

Отображение ядра BEEP на TCP

Mapping the BEEP Core onto TCP

Статус документа

Данный документ содержит спецификацию протокола, предложенного сообществу Internet, и служит запросом к дискуссии в целях развития протокола. Информацию о статусе данного протокола можно найти в текущей редакции документа "Internet Official Protocol Standards" (STD 1). Документ может распространяться без ограничений.

Авторские права

Copyright (C) The Internet Society (2001). All Rights Reserved.

Тезисы

В этом документе описано, как сессия BEEP¹ отображается в одно соединение TCP (Transmission Control Protocol).

Оглавление

1. Введение.....	1
2. Управление сессиями.....	1
3. Обмен сообщениями.....	1
3.1 Управление потоком данных.....	2
3.1.1 Создание канала.....	2
3.1.2 Передача сообщений.....	2
3.1.3 Обработка кадров SEQ.....	2
3.1.4 Использование управления потоком данных.....	2
4. Вопросы безопасности.....	3
Литература.....	3
Адрес автора.....	3
Приложение А. Благодарности.....	3
Полное заявление авторских прав.....	3
Подтверждение.....	3

1. Введение

Этот документ описывает, как сессия BEEP [1] отображается в одно соединение TCP [2]. Требования к отображениям заданы в параграфе 2.5 документа [1].

2. Управление сессиями

Отображение поддержки сессий BEEP на услуги TCP достаточно прямолинейно.

Сессия BEEP организуется, когда создано соединение TCP между двумя партнерами BEEP:

- узел BEEP, задавший пассивный вызов TCP OPEN, называется «слушателем»;
- узел BEEP, задавший активный вызов TCP OPEN, называется «инициатором».

Одновременный вызов TCP OPEN² приведет к тому, что оба партнера BEEP будут считать себя инициаторами и не смогут организовать ни одного канала. Поэтому основанные на протоколе BEEP службы должны быть устроены так, что одновременных вызовов TCP OPEN не возникает.

Если оба партнера согласны освободить (завершить) сеанс BEEP (см. параграф 2.4 в [1]), узел передавший отклик ok, незамедлительно вызывает TCP CLOSE. При получении отклика другой партнер также незамедлительно вызывает TCP CLOSE.

Сессия BEEP прерывается при вызове любым из партнеров TCP ABORT и последующем разрыве соединения TCP.

3. Обмен сообщениями

Отображение обменов BEEP на услуги TCP менее прямолинейно.

Гарантированная отправка и получение обеспечиваются вызовами SEND и RECEIVE протокола TCP, это обеспечивает также упорядоченную доставку сообщений в одном канале.

TCP обеспечивает управление потоком данных в соединении, но при использовании в одной сессии BEEP множества каналов протокол BEEP должен обеспечить механизм для предотвращения «заторов» в соединении. Для этого BEEP использует механизм, принятый в TCP, — управление потоком данных на основе окон, когда каждый канал использует скользящее окно, показывающее число октетов данных, которые можно передать до следующего разрешения.

¹Blocks Extensible Exchange Protocol — расширяемый протокол обмена блоками.

²С учетом предыдущего абзаца корректней было бы написать «Одновременный активный вызов TCP OPEN ...». Прим. перев.

3.1 Управление потоком данных

Напомним из параграфа 2.2.1.2 документа [1], что каждый октет данных, переданный в каждом направлении канала имеет порядковый номер. Нумерация октетов данных в кадрах выполняется последовательно от начала поля данных (payload).

Пространство порядковых номеров конечно, хотя и весьма велико, - от 0 до 4294967295 ($2^{32} - 1$). Конечность пространства номеров приводит к использованию арифметики с модулем 2^{32} и при достижении максимального значения $2^{32} - 1$ отсчет номеров продолжается с 0. Арифметика порядковых номеров описана в разделах 2 - 5 документа [3].

3.1.1 Создание канала

При создании канала первый октет данных в первом кадре получает номер 0, а начальный размер окна устанавливается равным 4096 октетам. После создания канала узлы ВЕЕР могут менять размер окна путем передачи кадров SEQ (3.1.3 Обработка кадров SEQ).

Если узел ВЕЕР запросили создать канала, но он не может выделить для этого канала хотя бы 4096 октетов, узел должен отвергнуть создание канала, как указано в параграфе 2.3.1.2 документа [1]. Аналогично, если в процессе организации сессии ВЕЕР «слушатель» ВЕЕР не может выделить хотя бы 4096 октетов для канала 0, он должен вернуть негативный отклик, как указано в параграфе 2.4 документа [1], вместо передачи сообщения greeting.

3.1.2 Передача сообщений

До передачи сообщения узел ВЕЕР должен убедиться в том, что размер данных этого сообщения не выходит за пределы окна, заданного принимающим узлом ВЕЕР. Если это условие не выполняется, возможны три варианта:

- если окно позволяет передать хотя бы один октет данных (payload), узел ВЕЕР может сегментировать сообщение и начать передачу более мелкого кадра (в соответствии с размером окна);
- узел ВЕЕР может задержать отправку сообщения, пока окно не станет больше;
- узел ВЕЕР может сообщить приложению о невозможности отправки сообщения, что позволит тому повторить попытку позднее (например, после получения сигнала об увеличении доступного окна).

