

A Path Computation Element (PCE)-Based Architecture

Архитектура на основе элементов расчёта пути

Статус документа

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

Авторские права

Copyright (C) The Internet Society (2006).

Аннотация

Расчёт пути на основе ограничений является фундаментальным компонентом систем организации трафика, таких как сети с многопротокольной коммутацией по меткам (Multiprotocol Label Switching или MPLS) и обобщённой MPLS (Generalized Multiprotocol Label Switching или GMPLS). Расчёт пути в больших, многодоменных, включающих множество регионов и уровней сетях является сложным и может потребовать специальных вычислительных компонентов и кооперации между разными доменами сети. Этот документ задаёт архитектуру для модели, основанной на элементах расчёта пути (Path Computation Element или PCE), для решения этой задачи. Документ не является попыткой подробного описания всех компонентов архитектуры, а описывает набор блоков архитектуры PCE, из которых можно построить решения.

Оглавление

1. Введение.....	2
2. Термины.....	2
3. Определения.....	2
4. Мотивы разработки архитектуры на основе PCE.....	3
4.1. Требования к ресурсам CPU.....	3
4.2. Частичная видимость.....	3
4.3. Отсутствие TED или использование IGP без TE.....	4
4.4. Узел вне домена маршрутизации.....	4
4.5. Элемент сети без плоскости управления и возможности маршрутизации.....	4
4.6. Расчёт резервного пути для защиты пропускной способности.....	4
4.7. Многоуровневые сети.....	4
4.8. Правила выбора пути.....	4
4.9. Негативные аспекты.....	5
4.9.1. Internet в целом.....	5
4.9.2. Гарантии организации TE LSP.....	5
5. Обзор архитектуры на основе PCE.....	5
5.1. Композитный узел PCE.....	5
5.2. Внешний PCE.....	5
5.3. Расчёт пути с несколькими PCE.....	6
5.4. Расчёт пути с несколькими взаимодействующими PCE.....	6
5.5. Основанное на системе управления использование PCE.....	7
5.6. Области стандартизации.....	7
6. Архитектурные соображения для PCE.....	7
6.1. Централизованный расчёт.....	7
6.2. Распределённый расчёт.....	7
6.3. Синхронизация.....	8
6.4. Обнаружение PCE и распределение нагрузки между ними.....	8
6.5. Обнаружение живучести PCE.....	8
6.6. Расчёты PCC-PCE и PCE-PCE.....	9
6.7. Синхронизация PCE TED.....	9
6.8. PCE с учётом и без учёта состояния.....	10
6.9. Мониторинг.....	10
6.10. Конфиденциальность.....	10
6.11. Политика.....	11
6.11.1. Архитектура политики PCE.....	11
6.11.2. Реализация политики.....	12
6.11.3. Типы политики.....	12
6.11.4. Взаимодействие с сигнализацией.....	12

6.12. Незапрошенные взаимодействия.....	12
6.13. Взаимодействие с откликами.....	12
7. Взгляд клиента расчёта пути.....	13
8. Оценочные показатели.....	13
9. Вопросы управляемости.....	13
9.1. Управление работой и правилами.....	14
9.2. Модели данных и информационные модели.....	14
9.3. Проверка живучести.....	14
9.4. Проверка корректности работы.....	14
9.5. Требования к другим протоколам и функциональным компонентам.....	14
9.6. Влияние на работу сети.....	15
9.7. Другие вопросы.....	15
10. Вопросы безопасности.....	15
11. Благодарности.....	15
12. Литература.....	15

1. Введение

Расчёт пути на основе ограничений является фундаментальным компонентом систем организации трафика в сетях MPLS [RFC3209] и GMPLS [RFC3473]. В [RFC2702] описаны требования к организации трафика в сетях MPLS, а в [RFC4105] и [RFC4216] для сред с обменом между областями и AS, соответственно.

Расчёт пути в больших, многодоменных, включающих множество регионов и уровней сетях является сложным и может потребовать специальных вычислительных компонентов и кооперации между разными доменами сети. Этот документ задаёт архитектуру для модели, основанной на элементах расчёта пути (PCE), для решения этой задачи.

Документ описывает набор блоков архитектуры PCE, из которых можно построить решения. Например, рассматриваются основанные на PCE реализации, включая составной, внешний и множественный расчёт пути PCE. Кроме того, рассматриваются архитектурные соображения, включая централизованный и распределенный расчёт, синхронизацию, обнаружение PCE и распределение нагрузки между ними, обнаружение живучести PCE, взаимодействие между клиентами расчёта пути (Path Computation Client или PCC) и PCE (коммуникации PCC-PCE) и между самими PCE, синхронизация баз данных организации трафика (Traffic Engineering Database или TED), PCE с учётом и без учёта состояния, мониторинга, правила и конфиденциальность, а также показатели оценки.

Модель Internet предполагает распределение функциональности (например, маршрутизации) внутри сети. Функциональность PCE не противоречит этой модели и может применяться в точном соответствии с моделью, например, при сосуществовании в сети функциональности PCE с маршрутизацией по меткам (Label Switching Router или LSR). PCE также может дополнять функциональность в сети, где модель Internet не может предоставить адекватных решений, например, при отсутствии обмена данными организации трафика между доменами сети.

2. Термины

- CSPF**
Constraint-based Shortest Path First - кратчайший путь на основе ограничений.
- LER**
Label Edge Router - граничный маршрутизатор по меткам.
- LSDB**
Link State Database - база данных о состоянии каналов.
- LSP**
Label Switched Path - путь с коммутацией по меткам.
- LSR**
Label Switching Router - маршрутизатор с коммутацией по меткам.
- PCC**
Path Computation Client - клиент расчёта пути. Любое клиентское приложение, запрашивающее расчёт пути элементом PCE.
- PCE**
Path Computation Element - элемент расчёта пути. Любая сущность (компонент, приложение или узел сети), способная рассчитывать сетевой путь или маршрут на основе графа сети и применения расчётных ограничений (см. раздел 3).
- TED**
Traffic Engineering Database - база данных организации трафика, содержащая сведения о топологии и ресурсах домена. TED может получать сведения от расширений протоколов внутренней маршрутизации (Interior Gateway Protocol или IGP) и иных источников.
- TE LSP**
Traffic Engineering MPLS Label Switched Path - путь организации трафика с коммутацией по меткам MPLS.

3. Определения

Элемент расчёта пути (PCE) - это сущность, способная вычислять сетевой путь или маршрут по графу сети с применением расчётных ограничений в этом процессе. PCE - это приложение, которое может размещаться в узле сети, компоненте, сервере вне сети и т. п. Например, PCE может рассчитывать путь TE LSP, оперируя базой TED с учётом пропускной способности и иных ограничений, применимы к запросу услуг TE LSP.

Доменом считается любой набор элементов сети с единым управлением адресами и общей ответственностью за расчёт пути. Примера доменов включают области (area) IGP, автономные системы (Autonomous System или AS), множество AS внутри сети сервис-провайдера. Домен ответственности за расчёт пути может быть также субдоменом области или AS.

Для полной классификации PCE и уточнения приведённых определений нужно рассмотреть приведённые ниже соображения.

- 1) Расчёт пути может выполняться в контексте домена, между доменами и между уровнями.

- a. Расчёт пути между доменами может включать сведения о топологии, маршрутизации и правилах из нескольких доменов, которые позволяют выявить взаимосвязи, помогающие в расчёте пути.
- b. Расчёт пути между уровнями относится к использованию PCE, когда задействовано несколько уровней и цель заключается в расчёте пути на одном или нескольких уровнях со учётом сведений о топологии и ресурсах этих уровней.

Перекрывающиеся домена не рассматриваются в этом документе. При междоменных расчётах домены могут относиться к одному или нескольким сервис-провайдерам.

2) Расчёт с одним или множеством PCE.

- a. При расчёте пути с одним PCE применяется лишь один элемент для расчёта данного пути в домене, где может быть множество PCE.
- b. При расчёте пути с множеством PCE применяется несколько элементов для расчёта пути в домене.

3) Централизованная и распределенная модель.

- a. Централизованной называют модель, где все пути в домене рассчитываются одним элементом PCE.
- b. Распределенной называют модель расчётов, где пути в домене вычисляются множеством PCE.

Пути, проходящие через несколько доменов, могут рассчитываться по распределённой модели с одним или несколькими PCE, отвечающими за каждый домен, или централизованно путём задания домена, включающего все остальные домены.

Из этих определений следует, что централизованная модель использует расчёт пути с одним PCE. Однако в распределённой модели может применяться расчёт на одном или нескольких PCE. Централизованной модели, использующей расчёт на нескольких PCE.

