Internet Engineering Task Force (IETF)
Request for Comments: 8084
BCP: 208
Category: Best Current Practice
ISSN: 2070-1721

G. Fairhurst
University of Aberdeen

March 2017

Network Transport Circuit Breakers

Автоматические выключатели сетевого транспорта

Аннотация

В этом документе разъясняется термин «автоматический выключатель сетевого транспорта» (network transport Circuit Breaker). Описана необходимость автоматических выключателей (Circuit Breaker или CB) для сетевых туннелей и приложений при использовании трафика без контроля перегрузок и указано, где CB нужны и где не нужны. Заданы также требования к созданию CB и описаны ожидаемые результаты применения CB в Internet.

Статус документа

Документ относится к категории Internet Best Current Practice.

Документ является результатом работы IETF[1] и представляет согласованный взгляд сообщества IETF. Документ прошел открытое обсуждение и был одобрен для публикации IESG[2]. Дополнительную информацию о документах BCP можно найти в разделе 2 в RFC 7841.

Информацию о текущем статусе документа, ошибках и способах обратной связи можно найти по ссылке http://www.rfc-editor.org/info/rfc8084.

Оглавление

1. Введение

Термин «автоматический выключатель» (Circuit Breaker или CB) происходит из систем электроснабжения и никак не связан с сетевыми или виртуальными устройствами. В электроснабжении CB предназначены для защиты в крайних случаях. При нормальных условиях CB не должны срабатывать, они предназначены для защиты сети питания и

[1]Internet Engineering Task Force - комиссия по решению инженерных задач Internet.
[2]Internet Engineering Steering Group - комиссия по инженерным разработкам Internet.

подключённого оборудования от перегрузок. Люди не ожидают срабатывания СВ (или предохранителя) в своём доме, за исключением случаев неисправностей в проводке или проблем с электроприборами.

В сетях принцип СВ может использоваться как защитный механизм для крайних случаев, чтобы избежать сохраняющейся перегрузки, влияющей на другие потоки, использующие ту же сеть. Постоянные перегрузки были частыми на ранних этапах Internet в 1980 гг и приводили к тому, что избыточный трафик лишал другие соединения доступа в Internet. Этому противостояло требование использовать контроль перегрузок (congestion control или CC) в протоколе управления передачей (Transmission Control Protocol или TCP) [Jacobson88]. Эти механизмы работают на хостах Internet, вызывая «отключение» соединений TCP при перегрузке. Добавление контроля перегрузки в TCP (документировано в [RFC5681]) обеспечило стабильность Internet за счёт обнаружения и своевременного реагирования на перегрузки. Это было эффективно в Internet, где большинство потоков TCP были долгоживущими (это позволяло обнаруживать перегрузки и реагировать на них до прерывания потока). Хотя трафик TCP до сих пор был доминирующим, картина меняется и трафик без контроля перегрузки, включая трафик многих приложений на основе протокола пользовательских дейтаграмм (User Datagram Protocol или UDP), может составлять значительную часть общего трафика на канале. Для предотвращения сохраняющейся перегрузки в современной сети Internet требуется рассмотрение способов пересылки трафика без контроля перегрузок.

СВ для сетевого транспорта - это автоматический механизм, служащий для постоянного отслеживания потока или агрегата потоков. Механизм стремится обнаружить сохраняющуюся чрезмерную нагрузку в потоках. При обнаружении перегрузки СВ прерывает поток (или существенно снижает его скорость). Это является безопасной мерой предотвращения истощения сетевых ресурсов, препятствующего доступу других потоков в Internet. Такие меры необходимы для гетерогенной сети Internet и трафика, который сложно предвидеть заранее. Предотвращение сохраняющейся перегрузки важно для снижения вероятности «коллапса перегрузки» (Congestion Collapse) [RFC2914].

Имеются важные различия между транспортным СВ и методом контроля перегрузок. Контроль перегрузки (реализованный в TCP, SCTP[1] и DCCP[2]) работает в масштабе времени порядка интервала кругового обхода (Round-Trip Time или RTT) - времени передачи пакета от отправителя к адресату и обратно. Перегрузку на сетевом канале можно обнаружить с использованием явных уведомлений о перегрузке (Explicit Congestion Notification или ECN) [RFC3168], позволяющих сети указывать перегрузку путём установки в пакетах с поддержкой ECN маркеров перегрузки (Congestion Experienced или CE). Потери или получение пакетов с маркером CE считаются индикацией перегрузки. Методы контроля перегрузок способны реагировать на такие события, постоянно адаптируясь для снижения скорости передачи. Цель обычно состоит в ограничении скорости передачи до максимального значения, отражающего беспристрастное использование доступной на сетевом пути пропускной способности. Эти методы обычно работают на уровне отдельного потока трафика (задаётся квинтетом из адресов IP, протокола и портов).

СВ рекомендуются для потоков Internet без контроля перегрузок, например, для трафика, передаваемого с использованием сетевого туннеля. Они работают в масштабе времени, многократно превыщающем RTT, и срабатывают при аномальной (чрезмерной) перегрузке. Люди часто применяли то, что в этом документе называется СВ, в качестве специализированных мер защиты трафика Internet. Этот документ даёт рекомендации по развёртывания и использованию таких механизмов. В последующих параграфах приведены примеры, где СВ могут быть или не быть желательны.

СВ должен измерять некоторую часть трафика для обнаружения перегрузки в сети и надёжно срабатывать при наличии сохраняющейся перегрузки.

Триггер СВ часто использует серию последовательных измерений в точке входа и точке выхода (любая из которых может быть конечной точкой транспорта). Триггер должен работать в масштабе времени, многократно превышающем RTT (например, от единиц до многих десятков секунд). Длительный период нужен для обеспечения достаточного времени для работы транспортного контроля перегрузки или приложения с целью корректировки скорости при возникновении перегрузки, а также для стабилизации сети после корректировки. Перегрузки могут быть обычным явлением, когда транспорт с контролем перегрузки используется через канал, занятый почти полностью. Каждое событие перегрузки ведёт к снижению скорости транспортного потока, столкнувшегося с перегрузкой. Более длинный период позволяет избавиться от ненужных срабатываний СВ после 1 или даже нескольких событий перегрузки.

