Лекция 7. ПРОТОКОЛЫ UDP И TCP

Основные положения

1. **Место в стеке TCP\IP**

   Протоколы TCP (Transmission Control Protocol) и UDP (User Datagram Protocol) относятся к транспортному уровню стека TCP\IP.

   Задачей этих протоколов является связывание приложений на разных хостах, или, в частном случае, в пределах одного хоста.

   Один из протоколов, TCP, обеспечивает надежность доставки данных путем установления соединения и подтверждения приема каждого сообщения.

   Сообщения TCP называются сегментами, сообщения UDP – дейтаграммами.

   В качестве адресной информации используются двухбайтные идентификаторы – номера портов.

2. **Номера портов**

   Порт представляет из себя номер двунаправленной очереди, которую занимает использующее сетевой стек приложение. Все сообщения отправляются с этого номера пора и получаются из очереди с этим номером порта.

   Порт «занимается» приложением, клиентом или сервером. Модули протоколов TCP и UDP занимаются направлением инкапсулированных данных в нужный порт.

   Номера портов для UDP и TCP могут быть одинаковыми.

   Номер порта 0 разрешен для UDP и запрещен для TCP.

   Номера портов для серверов фиксированы и должны быть известны заранее. Для большинства прикладных протоколов номера портов определены в RFC. Но сетевая служба может быть запущена на любом номере порта. Если порт RFC, то при соединении приложение подставляет номер порта по умолчанию, например tcp:443 для протокола https. Но при этом, если сервер работает на порту tcp:10129, то при обращении к сервису, в качестве адреса следует указывать сокет `адрес_сервера:10129`.

   Номера портов клиентов динамические, они выбираются из диапазона > 1024.

   Таким образом на хосте может быть запущено множество клиентов одного прикладного протокола.

   Несколько одинаковых серверов может быть запущено на одном хосте при условии что они запущены на разных сокетах (могут отличаться IP адреса и номера портов).

3. **Протокол UDP**

Порт отправителя (16 бит)	Порт получателя (16 бит)
Длина датаграммы (16 бит)	Контрольная сумма (16 бит)
Данные	

Заголовок UDP состоит из четырех 2-байтных полей:

a. номер UDP-порта отправителя;
b. номер UDP-порта получателя;
c. контрольная сумма, рассчитываемая по заголовку;
d. длина дейтаграммы.

Контрольная сумма служит для диагностики, но не для исправления. Поврежденную дейтаграмму UDP просто отбрасывает.
Соединений UDP не устанавливает.

4. Протокол TCP

Transmission Control Protocol описывается набором RFC: 793, 1323, 1644, 2018, 2581, 2582, 2861, -2873, 2883, 2923, 2988, 3465,3481 и др.
В отличии от UDP он осуществляет доставку дейтограмм, называемых сегментами, в виде байтовых потоков с установлением соединения. Протокол TCP применяется в тех случаях, когда требуется гарантированная доставка сообщений.
Протокол TCP:

a. делит данные прикладного уровня на части – сегменты;
b. использует контрольные суммы по всему сегменту для проверки их целостности;
c. использует для этого алгоритм "скользящего" окна;
d. освобождает прикладные процессы от необходимости таймаутов и повторных передач для обеспечения надежности;
e. устанавливает и контролирует соединение;
f. осуществляет передачу данных с уведомлением.

5. Заголовок TCP

<table>
<tr><td colspan="3">Порт источника (16 бит)</td><td>Порт назначения (16 бит)</td></tr>
<tr><td colspan="4">Номер последовательности (32 бита)</td></tr>
<tr><td colspan="4">Номер подтверждения (32 бита)</td></tr>
<tr><td>Смещение данных (4 бита)</td><td>Зарезервировано (4 бита)</td><td>Флаги (4 бита)</td><td>Размер окна (16 бит)</td></tr>
<tr><td colspan="3">Контрольная сумма (16 бит)</td><td>Указатель важности (16 бит)</td></tr>
<tr><td colspan="4">Опции (32 бита)</td></tr>
<tr><td colspan="4">Данные</td></tr>
</table>

TCP заголовок

Порт источника - идентифицирует приложение клиента, с которого отправлены пакеты. Порт назначения - идентифицирует порт, на который отправлен пакет.
Номер последовательности - выполняет две задачи:
Если установлен флаг SYN, то это начальное значение номера последовательности — ISN (Initial Sequence Number). Первый байт данных, который будут передан в следующем пакете, будет иметь номер последовательности, равный ISN + 1. В противном случае, если SYN не установлен, первый байт данных, передаваемый в данном пакете, имеет этот номер последовательности.
Поскольку поток TCP в общем случае может быть длиннее, чем число различных состояний этого поля, то все операции с номером последовательности должны выполняться по модулю 2^32. Это накладывает практическое ограничение на использование TCP. Если скорость передачи коммуникационной системы такова, чтобы в течение MSL (максимального времени жизни сегмента) произошло переполнение номера последовательности, то в сети может появиться два сегмента

с одинаковым номером, относящихся к разным частям потока, и приёмник получит некорректные данные.

Номер подтверждения - если установлен флаг ACK, то это поле содержит номер последовательности, ожидаемый получателем в следующий раз.

Смещение данных - поле, определяющее размер заголовка пакета TCP в 4-байтных словах. Минимальный размер составляет 5 слов, а максимальный — 15, что составляет 20 и 60 байт соответственно. Смещение считается от начала заголовка TCP.

Зарезервировано – шести битное поле, для будущего использования, должно устанавливаться в ноль. Из них два (5-й и 6-й) уже определены:

CWR (Congestion Window Reduced) — Поле «Окно перегрузки уменьшено» — флаг установлен отправителем, чтоб указать, что получен пакет с установленным флагом ECE (RFC 3168)

ECE (ECN-Echo) — Поле «Эхо ECN» — указывает, что данный узел способен на ECN (явное уведомление перегрузки) и для указания отправителю о перегрузках в сети (RFC 3168)

Флаги (управляющие биты) - поле содержит 6 битовых флагов:

URG(англ. Urgent pointer field is significant) - поле «Указатель важности».

ACK(англ. Acknowledgement field is significant) - поле «Номер подтверждения».

PSH(англ. Push function) - сообщает о данных, накопившиеся в приемном буфере, в приложениях пользователя.

RST(англ. Reset the connection) – обрывает соединения, сбрасывает буфер.

SYN(англ. Synchronize sequence numbers) - Синхронизация номеров последовательности

FIN(англ. FIN bit used for connection termination) - флаг, будучи установлен, указывает на завершение соединения.

Окно - в этом поле содержится число, определяющее в байтах размер данных, которые получатель готов принять.

6. Установление и завершение соединения в TCP

Для установки соединения протокол TCP использует метод тройного рукопожатия.

Реализуется следующий порядок:

Клиент посылает SYN-сегмент (active open). В этом сегменте установлен флаг SYN.

Сервер откликается, посылая свой SYN-сегмент, с ISN (Initial Sequence Number) = 0 (passive open). В этом же сегменте стоит флаг ACK и подтверждается получение первого сегмента от клиента.

Клиент отправляет подтверждение получения SYN-сегмента от сервера сообщением с флагом ACK и с идентификатором равным ISN (сервера)+1..

Закрытие соединения происходит путем передачи сегмента с флагом SYN, подтверждением его сегментом с флагом ACK, передачей вслед за ним ответного сегмента с флагом FIN. После получения подтверждающего сегмента с флагом ACK соединение закрывается, после истечения таймаута.

7. Регулирование скорости в TCP

Скорость передачи данных по TCP не равномерная. Она меняется в зависимости от:

a. времени от начала передачи,
b. успешности передачи,
c. общих условий передачи.

Основным механизмом является алгоритм скользящего окна. Он подразумевает, что отправитель может не дожидаться подтверждения от получателя, а отправлять сегменты в пределах значения байт, указанного в поле размера окна.

Размер окна сначала равен одному сегменту, но по мере накопления положительной истории безошибочной передачи увеличивается до максимального размера. Этот размер определяется как

$$W=RTT*B$$

W – размер окна в битах

B - Полоса пропускания (бит/сек)

RTT (round-trip time) — время приема-передачи — это время на отправку сигнала + время на получение подтверждения.

Размеры окон для обоих участников обмена могут быть разными.

Еще одним параметром, виляющим на производительность отправки\получения модуля TCP, является congestion window – окно перегрузки. При некотором упрощении, можно сказать, что оно в большинстве случаев равно размеру окна передачи и определяет размер буфера данных, доступных для передачи. Окно перегрузки поддерживается отправителем и является средством предотвращения перегрузки канала между отправителем и получателем из-за слишком большого объема трафика. Когда соединение установлено:

a. окно перегрузки независимо на каждом хосте, устанавливается на небольшое значение, кратное MSS (Maximum Segment Size), разрешенному для этого соединения.
b. в дальнейшем, в зависимости от алгоритма управления перегрузкой, при условии, что все сегменты получены и подтверждения достигают отправителя вовремя, к размеру окна добавляется некоторая константа.
c. когда окно достигает ssthresh, окно перегрузки увеличивается линейно при каждом новом полученном подтверждении,
d. в современных ОС может достигать максимального значения в 65535
e. окно перегрузки схлопывается до 1 MSS при обнаруженных проблемах передачи, выключая таймаут прихода подтверждения.

8. Критика TCP

Протокол TCP в настоящее время подвергается критике из-за:

a. проблем медленного открытия соединений (что особенно неприятно для современных Web-приложений, открывающих множество соединений,
b. проблем медленного старта, когда скорость соединения медленно выходит на максимальные значения,
c. слабой адаптации TCP к беспроводным сетям, а также
d. критического влияние не пропускной способности канала, а задержки на канале.

Выход видится не в замене TCP, а в переходе на иные алгоритмы управления перегрузкой и, в переходе на передачу данных на транспортном уровне на UDP, с обеспечением надежности доставки введением дополнительных протоколов между UDP и прикладным уровнем, вроде QUIC.

Основные термины

1. TCP (анг. Transmission Control Protocol - протокол управления передачей). Протокол транспортного уровня, обеспечивающий установку двунаправленного соединения между процессами, идентифицирующимися по сокету (комбинации IP адреса и порта), передачу

потока сегментов внутри соединения с подтверждением приема, управление и завершение соединения. Сообщение TCP содержит в заголовке адреса сегментов в направленном потоке и контрольную сумму при расчете которой используется поле данных и заголовок. Для оптимизации передачи и предотвращения перегрузок сети используется механизм переменного окна, позволяющий вести передачу без получения подтверждения приема каждого сообщения.

2. UDP (англ. User Datagram Protocol — протокол пользовательских дейтаграмм). Протокол транспортного уровня, обеспечивающий передачу сообщений между процессами, идентифицирующийся по сокету (комбинации IP адреса и порта). Сеанс не устанавливается, подтверждения приема не осуществляется. В качестве адресной информации использует порт.
3. QUIC - QUIC (Quick UDP Internet Connections) - это новый протокол, созданный на основе Google QUIC (gQUIC) и утвержденный IETF в целях стандартизации. Предназначен для замены TCP+TLC.
4. Congestion Control – управления перегрузкой – процесс упрвления размером окна перегрузки. Реализуется множеством различных алгоритмов, имеющих различную эффективность в разных условиях.
5. RTT (round-trip time) — время приема-передачи — это время на отправку сигнала + время на получение подтверждения.

Дополнительная литература

1. https://www.rfc-editor.org
2. http://rfc.com.ru
3. http://book.itep.ru/4/44/tcp_443.htm
4. http://book.itep.ru/4/44/udp_442.htm
5. Олифер В., Олифер Н. Компьютерные сети. Принципы, технологии, протоколы. Юбилейное издание. СПб.: Питер, 2020 г.
6. Таненбаум Э., Уэзеролл Д. Компьютерные сети. СПб.: Питер, 2019 г.