Занятие №23 «Маршрутизация в IP сетях»

1. «Структура IP пакета»

Имеется прямая связь между количеством полей заголовка пакета и функциональной сложностью протокола, который работает с этим заголовком. Чем проще заголовок — тем проще соответствующий протокол. Большая часть действий протокола связана с обработкой той служебной информации, которая переносится в полях заголовка пакета. Изучая назначение каждого поля заголовка IP-пакета, мы получаем не только формальные знания о структуре пакета, но и знакомимся с основными функциями протокола IP.

IP-пакет состоит из заголовка и поля данных. Ниже перечислены поля заголовка.

Поле **номера версии** занимает 4 бита и идентифицирует версию протокола IP. Сейчас повсеместно используется версия 4 (IPv4), хотя все чаще встречается и новая версия (IPv6).

Значение **длины заголовка** IP-пакета также занимает 4 бита и измеряется в 32-битовых словах. Обычно заголовок имеет длину в 20 байт (пять 32-битовых слов), но при добавлении некоторой служебной информации это значение может быть увеличено за счет дополнительных байтов в поле параметров. Наибольшая длина заголовка составляет 60 байт.

Поле **типа сервиса** (Type of Service, ToS) имеет и другое, более современное название — байт дифференцированного обслуживания, или DS-байт. Этим двум названиям соответствуют два варианта интерпретации этого поля. В обоих случаях данное поле служит одной цели – хранению признаков, которые отражают требования к качеству обслуживания пакета. В прежнем варианте первые три бита содержат значение приоритета пакета: от самого низкого – 0 до самого высокого – 1. Маршрутизаторы и компьютеры могут принимать во внимание приоритет пакета и обрабатывать более важные пакеты в первую очередь. Следующие три бита поля ToS определяют критерий выбора маршрута. Если бит D (Delay – задержка) установлен в 1, то маршрут должен выбираться для минимизации задержки доставки данного пакета, установленный бит T (Throughput – пропускная способность) — для максимизации пропускной способности, а бит R (Reliability – надежность) – для максимизации надежности доставки. Оставшиеся два бита имеют нулевое значение.

Стандарты дифференцированного обслуживания, принятые в конце 90-х годов, дали новое название этому полю и переопределили назначение его битов. В DS-байте также используются только старшие 6 бит, а два младших бита остаются в качестве резерва. Назначение битов DS-байта будет рассмотрено в подразделе «Дифференцированное обслуживание» раздела «Стандарты QoS в IP-сетях» главы 20.

Поле **общей длины** занимает 2 байта и характеризует общую длину пакета с учетом заголовка и поля данных. Максимальная длина пакета ограничена разрядностью поля, определяющего эту величину, и составляет 65 535 байт, однако в большинстве компьютеров и сетей столь большие пакеты не используются. При передаче по сетям различного типа длина пакета выбирается с учетом максимальной длины пакета протокола нижнего уровня, несущего IP-пакеты. Если это кадры Ethernet, то выбираются пакеты с максимальной длиной 1500 байт, умещающиеся в поле данных кадра Ethernet. В стандартах TCP/IP предусматривается, что все хосты должны быть готовы принимать пакеты вплоть до 576 байт длиной (независимо от того, приходят ли они целиком или фрагментами).

Идентификатор пакета занимает 2 байта и используется для распознавания пакетов, образовавшихся путем деления на части (фрагментации) исходного пакета. Все части (фрагменты) одного пакета должны иметь одинаковое значение этого поля.

Флаги занимают 3 бита и содержат признаки, связанные с фрагментацией. Установленный в 1 бит DF (Do not Fragment — не фрагментировать) запрещает маршрутизатору фрагментировать данный пакет, а установленный в 1 бит MF (More Fragments — больше фрагментов) говорит о том, что данный пакет является промежуточным (не последним) фрагментом. Оставшийся бит зарезервирован.

Поле **смещения фрагмента** занимает 13 бит и задает смещение в байтах поля данных этого фрагмента относительно начала поля данных исходного нефраг-ментированного пакета. Используется при сборке/разборке фрагментов пакетов. Смещение должно быть кратно 8 байт.

