



一种 IPv6 流标签的编码方法及其应用探讨

李阳春,冯薇薇,邹 洁,谭景华

(中国电信股份有限公司广东研究院 广州 510630)

摘 要:通过重新定义 IPv6 报头 traffic class 字段中的预留位(CU)和 IPv6 流标签的结构化编码方法,定义了流标签携带地域信息、应用信息、端口信息的编码结构,提升了根据流标签值识别业务流量的能力,并探讨了在 IPv6 过渡环境中实现基于流标签的服务质量差异化方法。

关键词:IPv6;流标签;结构化编码;服务质量

doi: 10.3969/j.issn.1000-0801.2013.09.016

A Type of Coding Method of IPv6 Flow Label with QoS Support

Li Yangchun, Feng Weiwei, Zou Jie, Tan Jinghua

(Guangdong Research Institute of China Telecom Co., Ltd., Guangzhou 510630, China)

Abstract: By redefining the reserved CU bits in IPv6 header traffic class field and the structured code of IPv6 flow label, the region information, application information, port number or other information were defined. The capability of identifying traffic type based on flow label code value was improved, the way to use the IPv6 flow label for differentiated service in IPv6 transiting network was discussed.

Key words: IPv6, flow label, structured code, QoS

1 前言

IETF 在 20 世纪 90 年代提出 IPv6, 就引入了流标签(flow label), 至今已经形成 RFC3697、RFC6294、RFC6436、RFC6437 等与流标签相关的 RFC 标准文档。流标签的长度定义从最初的 28 bit, 调整为 24 bit, 最后确定为 20 bit。流标签的定义和应用仍在研究中, 还不成熟, 但形成了典型的 3 种应用: QoS、报文交换和提高安全性。这些应用对 IPv6 流标签定义要求各不相同, 甚至互相矛盾, 比如部分应用要求流标签值不可改变, 另一些应用要求流标签值可以改变; 部分要求采用非结构化编码, 另一些要求采用结构化编码等。这些矛盾的要求缺少协调机制, 流标签的定义也缺少应用类型标识, 网络设备难以确定报文的处理方式, 影响流标签的应用和部署。

本文重新定义了 IPv6 头的 TC 字段的预留位, 并提出

了流标签的结构化编码, 一方面解决了各种流标签定义之间的矛盾, 另一方面提高了 IPv6 流标签应用识别能力, 增强了 IPv6 QoS 的支持能力。

2 流标签研究现状及存在的问题

与 IPv4 报头相比, IPv6 报头的字段数从 14 个减少到 8 个, 删除了报头长度、标识符、标志、分段偏移、报头校验和、选项、填充等字段, 增加了流标签字段, 修改了流量类型命名等, 具体情况如图 1 所示。

RFC3697、RFC6437 等 IETF 标准探讨了流标签的潜在应用及其实现的要求, 但这些应用和要求在某些方面也存在相互矛盾的地方, 影响技术实现和使用。具体来说, 对流标签的要求有如下几方面特点。

(1) 结构化编码和非结构编码

非结构化流标签要求采用随机数, 取值范围为



版本号	报头长度	服务类型	数据分组长度	
标识符		标志	分段偏移	
生存时间	传输协议	报头校验和		
源IPv4地址(32 bit)				
目的IPv4地址(32 bit)				
选项			填充(8 bit)	

(a) IPv4 报文格式

版本号	流量类型	流标签	
净荷长度		下一报头	跳数限制
源IPv6地址(128 bit)			
目的IPv6地址(128 bit)			

(b) IPv6 报文格式

图1 IPv4 报头到 IPv6 报头结构的变化

00000~FFFFF,禁止对标签进行编码,比如 RFC3697 禁止采用数学或其他方法对标签编码,而 RFC6437 则建议采用离散一致分布随机数。

标签的结构化编码方法定义了标签值的编码结构,通常的结构化编码可以采用伪随机数,比如采用 label flag (LF, 1 bit)+label type (LT, 2 bit)+label 码值 (17 bit 随机数) 的形式;同时可以嵌入 QoS 要求,采用 QoS 参数类型 (3 bit 或 4 bit)+随机数,或者 IP Prec (3 bit)+PHB (16 bit)+预留 1 bit 等结构;在逐跳交换中,可以采用 FEC (3 bit)+类型表示 (4 bit)+随机数的编码结构。