- 4) PCE может (не обязательно) размещаться на головной стороне (head-end) пути. Например, традиционным внутридоменным решением является выполнение расчётов головным LSR из MPLS TE LSP, который в этом случае включает PCE. Имеются решения, где в расчёте пути должны участвовать другие узлы (например, нестрогие интервалы - hop), что делает их самостоятельными PCE. Возможен расчёт пути элементом PCE, физически не входящем в рассчитываемый путь.
- 5) Рассчитанный PCE путь может быть явным (полный путь от начала до получателя с включением строго списка узлов) или строгим/нестрогим (strict/loose), т. е. содержащим комбинацию строгих и нестрогих узлов (hop) хотя бы с одним нестрогим узлом, представляющим получателя, где узел может быть абстракцией (например, AS).
- 6) Модель расчёта путей на основе PCE не претендует на исключительность и может применяться в сочетании с другими моделями расчётов. Например путь TE LSP между AS может рассчитываться по модели PCE в некоторых AS и иными средствами (без PCE) - в других. Кроме того, для разных TE LSP могут применяться разные модели расчёта пути.
- 7) Данный документ не принимает каких-либо допущений о природе или реализации PCE. Элементы PCE могут размещаться в маршрутизаторах, LSR, выделенных серверах сети и т. п. Кроме того, функции PCE ортогональным возможностям пересылки на узле, где эти функции реализованы.

4. Мотивы разработки архитектуры на основе PCE

Ниже указано несколько побудительных причин разработки архитектуры на основе PCE (раздел 5). Список не является исчерпывающим и приведён для наглядности. Следует отметить, что целью этого раздела является представление некоторых примеров приложений, где могут подойти пути, основанные на PCE. Здесь также чётко указано, что модель не ставит цель замены имеющихся моделей расчёта пути и применяется в конкретных существующих и будущих ситуациях.

Из приведённых примеров видно, что PCE не заменяет собой имеющуюся модель Internet с распределением интеллекта по сети. Напротив, она опирается на эту модель и использует информационные и вычислительные возможности распределённых центров. Поэтому PCE следует считать не централизованным «всевидящим оракулом», а совместной работой распределённых функций для решения конкретных задач, таких как расчёт кратчайшего пути с учётом ограничений.

4.1. Требования к ресурсам CPU

Во многих ситуациях расчёт пути может требовать значительных ресурсов CPU, например для:

- размещения набора TE LSP в домене для оптимизации целевой функции (например, наибольшего снижения максимальной загрузки канала);
- расчёта пути по множеству критериев (например, задержка и загрузка канала, учёт возможностей коммутации, свойства адаптации и оптические ограничения в оптической сети GMPLS);
- расчёт деревьев «один со многими» (Point to Multipoint, дерево Steiner) с минимальной стоимостью.

В таких случаях для некоторых маршрутизаторов может оказаться невозможным или нежелательным выполнение расчёта путей из ограничений CPU, и желательно «выгрузить» (off-load) такие расчёты в другие PCE, которые могут быть маршрутизаторами или выделенными серверами PCE.

4.2. Частичная видимость

В некоторых случаях отвечающий за расчёт пути узел может иметь ограниченную видимость топологии пути к адресату. Это ограничение может возникнуть, например, при попытке входного маршрутизатора организовать TE LSP к получателю, находящемуся в отдельном домене, поскольку сведения TE не передаются через границы доменов. В таких случаях можно применять нестрогие маршруты для организации TE LSP, полагаясь на граничные

маршрутизаторы домена для организации следующей части пути. Однако невозможно гарантировать, что будет использован оптимальный (кратчайший) путь или даже найден жизнеспособный путь, кроме как методом проб и ошибок с использованием откликов (crankback) или иных сигнальных расширений.

Эта проблема расчёта междоменных путей, скорее всего, будет решена распределением расчётов с кооперацией PCE в каждом домене и возможным использованием откликов между доменами для динамического решения задач обеспечения. Как вариант можно использовать «всевидящий» централизованный PCE с доступом ко всем топологическим данным, но в этом случае возникает проблема расширяемости (в части размера TED и ответственности одного PCE за обработку запросов из множества доменов) и сохранения конфиденциальности, когда домены относятся к разным сервис-провайдерам.

Отметим, что описанные здесь задачи могут быть дополнительно рассмотрены в контексте повторной оптимизации TE LSP или создания разных TE LSP для защиты или распределения нагрузки.

4.3. Отсутствие TED или использование IGP без TE

База данных организации трафика (TED) может сильно истощать ресурсы узла сети (например, граничного маршрутизатора или LER). Для поддержки TED может потребоваться много памяти и ощутимая загрузка CPU. В таких случаях имеет смысл использовать PCE на отдельном узле для организации и поддержки TED и сделать его доступным для расчёта путей.

Протоколов IGP, работающий в некоторых сетях, недостаточно для создания полной базы TED. Например, в сети может работать OSPF/IS-IS без расширений OSPF-TE/ISIS-TE или часть маршрутизаторов сети может не поддерживать расширения TE. В таких случаях для успешного расчёта путей через сеть база TED должна создаваться или дополняться с помощью действий по настройке и обновляться по мере резервирования или освобождения ресурсов сети. Такую базу TED можно распределить по маршрутизаторам, которым нужно рассчитывать пути, или поддерживать централизованно (на другом узле с поддержкой PCE) для централизованных расчётов.

4.4. Узел вне домена маршрутизации

LER может не входить в домен маршрутизации административным причинам, например, при подключении граничного маршрутизатора клиента (customer-edge или CE) к граничному маршрутизатору провайдера (provider-edge или PE) в контексте MPLS VPN [RFC4364], когда желательно обеспечить путь TE LSP от CE к CE. В таких случаях предполагается решение, не включающее расчёт на входном маршрутизаторе (головной TE LSP, CE) и не полагающееся на настройку статических нестрогих узлов. Оптимальный кратчайший путь при этом не гарантируется. Помочь может использование отдельного PCE, который в этом случае может обеспечивать путь с нестрогими узлами.

4.5. Элемент сети без плоскости управления и возможности маршрутизации

В унаследованных оптических сетях элементы сети зачастую не имеют плоскости управления и возможности маршрутизации, а имеют лишь плоскость данных и все кросс-соединения выполняются в плоскости администрирования (management). В таких случаях желательно запустить расчёт пути на PCE и передавать команды кросс-соединений каждому узлу на рассчитанном пути. Таким образом, PCC будет элементом плоскости администрирования, возможно входящим в систему управления сетью (Network Management System или NMS) или систему поддержки операций (Operations Support System или OSS). Этот сценарий важен для оптических сетей с поддержкой автоматической коммутации (Automatically Switched Optical Network или ASON) и может также применяться для взаимодействия между сетями с поддержкой GMPLS и без такой поддержки.

4.6. Расчёт резервного пути для защиты пропускной способности

PCE может служить для расчёта резервных путей в контексте защиты с быстрой перемаршрутизацией TE LSP. В этой модели все резервные TE LSP, защищающие данный объект, рассчитываются координированно элементом PCE. Это позволяет полностью распределять пропускную способность между резервными туннелями, защищающими независимые элементы, избегая при этом каких-либо расширений сигнализации TE LSP. Здесь применимы централизованные и распределенные модели расчёта. В распределенном случае каждый LSR может быть PCE для расчёта путей резервных туннелей с целью защиты от отказов каналов или узлов соседней сети.

4.7. Многоуровневые сети

Сеть серверного уровня с одним средством коммутации может поддерживать несколько сетей с другим средством (более тонкой) коммутации. Например, сеть с мультиплексированием по времени (Time-Division Multiplexing или TDM) может предоставлять связность для сетей клиентского уровня, таких как IP, MPLS, L2 [MLN].

В сети серверного уровня маловероятно использование такой же парадигмы связности, как в клиентской сети, поэтому гранулярность пропускной способности в серверной сети может быть более грубой, чем в клиентской. Управления этими сетями будет, скорее всего, разделено и в сетях будут использоваться независимые пространства адресов. Кроме того при использовании одной серверной сети несколькими клиентскими сетями в этих сетях клиентов могут применяться независимые правила, параметры управления, адресные пространства и маршрутные предпочтения.

Разные сети клиентского и серверного уровня могут считаться различными областями расчёта путей в домене PCE, поэтому архитектура PCE полезна для обеспечения расчёта путей из одной сети клиентского уровня в другую через сеть серверного уровня. В этом случае PCE отвечают за решение проблемы адресных пространств, обработку различий в правилах и параметрах управления между сетями. Отметим, что из-за различий в гранулярности пропускной способности связность через сеть серверного уровня может обеспечиваться виртуальными каналами TE или смежностью при пересылке (Forwarding Adjacency) - PCE может предоставить точку управления, отвечающую за решение о предоставлении новых каналов TE или смежности при пересылке через сеть серверного уровня.

4.8. Правила выбора пути

PCE может иметь локальную политику, влияющую на расчёт и выбор пути по запросам на расчёт. Эти правила могут работать на основе сведений, представленных запрашивающим PCC. Результат применения таких правил включает,

например, отклонение запроса на расчёт пути или предоставление пути, не соответствующего всем запрошенным ограничениям. Кроме того, политика может поддерживать административно заданные пути или выбор между транзитными провайдерами. Включение политики в PCE может упростить применение правил в процессе расчёта и выбора пути.

PCC может применять локальные правила выбора PCE для расчёта конкретного пути и запрашиваемых ограничений.

В контексте PCE политика может быть чувствительна к типу пути, который будет рассчитан. Например, могут применяться разные наборы правил для пути внутри области или уровня и пути между областями или уровнями.

Отметим, что может потребоваться синхронизация правил между PCE или PCC и PCE. Такие вопросы выходят за рамки архитектуры PCE, но входят в сферу применения и структуры политики PCE, описанных в отдельном документе.

