Протоколы, сервисы и услуги
в Интернет и IP-сетях

Тема № 4
Протокол IPv4

доц. каф. СС и ПД, к.т.н. С. С. Владимиров

2017 г.

Протокол IP

Протокол IP

Маршрутизируемый протокол сетевого уровня стека TCP/IP. Создан в 1981 г. Назначение протокола — передача пакетов между хостами сетей TCP/IP. Основной стандарт для протокола IPv4 — RFC 791 (STD 5) с обновлениями в RFC 1349, 2474, 6864. Являясь одним из базовых протоколов современных компьютерных сетей, протокол IP затрагивается и в многих других стандартах RFC и STD. Неотъемлемой частью протокола IP является адресация сетевых устройств.

IP объединяет сегменты сети в единую сеть, обеспечивая доставку пакетов данных между любыми узлами сети через произвольное число промежуточных узлов (маршрутизаторов). IP не гарантирует надёжной доставки пакета до адресата — в частности, пакеты могут прийти не в том порядке, в котором были отправлены, продублироваться (приходят две копии одного пакета), оказаться повреждёнными (обычно повреждённые пакеты уничтожаются) или не прийти вовсе. Гарантию безошибочной доставки пакетов дают некоторые протоколы более высокого уровня — транспортного уровня сетевой модели OSI, — например, TCP, которые используют IP в качестве транспорта.

Версии протокола IP

1. IPv4 — 32-битная адресация. Основной сетевой протокол в настоящее время.
2. IPv6 — 128-битная адресация. Перспективный протокол. Доля использования IPv6 постепенно увеличивается. Широко используется в сетях Интернета вещей и сенсорных сетях.

Адресация IPv4

Адрес IPv4

Сетевой адрес устройства. В версии IPv4 имеет длину 4 байта (32 бита). IP-адрес можно разделить на две части — адрес сети и адрес узла в сети. Традиционной формой записи IPv4 адреса является запись в виде четырёх десятичных чисел (от 0 до 255), разделённых точками.

Структура адреса IPv4

В рамках одной сети биты адреса сети неизменны.

Число сетевых узлов в одной сети

Если в IP-адресе под адрес узла отведено n бит, то такая сеть содержит

$$N = 2^n - 2 \text{ узлов.}$$

Оставшиеся два адреса отведены под

- адрес сети — все биты адреса узла равны `0`
- широковещательный адрес — все биты адреса узла равны `1`

Виды адресации

- Классовая адресация (Classful network)
- Бесклассовая адресация (Classless Inter-Domain Routing, CIDR)

Классовая адресация IP сетей

Классовая адресация (Classful network)

Принцип сетевой адресации, использовавшийся в Интернете в период с 1981 по 1993 годы. Адресное пространство протокола IPv4 делится на пять классов адресов: A, B, C, D и E. Принадлежность адреса к конкретному классу задаётся первыми битами адреса. Каждый класс определяет либо соответствующий размер сети, то есть количество возможных адресов хостов внутри данной сети (классы A, B, C), либо сеть многоадресной передачи (класс D). Диапазон адресов пятого класса (E) был зарезервирован для будущих или экспериментальных целей.

Классы адресов

Класс	Первые биты	Распр. байт Сеть, Хост	Число сетей	Хостов в сети	Начальный адрес	Конечный адрес
A	0	C.X.X.X	128	16777214	0.0.0.0	127.255.255.255
B	10	C.C.X.X	16384	65534	128.0.0.0	191.255.255.255
C	110	C.C.C.X	2097152	254	192.0.0.0	223.255.255.255
D	1110	Групповой адрес			224.0.0.0	239.255.255.255
E	1111	Зарезервировано			240.0.0.0	255.255.255.255

Бесклассовая адресация IP сетей

Бесклассовая адресация (Classless Inter-Domain Routing, CIDR)

Метод IP-адресации, позволяющий гибко управлять пространством IP-адресов, не используя жёсткие рамки классовой адресации. Использование этого метода позволяет экономно использовать ограниченный ресурс IP-адресов, поскольку возможно применение различных *масок подсетей* к различным подсетям.

