58!)
8	cernh3.euro-hep.net	(193.172.24.10) 1808ms 1508ms 1830ms
9	cgate1.cern.ch	(192.65.185.1) 1116ms 1270ms 1278ms
10	crnvma.cern.ch	(128.141.2.4) 1132ms 1362ms 1524ms

Отсюда видно, что наиболее узкими участками маршрута являются Гамбург-Дюссельдорф-Берн-CERN. Следует иметь в виду, что канал МГУ-Гамбург является спутниковым и 500мс задержки здесь вносит время распространения сигнала до спутника и обратно. Участок Гамбург-Дюссельдорф (X.25, квота 256кбит/с) является общим также и для потока данных из Германии в США. Передача IP поверх X.25 также снижает эффективную широкополосность канала. Остальные связи обладают недостаточной пропускной способностью. Программа ping показывает для данных участков в часы пик высокую долю потерянных пакетов. Таким образом, имея карту связей и используя указанные процедуры, вы можете успешно диагностировать ситуацию в сети. Продвинутые же программисты могут легко написать свои диагностические программы, базирующиеся на ICMP, как для локальной сети, так и для "окрестного" Интернет.

При работе с субсетью важно знать маску этой субсети. Как уже отмечалось выше, IP-адрес содержит секцию адреса ЭВМ и секцию адреса сети. Для получения маски субсети ЭВМ может послать "запрос маски" в маршрутизатор и получить отклик, содержащий эту маску. ЭВМ может это сделать непосредственно, если ей известен адрес маршрутизатора, либо прибегнув к широковещательному запросу. Ниже описан формат таких запросов-откликов (рис. 4.4.4.11).

Рис. 4.4.4.11. Формат запроса (отклика) маски субсети

Поле *тип* здесь специфицирует модификацию сообщения, тип=17 - это запрос, а тип=18 - отклик. Поля *идентификатор* и *номер по порядку*, как обычно, обеспечивают взаимную привязку запроса и отклика, а поле *адресная маска* содержит 32-разрядную маску сети.

Previous: 4.4.3 Протокол TCP **UP:** 4.4 Интернет
Next: 4.4.5 Протокол XTP