

ICS 33.040.40

M 32

**YD**

# 中华人民共和国通信行业标准

YD/T 2934-2015

---

## 端口控制协议（PCP）技术要求

Technical specification for port control protocol (PCP)

2015-07-14 发布

2015-10-01 实施

---

中华人民共和国工业和信息化部 发布



# 目 次

前 言	II
1 范围	1
2 缩略语	1
3 概述	1
3.1 场景概述	1
3.2 PCP Server 和 NAT/防火墙之间的关系	3
4 PCP 原理和處理流程	3
4.1 PCP 原理	3
4.2 通用 Request 和 Response 消息头结构	4
4.3 PCP 处理流程	7
4.4 版本协商	11
4.5 MAP 和 PEER Opcodes 介绍	12
4.6 MAP Opcode 操作码	13
4.7 PEER Opcode	18
4.8 MAP Opcode 和 PEER Opcode 的选项 options	22
5 安全考虑	25
5.1 概述	25
5.2 简单威胁模型	25
5.2 高级威胁模型	25

## 前 言

本标准按照GB/T 1.1-2009 规则起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本标准由中国通信标准化协会提出并归口。

标准起草单位：中国电信集团公司、中兴通讯股份有限公司。

标准主要起草人：孙琼、孟伟、王翠。



## 端口控制协议（PCP）技术要求

### 1 范围

本标准规定了端口控制协议的技术要求，包括端口控制协议的概述、原理和安全考虑等。  
本标准适用于使用网络地址翻译（NAT）的应用场景。

### 2 缩略语

下列缩略语适用于本文件。

B4	Basic Bridging BroadBand	基本桥接宽带功能单元
CPE	customer premises equipment	用户驻地设备
CGN	Carrier-Grade NAT	运营级NAT
DS-Lite	Dual-Stack Lite	轻型双栈方案
EDF	Endpoint-Dependent Filter	端点相关过滤
EDM	Endpoint-Dependent Mapping	端点相关映射
EIF	Endpoint-Independent Filter	端点无关过滤
EIM	Endpoint-Independent Mapping	端点无关映射
NAT	Network Address Translation	网络地址翻译
NAT-PMT	NAT Path MTU	NAT映射表
NAT44	NAT IPv4 to IPv4	网络地址翻译IPv4到IPv4
NAT444	Carrier NAT	运营级NAT
NAT64	NAT IPv6 to IPv4	网络地址翻译IPv6到IPv4
P2P	Peer to Peer	点对点
PCP	Port Control Protocol	端口控制协议

### 3 概述

#### 3.1 场景描述

端口控制协议（PCP）提供了一种 NATs 和防火墙上动态创建映射条目从而转发入向包的机制，PCP 协议是和地址族无关的协议。它是 NAT-PMT 的扩展，可以满足 IPv6 和大规模网络地址转换（large-scale NAT）的需要。PCP 协议可以应用在 NATs（NAT44/NAT444/NAT64/CGN）设备和防火墙上。

图 1 所示是 DS-lite CGN 中的一种应用场景。

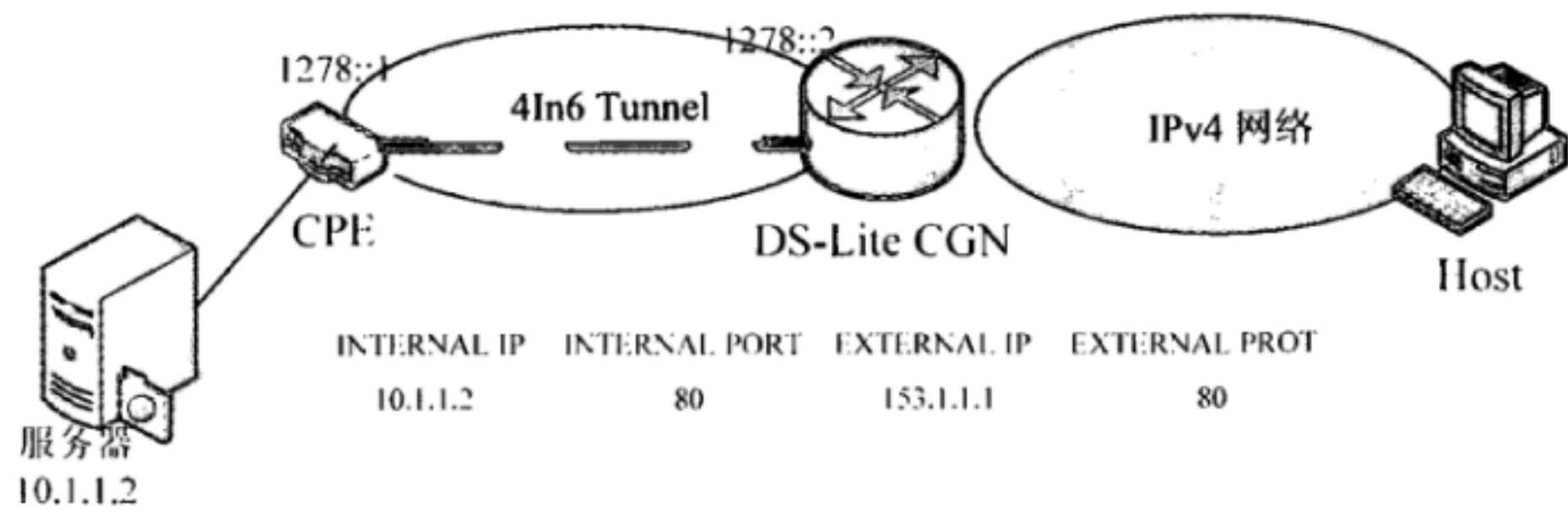


图 1 DS-lite CGN 中的应用场景

服务器 10.1.1.2 位于一个内网中，外部用户如果访问该服务器不能使用私网地址访问。在 CGN 设备设置一个内部地址、端口和外部地址、端口的映射表，把外部地址和外部端口提供给外部用户，则用户可以通过提供的外部地址和外部端口访问位于内网中的服务器。通常，这个映射表是通过手动配置的方式创建，PCP 协议则提供了一种动态创建映射表的机制。

服务器或者 CPE 作为 PCP 客户端（PCP client），CGN 作为 PCP 服务器端（PCP server），PCP client 向 PCP server 发送请求消息去创建、重建或者删除 PCP 映射，PCP server 向 PCP client 发送响应消息返回 PCP server 上生成的映射或者失败消息。

图 2 所示为 PCP 创建、更新、删除过程。

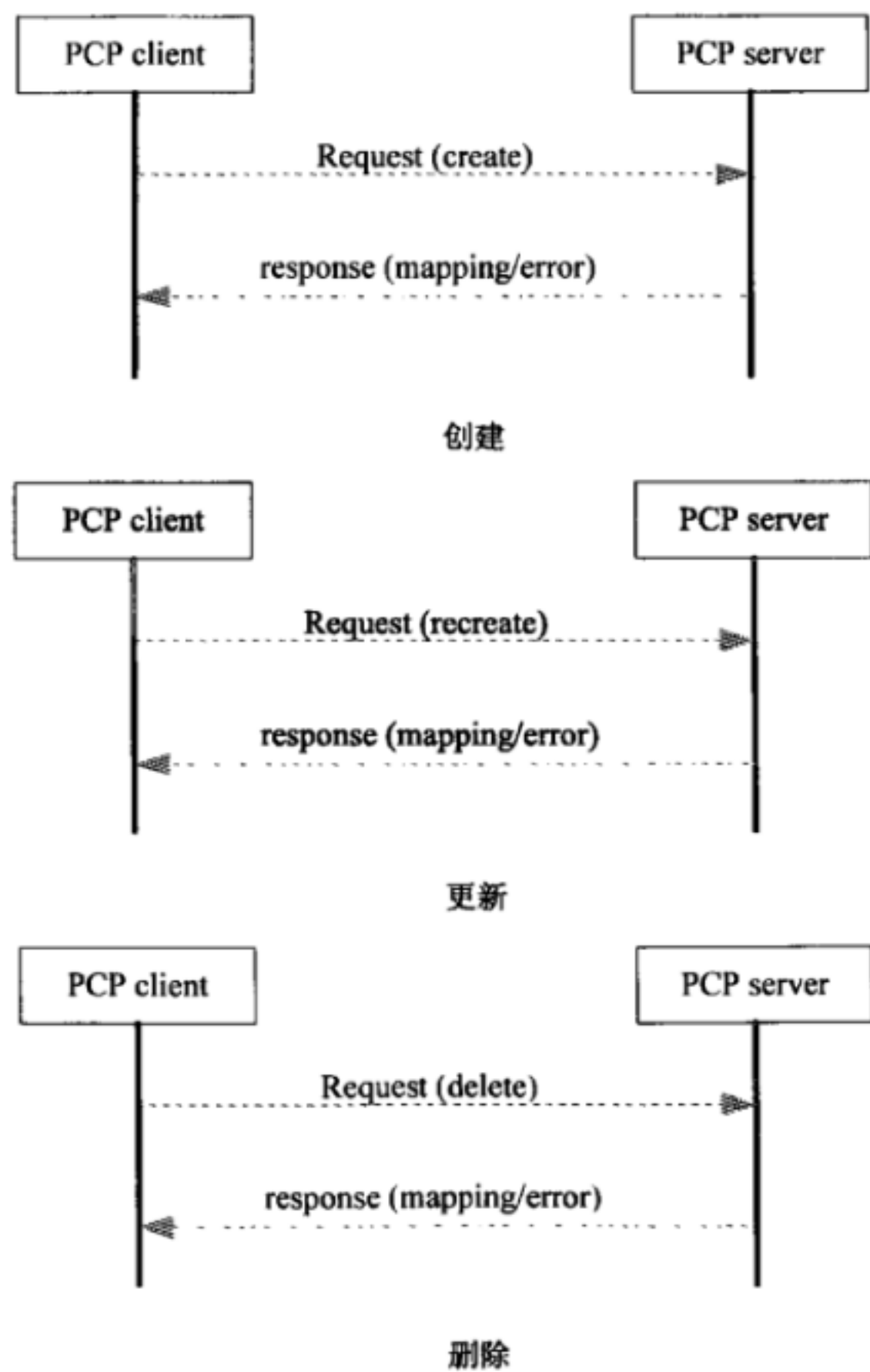


图 2 PCP 创建、更新、删除过程

PCP 除了上述场景外，基于 PCP client 和 sever 的交互过程，通过 PCP 的扩展，还有着非常广泛的应用场景。

在 NAT 应用中, B4 或 CGN 设备可以通过 PCP 扩展消息来向外部地址资源服务器获取和回收释放公网地址或端口, 用于进行公网地址和端口的转换。外部地址资源服务器 (即 PCP 服务器端) 对该请求消息做出响应, 对地址或端口进行分配或回收。例如在轻量级 4over6 技术中, B4 设备进行 NAT 转换并进行会话的维护, B4 作为 PCP 客户端, 并通过 PCP 的 MAP REQUEST 报文向 PCP 服务器获取公网地址和端口范围等信息。

在 NAT64 应用中, 当 IPv6 主机需要获取 IPv6 前缀来合成远端 IPv4 主机所转换的 IPv6 地址时, IPv6 主机作为 PCP 客户端, NAT 设备作为服务器, 通过 PCP 的 MAP REQUEST 报文, 即可将 IPv6 前缀 (Prefix64) 放在选项中带给 PCP 客户端也就是 IPv6 主机。

### 3.2 PCP Server 和 NAT/防火墙之间的关系

PCP Server 用来接收、处理和响应 PCP 请求。PCP Server 功能一般叠加在 NAT 或者防火墙设备上, 如图 3 所示。当然, PCP Server 也可以是独立于 NAT 或者防火墙的设备, 但是此时, PCP Server 和 NAT 或者防火墙之间需要有某种协议来交互映射信息。不管是叠加场景还是独立场景, 对于 PCP Client 来说没什么区别。

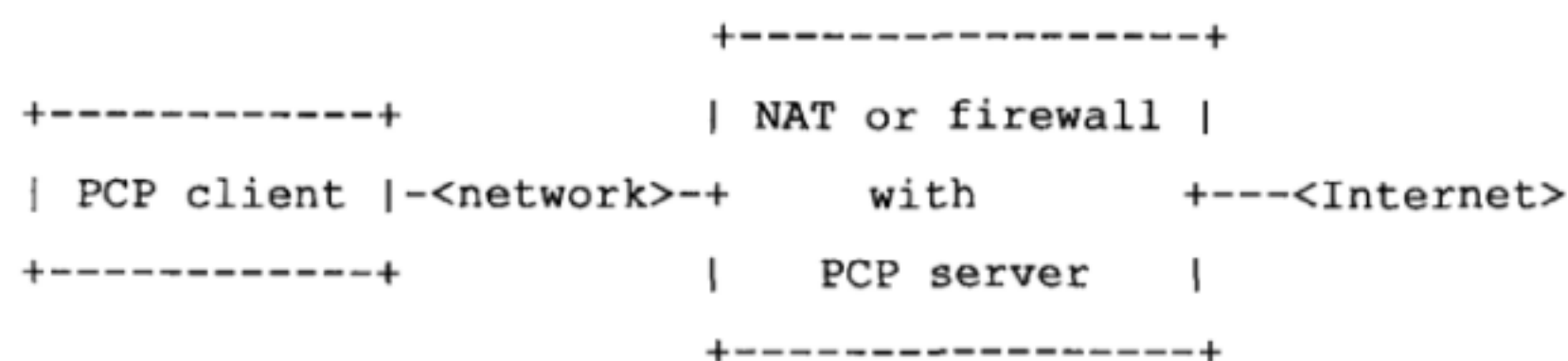


图 3 使能 PCP 的 NAT 或者防火墙设备

## 4 PCP 原理和处理流程

### 4.1 PCP 原理

PCP 是一种基于 UDP 的请求/响应 (Request/Response) 类的协议, 但是不是严格的一个请求紧跟着一个响应的格式。例如, PCP Client 发出的请求流可能因为包丢失等原因没有触发响应流; PCP Server 发出的响应流可能因为 Server 配置修改主动发送响应流, 而非响应之前的请求流。所以, PCP 协议是两种相反方向上的相对独立的消息流。

