Network Working Group
Request for Comments: 5325
Category: Informational

S. Burleigh
NASA/Jet Propulsion Laboratory
M. Ramadas
ISTRAC, ISRO
S. Farrell
Trinity College Dublin

September 2008

Протокол LTP - мотивация

Licklider Transmission Protocol - Motivation

Статус документа

В этом документе определён экспериментальный протокол для сообщества Internet. Документ не задаёт каких-либо стандартов Internet и служит приглашением к дискуссии в целях развития и совершенствования протокола.

Примечание IESG

Данный RFC не рассматривается в качестве кандидата на роль стандарта Internet. Документ является согласованным результатом работы исследовательской группы DTN[1] Комитета по исследованиям Internet (IRTF). Дополнительную информацию о процедурах можно найти в RFC 3932.

Аннотация

В этом документе рассмотрены мотивы разработки протокола LTP[2], предназначенного для обеспечения за счёт повторов надёжной передачи по каналам с экстремально большими значениями времени кругового обхода (RTT) для сообщений и/или частыми разрывами связи. Поскольку связь через межпланетное пространство является наиболее характерным примером такого сорта сред, протокол LTP, прежде всего, предназначен для обеспечения надёжной дальней связи в межпланетном пространстве. Однако этот протокол может применяться и в других средах.

В межпланетной сети Internet, использующей Bundle Protocol, который был разработан исследовательской группой DTN, протокол LTP предназначен для работы в качестве надёжного протокола уровня конвергенции между парами смежных узлов межпланетной сети, связанных по радиоканалам (RF) с одним интервалом. LTP выполняет автоматический повтор запросов (ARQ) на передачу данных за счёт использования селективных подтверждений доставки. Этот механизм учитывает состояние и не требует какого-либо дополнительного согласования.

Данный документ является результатом работы исследовательской группы DTN и согласован с членами этой группы. Публикация документа как RFC не встретила возражений.

Оглавление

1. Введение....1
2. Постановка задачи....2
2.1. Рабочая среда IPN....2
2.2. Почему не TCP или SCTP?....3
3. Обзор протокола....3
3.1. Штатная работа....3
3.1.1. Сигналы о состоянии канала....4
3.1.2. Отложенная передача....4
3.1.3. Таймеры....5
3.2. Повтор передачи....6
3.3. Ускоренный повтор передачи....7
3.4. Прерывание сессии....7
4. Вопросы безопасности....7
5. Взаимодействие с IANA....8
6. Благодарности....8
7. Литература....9
7.1. Информационные ссылки....9

1. Введение

Протокол быстрой передачи LTP разработан для обеспечения надёжной передачи данных по каналам с экстремально большими значениями времени кругового обхода и/или частыми разрывами связи. Связь в межпланетном пространстве является наиболее типичным примером таких сред и протокол LTP предназначен, прежде всего, для поддержки надёжной дальней связи через межпланетные радиоканалы. В частности, протокол LTP предназначен для обеспечения надёжного протокола уровня конвергенции, работающего «под» протоколом DTN [DTN] Bundle [BP], в разворачиваемых DTN средах, где каналы данных характеризуются очень большим временем кругового обхода.

В этом документе рассмотрены мотивы создания протокола LTP, возможности и функции протокола, а также его общее устройство. Этот документ является частью серии RFC, описывающих протокол LTP. Два других документа содержат спецификацию протокола [LTPSPEC] и описание протокольных расширений [LTPEXT].

[1]Delay Tolerant Networking - устойчивая к задержкам сеть.
[2]Licklider Transmission Protocol - протокол ускоренной передачи.

Протокол назван в честь пионера ARPA/Internet - JCR Licklider.

2. Постановка задачи

2.1. Рабочая среда IPN

Существует множество фундаментальных различий между средами наземных коммуникаций (такими, которые используются в Internet) и рабочими средами, которые могут применяться в межпланетных сетях (IPN[1]) [IPN].

Наиболее существенным различием при коммуникациях между точками, расположенными на Земле, и межпланетных коммуникациях является время кругового обхода. Причины такого различия имеют как физическую, так и экономическую природу.

Основной причиной задержки является конечная скорость распространения сигналов. Задержка при прохождении сигнала от Земли до Европы (спутник Юпитера) и обратно составляет от 66 до 100 минут.

Менее очевидными и более динамичными являются задержки, вносимые затмениями. Связи между планетами должна осуществляться посредством лучевой передачи, которая обычно возможна только при отсутствии препятствий между взаимодействующими объектами. Во время сеанса связи возможно возникновение пауз, связанных с перекрытием луча (затмением), в течение которых доставка сообщений становится не возможной и требуется помещать сообщения в очередь для последующей передачи.

Время кругового обхода и затмения могут быть рассчитаны на основе эфемерид взаимодействующих объектов. Дополнительная и менее предсказуемая задержка вносится перерывами в передаче данных, обусловленными экономическими причинами.

Межпланетная связь требует использования специализированного и дорогостоящего оборудования - гигантских антенн, высокочувствительных приёмников и т. п.

Для дальней космической связи, даже не относящейся к сфере деятельности NASA, в настоящее время используется поддерживаемая NASA сеть DSN[2] [DSN]. Коммуникационные ресурсы в настоящее время полностью расписаны и такое состояние сохраниться в обозримом будущем. Следовательно, радиообмен через DSN требуется планировать заблаговременно и выделяемые ресурсы весьма ограничены.

Постоянная полная загрузка означает, что на время кругового обхода для пакетов оказывает влияние не только задержка распространения сигналов и затмения, но и задержки, связанные с планированием и очередями, вносимыми системой управления наземными ресурсами. Пакеты, передаваемые по заданному адресу могут быть задержаны до следующего сеанса связи и величина задержки может измеряться часами, днями и даже неделями.