Поле **времени жизни** (Time To Live, TTL) занимает один байт и используется для задания предельного срока, в течение которого пакет может перемещаться по сети. Время жизни пакета измеряется в секундах и задается источником. Поскольку современные маршрутизаторы редко обрабатывают пакет дольше, чем за одну секунду, то время жизни можно интерпретировать как максимальное число транзитных узлов, которые разрешено пройти пакету. Если значение поля времени

жизни становится нулевым до того, как пакет достигает получателя, пакет уничтожается. Таким образом, время жизни является своего рода часовым механизмом самоуничтожения пакета.

Поле **протокола верхнего уровня** занимает один байт и содержит идентификатор, указывающий, какому протоколу верхнего уровня принадлежит информация, размещенная в поле данных пакета. Значения идентификаторов для разных протоколов приводятся в документе RFC 1700, доступном по адресу http://www. iana.org. Например, 6 означает, что в пакете находится сообщение TCP, 17 – сообщение UDP, 1 –сообщение ICMP.

Контрольная сумма заголовка занимает 2 байта (16 бит) и рассчитывается только по заголовку. Поскольку некоторые поля заголовка меняют свое значение в процессе передачи пакета по сети (например, поле времени жизни), контрольная сумма проверяется и повторно рассчитывается на каждом маршрутизаторе и конечном узле как дополнение к сумме всех 16-битовых слов заголовка. При вычислении контрольной суммы значение самого поля контрольной суммы устанавливается в нуль. Если контрольная сумма неверна, то пакет отбрасывается, как только обнаруживается ошибка.

Поля **IP-адресов источника** и **приемника** имеют одинаковую длину — 32 бита.

Поле **параметров** является необязательным и используется обычно только при отладке сети. Это поле состоит из нескольких подполей одного из восьми предопределенных типов. В этих подполях можно указывать точный маршрут, регистрировать проходимые пакетом маршрутизаторы, помещать данные системы безопасности или временные отметки.

Так как число подполей в поле параметров может быть произвольным, то в конце заголовка должно быть добавлено несколько нулевых байтов для **выравнивания** заголовка пакета по 32-битной границе.

2. «Таблицы маршрутизации в IP сетях»

Рассмотрим механизм IP-маршрутизации на примере составной сети, представленной на рис. 1. В этой сети 11 маршрутизаторов (изображенных в виде пронумерованных квадрантных блоков) объединяют 11 сетей в общую сеть; N1, N2, ..., N11 — это номера сетей. На каждом маршрутизаторе и конечных узлах A и B установлены протоколы IP.

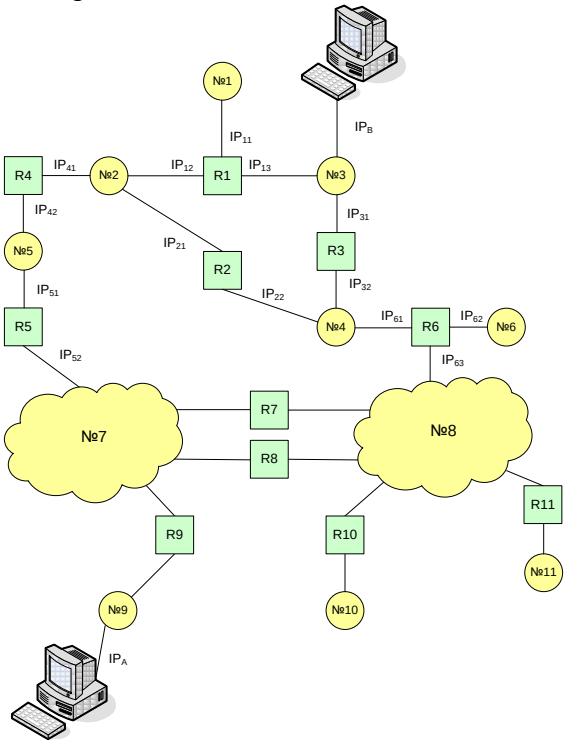

Рис. 1. Принципы маршрутизации в составной сети

Маршрутизаторы имеют по несколько интерфейсов (портов), к которым присоединяются сети. Каждый интерфейс маршрутизатора можно рассматривать как отдельный узел сети: он имеет сетевой адрес и локальный адрес в той подсети, которая к нему подключена. Например, маршрутизатор под номером 1 имеет три интерфейса, к которым подключены сети N1, N2, N3. На