一种是 PCP Client 向 PCP Server 发送的请求流, 即 PCP Client 希望 PCP Server 上映射条目的状态; 一种是 PCP Server 向 PCP Client 发送的响应流, 即通知 PCP Client 在 PCP Server 上真实映射条目的状态; 两种消息流轻度关联。

当 PCP Client 发起 PCP 请求创建消息, PCP Server 判断该消息的合法性, 处理该请求消息, 创建映射条目, 并根据实际情况发送响应消息给 PCP Client。

当 PCP Client 发起 PCP 请求重新创建的消息, PCP Server 判断该消息的合法性, 处理该请求消息, 重新创建映射条目, 并根据实际情况发送响应消息给 PCP Client。

当 PCP Client 发起 PCP 请求删除消息, PCP Server 判断该消息的合法性, 处理该请求消息, 删除相应映射条目或所有映射条目, 并根据实际情况发送响应消息给 PCP Client。

当 PCP Client 发起 PCP 请求, 并等待重传时间  $T_1$ , 如果  $T_1$  时间内没有收到来自 PCP Server 的响应时, PCP Client 会重新发送请求消息 (retransmission), 并设置等待时间翻倍为  $T_1 \times 2$ 。这个过程再重复三次, 并且每次的重传时间都翻倍。当第四次后, 还没有收到响应, 则终止请求过程。



即 PCP Client 在呈指数级递增的时间间隔内等待来自 PCP Server 的响应，如果没有响应，继续发送重传直至终止请求；具体重传时间算法在后续章节中规定。

当 PCP Client 收到来自 PCP Server 的响应时，会保留该映射条目并设置该条目的生命时间，当到了 1/2 至 5/8 生命时间时，PCP Client 会向 PCP Server 发起续租（renewal），如果没有收到来自 PCP Server 的响应，则在 3/4 至 3/4+1/16 生命周期时再次发起续租，如果仍然没有收到，则在 7/8 至 7/8+1/32 生命周期时再次发起续租，以此类推，PCP Client 在呈指数级递减的时间上发送续租消息直至生命时间结束，这个过程的限制是，续租请求的间隔必须大于 4s。具体续租时间算法在后续章节中规定。

对于 PCP Server 来说重传消息和续租消息是一模一样的，都标识了 Client 希望创建映射以及维持映射生命时间。

4.2 通用 Request 和 Response 消息头结构

4.2.1 概述

所有的 PCP 消息都是 UDP 报文，最大报文长度为 1024 字节。PCP 消息包含 Request 消息头或者 Response 消息头，每个消息头中包括一个操作码 Opcode，Opcode 包含相关的 Opcode-specific 信息以及相关的 option 信息。下面分别介绍通用的 Request 和 Response 消息头结构，PCP Client 和 PCP Server 的处理流程、以及两种主要的 Opcodes 应用。

4.2.2 通用 Request 和 Response 消息头结构

所有的请求消息使用如图 4 所示头结构。

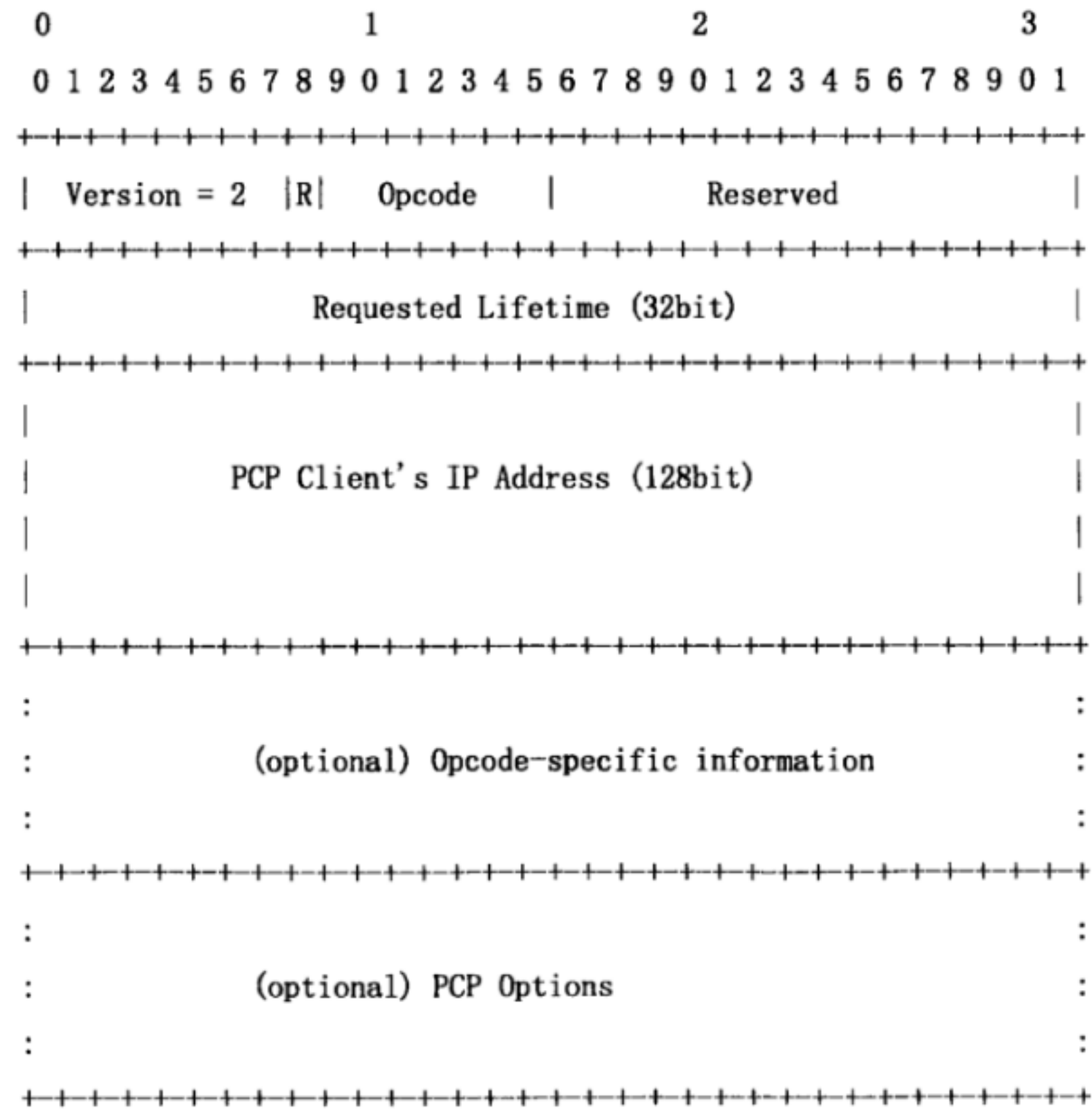


图 4 通用 Request 消息头格式

图 4 中各字段描述如下：

- Version: version 2, 这个字段用于 PCP Client 和 PCP Server 之间的版本协商;
- R: 0 表示 Request, 1 表示 Response;
- Opcode: 7bit, 用于标识是 MAP Opcode 还是 PEER Opcode;
- Reserved: 16bit 预留字段; 发送的时候该字段必须为 0, 接收时忽略不处理该字段;
- Request Lifetime: 32bit 整型, 单位是秒, 范围为  $0 \sim 2^{32}-1s$ , 用于标识映射条目的生命时间;
- PCP Client's IP Address: PCP 请求报文中 IP 头中的源 IPv4 地址或 IPv6 地址, IPv4 在这里标识为 IPv4-mapped IPv6 地址, 该字段用来检测 Client 和 Server 之间是否经过 NAT;
- Opcode-specific information: 对应 Opcode 的 payload 数据; 长度由操作码 Opcode 来决定;
- PCP Options: 对应该操作码 Opcode 的有效 option, 可能有 0 个、1 个或者多个。

#### 4.2.3 Response 消息头结构

所有的响应消息使用如图 5 所示的头结构。

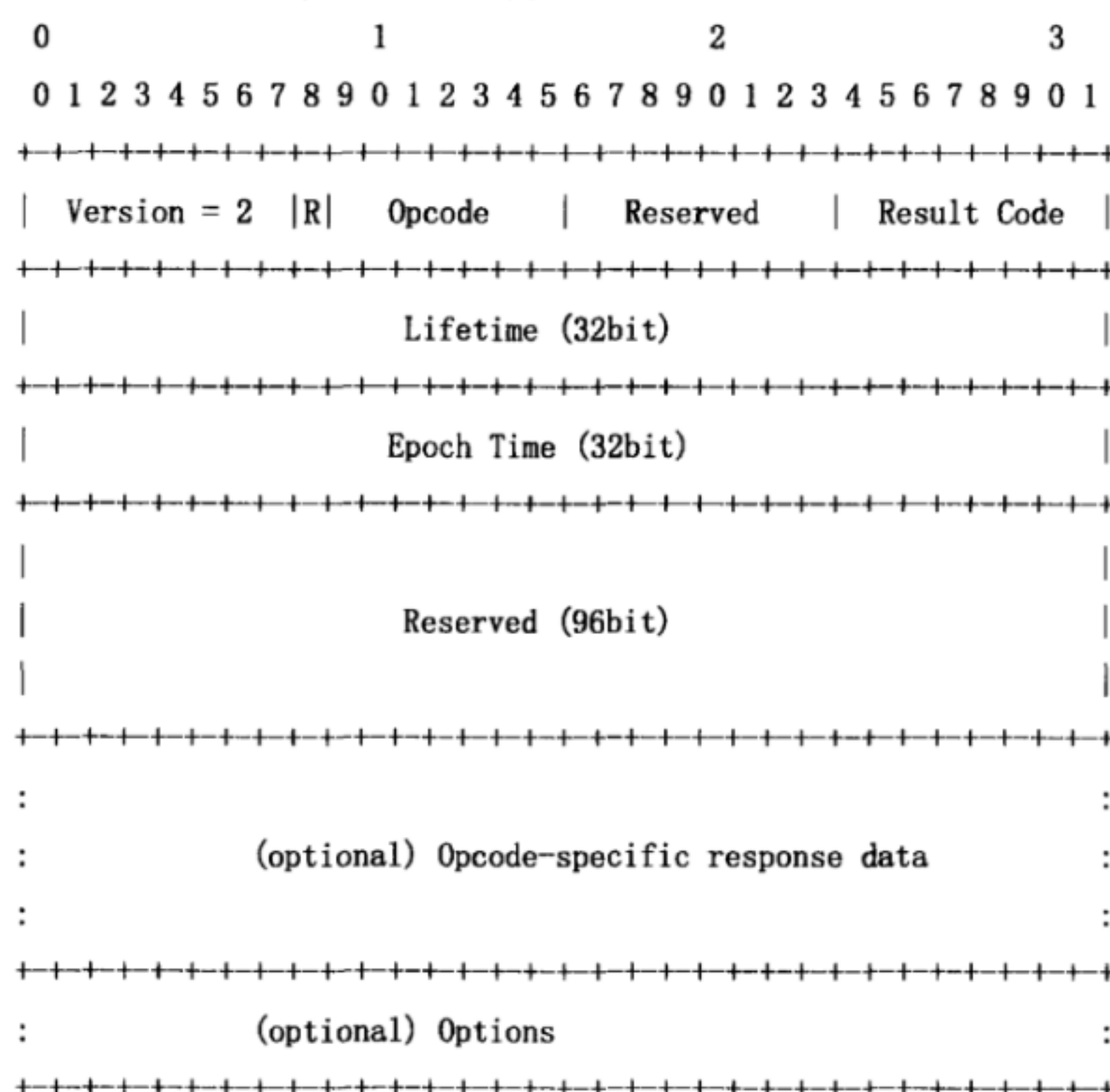


图 5 通用 Response 消息头结构

图 5 中各字段描述如下:

- Version: version 2, 这个字段用于 PCP Client 和 PCP Server 之间的版本协商;
- R: 0 表示 Request, 1 表示 Response; 所有的 Response 报文都必须为 1, 且由 Server 发送;
- Opcode: 7bit, 直接从 Request 消息对应的字段拷贝过来;
- Reserved: 8bit 预留字段; 发送的时候该字段必须为 0, 接收时忽略不处理该字段;
- Result Code: 8bit 响应的结果码, 具体见 4.2.5;
- Lifetime: 32bit 整型, 单位是 s, 范围为  $0 \sim 2^{32}-1s$ ; 在表示错误的 response 中, 该字段用来标识 PCP Client 上次和再次收到错误 response 之间的时间间隔, 在成功的 response 中, 该字段用于标识创建映射条目的生命时间;

- Epoch time: 由 PCP Server 设置的 Epoch 时间, 错误的 response 和成功的 response 中都包含该字段, 具体见 4.3.5;
- Reserved: 96bit 预留字段, 如果 Server 成功解析 Request 消息, 这个字段必须置 0; 如果 Server 未成功解析 Request 消息, PCP Server 拷贝 Request 消息中 PCP Client's IP Address 字段的后 96bit 到该字段作为响应;
- Opcode-specific information: 对应 Opcode 的 payload 数据; 长度由操作码 Opcode 来决定;
- PCP Options: 对应该操作码 Opcode 的有效 option, 可能有 0 个、1 个或者多个。