当非结构化编码的报文、不同结构化编码方法的报文同时在网上出现时,网络设备无法判断编码方法和执行相应的处理。现有文献对这个问题的讨论也不充分。

(2) 流标签标记能力的局限性

根据现网统计,10 Gbit/s 端口的并发流数最大可以达到 500 万个,40 Gbit/s 或 100 Gbit/s 等更大端口的流数将更多。流标签长 20 bit,最多可以标记大约 100 万个流,大约只占 10 Gbit/s 端口流数的 20%。因此,流标签不能用于对互联网络中的业务流进行一对一的标记,只能标记网络中少数业务流。

(3) 流标签的有效范围

流标签用于 QoS、报文交换和提高安全性等应用。流标签值可以由终端主机设置,也可以由网络设备设置。

当用于 QoS 时,终端主机设置的流标签随意性较大,网络设备通常会重新标记报文的标签值。为了支持服务质量差异化,流标签携带 QoS 信息,必须采用相应的编码方法,用流标签来区分业务流,并支持网络设备执行相应的

处理。当报文离开管理域,标签值就不再有意义。

当用于报文交换时,流标签表示报文的下一跳,实现方式和作用类似于 MPLS 标签。标签逐跳改变,由网络设备重新设置。当报文离开网络设备,标签值就不再有意义。

当用于提高安全性时,流标签值由终端主机设置,标签值在端到端路径上都有效,不允许网络设备修改。

3 支持 QoS 的流标签编码方法

这部分提出一种 IPv6 头中的 TC (traffic class) 字段预留位的编码和流标签编码方式,定义了通过流标签的编码携带地域信息、应用类型、端口号等信息的方法。

3.1 CU 位编码

在 DiffServ 模型中,网络设备接收报文,并根据 DSCP 取值确定报文的逐跳行为 (PHB),完成对报文的处理,比如识别、标识、排队、调度、丢弃等操作。在这个过程中,CU 位被忽略。

(1) DS 字段描述

RFC2474 定义了 IPv4 头和 IPv6 头的 DS 字段。在 IPv4 中,DS 字段是 TOS 字段;在 IPv6 中,DS 字段是 TC 字段,其结构如图 2 所示。

0	1	2	3	4	5	6	7
DSCP						CU	

图2 DS 字段格式

DS 字段的前 6 bit 定义 DSCP 码值,取值范围是 0~63,用于描述在每个节点的逐跳行为 (PHB)。DS 字段的后 2 位预留,未被使用。

(2) DS 字段 CU 编码方法

为了支持网络设备根据 IPv6 报头对流标签采用不同的处理方式,重新定义了 DS 字段中的 CU 码值,并描述了不同码值对应的流标签的不同应用类型,如图 3 所示。

00:缺省值,用于描述流标签的缺省状态;

01:QoS,用于描述支持 QoS 处理的流标签编码;

10:安全应用,用于描述支持安全应用的流标签编码;

0	1	2	3	4	5	6	7
DSCP						CU	
CU 码值		00	01	10	11		
流标签应用类型		缺省	QoS	安全	报文交换		

图3 CU 编码定义

11:报文交换,用于描述支持报文交换的流标签编码。

(3)网络节点的处理

当 IPv6 报文到达支持 IPv6 的网络设备时,该网络设备首先处理 TC 字段的 CU 位,再根据 CU 标志位解析流标签值,并执行相应的处理。网络设备根据 CU 标识位发起不同的处理流程,如图 4 所示。

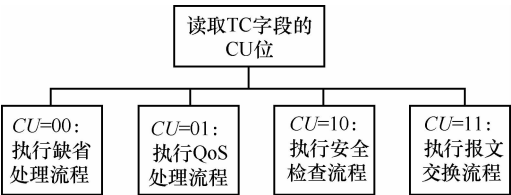


图 4 网络节点对 CU 处理的几种场景

当 CU 二进制值为 00,表示流标签未做任何设置,不需要网络设备做任何处理;

当 CU 二进制值为 01,流标签携带 QoS 信息,网络设备按照流标签值被流量、划分等级、排队和执行调度等策略;

当 CU 二进制值为 10,流标签携带用于提升安全性的流信息,网络设备执行安全性检查流程;

