Network Working Group
Request for Comments: 3746
Category: Informational

L. Yang
Intel Corp.
R. Dantu
Univ. of North Texas
T. Anderson
Intel Corp.
R. Gopal
Nokia

April 2004

Схема разделения пересылки и управления ForCES

Forwarding and Control Element Separation (ForCES) Framework

Статус документа

Этот документ содержит информацию для сообщества Internet. Документ не задаёт каких-либо стандартов Internet. Распространение документа не ограничивается.

Авторские права

Аннотация

Этот документ определяет архитектурную модель для сетевых элементов ForCES[1] и указывает связанные с ней объекты и их взаимодействия.

Оглавление

[1]Forwarding and Control Element Separation - разделение элементов управления и пересылки.

1. Определения

1.1. Уровни требований

Ключевые слова **необходимо** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не нужно** (SHALL NOT), **следует** (SHOULD), **не следует** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **не рекомендуется** (NOT RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе интерпретируются в соответствии с BCP 14, RFC 2119 [1].

1.2. Терминология

Термины, связанные с требованиям к ForCES, определены в [4] и в этом документе представлены наиболее тесно связанные с темой определения.

Addressable Entity (AE) - адресуемый объект (элемент)
Физическое сетевое устройство, которое непосредственно адресуется в данной технологии соединения. Например, в сетях IP - это устройства, к которым можно обращаться по адресу IP, а в матрице коммутации (switch fabric) - это устройства, к которым можно обращаться по номеру порта в матрице.

Physical Forwarding Element (PFE) - физический элемент пересылки
Элемент AE, который включает оборудование, использование для обработки и обслуживания каждого пакета. Это оборудование может включать сетевые процессоры ASIC, линейные платы с множеством микросхем, автономные устройства с процессорами общего назначения и др.

PFE Partition - раздел PFE
Логическая часть PFE, содержащая некоторое подмножество ресурсов (например, портов, памяти, записей таблицы пересылки), доступных в PFE. Эта концепция аналогична выделению ресурсов для виртуального элемента коммутации, описанного в [9].

Physical Control Element (PCE) - физический элемент управления
Элемент AE, который включает оборудование, используемое для обеспечения функций управления. Обычно это оборудование включает процессор общего назначения.

PCE Partition - раздел PCE
Логическая часть PCE, содержащая некое подмножество ресурсов, доступных в PCE.

Forwarding Element (FE) - элемент пересылки
Логический элемент, реализующий протокол ForCES. Элементы FE используют базовое оборудование для обработки каждого пакета и управляются (контролируются) одним или множеством CE по протоколу ForCES. FE может быть отдельным элементом (или PFE), часть PFE или множеством PFE.

Control Element (CE) - элемент управления
Логический объект, который реализует протокол ForCES и инструктирует один или множество FE по части обработки пакетов. Функциональность CE включает исполнение протоколов управления и сигнализации. CE может состоять из частей PCE или целых PCE.

ForCES Network Element (NE) - элемент сети ForCES
Объект, состоящий из одного или множества CE и одного или множества FE. NE обычно скрывает свою внутреннюю структуру от внешних наблюдателей и выглядит с их точки зрения единой точкой управления.

Pre-association Phase - фаза до объединения
Интервал времени, в течение которого менеджер FE (см. ниже) и менеджер CE (см. ниже) определяют каким FE и CE следует быть частью одного сетевого элемента. Некоторые элементы NE могут относится к фазе до объединения, а другие - к фазе после объединения.

Post-association Phase - фаза после объединения
Интервал времени, в течение которого FE знает управляющие им устройства CE и наоборот, включая интервал, в течение которого CE и FE организуют связи между собой.

ForCES Protocol - протокол ForCES
Хотя в архитектуре ForCES может применяться множество протоколов, термин «протокол ForCES» относится лишь к протоколу ForCES фазы после объединения (см. ниже).

ForCES Post-Association Phase Protocol - протокол ForCES после объединения
Протокол, используемый в коммуникациях между CE и FE после объединения. Этот протокол не применяется для коммуникаций CE-CE, FE-FE или между менеджерами FE и CE. Протокол ForCES работает в режиме «ведущий-ведомый» (master-slave) где FE являются ведомыми, а CE - ведущими. Этот протокол включает управление коммуникационным каналом (например, организацию соединения, обмен heartbeat) и сами управляющие сообщения. Протокол может представлять единое целое или набор совместно работающих протоколов. Эта информация будет представлена в отдельном документе.

FE Manager - менеджер FE
Логический элемент, которые работает в фазе до объединения и отвечает за определение CE, с которыми элементу FE следует взаимодействовать. Этот процесс называется обнаружением CE и может включать определение менеджером FE возможностей доступных CE. Менеджер FE может применять все, что угодно от статической конфигурации до протокола фазы до объединения (см. ниже) для определения используемого CE. Однако протокол фазы до объединения выходит за рамки документа. Будучи логическим устройством менеджер FE может быть физически объединён с любыми логическими элементами, упомянутыми в этом разделе.

CE Manager (CEM) - менеджер элементов управления

Логический объект, который отвечает за генерацию базовых задач управления CE. Используется, в частности, на этапе до объединения (pre-association phase) для определения FE, с которым CE следует взаимодействовать. Этот процесс называется обнаружением FE и может включать определение менеджером CE возможностей доступных FE. Менеджер CE может использовать все, что угодно от статической конфигурации до протокола фазы до объединения (см. ниже) для определения используемых FE. Протокол фазы до объединения выходит за рамки документа. Будучи логическим элементом, менеджер CE может быть физически объединён с любыми из логических элементов, упомянутых в этом разделе.

Pre-association Phase Protocol - протокол фазы до объединения

Протокол взаимодействия менеджеров FE и CE для определения используемых элементов CE и FE. Этот протокол может включать механизм определения возможностей CE и/или FE. Отметим, что этот процесс определения возможностей полностью отделен (и не служит заменой) от процесса, используемого протоколом ForCES. Однако эти два механизма определения возможностей могут использовать одну модель FE.

FE Model - модель элемента пересылки

Модель, описывающая функции логической обработки в FE.

ForCES Network Element (NE) - элемент сети ForCES

Объект, состоящий из одного или множества CE и одного или множества FE. Для внешних наблюдателей NE представляется единой точкой управления. Обычно NE скрывает свою внутреннюю структуру от внешних наблюдателей.

ForCES Protocol Element - элемент протокола ForCES

FE или CE.

Intra-FE topology - внутренняя топология FE

Представление реализации отдельного FE в виде комбинации множества логических функциональных блоков на множестве путей данных. Определяется моделью FE.

FE Topology - топология FE

Представление соединений между множеством FE в одном NE. Иногда это называют внешней топологией FE, чтобы отличать от внутренней топологии FE, используемой моделью FE.

Inter-FE topology - внешняя топология FE

См. FE Topology.

2. Введение в ForCES

Элементы сети IP (NE[1]) представляются внешнему наблюдателю монолитной частью сетевого оборудования, например, маршрутизатором, транслятором NAT, межсетевым экраном, балансировщиком нагрузки. Однако внутри NE (например, маршрутизатор) имеет множество логически разделённых элементов, которые совместно обеспечивают заданную функциональность (например, маршрутизацию). Существует два типа компонент в элементах сети - элементы управления CE[2] на уровне управления и элементы пересылки FE[3] на уровне управления (или уровне данных). Элементами пересылки обычно являются контроллеры ASIC, сетевые процессоры или устройства на основе процессоров общего назначения, которые обеспечивают операции пути данных для каждого пакета. Элементы управления обычно представляют собой процессоры общего назначения, которые реализуют функции управления типа протоколов сигнализации и маршрутизации.

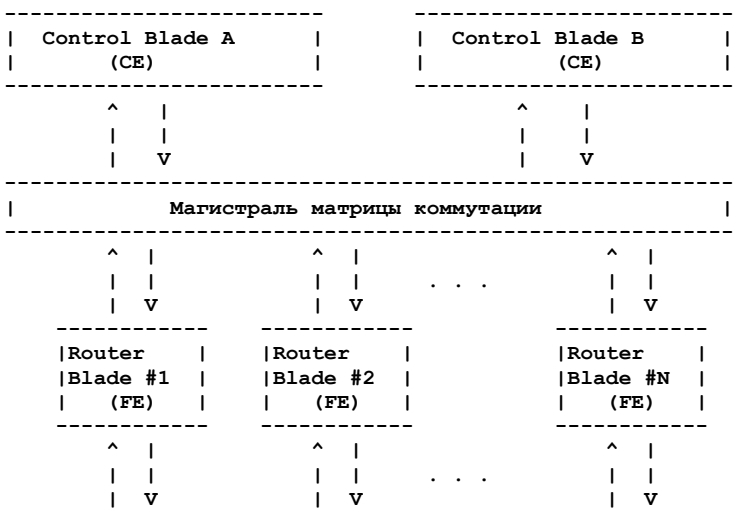


Рисунок 1. Пример конфигурации маршрутизатора с раздельными модулями.

Целью ForCES является определение основы (модели) и связанных с ней протоколов для стандартизации информационного обмена между уровнями управления и пересылки. Наличие стандартных механизмов позволяет элементам CE и FE стать физически разделёнными стандартными компонентами. Это физическое разделение обеспечивает архитектуре ForCES некоторые преимущества. Отдельные компоненты позволяют их производителям сосредоточится именно на них, не разбираясь в других компонентах. Стандартный протокол обеспечивает возможность взаимодействия между элементами CE и FE разных производителей и это позволяет разработчикам систем интегрировать CE и FE от разных производителей. Это взаимодействие транслируется в увеличение числа вариантов выбора и повышение гибкости систем. В целом ForCES позволит ускорить развитие уровней управления и пересылки с сохранением взаимодействия. Масштабирование в этой архитектуре также упрощается за счёт возможности добавления дополнительных компонент управления или пересылки в существующие элементы сети без их полного обновления.

Один из примеров такого разделения реализуется на уровне блейдов (модулей). На рисунке 1 показан пример конфигурации маршрутизатора с двумя модулями управления и множеством моделей пересылки, соединённых между

[1]Network element.
[2]Control element.
[3]Forwarding element.