4.2.4 Options 选项

Option 是 PCP Opcode 的扩展部分, 其主要的 Type-Length-Value 格式如下图 6 所示。

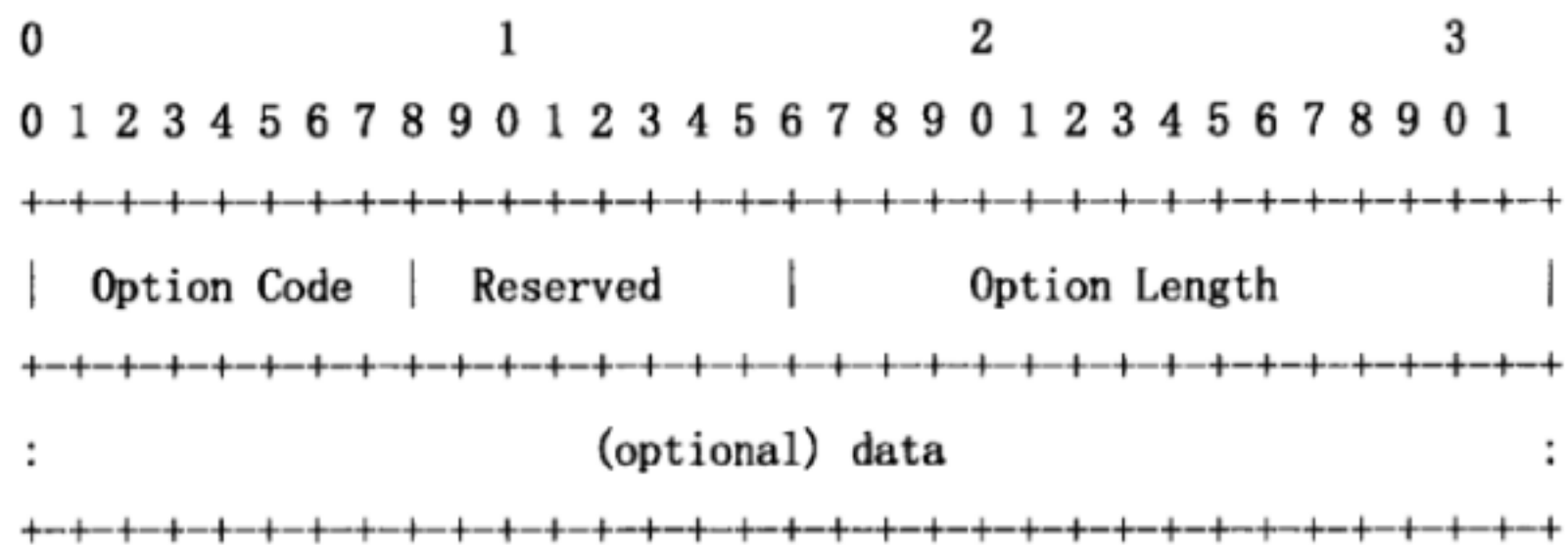


图 6 Option 头结构

图 6 中各字段描述如下:

- Option Code: 8bit, 其最高有效位置 0 标识了该 option 是强制的, 最高有效位置 1 标识了该 option 是可选的;
- Reserved: 8bit, 发送的时候该字段必须为 0, 接收时忽略不处理该字段;
- Option Length: 16bit, 标识后续 data 部分的长度, 以 8 位字节为单位; Option Length 可以为 0; 当 Option Length 不为 0 时, 必须是 4 字节的倍数, 不够 4 字节的后续以 0 填充;
- data: Option 的数据部分;
- 如果 PCP 请求消息中包括多个 options, PCP Server 收到报文时, 必须按照顺序逐个处理每个 option; PCP Client 必须保证 Server 有足够的空间来响应这个请求消息, 且 Server 响应消息不超过 1024bit, 如果响应消息超过 1024bit, Server 会响应一个错误消息;
- 当 PCP Server 在处理请求消息中的某个 option 时发生错误, PCP Server 会发送一个相应的错误响应消息; 比如, 当某个只允许出现一次的 option 在请求消息中出现了两次, PCP Server 必须发送响应报文携带 MALFORMED\_OPTION 结果码给 PCP Client; 当 PCP Server 接收到无效的 option (如 PCP option Length 比 UDP 包长还长), PCP Server 必须发送响应报文携带 MALFORMED\_OPTION 结果码给 PCP Client; 当 PCP Server 发现响应消息超过 1024bit, PCP Server 必须发送响应报文携带 MALFORMED\_REQUEST 结果码给 PCP Client;
- 所有成功解析并处理的 option 都包含在成功的响应消息中; 所有没有成功解析或处理的 option 都包含在错误的响应消息中;
- 对于不同的 Opcode 有不同的有效选项;
- THIRD\_PARTY Option 在 MAP Opcode 和 PEER Opcode 中都有效;
- FILTER Option 只在 MAP Opcode 中有效;



- PREFER\_FAILURE Option 只在 MAP Opcode 中有效。

#### 4.2.5 Result Codes 结果码

Result Codes 结果码枚举见下，只有 0 表示成功响应，其他都表示错误响应；当 PCP Server 在处理请求消息过程中遇到多个错误，PCP Server 应该使用最有代表性的错误码返回。每个错误的结果码划分为长生命时间的错误（long lifetime error）或短生命时间的错误（short lifetime error），长生命时间宜为 30min，短生命时间宜为 30s。

- 0 SUCCESS: 成功。
- 1 UPSUPP\_VERSION: 不支持的协议版本，属于长生命时间错误。
- 2 NOT\_AUTHORIZED: PCP Client 没有使能该请求操作，或者 PCP Client 请求的操作 PCP Server 因为安全策略等原因没办法执行，属于长生命时间错误。
- 3 MALFORMED\_REQUEST: PCP Server 不能成功解析该请求消息，属于长生命时间错误。
- 4 UNSUPP\_OPCODE: 不支持的 Opcode，属于长生命时间错误。
- 5 UNSUPP\_OPTION: 不支持的 Option，仅出现在某个 option 要求强制处理而 Server 不支持时，属于长生命时间错误。
- 6 MAIFORMED\_OPTION: 畸形的选项，如出现多次，或者无效长度等，属于长生命时间错误。
- 7 NETWORK\_FAILURE: PCP Server 正遭遇某种类型的网络故障，比如没有获取到外部 IP 地址等，属于短生命时间错误。
- 8 NO\_RESOURCES: 当前 PCP Server 上没有足够的资源完 PCP Client 请求的操作，比如，缺乏 CPU 资源或内存，或者其他临时性的状况；也许过段时间再发起相同的请求 PCP Server 能正确处理；这种属于系统性的错误，不同于 USER\_EX\_QUOTA，属于短生命时间错误。
- 9 UNSUPP\_PROTOCOL: 不支持的协议，属于长生命时间错误。
- 10 USER\_EX\_QUOTA: 试图创建一个新的映射条目但是超过了用户的端口配额，属于短生命时间错误。
- 11 CANNOT\_PROVIDE\_EXTERNAL: PCP Server 无法提供 PCP Client 建议的外部地址或者外部端口，这个错误限于处理 PEER 请求、带有 PREFER\_FAILURE Option 的 MAP 请求或者 SCTP 协议的 MAP 请求；属于哪种类型的生命时间错误依赖于发生故障的原因。
- 12 ADDRESS\_MISMATCH: 请求报文中的源 IP 地址和 PCP Client's IP Address 不匹配，属于长生命时间的错误。
- 13 EXCESSIVE\_REMOTE\_PEERS: PCP Server 无法创建过滤条目，只限于处理携带有 FILTER Option 的 MAP 请求，属于长生命时间错误。

### 4.3 PCP 处理流程

#### 4.3.1 PCP Client 产生和发送 Request 消息

PCP Client 发送第一个请求消息之前，需要先发现 PCP Server。发现 PCP Server 有以下几个步骤：

- 如果 PCP Server 已配置，例如，通过配置文件或者 DHCP 下发；则直接添加到 PCP Server 列表；否则

- 选择缺省路由添加到 PCP Server 列表；如果 PCP Client 既有 IPv4 地址又有 IPv6 地址，则选择 IPv4 缺省路由作为 IPv4 PCP Server 创建 IPv4 映射，选择 IPv6 缺省路由作为 IPv6 PCP Server 创建 IPv6 映射。

本标准中暂时只支持一个 PCP Server 地址。有了 PCP Server 地址，PCP Client 就可以发起 PCP 请求。

PCP 消息封装在 UDP 中，PCP Server 必须在 PCP Port 端口号上监听 PCP 请求。每个 PCP 请求都会产生一个 PCP 响应，所以 PCP 不用运行在一个可靠传输协议之上。图 7 所示是封装在以太帧中的 PCP 报文格式。

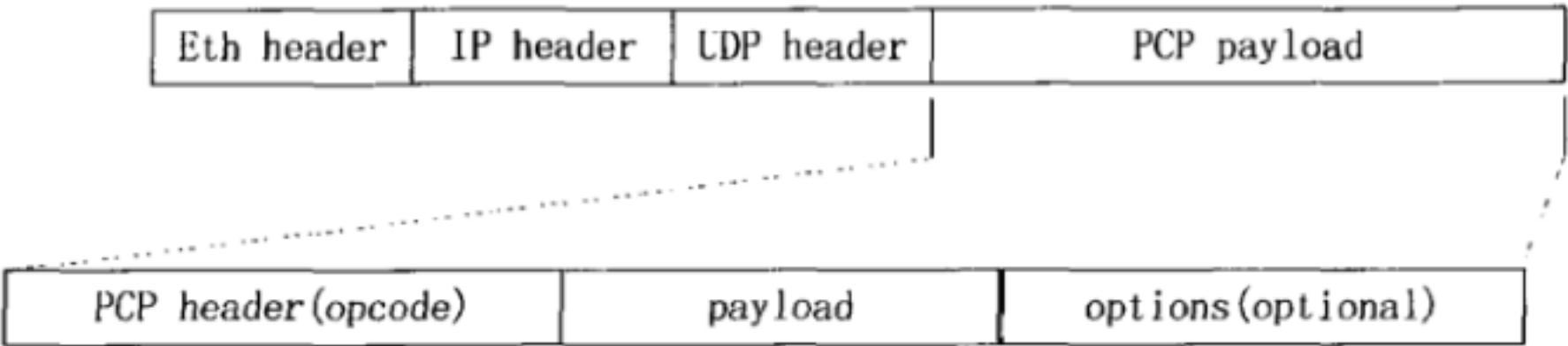


图 7 以太帧中的 PCP 报文格式

一个内部主机上可以有多个 PCP Clients，每个 Client 发起请求时使用不同的端口号，端口号是随机产生的，具体详见 IETF RFC6056。

### 4.3.2 PCP Client 重传 Request 消息

PCP Client 需要保证请求消息的正确传送。当 PCP Client 在期望时间内没有收到来自 PCP Server 的响应消息，PCP Client 必须重传这个请求消息，直到收到期望的响应消息。PCP Client 重传机制介绍如下，其中包括变量如下：

- RT: Retransmission timeout, 重传时间。
- IRT: Initial retransmission timeout, 初始化重传时间，可以为 3s。
- MRC: Maximum retransmission count, 最大重传次数，可以为 0（0 表示没有最大值）。
- MRT: Maximum retransmission time, 最大重传时间，可以为 1024s。
- MRD: Maximum retransmission duration, 最大重传持续时间，可以为 0（0 表示没有最大值）。
- RAND: Randomization factor, 随机因子。

每次请求消息发送或重传，PCP Client 设置 RT 时间如下规则，如果在等待了 RT 时间内仍然没有收到来自 PCP Server 的响应消息，则重新计算 RT 值并重传请求消息。

为了减少消息重传导致的消息同步，每次 RT 的计算都包含一个随机因子 RAND，RAND 的取值范围为 -0.1~+0.1。每个 PCP Client 在需要消息重传时都会随机选择一个 RAND 重新计算其各自的重传时间 RT。

RT 的初始化值如下：

$$RT = IRT + RAND \times IRT$$

下一次的重传时间基于上次重传时间，计算规则如下：

$$RT = 2 \times RT_{prev} + RAND \times RT_{prev}$$

最大重传时间 MRT 定义了重传时间 RT 的上限值，如果 MRT 值为 0，则表示没有最大重传时间，一直重传；当 MRT 值不为 0，当 RT > 最大重传时间 MRT 时，

$$RT = MRT + RAND \times MRT$$



最大重传次数 MRC 定义了重传次数的上限值，如果 MRC 值为 0，则表示没有最大重传次数，一直重传；当 MRC 值不为 0，则重传发生 MRC 次后不再发送重传。

最大重传持续时间 MRD 定义了重传时间的最大间隔，如果 MRD 值为 0，则表示没有最大重传持续时间，一直重传；当 MRD 值不为 0，则当从第一次发送请求到现在的时间超过这个最大重传持续时间 MRD 时，不再发生重传。

如果最大重传次数 MRC 和最大重传持续时间 MRD 都不为 0，其中任何一个变量达到限值都停止重传；如果最大重传次数 MRC 和最大重传持续时间 MRD 都为 0，PCP Client 继续发送重传直到收到来自 PCP Server 的响应消息。

如果 PCP Client 成功收到来自 PCP Server 的响应消息，则重新设置重传时间 RT 为初始化值；继续发送随后的 PCP 请求消息到 PCP Server。

#### 4.3.3 PCP Server 处理 Request 并响应 Response 消息

当收到一个 PCP 请求消息，PCP Server 解析并验证这个消息。一个有效的消息包括一个有效的 PCP Common Header，一个有效的 PCP Opcode，和 0 个或多个选项 options。假如验证失败，产生一个 error 响应并发送给 PCP Client。Error 响应和 success 响应有相同的包结构。对于 PCP Server 支持的 Opcode，error 响应会从请求包中拷贝部分域到 error 响应中，以保证 PCP Client 能够识别出哪个请求消息产生的 error 响应；对于 PCP Server 不支持的 Opcode，PCP Server 产生结果码为 UNSUPP\_OPCODE 响应消息并直接拷贝 PCP header 以及 PCP payload 到响应消息中，而不对包进行任何解析。

PCP Server 所有的响应消息都包含和请求消息一样的 Opcode，同时设置 Rbit 为 1。

error 响应的结果码非 0，其 error 响应报文构造如下：

- 拷贝整个 PCP 请求包并填充 0 以使得响应报文的长度为 4 字节的倍数；
- 设置 Rbit 为 1；
- 设置结果码 Result Code；
- 设置生命时间 lifetime、Epoch 值和 Reserved 字段；
- 更新响应消息中的其他字段。

Success 响应的结果码为 0；其 success 响应报文构造如下：

- 拷贝 PCP 请求包中的高 4 个字节到响应报文；
- 设置 Rbit 为 1；
- 设置结果码为 0；
- 设置生命时间 lifetime、Epoch 值和 Reserved 字段；
- 如果有的话设置 opcode-specific response data 字段；
- 添加已处理的 options。

如果收到的 PCP 请求报文长度小于 2 个字节，PCP Server 丢弃该请求报文。

如果收到的 PCP 请求报文中 Rbit 为 1，PCP Server 丢弃该请求报文。

如果收到的 PCP 请求报文中的版本字段是不支持的版本号，PCP Server 响应 Result Code 为 UNSUPP\_VERSION 的 error 响应报文。

如果 PCP Server 过载, PCP Server 会丢弃掉部分请求报文, 后续或者等待消息的重传; 或者发送 NO\_RESOURCES error 响应。