Такие условия работы вносят множество дополнительных ограничений для протоколов, способных обеспечивать надёжную связь по каналам дальней космической связи.

- Большое время кругового обхода означает существенную задержку между передачей блока данных и приёмом подтверждения доставки от получателя. Если LTP будет откладывать передачу последующих блоков данных до того момента, когда будет подтверждена доставка всех отправленных ранее блоков, дорогостоящие ресурсы каналов дальней космической связи (используемая полоса) будут практически полностью потеряны. Требуется организация множества параллельных «сеансов» передачи данных для предотвращения неполного использования канала связи.

- Подобно любому другому транспорту с гарантированной доставкой, использующему ARQ, протокол LTP учитывает состояния соединений. Для обеспечения гарантий получения переданного блока данных LTP должен сохранять возможность повторной передачи любой части блока данных, который ещё не был получен. Для реализации такой возможности протокол должен отслеживать доставку каждой части переданных блоков данных, чтобы иметь информацию о тех частях блока, которые уже доставлены и которые ещё не доставлены, а также дополнительную информацию, которая может потребоваться для повторной передачи любой части блока данных.

- В сетях IPN время кругового обхода может быть столь большим, а продолжительность сеансов столь малой, что согласование параметров (таких, как выбор скорости передачи) может требовать времени, превышающего продолжительность сеанса связи. Даже при непрерывной связи согласование параметров до начала передачи данных приводит к неэффективному использованию дорогих канальных ресурсов.

- Другое отличие LTP от протокола TCP заключается в том, что соединения TCP являются двухсторонними (блоки данных приложений могут передаваться по одному каналу в обоих направлениях), а сеансы LTP являются односторонними. Это связано с тем, что поток данных, передаваемых при межпланетных полётах, является однонаправленным по своей природе (большое время кругового обхода делает интерактивное взаимодействие неоправданно дорогим, поэтому космические аппараты достаточно автономны и поток команд с Земли весьма мал). Двухсторонний обмен данными в тех случаях, когда это возможно, организуется за счёт создания двух односторонних соединений в противоположных направлениях с разными скоростями.

- Кроме того, расчёт интервалов тайм-аута для сред, в которых обычно используется LTP, существенно отличается от расчёта этих параметров в Internet. Поскольку одновременно может существовать множество параллельных сессий, задержка передачи в любой отдельной сессии в ожидании тайм-аута не снижает производительности соединения в целом. Интервалы тайм-аутов, которые могут быть недопустимыми для TCP, в LTP могут даже не снижать эффективность использования полосы канала.
  Однако полудуплексная двухсторонняя передача данных в LTP делает невозможным использование статистического анализа времени кругового обхода для предсказания значений этого параметра. Время кругового обхода для переданного сегмента N может оказаться на несколько порядков больше аналогичного параметра для сегмента N-1 просто по тому, что между передачей этих сегментов наблюдался период отсутствия связи. Поэтому требуется иной механизм расчёта интервалов тайм-аута.

[1]Interplanetary Internet.
[2]Deep Space Network - сеть глубокого космоса.

2.2. Почему не TCP или SCTP?

Такие характеристики сред, как большие задержки со значительными вариациями, прерывающаяся связность и сравнительно высокая частота ошибок, делают невозможным использование обычного протокола TCP для сквозной передачи данных в IPN. Используя уравнение для расчёта пропускной способности TCP из работы [TFRC], мы можем определить частоту случаев потери (p), при которой может быть достигнута заданная полоса пропускания в установившемся состоянии. Предполагая, что минимальное значение RTT при связи между Землёй и Марсом равно 8 минутам (при минимальном расстоянии между планетами на прохождение света между планетами в одном направлении требуется 4 минуты), размер пакета составляет 1500 байтов, а получатель подтверждает каждый второй пакет, и игнорируя пренебрежимо малые компоненты высоких порядков в p (т. е., второй дополнительный член в знаменателе уравнения для пропускной способности TCP), мы получим значения частоты потерь, допустимой для достижения требуемой пропускной способности, показанные в таблице.

Пропускная способность	***Допустимая частота потерь (p)***
10 Мбит/с	$4{,}68*10^{-12}$
1 Мбит/с	$4{,}68*10^{-10}$
100 кбит/с	$4{,}68*10^{-8}$
10 кбит/с	$4{,}68*10^{-6}$

Отметим, что несмотря на трактовку множества потерь в течение одного периода RTT трактуется как одна потеря при расчёте пропускной способности TCP с помощью уравнения [TFRC], приведённые в таблице значения вероятности потерь являются недостижимыми для каналов в дальний космос.

В контексте нашего обсуждения более жёсткий расчёт пропускной способности TCP, используемый для скоростных каналов (HighSpeed TCP) [HSTCP], просто не рассматривался.

Характерной особенностью TCP является трехэтапное согласование на начальном этапе каждого нового соединения, после которого выполняется механизм замедленного старта. Эти операции создают дополнительные препятствия в работе, поскольку затраты времени на трехэтапное согласование и замедленный старт могут оказаться чрезмерными в средах с большими задержками.

Протокол SCTP[1] [SCTP] может мультиплексировать блоки[2] (объект данных приложения) для организации множества сессий в одном соединении (в SCTP это называется ассоциацией), как это делается в LTP, но в SCTP также затрачивается множество периодов кругового обхода до того, как начнётся передача данных приложений. Очевидно, что такое решение не подходит для использования в средах IPN.

3. Обзор протокола

3.1. Штатная работа

Ниже описана обычная последовательность событий для сеансов передачи LTP.

Работа начинается с момента, когда экземпляр клиента запрашивает у машины LTP передачу блока данных удалённому клиенту.