собой по шине матрицы коммутации (switch fabric backplane). В такой конфигурации шасси модули управления являются элементами CE, модули маршрутизации - элементами FE, а системная шина обеспечивает соединения между всеми модулями. Control blade A может быть первичным CE, а B может служить для резервирования. Возможно также использование резервной матрицы коммутации для повышения надёжности. Современные маршрутизаторы такого типа используют фирменные интерфейсы для обмена сообщениями между элементами CE и FE. Задачей ForCES является замена таких фирменных интерфейсов стандартным протоколом. С помощью стандартного протокола типа ForCES, реализованного во всех модулях (blade), становится возможным взаимодействие модулей управления от компании X с модулями пересылки от компании Y и их совместное использование в одном шасси.

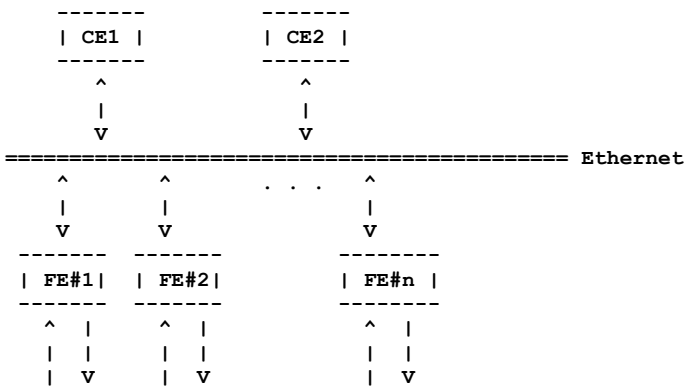


Рисунок 2. Пример конфигурации маршрутизатора в виде отдельных устройств.

Другой пример физического разделения CE и FE может быть реализован на уровне устройств (box). В такой конфигурации все CE и FE представляют собой автономные устройства, соединённые через скоростную ЛВС (например, Gigabit Ethernet). Эти раздельные CE и FE находятся в одной локальной сети (без промежуточных маршрутизаторов). Элементы CE и FE взаимодействуют между собой по протоколу ForCES и набор этих CE и FE становится единым маршрутизатором с точки зрения внешнего наблюдателя. Пример показан на рисунке 2.

В обоих примерах для соединений CE-FE и FE-FE используется одна физическая среда. Однако это совсем не обязательно. Одной из причин использования разных сред является то, что для соединений между CE и FE не требуется такой высокой скорости, как для соединений между элементами FE, которые существенно дороже. Раздельные соединения повышают также уровень надёжности. и резервирования для NE.

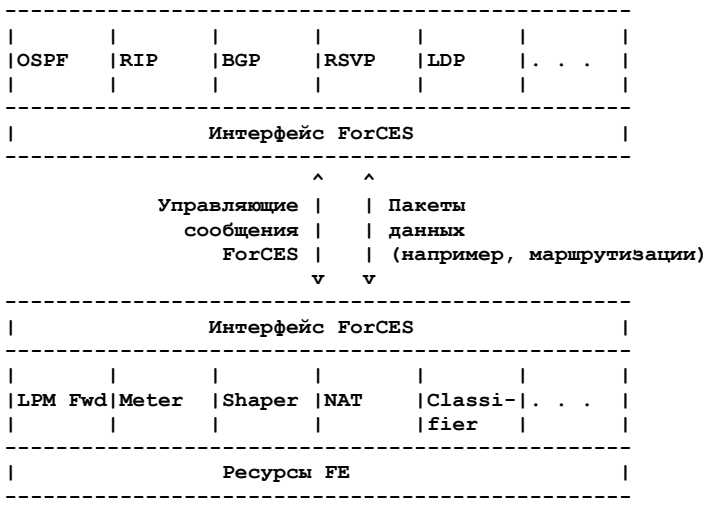


Рисунок 3. Примеры функций CE и FE.

Некоторые примеры функций управления, которые могут быть реализованы в элементах CE, включают протоколы маршрутизации типа RIP, OSPF и BGP, протоколы управления и сигнализации типа RSVP[1], LDP[2] для MPLS и т. п. Примерами функций пересылки в FE служат узлы пересылки LPM[3], классификаторы, формирователи трафика, измерители, NAT[4] и т. п. На рисунке 3 показаны примеры функций CE и FE. Любой данный элемент NE может содержать в себе один или множество таких CE и FE, функционирующих в нем. Рисунок также показывает использование протокола ForCES для транспортировки управляющих сообщений самого протокола ForCES и пакетов данных, которые адресованы функциям управления в элементах CE или созданы ими (например, пакеты протокола маршрутизации). Более подробная информация приведена в параграфе 4.2.4.

Набор требований для разделения пересылки и управления приведён в [4]. Этот документ описывает архитектуру ForCES, соответствующую архитектурным требованиям [4], и определяет основу для сетевых элементов ForCES и связанных с ними объектов для упрощения спецификации протокола. Когда это нужно, в документе используется множество примеров иллюстрирующих проблемы и/или возможные решения в ForCES. Это просто примеры и не следует считать их единственным или заданным путём решения той или иной задачи. Предполагается, что рабочая группа ForCES выпустит отдельный документ со спецификацией протокола ForCES.

[1]Resource Reservation Protocol - протокол резервирования ресурсов.
[2]Label Distribution Protocol - протокол распространения меток.
[3]Longest prefix match - самый длинный из совпадающих префиксов.
[4]Network Address Translator - транслятор сетевых адресов.

3. Архитектура

В этом разделе определяется архитектурная основа ForCES и связанные с ней логические компоненты. Схема ForCES определяет компоненты сетевых элементов ForCES (NE), включающие некоторые вспомогательные компоненты. Эти элементы могут соединяться между собой разными вариантами топологии для гибкой обработки пакетов.

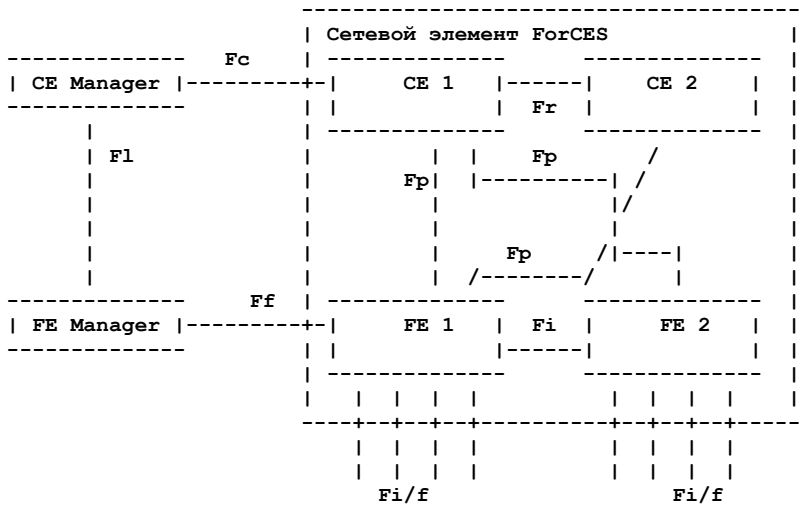


Рисунок 4. Архитектура ForCES.

На рисунке 4 показаны логические компоненты архитектуры ForCES и их взаимосвязи. Внутри сетевого элемента ForCES имеется два типа компонент - элементы управления CE и элементы пересылки FE. Схема предполагает наличие множества экземпляров CE и FE внутри одного NE. Каждый элемент FE включает один или множество интерфейсов с физической средой для приёма пакетов из внешнего мира и передачи своих пакетов вовне. Комбинация этих интерфейсов FE образует внешние интерфейсы элемента NE. Кроме внешних интерфейсов требуется также наличие того или иного типа соединений между элементами внутри NE, позволяющих CE и FE взаимодействовать (один элемент FE может пересылать пакеты другому FE). На рисунке также показаны два объекта, не относящихся к ForCES NE - CE Manager и FE Manager. Эти два дополнительных объекта обеспечивают настройку соответствующих CE и FE в фазе до их объединения (pre-association phase, см. параграф 4.1).

Для удобства логические взаимодействия между этими компонентами помечены опорными точками Fp, Fc, Ff, Fr, Fl и Fi, как показано на рисунке 4. Внешние интерфейсы FE имеют метку Fi/f. Более подробное описание этих интерфейсов приведено ниже. Все эти опорные точки важны для понимания архитектуры ForCES, однако протокол ForCES определён только для одного интерфейса Fp.

Интерфейс между двумя ForCES NE идентичен интерфейсу между двумя традиционными маршрутизаторами и эти два NE обмениваются протокольными пакетами через внешние интерфейсы Fi/f. Подключение ForCES NE к имеющимся маршрутизаторам «прозрачно[1]».

3.1. Элементы управления и интерфейс Fr

Не требуется определять какие-либо протоколы для интерфейса Fr, чтобы включить разделение управления и пересылки для простых конфигураций с одним CE и множеством FE. Однако архитектура допускает наличие в одном сетевом элементе множества CE. В случаях когда реализация использует множество CE, нужно поддерживать совместное представление элементов CE и FE как единого элемента NE.

Множество CE может использоваться для резервирования, распределения нагрузки, распределенного управления и других задач. В случае резервирования один или несколько элементов CE находятся в готовности принять на себя работу отказавшего активного CE. В случае распределения нагрузки одновременно активны два или более CE и все запросы, которые могут обслуживаться одним CE, могут также быть обработаны любым другим CE. Для резервирования и распределения нагрузки вовлечённые CE имеют эквивалентные возможности. Единственным различием для этих случаев является число одновременно активных CE. Для распределенного управления активны два или более CE одновременно, но каждый из них отвечает только за свою часть запросов.

При наличии множества CE в элементе ForCES NE, их внутренняя организация определяется реализацией и выходит за рамки ForCES. Элементы CE полностью отвечают за координацию своих действий чрез интерфейс Fr для обеспечения согласованности и синхронизации. Однако ForCES не определяет реализацию или протокол взаимодействия между CE и даже не задаёт распределение функций между ними. Тем не менее, ForCES будет поддерживать механизмы резервирования и отказоустойчивости CE и предполагается, что разработчики будут реализовать эти решения в рамках модели.