рисунке сетевые адреса этих портов обозначены IP_{11}; IP_{12} и IP_{13}. Интерфейс IP_{11} является узлом сети N1, и следовательно в поле номера сети порта IP_{11} содержится номер N1. Аналогично интерфейс IP_{12} – узел в сети N2, а порт IP_{13} – узел в сети N3. Таким образом, маршрутизатор можно рассматривать как совокупность нескольких узлов, каждый из которых входит в свою сеть. Как единое устройство маршрутизатор не имеет ни отдельного сетевого, ни локального адреса.

В сложных составных сетях почти всегда существуют несколько альтернативных маршрутов для передачи пакетов между двумя конечными узлами. Так, пакет, отправленный из узла А в узел В, может пройти через маршрутизаторы 9, 5, 4 и 1 или маршрутизаторы 9, 8, 6 и 3. Можно найти еще несколько маршрутов между узлами А и В.

Задачу выбора маршрута из нескольких возможных решают маршрутизаторы, а также конечные узлы. Маршрут выбирается на основании имеющейся у этих устройств информации о текущей конфигурации сети, а также на основании критерия выбора маршрута. В качестве критерия часто выступает задержка прохождения маршрута отдельным пакетом, средняя пропускная способность маршрута для последовательности пакетов или наиболее простой критерий, учитывающий только количество пройденных в маршруте промежуточных маршрутизаторов (ретрансляционных участков, или хопов). Полученная в результате анализа информация о маршрутах дальнейшего следования пакетов помещается в таблицу маршрутизации.

Упрощенная таблица маршрутизации

Используя условные обозначения для сетевых адресов маршрутизаторов и номеров сетей в том виде, как они приведены на рис. 1, посмотрим, как могла бы выглядеть таблица маршрутизации, например, в маршрутизаторе 4 (табл. 1).

Таблица 1. Таблица маршрутизации маршрутизатора 4

Адрес назначения	**Сетевой адрес следующего**	**Сетевой адрес выходного порта**	**Расстояние до сети назначения**
N1	IP_{12} (Rl)	IP_{41}	1
N2	-	IP_{41}	0(подсоединена)
N3	IP_{12} (Rl)	IP_{41}	1
N4	IP_{21} (R2)	IP_{41}	1
N5	-	IP_{42}	0(подсоединена)
N6	IP_{21} (R2)	IP_{41}	2
IPв	IP_{21} (R2)	IP_{41}	2
Маршрут по умолчанию	IP_{51} (R5)	IP_{42}	–

Первый столбец таблицы содержит адреса назначения пакетов.

В каждой строке таблицы следом за адресом назначения указывается сетевой адрес следующего маршрутизатора (точнее, сетевой адрес интерфейса следующего маршрутизатора), на который надо направить пакет, чтобы тот передвигался по направлению к заданному адресу по рациональному маршруту.

Перед тем как передать пакет следующему маршрутизатору, текущий маршрутизатор должен определить, на какой из нескольких собственных портов (IP_{41} или IP_{42}) он должен поместить данный пакет. Для этого служит третий столбец таблицы маршрутизации, содержащий сетевые адреса выходных интерфейсов.

Некоторые реализации сетевых протоколов допускают наличие в таблице маршрутизации сразу нескольких строк, соответствующих одному и тому же адресу назначения. В этом случае при выборе маршрута принимается во внимание столбец «расстояние до сети назначения». При этом расстояние измеряется в любой метрике, используемой в соответствии с заданным в сетевом пакете критерием. Расстояние может измеряться временем прохождения пакета по линиям связи, различными характеристиками надежности линий связи на данном маршруте, пропускной способностью или другой величиной, отражающей качество данного маршрута по отношению к заданному критерию. В табл. 1 расстояние между сетями измеряется хопами. Расстояние для сетей, непосредственно подключенных к портам маршрутизатора, здесь принимается равным 0, однако в некоторых реализациях отсчет расстояний начинается с 1.