При срабатывании СВ должен обеспечить функцию управления (реакция), которая исключает трафик из сети путём запрета потока или существенного снижения уровня трафика. Такая реакция обеспечивает защиту и предотвращает сохраняющуюся перегрузку для других потоков, проходящих (частично) по тому же пути через сеть.

В разделе 4 приведены требования к построению СВ.

Рабочие условия, вызывающие срабатывание СВ, следует считать аномальными. Примеры ситуаций, которые могут вызывать срабатывание СВ включают:

- аномальный трафик, превышающий предоставленную пропускную способность (или имеющий характеристики, превышающие порог, настроенный для СВ);
- трафик, генерируемый приложением в то время, когда выделенная пропускная способность используется для других целей;
- изменение маршрутизации, вызывающее передачу дополнительного трафика по пути, отслеживаемому СВ;
- некорректная настройка службы или сетевого устройства, когда доступной пропускной способности не хватает для поддержки текущего агрегата трафика;
- некорректная настройка контроллера доступа или ограничителя трафика, позволяющая передавать больше трафика, чем ожидалось на пути, отслеживаемом СВ.

Сетевым операторам могут быть доступны иные механизмы обнаружения перегрузки (например, избыточная загрузка порта на сетевом устройстве). Используя такие сведения, операционные механизмы могут реагировать, снижая сетевую нагрузку на время меньше, чем у транспортного СВ. Роль СВ на таких путях остаётся крайней мерой.

[1]Stream Control Transmission Protocol - протокол управления потоковой передачей.
[2]Datagram Congestion Control Protocol - протокол контроля перегрузки для дейтаграмм.

Поскольку СВ работает с более долгими интервалами, ему следует срабатывать лишь при отказе иных мер реагирования на перегрузку.

Во многих случаях причина срабатывания СВ не будет очевидна для источника трафика (пользоватя, приложения, конечной точки и т. п.). СВ можно применять для ограничения трафика от приложений, которые не могут или не хотят использовать контроль перегрузок, а также в случаях, нельзя полагаться на свойства контроля перегрузок (например, трафик передаётся через сетевой туннель). В таких ситуациях для СВ практически невозможно передать сигнал затронутым приложениям. Поэтому у приложений могут в некоторых случаях возникать сложности с определением срабатывания СВ и места в сети, где это произошло.

Разработчикам приложений рекомендуется по возможности разворачивать подходящие механизмы контроля перегрузок. Приложения, использующие контроль перегрузок будут знать о событиях перегрузки в сети. Это позволяет им регулировать нагрузку в условиях насыщения и, как ожидается, избегать срабатывания сетевых СВ. Для приложений, способных создавать эластичный трафик, это часто будет предпочтительным решением.

1.1. Типы СВ

Имеются разные формы СВ для сетевого транспорта, различающиеся в основном временем срабатывания, а также предоставляемой защитой.

Fast-Trip СВ - СВ с быстрым переключением
: Этот тип СВ использует сравнительно короткий масштаб времени для защиты сетевого трафика от одного потока или группы связанных потоков.

Slow-Trip СВ - СВ с медленным переключением
: Этот тип СВ использует более продолжительный масштаб времени для защиты сетевого трафика от перегрузки, вызываемой агрегатами трафика.

Managed СВ - управляемые СВ
: Использует операции и функции управления, которые могут присутствовать в управляемой службе, для реализации СВ.

Примеры всех типов СВ представлены в разделе 4.

2. Уровни требований

Ключевые слова **должно** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не следует** (SHALL NOT), **следует** (SHOULD), **не нужно** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **не рекомендуется** (NOT RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе интерпретируются в соответствии с [RFC2119].

3. Устройство СВ

Хотя СВ много лет обсуждались в IETF, до сих пор не было рекомендаций по применению и устройству СВ. В этом документе предпринята попытка обсудить эти вопросы.

СВ **рекомендуются** для протоколов IETF и туннелей, предающих потоки Internet без контроля перегрузки и агрегаты трафика, включая трафик через сетевой туннель. Разработчики других протоколов и туннельной инкапсуляции также должны рассмотреть использование этих методов как крайней меры для защиты трафика на общем пути через сеть.

Этот документ задаёт требования к устройству СВ и содержит примеры конструкций. В спецификациях протоколов и туннельной инкапсуляции нужно указывать детали протокольных механизмов, требуемых для реализации СВ. В параграфе 3.1 указаны функциональные компоненты СВ, а в параграфе 3.2 заданы требования к реализации СВ.

3.1. Функциональные компоненты

Базовая конструкция СВ включает связь между точками входа (отправитель) и выхода (получатель) сетевого потока или набора потоков. Простая иллюстрация работы приведена на рисунке 1, где показан набор маршрутизаторов (R), соединяющих множество конечных точек (Endpoint).

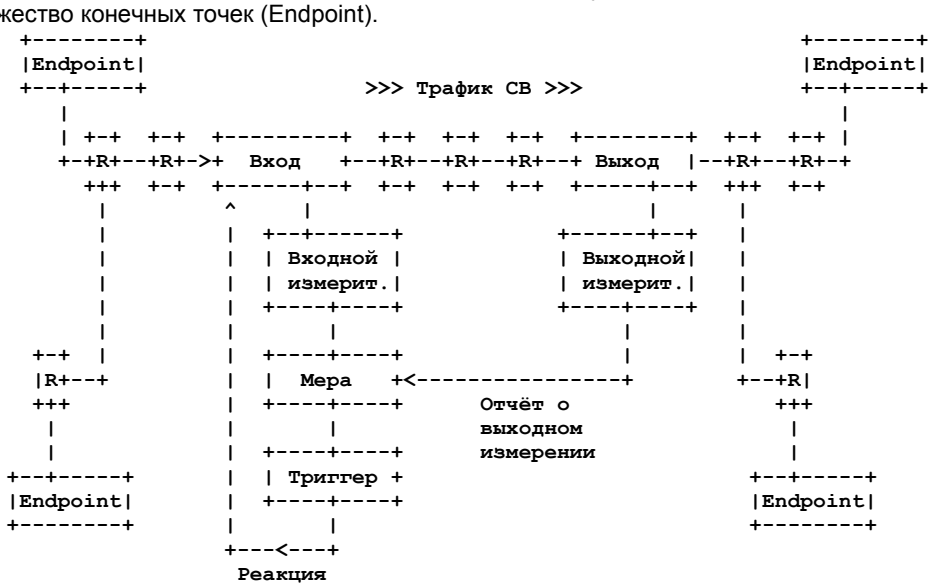

Рисунок 1. СВ, контролирующий часть сквозного пути между точками входа и выхода. В некоторых случаях функции триггера и измерений могут размещаться в других местах (например, в операционном центре сети).

СВ служит для управления сетевым трафиком, проходящим через часть этих маршрутизаторов, расположенных между сетевыми устройствами точек входа и выхода. Путь между входом и выходом может проходить через туннель или иное решение сетевого уровня. Одним из ожидаемых вариантов является применение на входе и выходе службы, где весь рассматриваемый трафик завершается за точкой выхода, поэтому на входе и выходе присутствуют общие потоки.

В контексте СВ входные и выходные функции могут быть реализованы в разных местах. Например, они могут размещаться в сетевых устройствах на входе и выходе туннеля. В некоторых случаях они могут размещаться на одной или обеих конечных точках сети (Рисунок 2) и могут быть реализованы как компоненты транспортного протокола.

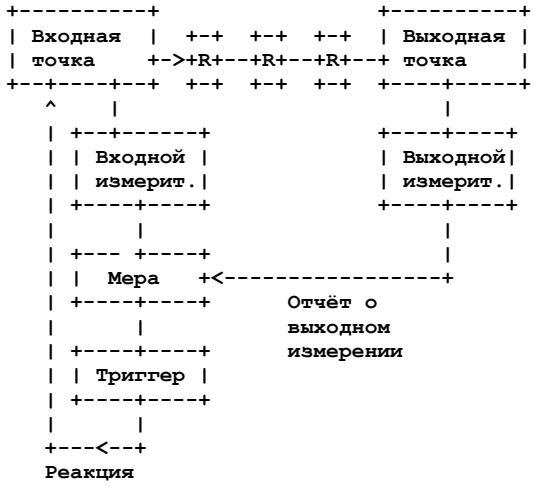

Рисунок 2. СВ конечной точки, реализованный у отправителя и получателя.

Ниже указаны компоненты, требуемые для реализации СВ.

1. Входной измеритель (у отправителя или на входе туннеля), записывающий число пакетов/байтов, переданных в каждом интервале измерения. Это указывает сетевую нагрузку для потока или набора потоков трафика. Например, интервал измерения может составлять много секунд (измерение каждые несколько десятков секунд или серия коротких последовательных измерений, объединяемых функцией измерения СВ).

2. Выходной измеритель (у получателя или на выходе туннеля), записывающий число пакетов/байтов, полученных в каждом интервале измерения. Он указывает поддерживаемую нагрузку для потока или набора потоков и может использовать другие сигналы обнаружения перегрузки (например, потери или маркировку [RFC3168] на пути). Измерения на выходе могут синхронизироваться (включая смещение для времени прохождения данных или привязку измерений к конкретному пакету), чтобы все счётчики относились к одному диапазону пакетов.

3. Метод передачи измеренных на входе и выходе значений функции СВ Measurement. Для этого можно применять несколько методов, включая передачу откликов на пакеты измерений (управляющих сообщений) от получателя для запуска функции у отправителя, реализацию с использованием OAM[1], передачу сигнальных дейтаграмм по основному каналу. Можно также реализовать это полностью как функцию плоскости управления, например, с использованием контроллера программно-определяемой сети.

4. Функция измерения, объединяющая входные и выходные измерения для оценки текущего уровня перегрузки сети (например, частота потерь в каждом интервале измерения может быть выведена из разности значений входных и выходных счётчиков). Отметим, что метод не требует высокой точности для измерительного интервала (и для измеренных значений, поскольку изолированные и редкие потери нужно игнорировать).

5. Функция триггера, которая определяет, указывают ли измерения сохраняющуюся перегрузку. Эта функция задаёт подходящий порог для определения наличия сохраняющейся перегрузки между входом и выходом. При этом предпочтительно учитывать скорость или отношение, а не абсолютное значение (например, потери выше 10% или иные значения, основанные на скорости передачи и частоте потерь). СВ переключается, когда порог превышен в течение нескольких интервалов измерения (например, при 3 последовательных измерениях). Функция должна быть надёжной, чтобы отдельные или ложные события не вызывали реакции.

6. Реакция, применяемая на входе в случае переключения СВ. Реакция стремится автоматически исключить трафик, вызвавший сохраняющуюся перегрузку.

7. Механизм управления с обратной связью, который срабатывает при недоступности входных или выходных измерений, поскольку это может указывать потерю пакетов управления (признак сильной перегрузки или неспособности управлять нагрузкой).

3.2. Другие топологии сетей

СВ могут применяться в сетях с топологией, отличающейся от показанной на рисунках 1 и 2. В этом параграфе приведены примеры такого использования и указаны возможные места реализации функций.

3.2.1. Использование с групповым протоколом управления/маршрутизации

На рисунке 3 показан пример реализации группового (multicast) СВ на паре групповых конечных точек (например, для реализации Fast-Trip СВ, см. параграф 5.1). Входная конечная точка (отправитель группового трафика) измеряет нагрузку на входе, генерирует результаты измерений (например, записывает счётчики пакетов с временными метками) и передаёт их в multicast-группу вместе с измеряемым трафиком.

[1]Operations, Administration and Management - операции, администрирование, управление.

Маршрутизаторы на групповом пути пересылают multicast-трафик (включая входные измерения) всем активным получателям в конечных точках. Каждый маршрутизатор последнего интервала (last hop, выход) пересылает трафик одной или нескольким выходным конечным точкам.

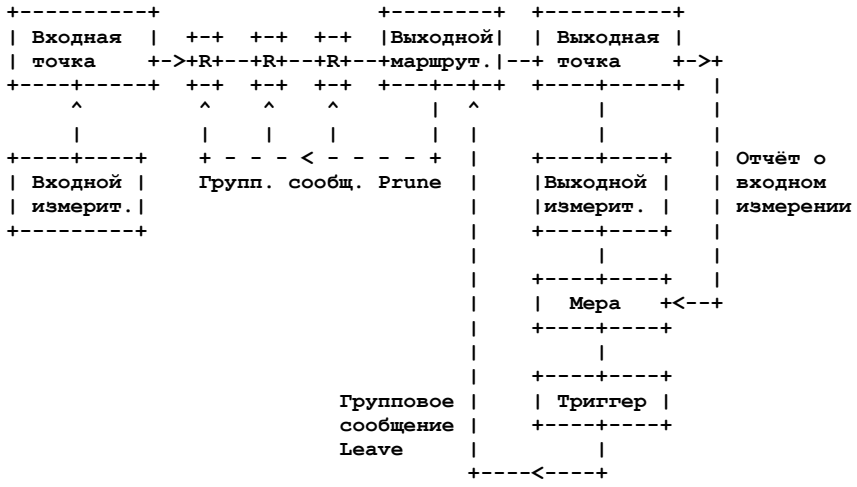

Рисунок 3. Пример группового СВ, контролирующего сквозной путь между входом и выходом.

На рисунке 3 каждая конечная точка включает измеритель для локальной выходной нагрузки. Конечная точка также извлекает из трафика полученные входные измерения и сравнивает входные и выходные значения для определения необходимости переключения СВ. Эти измерения должны быть устойчивы к потерям (см. предыдущий параграф). При переключении СВ генерируется групповое сообщение Leave (выход из группы) для выходов (например, передаётся сообщение IGMP или MLD маршрутизатору last-hop), которое заставляет маршрутизатор восходящего направления прервать пересылку трафика выходной конечной точке [RFC1112].

Групповые маршрутизаторы, не имеющие активных получателей для конкретной группы, будут пересылать своему маршрутизатору восходящего направления сообщение Prune (обрезка). Это запускает процесс освобождения пропускной способности, используемой трафиком, являющийся стандартной функцией групповой маршрутизации (например, использование протокола PIM-SM[1] [RFC7761]). Каждый выход работает автономно и «реакция» СВ происходит в групповой плоскости управления (например, через PIM), не требуя явных сигналов от СВ по пути, используемому для управляющих сообщений. Отметим, что прямой связи со входом нет, поэтому переключение СВ влияет лишь на нисходящий трафик первого группового маршрутизатора (first-hop multicast). Переключение не останавливает трафик от источника к маршрутизатору first-hop, это обычная практика для групповой передачи.

Метод подходит также для групповых туннелей или подсетей (см., например, параграфы 5.2, 5.3), где измеритель на входе генерирует дополнительные управляющие сообщения, переносящие данные измерений в сторону выхода, где реализовано выходное измерение.

3.2.2. Использование с протоколами управления, поддерживающими полосу

Некоторые пути предоставляются с использованием протокола управления, например, потоки, предоставляемые службами многопротокольной коммутации по меткам (Multiprotocol Label Switching или MPLS), пути, предоставляемые с использованием протокола резервирования ресурсов (Resource Reservation Protocol или RSVP), сети, использующие функции SDN[2], или дифференцированные услуги (Differentiated Service) с контролем доступа. На рисунке 1 показан один из ожидаемых вариантов использования, где для измерения и переключения может применяться отдельное устройство. Реакция, вызываемая триггером может быть сообщением управления сетью, передаваемым на вход и/или другим элементам сети, для выполнения соответствующих действий по сигналу СВ. Примеры этого варианта применения представлены в параграфе 5.3. Управляемые СВ.

3.2.3. Однонаправленные СВ на контролируемых путях

СВ можно применять для управления однонаправленным трафиком UDP при наличии пути передачи управляющих сообщений, связывающих функциональные компоненты на входе и выходе. Этот путь для управляющих сообщений может существовать в сети, где поток трафика является односторонним. Например, групповой поток передаёт пакеты через путь Internet и может применять групповую маршрутизацию для отсечки потоков с целью снижения нагрузки на сеть. Некоторые другие типы подсетей также применяют протоколы управления для контроля потоков трафика.

4. Требования к СВ для сетевого транспорта

Ниже приведены требования к реализации СВ.

1. Нужен коммуникационный путь для управляющих сообщений с данными измерения от входного и выходного измерителя в точку измерения (требования 16-18 относятся к передаче управляющих сообщений).
2. СВ **требует** определить период измерения, в течение которого функция СВ Measurement определяет уровень перегрузки или потерь. Этот метод не обнаруживает потери отдельных пакетов, но **должен** иметь способ о потере и/или маркировке пакетов из потока трафика.
3. Измеритель может считать маркеры ECN [RFC3168] CE для измерения перегрузки, но в этом случае **должны** измеряться и потери для полного представления об уровне перегрузки. Для туннелей в [CONGESTION-FEEDBACK] описан способ учёта потерь и маркеров ECN, такие измерения могут применяться на сравнительно коротких интервалах времени для управления реакцией контроля перегрузок и/или

[1]Protocol Independent Multicast - Sparse Mode - независимая от протокола групповая передача - разреженный режим.
[2]Software-Defined Network - программно-определяемая сеть.

агрегирования за больший интервал с более высоким порогом срабатывания CB. В последующих пунктах обсуждается необходимость использовать более длинные интервалы и более высокие пороги триггера.

4. Используемый функцией CB Measurement период измерения **должен** быть больше времени, требуемого текущим алгоритмам контроля перегрузок для снижения скорости при обнаружении перегрузки. Это важно, поскольку алгоритмам сквозного контроля перегрузок требуется время не менее 1 интервала RTT для уведомления и регулировки трафика в случае перегрузки, а узкие места на пути могут использоваться для трафика с разными сквозными значениями RTT. Поэтому предполагается период измерения существенно больше значения RTT, наблюдаемого самим CB.

5. При необходимости CB **может** объединять последовательные выборки со входа и выхода, чтобы получить усреднённое значение за достаточно долгий интервал. Отметим, что при необходимости объединять выборки комбинация должна отражать сумму отдельных значений выборок, делённую на общее время (объем) измерений. Выборки из разных интервалов нельзя объединять напрямую для получения среднего значения.

6. Устройство CB **должно** предотвращать срабатывание при слабой или прерывистой перегрузке (см. 7 - 9).

7. От CB **требуется** установка порога для определения события чрезмерной перегрузки.

8. От CB **требуется** установка интервала переключения, задающая период в течение которого триггер использует собранные измерения. CB должны срабатывать на достаточно долгое время для предотвращения дополнительного «наказания» для потоков с большим значением RTT на пути.

9. CB **должен** быть устойчив к многочисленным событиям перегрузки. Обычно это определяет число наблюдаемых событий сохраняющейся перегрузки в течение периода срабатывания. Например, CB **может** комбинировать результаты нескольких периодов измерения для определения срабатывания CB (например, срабатывание лишь в случае сохраняющейся перегрузки, обнаруженной при трёх измерениях в интервале срабатывания, когда было собрано более 3 измерений).

10. При нормальной реакции на триггер **следует** отключать весь трафик, вызывающий перегрузку (иначе 11 и 12).

11. Реакция **должна** быть намного более жёсткой, чем у алгоритма контроля перегрузок (такого как алгоритм контроля перегрузок TCP [RFC5681] или TFRC[1] [RFC5348]), поскольку CB реагирует на более длительную перегрузку и работает в большем масштабе времени (т. е. перегрузка сохраняется дольше перед срабатыванием CB).

12. При реакции, приводящей к снижению нагрузки, **следует** сокращать трафика как минимум на порядок. Отклик путём прерывания потоков вместо случайного отбрасывания пакетов зачастую предпочтителен для пользователей службы. CB, снижающий скорость потока, **должен** продолжать отслеживать уровень перегрузки и **должен** дополнительно снижать скорость, если CB срабатывает снова.

13. Реакция на срабатывание CB **должна** сохраняться в течение времени не меньше интервала срабатывания. Обычно для восстановления трафика требуется вмешательство оператора. Если нужен автоматизированный отклик для сброса триггера, он не обязан быть незамедлительным. Автоматизированный механизм сброса **следует** делать достаточно консервативным, чтобы он не влиял на другие механизмы (включая другие алгоритмы CB, контролирующие трафик по общему пути). **Не следует** выполнять автоматизированный сброс при наличии признаков продолжающейся перегрузки.

14. Срабатывание триггера CB **следует** считать аномальным сетевым событием, которое **следует** записывать в журнал (log). Измерения, ведущие к срабатыванию CB, также **следует** записывать в журнал.

15. Управляющие коммуникации должны передавать измерения (п. 1), а в некоторых случаях нужна передача сообщений триггера на вход. Эти коммуникации могут работать по основному или отдельному каналу. Связь по основному каналу **рекомендуется**, когда возможен любой вариант устройства. Предпочтительным устройством CB является срабатывание при невозможности получить отчёты об измерениях, показывающие отсутствие перегрузки, в отличие от успешной передачи сигнала о перегрузке отправителю (этот сигнал может быть потерян при перегрузке).

    In Band - по основному каналу
    При управлении по основному каналу **следует** считать потерю управляющего сообщения индикацией возможной перегрузки на пути, а повторяющиеся потери должны вызывать срабатывание CB. Это решение обеспечивает «общую судьбу» для потоков трафика и управляющих сообщений. Эта общность ослабевает если весь или часть измеренного трафика передаётся по пути, отличающемуся от пути трафика управления (например, разные пути задаёт маршрутизация по множеству равноценных путей, организация трафика или туннели для определённых типов трафика).
    Out of Band - по отдельному каналу
    При управлении по отдельному каналу **не следует** вызывать реакцию CB при потере управляющих сообщений (например, измерений). Это предотвратит усиление/распространение отказов при независимых отказах на путях измерения и данных. Отказ на отдельном канале **следует** считать аномальным событием в сети и соответственно обрабатывать. Например, такие события **следует** записывать в журнал и оператор может выполнять дополнительные действия в зависимости от сети и вовлечённого трафика.

16. Передача управляющих сообщений **должна** быть устойчива к потере пакетов. Управляющие сообщения могут теряться при отказе в пути и это ожидается при перегрузке на пути. Это не предполагает желательность гарантированной доставки (например, через TCP), поскольку это может вносить дополнительную задержку при откликах на перегрузку. Подходящим механизмом может быть дублирование управляющих сообщений для повышения устойчивости к потерям и/или отказ от управляющего трафика как индикации продолжающейся перегрузки [RFC8085]. Если управляющие сообщения передаются по общему пути, **рекомендуется** приоритизировать этот трафик для снижения вероятности потерь при перегрузке. Трафик управления следует также учитывать при приоритизации сети, использующей CB.

[1] TCP-Friendly Rate Control - дружественный к TCP контроль скорости.

17. Имеются требования к защите управляющих сообщений между конечными точками и/или сетевыми устройствами (раздел 7). Подлинность источника и целостность управляющих сообщений (измерения и триггеры) **должны** защищаться от атак извне пути. При наличии риска атак в пути рекомендуется применять механизм криптографической аутентификации для всех сообщений управления и измерений.

5. Примеры CB

Имеется множество типов CB, которые могут применяться в разных вариантах развёртывания. В некоторых случаях поток может контролироваться несколькими CB (например, трафик сквозного потока, передаваемый через туннель в сети). В этом разделе представлены примеры разных типов CB.

5.1. Быстрый CB

В [RFC2309] обсуждается опасность не реагирующих на перегрузку потоков и сказано: «во все потоковые приложения UDP следует включать эффективные механизмы предотвращения перегрузок». Некоторые приложения не применяют полнофункциональный транспорт (TCP, SCTP, DCCP) и они (например, при использовании UDP или UDP-Lite) должны обеспечивать надлежащее предотвращение перегрузки. Рекомендации для приложений, не использующих транспорт с контролем перегрузок, приведены в [RFC8085]. Механизмы могут быть устроены так, чтобы реагировать в более коротком масштабе времени, чем CB, где контролируется лишь огибающая трафика (envelope). Методы контроля перегрузок могут взаимодействовать с приложением для более эффективного управления скоростью передачи.

Fast-trip CB является наиболее быстродействующим вариантом CB и обеспечивает время отклика, лишь незначительно превышающее время изменения контролируемого трафика. Метод подходит для трафика с понятными характеристиками (и может включать одну или несколько триггерных функций, адаптированных к типу трафика). Метод не подходит для произвольного сетевого трафика и может не подходить для агрегатов трафика, поскольку может срабатывать преждевременно (например, когда сочетание трафика нескольких потоков с контролем перегрузок вызывает кратковременную перегрузку).

Хотя механизм может быть реализован в сетевых устройствах с поддержкой RTP, он подходит и для реализации в конечных точках (например, как часть транспортной системы), где он может дополнять методы сквозного контроля перегрузок. Быстрый отклик позволяет таким CB переключаться быстрее других вариантов (например, CB, работающих с агрегатами трафика в точках сетевого пути).

5.1.1. Быстрый CB для RTP

Был определён набор методов Fast-Trip CB для использования с потоками протокола доставки в реальном масштабе времени (Real-time Transport Protocol или RTP), использующими профиль RTP/AVP [RFC8083]. Предполагается, что при отсутствии значительной перегрузки все приложения RTP работающие в сетях IP best-effort могту работать без переключения таких CB. Поэтому RTP Fast-Trip CB реализован для устойчивости к отказам (fail-safe) и срабатывание ведёт к прерыванию трафика RTP.

Конечная точка отслеживает получение по основному каналу блоков с отчётами протокола управления RTP (RTP Control Protocol или RTCP), содержащихся в пакетах отчётов отправителя (sender report или SR) или получателя (receiver report или RR), которые обеспечивают обратную связь для качества приёма. Это служит для измерения потерь (перегрузка), возможно, в сочетании с ECN [RFC6679].

Действие CB (остановка потока) вызывается при выполнении любого из указанных ниже условий.

1. RTP CB срабатывает при отчёте об отсутствии «продвижения» (progress).
2. RTP CB срабатывает при отсутствии принятых отчётов от получателя.
3. RTP CB срабатывает, когда долгосрочная пропускная способность RTP (за множество RTT) превышает жёстко заданный верхний порог, определяемый методом, напоминающим TFRC.
4. RTP CB включает понятие пригодности среды (Media Usability) и срабатывает при снижении качества среды передачи ниже некого минимально допустимого уровня.

5.2. Медленный CB

Slow-Trip CB можно реализовать на конечной точке или сетевом устройстве. Этот тип CB гораздо медленней реагирует на перегрузку, нежели Fast-Trip CB. Предполагается, что он будет применяться более часто.

Одним из случаев, когда требуется Slow-Trip CB, является использованием потоком или агрегатом потоков туннеля или инкапсуляции, при этом не все потоки в туннеле поддерживают контроль перегрузок в стиле TCP (например, TCP, SCTP, TFRC), см. параграф 3.1.3 в [RFC8085]. Это может возникать при развёртывании туннелей в сети Internet (а не в «контролируемой среде» провайдера Internet или корпоративной сети), особенно при прохождении туннеля через клиентский маршрутизатор доступа.

5.3. Управляемые CB

Управляемые CB реализуются в сигнальном протоколе или плоскости управления, связанных с контролируемым агрегатом трафика. Этот тип CB обычно применяется в «контролируемых средах».

CB требуется больше, чем просто способность определить, что сетевой путь пересылает данные, или измерить скорость на пути, что часто является обычными рабочими функциями сети. Необходимо определять уровень перегрузки сети на пути и инициировать реакцию при превышении порога, указывающего сохраняющуюся перегрузку.

Можно использовать управляющие сообщения, передаваемые по основному или дополнительному каналу.

5.3.1. Управляемый CB для псевдопроводов SAToP

В разделе 8 [RFC4553] SAToP Pseudowire Emulation Edge-to-Edge (PWE3) описан пример управляемого CB для изохронных потоков. Если такой поток идёт через заранее подготовленную инфраструктуру (например, MPLS), можно

предполагать, что в PW не будет возникать перегрузок, поскольку ожидается, что скорость потока не будет расти или снижаться. Если же трафик PW мультиплексируется с другим трафиком в Internet, перегрузки могут возникать. В [RFC4553] сказано: «Если SAToP PW работают через PSN с услугами best-effort, им **следует** отслеживать потери пакетов для обнаружения серьёзных перегрузок.» В настоящее время рекомендуемый интервал измерения составляет 1 сек. и триггер срабатывает при наблюдении не менее трёх SES[1] в интервале измерения. В [RFC4553] также отмечено, что: «при обнаружении таких условий SAToP PW должен на некоторое время отключаться в обоих направлениях ...». Идея состояла в том, чтобы при превышении коэффициентом потери пакетов (packet-loss ratio, перегрузка) заданного порога PW по умолчанию отключался. Этот вариант применения рассматривал передачу с фиксированной скоростью, когда у PW нет разумного способа сбросить нагрузку.

Триггер должен устанавливаться на частоту, при которой PW очевидно сталкивается с серьёзной перегрузкой, которая может делать обслуживание неприемлемым. В этот момент срабатывание CB удалит трафик, предотвращая неприемлемое влияние на трафик, реагирующий на перегрузки (например, TCP). Частично обоснование заключалось в том, что высокие потери обычно указывали на какую-либо «поломку» и должны были приводить к вмешательству оператора, поэтому нужно инициировать такое вмешательство.

Вовлечение оператора при срабатывании CB даёт возможность выполнить иные действия по восстановления качества обслуживания (например, отключение другой нагрузки или добавление пропускной способности) или сознательно не реагировать на срабатывание, пока не найдено инженерное решение проблемы. Это может потребовать от триггерной функции передачи управляющего сигнала третьей стороне, например, в центр управления сетью (network operations center или NOC), отвечающей за работу входа в туннель, а не за сам этот вход.

5.3.2. Управляемый CB для псевдопроводов

Псевдопровода (pseudowire или PW) [RFC3985] стали распространённым механизмом туннелирования трафика и могут конкурировать за сетевые ресурсы с другими PW и прочим (не PW) трафиком, таким как потоки TCP/IP.

В [RFC7893] рассмотрены условия перегрузки, которые могут возникать при конкурировании PW с эластичным (т. е. реагирующим на перегрузки) трафиком (например, TCP). Эластичные PW, передающие трафик IP (см.[RFC4448]) не вызывают серьёзных проблем, поскольку вовлечённый в них трафик реагирует на перегрузки в сети снижением скорости передачи.

Неэластичные PW (например, TDM[2] [RFC4553] [RFC5086] [RFC5087] с фиксированной пропускной способностью) могут препятствовать реагирующему на перегрузки трафику или способствовать перегрузке, поскольку они не меняют свою скорость в ответ на перегрузки. В [RFC7893] проанализированы TDM PW и сделан вывод, что TDM PW, работающие с уровнем потерь, который может создавать связанные с перегрузками проблемы, будут сталкиваться с неприемлемым качеством работы служб TDM. Поэтому данный документ предполагает, что управляемый CB, отключающий PW при сохраняющейся невозможности предоставить приемлемые услуги TDM, может быть полезным способом решения проблемы перегрузок (см. Приложение A к [RFC7893]).

6. Примеры, где CB могут не требоваться

CB не требуется для одиночного потока с контролем перегрузок, использующего TCP, SCTP, TFRC и т. п. В этих случаях методы контроля перегрузок уже предотвращают сохраняющиеся перегрузки.

6.1. CB на путях с заранее предоставленной полосой

Один из основных является вопрос о целесообразности CB при развёртывании туннеля в частной сети с заранее выделенной пропускной способностью. В этом случае трафик, не выходящий за пределы предоставленной пропускной способности, не должен вызывать сохраняющейся перегрузки, поэтому CB будет срабатывать лишь при возникновении несоответствующего трафика. Можно утверждать, что этого не должно происходить, но можно также сказать, что CB никогда не будет срабатывать. При реализации CB он будет обеспечивать соответствующий отклик, если в работающей сети возникнет сохраняющаяся перегрузка.

Внедрение CB не снизит производительность потоков, но при возникновении сохраняющейся перегрузки это защитит сетевой трафик, использующий общую полосу с перегружающими сеть потоками. Это также защитит сетевой трафик от отказов, когда трафик CB (пере)маршрутизируется, создавая дополнительную сетевую нагрузку на непредусмотренном пути.

6.2. CB в туннелях для трафика с контролем перегрузок

Для трафика IP обычно предполагается наличие контроля перегрузок, т. е. считается, что транспортные протоколы, создающие трафик IP у отправителя, уже реализуют механизмы, достаточные для решения проблем перегрузки на пути. Поэтому при развёртывании туннелей, в которых предполагается передача лишь агрегата трафика TCP (или иного трафика с контролем перегрузок), возникает вопрос о целесообразности применения CB.

Предполагается, что трафик TCP (и SCTP) в туннеле будет снижать свою скорость передачи при обнаружении перегрузки. Другой транспорт (например, UDP) может реализовать механизмы, достаточные для решения проблемы перегрузок на пути [RFC8085]. Однако даже при реализации в каждом потоке, проходящем через туннель, механизма контроля перегрузок общий трафик агрегата потоков все равно может столкнуться с перегрузкой. Например, большинству механизмов контроля перегрузок требуются долгосрочные потоки для реагирования путём снижения скорости. В агрегате может быть много краткосрочных потоков, завершающихся до возникновения перегрузки.

Зачастую поставщик туннельных услуг не знает, что туннель включает лишь трафик с контролем перегрузки (например, проверка заголовков пакетов может быть невозможна). Некоторые приложения IP могут не реализовать адекватных механизмов для решения проблемы перегрузок. Важно отметить, что если агрегат трафика не создаёт сохраняющейся перегрузки (влияющей на другие потоки), CB не будет срабатывать. Это предполагается в данном контексте, поэтому реализация CB не должна снижать производительность туннеля, но при возникновении сохраняющейся перегрузки CB защитит остальной трафик сети, использующий общую пропускную способность с туннельным трафиком.

[1]Severely Errored Second - секунда со значительным числом ошибок.
[2]Time Division Multiplex - мультиплексирование с разделением по времени.

6.3. СВ с односторонним трафиком без пути управления

Односторонний путь пересылки может не иметь связанного пути для пересылки управляющих сообщений, поэтому его нельзя контролировать с помощью СВ (сравн. с 3.2.3. Однонаправленные СВ на контролируемых путях).

Односторонняя услуга может предоставляться с использованием пути с предварительно обеспеченной пропускной способностью, которая не используется совместно с другими эластичными потоками Internet (например, потоками, меняющими свою скорость). Путь пересылки может также использоваться совместно с другими потоками. Одним из путей смягчения влияния трафика на другие потоки является контроль огибающей трафика с помощью правил на входе. Для поддержки этого типа трафика в Internet требуется отслеживание оператором с целью обнаружения сохраняющейся перегрузки и реагирования на неё.

7. Вопросы безопасности

Все механизмы СВ полагаются на координацию между входными и выходными измерителями и взаимодействие с функцией триггера. Обычно это достигается передачей данных управления сетью (или протокольных сообщений) через сеть. Своевременность операций СВ зависит от выбора периода измерений. Слишком длинный интервал у получателя снижает отзывчивость СВ и это влияет на способность СВ обнаруживать перегрузку и реагировать на неё. При слишком коротком интервале СВ будет срабатывать преждевременно, не оставляя времени на срабатывание других механизмов, что может приводить к ненужному прерыванию обслуживания.

Злоумышленник может воспользоваться СВ для организации DoS[1]-атак на трафик, контролируемый СВ. Поэтому нужны механизмы предотвращения атак на данные управления сетью, которые могут вести к DoS-атаке.

Подлинность источника и целостность сообщений протокола (измерения и триггеры) **должны** защищаться от атак извне пути. Без такой защиты злоумышленник легко может внедрить поддельные или изменённые сообщения контроля и/или измерений (например, указать высокий уровень потерь), вынуждающие СВ срабатывать, что позволяет организовать DoS-атаки нарушающие потоки.

Простую защиту можно обеспечить за счёт использования случайного порта источника или эквивалентного поля в заголовке пакетов (например, RTP SSRC или порядковый номер RTP), которые считаются неизвестными атакующему извне пути. Более сильную защиту можно обеспечить с помощью защищённого протокола аутентификации.

Атаки на управляющие сообщения сравнительно просты для злоумышленника, размещающегося на пути передачи, если эти сообщения не шифруются и не аутентифицируются. Применение криптографического механизма проверки подлинности для всех сообщений управления и измерений **рекомендуется** для смягчения этой проблемы, а также обеспечит защиту от атак извне пути. Имеется компромисс между издержками на криптографическую защиту управляющих соединений и желанием защитить управляющие взаимодействия. Для некоторых вариантов развёртывания дополнительная защита от DoS-атак потребует аутентификации всех управляющих сообщений.