Выбор варианта зависит от реализации, хотя использующим ВЕЕР приложениям рекомендуется представлять механизм влияния на выбор решения.

3.1.3 Обработка кадров SEQ

Поскольку приложение принимает на себя ответственность за входящие кадры данных, его партнеру ВЕЕР следует передавать кадры SEQ для анонсирования нового окна.

ABNF [4] для кадров SEQ имеет вид:

```
seq      = "SEQ" SP channel SP ackno SP window CR LF
ackno    = seqno
window   = size
; channel, seqno и size определены в параграфе 2.2.1 документа [1].
```

Кадр SEQ имеет три параметра:

- номер канала;
- номер подтверждения, указывающий следующий порядковый номер, который отправитель предполагает получить в этом канале;
- размер окна, указывающий число октетов данных (payload), начиная с указанного номером подтверждения, которые отправитель ожидает получить по данному каналу.

Для разделения компонент служит одиночный символ пробела (десятичный код 32). Кадр SEQ завершается парой символов CRLF.

Если при получении кадра SEQ номер канала, номер подтверждения или размер окна не удается определить или заданное значение неприемлемо, сессия ВЕЕР прерывается без генерации отклика. Рекомендуется записать это событие в системный журнал для диагностики.

3.1.4 Использование управления потоком данных

Ключом к успеху управления потоком данных в ВЕЕР является баланс между производительностью и беспристрастностью:

- большие сообщения следует сегментировать в кадры, не превышающие 2/3 от согласованного максимального размера сегмента TCP;
- кадры разных каналов, готовые к передаче, следует выбирать «по кругу»;
- каждый раз при получении кадра следует отправлять кадр SEQ, если размер окна составляет не менее половины доступного для канала буферного пространства;
- если транспортный протокол представляет узлу ВЕЕР множество кадров одновременно, можно передать один общий кадр SEQ.

Для предотвращения проблем взаимодействия с транспортным сервисом важно, чтобы узел ВЕЕР анонсировал окна с учетом доступного буферного пространства для считывания данных из транспортного соединения без задержки. Кадры SEQ для канала должны иметь более высокий приоритет по сравнению с сообщениями этого канала.

Реализации могут представлять использующему ВЕЕР приложению средства управления очередями, например, приоритеты для каналов, распределение буферов и т. п. В частности, разработчикам не следует позволять одному каналу монополизировать окно базового транспорта (например, при медленном считывании данных следует уменьшать размер окна).

Кроме того, реализациям, по возможности, следует поддерживать API транспортного уровня для передачи сведений о перегрузках. Это позволит реализации определить доступную полосу и будет информировать об изменении пропускной способности пути. Отметим, что при наличии в сессии ВЕЕР множества каналов, одновременно передающих большие сообщения, реализации без доступа к данным о загрузке могут сталкиваться с проблемами распределения ресурсов в периоды перегрузки сети.

Реализациям следует также соблюдать рекомендации соответствующих разделов RFC1122 [5], относящиеся к управлению потоками данных (и принимать во внимание, что повтор передачи не применим для этого документа, хотя он и взаимодействует с управлением потоками данных TCP). Например, в параграфе 4.2.2.16 документа RFC1122 [5] сказано: «Получателям TCP **не рекомендуется** отсекать часть окна (т. е., перемещать правый край окна влево),» - а затем рассматривается влияние этого правила на неподтвержденные данные. В контексте отображения ВЕЕР на одно соединение TCP следует реализовать лишь относящиеся к управлению потоком данных части.

4. Вопросы безопасности

Вопросы безопасности рассмотрены в разделе 9 документа [1].

Литература

- [1] Rose, M., "The Blocks Extensible Exchange Protocol Core", [RFC 3080](#), March 2001.
- [2] Postel, J., "Transmission Control Protocol", STD 7, [RFC 793](#), September 1981.
- [3] Elz, R. and R. Bush, "Serial Number Arithmetic", [RFC 1982](#), August 1996.
- [4] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", [RFC 2234](#), November 1997.
- [5] Braden, R., "Requirements for Internet Hosts — Communication Layers", STD 3, [RFC 1122](#), October 1989.

Адрес автора

Marshall T. Rose

Invisible Worlds, Inc.

1179 North McDowell Boulevard

Petaluma, CA 94954-6559

US

Phone: +1 707 789 3700

E-Mail: mrose@invisible.net

URI: <http://invisible.net/>

Перевод на русский язык

Николай Малых

nmalykh@gmail.com

Приложение А. Благодарности

Автор благодарен участникам работы Dave Crocker, Steve Harris, Eliot Lear, Keith McCloghrie, Craig Partridge, Vernon Schryver и Joe Touch. В частности, Dave Crocker внес полезные предложения о природе управления потоком данных для отображения.

Полное заявление авторских прав

Copyright (C) The Internet Society (2001). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Подтверждение

Финансирование функций RFC Editor обеспечено Internet Society.