Когда пакет поступает на маршрутизатор, модуль IP извлекает из поступившего заголовка кадра номер сети назначения и последовательно сравнивает его с номерами сетей из каждой

строки таблицы. Строка с совпавшим номером сети указывает ближайший маршрутизатор, на который следует направить пакет. Например, если на какой-либо порт маршрутизатора 4 поступает пакет, адресованный в сеть N6, то из таблицы маршрутизации следует, что адрес следующего маршрутизатора – IP_{21}, то есть очередным этапом движения данного пакета будет движение к порту 1 маршрутизатора 2.

Чаще всего в качестве адреса назначения в таблице указывается не весь IP-адрес, а только номер сети назначения. Таким образом, для всех пакетов, направляемых в одну и ту же сеть, протокол IP будет предлагать один и тот же маршрут (мы пока не принимаем во внимание возможные изменения в состоянии сети — отказы маршрутизаторов или обрывы кабелей). Однако в некоторых случаях возникает необходимость для одного из узлов сети определить специфический маршрут, отличающийся от маршрута, заданного для всех остальных узлов сети. Для этого в таблицу маршрутизации помещают для данного узла отдельную строку, содержащую его полный IP-адрес и соответствующую маршрутную информацию. Такого рода запись имеется в табл. 1 для узла B. Пусть, например, администратор маршрутизатора 4, руководствуясь соображениями безопасности, решил, что пакеты, следующие в узел B (полный адрес IPB), должны идти через маршрутизатор 2 (интерфейс IP2), а не маршрутизатор 1 (интерфейс IP2), через который передаются пакеты всем остальным узлам сети N3. Если в таблице имеются записи о маршрутах как к сети в целом, так и к ее отдельному узлу, то при поступлении пакета, адресованного данному узлу, маршрутизатор отдаст предпочтение специфическому маршруту.

Таблицы маршрутизации конечных узлов

Задачу маршрутизации решают не только промежуточные (маршрутизаторы), но и конечные узлы – компьютеры. Решение этой задачи начинается с того, что протокол IP, установленный на конечном узле, определяет, направляется ли пакет в другую сеть или адресован какому-нибудь узлу данной сети. Если номер сети назначения совпадает с номером данной сети, это означает, что пакет маршрутизировать не требуется. В противном случае маршрутизация нужна.

Структуры таблиц маршрутизации конечных узлов и транзитных маршрутизаторов аналогичны. Обратимся снова к сети, изображенной на рис. 1. Таблица маршрутизации конечного узла B, принадлежащего сети N3, могла бы выглядеть следующим образом (табл. 2). Здесь IP_B — сетевой адрес интерфейса компьютера B. На основании этой таблицы конечный узел B выбирает, на какой из двух имеющихся в локальной сети N3 маршрутизаторов (R1 или R3) следует посылать тот или иной пакет.

Таблица 2. Таблица маршрутизации конечного узла B

Номер сети назначения	Сетевой адрес следующего маршрутизатора	Сетевой адрес выходного порта	Расстояние до сети назначения
N1	IP_{13} (R1)	IPв	1
N2	IP_{13}(Rl)	IP_B	1
N3	–	IP_B	0
N4	IP_{31} (R3)	IP_B	1
N5	IP_{13} (Rl)	IP_B	2
N6	IP_{31} (R3)	IP_B	2
Маршрут по умолчанию	IP_{31} (R3)	IP_B	–

Конечные узлы в еще большей степени, чем маршрутизаторы, пользуются приемом маршрутизации по умолчанию. Хотя они также в общем случае имеют в своем распоряжении таблицу маршрутизации, ее объем обычно незначителен, что объясняется периферийным расположением всех конечных узлов. Конечный узел часто вообще работает без таблицы маршрутизации, имея только сведения об адресе маршрутизатора по умолчанию. При наличии одного маршрутизатора в локальной сети этот вариант – единственно возможный для всех конечных узлов. Но даже при наличии нескольких маршрутизаторов в локальной сети, когда перед конечным узлом стоит проблема их выбора, часто в компьютерах для повышения производительности прибегают к заданию маршрута по умолчанию.