4.9. Негативные аспекты

4.9.1. Internet в целом

PCE не рассматривается решением для всей сети Internet, т. е. применимость PCE ограничена набором доменов с известными взаимоотношениями. Это похоже на партнерское (пиринговое) взаимодействие сервис-провайдеров.

4.9.2. Гарантии организации TE LSP

При расчёте двух и более путей для TE LSP по одному набору сведений о состоянии каналов TE возможна конкуренция полученных в результате путей за ограниченные ресурсы сети. Это может привести к тому, что работать будет лишь указанный первым TE LSP или даже ни один TE LSP не удастся организовать.

Пакетная обработка запросов на расчёт, время отката (back-off), расчёт дополнительных путей и возвратный сигнал (spankback) могут помочь при решении таких проблем, а PCE может повысить шансы на успех при организации TE LSP. Однако один централизованный PCE не представляется решением, гарантирующим организацию TE LSP, поскольку сохраняются возможные отказы в сети или борьба за ресурсы, когда централизованная база TED не может полностью представить текущее (в реальном масштабе времени) состояние сети.

5. Обзор архитектуры на основе PCE

В этом разделе представлен обзор архитектуры модели PCE. Для полного представления о гибкости модели его следует прочесть вместе с последующими параграфами, описывающими детали.

5.1. Композитный узел PCE

На рисунке 1 показаны компоненты типового композитного узла PCE (т. е. маршрутизатора, реализующего функции PCE), реализующего расчёт путей. Для обмена сведениями TE, на которых основана база TED, служит протокол маршрутизации. Запросы услуг предоставления TE LSP принимаются узлом и преобразуются в сигнальные запросы, но для преобразования может потребоваться расчёт пути, который запрашивается у PCE. PCE работает с базой TED в соответствии с локальной политикой для получения отклика с запрошенным путём.

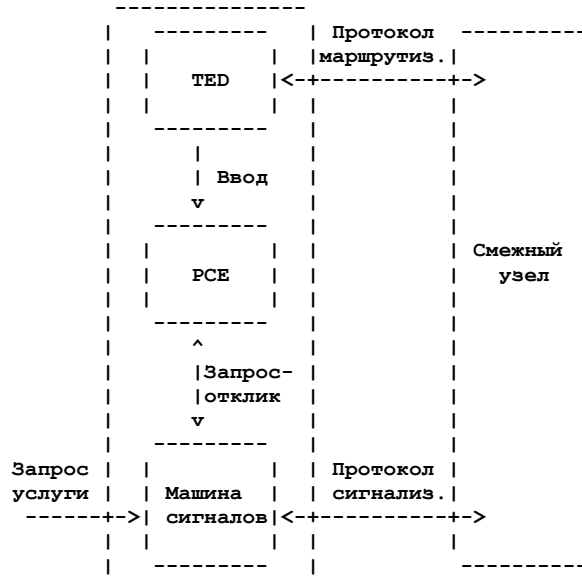


Рисунок 1. Композитный узел PCE.

Отметим, что маршрутизация между композитным узлом PCE и любым другим маршрутизатором может выполняться через прямое соединение или какой-либо туннель.

5.2. Внешний PCE

На рисунке 2 показан PCE, являющийся внешним для запрашивающего элемента сети. Запрос услуг получает головной узел, который перед возможностью инициировать сигнализацию делает запрос внешнему PCE на расчёт пути. PCE использует TED в соответствии с локальной политикой как входные данные для расчёта и возврата отклика.

Отметим, что в этом случае узел, поддерживающий функции PCE, может также быть LSR или маршрутизатором, выполняющим пересылку самостоятельно (композитный узел PCE), но эти функции ортогональны работе функции в рассматриваемом здесь примере.

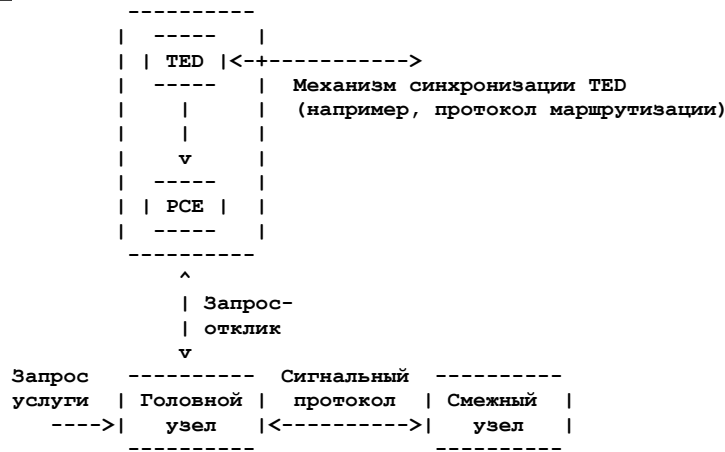


Рисунок 2. Внешний узел PCE.

5.3. Расчёт пути с несколькими PCE

На рисунке 3 показано, как на пути сигнальной службы может выполняться несколько расчётов пути на PCE. Как и в предыдущем примере головной PCC подаёт запрос к внешнему PCE, но возвращается частичный путь и следующий элемент сети должен выполнить дополнительный расчёт. Это может происходить в тех случаях, когда возвращённый путь не достигает предусмотренного адресата или рассчитанный путь является нестрогим (loose). Нисходящий элемент сети обращается к другому PCE для определения последующих интервалов (hop) пути. В этом случае все решения политики принимаются независимо на каждом PCE на основе сведений от PCC.

Отметим, что любой или оба PCE в этом случае могут быть композитными как в параграфе 5.1.

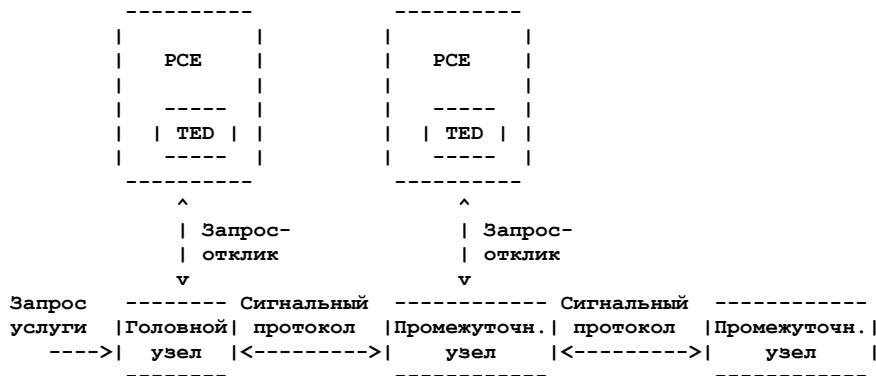


Рисунок 3. Расчёт пути с несколькими PCE.

5.4. Расчёт пути с несколькими взаимодействующими PCE

В параграфе 5.3 элементы PCE не способны предоставить полный путь для запрошенной услуги, поэтому смежному узлу требовалось передать свой запрос на расчёт. Эту проблему можно решить путём взаимодействия между PCE (рисунок 4), где PCE, к которому обращается головное устройство, подаёт запрос к другому PCE для оказания помощи при расчете.



Рисунок 4. Расчёт пути с множеством взаимодействующих PCE.

Расчёт пути с несколькими взаимодействующими PCE включает координацию между разными PCE так, что результат расчётов, выполняемых одним PCE, зависит от сведений о фрагменте пути, предоставленных другими PCE. Эта модель не указывает алгоритм распределённых вычислений, но позволяет разным PCE отвечать за расчёт частей (сегментов) пути. Взаимодействие между PCE дополнительно рассматривается в параграфе 6.6.

Отметим, что PCC может не отличать централизованный расчёт от расчёта пути с несколькими взаимодействующими PCE, т. е. узел сети или компонент PCC, запрашивающий расчёт, подаёт один запрос и получает в ответ полный путь, но на деле отклик обеспечивается координированными совместными расчётами нескольких PCE.

В этой модели все решения политики могут независимо приниматься каждым PCE на основе сведений о расчёте, полученных от предыдущего PCE. Как вариант, сведения о правилах могут явно передаваться между PCE.

5.5. Основанное на системе управления использование PCE

Следует отметить, что PCC не обязательно является LSR. Например, на рисунке 5 система NMS предоставляет головному LSR полностью рассчитанный путь для TE LSP, который тот организует через сигнализацию. NMS использует механизм плоскости управления для отправки запроса и кодирует данные с представлением в форме модуля TE MIB [RFC3812]. NMS создаёт явный путь, которые предоставляется головному LSR, используя сведения от оператора и обращаясь к PCE, который возвращает путь для использования системой NMS. Хотя на рисунке 5 элемент PCE показан как удалённый от NMS, он может размещаться и вместе с NMS.