LTP рассматривает каждый блок данных, как состоящий из двух частей - «красной», при доставке которой используются подтверждения и повторы передачи (при необходимости), и «зелёной», для которой предпринимаются попытки доставить без обеспечения гарантий. Размер каждой из частей может быть нулевым, т. е., любой юлок данных может трактоваться целиком, как «красный» или «зелёный». В первом случае попытки передачи блока повторяются, пока не будет получено подтверждение успешной доставки всего блока, а во втором просто предпринимается попытка передать данные без гарантии доставки. Таким образом, LTP может работать, подобно протоколу TCP и UDP одновременно в рамках одной сессии.

Отметим, что при передаче в режиме «красный-зеленый», «красная» часть **не** имеет какой-либо семантики для уровня важности или приоритета данных в «зелёной» части блока. «Красная» часть просто содержит данные, для которых пользователь запросил надёжную передачу и без которых возможно (но не обязательно) данные «зелёной» части становятся бесполезными. Например, «красная» часть может содержать заголовок прикладного уровня или другие метаданные.

Экземпляр клиента использует программный интерфейс LTP для идентификации LTP экземпляра удалённого клиента, которому нужно передать данные, местоположение передаваемых данных, общий размер передаваемой информации и размер данных в начальной части блока, которые передаются в качестве «красных». Передающая машина начинает сеанс передачи этого блока и уведомляет экземпляр клиента о начале сеанса. Отметим, что параметры коммуникационного сеанса LTP не согласуются, а просто заявляются в одностороннем порядке на уровне управления приложением; передающая машина не согласует начало сеанса с машиной удалённого клиента.

После этого передающая машина инициирует первоначальную передачу, помещая в очередь столько сегментов данных, сколько требуется для передачи блока целиком, с учётом ограничений на размер сегмента, вносимых нижележащим коммуникационным уровнем. Последний сегмент «красной» части блока маркируется, как EORP[3]; этот маркер говорит о завершении «красной части» и служит контрольной точкой (идентифицируется уникальным порядковым номером), показывающей, что приёмная машина должна отправить подтверждение доставки при получении этого сегмента. Последний сегмент блока маркируется, как EOB[4]; этот маркер говорит принимающей машине о возможности расчёта размера блока путём суммирования смещения и размера данных в сегменте.

Протокол LTP рассчитан на работу непосредственно «поверх» протокола канального уровня, но может также в некоторых случаях (например, при разработке программ или при использовании в частных ЛВС) работать «поверх» протокола UDP. В любом случае протокольный уровень, расположенный непосредственно под уровнем LTP мы будем называть локальным канальным уровнем[5].

[1]Stream Control Transmission Protocol - протокол управления потоковой передачей.
[2]В оригинале используется термин chunk. *Прим. перев.*
[3]End of red-part - конец «красной» части.
[4]End of block - конец блока.
[5]Local data-link layer.

На следующем этапе выделяется полоса для очереди, в которую помещены сегменты данных блока, помещённые в очередь сегменты, их размеры передаются локальному протоколу канального уровня (это может быть UDP/IP) через поддерживаемый этим протоколом интерфейс API для передачи машине LTP обслуживающей удалённый экземпляр клиента.

Таймер запускается в момент EORP, поэтому при отсутствии отклика он может быть автоматически запущен снова.

Каждый модуль данных локального протокола канального уровня (кадр канального уровня или дейтаграмма UDP) должен содержать целое число сегментов LTP и локальный протокол канального уровня никогда не должен доставлять неполные сегменты LTP приёмной машине LTP. Когда в качестве локального протокола канального уровня используется UDP, следует использовать расширение LTP для аутентификации [LTPEXT], чтобы гарантировать целостность данных в тех случаях. От использования этого расширения можно отказаться, если на всем пути обеспечивается пренебрежимо малая вероятность повреждения пакетов (как в некоторых приватных ЛВС) или последствиями повреждения данных в процессе передачи и/или обработки можно пренебречь (как в некоторых случаях при разработке приложений). Когда расширение LTP для аутентификации не используется, протокол LTP требует от локального протокола канального уровня контроля целостности всех принятых сегментов. В частности, локальный протокол канального уровня должен детектировать и отбрасывать принятые повреждённые сегменты.

Принятые сегменты, которые не были отброшены, передаются приёмной машине LTP через интерфейс API, поддерживаемый локальным протоколом канального уровня.

При получении первого сегмента данных для блока принимающая машина запускает сеанс приёма для этого блока и уведомляет локальный экземпляр соответствующего клиентского сервиса о начале сеанса. В обычных условиях все сегменты исходной передачи принимаются без ошибок. Следовательно, на получение сегмента данных EORP приёмная машина отвечает (a) помещением в очередь для отправки передающей машине сегмента отчёта, показывающего завершение приёма, и (b) доставкой крайней части блока локальному экземпляру клиентского сервиса; на получение каждого сегмента зелёной части приёмная машина реагирует незамедлительной доставкой принятых данных локальному экземпляру клиентского сервиса.

Все факты доставки данных и протокольные события передаются локальному экземпляру клиентского сервиса с использованием программного интерфейса реализации LTP.

Отметим, что по причине однонаправленности потока данных LTP подтверждения LTP (отчёты о получении) не могут прицепляться к сегментам данных TCP и передаются в отдельном типе сегментов.

На следующем этапе помещённый в очередь сегмент отчёта незамедлительно отправляется передающей машине и запускается таймер, по завершении отсчёта которого отчётный сегмент передаётся повторно, если не будет получен отклик.

Передающая машина получает сегмент отчёта, выключает таймер EORP, помещает в очередь для передачи приёмной машине сегмент подтверждения приёма отчёта и уведомляет локальный экземпляр клиентского сервера об успешной передаче красной части блока. На этом сеанс передачи красной части завершается.