如果收到的 PCP 请求超过 1024 字节, 且不是 4 字节的倍数, PCP Server 判断该报文无效并发送 MALFORMED\_REQUEST error 响应, 裁剪报文长度为 1024 字节。

如果收到的 PCP 请求报文中 IP 层的源 IP 地址和 PCP 报文中的 PCP Client's IP Address 字段不匹配, PCP Server 发送 ADDRESS\_MISMATCH error 响应报文; 这么做是为了检测和防止 PCP Client 和 PCP Server 之间存在一个 PCP-unware 的 NAT 设备。

#### 4.3.4 PCP Client 处理 Response 消息

当收到来自 PCP Server 的响应, PCP Client 首先判断报文中的源地址和源端口号是否是之前发送请求报文的 Server 地址, 如果不是, 丢弃该响应报文。

当收到来自 PCP Server 的响应报文的长度小于 4 字节时, PCP Client 丢弃该响应报文。

当收到来自 PCP Server 的响应报文的 Rbit 为 0 时, PCP Client 丢弃该响应报文。

当收到来自 PCP Server 的响应报文是 UNSUPP\_VERSION error 响应时, 处理细节见 4.4 节。

当收到来自 PCP Server 的响应报文长度小于 24 字节, 或者大于 1024 字节, 或者不是 4 字节的倍数时, PCP Client 判断该报文为无效, 丢弃该响应报文。

之后, 进一步判断响应报文中的 Opcode-specific data 字段和请求报文中的该字段, 如果成功; 再判断响应报文中的 Epoch 时间, 如果 Epoch 时间不在期望范围内, 则表明需要重新恢复 PCP Server 上的映射状态。

PCP Client 发出请求报文后, 可能会收到多个响应报文, 这个是允许的。比如, PCP Server 发送 success 响应报文给 PCP Client 后, PCP Server 上配置发生了变化, 会再次发送 NOT\_AUTHORIZED error 响应报文该 PCP Client, 此时, PCP Client 收到非请求的响应报文时, 按照异常报文默认丢弃掉。

当收到来自 PCP Server 的响应报文是 ADDRESS\_MISMATCH error 报文时, 表明 PCP Client 和 PCP Server 之间存在 PCP-unware 的 NAT 设备。

所有响应消息中都包含有生命时间 lifetime, 用来表示 PCP Server 上设置的相应的 PCP 请求的生命时间。

当收到来自 PCP Server 的响应报文中的结果码 Result Code 为 0 时, 表示请求成功;

当收到来自 PCP Server 的响应报文中的结果码 Result Code 不为 0 时, 表示请求失败; 如果 Result Code 是 NO\_RESOURCES, 则在生命时间 lifetime 以内, PCP Client 不应该再发生重传。

如果因为网络变化, PCP Client 发现了新的 PCP Server, 则 PCP Client 立即终止和之前 PCP Server 之间的消息交互, 立即开始和新的 PCP Server 之间的消息交互。

#### 4.3.5 Epoch

PCP Server 发送的每一个报文中都包含 Epoch time 字段, 当 PCP Server 准备好接收连接时, 必须设置 Epoch 时间为初始值, 一般为 0, 随着时间的递增 Epoch 值不停递增。一旦 PCP Server 由于重启、电力故障或者其他等原因, 丢失 PCP Server 上的显示动态映射条目的状态, 此时需要重新设置 Epoch 时间为 0, 重新计时。当 PCP Clients 发现收到的响应报文中包含无效的 Epoch 值时, PCP Clients 就认为该 PCP Server 上状态发生过丢失, 则立即发送续租 Renewal 请求报文以恢复 PCP Server



上的丢失的映射状态。类似的，当 PCP Server 上公网 IP 地址池发生变化，也会设置 Epoch 值为初始值（一般为 0）。一个 PCP Server 可以为所有 Clients 公用一个 Epoch 时间，也可以为每个 Client 单独使用一个 Epoch 时间，这个具体依赖于实现。

当 PCP Client 接收到来自 PCP Server 的响应报文，会对 Epoch 值进行有效性检查，具体检查过程如下：

- 如果是第一次接收到响应报文，则认为报文中的 Epoch 值是有效的；如果不是第一次，则：
  - a) 如果当前的 Epoch 时间值 `current_server_time` 小于上次响应报文中的 Epoch 时间值 `previous_server_time`，那么 PCP Client 认为本次 Epoch 值无效；否则
    - 1) PCP Client 计算 Client 上接收到两个响应消息之间的差值：
 
$$\text{client\_delta} = \text{current\_client\_time} - \text{previous\_client\_time}$$
    - 2) PCP Client 计算接收到两个响应消息里 Epoch 字段之间的差值：
 
$$\text{server\_delta} = \text{current\_server\_time} - \text{previous\_server\_time}$$
    - 3) 如果  $\text{client\_delta} + 2 < \text{server\_delta} - \text{server\_delta}/16$   
 或者  $\text{server\_delta} + 2 < \text{client\_delta} - \text{client\_delta}/16$ ，则 PCP Client 认为 Epoch 时间无效；否则，认为 Epoch 时间有效；
  - PCP Client 记录下当前时间，以供下次计算使用：
 
$$\text{previous\_client\_time} = \text{current\_client\_time}$$

$$\text{previous\_server\_time} = \text{current\_server\_time}$$

#### 4.4 版本协商

当前 PCP 请求消息使用版本号的值为 2，后续可能会定义新的消息格式，使用大于 2 的版本号，但是仍然需要支持版本号为 2 的消息类型。到时，如果收到 UNSUPP\_VERSION 的 error 响应，PCP Client 需要产生告警以告知用户使用的 Client 版本太老，需要更新以兼容新的 PCP Server。

如果后续的 PCP 版本大于 2，则版本协商过程如下：

- PCP Client 发送第一个请求消息时使用自身能支持的最高的版本号。
- 如果 PCP Server 支持这个版本号，正确响应。
- 如果 PCP Server 不支持这个版本号，则发送 UNSUPP\_VERSION error 响应报文，并携带自身支持的最接近请求报文中版本号的版本号；例如，如果 PCP Server 上支持的一序列大于 Client 请求的版本号，则返回这个版本号序列中最小值给 PCP Client；如果 PCP Server 上支持的一序列小于 Client 请求的版本号，则返回这个版本号序列中最大值给 PCP Client。
- 如果 PCP Client 收到来自 PCP Server 的携带有 PCP Server 支持版本号的 UNSUPP\_VERSION error 响应报文，且刚好 PCP Client 也支持这个版本号，后续发起的请求就使用这个版本号和 PCP Server 协商。
- 如果 PCP Client 收到来自 PCP Server 的携带有 PCP Server 支持版本号的 UNSUPP\_VERSION error 响应报文，但是 PCP Client 不支持这个版本号，则后续 PCP Client 会选择一个自身支持的且次大的版本号发起的请求，以此类推，直到有版本值协商成功，如果一直都没协商成功，直到 PCP Client 发送请求的版本号为 2 时，仍然协商失败的话，PCP Client 需要产生告警信息以告知用户 Client 上使用的版本太老，无法和 PCP Server 交互信息。

## 4.5 MAP 和 PEER Opcodes 介绍

### 4.5.1 MAP 和 PEER Opcode 的应用分类

本标准定义了 MAP 和 PEER Opcode 的四种应用，具体如下：

- 主机作为 server，希望接收入向流量建立入向连接；
- 主机同时作为 client 和 server，并使用相同的端口号；
- 主机作为 client，希望优化应用程序的保活流量；
- 主机作为 client，希望恢复 NAT 设备上丢失的映射状态。

不同于之前通用请求消息，MAP 和 PEER 请求消息包括一个 Suggested External IP Address 字段。很多 PCP-Controlled 设备，特别是 CGN，每个 CGN 设备上都有公网地址池，PCP MAP 和 PEER 请求消息可以指定地址池中的某个公网地址来创建动态映射。

当动态映射和静态映射并存时，PCP 请求报文中 Suggested External IP Address 为 0，而 PCP Server 上刚好有该请求报文内部 IP 地址和内部端口号对应的静态映射条目，则返回静态映射条目中的公网 IP 地址和公网端口号给该 PCP 请求。如果 PCP 请求报文中 Suggested External IP Address 字段不为 0，则不管 PCP Server 上是否有该请求报文内部 IP 地址和内部端口号对应的静态映射条目或者动态映射条目，均创建一条新的映射条目，映射条目的公网 IP 地址即为 Suggested External IP Address。如果 Suggested External IP Address 不在 PCP Server 的公网地址池范围内，则返回错误响应。

### 4.5.2 服务器端的操作

当主机作为某种应用程序的 Server，主机不会主动发起流量，而是在某个端口上侦听外部流量时。此时，对于需要穿越 NAT 设备或者防火墙的、来访问内部主机 Server 的外部流量，主机需要预先完成以下步骤：

- 第一步，在 NAT 设备或者防火墙上创建动态映射条目：内部主机 IP 地址+内部侦听端口号+协议，绑定外部公网 IP 地址+外部端口号+协议；
- 第二步，通过某种集中应用将该公网 IP 地址和外部端口号发布给集中服务器（比如 DNS/SIP 等）；
- 第三步，确保 NAT 设备或者防火墙上的过滤策略，不会过滤掉外部流量对内网服务器的访问。

通常，为了完成上述第一步，作为某种应用程序 server 的内部主机需要主动发起携带 MAP Opcode 的 PCP 请求消息，其中 Suggested External IP Address 字段根据需要填写全 0，或者填写需要的某个 External IP Address。PCP Server 对于 PCP 请求报文的处理和响应下面会详细介绍；PCP Client 对于 PCP Server 的响应消息的处理下面也会详细介绍。

### 4.5.3 对称的客户端/服务器端的操作

当主机同时作为某种应用程序的 Server 和 Client，并在同一个端口上侦听流量和发送流量时（例如 RTP 协议/SIP Symmetric Response Routing 协议），同上述步骤一样，作为某种应用程序 server 的内部主机需要主动发起携带 MAP Opcode 的 PCP 请求消息，在 NAT 设备或者防火墙上先创建映射条目，然后向集中服务器注册该映射条目对应的公网 IP 地址+外部端口号，最后接收来自该应用



程序 client 的外部流量。不同的是，该应用程序的某个 client 就是该内部主机，即本主机发起目的地址是该公网地址+外部端口号，源地址是该主机 IP 地址和相同端口号的流量。

#### 4.5.4 减少 NAT 或防火墙保活消息

当主机作为某个应用程序的 Client（比如 XMPP client/SIP client），该主机访问远端 Server 时会在经过的 NAT 设备或者防火墙上创建映射条目，这样来自 Server 的响应报文也能正确经过 NAT 设备或者防火墙到达内部主机。当内部主机 Client 暂时没有到外部 Server 的流量或者外部 Server 暂时没有到内部 Client 的流量时，NAT 设备或者防火墙上面的条目会老化，这样会导致后续的应用程序交互失败，为了防止这种情况的发生，很多应用程序会生成一种专门用于应用程序保持连接的保活流量，并周期性的发送，以保证 NAT 设备或者防火墙上映射条目不会老化。PCP 应用能够减少这种保活流量。

#### 4.5.5 TCP 动态映射的状态恢复

当 NAT 设备由于断电等原因丢失映射状态后，此时，如果 Clients 能主动发起请求在 NAT 设备上重新建立和已丢失映射一样的映射条目，这将大大提升网络的恢复能力，特别是对一些长时间的 TCP 连接或者需要有大量数据需要传送的 TCP 连接。PCP 中实现这种恢复能力如下：

- 第一步，主机先发起 TCP SYN 在 NAT 设备或者防火墙上建立隐式动态映射；
- 第二步，主机作为 PCP Client 再发送携带 Suggested External IP Address 为 0 的 PCP PEER 请求消息，作为 PCP Server 的 NAT 设备或者防火墙处理该消息并携带隐式动态映射条目对应的公网 IP 地址和端口信息响应给主机，此时，主机上记录下该映射条目；
- 第三步，当作为 PCP Server 的 NAT 设备上或者防火墙上因为断电等原因丢失映射条目时，主机作为 PCP Client 再发送携带有记录中 Suggested External IP Address 的 PCP PEER 请求报文到 PCP Server，PCP Server 上创建和第一步中一样的映射条目，从而恢复 TCP 连接继续发送 TCP 数据。

### 4.6 MAP Opcode 操作码

#### 4.6.1 概述

MAP Opcode: 创建内部 IP 地址+内部端口号和外部 IP 地址+外部端口号之间的显示动态映射关系。

PCP Server 上有对应的配置命令，用于使能/去使能 Server 上对于 MAP Opcode 请求的支持/不支持能力。

通过 PCP MAP Opcode 请求消息在 PCP Server 上创建的 NAT 条目的规则都是端点无关映射 EIM 和端点无关过滤 EIF（除非携带了 FILTER 选项），即使 NAT 设备上自身的配置是端点相关映射 EDM 和端点相关过滤 EDF。即对于没有携带 FILTER 选项的 PCP MAP Opcode 请求创建的映射条目接收外部任何端点的回归流量。

#### 4.6.2 MAP Opcode 包结构

MAP Opcode 包和通用请求包以及通用响应包有相似的结构。如果响应包中 Assigned External IP Address 和 Assigned External Port 和请求包中 Internal IP Address 和 Internal Port 一致，则说明 PCP Server 是防火墙，否则，说明 PCP Server 是个 NAT 设备。