当 CU 二进制值为 11,流标签携带的信息用于支持报文交换,网络设备用下一跳的标签替换现有标签,并把报文转发到下一跳。实现方式类似 MPLS 标签的处理。

对于不同的处理流程,本文不详细讨论。本文只关注 CU=01 时的情况,即流标签用于携带 QoS 信息、支持服务质量差异化的使用情况。

3.2 流标签编码及其实现

传统的 QoS 机制主要基于链路层、网络层或者传输层信息识别流量,比如 IP 地址、IP precedence/EXP/DSCP、interface/VLAN ID、IEEE 802.1Q 等,难以识别会话层以及应用层等各种上层信息,也无法基于这些信息划分流量类型,执行 QoS 策略。当 IPv6 报文中存在多个扩展头时,网络设备识别传输层信息(比如 TCP/UDP 端口号)的难度也会变大,同样影响网络设备执行 QoS 策略。

当流标签用于支持 QoS,流标签可以用于携带上层应用信息,比如标准化的应用信息(包括端口号等)、地域信息、自定义的应用类型等信息,识别和标记业务流量,支持网络设备执行 QoS 策略。

3.2.1 流标签的结构化编码

为了描述上层应用信息,这里首先定义流标签如图 5 的编码方法。



图 5 流标签的编码结构

如图 5 所示,流标签分成标签类型和信息编码两部分。这两部分的具体描述如下。

- 标签类型:长度 4 bit,取值范围 0~15,描述携带的上层应用类型,比如端口号、地域以及其他自定义的应用类型。
- 信息编码:长度 16 bit,取值范围 0~65 535,描述携带的上层应用信息的码值,比如端口号、地域信息以及其他自定义的应用信息。

当报文到达网络设备,网络设备根据标签类型判断流标签携带的信息类型,并按照不同的信息类型执行不同的处理,比如对比地域编码,识别流量的归属地;识别应用编码,确定业务流量所述的应用类型等。

3.2.2 流标签的编码定义

根据第 3.1 节的定义,当 CU 二进制值为 01,网络节点需要解析流标签编码,确定相关信息,并触发 QoS 处理流程。

这里首先定义一种流标签中的标签类型的编码方式,并用不同的编码制表示后 16 位编码携带的信息的类型,具体定义见表 1。

表 1 不同的编码制表示后 16 位编码携带的信息的类型

标签类型值	后 16 位携带信息 的类型	使用方式
0000	携带 TCP 端口号	用于复制 TCP 头中的端口号
0001	携带 UDP 端口号	用于复制 UDP 头中的端口号
...	不同应用类型的 标签码值	用于表示不同的应用类型,如 HTTP 应用类型、SIP 应用、P2P 应用、iTV 应用、E-mail
1111	携带地域信息	与地域信息编码表比较,可以 识别用户归属的地理位置,比 如广州天河区等

由于标签类型占了 4 bit,有 16 种不同的标签类型值,可以描述 16 种不同的信息类型。表 1 只定义了 3 类,即 0000、0001、1111 等类型的编码值,没有为不同应用类型定义相应的标签码值。在实际应用中,对具体应用的标签编码值可以由网络设备按照统一的策略定义,比如定义 HTTP 应用采用 0010,E-mail 应用采用 0011 等。

(1)标签携带 TCP/UDP 端口号

当标签的二进制类型值是 0000 和 0001,流标签的后 16 bit 可以用于携带 TCP 端口号和 UDP 端口号,流标签的



编码结构如图 6 所示。

4				16											
0	0	0	T/U	端口号											

图 6 TCP/UDP 端口号的标签编码

端口号:由 RFC6335 描述的 TCP、UDP 端口号组成。

T/U:0 表示 TCP;1 表示 UDP,指端口号的类型。

这种方式把传输层的 TCP/UDP 端口号复制到流标签中,由网络层设备根据端口号确定上层应用类型,实现对应用类型的识别。这种方式可以避免多个 IPv6 扩展头对网络设备读取传输层信息的影响。具体来说,在流标签中携带 TCP/UDP 端口号可以由于如下目的。

- 增加流信息。与 IPv4 相比,IPv6 只采用源/目的地址、流标签等三元组信息表示流,信息相对较少。当在流标签中携带 TCP/UDP 端口号,可以增加上层信息,帮助网络设备识别业务类型。
- 降低网络设备的处理量。当 IPv6 报文中有多扩展头时,网络节点要确定传输层的 TCP 端口号和 UDP 端口号,需要跨过多个扩展头,处理量增大。为了降低网络节点的负担,用流标签携带 TCP/UDP 端口号可以降低网络设备的处理量,又可以增加上层信息,提高业务识别能力。