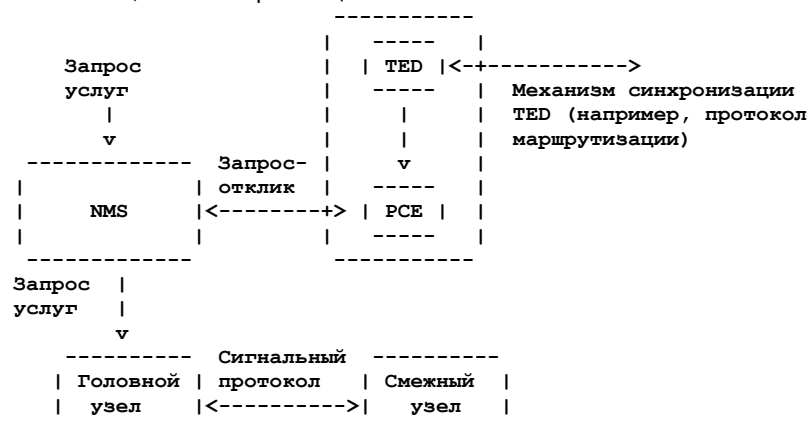


Рисунок 5. Управление на основе использования PCE.

5.6. Области стандартизации

Ниже указаны области, требующие стандартизации в архитектуре PCE.

- Взаимодействие между PCC и PCE, а также парой PCE, включая обмен сведениями о правилах.
- Требования к расширению имеющихся протоколов маршрутизации и сигнализации для поддержки обнаружения PCE и сигнализации о междоменных путях.
- Определение показателей для оценки качества, расширяемости, отзывчивости, отказоустойчивости и поддержки правил в моделях расчёта путей.
- Модули MIB, относящиеся к коммуникационным протоколам, расширениям маршрутизации и сигнализации, показателям и данным мониторинга PCE.

6. Архитектурные соображения для PCE

В этом разделе приведён список архитектурных компонентов PCE. Рассмотрение деталей конкретных реализаций PCE (конечных автоматов или алгоритмов) выходит за рамки этого документа. Отметим, что основанный на PCE расчёт путей никак не влияет на использование рассчитанных путей. Например, использование PCE не меняет способов сигнализации, поддержки и уничтожения TE LSP и относится лишь к аспектам расчёта таких TE LSP.

Этот раздел даёт архитектурное представление PCE, т. е. описывает имеющиеся компоненты и их взаимодействие. Отметим, что архитектурную модель (в частности, функциональную) разные компоненты системы PCE могут воспринимать по-разному. Например, PCC не знает, консультирует ли элемент PCE другие PCE (взгляд PCC на архитектуру PCE рассматривается в разделе 7).

6.1. Централизованный расчёт

В модели централизованного расчёта что все расчёты путей для данного домена выполняются одним централизованным PCE. Это может быть выделенный сервер (например, внешний узел PCE) или назначенный маршрутизатор (например, композитный узел PCE). В этой модели все PCC домена будут направлять свои запросы расчёта пути центральному PCE. Хотя доменом в этом контексте может быть область IGP или AS, им может быть группа узлов сети, определяемая их зависимостью от PCE.

В этой модели имеется критическая точка отказа - PCE. Для решения проблемы в модели централизованных расчётов может назначаться резервный PCE, принимающий ответственность за расчёт контролируемым способом при отказе основного PCE. Любые правила, имеющиеся на основном PCE, следует также включать в резервный, хотя правила основного узла сами могут задавать их реализацию на резервном узле. Отметим, что в каждый момент времени в домене активен лишь один элемент PCE.

6.2. Распределённый расчёт

В модели распределённых расчётов домен или сеть может включать множество PCE, между которыми распределяется расчёт путей. Данный путь может рассчитывать одним или несколькими PCE. Узел PCC может быть связан с конкретным PCE или свободно выбирать из нескольких PCE. Метод выбора выходит за рамки этого документа, но в параграфе 6.4 рассматривается обнаружение PCE, которое влияет на выбор. Реализацию правил следует согласовывать на всех доступных PCE.

Зачастую данный путь полностью рассчитывает один элемент PCE. Например, это обычно применяется в случае MPLS TE внутри одной области IGP, где выходной (LSR/композиционный) узел PCE отвечает за расчёт пути или взаимодействие с внешним PCE. Расчёт пути на нескольких PCE предполагает участие в одном расчёте более одного PCE. Примером может служить преобразование нестрогих интервалов (loose hop), выполняемое транзитными узлами PCE (LSR/композиционный) на пути MPLS TE LSP. Другим примером является использование нескольких взаимодействующих PCE для расчёта пути одного TE LSP через несколько доменов.

6.3. Синхронизация

Зачастую для одной службы требуется организовать несколько путей (например, для резервирования или распределения нагрузки). PCC, видящий потребность в расчёте более одного пути, может передать PCE серию отдельных запросов. В этом случае (без синхронизации) PCE может выполнить несколько отдельных расчётов путей, а PCC может передавать свои запросы разным PCE.

Как вариант, PCC может передать PCE один запрос для расчёта набора путей, указав при этом допустимость расчёта без синхронизации. PCE может рассчитывать каждый путь как при отправке узлом PCC серии запросов, а также может передать некоторые расчёты другим PCE по своему усмотрению. Возможен совместный расчёт элементом PCE всех путей с синхронизацией, как описано ниже.

PCC может также подавать один запрос для расчёта всех путей с синхронизацией и PCE будет одновременно рассчитывать все запрошенные пути. Такой расчёт зачастую обеспечивает лучшие результаты.

Участие нескольких PCE в расчётах по своей природе не синхронизировано, однако несколько взаимодействующих PCE можно синхронизировать под управлением одного PCE. Например, PCC может передать PCE запрос, вызывающий специфичные для домена расчёты на других PCE, до предоставления результата узлу PCC.

Желательно добавлять в протокол PCC-PCE параметр для запроса предоставления элементом PCE набора вариантов пути для использования PCC, если организация TE LSP по основному пути не удалась. Хотя дополнительные пути не всегда ведут к успеху при отказе основного, включение их в отклик PCE может сокращать издержки по сравнению с повторением узлом PCC отдельных запросов при возникновении необходимости. Это применяется в некоторых реализациях CSPF.

6.4. Обнаружение PCE и распределение нагрузки между ними

Для эффективного взаимодействия PCC с элементом PCE нужно знать местоположение PCE. Принятое здесь архитектурное решение заключается в том, что запросы PCC направляются конкретному PCE, а не передаются широкоэвентально в сеть для ответа любым PCE. Это означает, что лишь выбранный PCE будет работать с любым отдельным запросом и это сократит расход ресурсов сети при распространении запроса и ресурсы PCE, которым не нужно отвечать на запрос. Узнать местоположение PCE можно из статической локальной конфигурации PCC или с помощью основанного на протоколе механизма обнаружения, который может управляться правилами.

Если PCC известно более одного элемента PCE, у него должно быть достаточно информации для выбора подходящего для его целей PCE под управлением правил. Такая процедура выбора позволит распределять нагрузку между PCE и поддерживать PCE с разными возможностями расчётов, включая различные области видимости. Поэтому доступная PCC информация должна включать подробные сведения о возможностях PCE, которые могут быть статическими или меняющимися со временем. PCC может узнать о возможностях PCE из статической конфигурации или определять их динамически. Отметим, что даже при задании местоположения PCE в конфигурации PCC может сохраняться динамическое обнаружение возможностей PCE, поскольку динамические свойства PCE невозможно задать в конфигурации и можно лишь обнаруживать их.

Анонсы наличия PCE через Pгоxy PCE являются жизнеспособным вариантом, когда PCE не способен анонсировать себя самостоятельно. В этом случае требуется, чтобы прокси адекватно анонсировал статус и возможности PCE одновременно и с синхронизацией.

Если для обслуживания конкретного запроса на расчёт пути доступно несколько PCE, PCC должен выбрать элемент PCE для выполнения запроса. Детали такого выбора (например, для эффективного распределения нагрузки между PCE или повторного запроса расчёта после неполного расчёта или отказа) определяются PCC, могут быть основаны на правилах и выходят за рамки этого документа.

Некоторые возможности PCE, которые могут настраиваться или анонсироваться, включают:

- наборы ограничений, которые элемент может учитывать (разнообразие, SRLG¹, оптические нарушения, непрерывность длины волны и т. п.);
- расчётные возможности (например, объем вычислений за секунду);
- число уровней возможности коммутации с их конкретизацией;
- число критериев выбора пути с их конкретизацией;
- учёт состояний в PCE или возможность отправки уведомлений о лучших путях, которые могут стать доступными в будущем;
- возможности расчёта деревьев P2MP с их конкретизацией;
- возможности совместного использования ресурсов между резервными туннелями.

Эти сведения помогут PCC выбрать элемент PCE для использования.

Требования к анонсированию PCE будут описаны в отдельном документе. Отметим, что архитектура не ограничивает способы анонсирования местоположения и возможностей, а эти элементы следует считать функционально разными.

PCC может запросить у PCE определённый тип услуги, не зная возможностей PCE, и получить отклик, где указана неспособность PCE обеспечить такую услугу. Отклик может указывать возможности PCE, а также предлагать другой элемент PCE, который поддерживает нужную услугу.

6.5. Обнаружение живучести PCE

Возможность определить живучесть PCE является обязательным элементом архитектуры и может быть достигнута разными способами. Если для обнаружения PCE применяется какая-то форма анонсов (например, через расширения IGP), предполагается, что живучесть PCE будет определяться из анонсов состояния (например, IGP LSA/LSP).

¹Shared risk link group - группа каналов с общими рисками.