MAP Opcode 请求报文中 Opcode-specific information 字段的格式如图 8 所示。

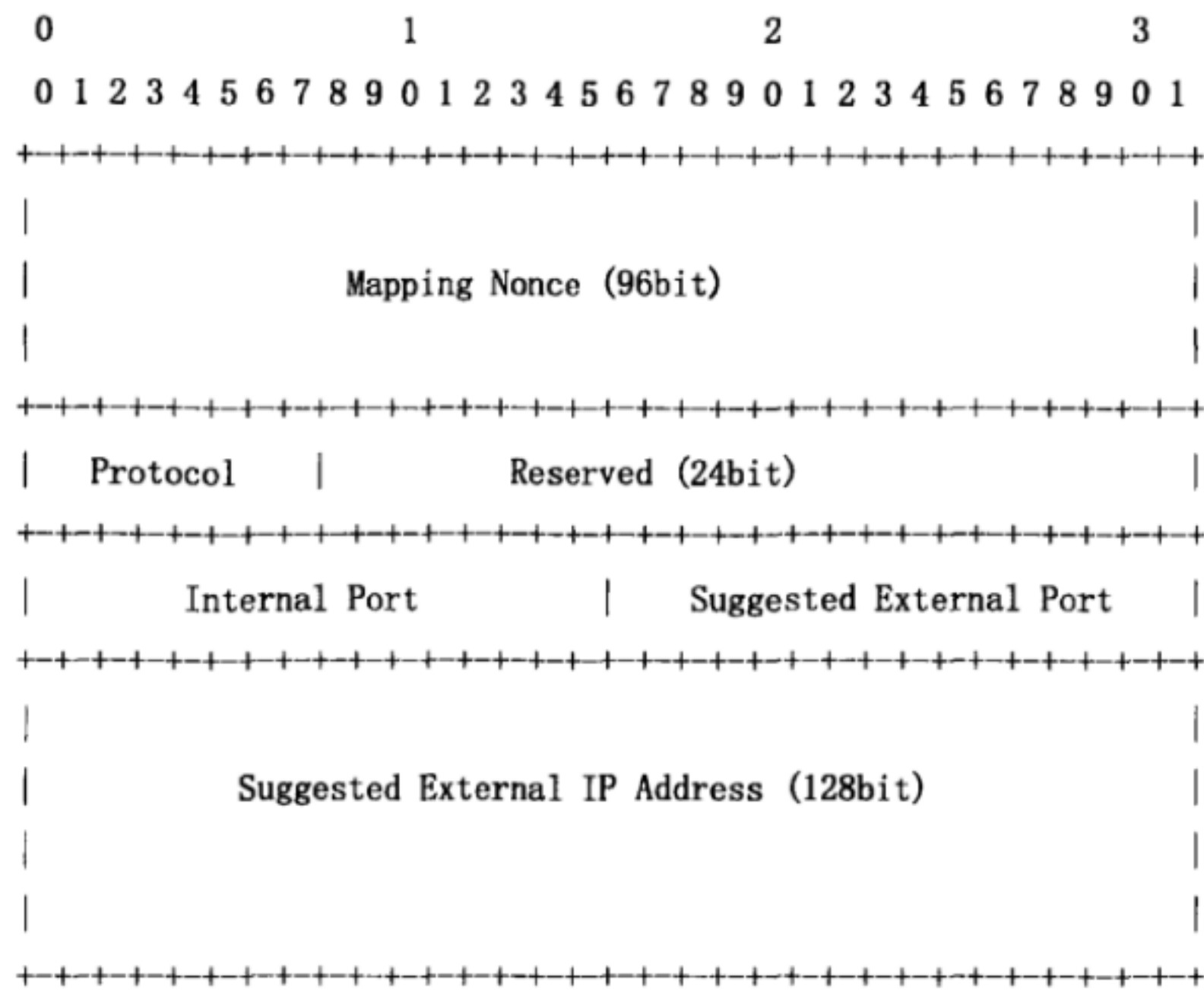


图 8 MAP Opcode 请求报文

各字段描述如下：

**Requested lifetime**（在通用请求报文头中）：请求的映射条目的生命时间，单位是 s，当值为 0 时表示删除该映射条目。

**Mapping Nonce**：PCP Client 在请求报文中填充的随机值字段。

**Protocol**：和 MAP Opcode 请求映射条目相关的协议类型；当值为 6 时表示请求创建一个 TCP 映射，当值为 0 时表示请求创建所有协议类型的映射条目，即创建的映射条目不拘泥于某种协议类型。

**Reserverd**：24bit 预留字段，发送时该字段必须设置为 0，接收时忽略该字段。

**Internal Port**：映射条目中的内部端口号；当值为 0 时表示所有端口，且当 Requested lifetime 为 0 时 Internal Port 为 0 也是合法的；当请求的协议类型报文中不需要 16bit 端口号时，该字段为 0。

**Suggested External Port**：指定的外部端口号；当 PCP Server 上丢失映射条目的状态时，PCP Client 可以发送指定外部端口号字段的请求报文到 PCP Server 请求更新丢失的映射条目；如果 PCP Client 不知道外部端口号信息，也没有偏好的外部端口号，则该字段必须为 0。

**Suggested External IP Address**：指定的外部 IPv4 地址或者 IPv6 地址；当 PCP Server 上丢失映射条目的状态时，PCP Client 可以发送指定外部 IPv4 地址或者 IPv6 地址的请求报文到 PCP Server 请求更新丢失的映射条目；如果 PCP Client 不知道外部 IP 地址信息，也没有偏好的外部 IP 地址，则该字段必须为 0。

请求报文的内部 IP 地址和请求报文 IP 头中的源 IP 地址一致，除非请求报文中携带了 THIRD\_PARTY 选项。

MAP Opcode 响应报文中 Opcode-specific information 字段的格式如图 9 所示。



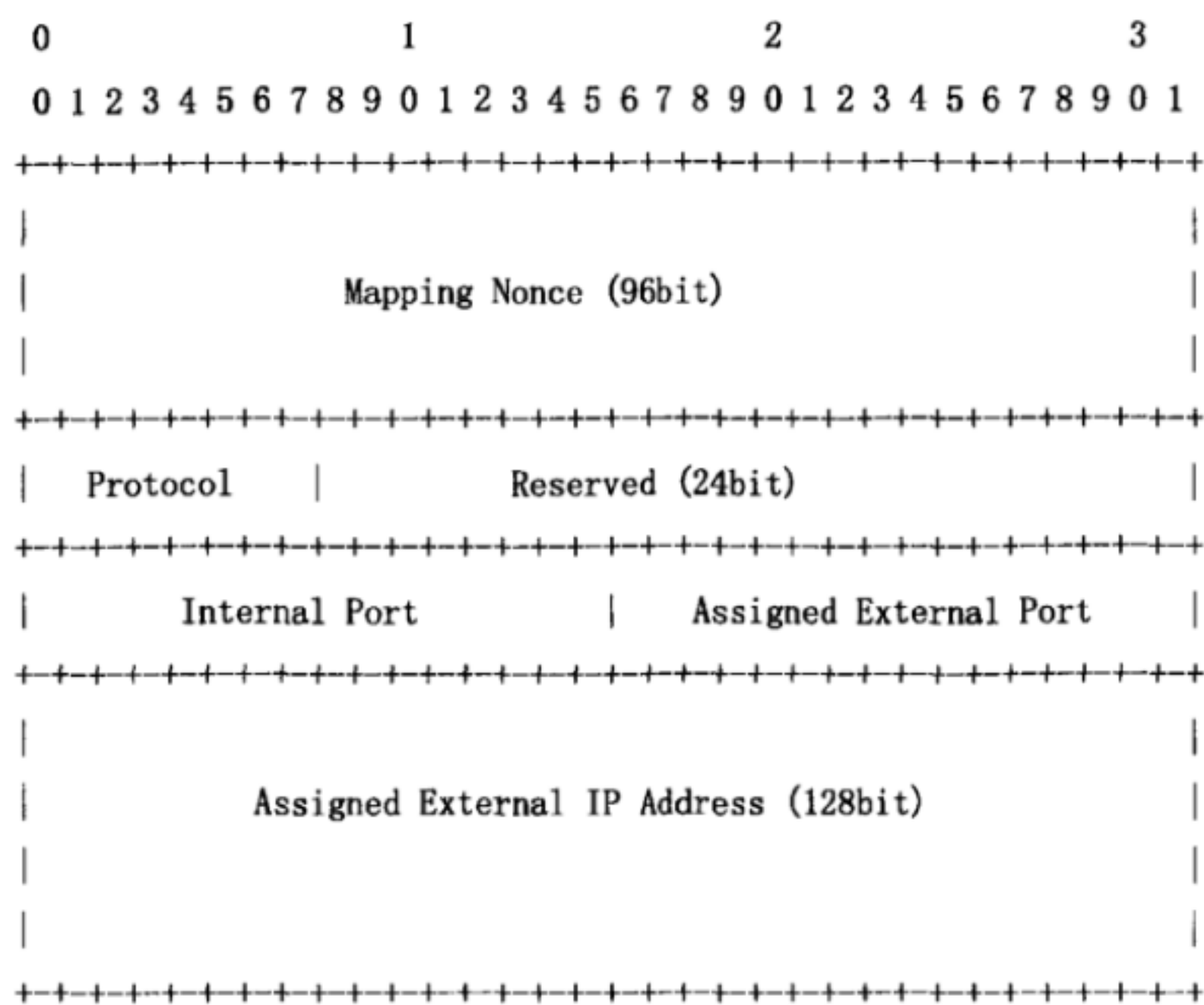


图 9 MAP Opcode 响应报文

各字段描述如下：

**Lifetime**(在通用请求报文头中)：在 **error** 响应报文中，该字段表示 PCP Client 再次发送请求报文需要等待的时间；在 **success** 响应报文中，该字段表示该映射条目的生命时间。

**Mapping Nonce**：随机值字段，直接从请求报文中拷贝获得。

**Protocol**：协议类型字段，直接从请求报文中拷贝获得。

**Reserved**：24bit 预留字段，发送时该字段必须设置为 0，接收时忽略该字段。

**Internal Port**：从请求报文中拷贝获得。

**Assigned External Port**：在 **success** 响应报文中，该字段表示映射条目生成的外部端口号；在 **error** 响应报文中，该字段直接从请求报文中拷贝获得。

**Assigned External IP Address**：在 **success** 响应报文中，该字段表示映射条目生成的外部 IPv4 地址或者 IPv6 地址，且使用 IPv4-mapped IPv6 地址标识 IPv4 地址；在 **error** 响应报文中，该字段直接从请求报文中拷贝获得。

4.6.3 PCP Client 产生和发送 MAP Request 消息

MAP 请求消息可能携带有具体参数值的 Suggested External IP Address 和 Suggested External Port 字段到 PCP Server，用于重建 PCP Server 上丢失的映射条目，并保持集中服务器上注册的外部信息不变。当然，也可能由于 PCP Server 上公网信息的变化导致 MAP 请求的条目无法正确创建，此时，PCP Server 上会重新创建一个新的映射条目并响应给 PCP Client，然后 PCP Client 通知集中服务器更新注册信息。

如果某种应用协议不需要使用 16bit 的端口字段，则设置端口字段为 0，这样，所有匹配该协议的流量都会被正确映射。

**Mapping Nonce**是PCP Client随机产生的，用来校验PCP响应报文的合法性。

4.6.4 PCP Client 续租映射的 Renew 消息

在映射条目老化之前，PCP Client 通过发送 Renew 消息来延长映射条目的生命时间。为了防止不必要的消息同步导致的丢包问题，PCP Client 考虑在续租时间上添加一个随机误差；推荐实现如下：

- 1/2 到 5/8 的生命时间内随机选择一个时间值，PCP Clients 发送第一个且仅发送一个 Renew 消息；如果没有收到 success 响应报文，则
- 3/4 到 3/4+1/16 的生命时间内再次发送 Renew 消息；如果没有收到 success 响应报文，则
- 7/8 到 7/8+1/32 的生命时间内再次发送 Renew 消息；如果没有收到 success 响应报文，则继续；
- 直到生命时间剩余不到 4s 时，不再发送 Renew 消息。

#### 4.6.5 PCP Server 处理 Request 并响应 Response 消息

PCP Server 接收到 MAP Opcode 请求报文时，处理细节如下：

PCP Server 拷贝请求报文中的协议字段、内部端口号字段和 Mapping Nonce 字段到 MAP 响应报文中；如果请求报文中存在 THIRD\_PARTY 选项，也拷贝该选项到 MAP 响应报文中。

如果 Requested Lifetime 不为 0，则接收到的请求报文是用来创建映射条目或者延长映射条目的生命时间；如果 PCP Server 不支持请求报文中的协议字段，则发送携带 UNSUPP\_PROTOCOLAL 的 error 响应报文给 PCP Client；如果 Requested Lifetime 不为 0，内部端口号为 0，协议号不为 0，则该请求表示在 PCP Server 上为该协议类型下的任何流量创建映射条目；如果 Requested Lifetime 不为 0，内部端口号为 0，协议号也为 0，则该请求表示在 PCP Server 上为任意协议类型下的任何流量创建映射条目；如果 Requested Lifetime 不为 0，内部端口号不为 0，协议号却为 0，则 PCP Server 携带 MALFORMED\_REQUEST 的 error 响应报文给 PCP Client。

如果 Requested Lifetime 为 0，则接收到的请求报文是用来删除某个映射条目或者删除某类映射条目的。

如果接收到 PCP Client 请求报文中某个选项的值是 128（表示必须处理该选项），但是该选项没有任何意义，比如携带 PREFER\_FAILURE 选项但是 Requested lifetime 为 0，此时，PCP Server 携带 MALFORMED\_REQUEST 的 error 响应报文给 PCP Client。

PCP Client 的 MAP 请求的映射条目在 PCP Server 上已经存在，此时，PCP Server 处理细节如下：

- 如果 MAP 请求报文中携带 PREFER\_FAILURE 选项，但是 MAP 请求报文中指定的外部地址和外部端口号和 PCP Server 上存在的映射条目不匹配，在 PCP Server 携带 CANNOT\_PROVIDE\_EXTERNAL 的 error 响应报文给 PCP Client；
- 如果 PCP Server 上存在的映射条目是静态的，则不管 MAP 请求报文中指定的外部 IP 地址和端口号，PCP Server 将静态条目的外部 IP 地址和外部端口号填写到 External Address 和 Port 字段发送给 PCP Client；并填写 Lifetime 为  $2^{32}-1$ ；
- 如果 PCP Server 上存在的映射条目是之前 MAP 请求创建的，则不管当前 MAP 请求报文中指定的外部 IP 地址和端口号，PCP Server 将已存在映射条目中的外部 IP 地址和外部端口号填写到 External Address 和 Port 字段发送给 PCP Client；并更新 Lifetime 字段；



– 如果 PCP Server 上存在的映射条目是之前流量创建的或者 PEER 请求创建的，则 PCP Server 上会重新创建一条新的 inbound 映射条目，inbound 条目中的端口信息和 IP 地址信息和已存在的 outbound 映射条目一致，且两个条目同时存在，直到新创建的 inbound 映射条目老化后 outbound 映射条目才继续生效。