На следующем этапе помещённый в очередь сегмент подтверждения отчёта незамедлительно передаётся приёмной машине.

Приёмная машина получает сегмент подтверждения отчёта и выключает для этого сегмента таймер повтора. Сеанс приёма красной части и передача блока на этом завершается.

3.1.1. Сигналы о состоянии канала

Организация межпланетного канала может повлечь за собой изменения аппаратной конфигурации на основе предполагаемого состояния удалённого коммуникационного устройства. Такие изменения могут включать, например:

- ориентацию приёмной антенны;
- подстройку транспондера на выбранные частоты передачи и/или приёма;
- точная подстройка питания транспондера в последний момент и отключение питания по окончании сеанса.

Мы, следовательно, предполагаем, что операционная среда, в которой работает протокол LTP, способна передавать протоколу LTP информацию о состоянии канала, говорящую протоколу с какой из удалённых машин LTP следует взаимодействовать локальной машине LTP. Операционная среда сама по себе должна иметь такую информацию для корректной настройки оборудования коммуникационного канала.

3.1.2. Отложенная передача

Информация о состоянии канала также уведомляет LTP о невозможности передачи сегментов. В межпланетной связи ни в какой момент нельзя быть уверенным в наличии двухсторонней связи. В LTP всегда возможна генерация исходящего, который невозможно передать незамедлительно, - в ответ на приём сегмента, по тайм-ауту или запросу клиентского сервиса. Эти сегменты должны помещаться в очередь для последующей передачи, когда канал будет организован, о чем укажет информация о состоянии канала.

Концептуально каждый исходящий сегмент LTP добавляется в конец одной или двух очередей (формируя набор очередей) трафика, привязанного к машине LTP, которая является получателем сегмента. Одна из таких очередей является «внутренней операционной очередью» данного набора, другая - очередью данных приложения. Извлечение сегмента из очереди всегда означает его доставку нижележащей коммуникационной системе для незамедлительной передачи. Когда внутренняя операционная очередь не пуста, наиболее старый сегмент в этой очереди является следующим сегментом, который будет извлечён из очереди для передачи адресату. В остальных случаях следующим сегментом для извлечения из очереди и передачи является наиболее старый сегмент очереди данных приложений.

Создание и размещение сегмента в очереди и последующая реальная передача этого сегмента, в принципе, совершенно не синхронизированы.

В случае, когда (a) коммуникационный канал к адресату активен, (b) очередь, в которую помещается данный исходящий сегмент, пуста и (c) эта очередь является внутренней операционной или последняя пуста, сегмент будет передаваться незамедлительно после его создания. Передача вновь созданных сегментов во всех других случаях откладывается.

Концептуально отмена постановки в очередь (de-queuing) сегментов из очередей трафика, привязанных к данному получателю, инициируется при получении сигнала о состоянии канала, указывающего, что базовая коммуникационная система передаёт данные этому адресату (т. е. канал к получателю активен). Это прекращается после получения сигнала о том, что базовая коммуникационная система больше не передаёт данных этому получателю (т. е. канал к адресату больше не активен).

3.1.3. Таймеры

LTP опирается на точный расчёт ожидаемого времени прибытия сегментов с отчётами и подтверждениями для того, чтобы знать, когда нужен упреждающий повтор передачи. Если расчётное время оказалось немного раньше, в результате могут возникнуть издержки от ненужного повтора передачи. С другой стороны, расчётное время прибытия может на несколько секунд позже безопасно и единственным «наказанием» за более поздний тайм-аут и повтор передачи будет незначительная задержка доставки данных и освобождения ресурсов сессии.

Поскольку статистику, основанную на истории интервалов кругового обхода, невозможно безопасно использовать для предсказания времени кругового обхода LTP, мы должны предположить, что время кругового обхода по крайней мере приблизительно детерминировано (т. е. достаточно точная оценка значения RTT может быть выполнена индивидуально в реальном масштабе времени на основе доступной информации).

Расчёт выполняется в два этапа, описанных ниже.

- Первое приближение RTT рассчитывается простым удвоением времени прохождения светового луча к получателю с добавлением произвольного значения предполагаемой дополнительной задержки (например, в очередях или при обработке в оконечных точках). Для операций в глубоком космосе такая добавка обычно составит несколько (немного) секунд. Хотя такая добавка представляется аномальной с точки зрения стандартов Internet, она невелика по сравнению со временем прохождения света туда и обратно. Мы предпочли риск запоздалого повтора передачи, который просто задержит доставку одного блока на сравнительно небольшое время, слишком раннему повтору, ведущему к неоправданному расходу пропускной способности и общему снижению производительности.
- Затем для учёта дополнительных задержек, вносимых перебоями в связи, таймеры динамически приостанавливаются на те периоды, когда относящиеся к делу удалённые машины LTP заведомо не способны передавать отклики. Предполагается, что состояние удалённых систем определяется на основе сигналов о состоянии канала от работающего оборудования.

Приведённое ниже обсуждение служит базой для расчётов ожидаемого времени прибытия LTP.

Общее время, требуемое для одного «кругового обхода» (передача и приём исходного сегмента, а затем передача и приём сегмента подтверждения) состоит из перечисленных ниже компонент.

- Протокольное время обработки. Время, затрачиваемое на выдачу исходного сегмента, его приём, генерацию и выдачу сегмента подтверждения, а также приём сегмента подтверждения.
- Задержка в выходных очередях. Задержка у отправителя исходного сегмента на ожидание в очереди на передачу и задержка у отправителя подтверждения на ожидание в очереди на передачу.
- Время на передачу. Время на выдачу (отправку) всех битов исходного сегмента и время на выдачу всех битов сегмента подтверждения (это важно лишь при очень низкой скорости передачи).
- Время кругового обхода со скоростью света. Время распространения сигнала со скоростью света в обоих направлениях.
- Задержка во входных очередях. Задержка во входной очереди у получателя исходного сегмента и задержка во входной очереди у получателя сегмента подтверждения.
- Задержка на передачу исходного сегмента или сегмента подтверждения по причине потери связи, т. е. в результате прекращения активности у отправителя любого из сегментов по причине затенения, запланированного выхода из зоны видимости и т. п.