3. «Алгоритмы маршрутизации в IP сетях»

В IP-сетях в качестве внутренних шлюзовых протоколов, то есть протоколов, применяемых внутри автономных систем, сегодня активно используются протоколы – RIP, OSPF. Внешним шлюзовым

протоколом, то есть протоколом выбора маршрута между автономными системами, сегодня является протокол BGP.

Протокол **RIP** (Routing Information Protocol – протокол маршрутной информации) является внутренним протоколом маршрутизации дистанционно-векторного типа, он представляет собой один из наиболее ранних протоколов обмена маршрутной информацией и до сих пор чрезвычайно распространен в вычислительных сетях ввиду простоты реализации.

В исходном состоянии в каждом маршрутизаторе автоматически создается минимальная таблица маршрутизации, в которой учитываются только непосредственно подсоединенные сети. Протокол RIP замещает запись о какой-либо сети только в том случае, если новая информация имеет лучшую метрику, чем имеющаяся. В результате в таблице маршрутизации о каждой сети остается только одна запись; если же имеется несколько записей, равнозначных в отношении расстояния путей к одной и той же сети, то все равно в таблице остается одна запись, которая пришла в маршрутизатор первая по времени. Для этого правила существует исключение – если худшая информация о какой-либо сети пришла от того же маршрутизатора, на основании сообщения которого была создана данная запись, то худшая информация замещает лучшую.

Для уведомления о том, что некоторый маршрут недействителен, используются два механизма: истечение времени жизни маршрута; указание специального (бесконечного) расстояния до сети, ставшей недоступной.

Протокол **OSPF** (Open Shortest Path First – выбор кратчайшего пути первым) ориентирован на применение в больших сетях.

На первом этапе каждый маршрутизатор строит граф связей сети, в котором вершинами графа являются маршрутизаторы и IP-сети, а ребрами – интерфейсы. Маршрутизаторы обмениваются со своими соседями информацией о графе сети, которой они располагают к данному моменту. Сообщения, с помощью которых распространяется топологическая информация, называются объявлениями о состоянии связей сети (Link State Advertisements, LSA). Все маршрутизаторы сети располагают идентичными сведениями о графе сети, которые хранятся в базе данных о топологии сети.

Второй этап состоит в нахождении оптимальных маршрутов с помощью полученного графа. Для этого используется алгоритм Дейкстры. Каждый маршрутизатор считает себя центром сети и ищет оптимальный маршрут до каждой известной ему сети. В каждом найденном таким образом маршруте запоминается только один шаг – до следующего маршрутизатора, в соответствии с принципом одношаговой маршрутизации. Данные об этом шаге и попадают в таблицу маршрутизации. Если несколько маршрутов имеют одинаковую метрику до сети назначения, то в таблице маршрутизации запоминаются первые шаги всех этих маршрутов.

Пограничный шлюзовой протокол (Border Gateway Protocol, **BGP**).

Маршрутизатор взаимодействует с другими маршрутизаторами по протоколу BGP только в том случае, если администратор явно указывает при конфигурировании, что эти маршрутизаторы являются его соседями.

Такой способ взаимодействия удобен в ситуации, когда маршрутизаторы, обменивающиеся маршрутной информацией, принадлежат разным поставщикам услуг (ISP). Администратор ISP может решать, с какими автономными системами он будет обмениваться трафиком, а с какими нет, задавая список соседей для своих внешних шлюзов.

Протоколы RIP и OSPF, разработанные для применения внутри автономной системы, обмениваются маршрутной информацией со всеми маршрутизаторами, находящимися в пределах их непосредственной досягаемости. Это означает, что информация обо всех сетях появляется в таблице маршрутизации каждого маршрутизатора, так что каждая сеть оказывается достижимой для каждой.

Для установления сеанса с указанными соседями BGP-маршрутизаторы используют протокол TCP, и разнообразные способы аутентификации маршрутизаторов, повышающие безопасность работы систем.

Основным сообщением протокола BGP является сообщение UPDATE (обновить), с помощью которого маршрутизатор сообщает маршрутизатору соседней автономной системы о достижимости сетей, относящихся к его собственной автономной системе.