(2)携带地域信息的标签编码

当流标签用于携带地域信息时,标签类型编码取值为 1111,即 10 进制 15。流标签的后 16 位编码用于表示地域信息,取值范围是 0~65 535。在实际使用中,流标签的地域信息编码可以采用电话区号或其他定义的信息。

当采用电话区号表示地域信息,地域信息编码可以采用如下十进制编码方式:

电话区号+地区编号

电话区号的十进制数值最大不超过 1 000,比如北京、上海、广州的电话区号分别是 010、021、020。地区编号可以是各城市的行政区的编号。以广州为例,如果天河区编号是 01,则天河区用户访问的地域编码可以采用 02001,对应的二进制流标签取值就变成 1111 0000 0011 1101 0001。

当网络节点需要对不同地域的业务流量采用不同的保障策略时,地域信息可以提供用户的归属信息,支持网络设备采用有针对性的策略。同时,地域信息也是分析网络流量流向的重要信息。在具体部署时,地域信息编码可以由运营商灵活定义,并根据统一策略设置。通常来说,终

端主机设置的地域编码可以在边缘接入设备重新设置。

(3)携带应用信息的标签编码

标签类型值除了可以描述携带的 TCP/UDP 端口号、携带地域信息,还可以用来表示 HTTP、SIP、P2P、iTV、E-mail 等应用类型。按照定义,标签类型值可以表示 13 种不同类型的应用。网络设备根据标签可以识别不同的应用类型,并提供不同等级的服务质量保证。

通过设置标签类型值描述不同应用类型,网络设备可以结合应用类型、IP 地址、端口号、precedence 等报文携带的信息,提供更准确、灵活的服务质量保障策略。

4 流标签的 QoS 应用

在 IPv6 QoS 中,流标签的主要用途是识别业务流量,并支持网络设备在 DiffServ 模型中实现服务质量差异化。这部分的讨论分为 3 部分:基于流标签地域信息的策略路由;基于流标签应用信息的服务质量保证;流标签在 IPv6 过渡技术中的应用。

4.1 基于流标签地域信息的策略路由

当流标签用于携带地域信息,网络设备可以根据流标签携带的地域信息识别属于同一个区域的业务流量,并部署策略路由。这里以实例说明基于流标签地域信息的策略路由。

如图 7 所示,假设地域 A、B、C 3 个区域的业务流量分别采用不同的策略。

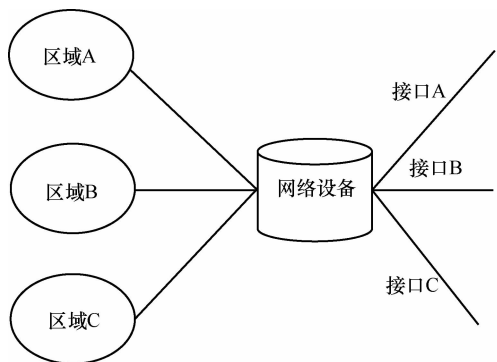


图 7 基于流标签地域信息的策略路由

对于来自区域 A、B、C 的报文,IPv6 头中的流标签的前 4 位都设置为 1111,即十进制的 15;后 16 位二进制值可以按照如表 2 设置(示例)。

与流标签的前 4 位二进制 1111 一起,流标签的后 16 位定义了携带的地域信息。表 2 中的区域 A、B、C 分别用十进制数 2001、2002、2003 表示,是广州的 3 个不同地区。

表 2 后 16 位二进制值

区域	后 16 位二进制值	后 16 位 10 进制值
区域 A	0000 0011 1101 0001	02001
区域 B	0000 0011 1101 0010	02002
区域 C	0000 0011 1101 0011	02003

根据流标签的区域信息,网络设备可以定义如下策略:

- 区域 A 的业务流量,从接口 A 转发;
- 区域 B 的业务流量,从接口 B 转发;
- 区域 C 的业务流量,从接口 C 转发。

当报文到达网络设备后,网络设备匹配报文的流标签值,并按照预先设置的策略从不同的端口转发报文(如图 7 所示)