3.2. Элементы пересылки и интерфейс Fi

FE является логическим объектом, который реализует протокол ForCES и использует базовое оборудование для обработки и обслуживания каждого пакета под управлением CE. Возможно разделение одного физического FE на множество логических элементов FE. Возможно также использование одним FE множества физических элементов FE. Отображение между физическими и логическими элементами FE выходит за рамки ForCES. Например, логическая

[1]Как подключение прочих сетевых устройств. *Прим. перев.*

часть физического FE может создаваться путём выделения некоторой части каждого из ресурсов (порты, память, записи таблицы пересылки), доступных ForCES в физическом FE, каждому логическому элементу FE. Такая концепция виртуализации FE аналогична виртуальным элементам коммутации, описанным в [9]. Если виртуализация FE происходит только до объединения, она не оказывает влияния на ForCES. Однако если виртуализация приводит к изменению ресурсов существующих FE (уже участвующих в ForCES после объединения), протокол ForCES должен обеспечить возможность информирования CE о таких изменениях с помощью асинхронных сообщений (требование 6 раздела 4[1] в [4]).

Элементы FE выполняют все функции обработки пакетов в соответствии с указаниями CE, не проявляя своей инициативы. Таким образом, элементы FE являются ведомыми и делают только то, что им сказано. Элементы FE могут взаимодействовать с одним или множеством CE через интерфейс Fp. FE не знают об избыточности CE, распределении нагрузки или распределенном управлении, они просто воспринимают команды от любого CE, имеющего полномочия управления и CE самостоятельно координируют свои действия для обеспечения резервирования, распределения нагрузки или распределенного управления. Идея состоит в том, чтобы сделать FE максимально простыми и «тупыми» и занимающимися только функциями обработки пакетов. Если протокольный обмен ForCES на задаёт иного, каждый элемент FE будет обрабатывать полномочные входящие команды, направленные ему, в порядке их поступления.

Например, на рисунке 5 элементы FE1 и FE2 могут быть настроены на восприятие команд от основного (CE1) и резервного (CE2) элементов CE. При обнаружении отказа CE1 (возможно через интерфейс Fr или Fp), CE2 будет принимать на себя функции CE1. Этот вопрос выходит за рамки ForCES и в дальнейшем не рассматривается.

Распределенное управления может быть реализовано похожим способом без привлечения интеллекта FE. Например, FE могут быть настроены на обнаружение пакетов протокола RSVP и BGP с пересылкой пакетов RSVP одному CE, а BGP - другому. Поэтому элементам FE могут потребоваться функции фильтрации для пересылки пакетов нужным CE.

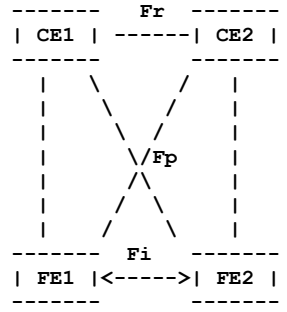


Рисунок 5. Пример избыточности CE.

Эта архитектура позволяет присутствовать множеству элементов FE в одном NE. В [4] требуется, чтобы протокол ForCES был способен обеспечивать масштабирование по меньшей мере до сотен FE (требование 11 раздела 4[1] в [4]). Каждый из этих FE может иметь свой набор функций обработки пакетов с разными интерфейсами в среду передачи. FE отвечают за базовую поддержку связности на канальном уровне с другими FE и внешними объектами. Многие среды канального уровня включают изощрённые протоколы управления. Протокол FORCES (через интерфейс Fp) будет способен переносить сообщения этих протоколов, чтобы в соответствии с «тупой» моделью FE элементы CE могли обеспечивать нужный интеллект и управление такими средами.

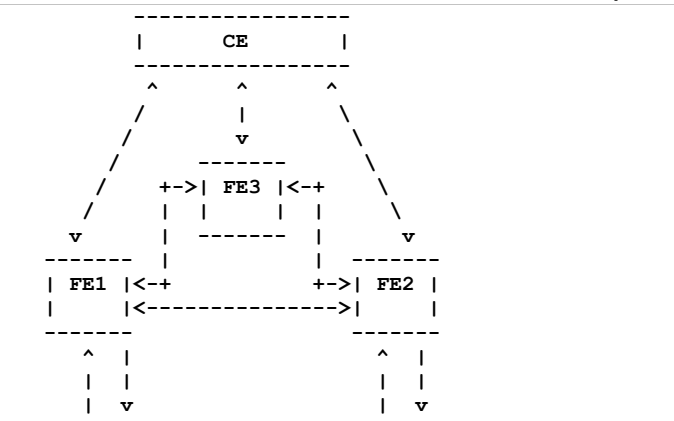
При наличии множества FE модель ForCES требует, чтобы пакеты могли поступать в NE через один элемент FE и покидать NE через другой FE (требование 3 раздела 4[1] в [4]). Пакеты, которые приходят в NE через один FE и выходят из NE через другой FE, передаются между элементами FE через интерфейс Fi. Опорная точка Fi может использоваться FE для определения своей (между FE) топологии, возможно в фазе до объединения. Интерфейс Fi отделен от интерфейса Fp и в настоящее время не определяется протоколом ForCES.

FE могут соединяться между собой в разных вариантах топологии и обработка пакета может проходить через несколько FE в этой топологии. Поэтому логическое течение (поток) пакетов может отличаться от физической топологии FE. На рисунке 6 представлены некоторые варианты топологии. Когда требуется пересылать пакеты между FE, элементам CE нужно знать топологию FE. Элемент CE может запросить у FE топологию по протоколу ForCES, но FE не обязаны представлять такую информацию элементам CE. Поэтому топологическая информация FE может собираться другими путями, выходящими за рамки протокола ForCES (например, протокол определения топологии inter-FE).

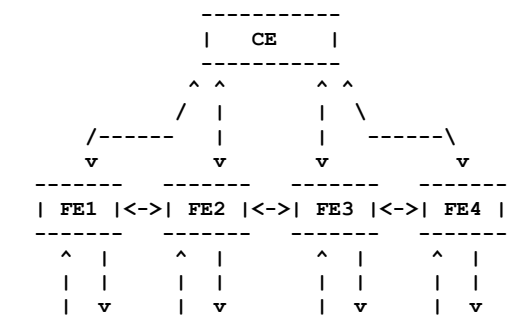
3.3. Менеджеры CE

Менеджеры CE отвечают за определение элементов FE, которыми CE следует управлять. Для менеджеров CE уместно жёсткое задание элементов FE, с которыми его элементам CE следует взаимодействовать. Менеджер CE может быть физически встроен в CE и реализован как простой механизм настройки конфигурации CE (например, с клавиатуры). Кроме того, менеджеры CE могут быть физически и логически отдельными объектами, которые задают в CE данные элементов FE с помощью механизмов типа COPS-PR [7] или SNMP [5].

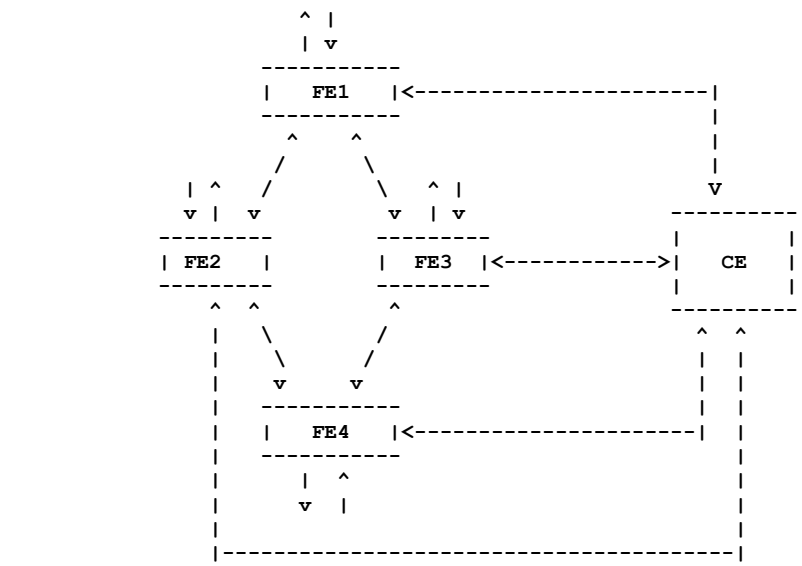
[1]В оригинале ошибочно указан раздел 5, см. https://www.rfc-editor.org/errata/eid5338. *Прим. перев.*



(a) Полносвязные соединения между FE1, FE2 и FE3



(b) Множество FE, соединённых в daisy-цепочку.



(c) Множество FE, соединённых в кольцо.

Рисунок 6. Примеры топологии FE.

3.4. Менеджеры FE

Менеджеры FE отвечают за определение элемента CE, с которым каждому отдельному FE следует начинать взаимодействие. Как и для менеджеров CE для них не задаётся ограничений по способу выбора CE взаимодействующего с его элементами FE, а также на способы реализации этих менеджеров. Каждому FE следует иметь единственный менеджер FE, хотя разные FE могут пользоваться общим менеджером FE. Каждый менеджер может существовать и работать независимо от других менеджеров.

4. Фазы работы

Элементы FE и CE требуют некоторой начальной настройки конфигурации до того, как начнётся обмен информации между ними и работа в качестве единого элемента сети. Данная схема предусматривает две фазы работы - до объединения (pre-association) и после объединения (post-association).

4.1. До объединения

Фазой Pre-association называется период, в течение которого FE Manager и CE Manager определяют каким FE и CE следует быть часть. Одного сетевого элемента. Протокол, используемый в этой фазе, может включать целиком или частично обмен сообщениями через интерфейсы Fl, Ff и Fc. Однако все это не обязательно и не входит в протокол ForCES.

4.1.1. Интерфейс Fl

Менеджеры CE и FE могут взаимодействовать через интерфейс Fl в фазе до объединения для того, чтобы определить отдельные CE и FE или множества этих элементов, которые нужно связать (объединить). Коммуникации через интерфейс Fl не являются в этой архитектуре обязательными. Требований к интерфейсу не предъявляется.

Менеджеры CE и FE могут управляться разными операторами. Оператор менеджера CE может не пожелать раскрытия каких-либо характеристик своих CE кому либо сверх указанного набора менеджеров FE. Аналогично и оператор менеджера FE может пожелать скрыть характеристики FE от всех, кроме уполномоченных объектов. Поэтому для менеджеров CE и FE может потребоваться проверка подлинности друг друга. Показанный ниже обмен сообщениями между менеджерами CE и FE может потребовать других функций защиты типа приватности, безотзывности, «свежести, и целостности.