Маска подсети (Variable length subnet mask, VLSM)

Битовая маска, определяющая, какая часть IP-адреса узла сети относится к адресу сети (эти биты в маске равны 1), а какая — к адресу самого узла в этой сети (биты маски, равные 0). Маска подсети не является частью IP-пакета. Она указывается в сетевых настройках узла сети.
Зная IP-адрес и маску подсети, можно определить, к какой сети относится данный IP-адрес. Для этого необходимо применить адресу и маске операцию поразрядной конъюнкции (логическое И).

Пример использования маски подсети

```
IP-адрес:       11000000.10101000.00000001.00000010 (192.168.1.2)
Маска подсети:  11111111.11111111.11111110.00000000 (255.255.254.0)
Адрес сети:     11000000.10101000.00000000.00000000 (192.168.0.0)
```

Способы записи маски подсети

- Десятичный: `255.255.254.0`
- Двоичный: `11111111.11111111.11111110.00000000`
- Постфиксный: `/23`

Специальные IP-адреса

Часть адресов IPv4 зарезервированы IANA для конкретных целей. Полный список этих адресов приведен в RFC 3330 и RFC 6890. Некоторые зарезервированные адреса имеют свои RFC, в которых приведено их подробное описание.

Некоторые важные специальные адреса IPv4

- `0.0.0.0/8` — адреса из этого блока относятся к адресам *этой же* сети (RFC 1700).
- `127.0.0.0/8` — диапазон адресов для петлевого (localhost, loopback) интерфейса (RFC 1700). Как правило используется адрес `127.0.0.1/32`. Использование loopback-адреса позволяет устанавливать соединение и передавать информацию для программ-серверов, работающих на том же компьютере, что и программа-клиент, независимо от конфигурации аппаратных сетевых средств компьютера (не требуется сетевая карта, модем, и прочее коммуникационное оборудование, интерфейс реализуется при помощи драйвера псевдоустройства в ядре операционной системы). Таким образом, для работы клиент-серверных приложений на одном компьютере не требуется изобретать дополнительные протоколы и дописывать программные модули.
- `169.254.0.0/16` — диапазон «link-local» адресов, предназначенных для использования в одной физической сети. Эти адреса автоматически присваиваются интерфейсу в тех случаях, когда другие способы автоматического присвоения адресов (DHCP) недоступны.
- `255.255.255.255/32` — широковещательный адрес («limited broadcast»). Используется только внутри одной физической сети.

Частные IP-адреса

Частный IP-адрес (Private IP address)

IP-адрес, принадлежащий к специальному диапазону, выделенному для локальных IP-сетей, т. е. не используемому в сети Интернет. Также его называют внутренним, внутрисетевым, локальным или «серым», в отличие от «белых» адресов, применяемых в сети Интернет. Распределение таких адресов никем не контролируется. В связи с дефицитом свободных IP-адресов, провайдеры как правило раздают своим абонентам именно внутрисетевые адреса — а не внешние.
Иногда частные адреса называют неанонсированными, внешние (так называемые «белые IP») — анонсированными.
Диапазоны частных адресов определены IANA и указаны в RFC 1918.

Диапазоны частных адресов IPv4

Начальный адрес диапазона	Конечный адрес диапазона	Маска подсети (десятичный вид)	Маска подсети (постфикс)
10.0.0.0	10.255.255.255	255.0.0.0	/8
172.16.0.0	172.31.255.255	255.240.0.0	/12
192.168.0.0	192.168.255.255	255.255.0.0	/16

Пакеты, идущие с внутренних IP-адресов или на них, магистральные маршрутизаторы не пропускают. То есть, внутрисетевые машины, если не предпринимать никаких мер, изолированы от Интернета. Тем не менее, есть несколько технологий, которые позволяют выходить таким машинам в Интернет.

Сервер-посредник (прокси-сервер). Сетевой туннель

Сервер-посредник (прокси-сервер)

Многие из старых интернет-служб (электронная почта, IRC, Usenet) специально спроектированы для машин, которые не имеют прямого выхода в Интернет. Для этого в самих протоколах предусмотрена эстафетная передача информации через сервер-посредник.