Неспособность PCE обслужить запрос (возможно из-за чрезмерной нагрузки) может быть указана PCC в сообщении об отказе, но сбой PCE или коммуникационного механизма в процессе обработки запроса таким способом не указать. Кроме того, при чрезмерной нагрузке у PCE может не оказаться ресурсов для передачи сообщения об отказе. Поэтому PCC следует использовать иные механизмы определения работоспособности PCE, такие как протокольные таймеры. Это особенно важно при расчёте междоменных путей, где живучесть PCE невозможно определить через протокол IGP, работающий в домене PCC.

6.6. Расчёты PCC-PCE и PCE-PCE

При выборе узлом PCC элемента PCE, не являющегося локальным для PCC, нужен протокол запрос-отклик для передачи запросов PCC на расчёт пути элементу PCE и возврата от того откликов. Требования безопасности и последствия применения такого протокола рассматриваются в разделе 10.

Запрос расчёта пути может включать многочисленные требования, включая:

- источник и адресат для пути;
- желаемую пропускную способность и другие параметры качества обслуживания (Quality of Service или QoS);
- ресурсы, их связанность (affinity), SRLG для использования или исключения;
- требуемое число несвязанных (disjoint) путей и приемлемую близость путей;
- степень отказоустойчивости, надёжности и прочности ресурсов пути;
- связанные с правилами сведения.

Степень отказоустойчивости ресурсов пути - это качественная установка уязвимости ресурсов, которые могут использоваться. Например, можно оценивать ресурсы на основе эмпирических данных (среднее время между отказами), известных рисков (например, выполнение строительных работ вблизи канала) или предубеждений (программы разработчика X постоянно ведут к отказам). PCC может запросить только надёжные ресурсы или разрешить использование любых.

В положительном отклике от PCE будет возвращаться один или несколько путей. В случае отказа при расчёте нужных путей возвращается ошибка и максимальный набор сведений о причинах неудачи, а также могут передаваться рекомендации о возможном смягчении ограничений, которое повысит вероятность успеха при повторном запросе. Отметим, что возвращаемый путь может состоять из набора строгих или нестрогих интервалов (hop) или комбинацию тех и других. Кроме того, интервал может быть задан в форме неявного абстрактного узла.

Протокол запросов и откликов нужен также для взаимодействия рассчитывающего путь PCE с другими PCE и возврата отклика в таких случаях. Запрос расчёта пути может включать большой набор требований, включая указанные выше. В положительном отклике от PCE будет возвращаться один или несколько путей. В случае отказа при расчёте нужных путей возвращается ошибка и максимальный набор сведений о причинах неудачи, а также могут передаваться рекомендации о возможном смягчении ограничений, которое повысит вероятность успеха при повторном запросе. Отметим, что возвращаемый путь может состоять из набора строгих или нестрогих интервалов (hop) или комбинацию тех и других. Кроме того, интервал может быть задан в форме неявного абстрактного узла.

Важной особенностью PCE, взаимодействующих при расчёте пути, является использование идентичных и совместимых алгоритмов расчёта и согласованных правил. Для этого может потребоваться координация путём обмена информацией между PCE. Отметим, что при взаимодействии нескольких PCE для расчёта пути важно, чтобы у них было согласованное представление о значениях таких ограничений, как стоимость, связанность (affinity) ресурсов и класс обслуживания. Это особенно важно в случаях, когда PCE отвечают за разные домены. Предполагается, что это обеспечивается согласованием правил между доменами и PCE.

В архитектуре не принимается каких-либо допущений об использовании общего протокола для взаимодействий PCC-PCE и PCE-PCE.

6.7. Синхронизация PCE TED

Как отмечено выше, PCE работает с базой TED. Сведения о состоянии сети для заполнения TED могут быть получены в домене несколькими способами.

- 1) Участие IGP в распространении сведений TE. Стандартным способом распространения информации TE внутри области IGP является использование расширений IGP [RFC3630, RFC3784¹]. Этот механизм позволяет участвующим узлам создать базу TED и служит стандартным методом, например, внутри одной области сети MPLS или GMPLS. Узел, где размещается PCE, может собирать сведения TE таким способом, поддерживая хотя бы одну маршрутную смежность с маршрутизатором в домене. Узел PCE может быть смежным или несмежным с маршрутизатором (через тот или иной метод туннелирования). Метод обеспечивает механизм, гарантирующий эффективную синхронизацию TED с состоянием сети и является обычным вариантом, например, при размещении PCE совместно с маршрутизаторами LSR в сети MPLS или GMPLS.
- 2) Синхронизация TED по отдельному каналу. Элементу PCE может быть неудобно или невозможно участвовать в IGP одного или нескольких доменов (например, доменов очень много, участие в IGP нежелательно или в некоторых доменах нет IGP с поддержкой TE). В этом случае может потребоваться определение некоего механизма, позволяющего узлу PCE извлекать TED из каждого домена. Механизм может быть инкрементным (как IGP в п. 1) или использовать передачу полной базы TED. Второй вариант может сильно ограничивать возможность синхронизации TED, что может приводить к росту частоты отказов при расчёте путей или возврату неоптимальных путей. Следует также учитывать влияние распределения TED на сеть и узел в домене, которому поручено распространять базу данных. Это особенно важно при частых изменениях в сети.
- 3) База TED может включать сведения, полученные из других источников (не IGP). Например, данные о правилах использования каналов могут быть заданы оператором. При расчёте путём может использоваться более широкий набор сведений, включающий данные о TE LSP, предоставляемых в сети. Эти данные могут включать

¹В оригинале ошибочно указан RFC 3748. Прим. перев.

маршруты TE LSP, зарезервированную пропускную способность и измеренный объем трафика через TE LSP. Такие сведения о TE LSP могут улучшить (повторную) оптимизацию TE LSP для обеспечения (ре)оптимизации «всей сети» и учесть флуктуации трафика. Подробные сведения о TE LSP могут также облегчить реконфигурацию топологии виртуальной сети (Virtual Network Topology или VNT) [MLN], где TE LSP нижележащего уровня (например, оптические пути) обеспечивают каналы TE для использования вышележащим уровнем, поскольку такая реконфигурация является также проблемой «всей сети».

Отметим, что методы синхронизации могут применяться как к внутридоменным, так и к междоменным базам TED. Кроме того, эти методы можно смешивать для использования в разных доменах. Степень синхронизации между PCE и сетью зависит от реализации и/или политики. Улучшение синхронизации обычно повышает успех путей.

Отметим также, что PCE может иметь доступ лишь к части TED, например, в случае расчёта междоменного пути, где каждый домен может управляться своей сущностью. В таких случаях у каждого PCE может быть доступ к части TED, а между PCE могут применяться кооперативные методы для расчёта сквозного пути без требования обработки каким-либо PCE полной базы TED относящейся к набору доменов в рассматриваемом TE LSP.

6.8. PCE с учётом и без учёта состояния

PCE может учитывать или не учитывать состояния. В первом случае имеется строгая синхронизация PCE не только с набором состояний сети (топология и сведения о ресурсах), но и с набором рассчитанных путей и зарезервированных ресурсов, используемых в сети. Иными словами, PCE использует при обработке запросов сведения из базы TED, а также информацию о существующих в сети путях (например, TE LSP). Хотя это позволяет рассчитать оптимальный путь и повышает вероятность успешного расчёта, для PCE с учётом состояний требуются надёжные механизмы синхронизации, поддержка большого объёма данных/состояний (например, полная связность TE LSP) и могут возникать значительные издержки в плоскости управления. Например, при наличии в домене одного PCE все расчёты TE LSP будут выполняться этим PCE, который может отслеживать все имеющиеся TE LSP и оставаться синхронизированным (все изменения состояний TE LSP должны отслеживаться элементом PCE). Для этого могут потребоваться значительные ресурсы плоскости управления. При наличии в сети множества PCE расчёт TE LSP и сведения о них распределяются между PCE, что ведёт и к распределению требуемых для расчётов ресурсов. Однако задачи синхронизации, рассмотренные в параграфе 6.7 сохраняют свою актуальность.

Поддержка базы данных с состояниями является нетривиальной задачей. Тем не менее, в среде с одним централизованным PCE этому элементу сравнительно просто запомнить все рассчитанные TE LSP, факты фактической организации TE LSP (если это можно узнать) и их уничтожения. Синхронизация TED по отдельному каналу также может быть сложной при наличии множества PCE в распределенной модели расчётов, при этом могут возникать «гонки» между PCE, проблемы расширяемости и т. п. Даже при наличии у PCE подробных сведений обо всех путях, приоритетах и уровнях учёт этих данных при расчёте пути может быть очень сложным. PCE могут синхронизировать состояния, взаимодействуя между собой, но при создании TE LSP с помощью распределенных по множеству PCE вычислений проблемы синхронизации и предотвращения гонки всё более усложняются.