PCP 协议占用端口号 5350 和 5351，故 PCP 映射条目不可以占用这两个端口号；当然，PCP Server 上也可以配置策略 deny 掉那些不支持的端口号，当 PCP Server 收到 MAP 请求报文携带 deny 范围内的端口号时，发送 NOT\_AUTHORIZED 的 error 响应报文给 PCP Client。

PCP Client 的 MAP 请求的映射条目如果在 PCP Server 上不存在，且 PCP Server 能够正确创建，则使用指定 IP 地址和端口号创建该映射条目，并发送 success 响应报文给 PCP Client；如果不能按照指定 IP 地址和端口号创建映射条目（MAP 请求报文中不携带 PREFER\_FAILURE 选项时），但是能分配新的 IP 地址和端口号给该映射条目，则创建新的映射条目，携带新的 IP 地址和端口号发送响应报文给 PCP Client。

对于 PCP Server 上无法分配指定 IP 地址和端口号的情况，总结如下：

- 指定 IP 地址和端口号已经分配给其他的映射条目；
- 指定 IP 地址和端口号已经被 NAT 设备用作他用；
- 指定 IP 地址和端口号无效，比如 127.0.0.1、组播地址、或者端口号和协议号不一致等；
- 指定的 IP 地址不在 NAT 地址池范围内；
- MAP 请求报文中携带 PREFER\_FAILURE 选项且指定 IP 地址和端口号是不能分配给该 Client 的地址和端口。

#### 4.6.6 PCP Client 处理 Response 消息

PCP Client 收到响应报文后，除了处理通用响应报文的步骤外，还需要比对响应报文中的 Internal IP Address 字段、Protocol 字段、Internal Port 字段和 Mapping Nonce 字段。

PCP Client 收到 success 响应报文时，记录下外部 IP 地址和端口号，并通过某种集中机制向集中服务器注册其对应的外部 IP 地址和端口信息，以实现后续和其他主机的通讯。

PCP Client 收到 error 响应报文时，在响应报文携带的 lifetime 时间内，PCP Client 不允许发送相同的 MAP 请求报文给 PCP Server。

当续租时间到，PCP Client 需要按照 4.6.4 所述，发送续租报文。

#### 4.6.7 映射的生存时间和映射的删除

PCP MAP 请求某个生命时间，PCP Server 可能会响应相应的生命时间，也可能会响应相对较短或者相对较长的生命时间。对于最大生命时间和最小生命时间，PCP Server 上应该可配；推荐最大生命时间为 24h，最小生命时间为 120s；当然具体实现时，需要根据实际场景去权衡出合适的最大最小值。

一旦 PCP Server 处理并携带生命时间 L1 响应了一个 PCP MAP 请求报文，则对应创建的映射条目的生命时间便是 L1，除非 PCP Client 再次发送 MAP 请求报文更新映射条目的生命时间，或者 PCP Server 上由于某种原因丢失了映射条目的状态。

当接收到 MAP 或者 PEER 响应报文中携带异常大的生命时间时, PCP Client 按照正常生命时间 (比如 24h) 进行处理, 并按照正常情况适时续租。

如果 Requested lifetime 为 0, 则:

- 如果内部端口号和协议均不为 0, 则该请求报文表示删除对应端口号和协议等的映射条目;
- 如果内部端口号和协议均为 0, 则该请求报文表示删除对应内部 IP 地址的所有映射条目; 当主机重启或者 IP 地址发生变化时, 这种请求消息有助于删除 NAT 设备或者防火墙上的之前创建的和原内部 IP 地址相关的所有映射条目;
- 如果内部端口号为 0, 协议号不为 0, 则该请求报文表示删除对应该协议的任何端口号的映射条目;
- 如果内部端口号不为 0, 协议号为 0, 则该请求报文无效, 返回携带 MALFORMED\_REQUEST 的 error 响应报文给 PCP Client。

如果 Requested lifetime 为 0, 则发送该请求报文时, Suggested External IP Address 和 Suggested External Port 字段必须为 0, PCP Server 上接收到该请求报文时, 忽略该字段。

PCP PEER 请求不支持删除映射条目, 也不支持缩短映射条目的生命时间; PCP Server 上接收到 PEER 请求, 且 Requested lifetime 小于当前存在的映射条目的生命时间, 则 PCP Server 发送携带当前存在映射条目的生命时间的 success 响应报文给 PCP Client。

PCP MAP 请求支持删除映射条目, 也支持缩短映射条目的生命时间; 如果一个 PCP Client 的 MAP 请求试图删除一个静态映射条目, 或者试图删除一个隐式动态映射条目, 或者试图删除一个 PEER 请求创建的映射条目, 则 PCP Server 携带 NOT\_AUTHORIZED 的 error 响应报文给 PCP Client; 如果 PCP Client 的 MAP 请求试图删除一个不存在的映射条目, PCP Server 发送 success 响应报文给 PCP Client。

如果删除请求正确处理并 success 响应, 则响应报文中 lifetime 字段为 0, 并拷贝请求报文中的协议字段和内部端口字段到响应报文中。

## 4.7 PEER Opcode

### 4.7.1 概述

PEER Opcode: 创建一个到远端 IP 地址和端口的动态 outbound 映射, 或者延长一个已经存在的 outbound 映射条目的生命时间。

PCP Server 上有对应的配置命令, 用于使能/去使能 Server 上对于 PEER Opcode 请求的支持/不支持能力。

通过 PCP PEER Opcode 请求消息在 PCP Server 上创建的 NAT 条目的规则和隐式动态映射创建的 NAT 条目规则一致, 即和 NAT 设备或者防火墙上的配置规则一致; 通过 PCP PEER Opcode 创建的映射规则可以是端点无关映射 EIM, 也可能是端点相关映射 EDM; 过滤规则可以是端点无关过滤 EIF, 也可能是端点相关过滤 EDF; 具体是哪种映射规则和过滤规则由 NAT 设备或者防火墙上配置决定。

### 4.7.2 PEER Opcode 包结构

PEER Opcode 请求报文格式如图 10 所示。



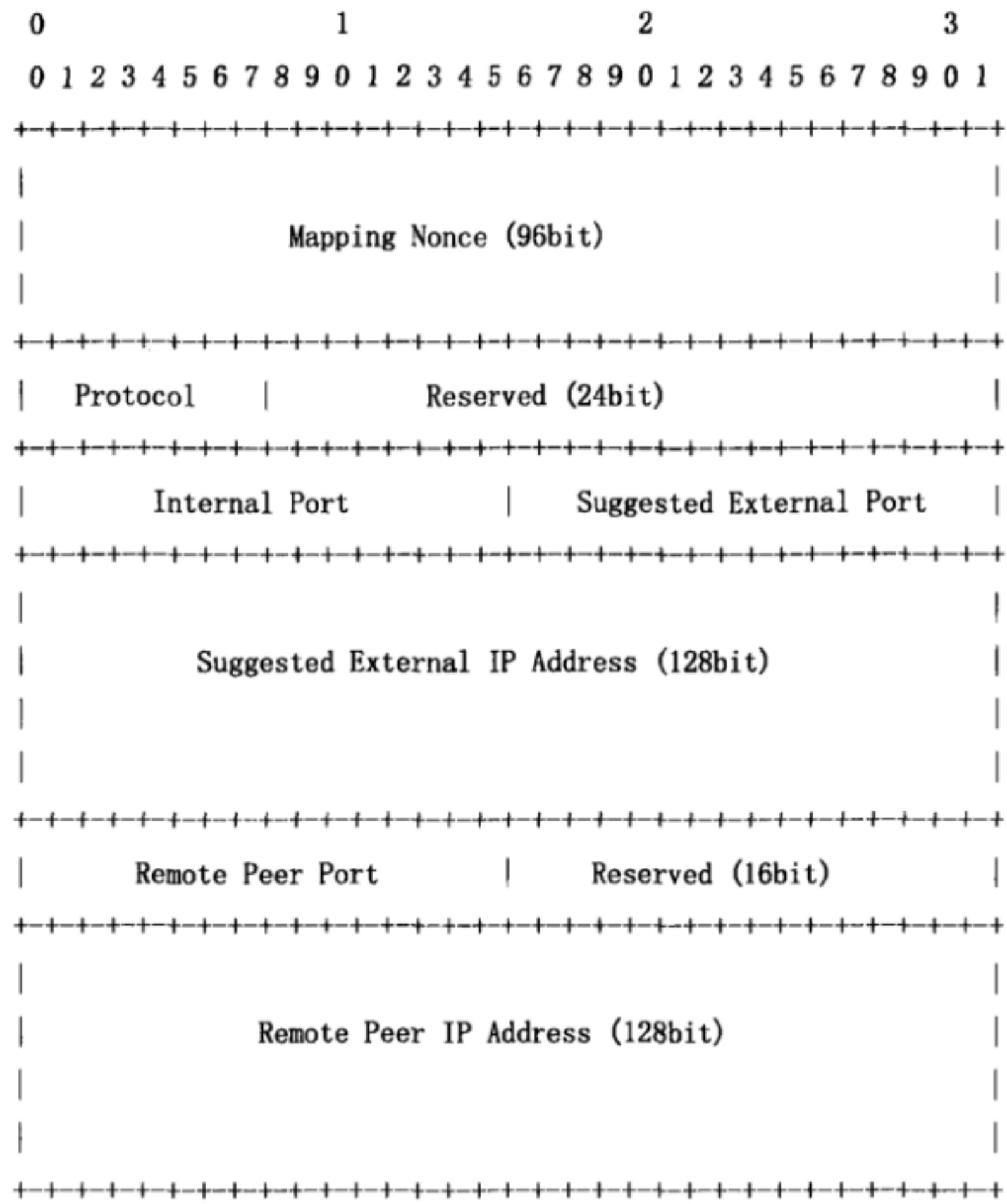


图 10 PEER Opcode 请求报文

各字段描述如下：

**Requested lifetime**(在通用请求报文头中)：请求的映射条目的生命时间，单位是秒，当值为 0 时表示删除该映射条目；当前不支持缩短映射条目的生命时间。

**Mapping Nonce**：PCP Client 在请求报文中填充的随机值字段。

**Protocal**：和 PEER Opcode 请求映射条目相关的协议类型；当值为 6 时表示请求创建一个 TCP 映射。

**Reserved**：24bit 预留字段，发送时该字段必须设置为 0，接收时忽略该字段。

**Internal Port**：映射条目中的内部端口号。

**Suggested External Port**：指定的外部端口号；如果 PCP Client 不知道外部端口号信息，也没有偏好的外部端口号，则该字段必须为 0。

**Suggested External IP Address**：指定的外部 IP 地址；如果 PCP Client 不知道外部 IP 地址信息，也没有偏好的外部 IP 地址，则该字段必须为 0。

**Remote Peer Port**：远端 Peer 端口信息。

**Reserved**：16bit 预留字段，发送时该字段必须设置为 0，接收时忽略该字段。

**Remote Peer IP Address**：PCP Client 角度上认为的远端 Peer 的 IP 地址信息；PCP Client 不需要关心这个地址是经过 NAT64 转换还是经过 NAT46 转换；通过该字段，PCP Server 可以识别相同内

部地址+内部端口号对应不同外部 Server 的多个连接；如果远端 Peer IP 地址实际上是 IPv4 地址，则使用 IPv4-mapped IPv6 地址填充该字段。

当 PCP Server 上映射条目丢失，需要重新创建时，PCP Client 会根据之前响应报文中的外部 IP 地址和外部端口号信息填充 PEER 请求报文中指定外部 IP 地址和指定外部端口号字段，在 PCP Server 上恢复映射条目；当第一次发送 PEER 请求报文时，指定外部 IP 地址和指定外部端口号字段为 0。

当 PEER 请求中指定外部 IP 地址或者指定外部端口号字段中有一个不为 0，且 PCP Server 无法分配指定外部 IP 地址或者指定外部端口号给该映射时，PCP Server 发送携带 CANNOT\_PROVIDE\_EXTERNAL 的 error 响应报文给 PCP Client。

PEER 请求报文中不允许携带 PREFER\_FAILURE 选项，如果 PCP Server 收到 PEER 请求报文中携带 PREFER\_FAILURE 选项，PCP Server 发送携带 MALFORMED\_REQUEST 的 error 响应报文给 PCP Client。