В этом контексте, где могут возникать ошибки продолжительностью в секунды и даже минуты, протокольное время обработки на каждой из сторон сессии считается пренебрежимо малым.

Задержка во входной очереди также считается пренебрежимо малой, поскольку даже на мелких космических станциях скорость обработки LTP значительно выше скорости передачи данных.

Применяется два механизма снижения задержки в выходных очередях до пренебрежимо малых значений.

- Ожидаемое время прибытия сегмента подтверждения не рассчитывается, пока базовая система связи не уведомит LTP о начале передачи исходного сегмента. Все задержки в выходной очереди для исходного сегмента к этому моменту уже состоялись.
- Модель отложенной передачи LTP минимизирует задержку доставки сегментов подтверждений (отчёты и подтверждения) базовой коммуникационной системе. Таким образом, сегменты подтверждения (концептуально) добавляются в конец внутренней очереди операций, а не в очередь данных приложения, поэтому они имеют более высокий приоритет передачи по сравнению с другими исходящими сегментами, т. е. они всегда должны извлекаться из очереди для первостепенной передачи. Это ограничивает задержку в выходной очереди для данного сегмента подтверждения до времени, требуемого для извлечения из очереди и отправки всех ранее созданных сегментов подтверждения, которые ещё находятся в очереди на передачу. Поскольку сегменты подтверждения передаются нечасто, и обычно малы, задержка данного сегмента подтверждения в выходной очереди будет, вероятно, минимальной.

Отсрочка расчёта ожидаемого времени прибытия сегмента подтверждения до момента, когда исходный сегмент передаётся в канал (излучается) обеспечивает дополнительный эффект отказа от рассмотрения любой задержки передачи исходного сегмента в результате потери связи на стороне отправителя этого сегмента.

Задержка при передаче на каждой стороне сессии - это просто размер сегмента, поделённый на скорость передачи. Она несущественна за исключением ситуаций, когда скорость передачи предельно мала (например, 10 бит/с) и применение LTP может оказаться нецелесообразным по иным причинам (например, издержки, связанные с заголовком LTP, могут быть слишком велики при такой скорости). Поэтому задержкой на излучение обычно пренебрегают.

Предполагается, что время распространения сигнала в одном направлении с точностью до секунды известно всегда (например, обеспечивается рабочей средой).

Поэтому начальное ожидаемое время прибытия для каждого сегмента подтверждения обычно рассчитывается путём добавления к текущему времени излучения исходного сегмента удвоенного времени прохождения сигнала в одном направлении и 2*N секунд запаса для учёта задержки в очередях и при обработке и длительности излучения на обеих сторонах. Параметр N устанавливается системой управления и значение 2 секунды представляется разумным по умолчанию.

Остаётся одна неизвестная величина - дополнительное время кругового обхода, вносимое потерей связности у получателя сегмента подтверждения. Для его учёта снова будем полагаться на сигналы состояния внешнего канала. Всякий раз, когда прерывание передачи на удалённой машине LTP указывается сигналом состояния канала, останавливаются таймеры обратного отсчёта для всех сегментов подтверждения, ожидаемых данной машиной. По сигналу восстановления передачи отсчёт таймеров возобновляется (фактически) добавляя к каждому ожидаемому времени прибытия интервала простоя, когда таймеры не работали.

3.2. Повтор передачи

Потеря или повреждение переданного сегмента может привести к отклонению работы LTP от обычной последовательности событий, описанной выше.

Потеря одного или нескольких сегментов данных красной части, отличных от сегмента EORP, вызывает повторную передачу данных - приёмная машина возвращает отчёт, указывающий все полученные непрерывные диапазоны красной части (в предположении, что не были получены описанные ниже дискреционные контрольные точки). Отчёт о получении обычно передаётся в одном отчётном сегменте, содержащем уникальный порядковый номер отчёта и охват красной части данных. Например, если данные красной части охватывали блок со смещениями [0:1000] и были получены все данные, кроме диапазона [500:600], сегмент отчёта будет содержать уникальный номер (скажем, 100), а охват [0:1000] будет указан двумя записями - (0:500) и (600:1000). Максимальный размер сегмента отчёта, как и всех сегментов LTP, ограничивается значением MTU в канале данных. Если было потеряно много дискретных сегментов одного большого блока и/или значение MTU в канале данных достаточно мало, может потребоваться создание множества сегментов отчёта. В таких случаях LTP создаёт нужное число сегментов данных и делит область охвата красной части между сегментами отчётов так, что каждый из них может сохранять самостоятельность. Например, при генерации трёх отчётных сегментов для красной части [0:1000000] они могут иметь вид: RS 19 с охватом [0:300000], RS 20 с охватом [300000:950000] и RS 21 с охватом [950000:1000000]. Во всех случаях таймер запускается при отправке каждого сегмента отчёта о получении данных.

При получении каждого отчётного сегмента передающая машина выполняет указанные ниже действия.