На примере электронной почты. Корпоративный почтовый сервер имеет два IP-адреса: внутренний и внешний. Для отправки почты пользователь по протоколу SMTP связывается с сервером. Сервер от своего имени выходит в интернет и переправляет почту дальше по цепочке. На этот же сервер по протоколу SMTP поступает входящая корреспонденция. Чтобы проверить ящик, пользователи соединяются с сервером по протоколу POP3.

Для Всемирной паутины была придумана технология «прокси-сервер». Машина с частным адресом обращается к прокси-серверу и посылает на него команды HTTP. Прокси-сервер связывается с веб-сервером от своего имени.

Такая конструкция удовлетворила важнейшие нужды внутрисетевых пользователей. Минусом является сложная архитектура сервера-посредника — он должен поддерживать множество разных протоколов, либо необходимо использовать отдельный сервер-посредник для каждого протокола. А по протоколам, которые посредник не поддерживает или которые не рассчитаны на эстафетную передачу, выход в Интернет невозможен. Одни программы (ICQ, Skype, P2P-часть протокола BitTorrent) проходят сквозь прокси-серверы, «заворачивая» свой протокол в HTTP-пакеты, другие (Subversion, связь с трекером в протоколе BitTorrent) — изначально реализуют свой протокол поверх HTTP. Следующая технология, NAT, позволила внутрисетевым машинам выходить в интернет по любому прикладному протоколу.

Прокси-серверы работают на прикладном уровне и потому могут налаживать цензуру сайтов и кэшировать страницы для экономии трафика — поэтому прокси-серверы широко применяются в корпоративных сетях (даже если другие протоколы работают через NAT). Кроме того, прокси-серверы применяются для особых задач, на которые NAT не способен (например, для передачи файлов в мессенджерах, когда обе машины за NAT).

Сетевой туннель

Технология, когда пакеты сетевого уровня «заворачиваются» в пакеты более высоких уровней (например, транспортного). Это позволяет наладить виртуальную локальную сеть поверх сети совсем другого устройства. Существует много технологий туннелирования (PPPoE, VPN, Hamachi и другие), со своими областями применения.

Трансляция сетевых адресов (NAT)

Технология была задокументирована в 1994 г. (RFC 1631, заменен на RFC 3022 в 2001 г.). Как правило, под NAT понимают так называемый Source NAT (SNAT), при котором маршрутизатор, реализующий NAT, пропуская идущий из локальной сети пакет, заменяет адрес отправителя своим и меняет номер порта, чтобы различать ответные пакеты, адресованные разным локальным компьютерам. Комбинацию, нужную для обратной подстановки, роутер сохраняет у себя во временной NAT-таблице. Когда маршрутизатор получает ответ от сервера, он по таблице открытых соединений восстанавливает адресата и ретранслирует ему ответ.
Через NAT внутрисетевой компьютер может налаживать связь с любым сервером Интернета по любому прикладному протоколу. Но у NAT есть и недостатки. С машиной с частным IP-адресом связаться можно только изнутри локальной сети. С одной стороны, это делает локальную сеть недоступной для многих атак извне. С другой стороны, в некоторых службах Интернета (одноранговых сетях, сетевых играх, передаче файлов в мессенджерах) это создаёт проблемы: если у одного из компьютеров IP-адрес частный, а у другого внешний, инициатором соединения будет клиент с частным IP; если частные у обоих — прямой обмен между ними затруднён. Для решения этой задачи используется так называемый Destination NAT (DNAT).

Концепции DNAT

1. *Статический NAT.* Отображение незарегистрированного IP-адреса на зарегистрированный IP-адрес на основании один к одному.
2. *Динамический NAT.* Отображает незарегистрированный IP-адрес на зарегистрированный адрес из группы зарегистрированных IP-адресов. Динамический NAT также устанавливает непосредственное отображение между незарегистрированным и зарегистрированным адресом, но отображение может меняться в зависимости от зарегистрированного адреса, доступного в пуле адресов, во время коммуникации.
3. *Перегруженный NAT* (NAPT, NAT Overload, PAT, маскарадинг). Форма динамического NAT, который отображает несколько незарегистрированных адресов в единственный зарегистрированный IP-адрес, используя различные порты. Известен также как PAT (Port Address Translation). При перегрузке каждый компьютер в частной сети транслируется в тот же самый адрес, но с различным номером порта.