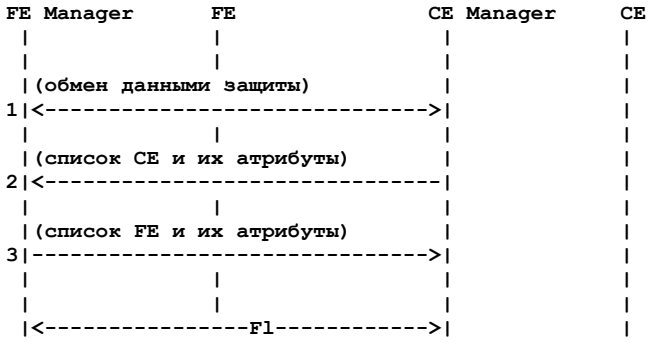


Рисунок 7. Пример обмена сообщениями через интерфейс Fl.

После выполнения требуемых функций защиты менеджеры CE и FE начинают взаимодействие для определения элементов CE и FE между которыми следует организовать взаимодействие (объединить). Менеджеры CE и FE должны, как минимум, знать о наличии доступных CE и FE. Этот процесс обнаружения может включать определение одним или обоими менеджерами возможностей обнаруженных протокольных элементов ForCES. На рисунке 7 показан пример результативного обмена сообщениями между менеджерами CE и FE через интерфейс Fl.

4.1.2. Интерфейс Ff

Интерфейс Ff служит для информирования элементов пересылки о решениях по связыванию, принятых менеджером FE в фазе pre-association. Только уполномоченные элементы могут давать FE инструкции по части CE, которым следует управлять им. Поэтому требуется функции защиты конфиденциальности, целостности, «свежести» и проверка подлинности между менеджером FE и элементами FE, когда менеджер удалён от элементов FE. После организации приемлемой защиты менеджер FE через интерфейс информирует элементы FE о присоединении к новому NE или выходе из существующего NE. Менеджер FE может также задать уникальные идентификаторы элементам FE, которые будут использовать этот интерфейс. Идентификаторы FE полезны после объединения для представления топологии FE. На рисунке 8 показан пример обмена сообщениями через опорную точку Ff.

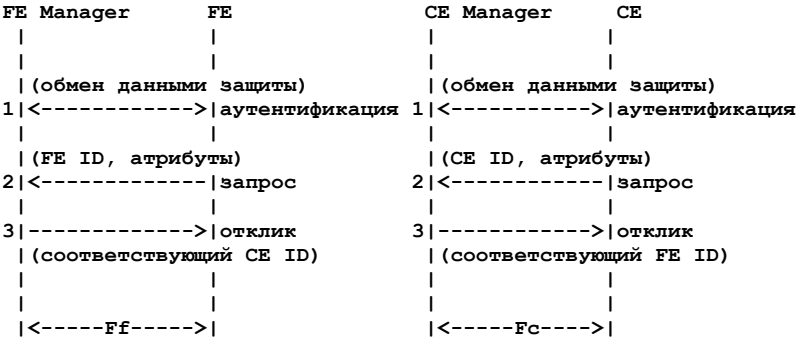


Рисунок 8. Примеры обмена сообщениями через интерфейсы Ff и Fc.

Отметим, что функциональность менеджера FE может быть реализована там же, где FE (например, ручной ввод IP-адресов CE с клавиатуры) и в этом случае интерфейс сводится к встроенной функции.

4.1.3. Интерфейс Fc

Интерфейс Fc служит для информирования элементов управления о решениях по связи (объединению) принятых менеджерами CE в фазе pre-association. Для случая удалённого менеджера CE только уполномоченные объекты могут давать CE инструкции по управления элементами FE. Защита конфиденциальности и целостности, а также «свежесть» и проверка подлинности требуются на интерфейсе в такой конфигурации. После обеспечения подходящей защиты менеджер CE инструктирует элементы CE, какими FE им следует управлять и как следует это делать. На рисунке 8 показан пример обмена сообщениями через интерфейс Fc.

Как и для FE, возможны ситуации совместного размещения менеджера и элементов CE без использования протокола для реализации этой функции.

4.2. Фаза после объединения и интерфейс Fp

Фаза Post-association (после объединения) представляет собой период времени, когда элементы FE и CE уже имеют информацию (настроены) для взаимодействия, и включает организацию связи и коммуникации в установившемся состоянии. Обмен информацией между CE и FE происходит через интерфейс Fp (p означает протокол). Протокол ForCES является единственным протоколом обмена данными через интерфейс Fp.

4.2.1. Близость элементов CE и FE и соединения между ними

Рабочая группа ForCES осознанно выбрала для первое версии ForCES «очень близкое» размещение элементов CE и FE в сетях IP. Очень близкими считаются элементы управления и пересылки, расположенные в одном физическом устройстве (box) или разделённые одним интервалом пересылки через локальную сеть[1] ([8]). Элементы CE и FE могут быть связаны с использованием разных технологий, включая соединения Ethernet, системные шины (backplane), матрицы коммутации ATM (ячейки) и т. п. Протоколу ForCES следует поддерживать все эти технологии (см. требование 1 в разделе 4[2] документа [4]). Когда элементы CE и FE разделены более чем одним интервалом маршрутизации L3 (hop), протокол ForCES будет использовать существующий протокол L4, соответствующий RFC 2914 [3], с подходящими гарантиями, защитой и контролем перегрузок (например, TCP или SCTP) в качестве транспорта.

4.2.2. Организация связи (объединение)

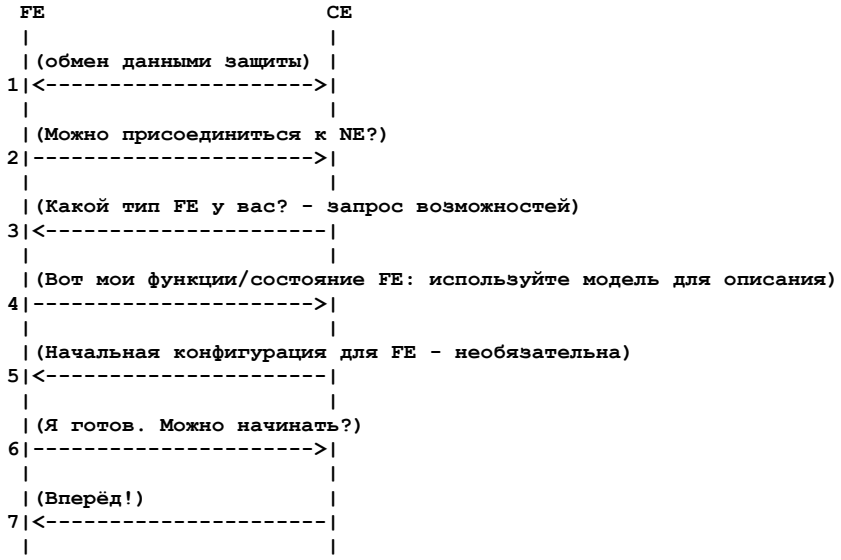


Рисунок 9. Пример обмена сообщениями между CE и FE через интерфейс Fp для организации NE.

Пример на рисунке 9 показывает некий обмен сообщениями перед тем, как организация связи (объединение) между CE и FE будет завершена. Инициатором соединения может быть CE или FE.

Согласование защиты требуется для взаимной проверки подлинности взаимодействующих конечных точек до начала обмена сообщениями. В это согласование следует включать взаимную аутентификацию и проверку полномочий CE и FE, но детали этого согласования зависят от решения для безопасности выбранного протоколом ForCES. Проверка полномочий может быть простым просмотром списка разрешённых точек, представленного менеджером FE или CE в фазе до объединения. Проверки подлинности и полномочий должны завершиться успешно до того, как будет организована ассоциация (объединение). При отказе аутентификации или проверки полномочий конечной точке не разрешается присоединение к NE. После успешного согласования защиты необходимость аутентификации и защиты конфиденциальности сообщений сохраняется для обмена информацией между CE и FE, если не существует той или иной формы физической защиты. Если пакет не проходит проверку подлинности, он должен быть отброшен и может передаваться уведомление для сигнала отправителю о возможной атаке. Более подробное рассмотрение вопросов безопасности для ForCES приведено в разделе 8.

После успешного согласования защиту FE нужно сообщить элементу CE свои возможности и можно также проинформировать о топологии соединений с другими FE. Возможности FE следует представлять моделью FE в соответствии с требованиями [4] (раздел 6, требование 1). Модель позволит FE описать тип выполняемых им функций обработки пакетов, собираемой статистики, возможных событий и т. п. После того, как эта информация станет доступна CE, тот может передать элементу FE начальную конфигурацию, чтобы FE мог корректно начать приём и обработку пакетов. По окончании обмена требуемой начальной информацией процесс объединения завершается. Обработка и пересылка пакетов в FE не может начаться до организации связи (объединения). После объединения элементы CE и FE переходят в режим установившейся связи (steady-state communication).

4.2.3. Установившаяся связь

После организации связи между CE и FE эти элементы используют протокол ForCES для обмена через интерфейс Fp информацией, облегчающей обработку пакетов.

На основе информации, полученной в результате обработки данных управления в CE, элементам CE зачастую будет требоваться изменять поведение своих FE при обработке пакетов путём отправки инструкций этим FE. На рисунке 10 показан пример обмена сообщениями, в котором CE передаёт новые маршруты элементу FE, чтобы тот мог добавить их в свою таблицу пересылки. CE может запросить у FE собранную тем статистику, а FE может уведомлять CE о важных событиях типа отказа порта.

4.2.4. Пакеты данных через Fp

[1] Без маршрутизаторов между ними. *Прим. перев.*
[2] В оригинале ошибочно указан раздел 5, см. https://www.rfc-editor.org/errata/eid5339. *Прим. перев.*

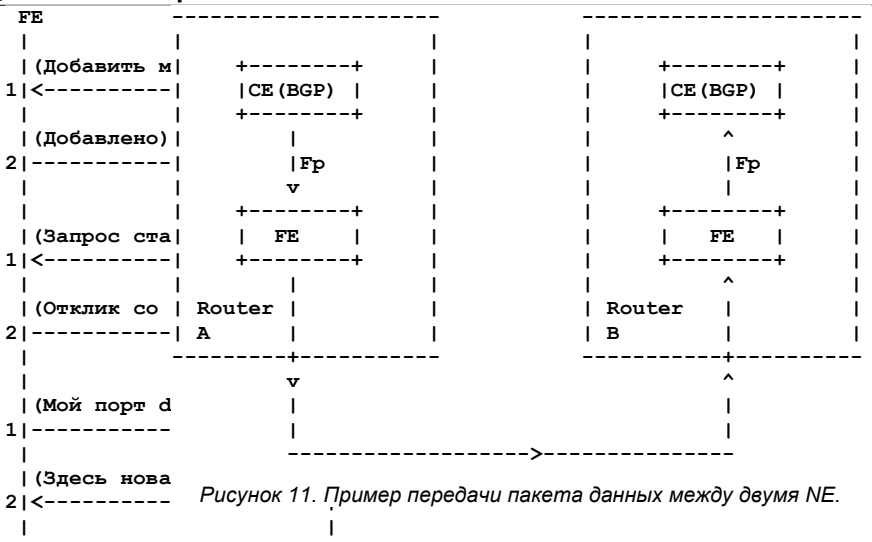


Рисунок 11. Пример передачи пакета данных между двумя NE.