Знание наличия и маршрутизации TE LSP полезно для поддержки таких приложений, как размещение высокоприоритетного TE LSP в переполненной сети так, чтобы он вытеснял как можно меньше других TE LSP (проблема минимальных возмущений). Отметим, что вытеснение на основе минимального числа каналов может не обеспечивать минимального числа нарушаемых TE LSP. Ещё одно применение связано с созданием и поддержкой топологии виртуальной сети [MLN]. Полезно также понимать, какие ещё TE LSP имеются в сети, чтобы решить как управлять смежностями по пересылке, которые имеются или должны быть созданы. Оценка издержек и преимуществ расчётов PCE с учётом состояний будет полезна для сопоставления выгод при расчёте путей с дополнительной загрузкой сети и расчётных ресурсов.

PCE без учёта состояний не помнят рассчитанных путей и каждый набор запросов обрабатывается независимо от других. Например, PCE без учёта состояния могут рассчитывать пути на основе текущих данных TED, которые могут быть не синхронизированными с фактическим состоянием сети, которое могло быть изменено недавними расчётами путей другими PCE. Отметим, что PCC может включить в свой запрос набор рассчитанных ранее путей, чтобы они были учтены, например, для предотвращения двойного учёта пропускной способности или попытки минимизировать изменения (проблема минимальных возмущений).

PCE без учёта состояний работают со сведениями о состоянии сети. База TED содержит сведения о состоянии каналов и доступности пропускной способности, расширяемые IGP или полученные иным путём. Эта информация дополнена деталями, например, текущей загрузкой пропускной способности некоторых каналов в соответствии с родством ресурсов или классами эквивалентной пересылки. Однако эти сведения не являются информацией о состоянии PCE, поэтому модель в контексте PCE всё равно считается не учитывающей состояния.

В PCE без учёта состояния может применяться ограниченная форма хранения состояний. PCE может сохранять некий контекст из недавно рассчитанных путей, чтобы не предложить использовать те же ресурсы для других TE LSP.

6.9. Мониторинг

Мониторинг PCE, несомненно, имеет важнейшее значение в любой архитектуре PCE и должен включать сбор переменных, связанных с состоянием и работой PCE. Например, требуется понимать, каким способом сохраняется синхронизация TED, с какой скоростью поступают новые запросы и какое время занимает расчёт, какие PCC используют PCE, а также знать операции протокола PCC-PCE.

6.10. Конфиденциальность

Как указано в [RFC4216], при расчёте TE LSP между провайдерами требуется возможность рассчитать путь с сохранением конфиденциальности между ядрами сетей нескольких сервис-провайдеров. Провайдеру недопустимо раскрывать сведения о своих ресурсах и топологии для поддержки расчёта межпровайдерских путей TE LSP. Таким образом, любое решение архитектуры PCE должно поддерживать возможность возврата частичных путей с помощью нестрогих переходов (например, когда каждый такой переход будет указывать граничный LSR). Это требование не относится к вопросам безопасности и связано с политикой сервис-провайдера. Конфиденциальность, целостность и проверка подлинности сообщений PCC-PCE и PCE-PCE должны обеспечиваться и описаны в разделе 10.

Желательна также возможность расчёта пути по запросу головного PCC с предоставлением этого пути в форме сегментов пути к граничному PCC домена.

6.11. Политика

Политика влияет на многие аспекты архитектуры PCE. Рассматриваются два варианта применения правил:

- в элементах архитектуры (PCC или PCE);
- в связанных с PCE коммуникациях.

Поскольку политика напрямую применима к TE LSP, она является частью сигнального механизма организации TE LSP и не описана здесь. Предполагается, что правила будут применяться в основном локально в рамках каждого PCC и PCE. Однако в этом документе необходимо задать модели политики, которые могут поддерживаться в архитектуре PCE и связанных с PCE коммуникациях. Примеры политики включают:

- выбор элемента PCE клиентом PCC;
- отклонение запросов элементом PCE на основе идентификации запрашивающего PCC;
- выбор пути элементом PCE или применение дополнительных ограничений в расчётах в зависимости от PCC, цели расчёта, времени суток и т. п.

6.11.1. Архитектура политики PCE

Два примера применения компонентов политики в архитектуре PCE показаны на рисунках 6 и 7. Компоненты политики могут также применяться в других конфигурациях PCE, показанных в разделе 5. В каждой конфигурации правила могут рассматриваться до предоставления отклика элементом PCE с возможным обращением к PCC/PCE, получающему отклик. У PCE могут быть локальные правила, влияющие на пути, выбранные по конкретному запросу PCE. Правила могут применяться на основе любых сведений, полученных от PCC.

На рисунке 6 показан компонент политики, обеспечивающий ввод для компонента PCE. Этот компонент политики может обращаться к внешней базе данных, но это выходит за рамки данного документа.



Рисунок 6. Компонент правил в комбинированном узле PCE.

Отметим, что данные политики могут передаваться как через внутренние интерфейсы, так и через внешние протокольные интерфейсы.

На рисунке 7 показана отдельная функция PCE на примере нескольких взаимодействующих PCE (сравн. с рисунком 4). Каждый PCE принимает данные локальной политики как часть процесса вычисления/определения маршрутизатора. Компоненты локальной политики могут обращаться к внешним компонентам политики или базам данных, но это выходит за рамки документа.

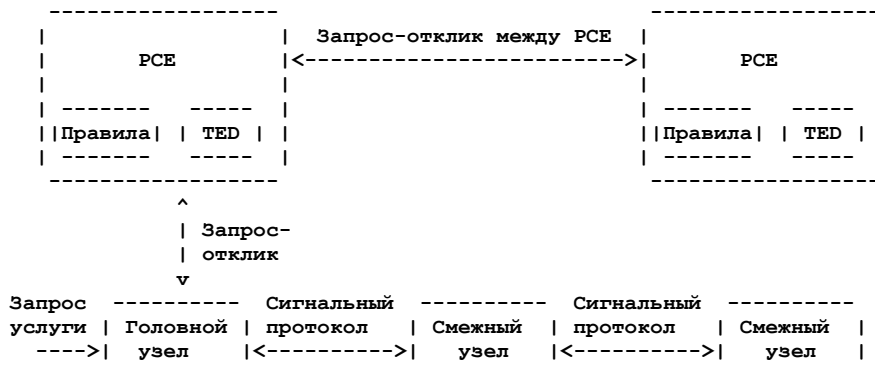


Рисунок 7. Компоненты правил для нескольких PCE.

Данные политики могут передаваться по внешним протокольным интерфейсам, включая интерфейс между PCE.

6.11.2. Реализация политики

Имеется несколько вариантов координации сведений политики, как указано ниже.

- Решения могут принимать РСС до консультации с элементами РСЕ. Этот тип решений включает выбор РСЕ, применение ограничений и интерпретацию запросов услуг.
- Решения могут приниматься независимо на РСЕ или каждом из взаимодействующих РСЕ, т. е. каждый элемент РСЕ принимает свои решения в части правил независимо от решений РСС и других РСЕ.
- Может происходить явная передача сведений политики между РСС и РСЕ или между РСЕ для обеспечения некоторого уровня координации между субъектами политики. Тип сведений, передаваемых для поддержки политики, оказывает важное влияние на правила, которые могут быть применены каждым РСЕ, а требования к обмену данными политики определяют выбор или реализацию коммуникационных протоколов, включая РСС-РСЕ, РСЕ-РСЕ и протоколы обнаружения.

6.11.3. Типы политики

В контексте РСЕ имеется несколько типов политики, указанных ниже.

- Связанные с пользователями правила оперируют сведениями, специфичными для пользователя услуги или самой услуги, т. е. сервиса, для которого рассчитывается путь, а не службы расчёта. Примеры таких данных включают содержимое объектов сообщений сигнализации или обеспечения, идентификатор порта, через который принято сообщение, VPN ID, тип опорной точки, отождествление пользователя, инициировавшего запрос. Такие правила могут применяться клиентом РСС при создании запроса на расчёт пути или элементом РСЕ при обработке запроса, если РСС предоставил элементу РСЕ достаточно информации.
- Связанные с запросом правила оперируют сведениями, относящимися к запросу на расчёт пути, которые включены в запрос. Примеры таких данных включают ограничения, разнообразие, стратегии их смягчения, а также функции оптимизации. Такие правила напрямую влияют на процесс выбора пути, поскольку они задают каналы, узлы, сегменты пути, и/или неприемлемые сегменты пути, а также сегменты, которые желательны в результирующем пути.
- Связанные с доменом правила оперируют идентификацией домена, где размещается запрашивающий РСС, и доменов, через которые проходит результирующий путь. Эти правила имеют такой же эффект, как связанные с пользователем правила, но могут применяться к группе, а не к отдельному пользователю. Примером такой политики является ограничение на информацию, публикуемую РСЕ в данном домене. РСЕ в некоторых доменах разрешается анонсировать лишь своё присутствие, тогда как другим может разрешаться анонсирование своих возможностей, процесса аутентификации клиентов и доступности расчётных ресурсов.

6.11.4. Взаимодействие с сигнализацией

При расчёте пути для междоменного TE LSP не требуется учитывать правила сигнальной плоскости, однако это может приводить к отказу при организации TE LSP или выделению меньшего, чем предполагалось, объёма ресурсов, что приведёт к некачественному обслуживанию. Поэтому, если РСЕ, вызываемый головным LSR, может видеть другие домены, ему следует учитывать политику при расчётах и знать междоменные соглашения о правилах. Если расчёт пути производится несколькими взаимодействующими РСЕ, каждый из которых отвечает за определённый домен, при расчётах, по возможности, следует учитывать правила, чтобы повысить вероятность успешной сигнализации TE LSP. В этом контексте нарушение правил при расчёте междоменного TE LSP может прерывать расчёт, о чём следует сообщать запрашивающему.

6.12. Незапрошенные взаимодействия

Взаимодействие РСС-РСЕ (см. параграф 6.6) может быть расширено за пределы простого обмена запрос-отклик. Например, РСЕ и РСС могут обмениваться сведениями о возможностях с помощью этого протокола. Кроме того, протокол может служить для сбора и передачи информации в поддержку РСЕ с учётом состояния.

В некоторых случаях РСЕ может оказаться способным обновить рассчитанный ранее путь (возможно, в ответ на изменения в сети или правилах) и тогда взаимодействие РСЕ-РСС может поддерживать «незапрошенные» сообщения о расчёте пути для представления нового пути клиенту РСС. Однако для этого может потребоваться хранение в РСЕ записи о предыдущих расчётах и наличия чёткого триггера для выполнения повторного расчёта. У РСС также должна быть возможность связать новый путь со старым и определить, следует ли использовать новый путь. Кроме того, РСС следует иметь возможность сообщать о смене пути запрашивающему РСЕ. Отметим, что взаимодействие РСЕ-РСС не является управляющим и РСС не обязан использовать дополнительный путь, полученный от РСЕ.

Эти функции легко вписываются в рассматриваемую архитектуру, но оставлены для дальнейшего обсуждения в отдельных документах, посвящённых требованиям.

6.13. Взаимодействие с откликами

Маршрутизация по откликам (crankback) - это механизм, с помощью которого отказ при создании пути или сбоя имеющегося пути может быть исправлен путём расчёта нового пути и освежения сигнализации. Такая маршрутизация основана на распространении отклика (crankback) вместе с уведомлением об отказе, чтобы можно было выполнить новый расчёт в обход места сбоя или блокировки.

В контексте РСЕ такие отклики (crankback) могут передаваться назад к головному узлу, где процесс расчёта и сигнализации повторяется с исключением из расчёта отказавшего ресурса. Однако отклик может служить для попытки устранения проблемы на промежуточных точках пути. Такие узлы повторного расчёта по отклику будут, скорее всего, границами доменов, где клиент РСС уже вызывал РСЕ. Таким образом, отказ в домене указывается входной границе домена, где выполняется попытка расчёта другого пути через домен. При неудаче о проблеме может быть проинформирован предыдущий домен, на входной границе которого может быть выполнена попытка выбрать более подходящий путь путём смены точки входа в следующий домен или выбора маршрута через иной набор доменов.

7. Взгляд клиента расчёта пути

Взгляд на архитектуру PCE, особенно на функциональную модель, слегка отличается от точки зрения PCC. Отчасти это связано с ограниченным знанием клиентом PCC о способах взаимодействия элементов PCE для ответов на его запросы, но в большей от того, что PCC интересны другие вопросы, указанные ниже.

- Выбор PCE, способного оперативно предоставить рассчитанный путь с учётом заданных ограничений.
- Сколько запросов на расчёт придётся отправить PCC? Будет ли нужный путь рассчитан первым элементом PCE (возможно во взаимодействии с другими PCE), которому отправлен запрос, или клиенту PCC придётся обращаться к другим PCE для заполнения пропусков в пути?
- Сколько ещё расчётов пути потребуется для организации в сети TE LSP?

Последний вопрос может считаться не относящимся к компетенции головного LSR, но важным ограничением, которое PCC может применить является полное вычисление пути и его представление без нестрогих интервалов (loose hop) и непростых абстрактных узлов.

Таким образом, с учётом ограниченного кругозора клиент PCC будет считать расчёт пути несколькими PCE (параграф 5.3) важным и выделит два случая. Первый случай показан на рисунке 3 с последующими запросами расчёта от других PCC на пути TE LSP. Во втором случае головной LSR подаёт несколько запросов на расчёт пути. С другой стороны, PCC не будет знать о расчёте пути несколькими PCE с взаимодействием между ними (параграф 5.4), который он не отличает от простого случая использования внешнего узла PCE (параграф 5.2). Поэтому PCC будет понимать, что модель централизованных расчётов (параграф 6.1) может по-прежнему требовать расчёта пути несколькими PCE, требуя от головного или последующих PCC передавать дополнительные запросы центральному PCE. PCC можно защитить от использования модели распределённых расчётов (параграф 6.2), поскольку первый элемент PCE, к которому клиент обращается, будет взаимодействовать с другими PCE для выполнения полного расчёта и другие запросы не потребуются.

Эти различия можно классифицировать, определив, включает ли отклик все требуемые пути и являются ли эти пути полностью явными (содержащими только строгие интервалы между простыми абстрактными узлами).

8. Оценочные показатели

Ниже указаны показатели, которые можно использовать для оценки применимости и эффективности любого решения на основе PCE. Эти показатели не применяются для определения путей и служат для оценки решений.

Оптимальность

Возможность максимального использования сети с минимальными затратами, учётом целей QoS, а также множества областей и уровней в сети. Отметим, что модели, требующие последовательного участия нескольких PCE (например, модель, описанная в параграфе 5.3), могут приводить к маршрутным петлям, если не применяются соответствующие правила.

Расширяемость

Влияние маршрутизации, сигнализации TE LSP и расчётных издержек PCE, включая число и размер сообщений (в том числе LSA, отклики (crankback), очереди, механизмы распространения и т. п.).

Распределение нагрузки

Возможность распределять расчёты путей между несколькими PCE, каждый из которых отвечает на определённую часть запросов расчёта пути.

Расчёт нескольких путей

Возможность расчёта нескольких потенциально разных путей для распределения нагрузки и защиты/восстановления, включая сквозное разнообразие и защиту в отдельных доменах.

Повторная оптимизация

Способность повторно оптимизировать TE LSP, включая возможность межуровневого сопоставления при повторной оптимизации на конкретном уровне.

Время расчёта пути

Время расчёта отдельных путей и множества разных путей, а также выполнения запросов массового расчёта путей. Отметим, что этот показатель может быть применён лишь к задачам, которые не являются NP-полными.

Стабильность сети

Способность минимизировать любые нарушения существующего состояния TE при расчёте и организации новых путей TE.

Синхронность

Способность поддерживать точную синхронизацию TED с топологией сети и состояниями ресурсов.

Скорость синхронизации

Скорость, с которой обеспечивается синхронизация TED.

Влияние на сеть

Воздействие процесса синхронизации на потоки данных в сети.

Обработка отсутствия пути

Способность обрабатывать ситуации, когда PCE не находит пути для заданного набора ограничений.

Правила

Применение правил к взаимодействиям PCC-PCE и PCE-PCE, а также к расчёту путей в соответствии с политикой организации междоменных TE LSP.

Могут учитываться и другие показатели. Оценочные показатели следует применять при рассмотрении определённой архитектуры на основе PCE. Следует оценивать также возможные компромиссы при оптимизации показателей (например, повышение оптимальности пути может влиять на время расчёта).

9. Вопросы управляемости

Архитектура PCE включает несколько элементов, которыми нужно управлять. Управление требуется для самого PCE, а также к его взаимодействию с PCC и другими PCE. Механизмы обнаружения PCE и PCC друг друга также подлежат управлению. Многие из вопросов управляемости уже рассмотрены в других параграфах этого документа.

9.1. Управление работой и правилами

Должна обеспечиваться возможность включать и отключать функцию PCE на элементе PCE, что приведёт к тому, что PCE будет воспринимать или отклонять запросы PCC или даже не получать их. Следует также рассмотреть возможность аккуратного отключения (graceful shutdown) функции PCE, чтобы в контролируемых ситуациях (например, при обновлении программ) PCE не просто «исчезал», а предупреждал своих клиентов PCC, аккуратно обслужив все запросы из очереди (возможно, выполняя, передавая другому PCE или отклоняя их).

Должна обеспечиваться возможность контролировать применение правил на PCE путём настройки конфигурации. Контроль может включать ограничение некоторых функций или алгоритмов, настройку прав доступа и приоритета клиентов PCC, а также отношения с другими PCE внутри и вне домена. Интерфейс настройки политики ещё предстоит определить. Он может быть целиком локальным или поддерживаться через стандартизированный интерфейс (например, модуль MIB).

9.2. Модели данных и информационные модели

Предполагается, что операции PCE и PCC будут моделироваться и контролироваться с помощью соответствующих модулей MIB. Таблицы в новых модулях MIB должны будут отражать взаимосвязи между объектами, а также контролировать настраиваемые параметры и сообщать из значения.

Сбор статистики будет важной частью работы PCE. У оператора должна быть возможность видеть историю взаимодействия PCC с его элементами PCE, производительность и долю успешных запросов. Для оператора важна возможность инспекции PCE и определения нагрузки на него, а также определения отдельного клиента PCC, ответственного за непропорционально большую нагрузку. Важна также возможность записи и проверки статистики взаимодействий между PCC и PCE, включая такие проблемы, как некорректно сформированные и несанкционированные сообщения, а также сообщения, отброшенные из-за перегрузки. В этом отношении управляемость явно пересекается с безопасностью. Статистика для архитектуры PCE может быть доступна через соответствующие таблицы в новых модулях MIB.

Новые модули MIB следует использовать также для уведомлений о достижении важных порогов или о важных событиях. Необходимо внимательно следить за тем, чтобы сеть не переполнялась уведомлениями простого протокола управления сетью (Simple Network Management Protocol или SNMP). Например, может быть нецелесообразно отправлять уведомление при получении элементом PCE каждого запроса на расчёт пути. В любом случае необходимо обеспечить полный контроль, чтобы уведомления можно было отключить с помощью, например, механизмов, заданных в модуле SNMP-NOTIFICATION-MIB [RFC3413].

9.3. Проверка живучести

В параграфе 6.5 обсуждалась важность способности PCC определять живучесть PCE. Методы взаимодействия PCE-PCC должны позволять PCC проверить живучесть PCE до отправки запроса а также в интервале между отправкой запроса и получением отклика.

Для PCE не так важно знать о живучести PCC и в рамках простой модели запрос-отклик это полезно лишь для:

- получения прогноза о вероятной загрузке PCE в будущем;
- возможности PCE отказаться от обработки полученного запроса.

9.4. Проверка корректности работы

Корректной работой архитектуры PCE можно считать наличие корректной связности «точка-точка» между PCC и PCE, а также достоверность рассчитанных путей. Первый вопрос относится к безопасности, которая может быть усилена путём аутентификации и регистрации событий в журнале с их отслеживанием, как описано в параграфе 9.1. Это может также относиться к маршрутизации для обеспечения связности PCC-PCE.

Проверка рассчитанных путей сложнее, однако требуемые для этого сведения могут быть получены оператором из таблиц MIB, при условии ведения полных записей об ограничениях, переданных в запросе, и рассчитанных путей, а также дополнительных сведений, представленных элементом PCE (например, применённые смягчения ограничений).

9.5. Требования к другим протоколам и функциональным компонентам

На этапе архитектурного проектирования невозможно полностью разобраться с влиянием на другие протоколы и функциональные компоненты, поскольку работа на решении ещё не завершена. Однако можно отметить ряд наблюдений.

- Зависимость от базовых транспортных протоколов

Для взаимодействий PCE-PCC могут быть выбраны имеющиеся протоколы, обеспечивающие транспортные механизмы, и некоторые вопросы управляемости, отмеченные выше, можно передать этим протоколам.

- Использование имеющихся протоколов для обнаружения

Не превосходящая требований и решений по обнаружению PCE (см. параграф 6.4), можно допустить применение для этого имеющихся протоколов и передачу им некоторых вопросов управляемости (см. выше).

- Влияние на LSR и сигнализацию TE LSP

Основным примером PCC в этой архитектуре является LSR MPLS или GMPLS. Поэтому нужно уделить внимание управляемости LSR и добавочным ограничениям управляемости сигнальных протоколов TE LSP.

В дополнение к описанной выше возможности управления PCC нужна настройка LSR, указывающая, будет ли маршрутизатор использовать удалённый PCE (вариантами являются поэтапная маршрутизация или предоставление функции PCE. Вероятно, будет важна способность LSR различать TE LSP, предоставленные сигнальным сообщением от другого LSR, оператором или PCE, а для указанного сигнальным сообщением пути способность увидеть улучшение или расширение пути элементом PCE.

- Использование имеющихся моделей и механизмов политики

Поскольку механизмы поддержки политики могут быть очень разными, следует разобраться с возможностями применения имеющихся механизмов в PCE. Желание воспользоваться выполненной работой не следует считать требованием применять определённое решение или протокол.

9.6. Влияние на работу сети

Эта архитектура может оказывать двойное влияние на работу сети. Она увеличивает время организации TE LSP, поскольку запрос нужно передать и обработать на удалённом PCE, и может вызывать перегрузку в сети, если за короткое время будет подано большое число запросов на расчёт путей. Эти проблемы наиболее заметны в загруженных сетях и после отказов, хотя влияние проблем можно смягчить, если заранее рассчитать защитные пути или распределить расчётную нагрузку между элементами PCE.

Проблему возможных перегрузок в процессе восстановления после отказов можно смягчить за счёт применения рассчитанных заранее схем защиты, таких как быстрая перемаршрутизация (fast reroute).

Важен упреждающий контроль перегрузок в сети, поскольку реактивное управление в таких ситуациях может оказаться невозможным. Операторам следует предоставлять возможность ограничивать для клиентов PCC частоту отправки запросов элементам PCE, а PCE следует предоставлять возможность сообщать о надвигающейся перегрузке (по заданному порогу) как оператору, так и клиентам PCC.

9.7. Другие вопросы

Других вопросов управления не было отмечено.

10. Вопросы безопасности

Влияние применения архитектуры на основе PCE должно рассматриваться в свете воздействия на безопасность используемых в сети протоколов и методов маршрутизации и сигнализации. При использовании PCE лишь внутри домена это влияние может быть более слабым, но рост междоменных потоков информации и облегчение организации междоменных путей может повысить уязвимость к атакам.

Особое значение имеет влияние на конфиденциальность, присущее архитектуре на основе PCE в многодоменных сетях. Такое решение не обязательно будет снижать уровень защиты, но оно **должно** тщательно проверяться.

Предполагается, что заявления о применимости конкретных сочетаний сигнализации, маршрутизации и методов расчёта путей будут содержать подробные разделы о безопасности.

Отметим, что использование нелокального PCE (не размещённого вместе с PCC) создаёт дополнительные проблемы безопасности, из которых наиболее важны:

- перехват запросов или откликов PCE;
- выдача атакующего за PCE или PCC;
- фальсификация сведений TE, правил или возможностей PCE;
- атаки на отказ в обслуживании для коммуникационных механизмов PCE или PCE.

Предполагается, что в решениях PCE эти проблемы будут устранены с помощью методов аутентификации и защиты.

11. Благодарности

Авторы выражают искреннюю признательность Arthi Ayyangar, Zafar Ali, Lou Berger, Mohamed Boucadair, Igor Bryskin, Dean Cheng, Vivek Dubey, Kireeti Kompella, Jean-Louis Le Roux, Stephen Morris, Eiji Oki, Dimitri Papadimitriou, Richard Rabbat, Payam Torab, Takao Shimizu, Raymond Zhang (перечислены в алфавитном порядке) за их отзывы и предложения. Lou Berger внёс полезный и детальный вклад в обсуждение правил в этом документе.

Спасибо также Pekka Savola, Russ Housley и Dave Kessens за отзывы и конструктивные дискуссии на финальных этапах публикации.

12. Литература

- [RFC2702] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M., and J. McManus, "Requirements for Traffic Engineering Over MPLS", [RFC 2702](#), September 1999.
- [RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", [RFC 4364](#), February 2006.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.
- [RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.
- [RFC3413] Levi, D., Meyer, P., and B. Stewart, "Simple Network Management Protocol (SNMP) Applications", STD 62, RFC 3413, December 2002.
- [RFC3473] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.
- [RFC3784] Smit, H. and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", RFC 3784, June 2004.
- [RFC3812] Srinivasan, C., Viswanathan, A., and T. Nadeau, "Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Management Information Base (MIB)", RFC 3812, June 2004.

- [RFC4105] Le Roux, J.-L., Vasseur, J.-P., and J. Boyle, "Requirements for Inter-Area MPLS Traffic Engineering", RFC 4105, June 2005.
- [RFC4216] Zhang, R. and J.-P. Vasseur, "MPLS Inter-Autonomous System (AS) Traffic Engineering (TE) Requirements", RFC 4216, November 2005.
- [MLN] Shiomoto, K., Papdimitriou, D., Le Roux, J.-L., Vigoureux, M., and D. Brungard, "Requirements for GMPLS-based multi-region and multi-layer networks (MRN/MLN)", Work in Progress¹, June 2006.

Адреса авторов

Adrian Farrel

Old Dog Consulting
E-Mail: adrian@olddog.co.uk

Jean-Philippe Vasseur

1414 Massachusetts Avenue
Boxborough, MA 01719
USA
E-Mail: jpv@cisco.com

Jerry Ash

AT&T
Room MT D5-2A01
200 Laurel Avenue
Middletown, NJ 07748,
USA
Phone: (732)-420-4578
Fax: (732)-368-8659
E-Mail: gash@att.com

Полное заявление авторских прав

Copyright (C) The Internet Society (2006).

К этому документу применимы права, лицензии и ограничения, указанные в BCP 78, и, за исключением указанного там, авторы сохраняют свои права.

Этот документ и содержащаяся в нем информация представлены "как есть" и автор, организация, которую он/она представляет или которая выступает спонсором (если таковой имеется), Internet Society и IETF отказываются от каких-либо гарантий (явных или подразумеваемых), включая (но не ограничиваясь) любые гарантии того, что использование представленной здесь информации не будет нарушать чьих-либо прав, и любые предполагаемые гарантии коммерческого использования или применимости для тех или иных задач.

Интеллектуальная собственность

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

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

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

Подтверждение

Финансирование функций RFC Editor обеспечено IETF Administrative Support Activity (IASA).

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

¹Опубликовано в RFC 5212. Прим. перев.