Типы NAT и технология NAT loopback

Типы NAT

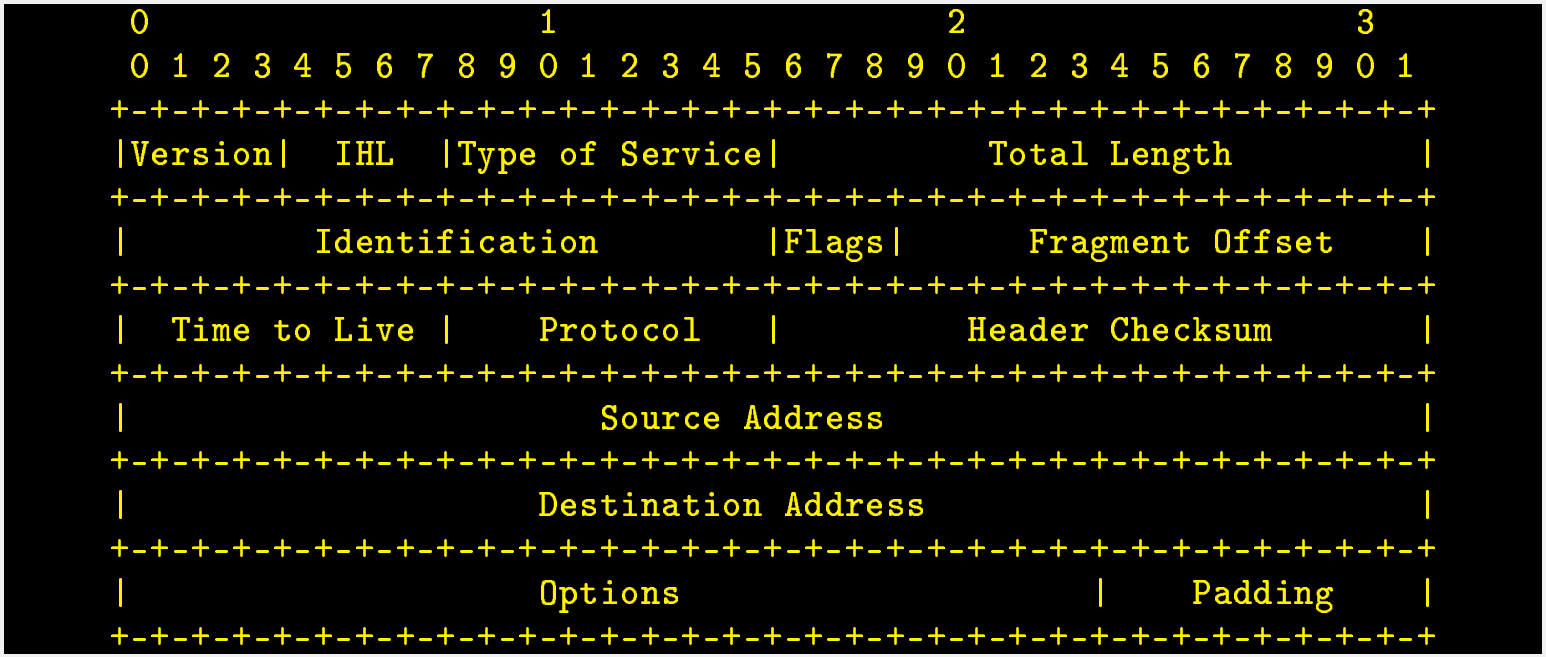
- Cone NAT, Full Cone NAT — Однозначная (взаимная) трансляция между парами «внутренний адрес: внутренний порт» и «публичный адрес: публичный порт». Любой внешний хост может инициировать соединение с внутренним хостом (если это разрешено в правилах межсетевого экрана).
- Address-Restricted cone NAT, Restricted cone NAT — Постоянная трансляция между парой «внутренний адрес: внутренний порт» и «публичный адрес: публичный порт». Любое соединение, инициированное с внутреннего адреса, позволяет в дальнейшем получать ему пакеты с любого порта того публичного хоста, к которому он отправлял пакеты ранее.
- Port-Restricted cone NAT — Трансляция между парой «внутренний адрес: внутренний порт» и «публичный адрес: публичный порт», при которой входящие пакеты проходят на внутренний хост только с одного порта публичного хоста — того, на который внутренний хост уже посылал пакет.
- Симметричный NAT (Symmetric NAT) — Трансляция, при которой каждое соединение, инициируемое парой «внутренний адрес: внутренний порт» преобразуется в свободную уникальную случайно выбранную пару «публичный адрес: публичный порт». При этом инициация соединения из публичной сети невозможна.