PEER Opcode 响应报文格式如图 11 所示。

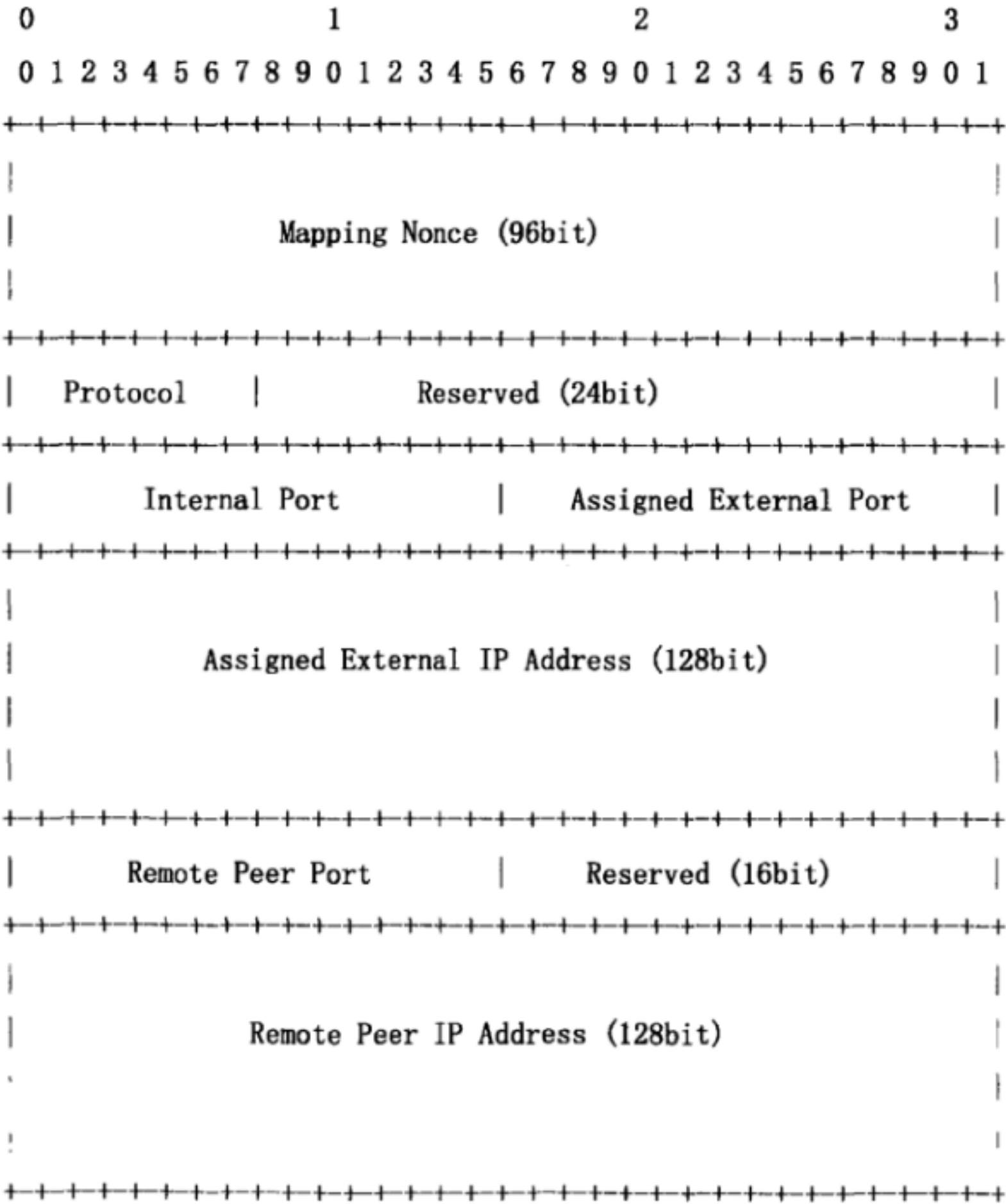


图 11 PEER Opcode 响应报文

各字段描述如下：

Lifetime(在通用请求报文头中)：在 error 响应报文中，该字段表示 PCP Client 再次发送请求报文需要等待的时间；在 success 响应报文中，该字段表示该映射条目的生命时间。

**Mapping Nonce:** 随机值字段，直接从请求报文中拷贝获得。

**Protocol:** 协议类型字段，直接从请求报文中拷贝获得。

**Reserved:** 24bit 预留字段，发送时该字段必须设置为 0，接收时忽略该字段。

**Internal Port:** 从请求报文中拷贝获得。

**Assigned External Port:** 在 success 响应报文中，该字段表示映射条目生成的外部端口号；在 error 响应报文中，该字段直接从请求报文中拷贝获得。

**Assigned External IP Address:** 在 success 响应报文中，该字段表示映射条目生成的外部 IPv4 地址或者 IPv6 地址，且使用 IPv4-mapped IPv6 地址标识 IPv4 地址；在 error 响应报文中，该字段直接从请求报文中拷贝获得。

**Remote Peer Port:** 远端 Peer 端口信息，直接从请求报文中拷贝获得。

**Reserved:** 16bit 预留字段，发送时该字段必须设置为 0，接收时忽略该字段。

**Remote Peer IP Address:** 直接从请求报文中拷贝获得。

#### 4.7.3 PCP Client 产生和发送 PEER Request 消息

PEER Opcode 请求可以发生在和远端 Peer 双向通信前或者通信后。

如果发生在双向通信前，PCP Server 上会根据 PEER 请求创建动态 outbound 映射条目；当 PCP Server 由于瞬断而丢失映射条目时，可以重新发送 PEER 请求恢复映射条目。

如果发生在双向通信后，则 PCP Server 上已经基于流创建了对应的隐式动态映射条目，当收到 PEER 请求报文时，PCP Server 携带隐式动态映射条目中的外部 IP 地址和外部端口信息以及生命时间，发送 success 响应报文告知 PCP Client。

Mapping Nonce 是 PCP Client 随机产生的，用来校验 PCP 响应报文的合法性。

PEER Opcode 请求报文携带远端 Peer 地址字段，该地址一般不是远端 Peer 的真实地址，而是经过 NAT 翻译后的地址。

#### 4.7.4 PCP Server 处理 Request 并响应 Response 消息

PCP Server 接收到 PEER Opcode 请求报文时，处理细节如下：

PCP Server 拷贝请求报文中的协议字段、内部端口号字段、远端 Peer IP 地址字段、远端 Peer 端口字段和 Mapping Nonce 字段到 PEER 响应报文中。

PCP Server 只需要记住每个映射条目的 Mapping Nonce，且最近请求报文中的 Mapping Nonce 覆盖之前请求报文中的 Mapping Nonce 值。

对于隐式动态映射条目的创建，一些 NAT 设备或者防火墙设备会判断报文目的地址的合法性，如果目的地址无效，比如 127.0.0.1，则不创建该流量对应的隐式动态映射条目；PCP 协议也有类似处理过程，PCP Server 接收到 PEER 请求报文后，会判断报文中的远端 Peer IP 地址字段、协议字段和端口字段是否有效，如果无效，则不创建显示动态映射条目；并发送携带 MALFORMED\_REQUEST 的 error 响应报文给 PCP Client。

PCP Server 接收到 PEER 请求报文，查映射表，如果请求的映射不存在，且 PCP Server 能够分配指定外部 IP 地址和指定外部端口信息给该 PEER 请求，则创建映射条目并发送 success 响应报文给 PCP Client；如果 PCP Server 不能分配指定外部 IP 地址和指定外部端口信息给该 PEER 请求，则



发送携带有 CANNOT\_PROVIDE\_EXTERNAL 的 error 响应报文给 PCP Client；如果查映射表发现请求的映射已存在，则更新该映射条目的生命时间。

PEER Opcode 请求支持延长已存在映射条目的生命时间；不支持缩短已存在映射条目的生命时间；如果 PEER 请求报文中的 Requested lifetime 小于已存在条目的生命时间，则返回已存在条目的生命时间给 PCP Client；如果大于已存在条目的生命时间，则更新映射条目的生命时间，并返回更新后的值给 PCP Client。

对于基于 PEER 请求创建的映射条目，需要使用 PEER Opcode 的续租消息去续租，如果没有续租，则和普通 NAT 条目一样，需要一直有流量触发才不会老化。

对于基于 PEER 请求创建的映射条目，任何时候接收到 TCP FIN 或者 TCP RST 消息，都会立即老化掉。

4.7.5 PCP Client 处理 Response 消息

PCP Client 收到响应报文后，除了处理通用响应报文的步骤外，还需要比对响应报文中的 Internal IP Address 字段、Protocol 字段、Internal Port 字段、Remote Peer Address 字段、Remote Peer Port 字段和 Mapping Nonce 字段。

PCP Client 收到 success 响应报文时，按照报文中的生命时间保持映射条目的状态；当续租时间到，PCP Client 可以按照 4.6.4 所述，发送续租报文，保活 PCP Server 上映射条目的状态；也可以通过连续的 outbound 流量保持 PCP Server 上映射条目的状态。

4.8 MAP Opcode 和 PEER Opcode 的选项 options

4.8.1 范围

下面介绍的这些选项是属于 MAP 和 PEER Opcode 的选项，这些选项不可以出现在其他 Opcode 报文中，除非其他 Opcode 也支持该选项。

4.8.2 MAP Opcode 和 PEER Opcode 的 THIRD\_PARTY Option

THIRD\_PARTY 选项用于 PCP Client 代替另一个内网主机管理其在 PCP Server 上映射条目的场景；MAP Opcode 和 PEER Opcode 都支持该选项。

出于安全考虑，必须有充分授信的 PCP Client 才能发送携带 THIRD\_PARTY 选项的 PCP 报文；实际应用中，使用该选项的 PCP Client 一般是个管理设备；该管理设备通过使用该选项代替另一个内网主机来管理其在 PCP Server 上的映射状态。

THIRD\_PARTY 选项包格式如图 12 所示。

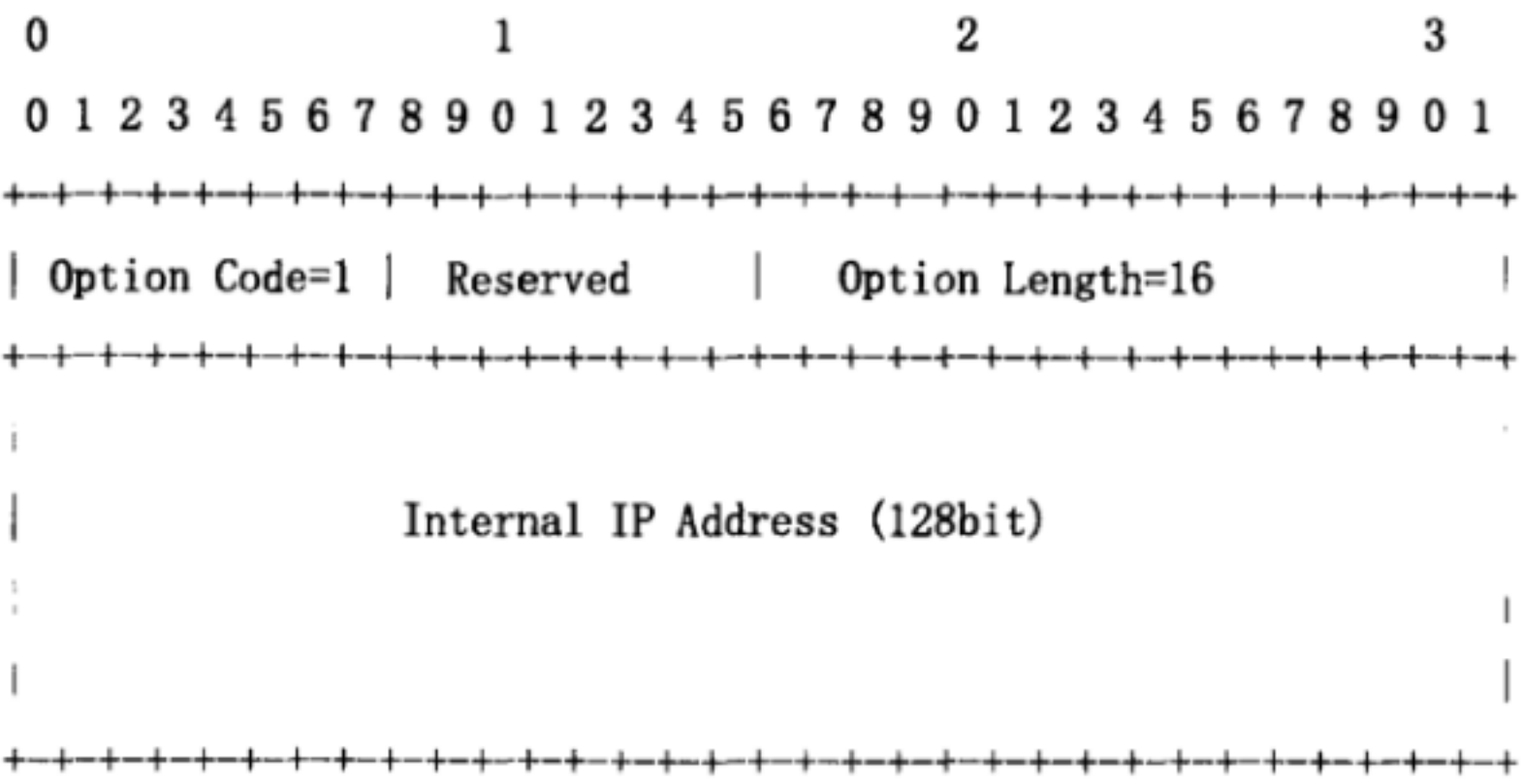


图 12 THIRD\_PARTY 选项

各字段描述如下：

**Option Length:** 选项长度，为 16 字节。

**Internal IP Address:** 创建映射的真实内部 IP 地址。

THIRD\_PARTY 选项携带的内部 IP 地址和请求报文 IP 头中的源 IP 地址不可以一样；当接收到携带 THIRD\_PARTY 选项报文中的两个 IP 地址一样时，PCP Server 发送携带 MALFORMED\_REQUEST 的 error 响应报文给 PCP Client。

通过配置上使能/去使能，PCP Server 上可以实现 THIRD\_PARTY 选项的支持/不支持能力。当 PCP Server 上支持该选项时，授信的 PCP Client 发起携带有 THIRD\_PARTY 选项的请求报文到 PCP Server 时，PCP Server 正确处理并返回携带有 THIRD\_PARTY 选项的响应报文给 PCP Client；当 PCP Server 上不支持该选项时，PCP Server 接收到携带 THIRD\_PARTY 选项的请求报文，发送携带 UNSUPP\_OPTION 的 error 响应报文给 PCP Client。

#### 4.8.3 MAP Opcode 的 PREFER\_FAILURE Option

只有 MAP Opcode 支持 PREFER\_FAILURE 选项。

当 MAP Opcode 请求报文指定外部 IP 地址和指定外部端口在 PCP Server 上创建映射而 PCP Server 不能按照指定信息创建映射时，如果请求报文中携带 PREFER\_FAILURE 选项，则不创建新的映射条目，并发送携带 CANNOT\_PROVIDE\_EXTERNAL 的 error 响应报文给 PCP Client；如果请求报文中不携带 PREFER\_FAILURE 选项，则使用另外可用的 IP 地址和端口号创建新的映射条目。