4.2 基于流标签应用信息的服务质量保证

基于流标签应用信息的服务质量保证用流标签携带应用信息,根据应用信息识别业务流量,支持服务质量差异化。这里以实例说明基于流标签地域信息的服务质量保证。

如图 8 所示,假设有 3 种流量类型:Web 网站访问、VoIP、E-mail,网络设备分别对这 3 种应用采用不同的服务质量保证策略。

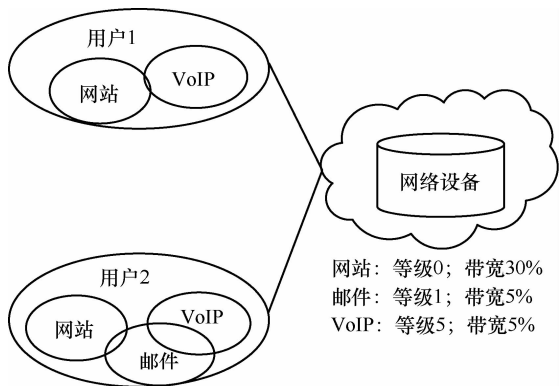


图 8 基于流标签应用信息的服务质量保证

对于来自用户 1、用户 2 的网站、VoIP、邮件等业务流量,IPv6 头中的流标签的前 4 位可以设置为表示互联网应用,比如 0010,即十进制的 2;后 16 位二进制值可以设置各种应用的代码(示例),如表 3 所示。

与流标签的前 4 位二进制 0010 一起,流标签的后 16 位定义了携带的应用类型信息。根据流标签的携带应用类型信息,网络设备可以定义如下策略。

- 访问网站的业务流量:服务等级为 0,占转发端口的 30%带宽。

表 3 后 16 位二进制值

应用类型	后 16 位二进制值	后 16 位 10 进制值
网站	0000 0000 0000 0000	0
E-mail	0000 0000 0000 0001	1
VoIP	0000 0000 0000 0010	2

- 邮件流量:服务等级为 1,占转发端口的 5%带宽。
- VoIP 流量:服务等级设为 5,栈转发端口的 5%带宽。

当报文到达网络设备后,网络设备匹配报文的流标签值,并按照预先设置的策略为不同类型的业务流量提供不同保障等级,支持业务流量的差异化,如图 8 所示。

4.3 过渡技术中的流标签应用

流标签是 IPv6 独有的字段。当流标签被用于携带地域、应用等信息,与 IPv4 相比,流标签提供了更多、更准确识别 IPv6 业务流量的能力,提高了 IPv6 服务质量的能力。与此相对,流标签难以用于 IPv4 报文,从而导致 IPv6 和 IPv4 业务流量具有不一致的服务质量保证能力。在双栈、隧道和 IPv6 与 IPv4 互通等 IPv6 过渡部署场景中,由于 IPv6 流标签的应用,IPv6 QoS 和 IPv4 QoS 在保障能力上表现出差异,如图 9 所示。

如图 9(a)所示的双栈网络环境。网络设备运行双栈,承载 IPv4 和 IPv6 业务流量。IPv4 按照现有机制划分多个等级,IPv6 继承了 IPv4 的实现机制,同时实现基于流标签的业务流量分类。流标签只能用于识别 IPv6 业务流量,无法用于识别 IPv4 业务流量。流标签可以用于提供更准确和精细的业务流量识别能力和 QoS 支持能力。

当 IPv4 封装在 IPv6 隧道中或者 IPv6 封装在 IPv4 隧道中,其应用场景如图 9 (b) 所示。R1 与 R2 之间建立了 IP-in-IP 的隧道。内层报头携带的应用信息应该复制到外层 IP 报头中,才能保证报文在隧道封装前后获得相同的服务质量保证。由于流标签是 IPv6 独有的,携带在 IPv6 头中的流标签信息(如地域信息、应用信息等)难以复制或映射到 IPv4 头,导致 IPv6-in-IPv4 隧道封装丢失了流标签的信息;与此同时,IPv4-in-IPv6 的隧道封装缺少对流标签设置的判据,也会影响到基于流标签的应用。

在 IPv6 与 IPv4 互通的场景中,如图 9(c)所示,IPv6 报文和 IPv4 报文之间需要互相翻译。IPv4 头与 IPv6 头之间进行如下翻译:

TOS DSCP↔TC DSCP

TOS/IP precedence↔TC/IP precedence

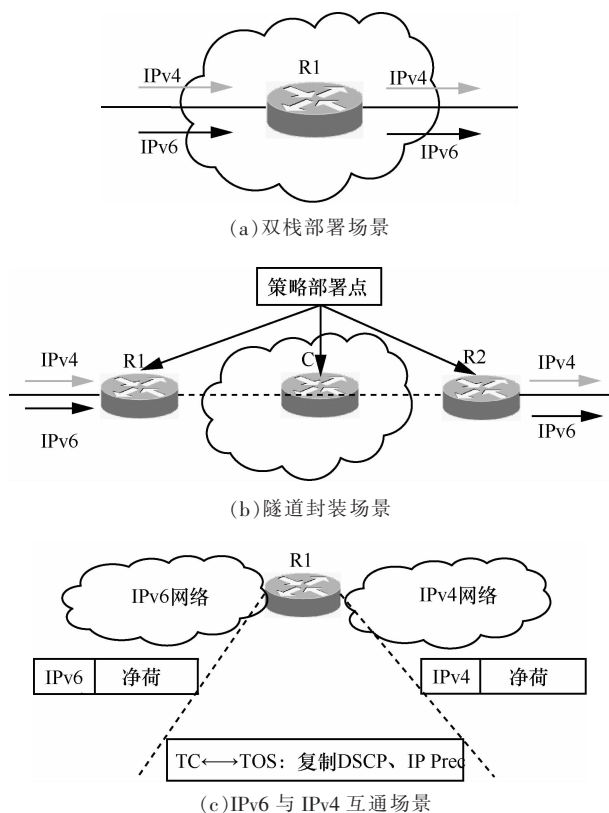


图9 IPv6 过渡部署场景中的流标签应用

相互之间只需复制 DSCP 和 IP precedence 值,就可以完成相互之间的信息映射。与此同时,流标签携带的地域信息和应用信息将丢失,影响保证 IPv6 业务流量的服务质量和用户体验。

5 结束语

本文通过重新定义 IPv6 报头 TC 字段 CU 保留位,标识流标签的不同应用方式:QoS、提高安全性、报文交换。通过对 CU 保留位的重新定义,网络设备可以根据 CU 位的不同编码值,确定流标签值的结构,并确定对 IPv6 报文应该执行处理流程。

本文把流标签分成标签类型值和信息编码两部分,用标签类型值描述信息编码代表的不同意义,用信息编码携带不同信息的编码值。虽然标签类型暂用 4 位,可以用于描述 16 类信息类型,但本文重点讨论流标签被用于携带地域信息、应用信息,提高服务质量保证能力的编码方法。在定义流标签的编码方式之后,本文以举例的方式探讨了流标签在 IPv6 QoS 方面的应用。

本文对流标签的定义有助于提高 IPv6 服务质量保障

能力,但还需要标准化这种编码方式,推动厂商设备支持。流标签的作用在于标识和识别业务流,可以支持业务流量分析,比如地域信息、应用信息可以用于流量流向分析、应用特征分析,为运营商的网络运营和规范提供基础数据。与此同时,在实际网络中如何应用和部署基于流标签的 QoS 策略还需要进一步研究。

参考文献

- 1 Rajahalme J, Conta A, Carpenter B, *et al.* IPv6 Flow Label Specification. RFC3697, 2004
- 2 Hu Q, Carpenter B. Survey of Proposed Use Cases for the IPv6 Flow Label. RFC6294, 2011
- 3 Amante S, Carpenter B, Jiang S. Rationale for Update to the IPv6 Flow Label Specification. RFC6436, 2011
- 4 Amante S, Carpenter B, Jiang S, *et al.* IPv6 Flow Label Specification. RFC6437, 2011
- 5 Westerlund M, Cheshire S. IANA Procedure for the Management of the Service Name and Transport Protocol Port Number Registry. RFC6335, 2011
- 6 Deering S, Hinden R. Internet Protocol Version 6 (IPv6) Specification. RFC2460, 1998
- 7 Nichols K, Blake S, Baker F, *et al.* Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers. RFC2474, 1998

[作者简介]



李阳春,男,博士,中国电信股份有限公司广东研究院高级工程师,长期从事互联网领域的技术研究,重点研究下一代互联网相关技术。

冯薇薇,女,中国电信股份有限公司广东研究院工程师,长期从事互联网领域的技术研究,重点研究移动互联网相关技术。

邹洁,女,中国电信股份有限公司广东研究院高级工程师、技术总监,长期从事互联网领域技术研究,重点研究下一代互联网相关技术。

谭景华,男,中国电信股份有限公司广东研究院高级工程师,长期从事互联网领域的技术研究工作,重点研究下一代互联网相关技术。

(收稿日期:2013-06-05)