Рисунок 10. Примеры обмена сообщениями между CE FE через Fp в установившемся состоянии.

Пакеты протокола уровня управления (например, RIP, сообщения OSPF), адресованные любому из интерфейсов NE, обычно перенаправляются принимающим элементом FE своему CE, а CE может быть источником пакетов, которые его FE затем доставляет другим элементам NE. Поэтому протокол ForCES через интерфейс Fp не только доставляет свои протокольные сообщения между CE и FE, но и инкапсулирует пакеты данных протоколов уровня управления. Кроме того, один элемент FE может контролироваться несколькими CE при распределенном управлении. В такой конфигурации протоколы управления, поддерживаемые элементами ForCES NE, могут проходить через несколько CE. Например, один элемент CE может поддерживать протоколы маршрутизации типа OSPF и BGP, а протоколы сигнализации и контроля доступа типа RSVP будут поддерживаться другим CE. Элементы FE настраиваются на распознавание и фильтрацию пакетов этих протоколов с их пересылкой соответствующим CE.

На рисунке 11 показано, как пакеты BGP от маршрутизатора A передаются в маршрутизатор B. В этом примере протокол ForCES служит для транспортировки пакетов из CE в FE внутри маршрутизатора A, затем из FE в CE внутри маршрутизатора B. В свете того, что протокол ForCES отвечает за транспортировку между CE и FE как управляющих сообщений, так и пакетов данных через интерфейс Fp, возможно использовать один или множество протоколов.

4.2.5. Proxy FE

В случае, когда физический элемент FE не может (например, по причине отсутствия CPU общего назначения) напрямую реализовать протокол ForCES, может применяться FE-посредник (proxy) для завершения Fp вместо физического FE. Это позволяет CE взаимодействовать с физическим FE через посредника по протоколу ForCES, тогда как посредник управляет физическим FE с помощью некой промежуточной формы коммуникаций (например, не являющийся ForCES протокол или DMA). В таких случаях комбинация прокси и физического FE становится одним логическим элементом FE. Посредник может также действовать от имени множества физических элементов FE.

Следует учитывать влияние FE на безопасность. Поскольку физический элемент FE не способен реализовать протокол ForCES, механизмы защиты ForCES могут обеспечивать безопасность только для коммуникационного канала между CE и FE-посредником, но не путь к физическому FE. Рекомендуется использовать другие механизмы защиты (включая физическую) для сквозного обеспечения безопасности взаимодействий между CE и физическим FE.

4.3. Повторная организация связи

Элементы FE и CE могут присоединяться к NE и выходить из них динамически (требование 12 раздела 4[1] в [4]). Когда FE или CE выходит из NE, ассоциация (связь) с NE разрывается. Если этот элемент снова присоединяется к данному NE, для восстановления связи ему может потребоваться повторение фазы pre-association. Потеря связи может также происходить неожиданно в результате разрыва соединения между CE и FE. Поэтому модель разрешает двух сторонние переходы между фазами «до объединения» и «после объединения», но протокол ForCES примени только в фазе post-association. Однако протоколу следует поддерживать механизм восстановления ассоциации. Это включает способность элементов CE и FE детектировать потерю связи, её восстановление и эффективные механизмы (ре)синхронизации (требование 7 раздела 5 в [4]). Отметим, что должны также восстанавливаться защитные связи (security association) и их состояния для гарантии того же уровня защиты после восстановления ассоциации.

Когда FE присоединяется или выходит из работающего элемента NE, элементу CE нужно осознать изменение топологии FE и принять соответствующие меры.

4.3.1. Изящный перезапуск CE

Отказ и перезапуск CE в маршрутизаторе могут вызвать серьёзные проблемы и нарушение работы уровня управления в сети, поскольку при любой перезагрузке CE маршрутизатор теряет отношения смежности и сессии со своими соседями по маршрутизации. Соседи, заметив потерю смежности, обычно рассчитывают новые маршруты и передают обновления своим соседям о потере связности. Эти соседи также обновляют свои маршруты и передают обновления своим соседям. В это же время перезагрузившийся маршрутизатор не может получать трафик от других маршрутизаторов, поскольку соседи прекратили использование ранее анонсированных им маршрутов. Когда перезагруженный маршрутизатор восстановит отношения смежности, его соседи опять пересчитают маршруты и распространят обновления. Перезагруженный маршрутизатор не способен пересылать пакеты, пока не будут восстановлены отношения смежности с соседями, получены обновления маршрутов от соседей и рассчитаны новые маршруты. Пока схождение маршрутов не завершится во всей сети, пакеты могут теряться во временных «чёрных дырах» или маршрутных петлях.

[1] В оригинале ошибочно указан раздел 5, см. https://www.rfc-editor.org/errata/eid5340. *Прим. перев.*

Был разработан механизм обеспечения высокой доступности, названный graceful restart (изящный перезапуск) для протоколов маршрутизации IP (OSPF [11], BGP [12], IS-IS [13]) и протокола распространения меток MPLS (LDP [10]), для снижения отрицательных воздействий на маршрутизацию в сети, вызываемых перезагрузкой маршрутизатора. Механизм позволяет предотвратить переключение маршрутов (route flap) на соседних маршрутизаторах и перезагруженный маршрутизатор будет пересылать пакеты, которые в ином случае просто бы отбрасывались.

Хотя детали различаются для разных протоколов, общая идея механизма изящного перезапуска сохраняется. Перезагружаемый маршрутизатор может информировать своих соседей о перезагрузке. Его соседи могут обнаружить потерю смежности, но не станут пересчитывать свои маршруты и передавать обновления своим соседям. Те, в свою очередь, тоже сохранят маршруты, полученные от перезагружающегося маршрутизатора, предполагая, что они остаются пригодными в течение некоторого ограниченного времени. В этом случае элементы FE перезагружаемого маршрутизатора могут продолжать получение и пересылку трафика в течение ограниченного интервала, используя уже имеющиеся у них таблицы. Затем перезагруженный маршрутизатор восстановит свои отношения смежности, загрузит обновлённые маршруты от своих соседей, рассчитает новые маршруты и использует их для замены прежних. Он будет передавать эти обновлённые маршруты своим соседям и сообщать о завершении процесса перезагрузки.

Безостановочная пересылка является требованием изящного перезапуска. Маршрутизатору необходимо продолжать пересылку пакетов, пока он загружает маршрутные данные и рассчитывает новые маршруты. Это обеспечивает предотвращение отбрасывания пакетов. Как можно видеть, одним из преимуществ разделения CE и FE является возможность продолжать пересылку пакетов в случае отказа или перезапуска CE. Поддержка динамического изменения ассоциаций CE-FE в ForCES также обеспечивает совместимость с механизмами высокой доступности типа изящной перезагрузки.

ForCES следует поддерживать возможность изящного перезапуска CE. При первоначальном создании ассоциации элемент CE должен информировать FE о том, что им следует делать в случае отказа CE. Если изящный перезапуск не поддерживается, элементам FE можно дать указание прекратить обработку пакетов в случае отказа CE. При поддержке изящного перезапуска элементам FE следует дать указание кэшировать и сохранять своё состояние (включая таблицы пересылки) на время перезагрузки CE. При это должен указываться тайм-аут, при наступлении которого действие кэшированных состояний должно прекращаться. Значения тайм-аутов следует делать настраиваемыми элементом CE.

4.3.2. Перезапуск FE

Для примера рисунке 5 предположим, что CE1 является в данный момент рабочим CE. Что же произойдёт, если один из FE (например, FE1) временно выйдет из состава NE? FE1 может по своей инициативе покинуть объединение. Кроме того, возможно просто прекращение работы FE1 в результате неожиданного отказа. В первом случае CE1 получит запрос leave-association от FE1, во втором случае CE1 обнаружит отказ FE1 каким-то иным способом. В обоих случаях CE1 должен информировать протоколы маршрутизации о произошедшем событии, что скорей всего приведёт в пересчёту доступности и SPF[1], а также связанной с этим загрузки новых FIB из CE1 в оставшиеся FE (в примере это FE2). Этот пересчёт и обновления FIB будет распространяться от CE1 к соседним NE, которые были связаны с FE1.

Когда FE1 решит вернуться в ассоциацию или перезагрузится после отказа, ему потребуется снова определить своего ведущего (CE). Это можно сделать несколькими способами. Можно снова перейти в фазу pre-association и получить нужную информацию от менеджера FE. Можно отыскать прежнюю информацию CE в своём кэше, если её свежесть может быть проверена. Как только элемент обнаружит свой CE, он начинает обмен сообщениями с ним для восстановления ассоциации, как показано на рисунке 9 (возможно при этом удастся обойти согласование транспорта). Предположим, что в FE1 сохранилась таблица маршрутизации и другие данные состояния из прежней ассоциации. Тогда вместо повторной передачи этих данных он сможет воспользоваться более эффективным механизмом для восстановления синхронизации состояния со своим CE, если такой механизм поддерживается протоколом ForCES. Например, контрольная сумма CRC-32 для состояния позволяет проверить его синхронизацию с CE. Сравнив своё состояние с CE, элемент может при необходимости передать лишь обновления. Протокол ForCES может реализовать похожий механизм для оптимизации, но может и не делать этого, поскольку требования не задают такой механизм.

5. Применимость RFC 1812

Требование 9 в разделе 4[2] [4] говорит «Любая предлагаемая архитектура ForCES **должна** объяснять, как она будет поддерживать все функции маршрутизации, определённые в [RFC1812].» В RFC 1812 [2] рассмотрено множество важных требований к маршрутизаторам IPv4 от канального до прикладного уровня. В этом разделе рассматриваются требования RFC 1812, связанные с реализациями маршрутизаторов IPv4 на основе архитектуры ForCES и показано, как эти требования выполняются в ForCES путём разделения функциональности, требуемой для уровней пересылки и управления.