- Выключение таймера для контрольной точки, указанной полученным сегментом отчёта (при наличии).
- Размещение в очереди подтверждения приёма отчётного сегмента для отключения таймера повтора на приёмной стороне). Этот сегмент передаётся при ближайшей возможности.
- Если приём данных, указанный в сегменте отчёта, говорит о том, что не все охваченные данные были получены, обычно инициируется повторная передача путём добавления в очередь всех не принятых сегментов. Последний из таких сегментов помечается как контрольная точка и содержит порядковый номер сегмента отчёта, для которого выполняется повтор передачи. Эти сегменты также отправляются при ближайшей возможности, но лишь после сегментов, ранее помещённых в очередь для передачи приёмной стороне. Запускается таймер для контрольной точки, чтобы повторить передачу, если не будет получен соответствующий сегмент отчёта.
- Если приём данных, указанный в сегменте отчёта, говорит о том, что все охваченные данные были получены, а объединение всех указанных в отчётах приёмов данных в этой сессии говорит о получении всей красной части блока, передающая машина уведомляет локального клиента о получении всей красной части блока и окончании красной части сессии.

При получении сегмента подтверждения отчёта приёмная сторона выключает таймер для этого сегмента.

При получении сегмента контрольной точки с отличным от 0 порядковым номером приёмная машина выполняет указанные ниже действия.

- Возвращение отчёта о приёме, содержащего нужное количество отчётных сегментов для информирования о всех принятых данных в сфере охвата упомянутого сегмента отчёта и запуск таймера для каждого сегмента.
- Если к этому моменту получены все данные красной части блока, приёмная машина доставляет полученный красный блок локальному экземпляру клиентского сервиса и при получении сегментов подтверждения приёма, подтверждающих все включённые в отчёт сегменты приём красной части сессии и передача блока завершается. В противном случае продолжаются циклы повторной передачи данных.

Потеря сегмента контрольной точки или отчётного сегмента, созданного в ответ, вызывает тайм-аут и в таких случаях передающая машина обычно повторяет сегмент контрольной точки. Точно так же потеря сегмента отчёта или соответствующего сегмента подтверждения отчёта ведёт к тайм-ауту и приёмная машина обычно повторяет сегмент отчёта.

Отметим, что избыточный сегмент отчёта (т. е. уже полученный и обработанный отправителем) передаётся повторно в результате потери соответствующего сегмента подтверждения отчёта, вызывает, например, передачу другого сегмента подтверждения отчёта, а в остальных случаях игнорируется. Если какой-либо из сегментов данных, повторно переданных в ответ на исходное получение сегмента отчёта, теряется, дальнейший повтор этих сегментов будет запрашиваться отчётом о получении, созданном в ответ на получение последний повторно переданных данных, помеченных как контрольная точка. Таким образом, ненужные повторы подавляются.

Отметим также, что ответственность за отклики на потерю сегмента в LTP разделена между отправителем и получателем блока. Отправитель повторно передаёт сегменты контрольных точек в ответ на тайм-аут для контрольной точки и повторно передаёт недостающие данные в ответ на получение отчёта, указывающего неполный приём, а получатель повторно передаёт сегменты отчётов при возникновении тайм-аута. Можно было сделать ответственным за все повторы лишь отправителя и в этом случае получатель не будет ожидать сегментов подтверждения отчёта и повторять отчётные сегменты. Однако с таким подходом связаны два недостатка.

Во-первых, по причине ограничений на размер сегмента, которые могут быть внесены базовой коммуникационной службой, (по меньшей мере, удалённо) возможно, что отклик на любую отдельную контрольную точку может содержать множество отчётных сегментов. Потребуется дополнительный механизм на стороне отправителя для обнаружения и надлежащего реагирования на потерю некоторого подмножества этих отчётов о приёме. Предложенное в документе решение представляется более простым.

Во-вторых, получающей блок машине нужен способ определения момента, когда сессия может быть закрыта. При отсутствии явного подтверждения финального отчёта (что влечёт за собой повторную отчёта в случае потери подтверждения отчёта) вариантами могут быть (a) немедленное завершение сеанса после передачи сегмента отчёта, который подтверждает полноту приёма или (b) завершение сеанса при получении явного указания от отправителя. В случае (a) потеря финального отчётного сегмента будет вызывать повторную передачу контрольной точки отправителем, но сессия уже будет закрыта к моменту прибытия переданной повторно контрольной точки. Контрольную точку можно резонно счесть первым сегментом данных нового блока, большая которого потеряна в пути и это приведёт к избыточному повтору передачи целого блока. В случае (b) явный сегмент прерывания сессии и ответное подтверждение получателя (требуется для отключения таймера сегмента завершения, который, в свою очередь, нужен на случай потери или повреждения завершающего сегмента в пути) несколько усложняют протокол, занимают избыточную полосу и замедляют освобождение ресурсов, использованных для состояния сессии на стороне отправителя. Здесь снова предложенный в документе вариант проще и эффективней.

3.3. Ускоренный повтор передачи

Повтор передачи сегмента данных происходит лишь при получении отчётного сегмента, указывающего неполный приём данных. Сегменты отчётов обычно передаются только в конце исходной передачи красного блока или в конце повторной передачи. Для некоторых приложений может быть желательным инициирование повторной передачи сегмента данных непосредственно в процессе передачи исходного красного блока, чтобы восстановить пропущенные сегменты быстрее. Это можно реализовать двумя способами, как описано ниже.

- Любой сегмент данных красной части до EORP можно дополнительно пометить как контрольную точку. Получение каждой такой дискреционной точки будет вызывать отправку приёмной машиной сегмента отчёта.
- В любой момент исходной передачи красной части блока (т. е. до получения любого сегмента данных зелёной части блока) принимающая машина может в одностороннем порядке передать дополнительные отчёты о получении. Отметим, что режим Immediate протокола CFDP является примером таких асинхронных отчётов о получении [CFDP]. Отчёты о получении создаются для ускорения повторной передачи точно так же, как обычные отчёты о получении.

3.4. Прерывание сессии

Сеанс передачи может быть прерван передающей или приёмной машиной в ответ на запрос экземпляра локального клиента или при отказе операции LTP, как указано выше.

Прерывающая сеанс машина удаляет все сегменты из очереди сессии и уведомляет локальный экземпляр соответствующего клиентского сервиса о прерывании сессии. Если никаких сегментов в этой сессии ещё не было передано или получено от соответствующей машины LTP, в этот момент прерывающая сессию машина просто закрывает все связанные с сессией состояния и на этом прерывание сессии завершается.

В остальных случаях в очередь помещается сегмент прерывания сессии. При следующей возможности помещённый в очередь сегмент прерывания передаётся машине LTP, обслуживающей экземпляр удалённого клиента. Для сегмента запускается таймер, чтобы иметь возможность автоматического повтора при отсутствии отклика.

Соответствующая машина принимает сегмент прерывания, помещает в очередь для прерывающей машины сегмент подтверждения, удаляет из очереди все другие сегменты указанной сессии, уведомляет локальный экземпляр клиента об отмене блока и закрывает запись состояния для этой сессии.

При следующей возможности помещённый в очередь сегмент подтверждения сегмента прерывания незамедлительно передаётся прерывающей сессию машине.

Прерывающая сеанс машина получает уведомление о разрыве сессии, выключает таймер сегмента прерывания и закрывает все записи состояния для сессии.

Потеря сегмента прерывания или соответствующего отклика на этот сегмент приводит к тайм-ауту. В этом случае прерывающая сессию машина повторно передаёт сегмент прерывания.

4. Вопросы безопасности

Существует явный риск того, что сторонний приёмник может прослушивать передачи LTP по спутниковым и другим широковещательным радиоканалам. Такие приёмники могут также при желании манипулировать сегментами LTP.

Поэтому имеется потребность в защите конфиденциальности и целостности, а также предотвращении DoS[1]-атак.

В частности, проблемы DoS более существенны для LTP по сравнению с типичными протоколами Internet, поскольку LTP по своей природе сохраняет состояния достаточно долго и использует продолжительное время ожидания. Кроме того, сброс узлов LTP для восстановления при атаке может быть достаточно сложен. Таким образом, любой злоумышленник, способный влиять на передачу LTP, может организовать серьёзную DoS-атаку на приёмник LTP.

Рассмотрим, например, наземную ситуацию, где LTP применяется в сети редко расположенных сенсоров. Здесь можно организовать DoS-атаки, в результате которых узлы потеряют критически важную информацию, такую как обновления расписания связи. В таких случаях одна успешная DoS-атака может полностью отключить узел от сети и потребовать его сброса при физическом посещении.

Даже при использовании LTP в глубоком космосе нужно учитывать организуемые с Земли атаки, в частности возможность вставки сообщений в текущий сеанс (обычно без просмотра байтов в предшествующих сообщениях сессии). Такие атаки вероятны при наличии сбоев межсетевых экранов на различных узлах сети или в результате применения троянских программ на легитимных хостах. Многие атаки со вставкой сообщений зависят от возможности атакующего корректно «предсказать» состояния узлов LTP, но опыт показывает, что сделать такие предсказания проще, чем это кажется [DDJ].

Ниже рассматриваются уровни, на которых могут быть реализованы механизмы защиты для повышения безопасности LTP, и связанные с этим компромиссы.

Уровень приложений (выше LTP)

Механизмы безопасности вышележащих уровней явно защищают данные LTP, но оставляют открытыми заголовки LTP. Это слабо защищает (или не защищает совсем) от DoS-атак на LTP, но вполне может сохранить целостность данных и обеспечить их конфиденциальность.

Уровень LTP

Заголовок аутентификации (подобно IPsec [AH]) может помочь в защите от атак с воспроизведением (replay) и других поддельных пакетов. Однако злоумышленнику все ещё доступны заголовки LTP, передаваемые через эфир. Такой подход также требует той или иной инфраструктуры управления ключами для обеспечения строгой проверки подлинности, что не всегда приемлемо. Заголовок аутентификации может ослабить многие DoS-атаки.

Можно задать защиту конфиденциальности для данных LTP и некоторых полей заголовков. Однако это представляется менее привлекательным, поскольку (a) защиту конфиденциальности можно более эффективно организовать на уровнях выше или ниже LTP, (b) управление ключами для такой защиты сложнее (в контексте больших задержек), нежели для защиты целостности и (c) требования к машинам LTP пытаться расшифровывать входящие сегменты само по себе открывает возможность для DoS-атак.

Кроме того, на уровне LTP можно применять различные решения для снижения вероятности успешных DoS-атак. В частности, можно потребовать случайного выбора некоторых полей заголовков (например, номеров сессий).

Уровень канала данных (ниже LTP)

Нижележащие уровни явно могут обеспечивать защиту целостности и конфиденциальности, хотя это может приводить к неоправданным издержкам, если криптографическая защита не требуется для всех данных. Например, это может усложнять управление нижележащими уровнями для шифрования лишь той части данных, для которой это требуется. Шифрование всех данных может вносить существенные издержки для некоторых вариантов применения LTP. Однако на нижележащих уровнях часто выполняется сжатие и исправление ошибок, поэтому они могут быть оптимальным местом для шифрования, поскольку решение вопросов о применении или отказе от сжатия и шифрования сталкивается с одинаковыми проблемами.

В свете сказанного LTP включает в себя механизмы защиты, перечисленные ниже.

Необязательный механизм LTP Authentication является расширением сегмента LTP с идентификатором шифра и необязательным идентификатором ключа перед содержимым сегмента, а также значением аутентификации (код аутентификации или цифровая подпись) после содержимого сегмента. Идентификатор шифра служит для указания размера и формата значения аутентификации. Механизм аутентификации служит для защиты целостности сегмента, а в зависимости от выбранного шифра и метода управления ключами может служить для контроля подлинности сегмента.