NAT loopback

Технология NAT loopback (NAT hairpinning) заключается в том, что пакет, приходящий из внутренней сети на внешний IP-адрес маршрутизатора, считается пришедшим извне — а значит, работают правила брандмауэра, относящиеся ко внешним соединениям. Если пакет успешно пройдёт сквозь брандмауэр, сработает NAT, взяв на себя посредничество между двумя внутрисетевыми машинами. Эта технология позволяет прямо изнутри локальной сети проверить, как настроены сетевые службы, и обеспечивает доступ к серверу, находящемуся в локальной сети, по доменному имени. Недостатком NAT loopback можно считать повышенную нагрузку на хаб и маршрутизатор (по сравнению с прямым доступом к серверу).

Формат пакета IPv4

Формат заголовка пакета IPv4



Заголовок пакета IPv4 содержит 14 полей, из которых 13 являются обязательными. Четырнадцатое поле предназначено для необязательных опций. Поля используют порядок байтов от старшего к младшему, старшие биты идут первыми. Первый бит имеет номер 0. Таким образом, например, поле с версией находится в четырёх старших битах первого байта.

Формат заголовка пакета IPv4. Поля заголовка

Версия протокола (Version)

Имеет размер в четыре бита. Для IPv4 значение этого поля равно 4.

Размер заголовка (Internet Header Length)

Имеет размер в четыре бита. Содержит размер заголовка пакета в 32-битных словах. Поскольку число опций не постоянно, указание размера важно для отделения заголовка от данных. Минимальное значение равно 5 ($5 \cdot 32 = 160$ бит, 20 байт), максимальное — 15 (60 байт).

Тип обслуживания» (Type of Service) или Differentiated Services Code Point (DSCP)

Изначально называлось «тип обслуживания» (Type of Service, ToS), в настоящее время определяется RFC 2474 как «Differentiated Services». Используется для разделения трафика на классы обслуживания, например для установки чувствительному к задержкам трафику, такому как VoIP, большего приоритета.

Размер пакета (Total Length)

16-битный полный размер пакета в байтах, включая заголовок и данные. Минимальный размер равен 20 байтам (заголовок без данных), максимальный — 65535 байт. Хосты должны поддерживать передачу пакетов размером до 576 байт, но современные реализации обычно поддерживают гораздо больший размер. Пакеты большего размера, чем поддерживает канал связи, фрагментируются.

Идентификатор (Identification)

Преимущественно используется для идентификации фрагментов пакета, если он был фрагментирован. Существуют эксперименты по его использованию для других целей, таких как добавление информации о трассировке пакета для упрощения отслеживания пути пакета с подделанным адресом источника.

Формат заголовка пакета IPv4. Поля заголовка

Флаги (Flags)

Поле размером три бита содержащее флаги контроля над фрагментацией. Биты, от старшего к младшему, означают:

0: Зарезервирован, должен быть равен 0.

1: Не фрагментировать

2: У пакета ещё есть фрагменты

Если установлен флаг «не фрагментировать», то в случае необходимости фрагментации такой пакет будет уничтожен. Может использоваться для передачи данных хостам, не имеющим достаточных ресурсов для обработки фрагментированных пакетов.
Флаг «есть фрагменты» должен быть установлен в 1 у всех фрагментов пакета, кроме последнего. У нефрагментированных устанавливается в 0 — такой пакет считается собственным последним фрагментом.

Смещение фрагмента (Fragment Offset)

Поле размером в 13 бит, указывает смещение текущего фрагмента от начала передачи фрагментированного пакета в блоках по 8 байт. Позволяет $(213 - 1) \cdot 8 = 65528$ байт смещения, что превышает максимальный размер пакета. Первый фрагмент в последовательности имеет нулевое смещение.

Время жизни (Time to Live, TTL) пакета

Позволяет предотвратить закольцовывание пакетов в сети путем уничтожения пакетов, превысивших время жизни. Указывается в секундах, интервалы менее секунды округляются до одной секунды. На практике каждый маршрутизатор уменьшает время жизни пакетов на единицу. Пакеты, время жизни которых стало равно нулю уничтожаются, а отправившему посылается сообщение ICMP Time Exceeded. На отправке пакетов с разным временем жизни основана трассировка их пути прохождения (traceroute).

Формат заголовка пакета IPv4. Поля заголовка

Протокол (Protocol)

Указывает, данные какого протокола содержит пакет (например, TCP или ICMP). Присвоенные номера протоколов можно найти на сайте IANA.

Контрольная сумма заголовка (Header Checksum)

16-битная контрольная сумма, используемая для проверки целостности заголовка. Каждый хост или маршрутизатор сравнивает контрольную сумму заголовка со значением этого поля и отбрасывает пакет, если они не совпадают. Целостность данных IP не проверяет — она проверяется протоколами более высоких уровней, которые тоже используют контрольные суммы. Поскольку TTL уменьшается на каждом шаге прохождения пакета, сумма тоже должна вычисляться на каждом шаге. Метод пересчета определён в RFC 1071.

Адрес источника (Source Address)

32-битный адрес отправителя пакета. Может не совпадать с настоящим адресом отправителя из-за трансляции адресов.

Адрес назначения (Destination Address)

32-битный адрес получателя пакета.

Опции (Options)

За адресом назначения может следовать поле дополнительных опций, но оно используется редко. Размер заголовка в этом случае должен быть достаточным чтобы вместить все опции (с учетом дополнения до целого числа 32-битных слов).

Контрольная сумма заголовка IPv4 (Header Checksum)

Контрольная сумма (КС, CS) заголовка IPv4 представляет собой 16-битовое поразрядное дополнение суммы всех 16-битовых слов заголовка. При вычислении контрольной суммы значение самого поля принимается нулевым.

Пример пакета IPv4. Заголовок IPv4 выделен зеленым. Поле КС — синим

```
0000: 00 50 FC 1E BF 8D 00 30 4F 0E 89 65 08 00 45 00
0010: 00 38 89 28 40 00 80 06 11 21 C0 A8 01 32 C3 13
0020: DB 88 04 50 00 15 00 4C 69 E7 3C 00 27 92 50 18
0030: 22 05 D3 39 00 00 55 53 45 52 20 61 6E 6F 6E 79
0040: 6D 6F 75 73 0D 0A
```

Вычисление КС заголовка IPv4

1. Заголовок разбивается на слова размером 16 бит (2 байта) каждое. Поле КС принимается равным нулю (не участвует в вычислении). Все полученные слова суммируются.

$$4500 + 0038 + 8928 + 4000 + 8006 + 0000 + C0A8 + 0132 + C313 + DB88 = 3EEDB$$

2. Если длина результата суммирования превышает 2 байта (4 шестнадцатеричные цифры), то результат делится на две части (правые 4 цифры и остаток), которые суммируются между собой.

$$0003 + EEDB = EEDE$$

3. Находится поразрядное дополнение от итоговой суммы. Результат и будет контрольной суммой заголовка пакета IPv4.

$$FFFF - EEDE = 1121 = CS_{IPv4}$$

Источники

- Материалы с сайта `https://wikipedia.org/`
- Материалы с сайта `https://www.rfc-editor.org/`
- Материалы с сайта `http://www.ibm.com/`
- Компьютерные сети. Принципы, технологии, протоколы: Учебник для вузов / В. Г. Олифер, Н. А. Олифер. — СПб. : Питер, 2010.
- Основы построения INTERNET : Электронный курс / Е. М. Доронин. URL: `http://opds.sut.ru/`
- RFC-791. Internet Protocol.
- RFC-1918. Address Allocation for Private Internets.
- RFC-3330. Special-Use IPv4 Addresses.