PREFER\_FAILURE 选项包格式如图 13 所示。

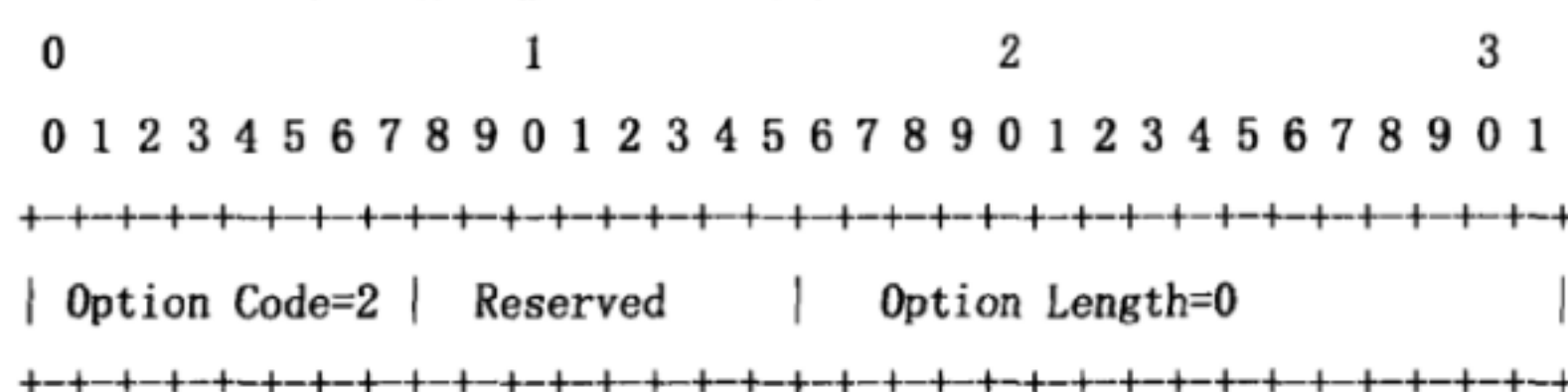


图 13 PREFER\_FAILURE 选项

通过配置上使能/去使能，PCP Server 上可以实现 PREFER\_FAILURE 选项的支持/不支持能力。当支持该选项的 PCP Server 上收到携带有该选项的 MAP 请求报文中指定外部 IP 地址或者指定外部端口号字段为 0 时，PCP Server 判断该报文无效，发送携带 MALFORMED\_OPTION 的 error 响应报文给 PCP Client。

#### 4.8.4 MAP Opcode 的 Addr-Port-Set Option

CGN 设备通过扩展 PCP 消息，向外部地址资源服务器请求分配外部地址资源、或者请求回收本 CGN 设备上的空闲外部地址资源。

地址资源是外部的地址资源服务器，根据 CGN 设备的请求，为 CGN 设备分配外部地址资源、或者对 CGN 设备请求回收的空闲外部地址资源进行回收操作。

地址资源包括地址和端口中间的一种或两种。

Addr-Port-Set Option 选项包格式如图 14 所示。

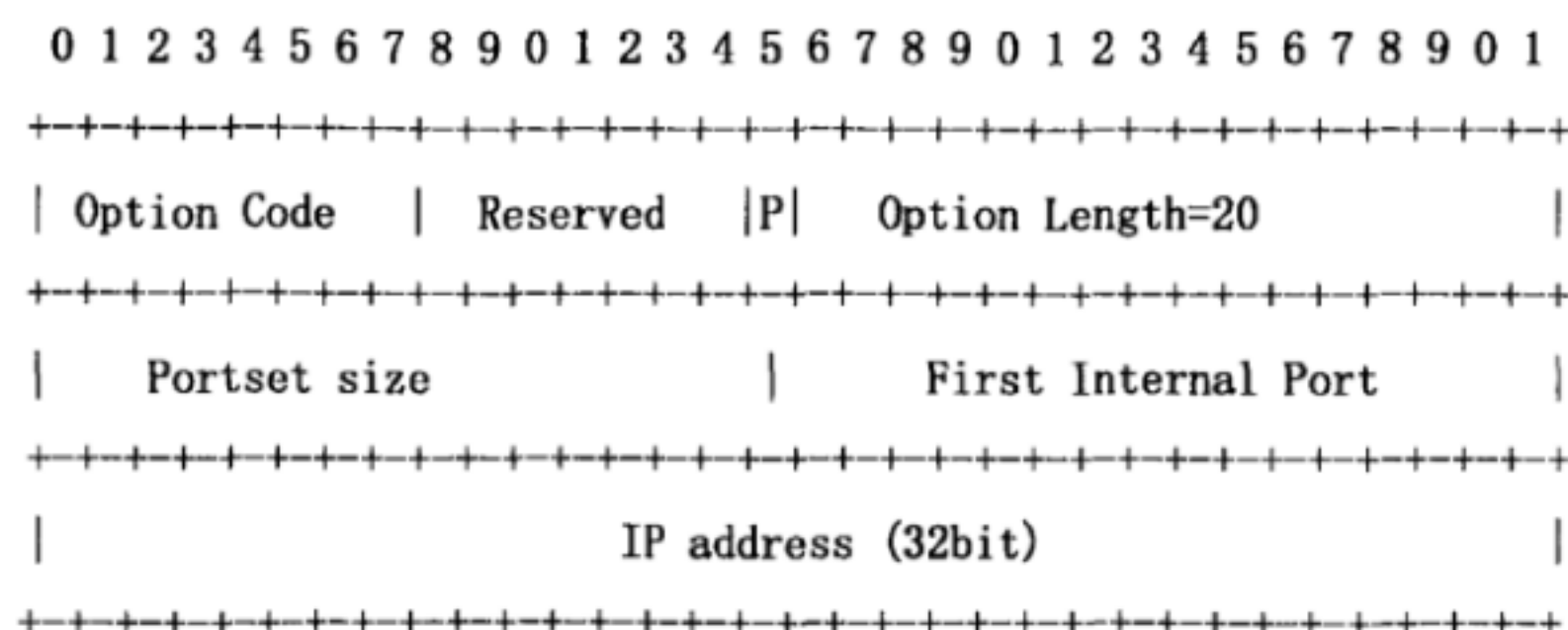


图 14 Addr-Port-Set Option 选项

#### 4.8.5 MAP Opcode 的 FILTER Option

只有 MAP Opcode 支持 PREFER\_FAILURE 选项。

FILTER 选项用于过滤 inbound 流量；MAP 请求报文中携带协议类型；FILTER 选项中携带远端 Peer IP 地址和远端 Peer 端口信息；分别表示允许接收 inbound 流量的协议类型、源 IP 地址和源端口号，来自其他源的流量一律过滤掉。

FILTER 选项包格式如图 15 所示。

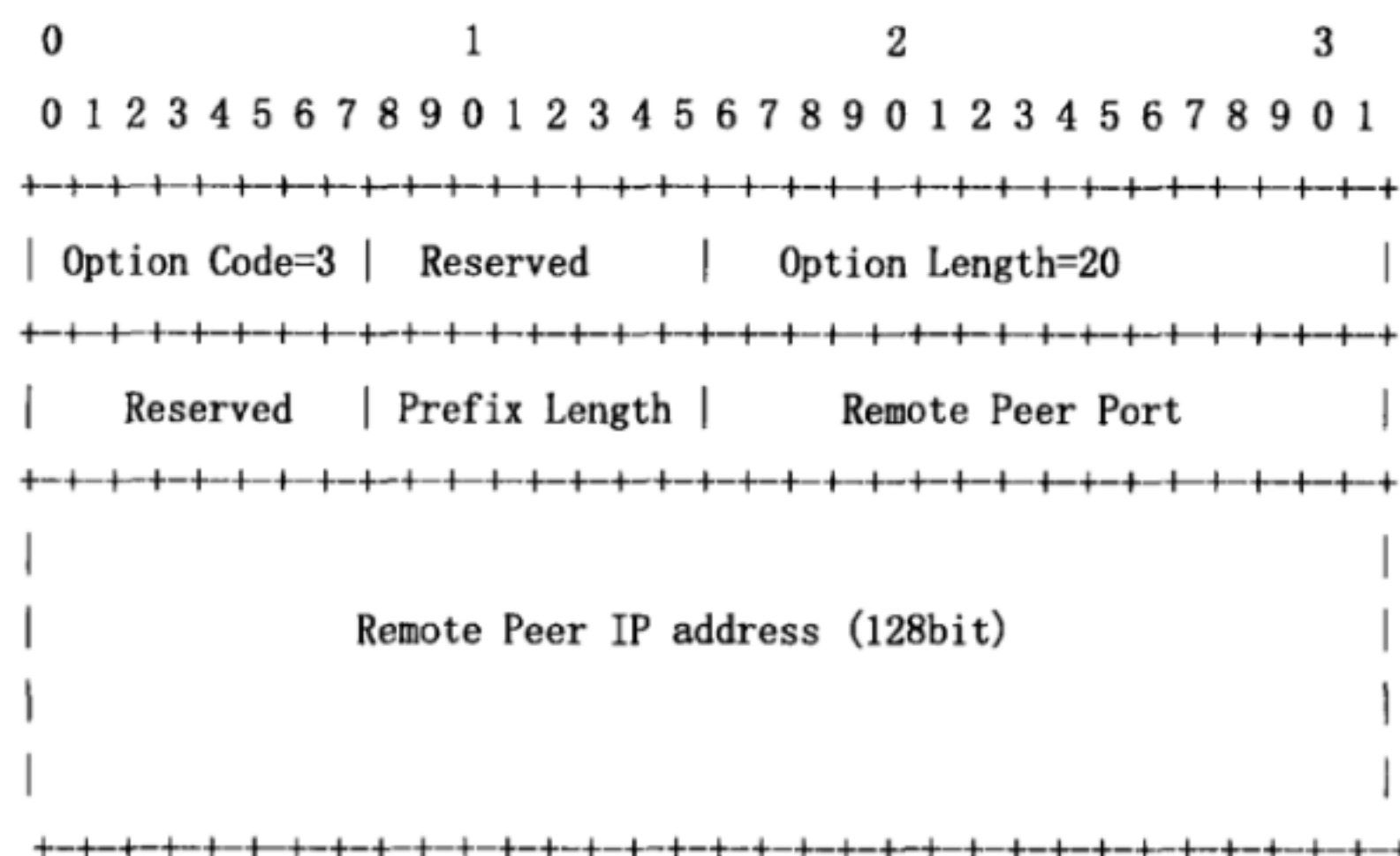


图 15 FILTER 选项

各字段描述如下：

**Option Length:** 选项长度，为 20 个字节。

**Reserved:** 8bit 预留字段，发送时必须设置为 0，接收时忽略该字段。

**Prefix Length:** 表示接收 IPv4 地址或者 IPv6 地址的前缀长度；该值为 0 时表示没有过滤并删除掉之前的所有过滤策略。

**Remote Peer Port:** 远端 Peer 的端口信息，该值为 0 时表示所有端口。

**Remote Peer IP Address:** 远端 Peer IP 地址。

当远端 Peer IP 地址是 IPv4 地址时，Prefix Length 的范围是 96~128bit；当远端 Peer IP 地址是 IPv6 地址时，Prefix Length 的范围是 0~128bit；当接收到 MAP 请求报文中范围不在对应范围内时，PCP Server 发送携带 MALFORMED\_OPTION 的 error 响应报文给 PCP Client。



当收到携带有多个 FILTER 选项的请求报文时, PCP Server 按照顺序逐个处理 FILTER 选项, 并添加过滤策略到已存在策略上; 如果其中任何一个 FILTER 选项处理失败, 则发送携带 MALFORMED\_OPTION 的 error 响应报文给 PCP Client。

当收到携带 FILTER 选项的请求报文时, 如果请求报文中 lifetime 为 0, 则 PCP Server 判断该报文无效, 发送携带 MALFORMED\_OPTION 的 error 响应报文给 PCP Client。

当收到携带 FILTER 选项的请求报文时, 如果选项中 Prefix Length 字段为 0, 则 PCP Server 上删除所有过滤策略; 当前暂不支持删除某一个策略。

如果期望改变当前过滤策略, 则 PCP Client 发送携带 2 个或 2 个以上 FILTER 选项的请求报文, 第一个选项中 Prefix Length 为 0, 用来删除当前 PCP Server 上所有存在的过滤策略; 第二个及以上的选项中携带新的远端 Peer IP 地址和远端 Peer 端口, 用来重新创建过滤策略。

## 5 安全考虑

### 5.1 概述

PCP 协议的目标是改善终端节点对其关联的 NAT 设备的控制能力, 提升错误处理机制。其安全目标是限制任何新的 DoS 攻击机会, 以避免因新的攻击导致未经授权的映射状态的变化。基于 PCP 协议的所有映射条目的创建、修改以及删除都需要经过 PCP Server 的授权。

### 5.2 简单威胁模型

对于 PCP 协议交互路径以外的能伪装成内部网络数据包的攻击者, PCP 协议是安全的; 对于 PCP 协议交互路径以外的能伪装成 PCP Server 的攻击者, PCP 协议也是安全的。除非, 攻击者能修改或者丢弃 PCP 协议交互路径上的报文信息, 或者重定向 PCP 协议交互路径上的信息到另一个主机。

简单威胁模型下, 要求 PCP Server 对于不存在的隐式映射条目, 不会通过显示映射去创建; 这就意味着该模型下, 当 NAT 设备或者防火墙上的 PCP Server 支持 PEER Opcode 时, 要求显示映射和隐式映射条目拥有相同的生命时间, PCP Server 上不支持删除映射条目或者减少已存在映射条目的生命时间, 且 PCP Server 不支持 THIRD\_PARTY 选项以及所有的内部主机都属于一个管理域或者多个有效分区的管理域; 当 NAT 设备或者防火墙上的 PCP Server 支持 MAP Opcode 时, PCP Server 上的安全策略要求支持隐式条目的端点无关过滤。

### 5.3 高级威胁模型

在高级威胁模型下, 要求 PCP 协议保证攻击者无法创建未授权的映射条目或者进行未授权的改动, 即需要 PCP 规定一种能够提供认证、授权等完整信令通道的安全机制。在下列情况下需要使用高级威胁模型的安全机制:

- 基础安全设备, 比如不创建隐式映射条目的公司防火墙;
- 缺乏有效分区机制的且管理多个管理域的设备, 比如 CGN 设备、服务提供商的防火墙设备等;
- 较之隐式映射的认证, 更倾向于显示映射的认证的实现场景;
- 任何不满足简单威胁模型的部署场景。





中华人民共和国  
通信行业标准  
端口控制协议（PCP）技术要求  
YD/T 2934-2015

\*

人民邮电出版社出版发行  
北京市丰台区成寿寺路1号邮电出版大厦  
邮政编码：100164  
北京康利胶印厂印刷  
版权所有 不得翻印

\*

开本：880×1230 1/16 2015年12月第1版  
印张：2 2015年12月北京第1次印刷  
字数：52千字

15115·863

定价：20元

本书如有印装质量问题，请与本社联系 电话：(010)81055492