В общем случае уровень пересылки выполняет большую часть обработки каждого пакета, которая требует высокой скорости, а уровень управления выполняет большую часть требующих сложных вычислений операций, которые типичны для протоколов управления и сигнализации. Однако провести чёткую линию раздела между обработкой в элементах CE и FE невозможно и архитектуре ForCES не следует ограничивать инновационные подходы к разделению пересылки и управления. По мере роста производительности обработки в FE некоторые функции управления, традиционно выполнявшиеся в CE, могут переноситься в элементы FE для повышения производительности и масштабируемости. Такие функции могут включать часть обработки ICMP или TCP, а также часть протоколов маршрутизации. Будучи перенесёнными на уровень пересылки, такие функции CE хотя и продолжают относиться к уровню управления, становятся реально функциями FE. Подобно другим функциям FE, эти перенесённые (off-loaded) функции должны выражаться как часть модели FE, чтобы элемент CE мог принять решение о наиболее эффективном использовании таких функций при их наличии в FE.

5.1. Общие требования к маршрутизатору

Маршрутизаторы имеют не менее двух логических интерфейсов. Когда элементы CE и FE разделены ForCES внутри одного NE, требуются некие дополнительные интерфейсы для коммуникаций внутри этого NE, как показано на рисунке

[1]Shortest Path First - выбор маршрута по кратчайшему пути.
[2]В оригинале ошибочно указан раздел 5, см. https://www.rfc-editor.org/errata/eid5337. *Прим. перев.*

12. В этом примере NE имеется один элемент CE и два FE. Каждый FE имеет 4 интерфейса, из которых два служат для приёма и передачи пакетов, связанных с внешним миром, а два оставшихся используются внутри NE. Элемент CE имеет два логических интерфейса 9 и 10, подключённых к интерфейсам 3 и 6 элементов FE1 и FE2, соответственно. Интерфейсы 4 и 5 служат для коммуникаций между FE1 и FE2. В результате маршрутизатор имеет 4 внешних интерфейса - 1, 2, 7 и 8.

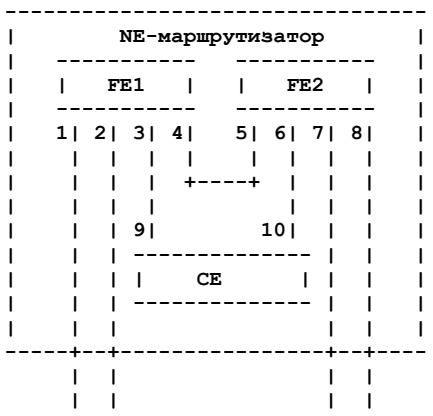


Рисунок 12. Пример NE-маршрутизатора с 4 интерфейсами.

Маршрутизаторы IPv4 должны реализовать протокол IP для поддержки функции пересылки пакетов, которая управляется таблицей FIB[1]. Эта пересылка уровня Internet (см. раздел 5 в RFC 1812 [2]) функционально относится к FE в архитектуре ForCES.

Маршрутизатор может реализовать протоколы транспортного уровня (типа TCP и UDP), которые требуются для поддержки протоколов прикладного уровня (см. раздел 6 в RFC 1812 [2]). Важным классом прикладных протоколов являются протоколы маршрутизации (см. раздел 7 в RFC 1812 [2]). В архитектуре ForCES протоколы маршрутизации относятся к элементам CE. Для протоколов маршрутизации требуется взаимодействие между маршрутизаторами. Это взаимодействие между CE в разных маршрутизаторах поддерживается в ForCES за счёт способности FE перенаправлять пакеты данных, адресованные маршрутизаторам (т. е. NE), и способности CE создавать пакеты, которые будут доставляться их элементами FE. Такие коммуникации осуществляются через интерфейс Fp внутри каждого маршрутизатора и между внешними интерфейсами соседних маршрутизаторов, как показано на рисунке 11.

5.2. Канальный уровень

Поскольку FE владеют всеми внешними интерфейсами маршрутизатора, элементы FE должны соответствовать требованиям RFC 1812 [2] в канальному уровню. Поддержка ARP может быть реализована в CE или FE. Как будет показано позднее, ряд категорий поведения, которые RFC 1812 относит к канальному уровню, может быть реализован реализован как в FE, так и в CE. Нужны общие рекомендации для обеспечения взаимодействия между разделёнными уровнями управления и пересылки. Рекомендации, предлагаемые здесь, заключаются в том, что CE **должны** быть способны выполнить эти типы операции, а FE **могут** реализовать их. Модели FE следует указывать возможности, для которых CE может принимать решение о месте реализации этих функций.

Параметры интерфейса, включая MTU, адрес IP и т. п., должны настраиваться элементами CE через ForCES. Элементы CE должны быть способны определить доступен ли физический интерфейс в FE для передачи пакетов. Элементы FE должны также информировать CE об изменении состояния своих интерфейсов (типа up/down) через ForCES.

5.3. Сетевой уровень

Элементы FE и CE должны реализовать протокол IP и все обязательные расширения, заданные RFC 1812. Элементам CE следует реализовать опции IP типа source route и record route, а FE могут реализовать их по своему усмотрению. Опцию timestamp следует реализовать в FE для аккуратной вставки временных меток. Элемент FE должен интерпретировать понятные ему опции IP и сохранять остальные неизменными для использования их элементами CE. FE и CE могут отбрасывать пакеты без отправки ошибок ICMP, но такие события следует учитывать и записывать в системный журнал. Элементы FE могут передавать такую статистику в CE через ForCES.

При участии в обработке пакета множества FE сохранение единства NE должно строго соблюдаться. Например, значение TTL[2] должно декрементироваться в данном элементе NE только один раз. Это можно делать, например, в выходной функции последнего FE.

FE должны принимать и обрабатывать любые пакеты с широковещательным или групповым адресом получателя, для которых маршрутизатор запросил приём. При поддержке в маршрутизаторах групповой адресации (IP multicast) протокол IGMP реализуется в CE. Элементы CE должны также поддерживать ICMP, а для FE такая поддержка не обязательна. Об наличии такой поддержки можно уведомить CE с помощью модели FE. Элементы FE всегда могут полагаться на CE в части передачи ошибок ICMP, но и сами FE также могут генерировать сообщения ICMP об ошибках.

5.4. Пересылка на сетевом уровне

Пересылка IP реализуется элементами FE. При обновлении таблицы маршрутизации в элементах CE используется ForCES для передачи новых маршрутов из CE в FE. Каждый элемент FE имеет свою таблицу пересылки и использует её для отправки пакетов на интерфейс следующего интервала пересылки (next hop).

При получении пакетов IP элемент FE проверяет заголовок IP и обрабатывает большинство опций IP. Некоторые опции не могут быть обработаны, пока не будет решения о маршрутизации, которое принимается после проверки адреса получателя. Если этот адрес относится к самому маршрутизатору, пакет фильтруется и обрабатывается локально или

[1]Forwarding Information Base - база данных для пересылки.
[2]Time-To-Live - время жизни.

пересылается CE в зависимости от инструкций, установленных CE. В остальных случаях FE определяет IP-адрес следующего маршрутизатора, просматривая свою таблицу пересылки. FE также определяет сетевой интерфейс, который он будет использовать для передачи пакетов. Иногда FE может потребоваться пересылка пакетов другому элементу FE до того, как пакет можно будет переслать следующему маршрутизатору. Непосредственно перед отправкой пакета следующему маршрутизатору FE уменьшает значение TTL на 1 и обрабатывает все опции IP, которые не были обработаны ранее. FE выполняет фрагментацию IP при необходимости, определяет адрес канального уровня (например, с помощью ARP), инкапсулирует дейтаграмму IP (или её фрагмент) в подходящий кадр канального уровня и помещает её в очередь выбранного выходного интерфейса.

Другие опции, упомянутые в RFC 1812 [2] для пересылки IP, также будут реализоваться в элементах FE (например, фильтрация пакетов).

FE обычно пересылают пакеты локальных получателей элементам CE. FE могут также отправлять элементам CE «исключительные» пакеты (про которые FE не знает куда отправить). Элементы CE должны обрабатывать пакеты по любым причинам пересланные им элементами FE. Для ForCES может потребоваться присоединение к таким пакетам неких метаданных, показывающих причину их пересылки от FE в CE. При получении от FE пакета с метаданными CE могут принять решение о самостоятельной обработке пакета или его передаче протоколам вышележащего уровня, включая протоколы маршрутизации и администрирования. Если элемент CE обрабатывает пакет самостоятельно, он может отбросить пакет или переслать его после изменения. Элементы CE могут также генерировать новые пакеты и доставлять их элементам FE для дальнейшей пересылки.

Любая смена состояния в процессе работы маршрутизатора также должна корректно обрабатываться в соответствии с RFC 1812. Например, когда FE прекращает пересылку, элемент NE в целом может продолжать пересылку пакетов, но ему нужно прекратить анонсирование маршрутов, на которые влиял отказавший элемент FE.

5.5. Транспортный уровень

Транспортный уровень обычно реализуется в CE для поддержки протоколов вышележащих уровней (например, протоколов маршрутизации) На практике это означает, что большинство CE реализуют протоколы TCP[1] и UDP[2].

Элементы CE и FE должны реализовать протокол ForCES. В некоторых транспортных протоколах L4, используемых для поддержки ForCES, элементы CE и FE должны реализовать транспорт L4 и протокол ForCES.

5.6. Прикладной уровень - протоколы маршрутизации

Протоколы внутренней и внешней маршрутизации реализуются в элементах CE. Пакеты протоколов маршрутизации, порождённые CE, пересылаются элементам FE для доставки. Результаты таких протоколов (типа обновлений таблиц пересылки) передаются FE через ForCES.

Из соображений производительности или масштабируемости часть функций уровня управления, требующих быстрого отклика, может быть «выгружена» из CE в элементы FE. Например, в протоколе OSPF периодически генерируются и обрабатываются пакеты Hello. При реализации этого в CE входящие пакеты Hello проходят с внешнего интерфейса FE на элемент CE по внутреннему каналу CE-FE. Аналогично, исходящие пакет пакеты Hello проходят от CE к FE и на внешние интерфейсы. Частый обмен обновлениями Hello создаёт высокую нагрузку на CE и может также перегружать канал CE-FE. Поскольку элементов FE в маршрутизаторе обычно больше, нежели элементов CE, «выгрузка» функций обработки Hello обеспечивает более эффективную распределенную и масштабируемую обработку. Указав такие «выгруженные» функции в модели FE, можно обеспечить взаимодействие. Однако точное описание «выгруженной» функциональности, соответствующее указанным в модели FE «выгруженным» функциям, не является частью самой модели и требует отдельной спецификации.

5.7. Прикладной уровень - протоколы управления сетью

В RFC 1812 [2] также сказано: «Маршрутизаторы **должны** поддерживать управление по протоколу SNMP.» В общем случае для фазы после объединения большинство внешних задач управления (включая SNMP) следует выполнять через взаимодействие с CE для поддержки представления в виде единого устройства. Поэтому рекомендуется реализовать агенты SNMP в элементах CE и перенаправлять сообщения SNMP, полученные элементами FE, в их CE. Здесь может быть использована модель AgentX, определённая в RFC 2741 [6], где CE выступают в роли ведущих агентов при обработке протокольных сообщений SNMP, а элементы FE служат субагентами, обеспечивающими доступ к базам MIB, размещённым в FE. Сообщения протокола AgentX между ведущим агентом (CE) и субагентами (FE) инкапсулируются и доставляются через ForCES как пакеты данных других протоколов прикладного уровня.

6. Заключение

Этот документ определяет архитектурную основу для ForCES. Он указывает относящиеся к делу сетевые элементы ForCES, включая (один или множество) FE, (один или множество) CE, а также необязательные менеджеры FE и CE. Документ также описывает взаимодействие этих компонент и опорные точки (интерфейсы) для такого взаимодействия. Важно подчеркнуть, что из числа этих опорных точек лишь интерфейс Fp между элементами CE и FE относится к ForCES. Возможностей ForCES может оказаться недостаточно для поддержки всех желаемых конфигураций NE. Однако мы надеемся, что ForCES через интерфейс Fp является самым важным элементом в физическом разделении и взаимодействии CE и FE, поэтому данный интерфейс нужно стандартизовать в первую очередь. Простые и полезные конфигурации можно будет реализовать, используя лишь стандартизованный интерфейс между CE и FE (например, один элемент CE с полносвязной сетью FE).

7. Благодарности

Joel M. Halpern представил множество полезных замечаний и предложений, а также отметил несколько важных проблем. T. Sridhar принадлежит предложение о возможности применения протокола AgentX с SNMP для управления сетевыми элементами ForCES. Susan Hares отметила проблему изящного перезапуска в ForCES. Russ Housley, Avri

[1] Transmission Control Protocol - протокол управления передачей.
[2]User Datagram Protocol - протокол пользовательских дейтаграмм.

Doria, Jamal Hadi Salim и многие другие участники почтовой конференции ForCES также представили полезные отклики.

8. Вопросы безопасности

Администратор NE волен выбирать конфигурацию защиты в своей среде. Например, ForCES может работать между CE и FE, соединёнными по шине внутри устройства. В этом случае физическая защита устройства предотвратит большинство атак типа MITM[1], отслеживания и подмены, поэтому архитектура ForCES может полагаться на физическую защиту устройства и протокольные механизмы можно отключить. Однако MITM-атаки через внешние интерфейсы, описанные в параграфе 8.1.8, создают угрозу даже в таких ситуациях и для их предотвращения нужен механизм ограничения скорости. Этот пример показывает важность вопросов защиты элементов сети с поддержкой ForCES в различных вариантах её развёртывания. Администраторам следует предоставлять возможность настройки уровня защиты, требуемого для протокола ForCES.

В общем случае физическое разделение двух объектов обычно использует потенциально небезопасное соединение между ними и требует более серьёзной защиты. Например, в параграфе 4.1 было указано, что проверка подлинности становится необходимой для взаимодействия между менеджерами CE и FE, элементами и менеджерами CE, а также элементами и менеджерами FE в некоторых конфигурациях. Физическое разделение элементов FE и CE также требует защиты для протокола ForCES, работающего через интерфейс Fp. В этом разделе предпринимается первая попытка описать угрозы, которые могут быть связаны с физическим разделением элементов FE и CE, а также приведены рекомендации и руководства по защите работы и администрирования протокола ForCES через интерфейс Fp на основе имеющихся стандартных решений для защиты.

8.1. Анализ возможных угроз, вносимых ForCES

В этом параграфе проводится анализ угроз для ForCES и особое внимание уделено интерфейсу Fp. Каждая угроза описана с деталями её воздействия на объекты протокола ForCES и/или NE в целом, а также рассмотрена полная функциональность, требуемая для защиты.

8.1.1. Лавинные рассылки сообщений Join и Remove для CE

Угрозы

Враждебный узел может передавать поток ложных запросов на включение в NE или выход из него от имени несуществующих или не имеющих полномочий FE легитимным элементам CE с очень высокой скоростью, создавая ненужную загрузку CE.

Воздействия

Если поддерживать состояния для отсутствующих или не уполномоченных элементов FE, CE может стать недоступным для других операций, что создаёт возможность организации DoS-атак[2], подобных TCP SYN DoS. При использовании множества CE ненужная информация о состояниях может также передаваться этим CE через интерфейс Fr (например, от активного CE к резервному CE) и DoS-атака коснётся множества CE.

Требования

Элементу CE получившему запрос join или remove, не следует создавать данных состояния для элемента, который не является аутентифицированным FE.

8.1.2. Атаки с подменой

Угрозы

Враждебный узел может прикинуться элементом CE или FE и передавать ложные сообщения.

Воздействия

Риску может быть подвержен элемент NE в целом.

Требования

Элементы CE и FE должны проверять подлинность сообщение от другого FE или CE по списку уполномоченных элементов ForCES (предоставляется менеджером CE или FE в фазе pre-association) до восприятия и обработки такого сообщения.

8.1.3. Replay-атаки

Угроза

Враждебный узел может повторно использовать (replay) полное сообщение, переданное раньше элементом FE или CE при проверке подлинности.

Воздействие

Риску может быть подвержен элемент NE в целом.

Требования

Механизм защиты от повторного использования сообщений должен быть включён в защитное решение.

8.1.4. Атаки при передаче управления

Угроза

Враждебный узел может воспользоваться механизмом CE fail-over для захвата управления элементом NE. Например, если два элемента CE - CE-A и CE-B управляют несколькими FE, CE-A может быть активным, а CE-B - резервным. При отказе CE-A управление принимает на себя CE-B и становится активным CE. Элементы FE уже имеют доверительные отношения с CE-A, но могут не иметь в момент передачи управления таких отношений с CE-B. Это даёт злоумышленнику возможность взять на себя роль CE-B, если доверительные отношения не были организованы до отказа или во время перехода.

Воздействие

Риску может быть подвержен элемент NE в целом после незащищённого переключения элементов CE.

Требования

Уровень доверия между резервным CE и элементами FE должен быть таким же, как для активного CE. Защищённая связь между FE и CE может быть организована заранее. Если этого не было сделано, такую связь требуется создать до передачи управления резервному CE.

[1]Man-in-the-middle - перехват и изменение данных с участием человек.
[2]Denial of service - атака, нацеленная на отказ в обслуживании легитимных клиентов.

8.1.5. Целостность данных

Угрозы

Враждебный узел может внедрять ложные сообщения для легитимных CE или FE.

Воздействие

FE или CE получит сфабрикованное сообщение и выполнит некорректную или недопустимую операцию.

Требования

Для сообщений протокола требуется защита целостности.

8.1.6. Конфиденциальность данных

Угроза

При физическом разделении FE и CE злоумышленник может перехватывать пакеты в пути. Некоторые сообщения могут играть критическую роль в работе сети в целом, а другие могут содержать конфиденциальные деловые данные. Утечка такой информации может создавать риск даже для непосредственно соединённых CE и FE.

Воздействие

Возможность раскрытия деликатной информации, передаваемой между CE и FE.

Требования

Должна быть доступна защита конфиденциальности данных, передаваемых между FE и CE.

8.1.7. Совместное использование параметров защиты

Угроза

В случае когда несколько FE взаимодействуют с одним CE, используя общие ключи аутентификации, компрометация любого элемента FE или CE подвергает риску и всех остальных.

Воздействие

Риску подвержен элемент NE в целом.

Требования

Для предотвращения таких побочных эффектов лучше использовать отдельные параметры защиты для каждой пары FE-CE, взаимодействующих через интерфейс Fp.

8.1.8. DoS-атаки через внешний интерфейс

Угроза

Когда FE получает пакет, адресованный CE, он пересылает этот пакет через интерфейс Fp. Злоумышленник может создать огромный поток сообщений типа пакетов протоколов маршрутизации и т. п. через внешний интерфейс Fi/f, которые FE будет обрабатывать и пересылать элементу CE через интерфейс Fp.

Воздействие

Элемент CE может столкнуться с нехваткой ресурсов и пропускной способности интерфейса Fp в результате чрезмерного числа пакетов от элементов FE.

Требования

Должен применяться тот или иной механизм ограничения скорости на элементах FE и CE. Ограничитель (Rate Limiter) **следует** делать настраиваемым на уровне FE для каждого типа сообщений, которые будут приниматься через интерфейс Fi/f.

8.2. Рекомендации по защите для ForCES

В требования [4] предложено для протокола ForCES обеспечивать гарантированную доставку для интерфейса Fp, но конкретный транспортный протокол для ForCES ещё не задан. Этот документ не предназначен для указания конкретного транспорта и содержит лишь рекомендации и руководства на основе имеющихся стандартных протоколов защиты [18], которые могут работать с наиболее вероятными транспортными протоколами для ForCES.

Здесь рассмотрены два защитных решения на основе IPsec[1] [15] и TLS[2] [14]. Протокол TLS работает с гарантированным транспортом типа TCP и SCTP для индивидуальной адресации, а IPsec можно использовать с любым транспортом (UDP, TCP, SCTP) и он поддерживает индивидуальную и групповую адресацию. TLS и IPsec подходят для выполнения всех требований защиты протокола ForCES. Кроме того, могут применяться другие решения, соответствующие требованиям безопасности, включая механизмы защиты L2 для используемой технологии канального уровня.

При использовании ForCES между CE и FE внутри устройства или в физически защищённом помещении проверка подлинности, защита конфиденциальности и целостности могут быть обеспечены физическими средствами и методами. Поэтому протокольные механизмы защиты могут быть отключены в зависимости от используемой топологии и административных правил. Однако важно помнить, что даже при реализации NE в одном устройстве возможны DoS-атаки, отмеченные в параграфе 8.1.8, через внешние интерфейсы Fi/f. Поэтому важно иметь соответствующие меры противодействия даже при развёртывании системы в одном устройстве.

8.2.1. Использование TLS с протоколом ForCES

TLS [14] можно применять если используется надёжный транспорт TCP или SCTP для протокола ForCES через интерфейс Fp. Протокол согласования TLS применяется при создании или восстановлении ассоциации для установки параметров сессии TLS между элементами CE и FE. После организации сессии используется протокол TLS record для защиты коммуникаций ForCES между CE и FE.

Ниже описывается базовая схема применения TLS с протоколом ForCES. Этапы 1) - 7) служат для согласования защиты, как показано на рисунке 9, а этап 8) - для обеспечения основных коммуникаций между CE и FE, включая сообщения, показанные на рисунке 9 после согласования защиты и коммуникации в установившемся состоянии (рисунок 10).

1) В фазе Pre-association для всех FE задаются элементы CE (включая активный и резервный CE).

2) FE организует соединение TLS с CE (ведущий) и согласует применяемый шифр.

[1]IP Security - защита IP.
[2]Transport Layer Security - защита транспортного уровня.

3) FE (ведомый) получает сертификат CE, проверяет подпись и срок действия, а также убеждается, что сертификат не отозван.

4) CE (ведущий) получает сертификат FE и выполняет проверки, указанные в п. 3).

5) При отказе любой из проверок пп. 3) и 4) конечная точка должна передать сообщение об ошибке и прервать согласование.

6) После успешной взаимной проверки подлинности организуется сессия TLS между элементами CE и FE.

7) FE передаёт сообщение join NE элементу CE.

8) FE и CE используют сессию TLS для последующих коммуникаций.

Отметим, что существуют разные способы проверки сертификатов элементами CE и FE. Один из способов заключается в настройке менеджера FE или CE или другой центральной компоненты в качестве удостоверяющего центра (CA), чтобы элементы CE и FE могли запрашивать у этого (известного заранее) CA проверку пригодности сертификата (отсутствие отзыва). Другим вариантов является прямое указание в CE и FE списка действующих сертификатов в фазе pre-association (до объединения).

Для случаев отказа активный и резервный CE отвечают за синхронизацию состояний ForCES, включая состояния TLS, для минимизации случаев повторной организации связи при отказах. Следует принять меры, гарантирующие проверку подлинности резервного CE таким же способом, как это выполнялось для основного CE, до отказа или в процессе смены ролей.

8.2.2. Использование IPsec с протоколом ForCES

IPsec [15] можно применять с любым транспортным протоколом (UDP, SCTP, TCP) через интерфейс Fp для протокола ForCES. При выборе IPsec рекомендуется использовать ESP в транспортном режиме, поскольку для ForCES нужна конфиденциальность сообщений.

IPsec можно использовать в ручном или автоматизированном режиме SA и управления криптографическими ключами. Однако при ручном управлении ключами механизмы IPsec для защиты от повторного использования (replay) не доступны. Поэтому рекомендуется автоматическое управления ключами, если функция защиты от повторного использования важна. В иных случаях ручного управления ключами может быть достаточно для некоторых вариантов развёртывания., особенно при сравнительно небольшом числе элементов CE и FE. Рекомендуется время от времени менять ключи даже при ручном управлении ими.

IPsec может поддерживать транспорт с индивидуальной и групповой адресацией. В момент публикации этого документа рабочая группа MSEC активно занималась стандартизацией протоколов для защиты при групповой адресации [17]. Решениям, основанным на групповой адресации в IPsec, следует указать, как они выполняют требования безопасности, указанные в [4].

В отличие от TLS, протокол IPsec обеспечивает услуги защиты между CE и FE на уровне IP, поэтому согласование защиты на рисунке 9 становится пустым для случая ручного управления ключами. Ниже показаны этапы действий для ForCES в таком случае.

1) В фазе Pre-association всем FE указываются элементы CE (включая активный и резервный CE), а также вручную задаются параметры SA.

2) FE передаёт сообщение join NE элементу CE. Это сообщение, а также все последующие, получает услуги защиты в соответствии с заданными вручную параметрами IPsec SA, но защиты от повторного использования (replay) не обеспечивается.

Администратор может выбрать использование общего ключа для множества коммуникаций FE-CE, но рекомендуется применять разные ключи. Аналогично рекомендуется использовать отдельные ключи для входящего и исходящего трафика.

Если требуется автоматическое управление ключами, можно использовать IKE [16], хотя подходят и другие методы распространения ключей типа Kerberos. Процесс обмена ключами согласует параметры защиты, как показано на рисунке 9. Ниже перечислены этапы процесса при использовании IKE с IPsec для протокола ForCES. Процессу согласования защиты на рисунке 9 соответствуют пп. 1) - 6).

1) В фазе Pre-association для всех FE задаются элементы CE (включая активный и резервный CE), правила IPsec и т. п.

2) FE запускает процесс IKE и пытается организовать IPsec SA с элементом CE (ведущий). FE (ведомый) в процессе согласования IKE получает сертификат CE, проверяет пригодность подписи, срок действия сертификата и отсутствие отзыва.

3) CE (ведущий) получает сертификат FE и выполняет те же проверки, что и FE в п. 2).

4) При отказе любой из проверок пп. 2) и 3) конечная точка должна передать сообщение об ошибке и прервать согласование.

5) После успешной взаимной аутентификации организуется сессия IPsec между элементами CE м FE.

6) FE передаёт сообщение join NE элементу CE. В FE пока не создано записи SADB.

7) Элементы FE и CE используют сессию IPsec для продолжения коммуникаций.

Менеджер FE или CE или иная централизованная компонента могут использоваться в качестве CA для проверки пригодности сертификатов CE и FE в процессе IKE. Другим вариантов является настройка CE и FE в фазе pre-association, с заданием требуемой информации типа сертификатов, паролей и т. п., в зависимости от выбранного администратором типа проверки подлинности CE и FE.

Для случаев отказа активный и резервный CE отвечают за синхронизацию состояний ForCES и IPsec для минимизации случаев повторной организации связи при отказах. Кроме того, FE при запуске нужно организовывать разные IPsec SA с каждым элементом CE. Это будет минимизировать периодическую передачу состояний с помощью уровня IPsec через интерфейс Fr (CE-CE).

9. Литература

9.1. Нормативные документы

[1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[2] Baker, F., Ed., "Requirements for IP Version 4 Routers", RFC 1812, June 1995.

[3] Floyd, S., "Congestion Control Principles", BCP 41, RFC 2914, September 2000.

[4] Khosravi, H. and Anderson, T., Eds., "Requirements for Separation of IP Control and Forwarding", RFC 3654, November 2003.

9.2. Дополнительная литература

[5] Case, J., Mundy, R., Partain, D. and B. Stewart, "Introduction and Applicability Statements for Internet Standard Management Framework", RFC 3410, December 2002.

[6] Daniele, M., Wijnen, B., Ellison, M. and D. Francisco, "Agent Extensibility (AgentX) Protocol Version 1", RFC 2741, January 2000.

[7] Chan, K., Seligson, J., Durham, D., Gai, S., McCloghrie, K., Herzog, S., Reichmeyer, F., Yavatkar, R. and A. Smith, "COPS Usage for Policy Provisioning (COPS-PR)", RFC 3084, March 2001.

[8] Crouch, A. et al., "ForCES Applicability Statement", Work in Progress[1].

[9] Anderson, T. and J. Buerkle, "Requirements for the Dynamic Partitioning of Switching Elements", RFC 3532, May 2003.

[10] Leelanivas, M., Rekhter, Y. and R. Aggarwal, "Graceful Restart Mechanism for Label Distribution Protocol", RFC 3478, February 2003.

[11] Moy, J., Pillay-Esnault, P. and A. Lindem, "Graceful OSPF Restart", RFC 3623, November 2003.

[12] Sangli, S. et al., "Graceful Restart Mechanism for BGP", Work in Progress[2].

[13] Shand, M. and L. Ginsberg, "Restart Signaling for IS-IS", Work in Progress[3].

[14] Dierks, T. and C. Allen, "The TLS Protocol Version 1.0", RFC 2246, January 1999.

[15] Kent, S. and R. Atkinson, "Security Architecture for the Internet Protocol", RFC 2401[4], November 1998.

[16] Harkins, D. and D. Carrel, "The Internet Key Exchange (IKE)", RFC 2409[5], November 1998.

[17] Hardjono, T. and Weis, B. "The Multicast Group Security Architecture", RFC 3740, March 2004.

[18] Bellovin, S., Schiller, J. and C. Kaufman, Eds., "Security Mechanisms for the Internet", RFC 3631, December 2003.

10. Адреса авторов

L. Lily Yang

Intel Corp., MS JF3-206,

2111 NE 25th Avenue

Hillsboro, OR 97124, USA

Phone: +1 503 264 8813

EMail: lily.l.yang@intel.com

Ram Dantu

Department of Computer Science,

University of North Texas,

Denton, TX 76203, USA

Phone: +1 940 565 2822

EMail: rdantu@unt.edu

Todd A. Anderson

Intel Corp.

[1]Работа опубликована в RFC 6041. *Прим. перев.*
[2]Работа опубликована в RFC 4724. *Прим. перев.*
[3]Работа опубликована в RFC 3847. *Прим. перев.*
[4]Этот документ заменён RFC 4301. *Прим. перев.*
[5]Этот документ заменён RFC 4306. *Прим. перев.*

2111 NE 25th Avenue

Hillsboro, OR 97124, USA

Phone: +1 503 712 1760

EMail: todd.a.anderson@intel.com

Ram Gopal

Nokia Research Center

5, Wayside Road,

Burlington, MA 01803, USA

Phone: +1 781 993 3685

EMail: ram.gopal@nokia.com

Перевод на русский язык

Николай Малых

nmalykh@gmail.com

11. Полное заявление авторских прав

Подтверждение

Финансирование функций RFC Editor обеспечено Internet Society.