Необязательный механизм LTP является расширением сегмента LTP со случайным числовым значением cookie перед содержимым сегмента. За счёт увеличения числа байтов в сегменте, которое не может быть точно предсказано поддельным источником данных, и требования отбрасывать сегменты, в которых нет корректного значения этих байтов, механизм cookie усложняет организацию DoS-атак на машины LTP.

Упомянутые механизмы подробно описаны в документе расширений LTP [LTPEXT].

В дополнение к этому порядковые номера в контрольных точках и отчётах LTP должны иметь случайные целочисленные значения, а разработчикам рекомендуется также выбирать случайные номера для сессий. Это повышает уровень защищенности от DoS-атак. Рекомендации по выбору случайных значений приведены в [PRNG].

5. Взаимодействие с IANA

См. одноимённые разделы в документах [LTPSPEC] и [LTPEXT].

6. Благодарности

Большое спасибо Tim Ray, Vint Cerf, Bob Durst, Kevin Fall, Adrian Hooke, Keith Scott, Leigh Torgerson, Eric Travis, and Howie Weiss за их вклад в разработку протокола и архитектуры устойчивых к задержкам сетей (DTN).

[1]Denial of Service - атака, нацеленная на отказ в обслуживании.

Часть описанного здесь исследования была выполнена в Jet Propulsion Laboratory, California Institute of Technology по контракту с National Aeronautics and Space Administration (DOD Contract DAA-B07-00-CC201, DARPA AO H912; JPL Task Plan No. 80-5045, DARPA AO H870; NASA Contract NAS7-1407).

Спасибо также Shawn Ostermann, Hans Kruse и Dovel Myers из Ohio University за их предложения и советы при выборе решений. Эта работа была выполнена, когда Manikantan Ramadas был аспирантом EECS Dept., Ohio University в Internetworking Research Group Laboratory.

Часть этой работы была выполнена в Trinity College Dublin в рамках контракта SeNDT, финансируемого исследовательским инновационным фондом Enterprise Ireland.

7. Литература

7.1. Информационные ссылки

[LTPSPEC] Ramadas, M., Burleigh, S., and S. Farrell, "Licklider Transmission Protocol - Specification", RFC 5326, September 2008.

[LTPEXT] Farrell, S., Ramadas, M., and S. Burleigh, "Licklider Transmission Protocol - Security Extensions", RFC 5327, September 2008.

[AH] Kent, S., "IP Authentication Header", RFC 4302, December 2005.

[BP] Scott, K. and S. Burleigh, "Bundle Protocol Specification", RFC 5050, November 2007.

[CFDP] CCSDS File Delivery Protocol (CFDP). Recommendation for Space Data System Standards, CCSDS 727.0-B-2 BLUE BOOK Issue 1, October 2002.

[DDJ] I. Goldberg and E. Wagner, "Randomness and the Netscape Browser", Dr. Dobb's Journal, 1996, (pages 66-70).

[DSN] Deep Space Mission Systems Telecommunications Link Design Handbook (810-005) web-page, "http://eis.jpl.nasa.gov/deepspace/dsndocs/810-005/"

[DTN] K. Fall, "A Delay-Tolerant Network Architecture for Challenged Internets", In Proceedings of ACM SIGCOMM 2003, Karlsruhe, Germany, Aug 2003.

[IPN] InterPlanetary Internet Special Interest Group web page, "http://www.ipnsig.org".

[TFRC] Handley, M., Floyd, S., Padhye, J., and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification", RFC 3448, January 2003.

[HSTCP] Floyd, S., "HighSpeed TCP for Large Congestion Windows", RFC 3649, December 2003.

[SCTP] Stewart, R., Ed., "Stream Control Transmission Protocol", RFC 4960, September 2007.

[PRNG] Eastlake, D., 3rd, Schiller, J., and S. Crocker, "Randomness Requirements for Security", BCP 106, RFC 4086, June 2005.

Адреса авторов

Scott C. Burleigh
Jet Propulsion Laboratory
4800 Oak Grove Drive
M/S: 301-485B
Pasadena, CA 91109-8099
Telephone: +1 (818) 393-3353
Fax: +1 (818) 354-1075
EMail: Scott.Burleigh@jpl.nasa.gov

Manikantan Ramadas
ISRO Telemetry Tracking and Command Network (ISTRAC)
Indian Space Research Organization (ISRO)
Plot # 12 & 13, 3rd Main, 2nd Phase
Peenya Industrial Area
Bangalore 560097
India
Telephone: +91 80 2364 2602
EMail: mramadas@gmail.com

Stephen Farrell
Computer Science Department
Trinity College Dublin
Ireland
Telephone: +353-1-896-1761
EMail: stephen.farrell@cs.tcd.ie

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

Полное заявление авторских прав

Интеллектуальная собственность

IETF не принимает какой-либо позиции в отношении действительности или объема каких-либо прав интеллектуальной собственности (Intellectual Property Rights или IPR) или иных прав, которые, как может быть заявлено, относятся к реализации или использованию описанной в этом документе технологии, или степени, в которой любая лицензия, по которой права могут или не могут быть доступны, не заявляется также применение каких-либо усилий для определения таких прав. Сведения о процедурах IETF в отношении прав в документах RFC можно найти в BCP 78 и BCP 79.

Копии раскрытия IPR, предоставленные секретариату IETF, и любые гарантии доступности лицензий, а также результаты попыток получить общую лицензию или право на использование таких прав собственности разработчиками или пользователями этой спецификации, можно получить из сетевого репозитория IETF IPR по ссылке http://www.ietf.org/ipr.

IETF предлагает любой заинтересованной стороне обратить внимание на авторские права, патенты или использование патентов, а также иные права собственности, которые могут потребоваться для реализации этого стандарта. Информацию следует направлять в IETF по адресу ietf-ipr@ietf.org.