На передачу сообщений управления сетью расходуется пропускная способность. Такой трафик управления необходимо учитывать при разработке СВ и он может вносить вклад в перегрузку сети. При передаче такого трафика по общему пути **рекомендуется** приоритизировать его для снижения вероятности потерь при перегрузке. Трафик управления необходимо учитывать при настройке сети, использующей СВ.

СВ **должны** быть устойчивы к потере пакетов, которая может возникать при перегрузках. Потеря управляющих сообщений может быть побочным эффектом перегрузки сети или вызываться иными причинами (см. раздел 4).

Влияние на безопасность зависит от устройства механизмов, типа контролируемого трафика и предусмотренного варианта развёртывания. Поэтому каждое решение для СВ **должно** оценивать влияние конкретного механизма СВ на безопасность.

8. Литература

8.1. Нормативные документы

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, DOI 10.17487/RFC3168, September 2001, <http://www.rfc-editor.org/info/rfc3168>.

[RFC8085] Eggert, L., Fairhurst, G., and G. Shepherd, "UDP Usage Guidelines", BCP 145, RFC 8085, DOI 10.17487/RFC8085, March 2017, <http://www.rfc-editor.org/info/rfc8085>.

8.2. Дополнительная литература

[CONGESTION-FEEDBACK] Wei, X., Zhu, L., and L. Deng, "Tunnel Congestion Feedback", Work in Progress, draft-ietf-tsvwg-tunnel-congestion-feedback-04, January 2017.

[Jacobson88] Jacobson, V., "Congestion Avoidance and Control", SIGCOMM Symposium proceedings on Communications architectures and protocols, August 1988.

[RFC1112] Deering, S., "Host extensions for IP multicasting", STD 5, RFC 1112, DOI 10.17487/RFC1112, August 1989, <http://www.rfc-editor.org/info/rfc1112>.

[RFC2309] Braden, B., Clark, D., Crowcroft, J., Davie, B., Deering, S., Estrin, D., Floyd, S., Jacobson, V., Minshall, G., Partridge, C., Peterson, L., Ramakrishnan, K., Shenker, S., Wroclawski, J., and L. Zhang, "Recommendations on Queue Management and Congestion Avoidance in the Internet", RFC 2309, DOI 10.17487/RFC2309, April 1998, <http://www.rfc-editor.org/info/rfc2309>.

[RFC2914] Floyd, S., "Congestion Control Principles", BCP 41, RFC 2914, DOI 10.17487/RFC2914, September 2000, <http://www.rfc-editor.org/info/rfc2914>.

[1]Denial-of-Service - отказ в обслуживании.

[RFC3985] Bryant, S., Ed. and P. Pate, Ed., "Pseudo Wire Emulation Edge-to-Edge (PWE3) Architecture", RFC 3985, DOI 10.17487/RFC3985, March 2005, <http://www.rfc-editor.org/info/rfc3985>.

[RFC4448] Martini, L., Ed., Rosen, E., El-Aawar, N., and G. Heron, "Encapsulation Methods for Transport of Ethernet over MPLS Networks", RFC 4448, DOI 10.17487/RFC4448, April 2006, <http://www.rfc-editor.org/info/rfc4448>.

[RFC4553] Vainshtein, A., Ed. and YJ. Stein, Ed., "Structure-Agnostic Time Division Multiplexing (TDM) over Packet (SAToP)", RFC 4553, DOI 10.17487/RFC4553, June 2006, <http://www.rfc-editor.org/info/rfc4553>.

[RFC5086] Vainshtein, A., Ed., Sasson, I., Metz, E., Frost, T., and P. Pate, "Structure-Aware Time Division Multiplexed (TDM) Circuit Emulation Service over Packet Switched Network (CESoPSN)", RFC 5086, DOI 10.17487/RFC5086, December 2007, <http://www.rfc-editor.org/info/rfc5086>.

[RFC5087] Stein, Y(J)., Shashoua, R., Insler, R., and M. Anavi, "Time Division Multiplexing over IP (TDMoIP)", RFC 5087, DOI 10.17487/RFC5087, December 2007, <http://www.rfc-editor.org/info/rfc5087>.

[RFC5348] Floyd, S., Handley, M., Padhye, J., and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification", RFC 5348, DOI 10.17487/RFC5348, September 2008, <http://www.rfc-editor.org/info/rfc5348>.

[RFC5681] Allman, M., Paxson, V., and E. Blanton, "TCP Congestion Control", RFC 5681, DOI 10.17487/RFC5681, September 2009, <http://www.rfc-editor.org/info/rfc5681>.

[RFC6679] Westerlund, M., Johansson, I., Perkins, C., O'Hanlon, P., and K. Carlberg, "Explicit Congestion Notification (ECN) for RTP over UDP", RFC 6679, DOI 10.17487/RFC6679, August 2012, <http://www.rfc-editor.org/info/rfc6679>.

[RFC7761] Fenner, B., Handley, M., Holbrook, H., Kouvelas, I., Parekh, R., Zhang, Z., and L. Zheng, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", STD 83, RFC 7761, DOI 10.17487/RFC7761, March 2016, <http://www.rfc-editor.org/info/rfc7761>.

[RFC7893] Stein, Y(J)., Black, D., and B. Briscoe, "Pseudowire Congestion Considerations", RFC 7893, DOI 10.17487/RFC7893, June 2016, <http://www.rfc-editor.org/info/rfc7893>.

[RFC8083] Perkins, C. and V. Singh, "Multimedia Congestion Control: Circuit Breakers for Unicast RTP Sessions", RFC 8083, DOI 10.17487/RFC8083, March 2017, <http://www.rfc-editor.org/info/rfc8083>.

Благодарности

Много людей обсуждало и описывало проблемы, ставшие мотивами этого документа. Вклад и комментарии представили Lars Eggert, Colin Perkins, David Black, Matt Mathis, Andrew McGregor, Bob Briscoe, Eliot Lear. Эта работа частично финансировалась Европейской комиссией по программе Seventh Framework в рамках проекта Reducing Internet Transport Latency (RITE) (ICT-317700).

Адрес автора

Godred Fairhurst
University of Aberdeen
School of Engineering
Fraser Noble Building
Aberdeen, Scotland AB24 3UE
United Kingdom
Email: gorry@erg.abdn.ac.uk
URI: http://www.erg.abdn.ac.uk

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru