Internet Engineering Task Force (IETF)
Request for Comments: 6887
Category: Standards Track
ISSN: 2070-1721

D. Wing, Ed.
Cisco
S. Cheshire
Apple
M. Boucadair
France Telecom
R. Penno
Cisco
P. Selkirk
ISC

April 2013

Протокол управления портом (PCP)

Port Control Protocol (PCP)

Аннотация

Протокол управления портом (PCP) позволяет хостам IPv6 или IPv4 управлять трансляцией и пересылкой входящих пакетов IPv6 или IPv4 на устройствах NAT[1] или простых межсетевых экранах (МСЭ), а также оптимизировать свои исходящие сообщения NAT keepalive.

Статус документа

Этот документ относится к категории проектов стандартов (Internet Standards Track).

Документ является результатом работы IETF[2] и представляет согласованный взгляд сообщества IETF. Документ прошёл открытое обсуждение и был одобрен для публикации IESG[3]. Дополнительная информация о документах BCP представлена в разделе 2 документа RFC 5741.

Информация о текущем статусе этого документа, обнаруженных ошибках и способах обратной связи может быть найдена по ссылке http://www.rfc-editor.org/info/rfc6887.

Оглавление

[1]Network Address Translator - транслятор сетевых адресов.
[2]Internet Engineering Task Force - комиссия по исследованиям Internet.
[3]Internet Engineering Steering Group - комиссия по решению инженерных задач Internet.

1 Введение

Протокол PCP обеспечивает механизм контроля за пересылкой входящих пакетов устройствами восходящего направления (upstream) типа NAT64[1], NAT44[2] и межсетевых экранов IPv6 и IPv4, а также механизм снижения трафика keepalive от приложений. Протокол PCP разработан для реализации в контексте CGN[3] и небольших NAT (например, домашних), а также маршрутизаторов CPE[4] с поддержкой протоколов IPv6/IPv4 или только IPv6 и всех существующих в настоящее время сценариев перехода к маршрутизаторам CPE с поддержкой только IPv6. PCP позволяет хостам поддерживать серверы, работающие достаточно долго (например, видеокамера с сетевым подключением) или в течение короткого времени (например, сеанс игры или телефонный звонок) при нахождении за устройством NAT, включая устройства CGN у провайдера Internet или МСЭ IPv6, встроенные в маршрутизаторы CPE.

[1]Network Address Translator IPv6/IPv4 - транслятор адресов IPv6/IPv4.
[2]Network Address Translator IPv4/IPv4 - транслятор адресов IPv4/IPv4.
[3]Carrier-Grade NAT - транслятор адресов операторского уровня.
[4]Customer Premises Equipment - пользовательское оборудование.

PCP позволяет приложениям создавать отображения внешних адресов IP, протоколов и портов на внутренние адреса IP, протоколы и порты. Эти отображения нужны для входящих соединений с машинами, расположенными за устройством NAT или МСЭ.

После организации отображения для входящих соединений необходимо сообщить удаленным компьютерам адрес IP, протокол и порт для входящих соединений. Обычно такое оповещение зависит от приложения. Например, компьютерная игра может использовать специальный рандеву-сервер, обслуживающий эту игру (или поддерживаемый её разработчиком), телефон SIP будет использовать SIP-прокси, а клиент обнаружения служб на базе DNS[1] [RFC6763] будет применять [RFC2136] [RFC3007]. PCP не поддерживает функций «сводника». Такие функции могут работать с протоколом IPv4, IPv6 или поддерживать оба протокола. В зависимости от этого, а также от поддержки приложениями IPv4 или IPv6, клиенту PCP может потребоваться отображение для IPv4, IPv6 или обоих протоколов.

Многие дружественные к NAT приложения часто передают сообщения прикладного уровня с целью предотвращения разрыва сессий на устройстве NAT. Такие сообщения обычно называют NAT keepalive, хотя они не передаются непосредственно устройству NAT (они проходят через NAT). Такие приложения могут снизить частоту передачи сообщений NAT keepalive за счёт использования PCP для определения (и влияния) времени жизни отображений NAT. Это позволяет снизить расход полосы в пользовательской сети доступа, трафик к серверам и расход батарей для мобильных устройств.

Многие устройства NAT и включают в свой состав шлюз прикладного уровня (ALG[2]) с целью создания отображений для приложений, которые создают дополнительные потоки или принимают входящие соединения. ALG, встроенные в NAT могут также изменять содержимое пакетов приложений. Опыт производителей оборудования показывает, что такие ALG мешают развитию протоколов. PCP обеспечивает приложениям возможность создавать свои отображения в устройствах NAT и МСЭ для снижения уровня использования ALG в МСЭ и NAT.

2 Сфера применения

2.1 Сценарии развёртывания

PCP может использоваться в разных сценариях, включая:

- базовую трансляцию NAT [RFC3022];
- трансляцию сетевых адресов и портов [RFC3022], широко используемую в домашних устройствах NAT;
- операторские NAT [RFC6888];
- DS-Lite [RFC6333];
- NAT с поддержкой уровня 2 [L2NAT];
- Dual-Stack Extra Lite [RFC6619]
- NAT64 без учёта [RFC6145] и с учётом [RFC6146] состояния;
- управление МСЭ IPv4 и IPv6 [RFC6092];
- трансляцию сетевых префиксов IPv6-IPv6 (NPTv6) [RFC6296].

2.2 Поддерживаемые протоколы

Коды операций PCP (Opcode), определённые в данном документе, разработаны для поддержки протоколов транспортного уровня с 16-битовой нумерацией портов (например, TCP, UDP, SCTP[3] [RFC4960] и DCCP[4] [RFC4340]). Протоколы, которые не пользуются номерами портов (например, RSVP[5], ESP[6] [RFC4303], ICMP, ICMPv6) поддерживаются для межсетевых экранов IPv4 и IPv6, а также NPTv6, но не входят в сферу действия функций NAT.

2.3 Пользовательская сеть с одним подключением

В PCP предполагается модель адресации IP с однодомным подключением. Т. е., для данного адреса IP существует только один принятый по умолчанию маршрут для доступа к другим хостам Internet с данного IP-адреса. Это важно, поскольку после организации отображения PCP, изменения входящих пакетов (например, TCP SYN) и их доставки хосту, исходящие отклики (например, TCP SYNACK) будут идти по тому же (обратному) пути, на котором пройдут через то же устройство NAT, которое сможет соответствующим образом изменить исходящие пакеты. Это ограничение связано с тем, что в противном случае потребовалось бы поддерживающее PCP устройство NAT для каждого исходящего пути (поскольку хост не может достоверно знать исходящий путь), а клиент должен будет организовать одинаковые отображения «внешний-внутренний» для каждого шлюза NAT, что в общем случае не представляется возможным (поскольку на других устройствах NAT требуемый внешний порт уже может быть отображён на другой хост).

3 Терминология

Ключевые слова **необходимо** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не нужно** (SHALL NOT), **следует** (SHOULD), **не следует** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе должны интерпретироваться в соответствии с «Key words for use in RFCs to Indicate Requirement Levels» [RFC2119].

[1] DNS-Based Service Discovery.
[2] Application Layer Gateway.
[3] Stream Control Transmission Protocol.
[4] Datagram Congestion Control Protocol.
[5] Resource Reservation Protocol.
[6] IP Encapsulating Security Payload.

Internal Host - внутренний хост

Хост, обслуживаемый шлюзом NAT или защищённый МСЭ. Это хост, который будет получать входящий трафик в результате запроса на отображение PCP, или хост, инициировавший динамическое исходящее отображение (например, передал TCP SYN) через межсетевой экран или NAT.

Remote Peer Host - удалённый хост-партнер

Хост, с которым обменивается данными внутренний хост. Это может быть другой (или даже тот же самый) внутренний хост, в этом случае при использовании NAT устройство NAT должно будет «заворачивать» трафик назад [RFC4787].

Internal Address - внутренний адрес

Адрес внутреннего хоста, обслуживаемого шлюзом NAT или защищённого МСЭ.

External Address - внешний адрес

Адрес внутреннего хоста с точки зрения других внешних партнёров в Internet, с которыми данный хост взаимодействует после трансляции во всех устройствах NAT на пути. Внешний адрес в общем случае является глобально маршрутизируемым (т. е., не приватным). Если внутренний хост защищён только МСЭ и на пути нет трансляции адресов, внешний адрес будет совпадать с внутренним.

Endpoint-Dependent Mapping (EDM) - зависящее от конечной точки отображение

Операция NAT, в которой для неявного отображения, организуемого исходящим трафиком (например, TCP SYN) от одного внутреннего адреса, протокола и порта к удаленным партнёрам и портам, могут выделяться разные внешние порты, а при следующем запросе отображения PCP для данного внутреннего адреса, протокола и порта может быть выделен ещё один внешний порт. Эта операция может применяться для отображений, зависящих от адреса (Address-Dependent Mapping) или адреса и порта (Address and Port-Dependent Mapping) [RFC4787].

Endpoint-Independent Mapping (EIM) - независимое от конечной точки отображение

Операция NAT, при которой для всех отображений от одного внутреннего адреса, протокола и порта выделяется один внешний адрес ии порт.

Remote Peer Address - адрес удалённого партнёра

Адрес удалённого партнёра с точки зрения внутреннего хоста. Обычно это глобально маршрутизируемый адрес. В тех случаях, когда удалённый партнёр сам обслуживается шлюзом NAT, адрес удалённого партнёра фактически является его внешним адресом, но, поскольку удалённая трансляция не видна для программ внутреннего хоста, это различие можно игнорировать в контексте данного документа.

Third Party - третья сторона

В общем случае внутренний хост сам управляет своими отображениями, используя запросы PCP, и внутренний адрес в этих отображениях совпадает с адресом отправителя запросов PCP.

В некоторых случаях устройство может организовывать отображения от имени других устройств, которые не поддерживают PCP, - присутствие опции THIRD_PARTY в запросе MAP говорит, что в качестве внутреннего адреса следует использовать указанный адрес вместо IP-адреса отправителя в запросе PCP.

Mapping, Port Mapping, Port Forwarding - отображение, отображение портов, перенаправление портов

Отображение NAT организует связь между внутренними и внешними адресами IP, протоколами и портами. Точнее говоря, создаётся правило трансляции получателя, в соответствии с которым пакеты **для внешнего** адреса IP, протокола и порта меняются так, чтобы они попадали **на внутренний** адрес IP, протокол и порт, а в обратном направлении происходит аналогичное преобразование параметров отправителя. Для «чистого» МСЭ отображение «транслирует» адрес IP, протокол и порт в те же значения (в дополнение применяются правила фильтрации).

Mapping Types - типы отображений

Имеется три критерия для классификации типов отображений: способ организации (явный/неявный), основное назначение (входное/выходное) и способ удаления (динамические/статические). Неявные отображения являются «побочным» результатом некоторых других операций, явные создаются с помощью специального механизма. Выходные отображения служат прежде всего для поддержки исходящих соединений, а входные - для входящих вызовов. Динамические отображения автоматически удаляются по истечении срока существования, а статические сохраняются до тех пор, пока не будут удалены пользователем.

- Неявные динамические отображения организуются, как «побочный» эффект исходящего трафика типа пакетов TCP SYN или UDP. Такие пакеты изначально не направлены на создание состояния в NAT (или МСЭ), но они могут вызывать такой эффект при прохождении через устройство NAT (или МСЭ). Неявные динамические отображения обычно существуют в течение ограниченного срока, хотя этот срок в общем случае не известен использующему отображение клиенту.

- Явные динамические отображения создаются в результате PCP-запросов MAP и PEER. Подобно аренде адресов DHCP явные динамические отображения имеют ограниченное время жизни и это время сообщается клиенту. Как и для DHCP клиент, желающий сохранять отображение, должен время от времени обновлять его, чтобы время жизни не истекло. Если клиент PCP «уходит», все созданные им отображения автоматически удаляются по истечении времени их жизни.

- Явные статические отображение организуются путём настройки конфигурации (например, с помощью команд или графического интерфейса) и существуют до тех пор, пока пользователь не изменит данную конфигурацию.

Явные и неявные динамические отображения создаются по запросу (явному или неявному) внутреннего хоста и существуют в течение ограниченного срока. По истечении этого срока отображение удаляется, если внутренний хост заранее не принял мер по продлению (например, передача дополнительного трафика или запроса PCP).

Статические отображения по своей природе всегда являются явными. Эти отображения отличаются от явных динамических отображений неограниченным сроком существования (до удаления вручную), в остальном не отличаясь от явных отображений MAP.

Хотя все отображения (по необходимости) являются двухсторонними (большинству коммуникаций Internet для работы нужен двухсторонний обмен данными), при обсуждении отображений может оказаться полезным их разделение по направлению преимущественного использования.

- Выходные отображения существуют, прежде всего, для поддержки исходящих соединений. Например, когда хост вызывает функцию **connect()** для организации исходящего соединения, шлюз NAT будет создавать неявное динамическое отображение для поддержки такого соединения.

- Входные отображения существуют, прежде всего, для обеспечения серверам возможности приёма входящих соединений. В общем случае, когда хост вызывает функцию **listen()** для прослушивания входящих

соединений, устройство NAT будет неявно создавать отображение для поддержки входящих соединений. Для явного создания динамического входного отображения может использоваться запрос PCP MAP.

Явные статические (вручную) и динамические (MAP) отображения позволяют внутренним хостам получать входящий трафик, который не является откликом для каких-либо предшествующих исходящих коммуникаций (т. е., позволяет внутренним хостам функционировать в качестве доступных из сети Internet серверов).

PCP Client - клиент PCP

Программный экземпляр PCP для отправки запросов PCP серверу PCP. На одном хосте может быть одновременно запущено несколько клиентов PCP. В одной локальной сети может размещаться несколько клиентов PCP. Клиент PCP может делать запросы PCP от имени другого устройства, предоставившего ему такие полномочия. Примером клиента PCP могут служить сетевые функции устройств UPnP IGDv1[1] [IGDv1]. Сервер PCP на шлюзе NAT, который сам по себе является клиентом другого шлюза NAT (вложенная трансляция), может быть PCP-клиентом восходящего NAT.

PCP-Controlled Device - управляемое PCP устройство

Устройство NAT или МСЭ, которое контролирует и изменяет потоки пакетов между внутренними хостами и их удалёнными партнёрами. Протокол PCP управляет отображениями на этом устройстве.

PCP Server - сервер PCP

Программный экземпляр PCP, размещённый на управляемом PCP устройстве, который принимает запросы PCP от клиентов PCP и организует в ответ на эти запросы соответствующие состояния.

Subscriber - абонент (подписчик)

Субъект коммерческого учёта ISP. Подписчик может получать от коммерческого ISP один адрес IP (этот адрес может совместно использоваться множеством хостов, расположенных за шлюзом NAT, которые с точки зрения ISP будут выглядеть, как один хост) или множество таких адресов. В любом случае предоставленные коммерческим ISP адреса IP могут впоследствии транслироваться шлюзом CGN на стороне ISP.

4 Отношения сервера PCP с его устройствами, управляемыми PCP

Сервер PCP получает запросы PCP и выдаёт отклики на них. Серверы PCP обычно совмещаются с устройствами NAT или МСЭ, как показано на рисунке 1. Функциональность PCP может также обеспечиваться неким другим устройством, которое взаимодействует с устройством(ами) NAT или МСЭ на основе фирменного (proprietary) механизма. С точки зрения клиента PCP такое «расщеплённое» устройство неотличимо от интегрированного.

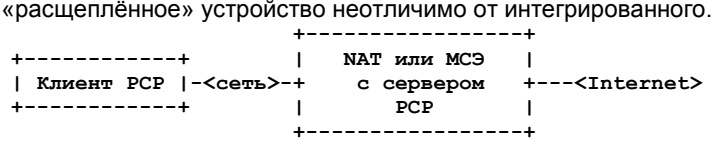

Рисунок 1. Межсетевой экран или устройство NAT с поддержкой PCP.

Устройство NAT или МСЭ между клиентом PCP и сетью Internet может реализовать простой или расширенный набор функций МСЭ. Как дополнительный эффект используемой технологии (например, транслятор адресов и портов обычно будет требовать организации соединений из внутренней сети в направлении Internet просто в силу того, что при трансляции изменяются номера портов) или в соответствии с явно заданными правилами незапрошенный трафик из Internet будет отвергаться. Некоторые МСЭ отвергают часть незапрошенного трафика из Internet (например, для большинства портов TCP и UDP), принимая другую его часть (например, UDP 500 и IP ESP) [RFC6092]. Рассмотрение используемой по умолчанию фильтрации выходит за рамки PCP. Если клиентское устройство хочет получать трафик и поддерживает PCP, но не знает заранее принятой по умолчанию политики фильтрации, ему **следует** использовать PCP для запроса отображения, позволяющего получать желаемый трафик.

5 Адреса фиксированного размера

Для простоты генерации и разбора пакетов с запросами и откликами PCP всегда использует 128-битовые адреса IP как для IPv6, так и для IPv4.

При работе с IPv6 128-битовые поля адресов IP просто содержат адреса IPv6 без каких-либо изменений.

Для IPv4 адреса IPv4 отображаются на адреса IPv6 в соответствии с [RFC4291] (::ffff:0:0/96). Первые 80 битов имеют нулевые значения, следующие 16 битов - 1, а в последних 32 битах размещается адрес IPv4. Это позволяет однозначно идентифицировать естественные адреса IPv6, поскольку отображённые на IPv4 адреса IPv6 [RFC4291] не будут корректными для отображения.

При проверке отображённых на IPv4 адресов IPv6 **должны** проверяться все первые 96 битов, недостаточно убедиться в том, что биты 81-96 имеют значение 1.

Адрес all-zeros IPv6 **должен** представляться заполнением всего 128-битового поля адреса IP нулями (::).

Адрес all-zeros IPv4 **должен** содержать 80 битов нулей, 16 битов единиц и снова 32 бита нулей (::ffff:0:0).

6 Устройство протокола

PCP можно рассматривать, как протокол «запрос-отклик», он похож на другие протоколы такого типа, работающие на основе транспорта UDP и может быть реализован столь же хорошо. Можно рассматривать протокол и как взаимодействие вида «намек-уведомление» (hint/notification), что может упростить реализацию.

Взаимодействие между клиентами и серверами PCP следует рассматривать не как поток пар «запрос-отклик», а скорее как два связанных потока сообщений, передаваемых во встречных направлениях:

- от клиента к серверу передаётся поток «намёков», в которых клиент показывает серверу, какое отображение он хотел бы получить;
- поток уведомлений от сервера PCP, в которых сервер информирует клиента о созданных отображениях.

[1]Universal Plug and Play Internet Gateway Device.

В той или иной степени такой подход требуется для всех протоколов «запрос-отклик» на основе UDP, поскольку пакеты UDP могут теряться, дублироваться и доставляться с нарушением порядка.

В таком представлении протокола клиент передаёт серверу свои «намёки» с разными интервалами, сообщая о своих желаниях, а сервер передаёт уведомления с актуальным состоянием отображений для данного клиента. Эти два потока сообщений коррелируют между собой в том смысле, что запрос (пожелание) клиента обычно вызывает отклик (уведомление) от сервера. Однако корреляция достаточно слабая, поскольку запрос клиента может не породить серверного отклика (в случае потери пакета) или отклик может быть сгенерирован без предшествующего ему запроса (изменение конфигурации сервера - например, смена внешнего адреса IP на шлюзе NAT).

Точное число передаваемых клиентом сообщений зависит от синхронизации состояний на стороне клиента с учётом (i) наличия отправленных запросов, на которые от сервера ещё не получено отклика, (ii) получения от сервера отклика, указывающего оставшийся срок существования отображения. Это является причиной того, что повторы передачи PCP и обновления в точности совпадают с переданными ранее переданными пакетами. Обычно повторные пакеты передаются с экспоненциально нарастающими интервалами при ожидании ответа от сервера, а пакеты обновления передаются с экспоненциально уменьшающимися интервалами по мере приближения срока завершения отображения. Однако с точки зрения сервера эти пакеты идентичны и сообщают о желании клиента создать или сохранить отображение.

Сервер PCP обычно передаёт отклик, как прямой результат запроса от клиента, но это выполняется не всегда. Например, если сервер перегружен и не сможет ответить сразу, ему позволено просто игнорировать запрос клиента. Кроме того, при изменении конфигурации шлюза NAT или МСЭ в силу внешних причин сервер PCP может отправлять клиентам незапрошенные отклики, информирующие их о новом состоянии отображений. Такие случаи предполагаются достаточно редкими, поскольку они могут нарушать работу клиентов. Однако при возникновении такой ситуации протокол PCP обеспечивает серверам способ своевременного оповещения клиентов без ожидания от них запросов на периодическое обновление.

Сказанное выше помогает понять, почему в запросах и откликах PCP нет идентификаторов транзакций. Эти идентификаторы просто не нужны и будут вносить неоправданные ограничения в работу протокола, а также осложнять его реализацию. Отклик сервера PCP (т. е., уведомление) является самодостаточным и полным. Он включает внутренний и внешний адрес, протокол и порты для отображения, а также указывает время жизни отображения. Если клиенту нужно такое отображение, он просто обновляет своё состояние на основе полученных данных. Если отображение не нужно клиенту, он может просто игнорировать полученное сообщение. При получении незапрошенных уведомлений клиент не обязан выполнять какие-либо действия. В современных сетях шлюзы NAT могут использовать статические отображения, о которых клиент не имеет явных сведений и не может как-то влиять на них. Клиентское устройство может быть напрямую подключено к сети Internet с глобально маршрутизируемым адресом IP и в таком случае он имеет «отображение» для всех прослушиваемых им портов. Такое устройство само несёт ответственность за свою безопасность и не может предполагать, что какое-то иное устройство в сети блокирует входящие пакеты.

7 Общий формат заголовков запросов и откликов

Все сообщения PCP передаются по протоколу UDP с максимальным размером поля данных UDP в 1100 октетов. Сообщения PCP содержат заголовок запроса или отклика, включающий код операции Opcode, относящиеся к операции данные (Opcode-specific information) и может также включать одну или несколько опций. Все числовые значения, размер которых превышает 1 октет (например, коды результата, время жизни, время Epoch и т. п.) используют сетевой порядок IETF (т. е., сначала передаётся старший октет). Нечисловые значения представляются как на всех платформах без перестановки байтов (например, адреса IP и номера портов размещаются в сообщениях PCP так же, как это делается в заголовках IP или TCP).

Схема пакета использует общий заголовок, операции клиентов и серверов PCP описаны ниже. Приведённая в данном разделе информация применима для всех операций (Opcode). Операции, определённые в данном документе, описаны в разделах 10, 11 и 12.

7.1 Заголовок запроса

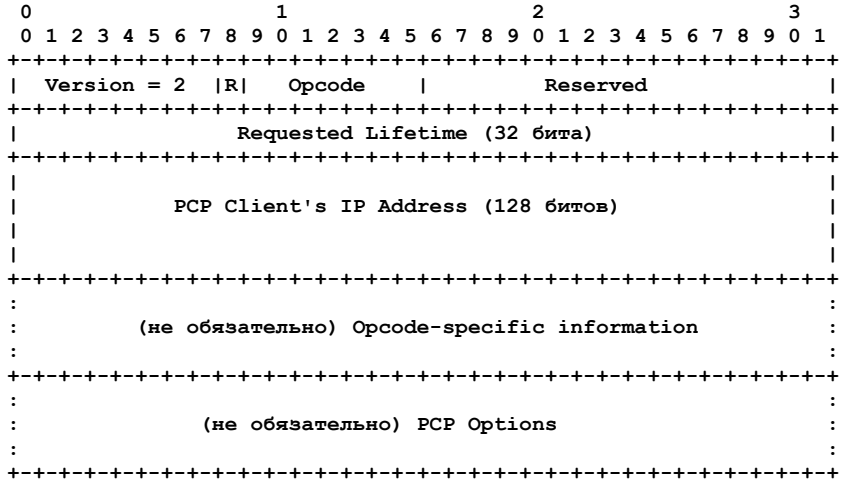

Рисунок 2. Формат запроса.

Формат запросов показан на рисунке 2.

Version - версия

Данный документ задаёт для протокола номер версии 2. Соответствующие данной спецификации клиенты и серверы используют значение 2. Это поле служит для согласования версии, описанного в разделе 9.

R - запрос/отклик

Указывает запрос (0) или отклик (1).

Opcode - код операции

7-битовое значение, указывающее код выполняемой операции. Коды операций MAP и PEER определены в разделах 11 и 12.

Reserved - резерв

16 резервных битов. При передаче **должны** устанавливаться в 0, а при получении **должны** игнорироваться.

Requested Lifetime - запрошенное время жизни

32-битовое целое число без знака, определяющее срок жизни отображения в секундах (0 - 2^{32}-1). Это значение используется определёнными в данном документе операциями MAP и PEER.

PCP Client's IP Address - IP-адрес клиента PCP

Адрес отправителя IPv4 или IPv6 в заголовке IP, использованном клиентом PCP при передаче данного запроса PCP. Адреса IPv4 отображаются на адреса IPv6. Поле PCP Client's IP Address из заголовка сообщения PCP используется для обнаружения неожиданных устройств NAT на пути между клиентом PCP и управляемым PCP устройством NAT или МСЭ (см. параграф 8.1).

Opcode-specific information – специфические для операции данные

Дополнительные данные для Opcode. Размер данных определяется значением Opcode.

PCP Options - опции PCP

Необязательное поле с одной или множеством опций PCP, применимых для запроса и кода (см. параграф 7.3).

7.2 Заголовок отклика

Формат откликов показан на рисунке 3.

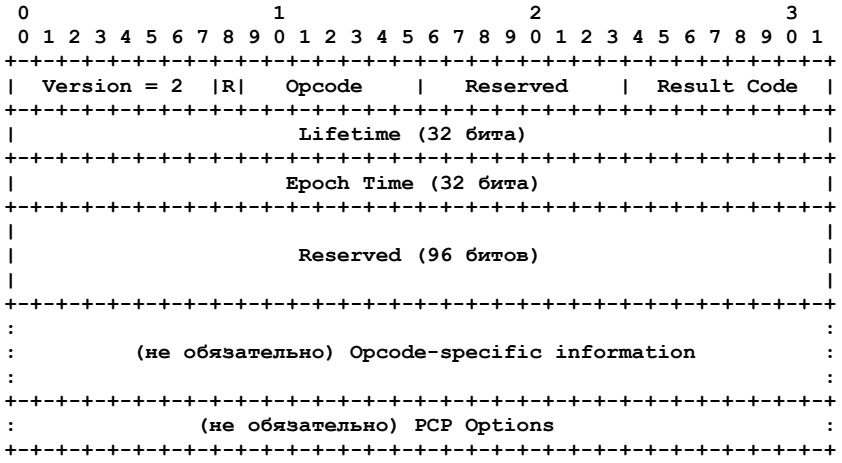

Рисунок 3. Формат отклика.

Version - версия

Отклики серверов, соответствующих данной спецификации, **должны** указывать версию 2. Устанавливается сервером.

R - запрос/отклик

Указывает запрос (0) или отклик (1). Во всех откликах **должно** использоваться значение 1. Устанавливается сервером.

Opcode - код операции

7-битовое значение, указывающее код выполняемой операции. Сервер копирует это значение из запроса.

Reserved - резерв

8 резервных битов. При передаче **должны** устанавливаться в 0, а при получении **должны** игнорироваться. Устанавливаются сервером.

Result Code - код результата

Код результата для данного отклика. Возможные значения описаны в параграфе 7.4. Устанавливается сервером.

Lifetime - время жизни

32-битовое целое число без знака, определяющее срок жизни отображения в секундах (0 - 2^{32}-1). В откликах об ошибках это значение указывает предполагаемое время, в течение которого ошибка на сервере PCP при повторении запроса также будет повторяться. При успешном выполнении запроса для кодов, создающих отображение (MAP и PEER), поле Lifetime указывает срок жизни отображения. Устанавливается сервером.

Epoch Time

Значение параметра Epoch Time на сервере (см. параграф 8.5). Устанавливается сервером.

Reserved - резерв

96 резервных битов. Для запросов, которые были успешно разобраны все биты **должны** иметь значение 0 (устанавливается сервером) и игнорироваться при получении. Если разобрать запрос не удалось, сервер копирует последние 96 битов поля PCP Client's IP Address из соответствующего запроса.

Opcode-specific information – специфические для операции данные

Дополнительные данные для Opcode. Размер данных определяется значением Opcode.

PCP Options - опции PCP

Необязательное поле с одной или множеством опций PCP, применимых для запроса и кода (см. параграф 7.3).

7.3 Опции

PCP Opcode может сопровождаться одной или множеством опций, которые могут применяться как в запросах, так и в откликах. Предлагаемое в данной спецификации решение по размещению необязательной информации основано на компромиссе между упрощением программного кода и увеличением размера пакетов. Размещение информации, которая требуется часто (или всегда), в фиксированных данных Opcode позволяет упростить код генерации и разбора пакетов, поскольку нужные данные всегда находятся в одном месте Opcode, но при этом пространство пакетов расходуется впустую, если информация не передаётся или не имеет отношения к делу. Размещение редко используемой информации в необязательных полях (опции) несколько усложняет код генерации и разбора пакетов, но

позволяет экономить пространство в пакетах, если данная информация не нужна. Размещение информации в форме опций обеспечивает не использующим такие данные реализациям просто не включать код генерации и разбора таких данных. Например, клиенту, который никогда не запрашивает отображения от имени других устройств, не требуется код для генерации опции THIRD_PARTY, а сервер PCP, который не поддерживает средств защиты для таких отображений может не включать код разбора опции THIRD_PARTY.

Опции используют формат TLV[1], показанный ниже:

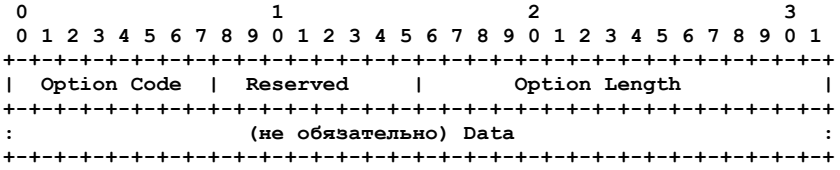

Рисунок 4. Заголовок опций.

Option Code – код опции

8 битов. Старший бит кода показывает является обработка этой опции обязательной (0) или не обязательной (1).

Reserved - резерв

8 битов. Поле **должно** устанавливаться в 0 при передаче, а на приёмной стороне **должно** игнорироваться.

Option Length - размер опции

16 битов. Указывает размер поля данных опции в октетах. Данные опции могут иметь нулевой размер (0). Если размер опции не кратен 4, в неё включается от 1 до 3 октетов заполнения со значениям 0 для выравнивания опции по 4-октетной границе. В поле Option Length указывается истинный размер опции без учёта заполнения.

Data - данные

Данные опции.

При включении в запрос PCP нескольких опций, клиент PCP **может** размещать их в произвольном порядке, но сервер PCP **должен** обрабатывать опции в порядке их размещения. Клиент PCP должен принимать во внимание, что размер отклика на его запрос не может превышать 1100 октетов, в противном случае сервер будет генерировать сообщение об ошибке.

Если в процессе обработки запроса PCP, включая его опции, возникает ошибка, это приводит к генерации сообщения PCP об ошибке, а состояние сервера PCP или управляемого PCP устройства **должно** остаться неизменным (т. е., все созданные при обработке изменения должны быть отменены). Сообщение об ошибке **должно** включать полную копию вызвавшего ошибку запроса с кодом ошибки и другими имеющими отношение к делу полями в заголовке.

В запросах и откликах опция **может** присутствовать в нескольких экземплярах, если определение опции позволяет это. Если же определение не допускает множественных экземпляров опции, но эта опция появляется в запросе неоднократно и понятна серверу PCP, этот сервер **должен** возвратить отклик с кодом результата MALFORMED_OPTION. Если сервер PCP получает некорректную опцию (например, значение размера опции PCP превышает размер пакета UDP), **следует** возвращать код MALFORMED_OPTION (а не MALFORMED_REQUEST), поскольку это даст клиенту более точную информацию об ошибке. Если размер отклика PCP будет превышать максимальный размер сообщения, серверу PCP **следует** возвращать код MALFORMED_REQUEST.

Если структуру опций в целом разобрать не удалось (например, не имеющее смысла поле размера), сервер PCP **должен** генерировать сообщение об ошибке с кодом MALFORMED_OPTION.

Если структура опций корректна, обработка каждой опции определяется старшим битом кода опции. Если этот бит установлен, обработка данной опции не обязательна и сервер PCP **может** обработать или игнорировать данную опцию по своему усмотрению. Если старший бит кода опции не установлен (0), обработка опции обязательна и сервер PCP **должен** возвращать код MALFORMED_OPTION для опций с некорректным форматом, UNSUPP_OPTION для не распознанных, не реализованных или отключённых опций, а также в тех случаях, когда клиент не уполномочен использовать данную опцию. В сообщениях об ошибках возвращаются все опции. При успешной обработке возвращаются только обработанные опции.

Поскольку клиент PCP не может отвергнуть отклик, содержащий опции, он **должен** игнорировать непонятные опции в откликах (включая опции с флагом обязательной обработки). Если клиент явно запрашивает опцию и корректное выполнение данной опции требует обработки данных опции в отклике на запрос, такому клиенту, естественно, **следует** поддерживать код для такой обработки.

Для разных кодов (Opcode) корректны различные опции. Например,

- опция THIRD_PARTY применима для кодов MAP и PEER;
- опция FILTER корректна только для MAP Opcode (для PEER Opcode она просто не имеет смысла);
- опция PREFER_FAILURE применима только для MAP (для PEER аналогичная семантика применяется автоматически).

7.4 Коды результата

Перечисленные ниже коды результатов могут возвращаться для любых кодов Opcode, полученных сервером PCP. Единственным кодом успешного выполнения операции является 0, а все остальные значения говорят о тех или иных ошибках. Если сервер PCP сталкивается с множеством ошибок в процессе выполнения запроса, ему **следует** возвращать наиболее специфичный код ошибки. Каждый из приведённых ниже кодов связан с продолжительной или краткосрочной ошибкой и эту информацию разработчики серверов PCP могут использовать в качестве значения поля Lifetime в сообщениях об ошибках. **Рекомендуется** использовать для краткосрочных ошибок срок в 30 секунд, а для долгосрочных - 30 минут.

0 SUCCESS - успешное выполнение

Операция выполнена без ошибок.

[1]Type-Length-Value - тип-размер-значение.

1 UNSUPP_VERSION - неподдерживаемая версия

Номер версии, указанный в начале заголовка PCP Request, не распознан данным сервером PCP. Эта ошибка относится к долгосрочным. Данный документ описывает PCP версии 2.

2 NOT_AUTHORIZED - нет полномочий

Запрошенная операция запрещена для данного клиента PCP или клиент запросил операцию, которая не может быть выполнена в соответствии с политикой безопасности сервера PCP. Эта ошибка относится к долгосрочным.

3 MALFORMED_REQUEST - некорректный по форме запрос

Запрос не удалось разобрать. Эта ошибка относится к долгосрочным.

4 UNSUPP_OPCODE - неподдерживаемая операция

Операция с указанным Opcode не поддерживается. Эта ошибка относится к долгосрочным.

5 UNSUPP_OPTION - неподдерживаемая опция

Опция не поддерживается. Сообщение о такой ошибке генерируется только для обязательных к обработке опций. Эта ошибка относится к долгосрочным.

6 MALFORMED_OPTION - некорректный формат опции

Ошибка в формате опции (например, избыточные повторы, некорректный размер). Ошибка является долгосрочной.

7 NETWORK_FAILURE - отказ в сети

Сервер PCP или управляемое им устройство столкнулись с тем или иным отказом в сети (например, внешний адрес IP ещё не был получен). Эта ошибка относится к краткосрочным.

8 NO_RESOURCES - недостаточно ресурсов

Запрос имеет верный формат и корректен, но у сервера не достаточно ресурсов для его выполнения в настоящее время. Например, устройство NAT не может в данный момент создать новое отображение по причине недостаточной производительности процессора или нехватки оперативной памяти или иных обстоятельств, которые с течением времени перестанут действовать. В будущем такой же запрос может быть выполнен без ошибок. Эта ошибка относится к системным в отличие от USER_EX_QUOTA. Этот код может использоваться в тех случаях, когда остальные коды не описывают ошибку. Эта ошибка относится к краткосрочным.

9 UNSUPP_PROTOCOL - неподдерживаемый протокол

Неподдерживаемый транспортный протокол (например, SCTP для устройства NAT, которое обслуживает лишь UDP и TCP). Эта ошибка относится к долгосрочным.

10 USER_EX_QUOTA - превышение квоты пользователем

Попытка создания нового отображения, которое превышает квоту для данного абонентского порта. Эта ошибка относится к краткосрочным.

11 CANNOT_PROVIDE_EXTERNAL - невозможно предоставить внешний порт или адрес

Предложенный внешний порт и/или адрес не может быть предоставлен. Код **должен** возвращаться только для:

- запросов MAP с опцией PREFER_FAILURE (обычные запросы MAP возвращают доступный внешний порт);
- запросов MAP для протокола SCTP (подразумевается PREFER_FAILURE);
- запросов PEER.

Описание опции PREFER_FAILURE дано в параграфе 13.2. Длительность ошибки зависит от причины отказа.

12 ADDRESS_MISMATCH - несоответствие адресов

IP-адрес отправителя в пакете с запросом не соответствует содержимому поля PCP Client's IP Address по причине наличия не ожидаемого устройства NAT на пути между клиентом PCP и управляемым PCP устройством NAT или МСЭ. Эта ошибка относится к долгосрочным.

13 EXCESSIVE_REMOTE_PEERS

Сервер PCP не способен создать фильтры для данного запроса. Этот код **должен** возвращаться только для запросов MAP с опцией FILTER, описанной в параграфе 13.3. Эта ошибка относится к долгосрочным.

8 Функционирование PCP

Сообщения PCP **должны** передаваться по протоколу UDP [RFC0768]. На каждый запрос PCP создаётся по крайней мере один отклик, поэтому протоколу PCP не требуется транспорт с гарантированной доставкой.

При получении множества идентичных запросов сервер PCP обычно будет генерировать идентичные отклики за исключением тех случаев, когда состояние сервера PCP меняется между запросами в результате иных действий. Примером такой смены состояния может служить получение запроса в то время, когда управляемое PCP устройство не имеет доступных отображений и сервер PCP генерирует отклик с сообщением об ошибке. Если отображение становится доступным и поступает другая копия того же запроса (возможно в результате дублирования в сети), сервер PCP будет создавать отображение и возвращать отклик без ошибки. Клиент PCP **должен** обрабатывать такие обновлённые отклики для всех своих запросов (прежде всего с целью поддержки быстрого восстановления, описанного в параграфе 14). Дополнительная информация о работе протокола приведена в разделе 6 Устройство протокола.

8.1 Клиент PCP - генерация запроса

В этом параграфе рассматриваются действия клиента PCP для всех значений Opcode. Процедуры для операции MAP описаны в разделе 11, а для операции PEER - в разделе 12.

Прежде, чем передать своё первое сообщение PCP, клиент PCP определяет сервер, которым он будет пользоваться. Для этого клиент выполняет следующие действия:

1. Если сервер PCP задан в конфигурации (например, указан в файле или получен от DHCP), эта информация служит единственным списком серверов PCP.
2. В противном случае в качестве списка серверов PCP используется список принятых по умолчанию маршрутизаторов (для IPv4 и IPv6). Таким образом при наличии у клиента PCP одновременно адресов IPv4 и IPv6 он будет иметь сервер IPv4 PCP (маршрутизатор IPv4, установленный по умолчанию) для отображений IPv4 и сервер IPv6 PCP (маршрутизатор IPv6, установленный по умолчанию) для отображений IPv6.

В соответствии с настоящим документом поддерживается адрес только одного сервера PCP. Если в будущих спецификациях число адресов серверов возрастёт, эти спецификации должны будут определить для клиентов способ выбора из списка одного или нескольких серверов.

С известным адресом сервера PCP клиент генерирует свой запрос PCP. Этот запрос включает общий заголовок PCP, PCP Opcode и данные, а также может включать опции. Как для любых программных клиентов UDP во всех операционных системах при наличии на хосте нескольких независимых клиентов PCP каждый из них будет использовать свой порт-источник, чтобы запросы и соответствующие им отклики не путались. Номер порта-источника для клиентов PCP **следует** генерировать случайным образом [RFC6056].

Клиент PCP **должен** включать IP-адрес отправителя сообщения PCP в свой запрос PCP. Обычно это будет IP-адрес клиента; представление адреса описано в параграфе 16.4. Наличие этого адреса позволяет определить наличие неожиданных устройств NAT на пути между клиентом PCP и управляемым PCP устройством NAT или МСЭ для того, чтобы предотвратить затрату ресурсов управляемого PCP устройства NAT на создание неработающих отображений. При обнаружении такого внутреннего устройства NAT, не поддерживающего PCP, сначала требуется создать с использованием неких иных средств отображение на внутреннем устройстве NAT, а потом с учётом этого отображения создавать соответствующее отображение на внешнем устройстве NAT, поддерживающем PCP. Получив с помощью того или иного метода отображение на внутреннем NAT, клиенту PCP следует использовать внешний адрес этого внутреннего устройства NAT в качестве клиентского адреса IP, чтобы сообщить внешнему устройству NAT под управлением PCP о том, что клиенту известно наличие промежуточного устройства NAT и приняты меры по созданию соответствующего отображения на этом устройстве, а полученное от внешнего устройства NAT отображение не станет бессмысленным.

8.1.1 Повтор запроса клиентом

Клиент PCP сам отвечает за доставку запросов PCP. Если клиент PCP не получает ожидаемого отклика от сервера, он должен повторить передачу своего сообщения. При повторе **должно** использоваться то же значение Mapping Nonce (см. параграфы 11.1 и 12.1). Клиент начинает обмен сообщениями с передачи сообщения серверу. Обмен продолжается до тех пор, пока у клиента сохраняется потребность в отображении и прерывается, когда транзакция PCP становится ненужной клиенту (например, запросившей отображение программе это отображение больше не требуется) или (опционально) принимается решение о возникновении отказа при обмене сообщениями в соответствии с описанными ниже правилами.

Поведение клиента при повторе передач управляется и описывается следующими переменными:

RT - тайм-аут повтора передачи, рассчитываемый в соответствии с приведённым ниже описанием;

IRT - изначальный интервал повтора (**следует** устанавливать 3 секунды);

MRC - максимальное число повторов (**следует** устанавливать значение 0 - неограниченное число повторов);

MRT - максимальный интервал повтора (**следует** устанавливать 1024 секунды);

MRD - максимальная продолжительность повторов (**следует** устанавливать 0 - неограниченное время);

RAND - случайный коэффициент, рассчитываемый в соответствии с приведённым ниже описанием.

При каждой передаче или повторе сообщения клиент устанавливает значение RT в соответствии с приведёнными ниже правилами. Если RT истекает до получения отклика, клиент повторяет передачу и рассчитывает новое значение RT.

Каждый расчёт нового значения RT включает новый случайный коэффициент (RAND), который представляет собой случайное значение выбранное из однородного распределения в диапазоне от -0,1 до +0,1. Случайный коэффициент служит для рассинхронизации сообщений, передаваемых клиентами PCP. К алгоритму выбора случайного значения не предъявляется криптографических требований. Алгоритму **следует** создавать разные последовательности случайных чисел для каждого вызова клиента PCP.

Начальное значение RT задаётся на основе IRT:

```
RT = (1 + RAND) * IRT
```

Для каждой последующей передачи сообщения значение RT базируется на предшествующем значении тайм-аута с учётом верхней границы RT, заданной значением MRT. Если MRT = 0, это означает отсутствие верхней границы для RT (MRT трактуется как «бесконечность»). Новое значение RT вычисляется по формуле (RTprev - текущее значение RT):

```
RT = (1 + RAND) * MIN (2 * RTprev, MRT)
```

MRC задаёт верхнюю границу числа повторов передачи сообщения клиентом. Если значение MRC отлично от нуля, после передачи клиентом сообщения MRC раз принимается решение об отказе.

MRD задаёт верхнюю границу временного интервала, в течение которого клиент может повторять передачу сообщения. Если значение MRD отлично от нуля, по прошествии MRD секунд с момента передачи первого сообщения клиентом принимается решение об отказе.

Если оба значения MRC и MRD отличны от нуля, отказ при обмене сообщениями фиксируется при выполнении любого из указанных выше условий. Если оба параметра MRC и MRD равны 0, клиент продолжает повторять передачу сообщения, пока не будет получен отклик или отпадёт потребность в отображении.

После того, как клиент PCP получит отклик об успешном завершении операции от сервера PCP на данном интерфейсе, он устанавливает для RT случайное значение из диапазона 1/2 - 5/8 от времени жизни отображения, как описано в параграфе 11.2.1 Обновление отображения и передаёт последующие запросы PCP для отображений тому же серверу.

Примечание. Если состояние сервера меняется в интервале между повторами сообщений, а отклик сервера задерживается или теряется, состояния сервера и клиента PCP могут оказаться рассинхронизированными. Это не является особенностью PCP и случается также в других протоколах (например, TCP). Если такая маловероятная рассинхронизация происходит, PCP «излечивается» самостоятельно по истечении времени жизни отображения.

8.2 Сервер PCP - обработка запроса

В этом параграфе более подробно рассматриваются операции на сервере PCP. Обработку **следует** выполнять в описанном ниже порядке.

Сервер PCP **должен** принимать только нормальные (не THIRD_PARTY) запросы PCP от клиента на том же интерфейсе, через который он обычно получает пакеты от данного клиента, и **должен** без уведомлений игнорировать запросы PCP на любом другом интерфейсе. Например, домашний шлюз NAT принимает запросы PCP только на внутреннем (LAN) интерфейсе и игнорирует без уведомлений все запросы PCP на внешнем (WAN) интерфейсе. Сервер PCP, поддерживающий запросы THIRD_PARTY, **может** быть настроен на восприятие запросов THIRD_PARTY на других интерфейсах (см. описание опции THIRD_PARTY в параграфе 13.1).

При получении запроса сервер PCP разбирает запрос и проверяет его пригодность. Пригодный запрос содержит корректный общий заголовок PCP, один пригодный код операции (PCP Opcode), а также может включать опции (которые сервер может и не поддерживать). Если при обработке запроса возникает ошибка, сервер генерирует отклик с информацией об ошибке и передаёт его клиенту PCP. Обработка кода операции и опций зависит от значения Opcode.

Сообщения об ошибках имеют такую же структуру пакета, как и отклики о нормальном завершении, включая в себя копии некоторых полей из запроса, а также поля, заданные сервером PCP, как показано на рисунке 3.

Копирование полей из запроса в отклик имеет важное значение, поскольку копии позволяют клиенту идентифицировать запрос, к которому относится данный отклик. Для значений Opcode, которые понятны серверу PCP, при копировании полей сервер будет следовать требованиям для соответствующего Opcode. Если код серверу не понятен, он просто генерирует отклик UNSUPP_OPCODE, копирует поля заголовка и остальные данные PCP без изменения (и попытки их интерпретации).

Все отклики (как при успехе, так и в случае ошибки) содержат Opcode из запроса, но с установленным битом R.

Все сообщения об ошибках имеют отличный от 0 код результата и создаются путём:

- копирования всех данных UDP или 1100 октетов (если данных больше) и дополнения нулями для выравнивания по 4-октетной границе (при необходимости);
- установки бита R;
- установки кода результата;
- установки значений полей Lifetime, Epoch Time и Reserved;
- обновления других полей отклика, для которых в описании поля указано «устанавливается сервером».

Отклик об успешном завершении имеет нулевой код результата и создаётся путём:

- копирования 4-х первых октетов заголовка пакета с запросом;
- установки бита R;
- установки кода результата в 0;
- установки значений полей Lifetime, Epoch Time и Reserved;
- установки связанных с опциями данных в тех случаях, когда это требуется;
- добавления опций обработки отклика.

Если размер полученного запроса PCP меньше двух октетов, сообщение отбрасывается без уведомления.

Если в запросе установлен бит R, сообщение отбрасывается без уведомления.

Если в первом октете (версия) указан не поддерживаемый номер версии, генерируется отклик с кодом результата UNSUPP_VERSION, после чего выполняются действия по согласованию версии, описанные в разделе 9.

Если версия поддерживается, но полученное сообщение короче 24 октетов, оно отбрасывается без уведомления.

Если сервер перегружен запросами (от данного клиента или от всех клиентов), он **может** просто отбрасывать запросы без уведомлений, поскольку клиенты PCP будут повторять запросы, или генерировать отклики NO_RESOURCES.

Если размер сообщения превышает 1100 октетов, не кратен 4 октетам или слишком мал для указанного Opcode, он считается недопустимым и сервер генерирует отклик MALFORMED_REQUEST, уменьшая размер сообщения до 1100 октетов.

Сервер PCP сравнивает адрес IP (из IP-заголовка принятого пакета) с полем PCP Client IP Address. Если адреса не совпадают, **должен** генерироваться отклик ADDRESS_MISMATCH. Это делается для обнаружения и предотвращения случайного использования PCP в тех случаях, когда на пути между клиентом и сервером имеется не поддерживающее PCP устройство NAT. Если клиент PCP хочет получить отображение в такой ситуации, он должен обеспечить соответствие поля PCP адресу IP, который будет видеть сервер PCP.

8.3 Клиент PCP - обработка отклика

Клиент PCP получает отклик и проверяет соответствие IP-адреса отправителя и номера порта серверу PCP, которому был отправлен запрос PCP. При обнаружении несоответствия отклик отбрасывается без уведомления.

При получении отклика PCP размером менее 4 октетов такой отклик отбрасывается без уведомления.

Если бит R в отклике не установлен, такой отклик отбрасывается без уведомления.

Если в отклике указан код результата UNSUPP_VERSION, выполняется согласование версий, описанное в разделе 9.

Отклики, размер которых меньше 24 октетов, больше 1100 или не кратен 4, являются некорректными и игнорируются.

Далее клиент PCP проверяет соответствие значения Opcode отправленному ранее запросу PCP. Если отклик не соответствует отправленному запросу PCP, он игнорируется. После этого проводится сравнение специфических для

Opcode полей данных в запросе и отклике в соответствии с описанием обработки для данного Opcode. При наличии расхождений отклик игнорируется.

После успешного завершения проверок клиент PCP просматривает поле Epoch Time (см. параграф 8.5) для определения потребности в обновлении своего состояния на сервере PCP. Клиенту PCP **следует** быть готовым к получению множества откликов от сервера PCP в любой момент после отправки одного запроса. Это позволяет серверу PCP информировать клиента о таких изменениях отображений, как обновление или удаление. Например, сервер PCP может передать отклик SUCCESS, а потом в результате изменения своей конфигурации, передать отклик NOT_AUTHORIZED. Клиент PCP **должен** быть готов к получению откликов на запросы, которых он не передавал (они могли быть переданы предшествующим экземпляром PCP на этом же хосте, другим хостом, который использовал такой же адрес IP, или атакующим систему злоумышленником) - такие сообщения просто игнорируются.

Получение кода результата ADDRESS_MISMATCH говорит о присутствии устройства NAT между клиентом и сервером PCP. Процедуры обработки таких ситуаций выходят за рамки данного документа.

Во всех откликах возвращается поле времени жизни (Lifetime). Значение этого поля показывает продолжительность интервала, в течение которого клиенту следует считать данных отклик действующим (при позитивном отклике это будет время действия созданного отображения, при отказах - время сохранения условий, вызвавших ошибку). Клиенту PCP **следует** ограничивать сверху значение времени жизни (для предотвращения абсурдно больших значений) в соответствии с параграфом 15 Срок действия и удаление отображений.

Если код результата имеет значение 0 (SUCCESS), запрос завершился успешно.

Отличный от 0 код результата говорит о возникновении отказа и клиенту PCP **не следует** повторять передачу того же запроса в течение указанного полем Lifetime срока (с учётом ограничений, описанных в разделе 15).

Если клиент PCP обнаруживает новый сервер PCP (например, подключенный к новой сети), этот клиент **может** незамедлительно начать взаимодействие с этим сервером, независимо от параметров ожидания, полученных от предыдущего сервера PCP.

8.4 Множественные интерфейсы

Хосты, которым требуется отображение PCP, могут иметь множество интерфейсов (логических и/или физических). На практике хост может использовать несколько адресов IPv4 (например, Wi-Fi и Ethernet) или адреса IPv4 и IPv6. Эти адреса могут иметь разные области видимости (например, IPv6 с глобальной в случае GUA[1] [RFC3587] или ограниченной в случае ULA[2] [RFC4193] доступностью).

Адреса IPv6 с глобальной доступностью (например, GUA) **следует** использовать в качестве адреса отправителя при генерации запроса PCP. Адреса IPv6 без глобальной доступности (например, ULA) **не следует** применять в качестве исходного интерфейса при генерации запроса PCP. Если для отображений PCP применялся приватный адрес IPv6 [RFC4941], потребуется новый запрос PCP в случае изменения приватного адреса IPv6. Такой запрос PCP **следует** передавать непосредственно с приватного адреса IPv6. Клиентам **рекомендуется** удалять отображения для предыдущего приватного адреса после того, как потребность в них отпадёт.

По причине повсеместного использования IPv4 NAT, адреса IPv4 с ограниченной доступностью (например, приватные адреса [RFC1918]) **могут** использоваться в качестве адреса источника при генерации запросов PCP.

8.5 Эпоха

Каждый отклик PCP, передаваемый сервером PCP, включает поле Epoch Time. Значение этого параметра на сервере инкрементируется каждую секунду. Аномалии в полученном клиентом значении Epoch Time подсказывают клиенту, что состояние на сервере PCP могло быть потеряно. Клиенты отвечают на такую подсказку быстрым запросом обновления для данного отображения, что позволяет без промедления восстановить потерянное состояние на сервере PCP.

Если сервер PCP сбрасывает или теряет состояние для явного динамического отображения (т. е., отображения, созданного по запросу PCP) в результате перезагрузки, сбоя по питанию или по иной причине, он **должен** сбросить своё значение Epoch Time к первоначальному (обычно 0), чтобы обеспечить подсказку своим клиентам PCP. После сброса параметра Epoch Time сервер PCP продолжает инкрементировать это значение каждую секунду.

Аналогично, при смене внешнего адреса IP на устройстве NAT (контролируемом сервером PCP) значение поля Epoch Time **должно** быть сброшено. Сервер PCP **может** поддерживать одно значение Epoch для всех клиентов PCP, а **может** установить разные значения Epoch Time (например, по клиентам, интерфейсам или иным критериям); решение этого вопроса отдано разработчикам.

При получении отклика PCP клиент проверяет значение в соответствии с описанной ниже процедурой, используя целочисленную арифметику.

- Если это первый отклик, полученный клиентом PCP от данного сервера PCP, значение Epoch Time трактуется, как обязательно корректное. В остальных случаях:
  - Если значение Epoch Time для текущего сервера PCP (curr_server_time) меньше предыдущего значения Epoch Time (prev_server_time) более, чем на 1 секунду, клиент трактует значение Epoch Time, как явно некорректное (время не может уменьшаться). Разница значений Epoch Time не более 1 секунды, считается допустимой, и такое незначительное нарушение порядка доставки на пути между сервером и клиентом PCP не вызовет каскада ненужных обновлений отображений. Если значение Epoch Time проходит описанную проверку, выполняются следующие операции:
    - Клиент вычисляет разницу между локальным текущим временем (curr_client_time) и временем получения предыдущего отклика от данного сервера PCP (prev_client_time):

```
client_delta = curr_client_time - prev_client_time;
```

[1]Global Unicast Address - глобальный индивидуальный адрес.
[2]Unique Local Address - уникальный локальный адрес.

- Клиент вычисляет разницу между текущим значением Epoch Time сервера PCP (curr_server_time) и предшествующим значением Epoch Time (prev_server_time):

```
server_delta = curr_server_time - prev_server_time;
```

- Если **client_delta+2 < server_delta - server_delta/16 or server_delta+2 < client_delta - client_delta/16**, клиент считает значение Epoch Time недопустимым.

- Клиент сохраняет значение текущего времени для использования в будущих сравнениях:

```
prev_client_time = curr_client_time
prev_server_time = curr_server_time
```

Если клиент PCP счёл полученное значение Epoch Time недопустимым, он предполагает, что сервер PCP мог потерять состояние и быстро обновляет все свои активные отображения в соответствии с процедурой, описанной в параграфе 16.3.1.

Примечания.

- Время на часах клиента **должно** монотонно возрастать. Если **curr_client_time < prev_client_time**, это говорит о наличии ошибки на стороне клиента. Обработка таких ошибок зависит от реализации клиента.

- Приведённые выше расчёты позволят определять значения client_delta и server_delta, как целые числа.

- «+2» в приведённом выше расчёте служит для компенсации возможных погрешностей квантования у клиента и сервера (возможна погрешность до 1 секунды на каждой из сторон).

- Деление на 16 (/16) в приведённом выше расчёте служит для компенсации погрешностей часов в дешёвых устройствах. Это позволяет компенсировать расхождение в 1/16 (6,25%) рассматривать, как погрешность (например, если на одной стороне часы спешат на 3%, а на другой отстают на 3%), не считая его аномалией или свидетельством перезапуска. Это допущение достаточно строго для того, чтобы перезагрузки эффективно детектировались и достаточно мягко, чтобы не возникало ложных срабатываний.

9 Согласование версий

Клиент PCP отправляет свои запросы, используя номер версии PCP 2. Впоследствии данная спецификация может быть обновлена с изменением формата сообщений и номером версии больше 2. Предполагается, что в этом случае серверы PCP будут поддерживать версию 2 наряду с более новой версией. Однако в случае возврата сервером кода результата UNSUPP_VERSION, клиент **может** генерировать сообщение об ошибке, информирующее пользователя о невозможности работы с данным сервером.

Если последующие обновления спецификации будут использовать иной формат сообщений с номером версии больше 2 и будет желательна совместимость с более старыми версиями, первый октет сообщения можно использовать для обеспечения совместимости.

При появлении спецификаций с номером версии больше 2 согласование версий следует выполнять в соответствии с приведённым ниже описанием.

1. Клиент отправляет первый запрос с указанием наибольшего (т. е., лучшего) поддерживаемого номера версии.
2. Если сервер поддерживает эту версию, он возвращает обычный отклик.
3. Если сервер не поддерживает указанную клиентом версию, он генерирует отклик с кодом результата UNSUPP_VERSION и указывает в нем ближайший поддерживаемый номер версии (если сервер поддерживает версии с номером больше указанного клиентом, он возвращает наименьший из поддерживаемых номеров, если же номера поддерживаемых сервером версий меньше указанного клиентом, он возвращает наибольший из поддерживаемых номеров).
4. Если клиент получает отклик с кодом UNSUPP_VERSION и поддерживаемым им номером версии, он записывает этот факт для использования указанного номера при последующих контактах с данным сервером PCP (пока от этого сервера вновь не будет получен отклик UNSUPP_VERSION в результате его обновления). Если в отклике UNSUPP_VERSION указан номер версии 0, это означает, что сервер NAT-PMP [RFC6886] и клиент **могут** взаимодействовать по устаревшему протоколу NAT-PMP, как описано в Приложении A.
5. Если клиент получает отклик с кодом UNSUPP_VERSION и неподдерживаемым им номером версии, клиенту **следует** перейти к использованию ближайшего меньшего номера версии из числа поддерживаемых и повторить запрос. Попытки снижения номера версии могут повторяться до тех пор, пока не будет достигнут номер 2. Если попытка использования версии 2 завершилась отказом, клиент **может** сгенерировать для пользователя сообщение о невозможности работы с данным сервером. Клиенту также **следует** установить для таймера повтора меньшее из значение 30 минут или возвращённое сервером значение Lifetime. Автоматический повтор через 30 служит для ситуаций, когда согласовать версии не удаётся по причине происходящего обновления сервера PCP.

10 Операции MAP и PEER

В данном документе определены четыре случая использования операций с кодами MAP и PEER:

- хост-сервер, желающий принимать входящие соединения (параграф 10.1);
- хост, использующий один порт в качестве клиента и сервера (параграф 10.2);
- хост-клиент, желающий оптимизировать пользовательский трафик keepalive (параграф 10.3);
- хост-клиент, желающий восстановить утраченное состояние в NAT (параграф 10.4).

Эти варианты описаны ниже и показаны на (ненормативной) диаграмме состояний в параграфе 16.5.

При работе в качестве сервера (см. параграфы 10.1 и 10.2) клиент PCP знает, какие адреса он будет прослушивать из Internet - IPv4, IPv6 или оба типа. Клиент PCP также знает о наличии адресов IPv4 и IPv6 на своих интерфейсах. На основе этой информации клиент принимает решение о выборе серверов PCP для отправки запросов (например, запрос адреса IPv4 или IPv6) и числе (1 или 2) запросов MAP для каждого из своих интерфейсов (например, если клиент PCP имеет только адрес IPv4, но хочет прослушивать IPv6 и IPv4, он отправляет запрос MAP, содержащий адрес IPv6 «все нули» в поле Suggested External Address, и запрос MAP с адресом IPv4 «все нули» в поле Suggested External Address). Если клиент PCP имеет адреса IPv4 и IPv6, но хочет прослушивать лишь IPv4, он отправляет один запрос MAP со своего адреса IPv4 (если сервер PCP поддерживает NAT44 или IPv4 МСЭ) или один запрос MAP со своего адреса IPv6 (если сервер PCP поддерживает NAT64). Клиент PCP может просто запросить желаемое отображение для того, чтобы определить поддерживает ли сервер PCP это отображение. Это будут полезно для приложений, которые включают адреса IP в поля данных (например, FTP или SIP), поскольку позволяет избавиться от трансляции адресов между разными семействами.

Запросы MAP и PEER включают поле Suggested External IP Address. Некоторые устройства, контролируемые PCP, в частности, CGN, а также многодомные сети NPTv6 имеют пул публично доступных адресов IP. PCP позволяет клиенту указать, нужно ли ему отображение на конкретный или любой из адресов данного пула. У некоторых приложений (например, FTP в активном режиме) могут возникать проблемы при создании отображений на другой адрес IP, поэтому таким приложениям следует принимать во внимание последствия использования такой возможности. Для такого внутреннего адреса могут существовать статические отображения (например, созданные командой на сервере PCP или управляемом PCP устройстве) на некие внешние адреса и, если предложен внешний адрес «все нули» IPv4 или IPv6, PCP **следует** организовать отображение на тот же адрес, поскольку это позволит приложению использовать комбинацию статических и созданных PCP отображений. С другой стороны, если предложен ненулевой адрес IP, серверу PCP **следует** организовать отображение на этот внешний адрес даже если для других отображений с того же внутреннего адреса используются иные внешние адреса. Когда для внутреннего адреса нет неявного или явного динамического отображения в управляемом PCP устройстве, для последующего явного или неявного отображения данного адреса может использоваться другой внешний адрес. Как правило, переназначение будет происходить в тех случаях, когда устройство CGN с балансировкой нагрузки вновь увидит эти внутренние адреса для своего пула внешних адресов.

В приведённой ниже таблице показано использование адресов IPv6 и IPv4 в типовых реализациях PCP.

Внутренний адрес «неявно» совпадает с IP-адресом отправителя запроса PCP за исключением случаев использования опции THIRD_PARTY.

Внешний адрес указывается в поле Suggested External Address запросов MAP и PEER, а семейство, к которому он относится, обычно совпадает с семейством внутреннего адреса, за исключением случаев использования технологий типа NAT64.

Адресом удалённого партнёра PCP является IP-адрес удалённого партнёра в запросе PEER или опции FILTER запроса MAP и всегда относится к тому же семейству, что и внутренний адрес (даже при использовании NAT64). В NAT64 клиент IPv6 PCP может быть не осведомлен о NAT64 или реальном адресе удалённого партнёра IPv4, поэтому он выражает адрес IPv6 со своей точки зрения, как показано на рисунке 5.

	Внутрен. адрес	***Внешний адрес***	***Адрес удаленного партнера PCP***	***Реальный адрес удаленного партнера***
IPv4 МСЭ	IPv4	IPv4	IPv4	IPv4
IPv6 МСЭ	IPv6	IPv6	IPv6	IPv6
NAT44	IPv4	IPv4	IPv4	IPv4
NAT46	IPv4	IPv6	IPv4	IPv6
NAT64	IPv6	IPv4	IPv6	IPv4
NPTv6	IPv6	IPv6	IPv6	IPv6

Рисунок 5. Семейства адресов в MAP и PEER.

Отметим, что внутренний адрес и адрес удалённого партнёра всегда относятся к одному семейству, равно как внешний адрес и реальный адрес удалённого партнёра.

10.1 Для сервера

Хост-сервер (например, сервер web) прослушивает трафик на определённом порту, но никогда не инициирует трафик через этот порт. Для работы через NAT или МСЭ хосту нужно (a) организовать динамическое отображение с публичного адреса IP, протокола и порта на самого себя, используя операцию MAP, как описано в параграфе 11, (b) опубликовать адрес IP, протокол и номер порта через тот или иной рандеву-сервер (например, DNS, сообщение SIP или фирменный протокол) и (c) обеспечить отсутствие препятствий в передаче трафика со стороны промежуточных устройств, не поддерживающих PCP (например, МСЭ или NAT). Публикация адреса IP и номера порта выходит за рамки данной спецификации. Для выполнения п. (a) хост должен следовать инструкциям данного параграфа.

Обычно приложению нужно начать с прослушивания порта. После этого приложение создаёт сообщение PCP с кодом операции MAP, указывая в качестве внешнего подходящий адрес «все нули» (IPv4 или IPv6).

Приведённый ниже псевдокод показывает применение PCP для организации доступа к серверу.

10.2 Для сервера-клиента

Хост, использующий один порт для работы в качестве клиента и сервера (например, Symmetric RTP [RFC4961] или SIP rport[1] [RFC3581]), организует локальное прослушивание порта, (обычно) передаёт локальный и публичный адрес IP, протокол и порт службе «рандеву» (выходит за рамки данного документа) и инициирует исходящее соединение с указанного адреса и порта. Для решения этой задачи хост должен следовать описанным ниже процедурам.

Приложение, которое использует один и тот же порт для входящих и исходящих соединений, **должно** сначала обозначить свою работу в качестве сервера, используя операцию MAP, как описано в разделе 11, и получить позитивный отклик PCP до начала передачи каких-либо пакетов через этот порт.

[1]Symmetric Response Routing - симметричная маршрутизация откликов.

```
/* запуск прослушивания локального порта на сервере */
int s = socket(...);
bind(s, ...);
listen(s, ...);
getsockname(s, &internal_sockaddr, ...);
bzero(&external_sockaddr, sizeof(external_sockaddr));
while (1)
    {
    /* Примечание. Проверка time_to_send_pcp_request() включает:
     * 1. Передачу первого запроса.
     * 2. Повтор запроса в результате потери пакета.
     * 3. Повтор запроса в связи с завершением срока аренды.
     * 4. Повтор запроса по причине потери состояния на сервере.
     * Во всех 4 случаях передаются идентичные пакеты PCP и с точки
     * зрения сервера PCP это будет одной операцией.
     * Предлагаемые внешний адрес и порт могут изменяться в течение
     * срока существования отображения. Остальные поля сохраняются
     * неизменными.
     */
    if (time_to_send_pcp_request())
        pcp_send_map_request(internal_sockaddr.sin_port,
            internal_sockaddr.sin_addr,
            &external_sockaddr, /* 0 для первого раза */
            requested_lifetime, &assigned_lifetime);
    if (pcp_response_received())
        update_rendezvous_server("Client Ident", external_sockaddr);
    if (received_incoming_connection_or_packet())
        process_it(s);
    if (other_work_to_do())
        do_it();
    /* ... */
    block_until_we_need_to_do_something_else();
    }
```

Рисунок 6. Псевдокод использования PCP для работы сервера.

Обсуждение. В общем случае клиент PCP заранее не знает, что он находится за устройством NAT или МСЭ. При обнаружении подключения к сети клиент PCP может попытаться запросить отображение с помощью PCP. В случае успеха клиент будет знать отображение адресов. Если после нескольких попыток отклик не будет получен, возможны два варианта - клиент не находится за устройством NAT/МСЭ или это устройство не поддерживает PCP (подключение клиента может сохраняться на основе созданного пользователем вручную отображения). Повтор запросов PCP перед принятием решения в таких случаях увеличивает задержку. Инициирование исходящего соединения TCP без ожидания PCP позволяет избавиться от этой задержки и будет работать, если поведение устройства NAT не зависит от конечной точки (EIM[1]), но может приводить к отказу, если отображение NAT зависит от конечной точки (EDM[2]). Достаточно долгое ожидание, позволяющее создать явное отображение PCP MAP (если оно возможно) обеспечивает использование того же внешнего порта для всех последующих неявных динамических отображений (например, TCP SYN) для указанного внутреннего адреса, протокола и порта. PCP поддерживает устройства NAT как EIM, так и EDM, поэтому клиенты могут предполагать работу с EDM NAT. В этом случае организация соединения для клиента будет более гарантированной, если он попытается использовать запросы PCP MAP до попытки организации исходящих соединений TCP для данного адреса и порта. Дополнительная информация об использовании PCP с устройствами EDM NAT приведена в параграфе 16.1.

Приведённый ниже псевдокод показывает использование PCP для работы в режиме сервера и клиента.

10.3 Снижение вспомогательного трафика

Хост-клиент (например, XMPP, SIP) передаёт передаёт пакеты в порт и может получать на них отклики, но никогда не воспринимает на этом порту входящих соединений от других удалённых партнёров. Такому хосту требуется предотвращать прерывание потока данных между ним и удаленным партнёром (в результате отсутствия активности) через устройство NAT или МСЭ. Для решения этой задачи приложения используют описанную здесь процедуру.

Промежуточным устройствам (NAT или МСЭ) в общем случае нужно видеть тот или иной трафик, иначе они будут разрывать сессии (удалять состояние) и это приведёт к отказу приложения. Для предотвращения этого многие приложения генерируют вспомогательный трафик keepalive (сохранение жизненности) с основной (или единственной) целью - сохранить состояние на промежуточном устройстве. Такие приложения могут снизить объем трафика keepalive, используя PCP.

Примечание. По причинам, не относящимся к NAT, пакеты keepalive могут быть нужны приложению для детектирования обрывов на пути между клиентом и сервером, сохранения состояния удалённого партнёра или детектирования отключения клиента. Такие пакеты не связаны с поддержкой состояния на промежуточном устройстве и PCP не поможет снизить объем трафика keepalive в этом случае.

Для использования PCP с целью снижения уровня вспомогательного трафика приложение сначала обычным способом соединяется с сервером. После этого приложение генерирует запрос PCP с кодом операции PEER (см. раздел 12) для определения срока жизни созданного отображения.

Приведённый ниже псевдокод показывает, как можно использовать PCP для динамического сокета с целью снижения уровня вспомогательного трафика.

[1]Endpoint-independent mapping - независимое от конечной точки отображение.
[2]Endpoint-dependent mapping - зависящее от конечной точки отображение.

```
/* запуск прослушивания локального порта на сервере */
int s = socket(...);
bind(s, ...);
listen(s, ...);
getsockname(s, &internal_sockaddr, ...);
bzero(&external_sockaddr, sizeof(external_sockaddr));
while (1)
    {                                                              т:
    /* Примечание. Проверка time_to_send_pcp_request() включает:
     * 1. Передачу первого запроса.
     * 2. Повтор запроса в результате потери пакета.
     * 3. Повтор запроса в связи с завершением срока аренды.       е.
     * 4. Повтор запроса по причине потери состояния на сервере.
     */
    if (time_to_send_pcp_request())
        pcp_send_map_request(internal_sockaddr.sin_port,
            internal_sockaddr.sin_addr,
            &external_sockaddr, /* 0 для первого раза */           е);
            requested_lifetime, &assigned_lifetime);
    if (pcp_response_received())
        update_rendezvous_server("Client Ident", external_sockaddr);
    if (received_incoming_connection_or_packet())
        process_it(s);
    if (need_to_make_outgoing_connection())
        make_outgoing_connection(s, ...);
    if (data_to_send())
        send_it(s);
    if (other_work_to_do())
        do_it();
    /* ... */
    block_until_we_need_to_do_something_else();
    }
```

Рисунок 7. Псевдокод использования PCP для работы в режиме клиента и сервера.

10.4 Восстановление утраченного состояния для неявного отображения TCP

После потери состояния NAT (например, в результате ошибки или отказа питания) клиентам полезно восстанавливать отображения TCP в устройстве NAT. Это позволяет серверам Internet видеть трафик клиента с тем же адресом IP и портом, что позволяет восстановить сессию в точке обрыва. Это может оказаться полезным для долгоживущих соединений (например, чат-системы) или соединений с большим объёмом данных (например, FTP). Этого можно достигнуть путём обычной организации соединения TCP с последующей передачей запросов/откликов PEER и запоминанием внешнего адреса и порта. Позднее при потере состояния NAT клиент может передать запрос PEER с сохранёнными для предыдущей сессии адресом и номером порта, что позволит создать в NAT отображение, соответствующее неявному динамическому отображению. После этого клиент может продолжить сессию TCP с сервером.

Примечание. Эта процедура хорошо работает для TCP, если:

(i) NAT создаёт новое неявное динамическое исходящее отображение только для исходящих сегментов TCP с установленным битом SYN (т. е., перезагруженное устройство NAT без уведомления отбросит исходящие сегменты, если в NAT нет активного отображения для них;

(ii) перезагруженное устройство NAT не передаёт сегментов TCP RST в ответ на получение неожиданных входящих сегментов TCP.

Для UDP эта процедура работает не так хорошо, новый исходящий трафик UDP создаёт на устройстве NAT новое неявное исходящее отображение (возможно, через другой порт).

11 Операция MAP

В этом разделе описана операция, контролирующая входящую пересылку от устройства NAT (или) на внутренний хост.

MAP

Создаёт явное динамическое отображение между внутренними и внешними парами (адрес + порт).

Серверам PCP **следует** поддерживать конфигурационную опцию, позволяющую администраторам отключить поддержку MAP.

Отображения, создаваемые запросами PCP MAP, по определению не зависят от конечных точек (EIM), включая независимость фильтрации от конечной точки (EIF[1]) (пока не используется опция FILTER), даже на устройствах NAT, которые обычно создают зависимые от конечной точки отображения (EDM) или фильтры (EDF[2]) для исходящих соединений, поскольку целью отображения MAP (без фильтров) является получение входящего трафика от любой удалённой точки, а не только от указанной конкретно.

Отметим также, что все отображения NAT (созданные PCP или иным путём) являются обязательно двухсторонними и симметричными. Для любого пакета в одном направлении (внутрь или наружу) и проходящего через NAT, идущий в противоположном направлении отклик должен пройти соответствующую трансляцию, чтобы попасть на нужную конечную точку. Это означает, что если клиент создает отображение MAP и потом отправляет исходящие пакеты, используя отображённые внутренний адрес, протокол и порт, NAT следует транслировать внутренний адрес и порт таких пакетов на используемые в отображении внешний адрес и порт, чтобы отклики, направленные на внешний адрес и порт корректно транслировались обратно на внутренние адрес и порт.

В операционных системах, которые позволяют связать множество прослушивающих серверов с одним внутренним адресом, протоколом и портом, **должны** обеспечить себе исключительное использование данных внутреннего адреса,

[1]Endpoint-independent filtering.
[2]Endpoint-dependent filtering.

протокола и порта (например, путём привязки порта с помощью INADDR_ANY, SO_EXCLUSIVEADDRUSE или иным способом) до отправки своего запроса PCP MAP, чтобы гарантировать, что на этой машине нет других клиентов PCP, прослушивающих те же самые адрес, протокол и порт.

Как побочный эффект организации отображения сообщения ICMP, связанные с отображением, **должны** пересылаться (и транслироваться при необходимости) в течение срока жизни отображения. Это делается для обеспечения хостам возможности использования сообщений ICMP без организации в приложениях или клиентских программах ICMP отдельных отображений для таких потоков.

Обработка операций с кодом MAP описана ниже.

11.1 Формат пакетов MAP

Пакеты запросов и откликов MAP имеют похожую структуру. Если выделенные внешний адрес IP и порт в отклике PCP всегда совпадают с внутренним адресом IP и портом в запросе PCP, в устройстве реализована только функциональность MCЭ, в противном случае устройство является транслятором адресов (NAT) и может также реализовать функции MCЭ.

На рисунке 9 показан формат специфических для операции данных в запросе MAP.

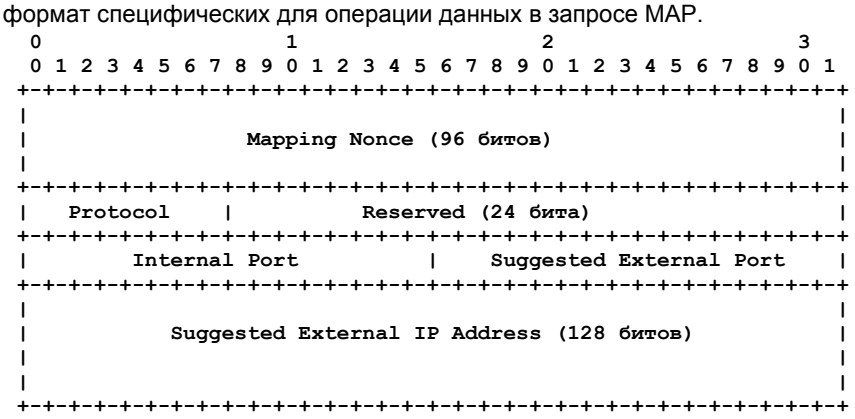

Рисунок 9. Запрос MAP.

Requested lifetime (в обычном заголовке) - запрошенное время жизни

Запрошенное время жизни отображения в секундах. Значение 0 служит для удаления отображения.

Mapping Nonce

Случайное значение, выбранное клиентом PCP (см. параграф 11.2 Генерация запроса MAP). Это поле может иметь нулевое значение (появление такого значения маловероятно - 1 раз примерно на 2^{96} запросов).

Protocol - протокол

Протокол вышележащего уровня, связанный с данным Opcode. Значения поля берутся из реестра протоколов IANA [proto_numbers]. Например, поле будет содержать значение 6 (TCP), если операция предназначена для создания отображения TCP или 17 (UDP), если создаётся отображение UDP. Значение 0 означает «все протоколы».

Reserved - резерв

24 резервных бита, которые **должны** устанавливаться в 0 при передаче и игнорироваться на приёме.

Internal Port - внутренний порт

Номер внутреннего порта для отображения. Значение 0 указывает «все порты» и допустимо для запросов с нулевым временем жизни (удалить отображение), если протокол не использует 16-битовых номеров портов или клиент запрашивает «все порты». Если для протокола указано значение 0 (все протоколы), номер внутреннего порта **должен** иметь нулевое значение при передаче и **должен** игнорироваться на приёмной стороне.

Suggested External Port предложенный внешний порт

Предлагаемый для отображения внешний порт. Это полезно при обновлении отображений, особенно в случае потери состояния на сервере PCP. Если клиент не знает внешний порт или не имеет предпочтений, он **должен** указать 0.

Suggested External IP Address - предложенный внешний адрес

Предлагаемый внешний адрес IPv4 или IPv6. Это полезно при обновлении отображений, особенно в случае потери состояния на сервере PCP. Если клиент не знает внешний адрес или не имеет предпочтений, он **должен** указать адрес «все нули» для соответствующего семейства (см. раздел 5).

Внутренним адресом для запроса является IP-адрес отправителя запроса PCP, если не используется опция THIRD_PARTY.

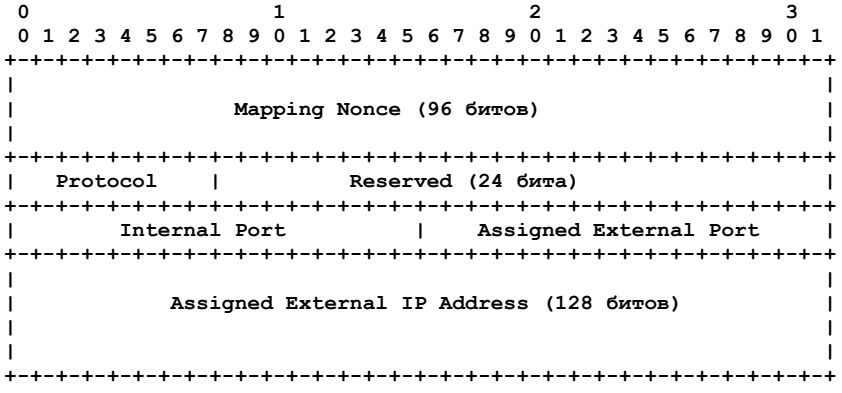

Рисунок 10. Отклик MAP.

На рисунке 10 показан формат специфических для операции данных в отклике MAP.

Lifetime (в обычном заголовке):
В откликах об ошибках это поле показывает клиенту в течение какого времени он будет получать от сервера PCP такой же отклик об ошибке при повторении запроса. В отклике об успешном отображении это поле показывает срок существования созданного отображения в секундах.

Mapping Nonce
Копируется из запроса.

Protocol - протокол
Копируется из запроса.

Reserved - резерв
24 резервных бита, которые **должны** устанавливаться в 0 при передаче и игнорироваться на приёме.

Internal Port - внутренний порт
Копируется из запроса.

Assigned External Port - выделенный внешний порт
Выделенный для отображения внешний порт. При возникновении ошибки значение этого поля копируется из запроса.

Assigned External IP Address - выделенный внешний адрес
Выделенный для отображения внешний адрес IPv4 или IPv6. Адреса IPv4 представляются с использованием отображённых на IPv4 адресов IPv6. При возникновении ошибки значение этого поля копируется из запроса.

11.2 Генерация запроса MAP

В этом параграфе описаны действия клиента PCP при передаче запросов с кодом MAP.

Запрос **может** содержать значения в полях Suggested External Port и Suggested External IP Address. Это позволяет клиенту PCP попытаться заново организовать потерянное состояние на сервере PCP, что повышает шансы сохранения существовавших соединений и позволяет клиенту PCP избавиться от необходимости изменения данных, поддерживаемых им на сервере rendezvous. Действия других элементов сети (например, запросы других пользователей или смена адресов в сети) могут воспрепятствовать предоставлению сервером PCP запрошенного внешнего адреса IP, протокола и порта. В таких случаях сервер будет выделять другой внешний адрес IP и порт.

Клиент PCP **должен** записать полученные данные, поскольку он может никогда не получить запрошенный внешний порт. В случае восстановления состояния после перезапуска шлюза NAT, предлагая полученный ранее внешний порт, клиент с большей вероятностью сможет получить его для продолжения использования. В остальных случаях клиент **должен** предполагать, что вероятность выделения ему запрошенного внешнего порта достаточно мала. Например, если множество абонентов пользуются общим шлюзом CGN, популярные порты 80, 443, 8080 будут пользоваться высоким спросом. Каждый из таких портов для каждого из внешних адресов IP может быть предоставлен только одному клиенту, а остальные получат динамически распределяемые внешние порты. На практике некоторые ISP могут (в соответствии со своей политикой) не предоставлять такие внешние порты ни одному из своих абонентов.

Если протокол не использует 16-битовых номеров портов (например, RSVP, протокол IP 46), для номера порта **должно** указываться значение 0. Это обеспечит отображение для всего трафика данного протокола.

Если клиент хочет получить отображение для всех протоколов, он использует протокол 0 и внутренний порт 0.

Значение Mapping Nonce клиент PCP выбирает случайным образом, следуя принятой политике генерации непредсказуемых случайных чисел [RFC4086], и использует это значение для проверки откликов PCP (см. ниже), а также обновлённых отображений от сервера PCP. Клиент **должен** использовать разные значения Mapping Nonce для каждого сервера PCP, с которым он взаимодействует, **рекомендуется** также выбирать новое случайное значение Mapping Nonce при каждой инициализации клиента PCP. Клиент **может** использовать разные значения Mapping Nonce для каждого отображения.

11.2.1 Обновление отображения

Срок жизни существующего отображение клиенту PCP **следует** продлять, если потребность в отображении сохраняется. Для продления срока клиент PCP отправляет новый запрос MAP, указывающий внутренний порт. В запрос PCP MAP **следует** также включать текущий присвоенный внешний адрес IP и номер порта в полях Suggested External IP Address и Suggested External Port, чтобы в случае потери состояния сервер PCP мог восстановить утраченное отображение с теми же параметрами.

Клиенту PCP **следует** обновлять отображение до завершения его срока жизни, иначе оно будет удалено сервером PCP (см. раздел 15 Срок действия и удаление отображений). Для снижения риска случайной синхронизации запросов на обновление следует включать флуктуацию (jitter) со случайным значением. Клиентам PCP **рекомендуется** передавать один пакет с запросом обновления в момент времени определяемый случайно выбранным значением из диапазона от 1/2 до 5/8 срока действия отображения. Если отклик SUCCESS на этот запрос не был получен, следующий запрос передаётся в интервале от 3/4 до 3/4 + 1/16, потом в интервале от 7/8 до 7/8 + 1/32 и т. д., с учётом того что интервал между запросами менее 4 секунд **недопустим** (клиентам PCP **недопустимо** передавать «лавину» часто повторяющихся запросов в течение последних нескольких секунд срока действия отображения).

11.3 Обработка запроса MAP

В этом параграфе описаны действия сервера PCP при обработке запроса с кодом операции MAP. Обработку **следует** выполнять в описанном ниже порядке.

Значения полей Protocol, Internal Port и Mapping Nonce из запроса MAP копируются в отклик MAP. Если опция THIRD_PARTY присутствует в запросе и обрабатывается сервером PCP, она тоже копируется в отклик MAP.

Если запрашиваемое время жизни отлично от нуля выполняются перечисленные ниже действия.

- Если значение полей протокола и внутреннего порта отличны от 0, это указывает на запрос для создания отображения или продления срока действия существующего отображения. Если сервер PCP или управляемое PCP устройство не поддерживает протокол, **должен** возвращаться код ошибки UNSUPP_PROTOCOL.

- Если значение поля протокола отлично от нуля, а внутренний порт равен 0, это указывает запрос на создание или продление срока действия отображения для всего входящего трафика указанного протокола. Если запрос не может быть выполнен полностью, **должна** возвращаться ошибка UNSUPP_PROTOCOL.

- Если поля протокола и внутреннего порта имеют значение 0, это указывает запрос на создание или продление срока действия отображения для всего входящего трафика любых протоколов (DMZ-хост). Если запрос не может быть выполнен полностью, **должна** возвращаться ошибка UNSUPP_PROTOCOL.

- Если поле протокола имеет значение0, а внутренний порт отличен от нуля, запрос является недопустимым и сервер PCP должен возвратить клиенту отклик с кодом ошибки MALFORMED_REQUEST.

Если запрошенное время жизни равно 0, это указывает запрос на удаление существующего отображения.

Дальнейшая обработка значений времени жизни описана в разделе 15 Срок действия и удаление отображений.

При работе в рамках простой модели угроз (Simple Threat Model - параграф 18.1), если внутренний порт, протокол и адрес соответствуют имеющемуся явному динамическому отображению, но Mapping Nonce не соответствует, запрос **должен** отбрасываться с кодом ошибки NOT_AUTHORIZED и указанием в поле lifetime срока жизни существующего отображения. Серверу PCP требуется помнить лишь одно значение Mapping Nonce для каждого явного динамического отображения. В данной спецификации не задаётся требований к Mapping Nonce для расширенной модели угроз.

Если внутренний порт, протокол и внутренний адрес соответствуют имеющемуся статическому отображению (которое не имеет nonce), передаётся отклик PCP с внешним адресом и портом данного статического отображения и используется значение nonce из запроса PCP. Сервер не записывает значение nonce.

Если имеется опция со значением меньше 128 (т. е, обязательная для обработки), но эта опция не имеет смысла (например, PREFER_FAILURE в запросе с lifetime=0), запрос считается недопустимым и возвращается ошибка MALFORMED_OPTION

Если управляемое PCP устройство не поддерживает информации о состоянии (т. е., не организует состояний для потоков или просто меняет адрес и/или порт по заданному алгоритму, включая отсутствие изменений), сервер PCP просто отвечает индикацией внешнего адреса IP и порта, выделенного алгоритмической трансляцией без поддержки состояния. Это позволяет клиенту PCP узнать свой внешний адрес IP и порт, которые видят удалённые партнёры. Примерами трансляторов без поддержки состояний могут служить NAT64, 1:1 NAT44, NPTv6 [RFC6296], которые меняют адреса, не меняя номеров портов, а также МСЭ, которые не меняют ни адресов, ни портов.

Возможно, что для указанного внутреннего адреса, протокола и порта уже существует отображение. В таких случаях сервер PCP выполняет перечисленные ниже действия.

1. Если запрос MAP включает опцию PREFER_FAILURE, а предложенные внешний адрес и порт не соответствуют существующим отображениям, сервер PCP **должен** возвращать CANNOT_PROVIDE_EXTERNAL.

2. Если существующее отображение является статическим (создано без участия PCP), сервер PCP **должен** возвращать в отклике адрес и порт существующего отображения, **следует** также указывать время жизни 2^{32}-1, независимо от предложенных в запросе внешнего адреса и порта.

3. Если существующее отображение является явным динамическим (создано предшествующим запросом MAP), сервер PCP **должен** возвращать в отклике внешний адрес и порт из этого отображения, независимо от предложенного в запросе внешнего адреса и порта. Кроме того, сервер PCP **должен** обновить время жизни существующего отображения в соответствии с разделом 15 Срок действия и удаление отображений.

4. Если существующее отображение является динамическим исходящим (создано исходящим трафиком или предшествующим запросом PEER), серверу PCP **следует** создать новое явное входное отображение, реплицируя порт и адрес из существующего исходящего отображения (это отображение сохраняется и продолжает существовать даже после удаления явного входного отображения).

Если для внутреннего адреса, протокола и порта нет отображения и сервер PCP может создать отображение с предложенным внешним адресом и портом, ему **следует** сделать это. Это обеспечивает преимущества при восстановлении потерянных состояний на сервере PCP (например, в результате перезагрузки). Существуют, однако, ситуации, когда сервер PCP не может создать отображение с заданным внешним адресом и портом:

- предложенные внешний адрес, протокол и порт уже выделены для явного или неявного отображения (т.е., уже используются для пересылки трафика на некоторый внутренний адрес и порт;

- предложенный внешний адрес, протокол и порт уже используются шлюзом NAT для своих целей (например, порт TCP 80 для web-интерфейса управления шлюзом NAT или порты UDP 5350 и 5351 для протокола PCP); серверам PCP **недопустимо** создавать отображения для своих внешних портов UDP 5350 и 5351;

- предложенный внешний адрес, протокол и порт запрещены для использования политикой сервера PCP;

- предложенный внешний адрес, протокол и порт являются недопустимыми или образуют недопустимую комбинацию (например, внешний адрес 127.0.0.1, ::1, групповой адрес или недопустимый для протокола порт);

- предложенный внешний адрес не принадлежит шлюзу NAT;

- предложенный внешний адрес не предназначен для использования в качестве внешнего адреса шлюза NAT или МСЭ.

Если сервер PCP не может выделить предложенный адрес, протокол и порт:

- Для запросов с опцией PREFER_FAILURE сервер PCP **должен** возвращать CANNOT_PROVIDE_EXTERNAL.

- Если запрос включает опции PREFER_FAILURE и сервер PCP может выделить некий другой внешний адрес и порт для заданного протокола, сервер **должен** выделить эти значения и возвратить их клиенту в отклике. Ни при каких условиях клиент не наказывается за «неверный» выбор предложенного внешнего адреса и порта.

Предлагаемые клиентом адрес и порт помогают серверу при выборе присваиваемых для клиента значений, но ни при каких обстоятельствах не являются препятствием для выделения отличных от предложенных внешнего адреса и порта. Присутствие в запросе отличных от нуля значений предлагаемого внешнего адреса и порта служат лишь намёком серверу и не могут наносить какого-либо вреда.

Управляемым PCP устройствам **недопустимо** создавать отображения для протоколов, не указанных в запросе. Например, если было запрошено отображение для TCP, автоматическое создание соответствующего отображения для UDP **не допустимо**.

Для отображений обычно расходуются состояния управляемого PCP устройство, поэтому на серверах PCP **рекомендуется** задавать ограничения на уровне хостов и/или абонентов для предотвращения нехватки ресурсов. При достижении заданного порога возвращается код результата USER_EX_QUOTA.

Если все операции при обработке выполнены успешно (не было откликов с ошибками), запрошенное отображение создаётся или обновляется в соответствии с запросом и генерируется отклик SUCCESS.

11.4 Обработка отклика MAP

В этом разделе описаны действия клиента PCP при обработке полученных откликов PCP с кодом операции MAP.

После выполнения обычной обработки PCP отклик сравнивается с переданным ранее запросом MAP на предмет соответствия внутреннего адреса IP (адрес получателя в отклике PCP или иной адрес, заданный через опцию THIRD_PARTY), протокола, внутреннего порта и nonce. Остальные поля не сравниваются, поскольку их значения задаёт сервер PCP. При изменении отображения (например, в результат смены адресов IP) сервер PCP будет передавать Mapping Update (параграф 14.2).

Если код результата имеет значение NO_RESOURCES, а запрос был передан для создания или обновления отображения, клиенту PCP **не следует** передавать новых запросов на отображения в течение указанного сервером PCP (ограниченного) срока (lifetime). Если код результата имеет значение NO_RESOURCES и запрашивалось удаление отображения, клиенту PCP **не следует** передавать каких-либо запросов данному серверу PCP в течение указанного сервером PCP (ограниченного) срока (lifetime).

При успешном завершении обработки запроса клиент PCP может использовать выделенный внешний адрес и порт. Обычно эти параметры применяются для обмена информацией с другим хостом при использовании специфического для приложения механизма rendezvous (например, записи DNS SRV).

После получения отклика о создании отображения (success) до того момента клиент PCP **должен** организовать таймер или иной способ уведомления о необходимости обновить отображение до завершения срока его действия. Обновление отображений выполняется путём передачи запроса MAP, как описано в параграфе 11.2, но в качестве предлагаемого внешнего адреса и порта **следует** указывать значения, полученные в отклике для действующего отображения. С точки зрения сервера PCP запрос MAP на обновление идентичен MAP-запросу на организацию отображения и обрабатывается так же. Действительно, при потере сервером состояния запрос клиента PCP на обновление отображения будет восприниматься сервером, как запрос на создание нового отображения, с предложенным внешним адресом и портом (которые сервер PCP выделил этому клиенту до потери состояния). Дополнительная информацию по части обновлений приведена в параграфе 16.3.1 Повторная организация отображений.

При получении отклика об ошибке клиенту **не следует** повторять тот же запрос тому же серверу PCP до истечения времени, указанного полем lifetime в полученном отклике.

11.5 Изменение адреса

Пользовательский маршрутизатор может получить новый внешний адрес IP в силу разных причин, включая перезагрузку, сбой питания, завершение срока аренды DHCP или действия ISP. В таких случаях трафик, направляемый по прежнему адресу хоста может быть доставлен другому хосту, получившему этот адрес. Это воздействует на все типы отображений (явные или неявные). Эта проблема возникла не сегодня и уже отмечена для устройств, получающих адреса IP без применения PCP или CGN. Решение проблемы заключается в предотвращении смены адреса хоста. Определённый в данном документе протокол PCP не предлагает механизмов для снижения остроты проблемы смены адреса хостов.

Когда внутренний хост меняет свой внутренний адрес IP (например, получает другой адрес от сервера DHCP), устройство NAT (или МСЭ) будет продолжать передачу пакетов по старому адресу. Обычно в таких случаях хост не получает такой трафик. Если хост хочет по-прежнему получать этот трафик, он должен организовать новое отображение для своего нового адреса IP. По всей вероятности прежнее значение Suggested External Port не будет принято сервером PCP, поскольку он продолжает передавать пакеты по старому адресу IP. Таким образом для нового отображения будет скорей всего выделен другой внешний порт и/или адрес IP. Отметим, что такая смена адресов не предполагается частой, поскольку большинство хостов будет продлять аренду DHCP (или запрашивать тот же адрес после перезагрузки) и большинство серверов DHCP смогут обеспечить хосту использованный ранее адрес.

На хосте могут добавляться или удаляться интерфейсы во время действия отображений (например, подключение или отключение кабеля Ethernet, подключение или отключение сети WiFi). В силу этого, если клиент PCP передаёт запрос PCP для поддержки состояния на сервере PCP, ему **следует** обеспечивает привязку запросов PCP к тому же интерфейсу (например, при обновлении отображения). Если клиент PCP передаёт запрос PCP для создания нового отображения на сервере PCP, он **может** использовать другой интерфейс-источник или иной адрес отправителя.

11.6 Самостоятельное определение внешнего адреса IP

NAT-PMP [RFC6886] включает механизм, который позволяет клиентам самостоятельно определить внешний адрес IP без запроса отображения. Протокол NAT-PMP был разработан для домашних шлюзов NAT, где такая операция имеет смысл, поскольку домашний шлюз NAT имеет лишь один внешний адрес IP. Сфера применения PCP шире и включает также устройства CGN, у которых может быть множество внешних адресов IP. Клиенту не может быть выделено ни одного внешнего адреса IP из такого пула, пока он не организует хотя бы одно (неявное, явное или статическое) отображение (и в этом случае адрес будет доступен лишь в течение срока действия отображения). Клиентское

приложение, которое просто хочет показать пользователю внешний адрес IP (в косметических целях), может просто запросить кратковременное отображение (например для службы Discard - TCP/9 или UDP/9 или другого порта) и вывести для пользователя полученный внешний адрес IP. Однако по завершении срока действия такого отображения любое последующее отображение (явное или неявное) может дать другой внешний адрес IP.

12 Операция PEER

В этом разделе описана операция, используемая для управления динамическими исходящими отображениями.

PEER

Операция создаёт динамическое исходящее отображение для адреса IP и порта удалённого партнёра или продлевает срок действия существующего исходящего отображения.

Использование операции описано в данном разделе.

Серверам PCP **следует** поддерживать конфигурационную опцию, позволяющую отключить операции PEER.

Поскольку отображения, создаваемые и поддерживаемые операцией PEER, ведут себя практически так же, как исходящие неявные динамические отображения, создаваемые по исходящим от хостов пакетам (например, TCP SYN), отображения PCP PEER могут быть независимыми (EIM) или зависимыми (EDM) от конечной точки с независимой (EIF) или зависимой (EDF) от конечной точки фильтрацией, в соответствии с поведением шлюза NAT или МСЭ по отношению к неявным исходящим отображениям, которые создаются по факту получения исходящего трафика от внутренних хостов.

12.1 Форматы пакетов PEER

Операция PEER позволяет клиенту PCP создать новое явное динамическое исходящее отображение (которое работает подобно неявно создаваемым отображениям в результате отправки хостом пакетов TCP SYN наружу) или продлить срок действия существующего исходящего отображения.

На рисунке показана структура пакетов операции PEER. Формат запросов и откликов PEER организован так, чтобы связанные логически поля имели одинаковое положение в пакетах.

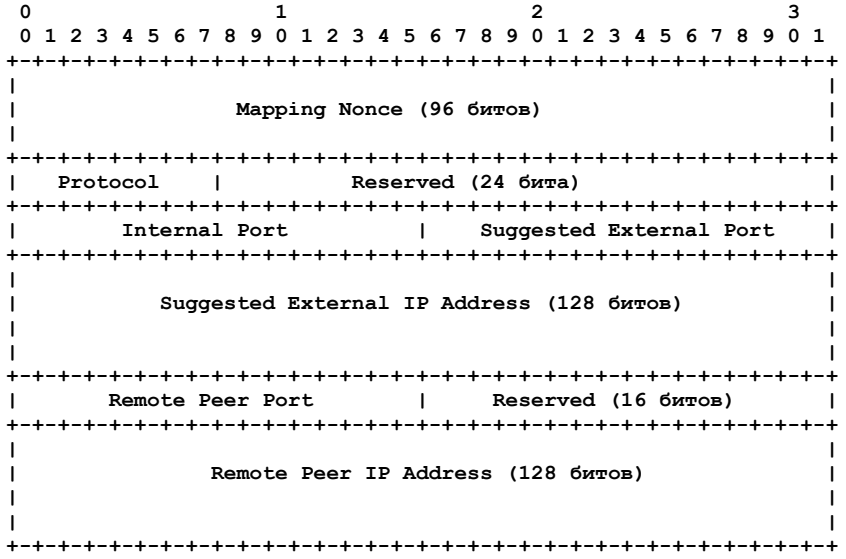

Рисунок 11. Запрос PEER.

Requested Lifetime (в общем заголовке) - срок действия

Запрашиваемое время действия отображения в секундах. Отметим, что с помощью запросов PEER невозможно сократить срок действия отображения (или удалить его с помощью lifetime=0).

Mapping Nonce - случайный идентификатор отображения

Случайное значение, выбираемое клиентом (см. параграф 12.2 Генерация запроса PEER). Допустимо использование значения 0 (вероятность его появления составляет примерно 2^{-96}).

Protocol - протокол

Протокол вышележащего уровня, связанный с этой операцией. Значения поля берутся из реестра протоколов IANA [proto_numbers]. Например, это поле будет содержать значение 6 (TCP), если операция связана с отображением для протокола TCP или 17 (UDP), если операция связана с отображением для UDP. **Недопустимо** использовать значение 0.

Reserved - резерв

24 резервных бита, которые **должны** устанавливаться в 0 при передаче и игнорироваться на приёме.

Internal Port - внутренний порт

Внутренний порт для отображения (нулевое значение **недопустимо)**.

Suggested External Port - предлагаемый внешний порт

Предлагаемый для отображения внешний порт. Если клиент PCP не знает номер внешнего порта или этот номер не имеет значения, в данном поле **должно** устанавливаться значение 0.

Suggested External IP Address - предлагаемый внешний адрес

Предлагаемый для отображения внешний адрес IP. Если клиент PCP не знает внешнего адреса или этот адрес не имеет значения, он **должен** использовать адрес «все нули», соответствующего семейства (см. раздел 5).

Remote Peer Port - порт удалённого партнёра

Порт удалённого партнёра для данного отображения (значение 0 **недопустимо**).

Reserved - резерв

16 резервных битов, которые **должны** устанавливаться в 0 при передаче и игнорироваться на приёме.

Remote Peer IP Address - адрес удалённого партнёра

IP-адрес удалённого партнёра с точки зрения клиента PCP. Это позволяет клиенту PCP не задумываться о наличии на пути преобразований NAT64 или NAT46, которые приводят к тому, что представляемый клиентом адрес IP отличается от реального IP-адреса удалённого хоста. Это поле позволяет избавиться от неоднозначностей, связанных с наличием множества соединений с одного порта внутреннего хоста с разными серверами. Адреса IPv6 в этом поле представляются напрямую, а адреса IPv4 с использованием синтаксиса IPv4-mapped (раздел 5).

При попытке восстановления утраченного отображения в качестве предлагаемых внешнего адреса и порта устанавливаются значения External IP Address и Port, полученные в предшествующем отклике PEER от сервера PCP. В первоначальном запросе PEER для внешнего адреса IP и порта устанавливаются значения 0.

Отметим, что для запросов PEER автоматически предполагается семантика, похожая на семантику опции PREFER_FAILURE. Если поле Suggested External IP Address или Suggested External Port отлично от нуля, а сервер PCP не может выделить предложенный внешний адрес IP, протокол или порт, сервер **должен** возвратить отклик об ошибке CANNOT_PROVIDE_EXTERNAL. Опция PREFER_FAILURE не нужна и не разрешена в запросах PEER и, если сервер PCP получает запрос PEER с опцией PREFER_FAILURE, он **должен** возвратить отклик MALFORMED_REQUEST.

На рисунке показана структура поле отклика PEER.

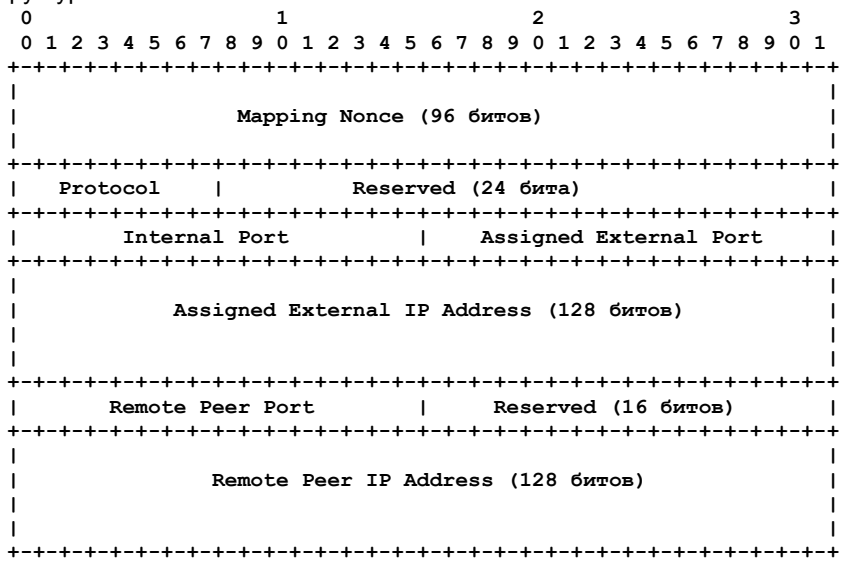

Рисунок 12. Отклик PEER.

Lifetime (в общем заголовке) - срок действия

При успешном завершении это поле указывает срок действия созданного отображения, в случае ошибки - время, в течение которого сервер PCP будет возвращать такую же ошибку при повторных запросах.

Mapping Nonce - случайный идентификатор отображения

Копируется из запроса.

Protocol - протокол

Копируется из запроса.

Reserved - резерв

24 резервных бита, которые **должны** устанавливаться в 0 при передаче и игнорироваться на приёме.

Internal Port - внутренний порт

Копируется из запроса.

Assigned External Port - выделенный внешний порт

При успешном завершении показывает выделенный для отображения внешний порт, в случае ошибки значение поля копируется из запроса.

Assigned External IP Address - выделенный внешний адрес

При успешном завершении показывает выделенный для отображения внешний адрес IPv4 или IPv6, в случае ошибки значение поля копируется из запроса.

Remote Peer Port - порт удалённого партнёра

Копируется из запроса.

Reserved - резерв

16 резервных битов, которые **должны** устанавливаться в 0 при передаче и игнорироваться на приёме.

Remote Peer IP Address - адрес удалённого партнёра

Копируется из запроса.

12.2 Генерация запроса PEER

В этом параграфе описаны действия клиента при генерации сообщения с кодом операции PEER.

Сообщение с кодом операции PEER **может** передаваться до организации двухстороннего взаимодействия с удаленным партнёром или после такой организации.

Если сообщение передаётся до организации соединения, оно рассматривается, как запрос PEER на создание исходящего динамического отображения на управляемом PCP устройстве.

Если сообщение передаётся после организации соединения, оно позволяет клиенту PCP узнать адрес IP, порт и срок действия имеющегося неявного динамического отображения и продлить срок действия отображения (для снижения объёма сообщений keepalive, как описано в параграфе 10.3).

Запросы PEER полезны также при восстановлении отображений после потери состояния отображений в NAT 9например, в результате перезапуска устройства).

Значение Mapping Nonce случайным образом выбирается клиентом в соответствии с рекомендациями по генерации случайных чисел [RFC4086] и применяется в процессе проверки откликов PCP (см. ниже) клиентом и проверки при обновлении отображений на сервере PCP. Клиент **должен** использовать своё значение nonce для каждого сервера PCP, с которым он взаимодействует и **рекомендуется** также выбирать новое значение nonce при каждой инициализации клиента PCP. Клиент **может** использовать разные значения mapping nonce для каждого отображения.

Сообщение PEER включает поле Remote Peer Address, адрес в котором указывается с точки зрения клиента PCP. Отметим, что в тех случаях, когда управляемое PCP устройство меняет семейство адресов (трансляция NAT46 или NAT64), адрес удалённого партнёра с точки зрения клиента PCP отличается от адреса, который виден с другой стороны устройства-транслятора.

12.3 Обработка запроса PEER

В этом параграфе описаны действия сервера при получении запросов с кодом операции PEER. Операции **следует** выполнять в порядке приведённого здесь описания.

Поля запроса PEER Protocol, Internal Port, Remote Peer IP Address, Remote Peer Port и Mapping Nonce копируются из запроса в отклик.

При создании неявных динамических отображений некоторые устройства NAT и МСЭ проверяют адреса получателей и не будут создавать неявных динамических отображений для недопустимых адресов (например, 127.0.0.1). Если управляемое PCP устройство выполняет такую проверку для неявных динамических отображений, ему **следует** также проверять допустимость полей удалённого адреса, протокола и порта для создаваемых операцией PEER явных динамических отображений. Если при проверке адреса удалённого партнёра в запросе PEER обнаружена его недопустимость, отображение не создаётся и клиенту возвращается отклик с кодом ошибки MALFORMED_REQUEST.

При получении запроса PEER сервер PCP проверяет свою таблицу отображений на предмет наличия отображения для {Protocol, Internal Address, Internal Port, Remote Peer Address, Remote Peer Port}.

Если соответствующего отображения не обнаружено в таблице, а предложенные внешний адрес и порт имеют значение 0 или могут быть выделены для указанного значения Protocol, создаётся новое отображение. За счёт создания таких отображений с помощью PEER мы избегаем «соперничества» между отправкой PEER и прибытием на устройство NAT или МСЭ первого исходящего пакета, а также обеспечиваем возможность использования операции PEER для восстановления утраченных исходящих динамических отображений (см. 16.3.1 Повторная организация отображений). После этого созданное с помощью операции PEER отображение трактуется, как неявное динамическое исходящее отображение (например, созданное в результате передачи клиентом пакета TCP SYN) и возвращается срок действия такого отображения (отметим, что на многих устройствах NAT и МСЭ время жизни таких отображений весьма мало и они удаляются после завершения двухстороннего трафика через NAT или МСЭ).

Если соответствующего отображения не найдено, а предложенный адрес и порт не могут быть выделены, новое отображение не создаётся и клиенту возвращается отклик с кодом результата CANNOT_PROVIDE_EXTERNAL.

Если соответствующее отображение найдено в таблице, но для него не зафиксировано успешной обработки PEER, значения полей Suggested External Address и Port в запросе игнорируются, срок жизни отображения устанавливается в соответствии приведённым ниже описанием и клиенту возвращается информация об имеющемся отображении. Это позволяет клиенту явно продлить срок действия отображения и/или узнать текущие значения для внешнего адреса, порта и срока действия отображения. Значение mapping_nonce для отображения сервер запоминает.

Если используется простая модель угроз (параграф 18.1) и отображение для внутреннего порта, протокола и внутреннего адреса уже организовано, но значение mapping nonce не соответствует (т. е., был обработан предшествующий запрос PEER), запрос **должен** быть отвергнут с возвратом клиенту кода ошибки NOT_AUTHORIZED и указанием в поле lifetime срока действия имеющегося отображения. Серверу PCP требуется лишь запоминать значения Mapping Nonce для каждого отображения. В данной спецификации не предъявляется требований к обработке Mapping Nonce для расширенной модели угроз.

Обработка значения Lifetime из PEER Opcode описана в разделе 15 Срок действия и удаление отображений. Отправка запросов PEER с очень малым сроком действия может применяться для запроса срока действия существующих отображений. Для того, чтобы клиенты PCP могли сократить частоту своих сообщений keepalive устройствам NAT и МСЭ, **рекомендуется** делать срок действия отображений, созданных или продлённых с помощью PEER, больше срока действия неявно созданных отображений.

Если все операции обработки завершились успешно (не было откликов об ошибках), создаётся отклик SUCCESS с полем Lifetime, указывающим срок действия отображения.

Если созданное или поддерживаемое с помощью PEER отображение не обновляется с применением PEER, оно возвращается к обычному для NAT поведению неявных отображений. Например, продолжающийся исходящий трафик будет сохранять отображение в соответствии с политикой NAT или МСЭ. Созданные или поддерживаемые с помощью PEER отображения могут быть прерваны в любой момент действиями клиента или сервера TCP (например, передачей TCP FIN или TCP RST), а также политикой NAT или МСЭ.

12.4 Обработка отклика PEER

В этом параграфе описаны действия клиента при обработке отклика с кодом операции PEER.

После обычной обработки отклика PCP отклик дополнительно сопоставляется с незавершённым запросом PEER путём сравнения внутреннего адреса IP (адрес получателя сообщения PCP или иной адрес IP, заданный опцией THIRD_PARTY), протокола, внутреннего порта, адреса и порта удалённого партнёра и nonce отображения. Остальные поля не сравниваются, поскольку сервер PCP устанавливает эти поля для предоставления информации об отображении, созданном Opcode. Сервер PCP будет передавать Mapping Update (параграф 14.2) при смене отображения (например, в результате изменения адреса IP).

Если результат имеет код NO_RESOURCES и запрос был сделан для создания или обновления отображения, клиенту PCP **не следует** передавать дополнительных запросов серверу PCP для новых отображений в течение (ограниченного) времени жизни.

При успешном отклике приложение может использовать выделенное значение Lifetime для снижения частоты сообщений keepalive от приложения для данного отображения NAT. Естественно, у приложения могут быть свои причины использовать более высокую частоту отправки keepalive. Например, назначенное PCP время жизни может составлять 1 час, а приложение может желать указывать состояние своего сервера (например, busy или away) более часто. Если отклик указывает неожиданный адрес IP или порт (например, в результате смены IP), клиент PCP захочет заново организовать соединение с удаленным сервером.

Если клиент PCP хочет сохранять данное отображение дольше указанного срока действия, он **может** полагаться на продолжающийся трафик изнутри наружу для уверенности в том, что отображение продолжает действовать, или **может** ввести новый запрос PCP до завершения срока. Рекомендуемое время обновления отображений PEER совпадает с описанным для отображений MAP в параграфе 11.2.1.

Примечание. Реализации должны ожидать, что сообщение PEER может содержать внешний адрес IP из другого семейства, нежели адрес удалённого партнёра, например, при использовании NAT64 или NAT46.

13 Опции MAP и PEER

В этом разделе описаны опции для операций MAP и PEER. Эти опции **недопустимо** применять с другими Opcode, если это явно не указано для соответствующего Opcode.

13.1 Опция THIRD_PARTY для MAP и PEER

Эта опция применяется, когда клиент PCP хочет контролировать отображение на другой внутренний хост. Опция используется с кодами MAP и PEER.

По причине проблем безопасности, связанных с THIRD_PARTY, эту опцию **недопустимо** реализовать или применять, если сеть, в которой передаются сообщения PCP, не является полностью доверенной. Например, если у клиента PCP установлены списки контроля доступа (ACL[1]), в соответствии с которыми разрешается доступ лишь к серверу PCP и сети между клиентом и сервером PCP.

Управляющее устройство будет применять эту опцию для управления сервером PCP от имени пользователей. Например, устройство управления, размещённое в центре сетевых операций, которое предоставляет интерфейс конечным пользователям или персоналу оператора и позволяет передавать запросы PCP с опцией THIRD_PARTY подходящему серверу PCP.

Ниже показан формат опции THIRD_PARTY.

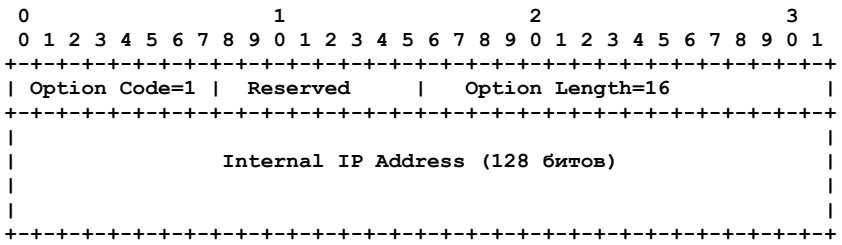

Рисунок 13. Опция THIRD_PARTY.

Internal IP Address
Внутренний IP-адрес для этого отображения.

Имя опции
THIRD_PARTY

Номер
1

Назначение
Указывает, что запрос MAP или PEER сделан не для того хоста, который передаёт опцию PCP.

Действительна для Opcode
MAP и PEER.

Размер
16 октетов

Может указываться
В запросах. В отклике может присутствовать лишь при наличии опции в соответствующем запросе.

Максимальное число экземпляров
1

В опцию THIRD_PARTY **недопустимо** включать тот же адрес, который указан в поле адреса отправителя пакета. Это связано с тем, что многие серверы PCP могут совсем не реализовать опцию THIRD_PARTY и с такими серверами избыточное использование клиентами опции THIRD_PARTY для указания своего адреса IP будет вызывать отказы при запросах отображения, которые в ином случае были бы успешны. Сервер PCP, получивший опцию THIRD_PARTY с тем же адресом, который является адресом отправителя пакета, **должен** возвращать код MALFORMED_REQUEST.

Сервер PCP **можно** настроить на разрешение или запрет использования опции THIRD_PARTY. Если эта опция разрешена, корректно уполномоченные клиенты могут выполнять эти операции от имени других хостов. Если опция запрещена и сервер PCP получает запрос PCP MAP с опцией THIRD_PARTY, он **должен** вернуть отклик UNSUPP_OPTION.

На абонентском оборудовании, реализующем сервер PCP, **рекомендуется** настроить по умолчанию запрет на отображения для других хостов. В таком случае, если пользователь хочет создать отображение для другого хоста, ему

[1]Access control list.

нужно взаимодействовать по отдельному каналу (out-of-band) с абонентским маршрутизатором (например, через административный интерфейс HTTP).

Провайдерским устройствам NAT и МСЭ, реализующим функции сервера PCP, **рекомендуется** разрешать опцию THIRD_PARTY, переданную корректно уполномоченным хостом. Если пакет приходит от неуполномоченного хоста, сервер PCP **должен** генерировать ошибку UNSUPP_OPTION.

Следует отметить, что опция THIRD_PARTY не требуется в распространённом современном варианте, когда ISP предоставляет клиенту один адрес IP и клиент применяет NAT для совместного использования данного адреса, поскольку в этом случае все хосты клиента с точки зрения ISP будут одним хостом.

Когда клиент PCP использует опцию THIRD_PARTY для организации и поддержки отображения от имени другого хоста, может быть полезно (при наличии возможности) клиенту PCP проверять реальное присутствие и активность этого хоста в сети. Иначе клиент PCP рискует сохранять это отображение в течение долгого времени после того, как устройство выйдет из сети. Это противоречит цели PCP создавать отображения с ограниченным сроком действия, автоматически удаляемые после того, как они станут не нужны.

13.2 Опция PREFER_FAILURE для MAP

Эта опция может применяться лишь с MAP Opcode.

Эта опция указывает, что серверу PCP не следует создавать отображение, если он не способен отобразить сразу предложенный внешний порт и предложенный внешний адрес. Без опции отображение будет создаваться.

PREFER_FAILURE никогда не требуется клиенту PCP для управления своими отображениями, а её применение требует дополнительной работы клиента и сервера PCP. Опция служит для взаимодействия с протоколами отображения, не поддерживающими PCP, которые применяют отличную от PCP семантику (например, UpnP IGDv1 [PNP-IGD-PCP], где семантика UPnP IGDv1 позволяет клиенту UPnP IGDv1 лишь задать отображение конкретного порта), или отдельными системами распределения портов между клиентами (например, доступный клиентам web-портал, который управляется тем же ISP что и сервер PCP). Сервер PCP **может** поддерживать эту опцию, если его разработчики считают нужным поддерживать такие нисходящие устройства или отдельные системы распределения портов. Серверы PCP, не предназначенные для взаимодействия с такими системами, могут не поддерживать опцию. Клиентам PCP, отличным от клиентов взаимодействия с UPnP IGDv1 или отдельных систем распределения портов, **не следует** применять эту опцию, поскольку она ведёт к неэффективной работе и нет уверенности в том, что все серверы PCP будут реализовать опцию. Ожидается, что в будущем опция будет отменена, поскольку все больше клиентов будут естественным путём поддерживать PCP.

Формат опции PREFER_FAILURE показан на рисунке 14.

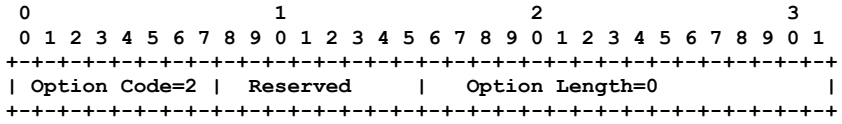

Рисунок 14. Опция PREFER_FAILURE.

Имя опции
PREFER_FAILURE

Номер
2

Назначение
Указывает, что серверу PCP не следует создавать другое отображение, если предложенные внешний порт и адрес не могут быть отображены.

Действительна для Opcode
MAP

Размер
0

Может указываться
В запросах. В отклике может присутствовать лишь при наличии опции в соответствующем запросе.

Максимальное число экземпляров
1

Если предложенные внешний адрес, протокол и порт не могут быть отображены, возвращается код CANNOT_PROVIDE_EXTERNAL. Это может быть в результате того, что внешний порт уже отдан динамическому выходному или входному или статическому отображению другого хоста или тот же внутренний адрес, протокол и порт уже имеют выходное динамическое отображение на другой внешний порт. Это может возникать также в результате того, что внешний адрес утратил доступность (например, в результате смены адресов). Сервер **может** указать в отклике оставшийся срок действия конфликтующего отображения + TIME_WAIT [RFC0793] с округлением до целого числа секунд в большую сторону.

Если запрос PCP включает опцию PREFER_FAILURE, а в поле Suggested External Port указан 0, запрос будет недействительным. Сервер PCP **должен** отвергнуть сообщение с возвратом кода ошибки MALFORMED_OPTION.

Серверы PCP **могут** ограничивать скорость обработки запросов PREFER_FAILURE с целью защиты от потока из 65535 последовательных запросов PREFER_FAILURE от клиентов, пытающихся проверить доступность внешних портов.

Может возникать конкуренция между MAP Opcode с опцией PREFER_FAILURE и Mapping Update (параграф 14.2). например, предыдущий хост локальной сети мог ранее использовать тот же внутренний адрес с отображением на тот же внутренний порт. Примерно в одно время с отправкой текущим хостом запроса MAP с опцией PREFER_FAILURE сервер PCP может передать спонтанное сообщение Mapping Update для старого отображения в результате внешнего изменения конфигурации, которое может выглядеть откликом на запрос нового отображения. По этой причине клиент PCP **должен** проверить свои внешний адрес IP, протокол, порт и nonce в отклике об успешном отображении на предмет совпадения со значениями, предложенными в запросе. Несоответствие говорит о том, что сообщение Mapping Update было передано до обработки запроса MAP.

13.3 Опция FILTER для MAP

Эта опция может применяться лишь с MAP Opcode.

Опция указывает желательность фильтрации входящих пакетов. Фильтруемый протокол указывается полем Protocol в запросе MAP Request, а IP-адрес и порт удалённого партнёра в опции FILTER указывают разрешённые адрес и порт отправителя для пакетов из Internet (остальной трафик блокируется). Размер префикса удалённого партнёра указывает значимые биты IP-адреса удалённого партнёра, это позволяет с помощью одной опции разрешить пакеты из всей подсети. После обработки запроса MAP с опцией FILTER и генерации отклика об успехе управляемое PCP устройство будет отбрасывать пакеты на своём внешнем интерфейсе, если они не совпадают с полями фильтра. Если политика безопасности разрешат, управляемое PCP устройство после отбрасывания пакета **может** передавать сообщение ICMP об ошибке.

Применение опции FILTER можно рассматривать как оптимизацию производительности. Поскольку все программы, использующие PCP для входящих соединений, фактически напрямую подключены к Internet и будут получать входящие соединения TCP и пакеты UDP, им желательно ограничить входящий трафик конкретным набором адресов источников для чего нужна проверка адресов отправителей входящего трафика и отклонение нежелательных пакетов. Однако опция FILTER особенно полезна для оптимизации производительности устройств с батарейным питанием, поскольку это позволяет снизить мощность, потребляемую на обработку нежелательного трафика.

Формат опции FILTER показан на рисунке 15.

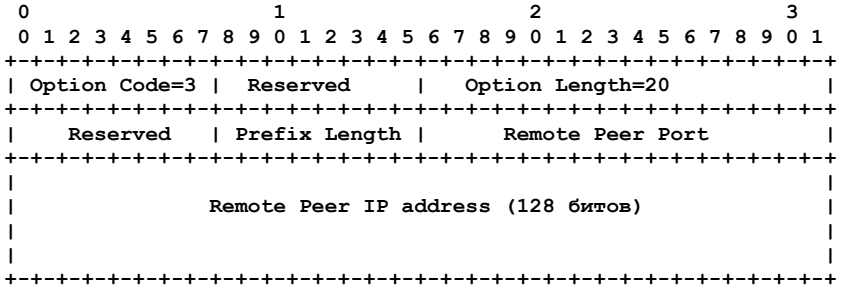

Рисунок 15. Схема опции FILTER.

Reserved
8 зарезервированных битов, которые **должны** устанавливаться в 0 при передаче и игнорироваться при получении.

Prefix Length
Указывает число битов адреса IPv4 или IPv6, используемых фильтром. Нулевое значение отменяет фильтрацию и удаляет все прежние фильтры (см. ниже).

Remote Peer Port
Номер порта удалённого партнёра (0 - все порты).

Remote Peer IP address
IP-адрес удалённого партнёра.

Имя опции
FILTER

Номер
3

Назначение
Задаёт фильтрацию входящих пакетов.

Действительна для Opcode
MAP

Размер
20 октетов

Может указываться
В запросах. В отклике может присутствовать лишь при наличии опции в соответствующем запросе.

Максимальное число экземпляров
Число фильтров ограничивается размером сообщения PCP.

Prefix Length указывает число битов адреса, учитываемых фильтром. Для адресов IPv4 (представляются в формате IPv4-mapped - ::FFFF:0:0/96) размер префикса может составлять от 97[1] до 128 битов, включительно, т. е. к реальному размеру префикса IPv4 добавляется 96. Для адресов IPv6 размер префикса может составлять от 1[1] до 128 битов, включительно. Значения, выходящие за указанные пределы заставят сервер PCP вернуть код MALFORMED_OPTION.

При включении множества опций FILTER в запрос MAP они обрабатываются в порядке получения (как при обычной обработке опций PCP) и запрошенные фильтры **могут** перекрываться. Если отображение уже есть (с фильтром или без него) и сервер получает запрос MAP с опцией FILTER, указанные в новой опции фильтры добавляются к имеющимся. Если для запроса MAP указан срок действия 0 и в нем имеется опция FILTER, возвращается код ошибки MALFORMED_OPTION.

Если при обработке какой-либо из опций FILTER в пакете запроса возник отказ, возвращается код ошибки (например, MALFORMED_OPTION, если одна из опций имеет некорректный формат) и (как для других ошибок PCP) состояние сервера PCP или управляемого PCP устройства не меняется.

Для удаления всех фильтров применяется Prefix Length=0. В полях Remote Peer Port и Remote Peer IP Address **должно** устанавливаться значение 0, а при получении они **должны** игнорироваться[2]. Механизмов удаления отдельных фильтров нет.

Для изменения имеющегося фильтра клиент PCP передаёт запрос MAP с двумя опциями FILTER, первая из которых указывает префикс размера 0 (удаление всех фильтров), а вторая задаёт новый адрес, протокол и порт удалённого партнёра. Если нужно указать множество партнёров, в запрос PCP включаются дополнительные опции FILTER.

[1] В оригинале ошибочно указано 96 и 0. См. https://www.rfc-editor.org/errata/eid3891. *Прим. перев.*
[2] В оригинале это предложение отсутствует. См. https://www.rfc-editor.org/errata/eid3891. *Прим. перев.*

Предполагается что серверы PCP и управляемые PCP устройства ограничивают число поддерживаемых удалённых партнёров (вплоть до одного). Если запрос MAP превышает установленное ограничение, этот запрос отвергается целиком с возвратом ошибки EXCESSIVE_REMOTE_PEERS, а состояние сервера PCP не меняется.

Все серверы PCP **должны** поддерживать хотя бы один фильтр на отображение MAP.

14 Быстрое восстановление

PCP включает функцию быстрого восстановления, которая позволяет клиентам PCP исправлять повреждённые отображения в течение нескольких секунд, а не минут или часов, которые требуются, если полагаться лишь на ожидание следующей процедуры обновления. Отказы отображений могут возникать при перезагрузке шлюзов NAT и потере состояния отображений или при смене IP-адреса шлюза NAT, делающей отображения недействительными.

Функция быстрого восстановления PCP позволяет пользователям, например, подключиться к удалённой машине по протоколу ssh, а затем перезагрузить своё устройство NAT или МСЭ (и даже заменить его физически на другое устройство) без потери соединения ssh.

Использование быстрого восстановления PCP является оптимизацией производительности процедуры самовосстановления PCP. Без этой функции клиенты PCP по-прежнему должны заново создавать нужное состояние при следующем обновлении их отображений, но эта процедура быстрого самовосстановления может занять часы, а не секунды, что вероятно не позволит предохранить активные соединения TCP от тайм-аутов.

Имеется два механизма быстрого восстановления, описанных ниже. Неспособность реализовать и развернуть механизм быстрого восстановления заставляет разработчиков приложений чаще обновлять состояние PCP, нежели это требуется, что ведёт к росту сетевого трафика. Поэтому сервер PCP, который может потерять состояние (например, при перезагрузке) или изменить отображение (например, в результаты смены адресов IP), **должен** реализовать механизм Announce Opcode или Mapping Update и **следует** реализовать оба.

14.1 Код операции ANNOUNCE

Этот механизм быстрого восстановления использует ANNOUNCE Opcode. Когда сервер PCP теряет своё состояние (например, при перезагрузке), он сбрасывает время Epoch в исходное стартовое значение (обычно 0) и передаёт отклик ANNOUNCE с групповым адресом в рамках канала (link-scoped, см. ниже), если на локальном интерфейсе имеется сеть с групповой адресацией, или при настройке на нем адресов IP и портов клиентов PCP передаёт индивидуальные сообщения ANNOUNCE в эти адреса и порты. Это означает, что сообщения ANNOUNCE могут быть доступны не всей сети (например, для сети без групповой адресации на канале между сервером и клиентами PCP). Кроме того, запрос ANNOUNCE может быть передан (индивидуально) клиентом PCP, желающим получить индивидуальный отклик ANNOUNCE, подобно любому другому Opcode.

При получении пакетов отклика PCP с аномальным значением Epoch клиенты узнают, что сервер PCP потерял состояние и восстанавливают свои утраченные отображения.

14.1.1 Операция ANNOUNCE

Запросы и отклики PCP ANNOUNCE Opcode не имеют специфического для этого кода содержимого (т. е. размер связанных с Opcode данных равен 0). Поле Requested Lifetime в запросах и поле Lifetime в откликах имеют значения 0 при передаче и игнорируются получателем.

Если сервер PCP получает запрос ANNOUNCE, он сначала анализирует его и генерирует отклик SUCCESS при успешном разборе и обработке ANNOUNCE. Если поле IP Address для клиента не совпадает с адресом отправителя пакета или в пакете имеется иная ошибка форматирования, например, размер пакета меньше 24 октетов, сервер возвращает ошибку. Отметим, что в будущем для PCP ANNOUNCE Opcode **могут** быть добавлены опции, поэтому клиентам и серверам PCP нужно быть готовыми принимать опции с ANNOUNCE Opcode.

> Обсуждение. Сообщения с запросами от клиента к серверу передаются из любого порта клиента в порт UDP 5351 на сервере, групповые уведомления клиентам от сервера передаются из серверного порта UDP (5351) в порт UDP 5350 на стороне клиента. Причина использования разных портов на приёмной стороне заключается в том, что устройство может одновременно выступать в нескольких ролях. Например, многофункциональный домашний шлюз может предоставлять услуги NAT (сервер PCP), а также совместный доступ к принтеру (клиент PCP) или домашний компьютер (клиент PCP) может обеспечивать совместный доступ в Internet (NAT), для чего он должен быть сервером PCP. Такие устройства выступают в качестве клиента и сервера одновременно и серверная программа PCP на устройстве может использовать не те компоненты, которые реализует клиент PCP. Серверным программам PCP нужно прослушивать индивидуальные пакеты с запросами клиентов, а клиентским - групповые анонсы перезапуска. Во многих сетевых API сложно или невозможно иметь 2 независимых клиента, прослушивающих одновременно индивидуальные и групповые пакеты на одном порту. Поэтому применяется два разных порта.

14.1.2 Генерация и обработка запрошенных сообщений ANNOUNCE

PCP ANNOUNCE Opcode **может** передавать клиент PCP (по индивидуальному адресу). Значение Requested Lifetime **должно** быть 0.

Когда сервер PCP получает ANNOUNCE Opcode, успешно анализирует и обрабатывает его, он генерирует отклик SUCCESS с нулевым временем жизни.

Эта функциональность позволяет клиенту PCP определить значение Epoch на сервере или убедиться в работоспособности сервера без изменения состояния сервера.

14.1.3 Генерация и обработка незапрошенных сообщений ANNOUNCE

При передаче незапрошенных откликов ANNOUNCE Opcode они **должны** иметь нулевой код результата (SUCCESS), а пакет **должен** передаваться с индивидуального адреса IP и из порта UDP, через который принимаются запросы PCP (чтобы описанная в параграфе 8.3 обработка откликов PCP воспринимала сообщение). Эти сообщения обычно передаются по групповому адресу, но могут быть и индивидуальными. Групповые анонсы перезапуска PCP

передаются по адресу 224.0.0.1:5350 и/или [ff02::1]:5350, как описано ниже. Передача анонсов перезапуска PCP по индивидуальному адресу требует от сервера PCP знать IP-адреса и порты своих слушающих клиентов, поэтому такая передача анонсов перезапуска применима лишь для тех серверов PCP, которые сохраняют адреса и порты своих клиентов даже после потери и восстановления своего состояния.

Когда серверное устройство PCP перезагружается, перезапускает свою машину NAT или иначе меняет своё состояние с потерей данных отображения из своего предшествующего состояния (или входит в состояние где даже неизвестно о наличии предыдущего состояния, которое было потеряно), сервер **должен** информировать клиентов PCP об этом факте путём индивидуальной или групповой передачи беспричинных пакетов откликов PCP ANNOUNCE Opcode, как показано ниже, через пути, по которым сервер принимает запросы PCP. Если передаётся групповое сообщение ANNOUNCE, серверное устройство PCP, которое воспринимает запросы PCP по протоколу IPv4 передаёт Restart Announcement по групповому адресу 224.0.0.1:5350 (групповой адрес для всех хостов - All Hosts), а устройство, воспринимающее запросы PCP по протоколу IPv6 передаёт Restart Announcement по групповому адресу IPv6 [ff02::1]:5350 (все хосты локального сегмента). Серверное устройство, воспринимающее запросы PCP по протоколам IPv4 и IPv6 передаёт пару сообщений Restart Announcements - по одному в каждый из указанных адресов. При передаче индивидуальных сообщений ANNOUNCE отклики направляются по адресам IP и портам клиентов PCP. Для компенсации потери пакетов серверное устройство PCP **может** передать такие пакеты (или пары пакетов) до 10 раз (с подходящим значением Epoch Time в каждом, отражающим время между отправкой сообщения) при этом интервал между первыми двумя сообщениями должен быть не менее 250 мсек, а затем как минимум удваивался для каждого последующего сообщения.

Клиент, который передаёт запросы PCP серверу PCP по пути с поддержкой групповой адресации, реализует функцию Restart Announcement и хочет получать эти анонсы, **должен** слушать эти PCP Restart Announcement (беспричинные пакеты откликов PCP ANNOUNCE Opcode) на подходящих интерфейсах с поддержкой групповой адресации, в которые он передаёт запросы PCP, и **может** также слушать индивидуальные анонсы от сервера (используя порт UDP, который он применяет для отправки индивидуальных запросов PCP и для приёма индивидуальных откликов от этого сервера PCP). Клиентское устройство PCP, которое передаёт запросы PCP, используя IPv4, слушает пакеты IPv4, переданные по групповому адресу 224.0.0.1:5350. Клиентское устройство PCP, которое передаёт запросы PCP, используя IPv6, слушает пакеты IPv6, переданные по групповому адресу [ff02::1]:5350. Клиентское устройство PCP, которое передаёт запросы PCP, используя IPv4 и IPv6, слушает оба типа Restart Announcement. Следует использовать опцию сокета SO_REUSEPORT или эквивалент для группового порта UDP, если ОС хоста требует множества независимых экземпляров для прослушивания одного группового порта UDP.

При получении индивидуального или группового пакета с откликом PCP ANNOUNCE Opcode клиент PCP **должен** (как и для всех полученных откликов PCP) проверить адрес отправителя анонса и если значение Epoch Time выходит за пределы ожидаемого для этого сервера, клиент **должен** выждать случайное время от 0 до 5 секунд (для предотвращения синхронизации всех клиентов PCP), а затем передать новые запросы PCP для всех отображений на этом сервере, чтобы восстановить утраченные состояния. Использование полей Suggested External IP Address и Suggested External Port fields в запросах на обновление позволяет клиенту напомнить перезагруженному серверному устройству PCP свои прежние отображения, что позволяет во многих случаях воссоздать их. Для серверных устройств PCP, которые перезагружаются сравнительно быстро, обычно можно восстановить потерянные состояния отображения достаточно быстро, чтобы существующие соединения TCP и обмен UDP не столкнулись с тайм-аутом и продолжили работать без сбоев. Если в откликах PCP значение Epoch Time находится внутри ожидаемого диапазона, клиент PCP не создаёт своих отображений заново. Как и для всех откликов PCP, после получения и проверки сообщения ANNOUNCE клиент обновляет своё значение Epoch для сервера, как описано в параграфе 8.5.

14.2 Обновление отображений PCP

Этот механизм быстрого восстановления применяется, когда сервер PCP помнит своё состояние и видит непригодность имеющихся отображений (например, изменился внешний адрес IP управляемого PCP устройства NAT).

Ожидается, что серверы, которые перенастраиваются администратором или имеют часто меняющийся адрес WAN (например, домашние маршрутизаторы CPE), будут поддерживать эту функцию. Предполагается также, что серверы, которые не перенастраиваются (например, операторский CGN), не будут реализовать эту функцию.

Если серверное устройство PCP не забыло своё состояние отображений, но по какой-то причине определило непригодность части отображений (например, домашнему шлюзу был назначен другой внешний адрес IPv4 восходящим сервером DHCP), этот сервер PCP автоматически исправляет свои отображение и уведомляет клиентов в соответствии с описанной ниже процедурой.

Для управляемых PCP отображений следует обновить внешний IP-адрес и порт подходящими доступными значениями для каждого серверного устройства PCP, а затем передать индивидуальные отклики PCP MAP или PEER (что подходит для отображения), чтобы информировать клиента PCP о новом внешнем адресе IP и номере порта. Такие незапрошенные отклики MAP или PEER обычно возвращаются в ответ на клиентский запрос MAP или PEER с новыми значениями External IP Address и External Port и передаются в тот же IP-адрес и порт клиента, которые сервер PCP использовал при предшествующем отклике для того же отображения. Если предшествующий связанный запрос включал опцию THIRD_PARTY, эта опция **должна** быть указана в Mapping Update, поскольку она требуется клиенту PCP для устранения неоднозначности отклика. Если предшествующий связанный запрос включал опцию PREFER_FAILURE, но прежние внешний адрес IP, протокол и порт не могут быть предоставлены, **следует** возвращать ошибку CANNOT_PROVIDE_EXTERNAL. Если предшествующий связанный запрос включал опцию FILTER, фильтры переносятся в новое отображение и опция FILTER передаётся в отклике Mapping Update. Необязательные опции **не следует** передавать в отклике Mapping Update.

> Обсуждение. Можно было бы сделать так - сервер PCP (1) передаёт ANNOUNCE Opcode клиенту PCP, клиент отвечает на это (2) отправкой нового запроса MAP и (3) получает отклик MAP. Вместо этого сервер использует более рациональный вариант, просто отправляя сообщение, которое он передал бы в (3).

Для компенсации потери пакетов серверному устройству PCP **следует** передавать такие пакеты 3 раза с соответствующей корректировкой значений Epoch Time в каждом пакете для учёта прошедшего времени. Интервал между первым и вторым уведомлением **должен** быть не менее 250 мсек, между вторым и третьим - не менее 500 мсек.

После получения сервером PCP обновлённого состояния для этого отображения серверу **следует** прекратить повтор передачи этих отображений, поскольку этого больше не требуется.

При получении таких обновлённых откликов MAP или PEER клиент PCP использует информацию из откликов для настройки сервера rendezvous или повторного подключения к серверу, соответственно. Для MAP это будет означать, что записи DNS и другие данные об адресах и портах записаны с каким-то зависящим от приложения сервером rendezvous. Для откликов PEER, указывающих ошибку CANNOT_PROVIDE_EXTERNAL, это обычно означает организацию новых соединений с серверами. При каждом изменении внешнего адреса или порта имеющиеся соединения TCP и UDP будут теряться - PCP не может это предотвратить, но пытается обеспечить механизм уведомлений для снижения негативного влияния.

15 Срок действия и удаление отображений

Клиент PCP запрашивает определённый срок действия отображений, а сервер PCP отвечает назначенным сроком действия. Сервер PCP **может** предоставить срок действия, отличающийся от запрошенного. Серверу PCP **следует** поддерживать возможность настройки максимального и минимального срока действия и в качестве минимального **следует** устанавливать срок 120 секунд. В качестве максимального срока **следует** устанавливать оставшееся время действия IP-адреса, назначенного клиенту, если эта информация доступна (например, от сервера DHCP), половину от срока выделения IP-адресов в данной сети, если оставшийся срок действия адреса недоступен, или 24 часа. Слишком большой срок действия может приводить к неоправданному расходу портов, если внутренним хостам уже не требуется получение трафика или они отключились от сети. Эти рекомендации не являются строгими и администраторам следует оценить компромиссы для выбора минимального и максимального срока действия отображения (Lifetime) в их сети.

После положительного ответа сервера PCP на запрос MAP для того или иного срока действия отображение портов сохраняется в течение заданного срока, пока это время не снижается клиентом PCP (до меньшего значения или 0) или сервер PCP не теряет состояние (например, отказ). Отображения, созданные запросами PCP MAP, не являются особыми и не отличаются от отображения, созданных иными способами. В частности, реализация может продлить срок действия, если исходящий трафик выходит за рамки срока действия, назначенного PCP. Клиентам PCP **недопустимо** зависеть от этого поведения для сохранения активности отображений и они **должны** явно обновлять свои отображения, как того требует поле Lifetime в откликах PCP.

При получении отклика PCP с абсурдно большим сроком действия клиенту PCP **следует** вести себя как при получении более адекватного значения (например, 24 часа) и соответственно обновлять отображение, чтобы на случай удаления статического отображения клиент поддерживал желаемое отображение.

Приложение, которое забыло назначенные ему через PCP отображения (например при отказе приложения или ОС), будет запрашивать новые отображения PCP. Это приведёт к расходу портов, если приложение будет привязано к другому порту. Приложение также будет, вероятно, инициировать новые исходящие соединения TCP, которые будут создавать динамические исходящие отображения без применения PCP, также расходующие порты. Если для внутренних хостов установлены ограничения на число используемых портов, это может создавать проблемы.

Для оказания помощи в очистке состояний PCP при размещении управляемого PCP устройства вместе с сервером назначения адресов (DHCP), что типично для домашних CPE, **рекомендуется** следующее поведение - если адрес IP недействителен (например, завершение аренды DHCP или отправка клиентом DHCP явного сообщения DHCP RELEASE), управляемому PCP устройству **следует** отбросить динамические отображения, связанные с этим адресом.

При использовании NAT один внешний порт может в разное время назначаться разным внутренним хостам. Например, если внутренний хост, использующий внешний порт, перестаёт передавать трафик через этот порт, срок отображения может закончиться и тот же порт позднее может быть выделен другому внутреннему хосту. В результате этот хост может получить входящий трафик, адресованный прежнему хосту. Обычно это происходит непреднамеренно и переназначение внешнего порта выполняется лишь после того, как текущий владелец порта не пользуется им достаточно долго. Была бы неприемлемой возможность использования PCP злоумышленниками для преднамеренного ускорения переназначения внешнего порта с целью захвата трафика, предназначенного текущему владельцу, путём (i) подделки запросов PCP с использованием адреса IP и nonce текущего владельца для ускоренного удаления или сокращения срока действия и (ii) последующего запроса этого порта для себя.

Поэтому в простой модели безопасности для защиты от таких атак PCP **недопустимо** разрешать запросы PCP (даже если они представляется исходящими от текущего владельца отображения), вынуждающие сократить срок действия отображения при наличии для этого отображения исходящего трафика. Сервер PCP **должен** установить для отображения срок действия не меньше времени, остающегося до окончанию отображения при отсутствии для того исходящего трафика. Это означает, что запросы MAP или PEER с нулевым сроком действия будут устанавливать 0 для назначенного времени (т. е. удалять отображение) лишь в том случае, когда использующий это отображение хост не передаёт пакетов в течение заданного времени ожидания, в противном случае назначенное время будет устанавливаться в соответствии с оставшимся временем допустимого бездействия.

Наконец, для снижения уровня нежелательного трафика и повреждения данных TCP и UDP назначенный внешний порт (с помощью MAP Opcode или PEER Opcode) **не следует** повторно выделять в течение интервала, равного ограничению времени повторного использования в NAT для неявных динамических отображений (обычно, максимальное время жизни сегмента TCP - 2 минуты [RFC0793]). Кроме того, для подавления атак с захватом портов выделенный внешний порт **не следует** снова использовать до истечения интервала, равного времени, в течение которого управляемое PCP устройство может обычно поддерживать бездействующие (нет трафика) неявные динамические отображения (например, 2 минуты для UDP [RFC4787] и 124 минуты для TCP [RFC5382]). Однако в этом интервале серверу PCP **следует** разрешать повторное заявление порта тем же клиентом (тот же в данном случае означает совпадение внутреннего адреса IP, внутреннего порта и nonce для отображения).

15.1 Обработка срока действия для MAP

При нулевом значении запрошенного срока действия выполняется один из указанных ниже вариантов.

- Если поля протокола и внутреннего порта отличны от 0, это означает запрос незамедлительного удаления указанного отображения.

- Если поле протокола отлично от нуля, а внутренний порт имеет значение 0, это означает запрос на удаление предыдущего «шаблонного» (все порты) отображения для этого протокола. Значение nonce **должно** совпадать с указанным при создании «шаблонного» отображения nonce.

- Если поля протокола и внутреннего порта имеют значение 0, это указывает запрос на удаление предыдущего отображения DMZ host (весь входящий трафик для всех протоколов). Значение nonce **должно** совпадать с указанным при создании отображения DMZ host.

- Если поле протокола имеет значение 0, а внутренний порт - ненулевое значение, это говорит о недействительности запроса и сервер PCP **должен** вернуть клиенту код ошибки MALFORMED_REQUEST.

В запросах, где Lifetime = 0, поля Suggested External Address и Suggested External Port **должны** устанавливаться в 0 при передаче и **должны** игнорироваться принимающей стороной. В поле Suggested External Address должен быть указан подходящий адрес, содержащий только нули, в зависимости от того, запрашивается удаление для внешнего адреса IPv4 или IPv6. Значения поле Suggested External Address и Suggested External Port копируются в поля назначенных внешнего адреса и внешнего порта в отклике[1].

Запросы PCP MAP могут удалять или сокращать срок действия лишь для отображений, созданных с помощью MAP. Если клиент PCP пытается удалить статическое (т. е. созданное за пределами PCP) или исходящее (неявное или PEER) отображение, сервер PCP **должен** возвращать NOT_AUTHORIZED. Если клиент PCP пытается удалить отсутствующее отображение, возвращается код SUCCESS (это требуется ля того, чтобы PCP возвращал один отклик для повторов или дубликатов запроса). Если запрос на удаление сформирован корректно и успешно обработан, создаётся отклик SUCCESS с полями протокола и внутреннего порта из запроса и сроком действия 0. Для входящих отображения (т. е. статических или динамически созданных MAP) **недопустимо** снижение срока действия с помощью сообщений транспортного протокола (например, TCP RST, TCP FIN). Отметим, что опций THIRD_PARTY (параграф 13.1), если она разрешена, также может удалять созданные PCP отображения MAP.

16 Вопросы реализации

Раздел 16 содержит рекомендации, которые могут помочь разработчикам, но не являются нормативными.

16.1 Реализация MAP с EDM Port-Mapping NAT

Для неявных динамических отображений поведение некоторых устройств NAT не зависит от конечных точек (EIM[2]), тогда как поведение других зависит (EDM[3]). Устройства NAT с поведением EIM не подвержены описанной здесь проблеме. IETF настоятельно рекомендует поведение EIM [RFC4787][RFC5382].

В устройствах EDM NAT один внешний порт может применяться для входящего и исходящего динамических отображений (для одного или разных внутренних хостов). Это осложняет взаимодействие с MAP Opcode. Для таких устройств NAT предусмотрено два способа, описанных ниже.

1. Использовать для входящих (например, созданных MAP) и исходящих отображений разные наборы портов для смягчения проблем взаимодействия между ними.

2. При получении пакета (входящего из Internet или исходящего от внутреннего хоста), сначала предпринимается попытка использовать для его обработки динамическое выходное отображение. Если соответствия не найдено, применяется входящее динамическое отображение. Это по сути даёт входящим отображениям более высокий приоритет.

16.2 Срок действия явных и неявных динамических отображений

Независимо от режима NAT EIM или EDM, возможно, что одно (или несколько) выходных отображений, использующих тот же внутренний порт на внутреннем хосте, могут быть созданы до или после запроса MAP. В таких случаях важно соблюдение NAT срока действия, указанного в отклике MAP. В частности при создании входного отображения с помощью MAP Opcode реализация должна гарантировать, что прерывание выходного отображения (например, с помощью TCP FIN) не нарушит созданное MAP входное отображение.

16.3 Восстановление при отказах PCP

Если происходит событие, в результате которого сервер PCP теряет состояние динамических отображений (например, отказ или потеря питания), созданные PCP отображения теряются. Случайные потери состояния могут быть неизбежными в домашних устройствах NAT, которые не записывают временные данные в энергонезависимую память. Предполагается, что в средах сервис-провайдеров потери состояния будут редкими (резервное питание, запись данных на диск и т. п.). Конечно при полном отказе оборудования сервис-провайдера (например, программный сбой) состояние все равно может теряться.

Время Epoch позволяет клиенту сделать вывод о возможной потере состояния сервером PCP. Когда наблюдаемое значение Epoch Time выходит за пределы ожидаемого диапазона, клиент PCP может попытаться заново создать отображения в соответствии с описанными в последующих параграфах процедурами.

Дополнительный анализ сценариев отказа PCP планируется рассмотреть в отдельном документе [PCP-FAIL].

16.3.1 Повторная организация отображений

Пакет обновления отображения форматируется аналогично исходному запросу отображения. К точки зрения клиента это будет обновлением имеющегося отображения, однако для недавно загруженного сервера PCP это представится новым запросом отображения. В обычном процессе регулярного обновления своих отображений до завершения срока их действия клиент PCP будет автоматически восстанавливать потерянные отображения.

[1] В оригинале этот абзац был сформулирован нечетко. См. https://www.rfc-editor.org/errata/eid3621. Прим. перев.
[2] Endpoint-independent mapping - независимое от конечной точки отображение.
[3] Endpoint-dependent mapping - зависимое от конечной точки отображение.

Когда сервер PCP теряет состояние и начинает обработку новых сообщений PCP, его время Epoch сбрасывается и отсчёт начинается заново. В результате получения пакета с полем Epoch Time, выходящим за пределы ожидаемого диапазона (параграф 8.5), что указывает на перезагрузку или другую потерю состояния сервером, клиент может начать обновление своих отображений быстрее, нежели при обычной процедуре обновления.

16.3.2 Поддержка отображений

Клиент PCP обновляет отображение путём передачи нового запроса PCP с информацией, полученной из прежнего отклика PCP. Сервер PCP будет отвечать с указанием нового срока действия. В результате перенастройки или отказа сервера PCP может смениться внешний адрес IP и/или внешний порт, а также сам сервер PCP (по причине нового маршрута к другому серверу PCP). Такие события редки и не являются ошибкой. Сервер PCP будет просто возвращать клиенту новый внешний адрес и/или порт, а клиенту следует записать новый внешний адрес и порт для своей службы rendezvous. Для более быстрого обнаружения таких событий сервер, которому нужна максимально возможная доступность, может предпочесть указание более коротких сроков действия в запросах PCP, чтобы взаимодействие с этим сервером происходило чаще. Это технический компромисс, основанный на учёте (i) приемлемого времени простоя для соответствующей службы, (ii) ожидаемой вероятности потери состояния устройствами NAT или МСЭ и (iii) приемлемого уровня служебного трафика PCP.

Если у клиента PCP имеется несколько отображений, значение Epoch Time достаточно извлечь лишь для одного из них, чтобы определить, действительно ли сервер PCP потерял состояние. Если клиент хочет проверить время Epoch для сервера PCP, он передаёт запрос PCP для любого из своих отображений. Отклик будет содержать текущее значение Epoch Time. В таком запросе клиент может продлить срок действия отображения (запросив дополнительное время) или сохранить текущий (запросив оставшееся число секунд действия отображения).

Если клиент PCP меняет свой внутренний адрес IP (например, при перемещении внутреннего хоста в другую сеть) и хочет по-прежнему получать входящий трафик, ему нужно создать новое отображение для этого адреса. Обычно новое отображение требует также обновления связанного с приложением сервера rendezvous, если внешний адрес и порт отличаются от прежних значений (см. параграфы 10.1 и 11.5).

16.3.3 SCTP

Хотя SCTP использует номера портов как TCP и UDP, протокол SCTP работает иначе при расположении за устройством NAT с общим адресом, поскольку номера портов SCTP не меняются [SCTPNAT]. Исходящее динамическое отображение SCTP использует тег проверки ассоциации вместо номеров порта локального и удалённого партнёра. Как и в TCP, явные исходящие отображения могут создаваться для снижения интервалов keepalive, а явные входные отображения могут создаваться пассивными слушателями для приёма новых ассоциаций на внешнем порту.

Поскольку поддерживающие SCTP устройства NAT (в настоящее время) не меняют номера портов SCTP, они не могут назначать внешний порт, отличающийся от внутреннего порта клиента. Клиент PCP, делающий запрос MAP для SCTP, должен учитывать это ограничение. Клиенту PCP **следует** делать запрос SCTP MAP как для случая TCP MAP - в начальном запросе PCP MAP **следует** указать 0 для внешнего адреса и порта, а в последующих запросах обновления **следует** указывать назначенный внешний адрес и порт. Однако с учётом того, что современные устройства SCTP NAT могут назначать только совпадающий с внутренним внешний порт, оно может оказаться неспособным назначить внешний порт, поскольку он уже выделен другому клиенту PCP. Такое событие вероятно при наличии в локальной сети нескольких экземпляров данного сервиса SCTP, поскольку два экземпляра могут прослушивать один общеизвестный порт SCTP на соответствующих хостах, но они не смогут использовать один порт на внешнем адресе шлюза NAT. Определённый внешний порт может оказаться недоступным и по иным причинам, таким как использование самим устройством NAT или запрет политикой, как указано в параграфе 11.3 Обработка запроса MAP. Если внешний порт, совпадающий с внутренним, не может быть назначен (и SCTP NAT не меняет порты SCTP), устройство SCTP NAT **должно** возвращать клиенту ошибку CANNOT_PROVIDE_EXTERNAL. Отметим, что это ограничение создаёт дополнительную нагрузку на сервер SCTP, чьи запросы MAP отвергаются, поскольку он должен закрыть имеющийся приёмный сокет и пытаться применить другой порт, пока не будет найден подходящий, которые свободен снаружи.

Описанные сложности SCTP связаны с совместным использованием адреса, которого можно избежать (например, с помощью отображения 1:1 в NAT или МСЭ).

16.4 Репликация адреса отправителя в заголовок PCP

Все запросы PCP повторяют IP-адрес клиента PCP в заголовке PCP. Это служит для обнаружения неожиданного изменения адреса (NAT) в пути между клиентом и сервером PCP. В ОС с поддержкой API сокетов клиентам PCP **рекомендуется** выполнять указанные ниже действия для вставки корректного адреса в заголовки PCP.

1. Создать сокет UDP.
2. Вызвать функцию connect на этом сокете UDP, используя адрес и порт нужного сервера PCP.
3. Вызвать функцию getsockname() для получения sockaddr с адресом отправителя, который ядро будет указывать в пакетах UDP, передаваемых через этот сокет.
4. Если адрес относится к IPv4, он кодируется в IPv4-mapped адрес IPv6. Адрес IPv4-mapped или естественный адрес IPv6 помещается в поле клиентского адреса IP в заголовке PCP.
5. Запрос PCP передаётся с использованием подключённого сокета UDP.

16.5 Диаграмма состояний

Каждая запись отображения в управляемом PCP устройстве проходит через конечный автомат, показанный ниже. Эта диаграмма состояний не является нормативной.

NO_ENTRY
Недействительное состояние представляет отсутствующую запись (Entry) и является единственным при старте.

M-R
Запрос MAP.

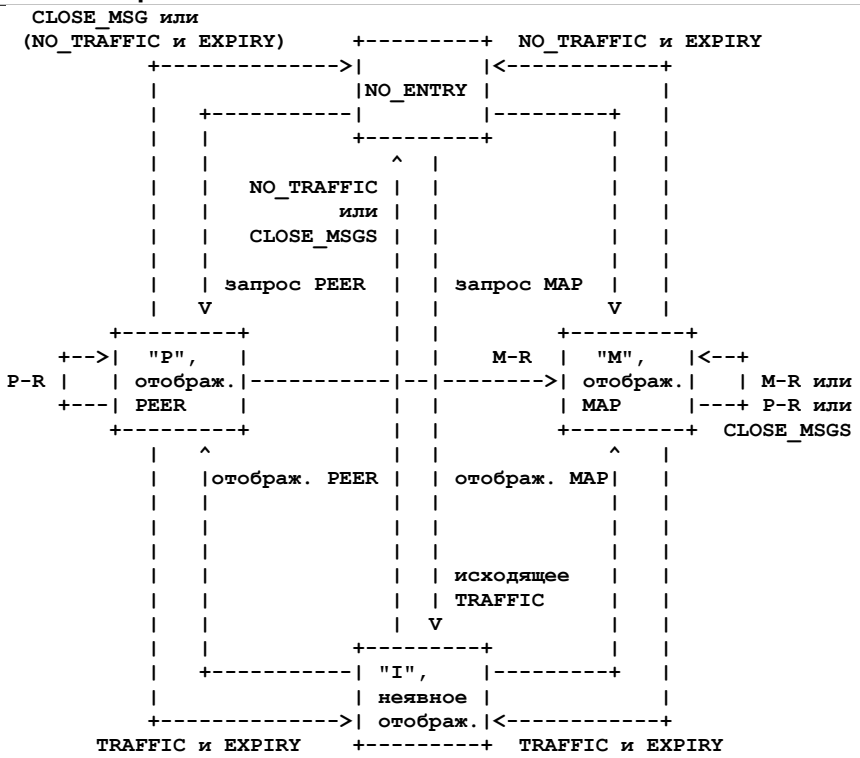

Рисунок 16. Диаграмма состояний PCP.

P-R
Запрос PEER.

M
Запись отображения при создании по запросу MAP.

P
Запись отображения, созданная/поддерживаемая отображением PEER.

I
Неявное отображение, созданное исходящим пакетом от клиента (например, TCP SYN), или состояние, когда срок действия созданного PCP истек, но сохраняется активный трафик.

EXPIRY
Закончился срок действия отображения PEER или MAP.

TRAFFIC
Трафик, который видит управляемое PCP устройство, использующий эту запись в течение срока её действия. Трафик может быть входящим или исходящим.

NO_TRAFFIC
Указывает отсутствие трафика (TRAFFIC).

CLOSE_MSG
Протокольные сообщения от клиента или сервера для завершения сессии (например, TCP FIN или TCP RST) в соответствии с обработкой таких протокольных сообщений устройством NAT или МСЭ.

Примечания к рисунку.

1. «И» указывает, что для перехода нужно выполнение условий по обе стороны «и», «или» указывает, что достаточно любого из событий.

2. Переход из состояния M в состояние I зависит от реализации.

17 Вопросы развёртывания

17.1 Фильтрация на входе

Как и неявные динамические отображения, создаваемые исходящими пакетами TCP SYN, явные динамические отображения PCP используют IP-адрес отправителя пакета в качестве внутреннего адреса для отображения. Поэтому **следует** применять входную фильтрацию [RFC2827] на пути между внутренним хостом и сервером PCP для предотвращения вставки обманных пакетов.

17.2 Квотирование отображений

На управляемых PCP устройствах, создающих состояние при организации отображения (например, NAT), серверу PCP **следует** поддерживать квоты для отображений на уровне хоста и/или абонента. Конкретная реализация зависит от применения сервером PCP раздельных квот для неявных, явных и статических отображений или иной политики.

18 Вопросы безопасности

Целью PCP является расширение возможности конечных узлов управлять своим состоянием NAT, более эффективная обработка и контроль ошибок отображений NAT по сравнению с имеющимися механизмами неявного отображения в устройствах NAT и МСЭ с учётом состояния. Защитной целью протокола PCP является ограничение возможности организации новых атак, нацеленных на отказ в обслуживании (DoS[1]) и несанкционированное изменение состояния отображений. Одним из наиболее серьёзных следствий несанкционированного изменения отображений является

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

кража трафика. Все отображения, которые конкретный хост может создать с помощью неявных механизмов, считаются полномочными. Конфиденциальность отображений не является обязательной даже в случаях передачи сообщений PCP по путям, которые не будут использоваться отображённым трафиком.

18.1 Простая модель угроз

Серверы PCP защищены от атак, где злоумышленник не может подделать пакет так, чтобы он казался исходящим из внутренней сети. Клиенты PCP защищены от атак, где могут применяться пакеты с поддельным адресом сервера PCP.

Защита от атакующих, которые могут изменять или отбрасывать пакеты между внутренней сетью и сервером PCP или внедрять обманные пакеты, которые представляются исходящими из внутренней сети, не рассматривается. Такие злоумышленники могут перенаправить трафик на нужный им хост.

В простой модели угроз (Simple Threat Model) сервер PCP защищён, если он ограничен так, что не создаёт явных отображений, которые он будет настраивать неявно. В большинстве случаев это означает, что серверы PCP на устройствах NAT и МСЭ с учётом состояния, поддерживающие запросы PEER и MAP могут быть защищены в этой модели угроз, если (1) все их хосты находятся в одном административном домене (или могут быть безопасно распределены по разным административным доменам, как в случае DS-Lite B4), (2) явные отображения создаются с таким же сроком действия, как у неявных и (3) опция THIRD_PARTY не поддерживается. Серверы PCP могут также безопасно поддерживать MAP Opcode в этой модели угроз, если политика безопасности устройства, где работает сервер PCP, разрешает независимую от конечных точек фильтрацию неявных отображений.

Серверы PCP, соответствующие простой модели угроз и не реализующие механизмов защиты PCP, описанных в параграфе 18.2, **должны** применять ограничения, описанные выше.

18.1.1 Предполагаемые атаки

- Если множеству административных доменов разрешено передавать запросы серверу PCP, который не разграничивает эти домены, узел одного домена может организовать атаку для отказа обслуживания других доменов или захватывать трафик, направленный узлу из другого домена.
- Если явное отображение действует дольше неявных, организовать DoS-атаку будет проще, чем при отсутствии сервера PCP.
- Если сервер PCP поддерживает удаление или сокращение срока действия имеющихся отображений, это позволяет атакующему захватить имеющееся отображение и получить трафик, направленный другому узлу.
- Если поддерживается опция THIRD_PARTY, атакующий имеет возможность открыть окно внешнему узлу для атаки на внутренний, может красть чужой трафик и организовывать DoS-атаки. Примером того, как опция THIRD_PARTY может расширить возможности атакующего по сравнению с поддельным неявным отображением является то, что сервер PCP (особенно при размещении в сети сервис-провайдера) может не знать о внутренней фильтрации (например, между гостевой и корпоративной сетью), которая предотвратит подделку эквивалентного неявного отображения.
- Если опция MAP Opcode поддерживается сервером PCP в тех случаях, когда политика безопасности не поддерживает независимой от конечных точек фильтрации неявных отображений, MAP Opcode меняет защитные свойства устройства, на котором работает сервер PCP, разрешая явные отображения, нарушающие политику безопасности.

18.1.2 Примеры развёртывания для простой модели угроз

В этом параграфе приведены два примера поддержки Simple Threat Model в реальных системах.

18.1.2.1 Развёртывание домашнего шлюза

Аналогию со многими современными домашними шлюзами может обеспечить использование сервера PCP с ограничениями, описанными в параграфе 18.1.

18.2 Расширенная модель угроз

В расширенной модели угроз (Advanced Threat Model) протокол PCP гарантирует, что злоумышленники (на пути или вне его) не смогут создать несанкционированных отображений или несанкционированно изменить имеющиеся отображения. Протокол также должен препятствовать атакующим в пути или вне его организовать DoS-атаку.

Расширенная модель безопасности требуется в перечисленных ниже случаях.

- Оборудование инфраструктуры защиты (скажем, корпоративные МСЭ), не создающее неявных отображений.
- Оборудование (такое как CGN и провайдерские МСЭ), обслуживает множество административных доменов и не имеет механизмов безопасного разделения трафика между этими доменами.
- Реализация хочет быть более терпимой при разрешении явных отображений по сравнению с неявными.
- Реализация хочет поддерживать вариант развёртывания, не соответствующий ограничениями параграфа 18.1.

Для защиты от атак в этой модели угроз должен быть задан механизм защиты PCP, обеспечивающий аутентифицированный канал сигнализации с защитой целостности.

Серверы PCP, реализующие механизм защиты PCP, **могут** не проверять подлинность запросов. В принятой по умолчанию конфигурации сервер PCP, реализующий механизм защиты PCP, **должен** применять ограничения, заданные в параграфе 18.1, при обработке запросов без аутентификации.

18.3 Остаточные угрозы

В этом параграфе рассмотрены некоторые угрозы, которые не устранены в описанных выше моделях безопасности, и даны рекомендации по смягчению угроз.

18.3.1 Отказ в обслуживании

Состояние, создаваемое в NAT или МСЭ, будут вероятно вносить ограничения (квоты) на уровне хоста или абонента для явных и неявных динамических отображений. Хост может создавать чрезмерное число явных или неявных динамических отображений, потребляя много портов, что препятствует обслуживанию других хостов. Поэтому параграф 17.2 рекомендует ограничивать число явных динамических отображений разумным значением.

Атакующий на пути между клиентом и сервером PCP может отбрасывать запросы и/или отклики PCP, или вводить фиктивные ошибки PCP - все это фактически создаёт DoS-атаку. Из-за таких действий клиент PCP может не узнать, что сервер PCP в действительности обработал запрос PCP. Передавая ошибку NO_RESOURCES, атакующий может вынудить клиента PCP некоторое время не передавать сообщений этому серверу. Для таких атак нет путей смягчения.

18.3.2 Фильтрация на входе

Важно не допустить создания, удаления или обновления (или фильтрации) мошенническим узлом для другого хоста, поскольку это может открыть хост-жертву для нежелательного трафика или препятствовать доставке нужного, а также расходовать квоту отображений для этого хоста. Явные и неявные динамические отображения создаются на основе IP-адреса отправителя в пакете, поэтому они зависят от входной фильтрации для защиты от атак с фиктивными адресами отправителей.

18.3.3 Захват отображения

В интервале между потерей состояния сервером PCP и получением значения Epoch Time меньше ожидаемого клиентом PCP возможен захват отображений клиента PCP другим хостом (с помощью явного или неявного динамического отображения). Это означает передачу входящего трафика другому хосту («кража»). Быстрое восстановление уменьшает этот интервал, но не предотвращает кражи совсем. Клиент PCP может уменьшит интервал, используя сравнительно короткий срок действия, однако это повышает объем трафика PCP. Данная угроза снижается использованием постоянного хранилища явных динамических отображения на сервере PCP (в результате не будет теряться состояние явных динамических отображений) или за счёт предотвращения назначения прежнего внешнего адреса IP, протокола и порта другому хосту (например, путём использования другого пула адресов IP).

18.3.4 Атаки на обнаружение серверов

Этот документ не задаёт обнаружения сервера кроме связи с принятым по умолчанию шлюзом.

19 Согласование с IANA

Агентство IANA выполняет указанные ниже действия.

19.1 Номер порта

PCP использует порты 5350 и 5351, ранее выделенные IANA для NAT-PMP [RFC6886]. Сейчас эти порты переназначены PCP.

19.2 Коды операций

Агентство IANA создало реестр PCP Opcode с номерами 0-127 и начальными значениями, указанными в таблице.

Значение	***Opcode***	
0	ANNOUNCE	
1	MAP	
2	PEER	
3 - 31	Standards Action [RFC5226]	Стандартизация
32 - 63	Specification Required [RFC5226]	Нужна спецификация
96 - 126	Reserved for Private Use [RFC5226]	Резерв для приватного использования
127	Reserved, Standards Action [RFC5226]	Резерв для стандартизации

Значение 127 зарезервировано и может быть назначено по процедуре Standards Action [RFC5226]. Значения 3-31 могут быть назначены по процедуре Standards Action [RFC5226], 32-63 - по процедуре Specification Required [RFC5226], а 96-126 предназначены для приватного использования (Private Use) [RFC5226].

19.3 Коды результатов

Агентство IANA создало новый реестр для кодов результата PCP со значениями 0-255, начальные значения которого взяты из параграфа 7.4. Значение 255 зарезервировано и может быть назначено по процедуре Standards Action [RFC5226].

Значения 14-127 могут быть назначены по процедуре Standards Action [RFC5226], 128-191 - по процедуре Specification Required [RFC5226], а 191-254 предназначены для приватного использования [RFC5226].

19.4 Опции

Агентство IANA создало новый реестр для опций PCP со значениями 0-255, для каждого из которых имеется мнемоническое обозначение. Значения 0-127 обязательны для обработки, а 128-255 необязательны. Начальные значения опция для реестра взяты из раздела 13. Коды 0, 127 и 255 зарезервированы и могут быть назначены по процедуре Standards Action [RFC5226].

Дополнительные коды опций PCP 4-63 и 128-191 могут быть назначены по процедуре Standards Action [RFC5226], 64-95 и 192-223 - по процедуре Specification Required [RFC5226], а 96-126 и 224-254 оставлены для приватного

использования [RFC5226].

В документах, описывающих опции, следует указывать их обработку клиентом и сервером PCP, а также перечисленную ниже информацию.

Имя опции: <мнемоника>

Номер: <число>

Назначение: <текстовое описание>

Пригодность для Opcode: <список Opcode>

Размер: <правила для размера опции>

Может включаться в: <запрос/отклик/оба>

Максимальное число экземпляров: <значение>

20 Благодарности

Спасибо Xiaohong Deng, Alain Durand, Christian Jacquenet, Jacni Qin, Simon Perreault, James Yu, Tina TSOU (Ting ZOU), Felipe Miranda Costa, James Woodyatt, Dave Thaler, Masataka Ohta, Vijay K. Gurbani, Loa Andersson, Richard Barnes, Russ Housley, Adrian Farrel, Pete Resnick, Pasi Sarolahti, Robert Sparks, Wesley Eddy, Dan Harkins, Peter Saint-Andre, Stephen Farrell, Ralph Droms, Felipe Miranda Costa, Amit Jain и Wim Henderickx за их комментарии и рецензии.

Спасибо Simon Perreault за то, что он выделил взаимодействие динамических соединений с созданными PCP отображениями, а также за множество других комментариев.

Спасибо Francis Dupont за тщательное рецензирование спецификации, которое существенно улучшило протокол.

Спасибо T. S. Ranganathan за диаграмму состояний.

Спасибо Peter Lothberg за информацию о сдвиге часов, которая помогла выбрать уровни допусков при решении вопроса об аномальных значениях времени Epoch.

Спасибо Margaret Wasserman и Sam Hartman за написание раздела «Вопросы безопасности».

Спасибо авторам DHCPv6 за текст о повторах передачи.

21 Литература

21.1 Нормативные документы

[RFC0768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, August 1980.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, May 2000.

[RFC4086] Eastlake, D., Schiller, J., and S. Crocker, "Randomness Requirements for Security", BCP 106, RFC 4086, June 2005.

[RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, October 2005.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

[RFC6056] Larsen, M. and F. Gont, "Recommendations for Transport-Protocol Port Randomization", BCP 156, RFC 6056, January 2011.

[proto_numbers] IANA, "Protocol Numbers", 2011, <http://www.iana.org/assignments/protocol-numbers>.

21.2 Дополнительная литература

[IGDv1] UPnP Gateway Committee, "WANIPConnection:1", November 2001, <http://upnp.org/specs/gw/UPnP-gw-WANIPConnection-v1-Service.pdf>.

[L2NAT] Miles, D. and M. Townsley, "Layer2-Aware NAT", Work in Progress, March 2009.

[PCP-FAIL] Boucadair, M., Dupont, F., and R. Penno, "Port Control Protocol (PCP) Failure Scenarios", Work in Progress, August 2012.

[PNP-IGD-PCP] Boucadair, M., Penno, R., and D. Wing, "Universal Plug and Play (UPnP) Internet Gateway Device (IGD)-Port Control Protocol (PCP) Interworking Function", Work in Progress[1], December 2012.

[RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981.

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

[RFC2136] Vixie, P., Thomson, S., Rekhter, Y., and J. Bound, "Dynamic Updates in the Domain Name System (DNS UPDATE)", RFC 2136, April 1997.

[RFC3007] Wellington, B., "Secure Domain Name System (DNS) Dynamic Update", RFC 3007, November 2000.

[1]Работа опубликована в RFC 6970. *Прим. перев.*

[RFC3022] Srisuresh, P. and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", RFC 3022, January 2001.

[RFC3581] Rosenberg, J. and H. Schulzrinne, "An Extension to the Session Initiation Protocol (SIP) for Symmetric Response Routing", RFC 3581, August 2003.

[RFC3587] Hinden, R., Deering, S., and E. Nordmark, "IPv6 Global Unicast Address Format", RFC 3587, August 2003.

[RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)", RFC 4303, December 2005.

[RFC4340] Kohler, E., Handley, M., and S. Floyd, "Datagram Congestion Control Protocol (DCCP)", RFC 4340, March 2006.

[RFC4787] Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", BCP 127, RFC 4787, January 2007.

[RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 4941, September 2007.

[RFC4960] Stewart, R., "Stream Control Transmission Protocol", RFC 4960, September 2007.

[RFC4961] Wing, D., "Symmetric RTP / RTP Control Protocol (RTCP)", BCP 131, RFC 4961, July 2007.

[RFC5382] Guha, S., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, "NAT Behavioral Requirements for TCP", BCP 142, RFC 5382, October 2008.

[RFC6092] Woodyatt, J., "Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service", RFC 6092, January 2011.

[RFC6145] Li, X., Bao, C., and F. Baker, "IP/ICMP Translation Algorithm", RFC 6145, April 2011.

[RFC6146] Bagnulo, M., Matthews, P., and I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", RFC 6146, April 2011.

[RFC6296] Wasserman, M. and F. Baker, "IPv6-to-IPv6 Network Prefix Translation", RFC 6296, June 2011.

[RFC6333] Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", RFC 6333, August 2011.

[RFC6619] Arkko, J., Eggert, L., and M. Townsley, "Scalable Operation of Address Translators with Per-Interface Bindings", RFC 6619, June 2012.

[RFC6763] Cheshire, S. and M. Krochmal, "DNS-Based Service Discovery", RFC 6763, February 2013.

[RFC6886] Cheshire, S. and M. Krochmal, "NAT Port Mapping Protocol (NAT-PMP)", RFC 6886, April 2013.

[RFC6888] Perreault, S., Ed., Yamagata, I., Miyakawa, S., Nakagawa, A., and H. Ashida, "Common Requirements for Carrier-Grade NATs (CGNs)", BCP 127, RFC 6888, April 2013.

[SCTPNAT] Stewart, R., Tuexen, M., and I. Ruengeler, "Stream Control Transmission Protocol (SCTP) Network Address Translation", Work in Progress, February 2013.

Приложение A Переход от NAT-PMP

Протокол PCP является преемником протокола отображения портов NAT-PMP[1] [RFC6886] и использует близкую семантику, концепции и форматы пакетов. Протоколы NAT-PMP и PCP работают через одинаковые порты и применяют средства согласования версий NAT-PMP и PCP для определения используемой версии. В этом приложении описано, как можно реализовать упорядоченный переход от NAT-PMP к PCP.

Клиенту, поддерживающему NAT-PMP и PCP, **следует** передавать свои запросы в формате PCP, которые будут приниматься сервером NAT-PMP или PCP. Сервер NAT-PMP будет возвращать отклик UNSUPP_VERSION, как указано в спецификации NAT-PMP [RFC6886], который заставит клиента вернуться к протоколу NAT-PMP и повторить запрос в формате NAT-PMP. Сервер PCP будет передавать отклик в соответствии с этим документом и обработка будет продолжаться как обычно.

Сервер PCP, поддерживающий одновременно NAT-PMP и PCP, может обслуживать запросы в любом формате. Первый октет в пакете указывает NAT-PMP (0) или PCP (не 0).

Шлюз, поддерживающий лишь PCP, при получении запроса NAT-PMP (0 в первом октете) будет считать этот запрос несоответствием версий. Обычная обработка PCP выдаёт отклики PCP, совместимые с NAT-PMP, без какой-либо специальной обработки на сервере PCP.

Адреса авторов

Dan Wing (editor)

Cisco Systems, Inc.

170 West Tasman Drive

San Jose, California 95134

USA

EMail: dwing@cisco.com

[1]NAT Port Mapping Protocol.

Stuart Cheshire

Apple Inc.

1 Infinite Loop

Cupertino, California 95014

USA

Phone: +1 408 974 3207

EMail: cheshire@apple.com

Mohamed Boucadair

France Telecom

Rennes 35000

France

EMail: mohamed.boucadair@orange.com

Reinaldo Penno

Cisco Systems, Inc.

170 West Tasman Drive

San Jose, California 95134

USA

EMail: repenno@cisco.com

Paul Selkirk

Internet Systems Consortium

950 Charter Street

Redwood City, California 94063

USA

EMail: pselkirk@isc.org

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru