

ICS 33.040.40
M 32



中华人民共和国通信行业标准

YD/T 1097-2009

代替 YD/T 1097-2001

路由器设备技术要求 核心路由器

Equipment technical specification Core router

2009-06-15 发布

2009-09-01 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 规范性引用文件	1
3 术语、定义和缩略语	4
4 核心路由器功能要求	13
5 接口类型及特性	15
6 通信规程	22
7 协议要求	25
8 安全性要求	53
9 性能指标要求	53
10 定时和同步要求	55
11 操作管理维护要求	58
12 硬件要求	65
13 软件要求	65
14 机械结构与工艺要求	66
15 过流过压保护要求	67
16 环境要求	68
17 电源与接地要求	70
附录 A（规范性附录）SDH 上传送 IP 的技术要求	72
附录 B（规范性附录）在 ATM 上支持传统 IP 及地址解析（ARP）的技术要求	79
附录 C（资料性附录）区分业务	87

前 言

本标准代替 YD/T 1097-2001《路由器设备技术规范——高端路由器》。

本标准与 YD/T 1097-2001 相比主要变化如下：

- 标准名称由《路由器设备技术规范——高端路由器》修改为《路由器设备技术要求 核心路由器》；
- 修改了对核心路由器的界定（见第 1 章）；
- 在功能要求中增加了 MPLS、QoS、IPsec 及 VPN 等功能要求（见第 4 章）；
- 第 5 章新增 10GE、10G POS 接口要求（见 5.2.3 节），删除 E3 接口（2001 年版的 5.2.4 节）；
- 在协议要求中增加了 MPLS 要求（见 7.7 节），增加了 IPsec 协议要求（见 7.8 节），增加了 VPN 要求（见 7.9 节）；
- 增加了安全性要求（见第 8 章）；
- 在性能指标中增加或修改以下技术指标：将系统双向交换容量改为 60Gbit/s，删除对吞吐量的指标要求，删除对丢包率的指标要求，删除设备处理能力的要求，增加对 BGP、IGP 邻居的数量要求，增加对标记交换吞吐量、标记交换转发延迟指标（见第 9 章）；
- 删除关于 IPv6 的参考性附录（2001 年版的附录 D）。

本标准是《支持IPv4的路由器设备》系列标准之一，本系列的标准结构和名称预计如下：

1. YD/T 1096-2009 路由器设备技术要求 边缘路由器；
2. YD/T 1097-2009 路由器设备技术要求 核心路由器；
3. YD/T 1098-2009 路由器设备测试方法 边缘路由器；
4. YD/T 1156-2009 路由器设备测试方法 核心路由器。

与本系列标准相关的标准还有《支持IPv6的路由器设备》系列标准，该系列的标准结构和名称如下：

1. YD/T 1452-2006 IPv6 网络设备技术要求——支持 IPv6 的边缘路由器；
2. YD/T 1453-2006 IPv6 网络设备测试方法——支持 IPv6 的边缘路由器；
3. YD/T 1454-2006 IPv6 网络设备技术要求——支持 IPv6 的核心路由器；
4. YD/T 1455-2006 IPv6 网络设备测试方法——支持 IPv6 的核心路由器。

本标准的附录 A、附录 B 均为规范性附录。

本标准的附录 C 是资料性附录。

本标准由中国通信标准化协会提出并归口。

本标准起草单位：工业和信息化部电信研究院、华为技术有限公司、中兴通讯股份有限公司、上海贝尔阿尔卡特股份有限公司

本标准主要起草人：田 辉、高 巍、马 科、马军锋、赵 锋、刘 述、唐 浩

本标准于 2001 年首次发布，本次为第一次修订。

路由器设备技术要求 核心路由器

1 范围

本标准规定了核心路由器的技术要求，包括功能特性、通信规程、协议要求、路由协议、接口类型及特性、性能指标要求、定时和同步要求、可靠性和可用性要求、操作维护功能以及安全、环境等内容。

本标准适用于支持IPv4的核心路由器设备。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件，其随后所有的修改单（不包括勘误的内容）或修订版均不适用于本标准。然而，鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件，其最新版本适用于本标准。

GB/T 7611	数字网系列比特率电接口特性
GB/T 20185-2006	同步数字体系设备和系统的光接口技术要求
YD/T 877-1996	同步数字体系（SDH）复用设备和系统的电接口技术要求
YD/T 976-1999	B-ISDN 用户—网络接口（UNI）物理层技术规范
YD/T 1061-2000	SDH 上传送 IP 的 LAPS 技术要求
YD/T 1177-2002	IP 组播路由协议
YD/T 1162.1-2005	多协议标记交换（MPLS）技术要求
YD/T 1358-2005	路由器设备安全技术要求—中低端路由器（基于 IPv4）
YD/T 1359-2005	路由器设备安全技术要求—高端路由器（基于 IPv4）
YD/T 1476-2006	基于边界网关协议/多协议标记交换的虚拟专用网（BGP/MPLS VPN）技术要求
YD/T 1466-2006	IP 安全协议（IPSec）技术要求
YDN 052-1998	B-ISDN ATM 层技术规范
YDN 053.4-1998	B-ISDN ATM 适配层（AAL）类型 5 技术规范
YDN 067-1998	ATM 交换机设备技术规范
YDN 081-1998	宽带综合业务数字网 DSS2 技术规范—用户呼叫控制部分（点到点）
YDN 082-1998	宽带综合业务数字网 DSS2 技术规范—用户呼叫控制部分（点到多点）
YDN 083.1-1998	宽带综合业务数字网（B-ISDB）2 号数字用户信令系统（DSS2）技术规范—适配层 第一部分：业务特定面向连接协议
YDN 083.3-1998	宽带综合业务数字网（B-ISDB）2 号数字用户信令系统（DSS2）技术规范—适配层 第三部分：支持用户—网络接口
YDN 099-1998	光同步传送网技术体制
YDN 120-1999	光波分复用（WDM）系统总体技术要求
IEEE802.1Q（1998）	虚拟桥接局域网（VLAN）
IEEE802.2/3（1985）	局域网协议标准
IEEE802.3ab（1999）	用于操作在 4 对 5 类线平衡铜缆上的 1000BASE-T 物理层参数和规范

YD/T 1097-2009

IEEE802.3ae	10G 以太网标准
IEEE802.3u	百兆以太网标准 (100Base-TX)
IEEE802.3z (1998)	千兆以太网标准 (1000Base-LX/1000Base-SX)
ITU-T E.164 (1993)	国际公用电信编号计划
ITU-T G.704	用于 1544、6312、2048、8448 以及 44736 kbit/s 系列的同步帧结构
ITU-T G.707 (1996)	用于 SDH 的网络节点接口
ITU-T G.782	SDH 设备的类型和总体特征
ITU-T G.783 (1991)	SDH 设备功能组特性
ITU-T G.784	同步数字体系 (SDH) 的管理
ITU-T G.804	ATM 信元到 PDH 的映射
ITU-T G.832	34368kbit/s 信号虚串联的有效负载类型代码
ITU-T I.432 (1995)	宽带综合业务数字网 (B-ISDN) 用户网络接口物理层规范
IETF RFC768 (1980)	用户数据包协议
IETF RFC791 (1981)	互联网协议
IETF RFC792 (1981)	互联网控制消息协议
IETF RFC793 (1981)	传输控制协议
IETF RFC795 (1981)	服务映射
IETF RFC796 (1981)	地址映射
IETF RFC826 (1982)	以太网地址解释协议 (ARP)
IETF RFC922 (1984)	存在子网时的互联网数据包广播
IETF RFC950 (1985)	互联网标准子网划分规程
IETF RFC1075 (1988)	距离矢量组播路由协议
IETF RFC1089 (1989)	以太网上的SNMP
IETF RFC1108 (1991)	IP安全任选域
IETF RFC1112 (1989)	IP组播主机扩展
IETF RFC1122 (1989)	互联网主机要求——通信层
IETF RFC1123	互联网主机要求——应用和支持
IETF RFC1142 (1990)	IS-IS域内路由协议
IETF RFC1154	互联网消息的编码头字段
IETF RFC1155 (1990)	用于TCP/IP互联网的管理信息的结构和标识
IETF RFC1191 (1990)	路径MTU发现
IETF RFC1195 (1990)	在TCP/IP和双重环境路由中使用OSI的IS-IS
IETF RFC1212	简要管理信息库 (MIB) 定义
IETF RFC1213 (1991)	管理信息库 (MIB-II)
IETF RFC1224 (1991)	用于管理异步产生的告警的技术要求
IETF RFC1256 (1991)	ICMP路由发现消息
IETF RFC1285 (1992)	FDDI管理信息库

IETF RFC1332 (1992)	PPP互联网协议控制协议 (IPCP)
IETF RFC1334 (1992)	PPP认证协议 (PAP)
IETF RFC1349 (1992)	互联网协议组中的服务类型
IETF RFC1418 (1993)	OSI上的SNMP
IETF RFC1471 (1993)	对PPP链路控制协议管理对象的定义
IETF RFC1472 (1993)	对PPP安全协议管理对象的定义
IETF RFC1473 (1993)	对PPP IP网控制协议管理对象的定义
IETF RFC1483 (1993)	ATM AAL5的多协议封装
IETF RFC1512 (1993)	FDDI管理信息库
IETF RFC1542 (1993)	BOOTP的澄清及扩展
IETF RFC1619 (1994)	SONET/SDH上的PPP技术要求
IETF RFC1659 (1994)	对RS-232链路硬件设备管理对象的定义
IETF RFC1661 (1994)	点到点协议 (PPP)
IETF RFC1662 (1994)	HDLC帧中的PPP
IETF RFC1700 (1994)	授权号
IETF RFC1724 (1994)	RIPv2 MIB扩展
IETF RFC1755 (1995)	ATM上支持IP的ATM信令
IETF RFC1772 (1995)	边缘网关协议在互联网中的应用
IETF RFC1812 (1995)	IPv4路由器技术要求
IETF RFC1994 (1996)	PPP握手认证协议 (CHAP)
IETF RFC1997 (1996)	BGP团体 (community) 属性
IETF RFC2132 (1997)	DHCP选项域及BOOTP厂商扩展
IETF RFC2225 (1998)	ATM上支持传统IP及ARP
IETF RFC2236 (1997)	互联网组管理协议IGMP (版本2)
IETF RFC2320 (1998)	用于ATM上支持传统IP及ARP的使用SMIv2的管理对象的定义
IETF RFC2328 (1998)	开放式最短路径优先 (版本2)
IETF RFC2332 (1998)	下一跳解释协议 (NHRP)
IETF RFC2439 (1998)	BGP4路由振荡抑制
IETF RFC2474 (1998)	在互联网协议族中服务类型
IETF RFC2515 (1999)	用于ATM管理的管理对象的定义
IETF RFC2615 (1999)	SONET/SDH上的PPP技术要求
IETF RFC2644 (1999)	路由器中定向广播缺省值的修改
IETF RFC2684 (1999)	AAL5上的多协议封装
IETF RFC2778 (2000)	虚拟冗余路由器协议的MIB
IETF RFC2819 (2000)	远程网络监控MIB
IETF RFC2863 (2000)	使用SMI v2的接口组MIB
IETF RFC3416 (2002)	SNMPv2的协议操作

IETF RFC3417 (2002)	SNMP的传送映射
IETF RFC3418 (2002)	用于SNMPv2的MIB
IETF RFC3592 (2003)	用于SDH接口类型的管理对象的定义
IETF RFC3635 (2003)	以太网接口类型的管理对象定义
IETF RFC3768 (2004)	虚拟冗余路由器协议 (VRRP)
IETF RFC3896 (2004)	对E3/DS3接口类型管理对象的定义
IETF RFC4022 (2005)	使用SMIv2对TCP的MIB
IETF RFC4113 (2005)	使用SMIv2对UDP的MIB
IETF RFC4271 (2006)	边缘网关协议第4版本 (BGP4)
IETF RFC4273 (2006)	BGP4管理对象的定义
IETF RFC4292 (2006)	IP转发表MIB
IETF RFC4293 (2006)	使用SMIv2对IP的MIB
IETF RFC4456 (2006)	BGP路由反射
IETF RFC4502 (2006)	远程网络监控MIB
IETF RFC4632 (2006)	无类域间路由选择 (CIDR): 地址分配及拥塞策略
IETF RFC4750 (2006)	OSPF v2管理信息库
IETF RFC4760 (2007)	BGP4多协议扩展
IETF RFC4762 (2007)	使用标签分发协议 (LDP) 作为信令的虚拟专用局域网业务 (VPLS)
IETF RFC4906 (2007)	MPLS上传送二层帧
ATM FORUM UNI 3.1	ATM论坛UNI规范3.1版
ATM FORUM UNI 4.0	ATM论坛UNI规范4.0版

3 术语、定义和缩略语

3.1 术语和定义

下列术语和定义适用于本标准。

3.1.1

路由器 Router

路由器是通过转发数据包来实现网络互联的设备。路由器可以支持多种网络层协议(例如 TCP/IP 等), 可以在多个网络层次上转发数据包(例如数据链路层、网络层、应用层)。

路由器需要至少拥有一个物理端口, 连接两个或多个由 IP 子网或无编号点到点线路标识的逻辑端口。路由器根据收到的数据包中的网络层地址以及路由器内部维护的路由表, 选择下一跳路由器或主机(最后一跳时)的地址和相关接口, 并重写链路层数据包头。

路由表应动态维护以反映当前的网络拓扑。路由器通常通过与其他路由器交换路由信息来完成动态维护路由表。

路由器可以提供数据包传输服务。为实现路由选择的灵活性和鲁棒性 (Robust), 路由器可使用最少状态信息以维持数据包传输服务。

路由器还可以支持多种业务 (L2/L3 MPLS VPN、组播等)。

3.1.2

核心路由器 Core Router

通常位于网络骨干层，用作扩大互联网的路由处理能力和传输带宽的路由器。在本标准中，要求核心路由器的系统双向交换容量至少达到 60Gbit/s。

3.1.3

协议分层

协议分层通常按照互联网的 5 层结构或者开放系统互联（OSI）7 层参考协议描述。本标准按照互联网 5 层结构来描述。在互联网上的协议分层如下所述：

① 应用层（Application Layer）

应用层位于互联网协议栈中最高层。应用层通常包括 OSI 7 层参考模型中的表达层和应用层的功能，以及会话层的部分功能。

应用层协议可以分为直接为用户提供服务的用户协议和提供通用系统功能的支持协议。用户协议包括 Telnet（远程登录），FTP（文件传输协议），SMTP（简单邮件传递协议）等。支持协议可以包括 SNMP（简单网络管理协议），BOOTP（自举协议），TFTP（简单文件传输协议）和大量的路由协议。

② 传输层（Transport Layer）

传输层协议提供端到端的通信服务。该层协议除完成 OSI 7 层参考模型中传输层功能外，还包含少量的会话层功能。

目前，主要有两种传输层协议：

- 传输控制协议（TCP）；
- 用户数据包协议（UDP）。

TCP 协议提供面向连接的可靠传输服务，保证端到端传输的次序和可靠性，并提供流量控制。UDP 协议提供无连接的传输服务。

③ 互联网（Internet）层

所有互联网传输协议都使用互联网协议（IP）将数据从数据源传送到目的地。IP 是基于无连接或数据包的网际服务，不提供端到端的传递保证。IP 数据包到达时可能损坏、重复、失序或者部分丢失。需要时，IP 层以上的协议层负责可靠的数据传递。IP 协议包含寻址，服务类型规范，分装/重组以及安全性功能。该层次相当于 OSI 参考模型中的网络层。

互联网控制消息协议（ICMP）是一个控制协议。ICMP 位于 IP 层上，封装在 IP 数据包中。ICMP 提供差错报告，拥塞报告和第一跳路由器重定向等功能。

互联网组管理协议（IGMP）是为 IP 组播（multicasting）建立动态主机组的网络层协议。

④ 链路层

链路层协议包含了在物理层之上，网络层之下的所有功能，负责正确传递数据包。

互联网链路层标准通常只描述用于在指定链路层协议上传输 IP 数据包的地址解析原则。

3.1.4

自治系统 Autonomous System（AS）

自治系统（AS）包含一组由一系列路由器互联而成的子网（子网上连接主机），并构成网络拓扑一个可连接的分段。这些子网和路由器一般都由一个单一的操作维护（O&M）管理组织来控制维护。在一个 AS 内，路由器可以使用一个或多个内部路由协议。每个 AS 对外部网络一般都有一个统一的内部路由计

划，并提供精简的可达路由。一个 AS 由 AS 号来标识。

3.1.5

寻址结构

IPv4 数据包包含 32bit 源地址和目的地址，每个地址可以分成两部分：网络地址和主机地址。在正常分发数据包的情况下，最后一个路由器应将主机 IP 地址映射到主机的链路层地址。

① 传统的 IP 地址结构

网络地址划分应符合在 IETF RFC950 以及 IETF RFC796 中的规定。

② 无类域间路由选择 (CIDR)

CIDR 对 32bit 的互联网地址作更有效的利用。实现 CIDR 应符合 IETF RFC4632 的规定。CIDR 主要包含如下 3 个要素：

- 按拓扑结构分配地址；
- 具有汇聚网络层可达信息的路由协议；
- 一致的数据转发算法（“最长匹配”）。

3.1.6

IP 组播 Multicasting

IP 组播是链路层组播的扩展。使用组播技术，一个数据包能传送到多个主机（并非全部主机）。在扩展情况下，这些主机可以在不同的地址域中，这些主机的集合称为组播组。每个组播组由一个 D 类 IP 地址标识。发送给每个组播成员的 IP 数据包具有与 IP 单播业务流相同的服务质量。数据包发送者可以不属于组播组成员。

对 IP 组播的实现应符合 IETF RFC1112。

3.1.7

无编号线路及网络前缀

通常路由器或主机的每个端口都拥有 IP 地址，这种情况造成了 IP 地址的浪费——每个点到点链路都需要网络前缀。为解决这个问题提出了无编号线路（Unnumbered Line）——点到点链路不分配 IP 地址。实现无编号线路可以采用两种方法：

- 将点到点链路两端路由器作为一个虚拟路由器，点到点链路作为虚拟路由器内部总线。
- 将路由器 ID 作为路由器上无编号线路的地址。路由器 ID 使用路由器的一个地址（路由器至少拥有一个地址）。

由于作为虚拟路由器存在非标准性和兼容性问题，本标准建议采用第二种方法。

3.1.8

数据包 Datagram

从源到目的地互联网模块之间的一组传输单元称为数据包。

3.1.9

缺省路由 Default Route

当路由器路由表的条目中没有到某个目的地的明确路由时，用作向任意网络前缀转发的下一跳地址。

3.1.10

密集模式 Dense Mode

组播转发中的一种转发模式：所有的组播数据作为数据链路层组播向除接收端口之外所有端口发送，除非相邻路由器指示不这样做。

3.1.11

转发器 Forwarder

路由器中负责在各接口间交换数据包的逻辑实体。

3.1.12

转发 Forwarding

转发是路由器对每个收到数据包的处理过程。数据包可能由路由器自身接收，也可能送到另一个或多个端口，或者两者皆有。转发包括了决定如何处理数据包的过程——排入队列输出或者由自身接收。

3.1.13

转发信息表 FIB

该表包括转发 IP 数据包所需要的信息，在本标准中称为 FIB。该表中至少包含接口标识和到每一个可达目标网络前缀的下一跳信息。

3.1.14

分段 Fragment

包含上层数据包一部分内容的 IP 数据包，该上层数据包由于太大，不能整个放入输出网络的一个数据包中。

3.1.15

通用串口

一个能连接两个系统的物理媒体，能配置成点到点链路，同样也能支持使用例如 LAPS 协议的链路层网络。链路层网络连接到另一系统或交换机，在连接上可能存在高层复用虚电路通信。参见点到点线路。

3.1.16

内部网关协议 IGP

在自治系统内分发路由信息的协议。

3.1.17

接口 IP 地址 Interface IP Address

赋予路由器一个特定接口的 IP 地址及其网络前缀长度。

3.1.18

互联网地址 Internet Address

在互联网上标识一台主机的数字，包含两部分：IP 地址以及前缀长度。前缀长度是指网络地址的比特数。

3.1.19

IP

由 IETF RFC791 定义的包交换协议。IP 不提供可靠的通讯机制，即没有端到段，端到端的概念。

3.1.20

IP 数据包 IP Datagram

IP 数据包是互联网协议中端到端的传输单元。IP 数据包中包含 IP 头和高层数据（例如 TCP、UDP、ICMP 等）。IP 数据包包括 IP 头以及后面的消息部分。

IP 数据包包含一个或多个 IP 分段。

3.1.21

IP 分段 IP Fragment

IP 分段是 IP 数据包的一部分。IP 分段包括 IP 头以及 IP 数据包中高层信息的部分或全部。

一个或多个 IP 分段组成 IP 数据包。

3.1.22

IP 包 IP Packet

IP 数据包或 IP 分段。

3.1.23

逻辑（网络）接口 Logical [network] Interface

由不同 IP 地址标识，到达相联网络的逻辑路径。

3.1.24

Martian 地址过滤 Martian Filtering

包含无效源或目的地址的数据包称为 Martian，在转发处理中应被丢弃。

3.1.25

最大传输单元 Maximum Transmission Unit (MTU)

可通过逻辑接口收发的最大尺寸数据包。该数值包含 IP 头，不包含链路层帧封装。

3.1.26

组播 Multicast

组播是指在 IP 网络中，一个或多个主机以某个组播地址为标识构成一个集合（组播组），数据发送的源主机只需要将一份数据发送至这个组播地址，则组播组中的所有成员都可以收到这份数据。

3.1.27

组播地址 Multicast Address

由组播主机识别的一种特殊类型地址。

组播地址有时称为功能地址或组地址。

3.1.28

网络前缀 Network Prefix

IP 地址中标识网络的部分。设置地址中表示网络及子网部分的比特数。

3.1.29

始发 Originate

从路由器发出的包有两种，一种是收到后转发的包；另一种是路由器产生的包（例如路由通告）。由路由器产生的包称为始发于路由器。

3.1.30

包 Packet

在互联网层及链路层穿过网络接口传输的数据单元。包括 IP 头和数据。包可以是整个 IP 数据包或者

IP 数据包分段。

3.1.31

路径 Path

数据包从一个路由器到某个特定目标需要穿过的路由器及（子）网的序列。路径是单向的，在一对主机间的路径可能是不同的。

3.1.32

物理网络 Physical Network

物理网络是数据链路层之下的网络。其内部结构（如果存在）对互联网层是透明的。

3.1.33

物理网络接口 Physical Network Interface

连接网络的物理接口，拥有（可能惟一）链路层地址。在同一路由器上的物理网络地址可能共享同一个链路层地址，但同一网络上不同路由器的链路层地址应是惟一。

3.1.34

点到点线路 Point to Point Line

能且仅能连接两个系统的物理媒体。

3.1.35

反向路径转发 RPF

对广播和组播报文指定下一跳目标的方法。

3.1.36

悄悄丢弃 Silently Discard

路由器不作任何进一步处理而丢弃数据包，且不发 ICMP 差错消息。为了进行诊断差错，路由器应提供将差错及包内容写入日志，并对差错计数的能力。

3.1.37

稀疏模式 Sparse Mode

组播转发中的一种转发模式：在进行组播转发时，在初始状态下假设路由器所有连接的网络均不需要接收组播数据，只有当下游网络或设备表示要接收组播数据时才向其进行转发。

3.1.38

子网 Subnet

网络的一部分，在物理上可能是独立的，与网络的其他部分共享同一个网络地址，由子网号区分。

3.1.39

子网号 Subnet Number

互联网地址的一部分，用于区分子网。在互联网路由时不起作用，只用于企业网内路由。

3.1.40

服务类型 TOS

IP 头中的服务类型域，表示传输层或应用层希望网络层提供的可靠程度。

3.1.41

生存时间 TTL

IP 头中的生存时间域，表示数据包的有效期。该值与经过的路由器跳数及定时器值有关。

3.1.42

SDH 上传送 IP IP over SDH

是一种将 IP 与 SDH 网结合起来的数据通信体系结构，其物理层、数据链路层和网络层分别定义为 SDH、LAPS 和 IP 协议层。

3.1.43

LAPS

HDLCL 的一个子集，包括数据链路服务和协议规范，主要用于 SDH 上传送 IP。

3.1.44

WDM 上传送 IP IP over WDM

是一种将 IP 与 WDM 网结合起来的数据通信体系结构。

3.1.45

ATM 业务分类

有 4 种不同的 ATM 业务类型：

1) CBR

CBR 业务类型用于在连接的整个生存时期内需要恒定带宽的任何连接。信源允许在整个 CBR 连接时间内以 PCR 的吞吐量发送信元。CBR 适合需要最小信元延迟变化的实时应用，诸如语音、视频、电路交换等。

2) 实时 VBR

实时 VBR (rt-VBR) 业务类型用于具有突发特性的实时应用。它也适用于需要最小 CDV (信元延迟变化) 和 CTD (信元传输延迟) 的应用，诸如语音和视频。rt-VBR 连接的特征由 PCR (峰值信元速率)、SCR (统计信元速率) 和 MBS (最大突发长度) 共同决定。信源期望以可变速率发送信元。

3) 非实时 VBR

非实时 VBR (nrt-VBR) 业务类型用于具有突发特性的非实时应用。nrt-VBR 连接的特征是由 PCR、SCR 和 MBS 来决定的。使用 nrt-VBR 的应用期望在所协商范围内传输的信元具有低的 CLR (信元丢失率)。nrt-VBR 业务可能支持统计复用。nrt-VBR 业务类型不规定信元延迟的上限。

4) UBR

UBR 业务适用于不需要最小信元传输延迟、最小信元延迟以及最小信元延迟偏差的非实时应用，诸如文件传输或电子信件的计算机通信。UBR 业务不保证诸如 CLP 或 CTD 的 QoS (服务质量)。

3.1.46

MPLS 标记 MPLS label

MPLS 分组头中用于识别流的标记。

3.1.47

MPLS 节点 MPLS node

即运行 MPLS 协议的节点。MPLS 节点将理解 MPLS 控制协议，操作一个或多个 L3 协议并能够转发基于标记的分组；MPLS 节点也可以转发纯 L3 分组。

3.1.48

标记交换路由器 LSR

可以实施MPLS中描述的标记交换控制和转发的网络设备。

3.1.49

转发等价类 Forwarding Equivalence Class

以同一种方式转发的一组L3分组（例如，在同一路径上按同一转发处理对待），因此，转发等价类是可以安全地映射至同一标记上的L3分组集合。注意，由于某些原因可能会使来自单个等价类的分组映射至多个标记上（例如，当不使用流合并时）。

3.1.50

虚拟专用网 Virtual Private Network (VPN)

对连接到骨干网上的站点集合施加某种控制策略，生成站点的子集，当某一子集同时包含两个或更多的站点，且这些站点之间通过骨干网连接具有可达性时，称这个子集为VPN。

当VPN所有站点属于同一企业时，VPN被看作是内联网（intranet）。当VPN站点属于不同企业时，VPN被看作是外联网（extranet）。单个站点可以属于多个VPN，可以同时属于一个内联网或多个外联网。本标准的VPN中，不再区分内联网和外联网。

3.1.51

用户边缘设备 CE

用户边缘设备位于客户网络的边缘，它通过到一个或多个运营商边缘（PE）设备的数据连接链路为用户提供对运营商的接入。这里的连接可以是ATM、帧中继、以太网、PPP以及各种隧道等。

CE设备可以是一台主机、以太网交换机或路由器，通常情况下，CE设备是一台路由器，一个站点可能包含多个路由器，仅将连接到PE的路由器称为CE。CE与直连的PE设备建立路由邻接关系。CE路由器将站点的本地路由广播给PE路由器，并从PE路由器学习远端VPN路由。不同站点的CE路由器之间不能直接交换路由信息。

3.1.52

运营商边缘设备 PE

运营商边缘设备位于运营商网络的边缘，通常是路由器设备。PE路由器使用静态路由、RIPv2、OSPF、或BGP协议与CE路由器交换路由信息。

为了增强VPN的可扩展性，对于PE路由器来说只需维护与其直接相连的VPN路由信息，而不要求PE路由器维护运营商网络中所有VPN的路由信息。

当使用MPLS对VPN业务进行转发以穿越运营商网络时，入口PE路由器的作用相当于入口LSR，而出口PE路由器的作用相当于出口LSR。

3.2 缩略语

下列缩略语适用于本标准。

ACCM	Asynchronous Control Character Map	异步控制字符映射
ANSI	American National Standard Institute	美国国家标准学会
APS	Automatic Protection Switching	自动保护切换
ARP	Address Resolution Protocol	地址解析协议
AS	Autonomous System	自治系统

ATM	Asynchronous Transfer Mode	异步转移模式
BACP	Bandwidth Allocation Control Protocol	带宽分配控制协议
BAP	Bandwidth Allocation Protocol	带宽分配协议
BE	Best Effort	尽力而为
BGP	Border Gateway Protocol	边界网关协议
CHAP	Challenge-Handshake Authentication Protocol	握手认证协议
CIDR	Classless Inter Domain Routing	无类域间路由选择
CIPOA	Classic IP over ATM	ATM上支持传统IP及地址解析协议
CLP	Cell Loss Priority	信元丢失优先级
ECP	Encryption Control Protocol	加密控制协议
EGP	Exterior Gateway Protocol	外部网关协议
FCS	Frame Check Sequence	帧校验序列
FEC	Forwarding Equivalence Class	转发等价类
FIB	Forwarding Information Base	转发信息表
FTP	File Transmission Protocol	文件传输协议
HDLCL	High-Level Data Link Control	高级数据链路控制协议
ICMP	Internet Control Message Protocol	互联网控制消息协议
IGMP	Internet Group Management Protocol	互联网组管理协议
IGP	Interior Gateway Protocol	内部网关协议
IP	Internet Protocol	互联网协议
IPCP	IP Control Protocol	IP控制协议
IPv4	Internet Protocol version 4	互联网协议—第4版
IPXCP	The PPP Internetwork Packet Exchange Control Protocol	网间数据包交换控制协议
IS-IS	Intermediate System to Intermediate System	中间系统—中间系统
L2F	Layer 2 Forwarding	第2层发送协议
L2TP	Layer 2 Tunneling Protocol	第2层隧道协议
LAN	Local Area Network	局域网
LAPS	Link Access Procedure SDH	SDH上的链路接入协议
LCP	Link Control Protocol	链路控制协议
LQM	Link Quality Monitor	链路质量监视
LSP	Label Switched Path	标记交换路径
LSR	Label Switching Router	标记交换路由器
MBGP	Multicast Border Gateway Protocol	组播边界网关协议
MIB	Management Information Base	管理信息库
MPLS	MultiProtocol Label Switching	多协议标记交换
MRU	Maximum Receive Unit	最大接收单元
MTU	Maximum Transmission Unit	最大传输单元

NCP	Network Control Protocol	网络控制协议
NHRP	Next Hop Routing Protocol	下一跳路由协议
NIC	Network Interface Card	网络接口卡
NOC	Network Operation Center	网络运行中心
NTP	Network Time protocol	网络时间协议
O&M	Operation and Maintenance	运行与维护
OOB	Out of Band	带外
OSPF	Open Shortest Path First	开放最短路径优先
PAP	Password Authentication Protocol	密码认证协议
PDH	Plesiochronous Digital Hierarchy	准同步数字系列
PHB	Per-Hop Behavior	逐跳行为
POS	Packet over SONET/SDH	在SONET/SDH上传送IP包
PPP	Point to Point Protocol	点到点协议
PPTP	Point to Point Tunneling Protocol	点到点隧道协议
QoS	Quality of Service	服务质量
RPF	Reverse Path Forwarding	反向路径转发
SDH	Synchronous Digital Hierarchy	同步数字体系
SLA	Service Level Agreement	业务等级协商
SNMP	Simple Network Management Protocol	简单网络管理协议
TCP	Transmission Control Protocol	传输控制协议
TFTP	Trivial file transfer protocol	简单文件传输协议
TOS	Type of Service	服务类型
TTL	Time to Live	生存时间
UDP	User Datagram Protocol	用户数据包协议
UNI	User-Network Interface	用户网络接口
VPLS	Virtual Private LAN Service	虚拟专用局域网业务
VPN	Virtual Private Network	虚拟专用网
VRRP	Virtual Router Runday Protocol	虚拟路由器冗余协议
WAN	Wide Area Network	广域网
WDM	Wavelength division Multiplexing	波分复用
WFQ	Weighted Fair Queuing	加权的公平排队算法
WRED	Weighted Random Early Detection	加权的随机早期探测

4 核心路由器功能要求

4.1 核心路由器功能划分

核心路由器通常应用于骨干网，为骨干网转发数据提供路由处理能力和传输带宽。核心路由器一般采用模块化结构，端口密集，有较高的吞吐量。

核心路由器功能可分为以下几方面：

1) 接口功能要求

该功能用于连接路由器到网络或用于路由器间互联，可以分为局域网接口及广域网接口两种。局域网接口主要包括以太网接口；广域网接口主要包括 SDH、ATM、WDM 等网络接口。

核心路由器也可以提供用于网管的 RS232 接口。

2) 通信协议功能

该功能负责处理通信协议包括 TCP/IP、PPP、ATM、路由协议、MPLS 等协议。

3) 数据包转发功能

该功能主要负责按照路由表内容在各端口（包括逻辑端口）间转发数据包并且改写链路层数据包头信息。

4) 路由信息维护功能

该功能负责运行路由协议并维护路由表。路由协议可以包括 RIPv2、OSPFv2、BGP4、IS-IS、组播等协议。

5) 管理控制功能

路由器管理控制功能包括 5 个功能：SNMP 代理功能，Telnet 服务器功能，本地管理功能，远端监控功能和 RMON 功能。通过 5 种不同的途径对路由器进行控制管理，并且允许记录日志。

6) 安全功能

用于完成数据包过滤、地址转换、访问控制、数据加密、防火墙、地址分配等功能。

7) MPLS 功能

该功能主要负责处于 MPLS 域的 LSR 的标记交换、标记交换路径的建立和数据包的转发。标记交换协议包括 LDP 协议和 RSVP 协议。

8) VPN 功能

该功能主要负责 VPN 的建立、维护和数据转发，包括二层的 Martini、Kompella、VPLS 功能以及三层的 BGP MPLS VPN 和 IPSec VPN 功能。

9) 服务质量功能

该功能负责根据优先级标志对数据包的转发质量进行控制，包括 DiffServ、E-LSP 等。

4.2 核心路由器功能要求

路由器应实现以下基本功能：

1) 实现本标准规定的互联网协议，包括 IP，ICMP 以及其他相关的协议。

2) 对每个连接的网络，路由器应实现该网络所要求的功能。这些功能通常包括：

——IP 数据包的封装/解封装；

——根据该网络所支持的最大数据包大小发送或接收 IP 数据包，该大小是网络最大传输单元 (MTU)；

——将 IP 地址与相应网络的链路层地址相互转换，例如将 IP 地址转换成以太网硬件地址；

——响应网络支持的流量控制和差错指示。

3) 接收及转发数据包，并负责缓冲区管理，拥塞控制以及转发的公平性：

——应能辨认差错状态，并按要求产生 ICMP 差错消息。

——丢弃生存时间 (TTL) 域为 0 的数据包。

——当下一网络 MTU 较小时将数据包分段。

4) 按照路由表信息，为每个 IP 数据包选择下一跳目的地。

5) 支持 OSPFv2、IS-IS 和 RIPv2 等内部网关协议（IGP）与其他同一自治域中路由器交换路由信息及可达性信息。支持 BGP4 外部网关协议与其他自治域交换拓扑信息。

6) 支持基于 LDP 标记分发协议的 MPLS 功能。

7) 提供系统网络管理和控制机制，包括存储/上载配置、诊断、升级、状态报告、异常情况报告及控制等。

8) 提供物理层传输接口和适配功能。

9) 提供组播功能。

10) 提供拥塞控制功能。

11) 提供同步和定时功能。

12) 提供包数、字节数、端口、业务类型等信息统计功能。

13) 安全功能，见第 8 章。

14) VPN 功能，见第 7.9 节。

5 接口类型及特性

5.1 概述

本章规定核心路由器应支持的接口类型以及接口特性。

5.2 接口类型及特性

5.2.1 10/100Base-T 接口（可选）

核心路由器可以支持 10/100Mbit/s 自适应以太网接口。

10Mbit/s 以太网接口应符合 IEEE802.3，物理层接口上采用曼切斯特编码，用 0.85V 和 -0.85V 分别表示“1”和“0”。电缆可采用 10Base-T。

100Mbit/s 以太网接口应符合 IEEE802.3u。100Base-T 技术中可采用 3 类传输介质：100Base-T4、100Base-TX 和 100Base-FX。采用 4B/5B 编码方式。

5.2.2 千兆比以太网接口

路由器应支持吉比特以太网接口（符合 IEEE802.3z）。

1000Mbit/s 以太网物理接口可以支持 1000Base-SX，1000Base-LX 以及 1000Base-T。1000BaseT 接口应符合 IEEE802.3ab。

5.2.2.1 1000Base-SX 接口

1) 1000Base-SX 接口的使用范围见表 1。

表 1 1000Base-SX 接口的使用范围

光纤类型	模宽@850nm（最小满负发送）（MHz·km）	最小范围（m）
62.5μm MMF	160	2~220
62.5μm MMF	200	2~275
50μm MMF	400	2~500
50μm MMF	500	2~550

2) 1000Base-SX 接口的发送光功率见表 2。

表 2 1000Base-SX 接口的发送光功率

项 目	62.5μm MMF	50μm MMF	单 位
发送器类型	短波激光器		
信令速度（范围）	1.25±100ppm		GBd
波长（λ，范围）	770~860		nm
T _{rise} /T _{fall} （最大值；20%~80%；>830nm）	0.26		ns
T _{rise} /T _{fall} （最大值；20%~80%；≤830nm）	0.21		ns
RMS 谱宽（最大值）	0.85		nm
平均发送光功率（最大值）	a		dBm
平均发送光功率（最小值）	-9.5		dBm
发送器 OFF 时平均发送光功率（最大值）	-30		dBm
消光比（最小值）	9		dB
RIN（最大值）	-117		dB/Hz
耦合功率比（CPR）（最小值）	9<CPR		dB
a： 最大发送功率应取最大接收功率与 IEEE803.2 规定的 1 类安全限中的小值			

3) 1000Base-SX 接口的接收要求见表 3。

表 3 1000-BaseSX 接口的接收要求

项 目	62.5μm MMF	50μm MMF	单 位
信令速度（范围）	1.25±100ppm		GBd
波长（λ，范围）	770~860		nm
平均接收光功率（最大值）	0		dBm
接收灵敏度	-17		dBm
回损（最小值）	12		dB
加强接收灵敏度（最大值）	-12.5	-13.5	dBm
纵向眼图闭合代价	2.60	2.20	dB
接收电信号 3dB 高端截止频率（最大值）	1500		MHz

5.2.2.2 1000Base-LX 接口

1) 1000Base-LX 接口的使用范围见表 4。

表 4 1000Base-LX 接口的使用范围

光纤类型	模宽@850nm（最小满负发送）（MHz·km）	最小范围（m）
62.5μm MMF	500	2~550
50μm MMF	400	2~550
50μm MMF	500	2~550
9μm SMF	N/A	2~5000

2) 1000Base-LX 接口的发送光功率见表 5。

表 5 1000Base-LX 接口的发送光功率

项 目	62.5μm MF	50μm MMF	9μm SMF	单 位
发送器类型	长波激光器			
信令速度（范围）	1.25±100ppm			GBd
波长（λ，范围）	1270~1355			nm
T _{rise} /T _{fall} （最大值；20%~80%；>830nm）	0.26			ns
RMS 谱宽（最大值）	4			nm

表 5（续）

项 目	62.5μm MF	50μm MMF	9μm SMF	单 位
平均发送光功率（最大值）	-3			dBm
平均发送光功率（最小值）	-11.5	-11.5	-11.0	dBm
发送器 OFF 时平均发送光功率（最大值）	-30			dBm
消光比（最小值）	9			dB
RIN（最大值）	-120			dB/Hz
耦合功率比（CPR）（最小值）	28<CPR<40	12<CPR<20	N/A	dB
注：对 MMF 模式，需使用一段 SMF 模式调节尾纤				

3）1000Base-LX 接口的接收要求见表 6。

表 6 1000Base-LX 接口的接收要求

项 目	62.5μm MF	50μm MMF	9μm SMF	单 位
信令速度（范围）	1.25±100ppm			GBd
波长（λ，范围）	1270~1355			nm
平均接收光功率（最大值）	-3			dBm
接收灵敏度	-19			dBm
回损（最小值）	12			dB
加强接收灵敏度（最大值）	-14.4			dBm
纵向眼图闭合代价	2.60			dB
接收电信号 3dB 高端截止频率（最大值）	1500			MHz

5.2.2.3 1000Base-SX 及 1000Base-LX 抖动规范

1000Base-SX 及 1000Base-LX 抖动规范见表 7。

表 7 1000Base-SX 及 1000Base-LX 抖动规范

参考点 ^b	总抖动 ^a		确定抖动	
	UI	Ps	UI	ps
TP1	0.240	192	0.100	80
TP1-TP2	0.284	227	0.100	80
TP2	0.431	345	0.200	160
TP2-TP3	0.170	136	0.050	40
TP3	0.510	408	0.250	200
TP3-TP4	0.332	266	0.212	170
TP4	0.749	599	0.462	370
注 a：总的抖动包括确定抖动和自由抖动。表中规定的值为 637kHz 以上的高频抖动，不包括低频抖动。表中黑体字参数为必选值，其他参数为参考值。				
注 b：参考点的定义见 IEEE802.3z				

5.2.2.4 1000Base-T 接口

1000Base-T 接口应符合 IEEE802.3ab。

5.2.2.5 以太网层要求

- 1) 支持 MAC 层全双工操作；
- 2) 支持 8B/10B 编/解码；
- 3) 连接器类型：SC 双工；

4) 具有转发 64byte 长的以太网帧的能力。

5.2.2.6 分组层要求

支持以太网接口上的 MPLS (MPLS over Gigabit Ethernet) 。

5.2.2.7 软件特性

- 1) 在最高速接口上应支持长达 15min 的比特计数而不溢出。
- 2) 支持流控 (发送) 。
- 3) 支持优先级设定/映射。
- 4) 当用于局域网时, 应支持虚拟冗余路由器协议 (VRRP), 并符合 IETF RFC3768 的要求。
- 5) 当用于局域网时, 应支持 IEEE802.1P/Q。

5.2.3 10G 以太网接口 (可选)

5.2.3.1 概述

路由器可选支持 10G 以太网接口 (符合 IEEE802.3ae)。

10G以太网物理接口可以支持 10GBase-R、10GBase-W。

5.2.3.2 10GBase-S 接口

- 1) 10GBase-S 接口的使用范围见表 8。

表 8 10GBase-S 接口的使用范围

光纤类型	最小模宽@850nm (MHz·km)	最小范围 (m)
62.5μm MMF	160	2~26
	200	2~33
50μm MMF	400	2~66
	500	2~80
	2000	2~300

- 2) 10GBase-S 接口的发送光接口参数见表 9。

表 9 10GBase-S 接口的发送光接口参数

项 目	10GBase-SW	10GBase-SR	单 位
信号速率	9.95328	10.3125	GBd
信号速率最大偏差	±20	±100	ppm
波长 (λ, 范围)	840~860		nm
平均发送光功率 (最大值)	-1.0		dBm
平均发送光功率 (最小值)	-7.3		dBm
发送器 OFF 时平均发送光功率 (最大值)	-30		dBm
消光比 (最小值)	3		dB
回损 (最小值)	12		dB

- 3) 10GBase-S 接口的接收要求见表 10。

表 10 10GBase-S 接口的接收要求

项 目	10GBase-SW	10GBase-SR	单 位
信号速率	9.95328	10.3125	GBd
信号速率最大偏差	±100		ppm
波长 (λ, 范围)	840~860		nm
平均接收光功率 (最大值)	-1.0		dBm

表 10（续）

项 目	10GBase-SW	10GBase-SR
平均接收光功率（最小值）	-9.9	dBm
接收灵敏度	-11.1	dBm
加强接收灵敏度（最大值）	-7.5	dBm
纵向眼图闭合代价	3.5	dB
接收电信号 3dB 高端截止频率（最大值）	12.3	GHz

5.2.3.3 10GBase-L 接口

1) 10GBase-L 接口的使用范围见表 11。

表 11 10GBase-L 接口的使用范围

PMD 类型	正常波长（nm）	最小范围（m）
10GBase-L	1310	2~10000

2) 10GBase-L 接口的发送光接口参数见表 12 所示。

表 12 10GBase-L 接口的发送光接口参数

项 目	10GBase-LW	10GBase-LR	单 位
信号速率	9.95328	10.3125	GBd
信号速率最大偏差	±20	±100	ppm
波长（λ，范围）	1260~1355		nm
最小边模抑止比	30		dB
平均发送光功率（最大值）	0.5		dBm
平均发送光功率（最小值）	-8.2		dBm
发送器 OFF 时平均发送光功率（最大值）	-30		dBm
消光比（最小值）	3.5		dB
回损（最小值）	12		dB

3) 10GBase-L 接口的接收要求见表 13。

表 13 10GBase-L 接口的接收要求

项 目	10GBase-LW	10GBase-LR	单 位
信号速率	9.95328	10.3125	GBd
信号速率最大偏差	±100		ppm
波长（λ，范围）	1260~1355		nm
平均接收光功率（最大值）	0.5		dBm
平均接收光功率（最小值）	-14.4		dBm
接收灵敏度	-12.6		dBm
加强接收灵敏度（最大值）	-10.3		dBm
纵向眼图闭合代价	2.2		dB
接收电信号 3dB 高端截止频率（最大值）	12.3		GHz

5.2.3.4 10GBase-E 接口

1) 10GBase-E 接口的使用范围见表 14。

表 14 10GBase-E 接口的使用范围

PMD 类型	正常波长（nm）	最小范围
10GBase-E	1550	2m~30km

2) 10GBase-E 接口的发送光接口参数见表 15。

表 15 10GBase-E 接口的发送光接口参数

项 目	10GBase-EW	10GBase-ER	单 位
信号速率	9.95328	10.3125	GBd
信号速率最大偏差	±20	±100	ppm
波长 (λ, 范围)	1530~1565		nm
最小边模抑止比	30		dB
平均发送光功率 (最大值)	4.0		dBm
平均发送光功率 (最小值)	-4.7		dBm
发送器 OFF 时平均发送光功率 (最大值)	-30		dBm
消光比 (最小值)	3.0		dB
回损 (最小值)	21		dB

3) 10GBase-E 接口的接收要求见表 16。

表 16 10GBase-E 接口的接收要求

项 目	10GBase-EW	10GBase-ER	单 位
信号速率	9.95328	10.3125	GBd
信号速率最大偏差	±100		ppm
波长 (λ, 范围)	1530~1565		nm
平均接收光功率 (最大值)	-1.0		dBm
平均接收光功率 (最小值)	-15.8		dBm
接收灵敏度	-14.1		dBm
加强接收灵敏度 (最大值)	-11.3		dBm
纵向眼图闭合代价	2.7		dB
接收电信号 3dB 高端截止频率 (最大值)	12.3		GHz

5.2.3.5 10GBase-LX4 接口

1) 10GBase-LX4 接口的使用范围见表 17。

表 17 10GBase-LX4 接口的使用范围

光纤类型	最小模宽@850nm (MHz·km)	最小范围 (m)
62.5μm MMF	500	2~300
50μm MMF	400	2~240
50μm MMF	500	2~300
10μm SMF	n/a	2~10000

2) 10GBase-LX4 接口的发送光接口参数见表 18。

表 18 10GBase-LX4 接口的发送光接口参数

项 目	62.5μm 和 50μm MMF	10μm SMF	单 位
每通道的信号速率	3.125		GBd
信号速率最大偏差	±100		ppm
波长 (λ, 范围)	1269.0~1282.4 1293.5~1306.9 1318.0~1331.4 1342.5~1355.9		nm
Trise/Tfall (最大值; 20%~80%相应时间)	120		ps

表 18（续）

项 目	62.5μm 和 50μm MMF	10μm SMF	单 位
最小边模抑止比	0.0		dB
RMS 谱宽（最大值）	0.62		nm
四通道最大发送光功率	5.5		dBm
每通道最大发送光功率	-0.5		dBm
每通道发送器 OFF 时平均发送光功率（最大值）	-30		dBm
消光比（最小值）	3.5		dB

3) 10GBase-LX4 接口的接收要求见表 19。

表 19 10GBase-LX4 接口的接收要求

项 目	62.5μm 和 50μm MMF	10μm SMF	单 位
每通道的信号速率	3.125		GBd
信号速率最大偏差	±100		ppm
波长（λ，范围）	1269.0~1282.4 1293.5~1306.9 1318.0~1331.4 1342.5~1355.9		nm
四通道最大发送光功率	5.5		dBm
每通道最大发送光功率	-0.5		dBm
每通道接收灵敏度	-14.25	-14.45	dBm
加强的每通道接收灵敏度	-10.5	-13.4	dBm
每通道发送器 OFF 时平均发送光功率（最大值）	-30		dBm
消光比（最小值）	3.5		dB
回损	12		dB

5.2.3.6 以太网层要求

- 1) 支持 MAC 层全双工操作；
- 2) 支持 8B/10B 或 64B/66B 编/解码；
- 3) 连接器类型：SC/LC 等；
- 4) 具有转发 64byte 长的以太网帧的能力。

5.2.3.7 分组层要求

支持 10G 以太网接口上的 MPLS。

5.2.4 SDH 接口

5.2.4.1 接口类型

核心路由器应该支持 SDH STM-1 接口、SDH STM-4 接口、SDH STM-16 接口和 STM-64 接口中的一种或者几种。

STM-1 有光接和电接口两种，STM-1 电接口适用于局内，干扰信号弱的情况。STM-4，STM-16 和 STM-64 应采用光接口。

5.2.4.2 SDH 层要求

- 应符合 YDN 099-1998 和 ITU-T G.707。
- 应支持以下告警处理功能：LOS、LOF、LAIS、PAIS、LOP、SF、SD。

- 应支持性能监控。
- 应支持 B1, B2, B3 差错计数。
- 应支持本地（内部）或环路定时（从网络恢复时钟），并至少具有 20ppm 的时钟精度。
- 应支持保护倒换和本地环回（诊断）和网络环回功能。

5.2.4.3 链路层要求

SDH 接口上的链路层协议具体要求见第 6.4 节中的规定。

5.2.4.4 分组层要求

- 具有转发 40byte 长的 IP 包的能力。
- 支持 QoS 或 CoS。
- 基于 IP 的拥塞管理。

5.2.4.5 软件特性

- 通过调度排队（scheduling）算法提供 QoS（或 CoS）。
- 支持 POS 上的 MPLS（MPLS over POS）。

5.2.5 ATM 接口（可选）

5.2.5.1 接口类型

核心路由器应至少支持 ATM 155Mbit/s 接口和 ATM 622Mbit/s 光接口。ATM 155Mbit/s 接口分光接口和电接口两种，电接口适用于局内，干扰信号弱的情况。

ATM 155Mbit/s 接口和 ATM 622Mbit/s 光接口具体要求见 YDN 067-1998。

此外，作为任选，核心路由器还可以支持 ATM 2.5Gbit/s 光接口，具体要求待定。

5.2.5.2 分组层要求

- 具有转发 40byte 长的 IP 包的能力。
- 每个物理端口支持 4 个队列。
- 输出队列应支持对每条 VC 分段和业务整形功能。
- 支持 ATM 上的 MPLS（MPLS over ATM）。
- 应能在独立的 VC 上，或在单条 VC 上排队发送属于同一个下一跳路由器的区分业务流。

5.2.5.3 ATM 层要求

- 支持 PVC 和 SVC。
- 支持 AAL5，支持 CBR, UBR 和 VBR 业务，支持业务量整形。
- 支持 IETF RFC2684 规定的 AAL5 上的多协议封装。
- 支持 LLC/SNAP 和 IP 复用 PVC（路由协议的 LLC 封装）。
- 支持 F4/F5 OAM 信元处理。

5.2.6 WDM 接口（可选）

作为任选，核心路由器可以支持 WDM 接口。

有关 WDM 接口的具体要求见 YDN 120-1999。

6 通信规程

6.1 概述

本章规定核心路由器与广域网（WAN）相连接时应实现的通信规程的基本要求。

6.2 ATM 协议（可选）

为了与 ATM 交换机互联，路由器应支持 ATM 协议。

6.2.1 PVC 连接

路由器应支持 ATM PVC 连接，采用 AAL5 适配层，支持 ATM CBR、UBR 和 VBR 业务，支持业务量整形。具体要求应符合 YDN 067-1998。

6.2.2 SVC 连接

6.2.2.1 UNI 信令要求

1) 信令栈

用户—网络接口采用 ITU-T 建议时的信令标准协议栈如图 1 所示。其中 SSCF (UNI) 为 UNI 业务特定协调功能；SSCOP 为业务特定面向连接协议；CP-AAL 为 ATM 适配层公共部分；DSS2 为 2 号数字用户信令系统。

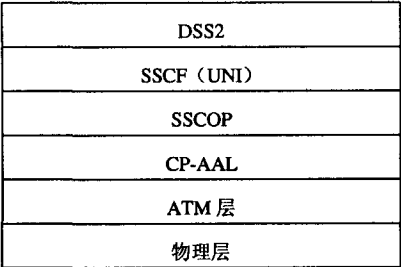


图 1 用户—网络接口协议层间的关系

2) 物理层的要求

UNI 的物理层应满足 YD/T 976-1998 的规定。

3) ATM 层的要求

UNI 的 ATM 层应满足 YDN 052-1998 的规定。

4) SAAL 的要求

UNI 的 SAAL 由 ATM 适配层公共部分 (CP-AAL)、业务特定协调功能 (SSCF)、业务特定面向连接协议 (SSCOP) 组成，它应满足 YDN 053.4-1998、YDN 083.1-1998、YDN 083.3-1998 的规定。

5) DSS2 的要求

DSS2 完成用户—网络接口呼叫/连接控制功能的程序。路由器应满足 YDN 081-1998 和 YDN 082-1998 的要求。YDN 081-1998 和 YDN 082-1998 与 ATM FORUM UNI 3.1 兼容，由于现有的某些设备暂时仅实现 UNI 3.1 规范要求的功能，故在初始阶段，该技术规定中包含的，而 ATM 论坛 UNI 3.1 中不包含的功能可作为任选特性。

6.3 ATM 上支持传统 IP 及地址解析协议 (CIPOA)（可选）

路由器可以支持 CIPOA 协议。

CIPOA 协议应符合 IETF RFC1483、IETF RFC2225 和 IETF RFC1755。

在 CIPOA 的方案中，IP 分组可以采用 LLC/SNAP 封装和基于 VC 的复用，其中 LLC/SNAP 封装是缺省封装，为了减少封装开销，则可以采用基于 VC 的复用。

CIPOA 网络部件示意图如图 2 所示。ATMARP 服务器为网络中的逻辑部件，其软件可在网络中的路由器、交换机等设备上实现。ATMARP 服务器应冗余备份，并使相互间在数据上同步。

ATMARP 客户机为网络中的逻辑部件，其软件可在用户侧的 ATM 终端、路由器等设备上实现。

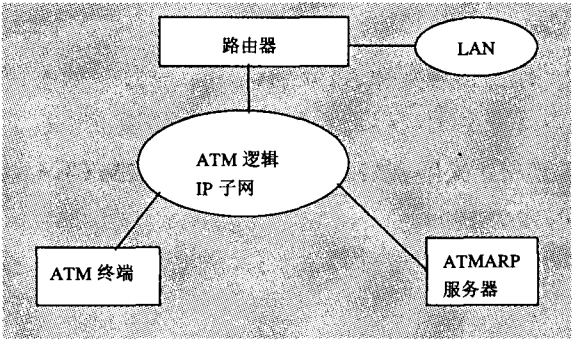


图 2 CIPOA 网络部件示意图

CIPOA技术的封装格式见图3所示，MTU的缺省长度是9180byte。

LLC 0xAA-AA-03
OUI 0x00-00-00
EtherType 0x08-00
IP PDU (最大 2^{16} - 9 个 byte)

图 3 用于承载 IP 数据包的 AAL5 CPCS-PDU 净荷的格式

CIPOA 应实现逻辑子网，在逻辑 IP 子网（LIS）方案中，每个独立的管理实体在一个封闭的逻辑 IP 子网内配置它的主机和路由器。每个 LIS 的运行和通信独立于同一个 ATM 网络中的其他 LIS。与 ATM 网络相连接的主机直接与同一 LIS 内的其他主机进行通信。与本地 LIS 以外主机的通信通过 IP 路由器进行。该路由器是一个与 ATM 网络相连接的 ATM 端点，并被配置成一个或多个 LIS 的成员。该配置导致在同一个 ATM 网络上运行多个分离的 LIS。属于不同 IP 子网的主机之间应通过一个中间 IP 路由器进行通信，即使在 ATM 网络上，这两个 IP 成员间可以建立一条直接的 VC 连接。在提供 SVC 的宽带网中，每个 LIS 中设有一个 ATM ARP 服务器，负责本 LIS 内所有成员的 IP/ATM 地址解析。

在一个 ATM LIS 配置中运行的 IP 成员（主机、路由器）的要求如下：

- 1) 所有成员具有同样的 IP 网络/子网号码和地址掩码；
- 2) 一个 LIS 内的所有成员都直接连接到 ATM 网络；
- 3) LIS 以外的所有成员通过路由器进行访问；
- 4) 当使用 SVC 的时候，一个 LIS 内的所有成员应能完成 ATMARP 功能，从目的 IP 地址中解析出目的 ATM 地址。一个 LIS 内的所有成员还应能完成 InATMARP 功能，在 ARP 服务器中进行地址登记；
- 5) 当使用 PVC 的时候，一个 LIS 的所有成员应能完成 InATMARP，从与之相连接的 VC 中解析出所有与该成员相连接的所有其他成员的 IP 地址；
- 6) LIS 内的所有成员应能够通过 ATM 与该 LIS 内的其他所有成员进行通信，即连接该子网成员的底层虚连接拓扑结构是全网状的。

每个与 ATM 网络相连接的 IP 端站中应实施一组特定的 ATM 参数：

- 1) ATM 硬件地址（atm\$ha）：每个 IP 端站的 ATM 地址。
- 2) ATMARP 请求地址（atm\$arp-req）：atm\$arp-req 是每个 LIS 内的 ATMARP 服务器的 ATM 地址。在 SVC 环境下，向该地址发送 ATMARP 请求，用来从目的协议地址中解析出的 ATM 地址。该服务器应有权负责解析该 LIS 内所有成员的 ATMARP 请求。如果 LIS 仅进行 PVC 操作，那么该参数被置为空并且 IP 端站不需要送 ATMARP 请求给 ATMARP 服务器。

在传统 IP 模型中，ATM 地址使用 E.164 UNI 地址，在 ATMARP 协议中，把 ATM 地址当作“硬件

地址”。

为了支持多个 LIS 之间的通信，CIPOA 功能中应实现 IETF RFC2332 中规定的 NHRP 协议。

有关 CIPOA 协议的详细要求见附录 B。

6.4 SDH 上传送 IP 的协议（IP over SDH）（必选）

核心路由器应支持 SDH 上传送 IP 的协议。

核心路由器应同时支持 IETF RFC1619/IETF RFC2615 以及 YD/T 1061-2000。

6.4.1 IETF RFC1619/ IETF RFC2615

SDH 上传送 IP（IP Over SDH）以 SDH 网络作为 IP 数据网络的物理传输网络，它使用链路适配及成帧协议对 IP 数据包进行封装，然后按字节同步的方式把封装后的 IP 数据包影射到 SDH 的同步净荷封装（SPE）中。

有关 IETF RFC1619/IETF RFC2615 规定的 SDH 上传送 IP 的技术要求见附录 A。

6.4.2 SDH 上传送 IP 的 LAPS 技术规范

具体要求见 YD/T 1061-2000。

6.5 WDM 上传送 IP 的协议（IP over WDM）（可选）

如果核心路由器支持 WDM 接口，核心路由器应支持 WDM 上传送 IP 的协议。

7 协议要求

7.1 概述

本章规定核心路由器实现的链路层、互联网层、传输层、路由协议、MPLS 等协议的基本要求。

7.2 链路层协议

7.2.1 链路层/互联网层接口要求

对每个收到的数据包，链路层应传输到上层的信息如下所示：

- 1) IP 包；
- 2) 链路层数据包的数据部分长度；
- 3) 收到该数据包的物理接口标识；
- 4) 数据包目的地物理地址的分类：单播包、广播包或组播包；
- 5) 源物理地址。

如果将要传送的数据包引起了链路层协议相关的差错，链路层应通知互联网层。

对每个需要传输的数据包，互联网层应提供：

- 1) IP 包；
- 2) IP 包长度；
- 3) 目的地物理地址；
- 4) 下一跳 IP 地址；
- 5) 链路层优先级值。

7.2.2 链路层附加要求

7.2.2.1 地址解析协议-ARP

实现 ARP 的路由器应符合 IETF RFC1122 中 ARP 部分。

如果仅因为 ARP 缓存中没有相应的目的地址，链路层不允许报告目的地不可达差错；链路层应在执

行 ARP 请求/响应序列时缓存数据包，只有在请求无结果后才报告目的地不可达。

如果 ARP 请求结果指示另一主机或路由器地址是广播或组播地址，路由器应丢弃该结果。

7.2.2.2 最大传输单元-MTU

每个逻辑端口的 MTU 值应在该接口合法的 MTU 值范围内配置。

由于一些链路层协议规定了可以发送的最大帧尺寸，在这种情况下，路由器不允许发送超过链路层最大帧尺寸的 MTU。路由器应允许接收尺寸大于 MTU 的链路层帧。

7.2.2.3 点到点协议-PPP

7.2.2.3.1 概述

对点到点协议的实现应符合 IETF RFC1661、IETF RFC1332、IETF RFC1334 和 IETF RFC1994。

点到点接口是使用点到点线路传输数据的接口。这样的接口包括专线和复合接口等。复合接口通常使用特殊的物理接口。

实现点到点接口或通用串行接口的路由器应实现 PPP。

路由器应在所有的通用串行接口上实现 PPP。路由器可以允许线路配置成点到点链路协议来替代 PPP。点到点链路接口应缺省使用 PPP（当使能时），在使能以前要求配置链路层协议。通用串行接口应要求在使能以前配置链路协议。

7.2.2.3.2 LCP 任选域

LCP 协议提供一系列可以协商的任选域，这些任选域包括地址和控制字段压缩、协议字段压缩、异步控制字符映射（ACCM）、最大接收单元（MRU）、链路质量监视（LQM）、幻数（用于环回检测）、PAP、CHAP 和 32 位的 FCS。

路由器可以在同步或异步链路上使用地址和控制字段压缩，协议字段压缩。如果路由器指示它能接收这些压缩，也应能接收非压缩的 IP 头信息。

路由器可以在异步 PPP 链路上协商 ACCM，但不应在同步 PPP 链路上协商 ACCM。如果路由器在同步 PPP 链路上收到一个 ACCM 协商试图，它应确认这个任选域，然后不理睬它。

路由器应正确协议最大接收单元（MRU）。如果路由器协商的 MRU 小于 1500byte，它应有能力接收一个 1500byte 的帧。

路由器不应支持链路质量监视（LQM）。

路由器应实现和协商用于环回检测的幻数。

路由器应支持 PAP 和 CHAP。

路由器应支持 16 位的 FCS，可以支持 32 位的 FCS。

有关 LCP 协议的具体要求见附录 A。

7.2.2.3.3 IPCP 任选域

路由器可以执行 IP 地址协商。如果对端不支持 IP 地址协商规程，路由器应采取正确的措施。

有关 IPCP 协议的具体要求见附录 A。

7.2.2.4 接口测试

路由器应提供一种机制允许路由软件决定一物理接口是否可用；在复合接口上，路由器通常需要判断虚电路是否可用。路由器应提供一种机制允许路由软件判断物理接口质量。路由器应提供一种机制来通知路由软件：由于管理原因接口是否可用。路由器应提供一种机制当检测到链路层接口可用或不可用

时应通知路由软件。

7.2.3 SDH 接口上的链路层协议

SDH 接口上的链路层协议具体要求见第 6.4 节中的规定。

7.2.4 WDM 接口上的链路层协议（可选）

WDM 接口上的链路层协议具体要求见第 6.5 节中的规定。

7.2.5 ATM 接口上的链路层协议（可选）

ATM 接口上的链路层应用 AAL5，其协议应符合 YDN 067-1998 的规定。

7.3 互联网层协议

本节描述的互联网协议包括 IP、ICMP、IGMP。

7.3.1 互联网协议-IP

7.3.1.1 定义

路由器应实现 IP 协议，并符合 IETF RFC791。路由器应实现与 IP 相关的子网（符合 IETF RFC950），IP 广播（符合 IETF RFC922）和无类域间路由选择（符合 IETF RFC4632）。

在某些情况下，要求路由器在丢弃数据包时不作任何处理（即不发 ICMP 差错消息），但是为了诊断故障，路由器应提供将差错写入日志（包括所丢弃数据包的内容）以及对丢弃数据包进行计数的能力。

7.3.1.2 协议概述

IP 协议在 IETF RFC791 中规定。

7.3.1.2.1 协议选项域

路由器收到数据包时，IP 层应解释它理解的 IP 选项域，将剩下的内容留给高层协议处理。

IP 层应为高层协议提供下列内容：

1) 安全性选项域（可选）

在某些环境下，每个始发或接收的数据包都要求包含安全性选项域。路由器可以任选实现 IETF RFC1108 中规定的安全性选项域。

2) 流标识选项域（可选）

已停止使用该选项域，路由器不应在发出的数据包中加入该选项域。路由器也应不理睬接收到的包中的该选项域。

3) 源路由选项域（可选）

路由器应能作为一源路由的最终目的地。如果路由器收到一个包含完整源路由的数据包，该数据包就已到达最终目的地。在这种情况下，指针指向最后字段之后，IP 头中的目的地址定址该路由器。收到该选项域时应传送到传输层协议来处理（或到 ICMP 消息来处理）。

一般情况下，对源路由数据包的正确响应是横截相同的路由。路由器应能够反转接收到的数据包中的源路由，并把它插入到路由器始发的数据包中。

当创建源路由选项域时，路由器应能够正确处理。

路由器应提供关闭源路由选项的功能。

4) 路由记录选项域（可选）

路由器可以在发送的数据包中支持路由记录选项域。

5) 时间戳选项域（可选）

路由器可以在发送的数据包中支持时间戳选项域，并使用以下规则：

① 当路由器发出一个带时间戳选项域的数据包时，如果出现下列情况，路由器应将时间戳记录在选项域中：

- 如果没有预先指定互联网地址域；
- 第一个预先指定的 IP 地址是即将发送该数据包端口的逻辑地址。

② 如果路由器收到一个带时间戳选项域的数据包，路由器应在传递给传输层协议或 ICMP 处理之前将当前时间插入到时间戳选项域中（如果还有空间）。如果选项域中没有空间，路由器应将选项域中的溢出计数器加 1。

③ 时间戳的值应符合 IETF RFC1122。

7.3.1.2.2 选项域中的地址（可选）

要求路由器将地址插入记录路由、严格源和记录路由、松散源和记录路由或时间戳选项域中。路由器应插入发送该数据包的逻辑接口的地址。当发包接口没有 IP 地址时（例如非编号接口），路由器应插入路由器 ID。路由器 ID 是路由器 IP 地址之一，可以是某接口地址。路由器 ID 应基于系统或链路指定。除非网络管理员修改，路由器 ID 一般不能改变（即使在重新启动之后）。拥有多个非编号接口的路由器可以拥有多个路由器 ID。每个非编号接口应与一特殊的路由器 ID 相关联。这种关联不能改变，除非重新配置了路由器。

7.3.1.2.3 IP 头中的未使用比特

IP 头包含两个未使用比特：一个在服务类型字节中，另一个在标志域中。在路由器产生的数据包中，不允许将上述任何比特置为 1。不允许路由器仅仅因为上述任何比特置为 1 而丢弃（即拒绝接收或拒绝转发）数据包，即路由器不应检查上述比特的值。

7.3.1.2.4 服务类型

IP 头中服务类型字节分为 3 个部分：优先级字段（最高 3byte），服务类型字段（TOS）（后续 4byte）以及保留比特（最后 1 个 byte）。

路由器不应实现 IETF RFC795 中规定的服务映射功能。

7.3.1.2.5 头校验和

IP 应核实收到的任一数据包的 IP 校验和，应丢弃包含无效校验和的数据包。路由器不允许提供关闭校验和核实功能的方法。

当 IP 头中仅 TTL 改变时，路由器可以使用递增的 IP 头校验和更新，这可以降低因路由器引起 IP 头毁坏的可能性。

7.3.1.2.6 不可识别的头选项域

路由器应不理睬不可识别的头选项域，即路由器应实现选项域列表结尾选项域和无操作选项域，因为上述选项域不包含长度。

7.3.1.2.7 分段

路由器应支持分段，分段应符合 IETF RFC791。

当路由器将 IP 包分段时，它应尽量减少分段数，并按顺序发送。当分段方法可能产生长度明显小于其他分段的一个分段时，则该分段为第 1 个 IP 分段，并首先发送。

7.3.1.2.8 重组

路由器应支持将发送给自己的分段 IP 包重组。

重组应符合 IETF RFC1122。

7.3.1.2.9 生存时间 (TTL)

路由器对产生或收到的 IP 包 TTL 的处理应按照 IETF RFC1122。

7.3.1.2.10 多子网广播

路由器不应支持针对所有子网广播，见 7.4.2.4.3 节。

7.3.1.2.11 地址

IP 地址分 5 类，从 A 类到 E 类。D 类用作 IP 组播，E 类保留。A、B、C 类地址一般用作单播网络前缀。

IP 组播地址是 28 比特的逻辑地址，它表示一组主机，可以是永久的或临时的。永久组播地址在 IETF RFC1700 中规定。临时地址可以为临时组动态分配，组员由 IGMP (IETF RFC1112) 动态决定。

使用如下方法表示一个单播 IP 地址：

{网络前缀，主机号}

其中：— 1 表示相应区域全 1byte；

— 0 表示相应区域全 0byte。

a) {0, 0}

本网络上的本主机。除在初始化规程外（例如路由器使用 BOOTP 调取配置信息），该地址不能用作路由器产生的数据包의 源地址。

如果路由器从本地分发收到上述数据包，且路由器实现了明确指出采取相应动作的相关协议时，路由器应接收该数据包，否则路由器应丢弃任何从本地分发处得到的源地址为 {0, 0} 的数据包。

b) {0, 主机号}

本网络上的指定主机。除在路由器初始化规程中作为源地址学习自身的 IP 地址外，路由器不能发送 {0, 主机号} 地址。

c) {-1, -1}

有限广播，不能用作源地址。相连物理网络上所有主机和路由器可以接收目的地址使用 {-1, -1} 的数据包，但不能转发到该网络之外。

d) {网络前缀, -1}

定向广播——对指定网络前缀的广播，不能用作源地址。路由器可以产生网络定向广播数据包，应接收网络定向广播数据包。路由器可以包含一个可配置的选项域允许路由器接收网络定向广播数据包。然而，该选项域应缺省设置成不允许接收网络定向广播数据包。这样路由器不必接收网络定向广播的数据包，除非端用户特别配置，具体见 IETF RFC2644。

e) {127, 任意值}

互联网主机环回地址，该形式地址不允许应用在主机之外。

当路由器产生 IP 包时，IP 源地址应是该路由器地址的其中之一（但不能是广播或组播地址）。惟一的例外是在路由器初始化期间。

在大多数情况下，目的地址为广播地址或组播地址的数据包的处理应与发往路由器某地址的数据包一样，即：

- 路由器应接收, 且正确处理带有一个广播目的地址的任何数据包;
- 路由器应接收, 并正确处理带有一个组播目的地址的任何数据包, 该数据包是要求路由器接收的。

路由器应丢弃源地址包含本节中所描述无效地址的数据包。这种检查可以在 IP 层或传输层进行。路由器应对因此丢弃的数据包进行计数。

7.3.1.3 特定规定

7.3.1.3.1 IP 广播地址

路由器应符合:

1) 将 255.255.255.255 或{网络前缀, -1}作为 IP 广播地址。

2) 应丢弃(即不传递到路由器上层应用)发给 0.0.0.0 或{网络前缀, 0}的数据包。如果不丢弃, 则这些数据包应作为 IP 广播包来处理。可以设置一个可更改的配置选项域来决定是否接受上述数据包。缺省情况是丢弃上述数据包。

3) 在直连网络/子网中发起 IP 广播时, 缺省情况下, 应使用有限广播地址(255.255.255.255)(发送 ICMP 地址掩码应答时除外)。路由器应接收有限广播。

4) 不应产生源地址为 0.0.0.0 或{网络前缀, 0}的数据包。可以设置一个可更改的配置选项域决定是否允许产生上述数据包。缺省情况是不产生上述数据包。

7.3.1.3.2 IP 组播

路由器应满足 IETF RFC1122 中描述的 IP 组播要求。路由器应在所有相连网络上支持本地 IP 组播。当已指定 IP 组播地址到链路层地址映射时, 路由器应使用该映射。路由器可以配置成使用链路层广播来替代上述映射。在所有点到点链路和其他端口上, 组播被封装成链路层广播。

支持本地 IP 组播可以包括始发组播包、加入组播组、接收组播包、离开组播组等。

7.3.1.3.3 发现路径 MTU

为消除 IP 包分段或使 IP 分段数目最少, 有必要寻找路径端到端的 MTU。路径 MTU 是端到端每一跳 MTU 中的最小值。在 IETF RFC1191 中描述了如何动态发现路径 MTU 的技术。当路径中包含不支持 IETF RFC1191 的路由器时, 该技术可能不能发现正确的路径 MTU, 但该技术能得到一个 MTU, 该 MTU 总是比使用旧技术得到的 MTU 更精确。

当路由器产生一个 IP 数据包时, 建议使用 IETF RFC1191 中描述的技术限制数据包的大小。当路由器到目的地的路由所使用的路由协议提供路径 MTU 信息时, 仍应使用 IETF RFC1191 描述的技术。

7.3.1.3.4 划分子网

在某些情况下, 路由器有必要支持网络划分子网, 这些子网可能由不属于该网络的网络互联, 上述情况称为支持非连续子网。路由器应支持非连续子网。

一个给定网络的地址块可能被划分成不同大小的子块, 不同子块的网络前缀可能长度不同, 路由器应在所有接口配置和路由数据库中支持不同长度网络前缀。

7.3.2 互联网控制消息协议

7.3.2.1 定义

ICMP 是辅助协议, 它为 IP 提供路由、诊断和差错处理功能。路由器应支持 IETF RFC792 中描述的 ICMP 协议。

ICMP 消息分成两类:

1) ICMP 差错消息

- 目的地不可达;
- 重定向;
- 源抑制;
- 超时;
- 参数问题。

2) ICMP 请求消息

- 回应 (echo);
- 信息;
- 时间戳;
- 地址掩码;
- 路由器发现。

7.3.2.2 一般规定

7.3.2.2.1 未知消息类型

如果收到未知类型的 ICMP 消息, 该消息应送到 ICMP 用户接口 (如果路由器存在用户接口) 或者被丢弃 (如果路由器不存在用户接口)。

7.3.2.2.2 ICMP 消息 TTL

当产生 ICMP 消息时, 路由器应初始化 TTL 值。ICMP 应答的 TTL 值不能从触发该应答的包中得到。

7.3.2.2.3 初始消息头

ICMP 数据包应在不超过 576byte 的条件下尽可能多包含原始数据包内容。返回的 IP 头 (和用户数据) 应与收到的完全一致, 除非路由器没有要求恢复所有对 IP 头的改变, 这些改变通常是在检查出差错以前转发数据包的例行工作 (例如 TTL 减 1, 更新选项域)。但对参数问题消息, 如果问题在更新的字段中, 路由器应恢复更新 (见 7.3.2.3.5 节)。

7.3.2.2.4 ICMP 消息源地址

除非特别指定, 由路由器产生的 ICMP 消息的源地址应是传输 ICMP 消息的物理端口的 IP 地址。如果该端口没有 IP 地址, 则使用路由器 ID 替代。

7.3.2.2.5 TOS 及优先级

ICMP 差错消息的 TOS 应设置成与触发该 ICMP 差错消息的数据包的 TOS 相同, 除非如果设置成该值会导致该差错消息因无法选路到目的地而被立即丢弃, 否则 ICMP 差错消息应将 TOS 设置成 0。ICMP 应答消息的 TOS 应设置成与引发该应答的 ICMP 请求相同。

7.3.2.2.6 不发送 ICMP 差错情况

在下面情况下可以不发送 ICMP 差错消息:

- 收到 ICMP 差错消息;
- IP 头有效性检验失败的数据包 (本节中指定允许发送 ICMP 差错消息时除外);
- 包的目的地地址是 IP 广播地址或组播地址;
- 作为链路层广播或组播发送的包;
- 源地址的网络前缀为 0 或者无效源地址的包;

— IP 分段后不是第 1 个 IP 分段（IP 头中分段偏移量为非 0）。

另外本标准规定丢弃包时不允许发送 ICMP 差错消息。

7.3.2.2.7 速率限制

发送源抑制消息的路由器应能限制产生该消息的速率。路由器应能限制发送其他类型 ICMP 消息（如目的地不可达、重定向、超时、参数问题等）的速率。发送速率限制应是路由器配置工作的一部分。限制的方式（每路由器或每端口）由设备厂商决定。

7.3.2.3 特定规定

7.3.2.3.1 目的地不可达（又称信宿不可达）

1) 当路由器因没有到指定目的地的路由（包含无缺省路由）而无法转发包时，路由器应产生一个“目的地不可达，编码为 0（网络不可达）”的 ICMP 消息。

2) 如果路由器中存在到目的网络的路由，但该路由指定的 TOS 既不是“0000”，也不是需要路由的数据包中的 TOS，路由器应产生一个“目的地不可达，编码为 11（网络因 TOS 不可达）”的 ICMP 消息。

3) 如果 IP 包需要被转发到直连在路由器上网络中主机（路由器是最后一跳路由器）且路由器确认没有到目标主机的路由，路由器应产生“目的地不可达，编码 1（主机不可达）”的 ICMP 消息。

4) 如果 IP 包需要被转发到直连在路由器上网络中主机但路由器因为没有到目的地的路由，该路由的 TOS 或者等于“0000”，或者等于需转发数据包的 TOS，不能转发 IP 包时，路由器应产生一个“目的地不可达，编码为 12（主机因 TOS 不可达）”的 ICMP 消息。

7.3.2.3.2 重定向

ICMP 重定向消息用作通知本地主机某特定流量需使用不同的下一跳路由器。

与 IETF RFC1122 相反，路由器在下面情况下可以不理睬 ICMP 重定向：当路由器运行路由协议，或者路由器正在发送数据包的端口上允许转发时，路由器可以为自身生成的数据包选择路径。

7.3.2.3.3 源抑制

路由器不应产生 ICMP 源抑制消息。如果路由器产生 ICMP 源抑制消息，则应能控制该消息产生的速率。

路由器可以不理睬收到的源抑制消息。

7.3.2.3.4 超时

当路由器转发一个 IP 包且 TTL 域减为 0 时，应符合 7.4.2.1 节中的规定。

当路由器重组一个发给该路由器的包时，路由器应符合 IETF RFC1122。

当路由器收到（即包目的地是该路由器）一个超时消息时，路由器应符合 IETF RFC1122。

7.3.2.3.5 参数问题

当路由器遇到其他 ICMP 消息没有覆盖的差错时，应产生参数问题消息。在该 ICMP 消息中，应包含未经改变的 IP 头及指针指向的参数域。7.3.2.2 节中规定了上述规定的一个例外。

一个参数问题消息的新变量已在 IETF RFC1122 规定：

编码 1= 所需选项域丢失。

7.3.2.3.6 Echo 请求/响应

路由器应实现 ICMP Echo 服务器功能：接收发给该路由器的 Echo 请求并发送相应的 Echo 响应。路由器应能够接收，重组及响应一个 ICMP Echo 请求，该请求数据包可能最大到 576byte 或者满足所有相

连网络的 MTU。

Echo 服务器功能可以选择不响应一个目的地址是 IP 广播或 IP 组播的 IP 包。

路由器应有一个可选项域，能使路由器丢弃 ICMP Echo 请求，如果提供该选项域，缺省情况下应设置成回应 Echo 请求。

出于路由器维护的目的，路由器应实现一个用户/应用层接口来发送/接收 Echo 以用作网络诊断。所有的 ICMP Echo 响应应传送到该接口。

ICMP Echo 响应中的源地址应与 ICMP Echo 请求中的目的地址相同。

在 ICMP Echo 请求中包含的数据应包含在相应的 Echo 响应中。

如果 ICMP Echo 请求中包含路由记录或/和时间戳选项域，这个/些选项域应被更新后包含在 ICMP Echo 响应消息中。这样，路由记录将得到整个路由轨迹。

如果 ICMP Echo 请求中包含源路由选项域，返回路由应将 Echo 请求消息中路由记录反向，除非该路由违反路由器的转发策略。

7.3.2.3.7 信息请求/响应

路由器不应产生或响应这些信息。

7.3.2.3.8 时间戳及时间戳响应

路由器应实现时间戳及时间戳响应：

— ICMP 时间戳服务器功能应响应每个收到的时间戳消息。路由器应设计成实现最小的延迟变化。

— 可以丢弃对 IP 广播或组播地址的 ICMP 时间戳请求消息。

— ICMP 时间戳响应消息中的 IP 源地址应与相应 ICMP 时间戳请求消息中目的地址相同。

— 如果 ICMP 时间戳请求中包含源路由选项域，返回路由应将时间戳请求消息中路由记录反向，除非该路由违反路由器的转发策略。

— 如果路由器提供发送 ICMP 时间戳请求的应用层接口，则收到的 ICMP 时间戳响应应传送到该用户接口。

— 时间戳值所期望的值（标准值）以百万分之一秒从午夜（环球时间）计数。但是很难提供百万分之一秒精确率的值，例如一些系统以现行频率更新时钟，每秒 50/60 次，所以允许标准值应至少每秒更新 16 次以及标准值的精确性应接近操作员设置的 CPU 时间。

7.3.2.3.9 地址掩码请求/响应

路由器应实现接收 ICMP 地址掩码请求消息并应答回应 ICMP 地址掩码响应，该消息在 IETF RFC950 中定义。

路由器应为每个逻辑端口提供是否允许响应该端口上地址掩码请求的选项域，该选项域应缺省设置成响应 ICMP 地址掩码请求。路由器在知道正确的地址掩码前不允许响应地址掩码请求。

当从一个包含多个逻辑端口的物理端口上得到源地址为 0.0.0.0 的地址掩码请求，且这些逻辑端口的地址掩码不同时，路由器不允许响应该地址掩码请求。

路由器应检查由地址掩码请求得到的掩码是否与路由器了解的掩码匹配。如果 ICMP 地址掩码请求是差错的，路由器应将地址掩码及发送方 IP 地址写入日志。不允许路由器使用 ICMP 地址掩码响应中的内容决定正确的地址掩码。

当路由器宕机时启动的主机可能无法得到地址掩码时，路由器在配置完它的地址掩码之后，可以在

每个逻辑端口上广播发送没有必要的 ICMP 地址掩码响应。但是上述特性在变长掩码网络环境中是危险的。所以如果实现该特性，在下面情况下不广播没有必要的地址掩码响应：

— 配置成不广播无必要的地址掩码请求。每一逻辑端口应拥有可配置的参数来决定是否广播，端口缺省配置应是不发送无必要的地址掩码响应。

— 共享包含（但不相同）网络前缀和物理接口。

{网络前缀，-1}格式的 IP 广播地址应在广播地址掩码中使用。

7.3.2.3.10 路由器宣告和请求

路由器应在该路由器支持 IP 广播或 IP 组播地址的相连网络上支持 ICMP 路由器发现协议（IETF RFC1256）中的路由器部分，而且应包含对路由器特定的、具有缺省值的所有可配置变量。

7.3.3 互联网组管理协议

IGMP是用于主机和组播路由器之间的协议，应用于一单个物理网络上以建立特定组播组中的主机成员关系。组播路由器使用该信息和组播路由协议一起支持互联网上的IP组播转发。

路由器可支持互联网组管理协议，应支持IGMPv1、IGMPv2、IGMPv3。

路由器应实现IGMP中路由器部分要求。路由器可实现IGMP中的主机部分要求。

7.4 互联网层转发协议

本节描述包转发进程。

7.4.1 转发描述

包转发过程具体规定见互联网层协议（IETF RFC791、IETF RFC950、IETF RFC922、IETF RFC792、IETF RFC2474）。

7.4.1.1 转发算法

转发是路由器的重要工作。转发算法的好坏对路由器的性能影响极大。路由器实现中最关键的是使路由器在最大速率转发数据包时仍得到最佳性能。由于转发算法与路由器的结构有关，详细的转发算法实现不在本标准考虑的范围之内。本标准仅指出下列步骤的依赖关系：

1) 路由器在按照 IP 头内容进行转发动作前应检验 IP 头的有效性，检验过程在 7.4.1.2 节中描述。该检验允许路由器在无效的包占用其他资源之前将其丢弃。

2) 在执行某些 IP 选项域时，要求路由器在选项域中插入路由器自身地址。如 7.4.1.4 节中的描述，路由器应将发送数据包逻辑端口的 IP 地址插入 IP 包，在非编号接口上则使用路由器 ID。所以在输出端口选择之后，才能完成上述 IP 选项域的处理。

3) 由于 7.3.1.2.9 节中指出的原因，路由器在确认数据包是否分发给该路由器自身之前不能检查 TTL 或为 TTL 递减。

4) 当 IP 包本地分发给路由器后，IP 头不能做任何改变（除非路由器被要求插入时间戳选项域），所以路由器在决定是否将包本地分发给路由器之前，路由器不能做任何无法恢复的更改。

7.4.1.2 IP 头确认

路由器在处理 IP 包之前，应作下述的有效性检查来确认该 IP 包头是否有意义。如果 IP 包头在下面任一项检查中失败，该 IP 包被丢弃，差错写入日志：

- 1) 链路层指示的 IP 包长度应足够大，能容纳最小的 IP 数据包合法长度（20byte）。
- 2) IP 校验和正确。

3) IP 版本号为 4。

4) IP 头长度与应足够大, 能容纳最小的 IP 数据包合法长度 (20byte=5Word)。

5) IP 数据包总长度应足够大能容纳 IP 数据包头, IP 包头的长度在 IP 包头长度字段中指示。

路由器中不允许有任何选项域来禁止上述任何一项检查。

如果数据包通过第 2), 第 3) 项检查, 则 IP 头长度字段至少为 4, IP 总长度字段以及链路层报告的数据包长至少为 16; 若未通过上述检查, 路由器可以用 ICMP 参数问题消息来响应, 其中指针指向 IP 头长度字段 (如果未通过第 4) 项测试) 或 IP 总长度字段 (如果未通过第 5) 项测试)。同时, 路由器仍应丢弃上述数据包且将差错写入日志。

上述规则只对 IP 协议版本 4 有效。这些规定不能用作禁止路由器支持其他版本协议。

另外, 路由器应检验链路层报告的包长应至少等于 IP 头中指示的 IP 包长度。如果包被截断, 该包应丢弃, 差错写入日志, 路由器应发出 ICMP 参数问题消息, 其中指针指向 IP 包总长度字段。

最后, 如果目的地址不是路由器的某 IP 地址, 路由器应检验 IP 包不包含严格源或路由记录选项域。如果 IP 包在该测试中失败 (如果包含严格源路由选项域), 路由器将差错写入日志, 发出 ICMP 参数问题消息, 其中指针指向 IP 目的地址域。

7.4.1.3 本地分发决定

路由器收到 IP 包时, 应决定该数据包是否发给本路由器 (应本地分发) 或另一系统 (应由转发处理)。上述两种情况可能同时发生, 某些 IP 广播或 IP 组播可能同时要求本地分发及转发。路由器应使用下面规则决定适用那一种情况:

1) 当选项域指针值没有超过源路由选项域中最后条目 (entry) 时称为未到期的源路由选项域。如果 IP 包包含一个未到期的源路由选项域, 选项域中的指针应前移, 直到指针超过该选项域中最后一个地址, 或下一地址已不是该路由器的地址。在后者情况下 (即一般情况下), 包被转发 (不是本地分发)。

2) 下列情况下, 包被本地分发, 不考虑转发:

- 数据包的目的地址完全匹配路由器的某个 IP 地址;
- 数据包的目的地址是有限广播地址 ($\{-1, -1\}$);
- 数据包的目的地址是 IP 组播地址, 该地址不应被转发 (例如 224.0.0.1 或 224.0.0.2), 且收到数据包的物理接口上至少有一个相关联的逻辑接口是目标组播组的成员。

3) 下列情况下, 包被转发及本地分发:

- 数据包的目的地址是 IP 广播地址, 该地址指向至少一个路由器的逻辑接口, 但是不指向收到该数据包的物理接口上相关联的逻辑接口。
- 数据包的目的地址是 IP 组播地址, 该地址允许转发, 且收到数据包的物理接口上至少有一个相关联的逻辑接口是目标组播组的成员。
- 如果数据包的目的地址是 IP 广播地址 (不是有限广播地址), 该地址至少是收到该数据包物理接口相关联的逻辑接口, 该 IP 包被本地分发, 除非收到该 IP 包的链路使用不区分广播与单播封装 (例如使用不同的链路层目标地址) 的 IP 封装, 否则, IP 包同时被转发。

在其他所有情况下, IP 包被转发。

7.4.1.4 决定下一跳地址

当路由器转发包时, 应决定将包直接发到目的地或者需要发到下一路由器。如果是后一种情况, 路

由器需要决定使用哪一个路由器作下一路由器。本节解释如何作决定。

本小节使用下列定义：

- 1) **LSRR-IP**：松散源路由及路由记录选项域。
- 2) **SSRR-IP**：严格源路由及路由记录选项域。
- 3) **源路由选项域**：**LSRR** 或 **SSRR**。
- 4) **最终目的地址**：包被发送到的地方——源路由包中源路由的最后一个地址，或者非源路由包的 **IP** 头中的目的地址。
- 5) **相邻**：不通过另一路由器到达。
- 6) **下一跳地址**：报应发送到的下一主机或路由器的 **IP** 地址。
- 7) **IP 目的地址**：最终目的地址，除非在源路由包中，源路由指定的下一地址。
- 8) **立即目标**：节点、系统、终系统、或任何 **IP** 目标地址所确定的系统。

7.4.1.4.1 IP 目的地址

如果：**IP** 包头中的目的地址是路由器的一个地址；数据包包含源路由选项域，且源路由选项域中指针没有超过选项域内容，则下一 **IP** 目标地址是由选项域中指针指向的地址。

如果：**IP** 头中的目的地址是路由器的一个地址；数据包包含源路由选项域，且源路由选项域中指针超过选项域内容，则消息指向分析该消息的系统。

路由器决定如何处理数据包时，应使用 **IP** 目的地址，不是最终目的地址（源路由选项域中最后一个地址）。

数据包中包含超过一个的源路由选项域是一个差错，如果收到这样的包，路由器应丢弃该包，响应一个 **ICMP** 参数问题消息，该消息的指针指向第二个源路由选项域的开始处。

7.4.1.4.2 本地/远程决定

当决定 **IP** 包需要按照 7.4.1.3 节指定的规则转发时，应使用下列算法决定立即目标是否可以直接访问（见 **IETF RFC950**）。

- 1) 对每个没有 **IP** 地址的网络地址（上文中描述的非编号线路），对比线路另一端路由器 **ID** 与 **IP** 目的地址。如果完全相等，包可以通过该接口传输。
- 2) 如果在第一步中没有选择网络接口，对赋予该路由器的每个 **IP** 地址：
 - a) 隔离该接口使用的网络前缀。
 - b) 将包中 **IP** 目的地址相应比特隔离。
 - c) 对比网络前缀，如果相等，该 **IP** 包经过相应的网络接口传送。
- 3) 如果目的地既不是非编号接口的相邻路由器 **ID** 也不是直联的网络前缀，该 **IP** 目的地只能通过其他路由器访问。对路由器及下一跳 **IP** 地址的选择在 7.4.1.4.3 节中描述。对于不作路由器的主机，该选择是配置的缺省路由器。

注：应考虑下面情况：如果在同一链路层网络上叠加多个 **IP** 子网，除非有策略约束，如果存在足够信息，即使不属于同一（子）网的两个路由器或主机也应该能利用链路层网络直接通信。下一跳路由协议（**NHRP**）允许 **IP** 实体决定在链路层网络向源端目标传送的“最佳”链路层地址。

- 4) 如果选择的“下一跳”可以通过配置 **NHRP** 的接口访问，则应用下列附加步骤：
 - a) 对比目的地 **IP** 地址与 **NHRP** 缓存中的目的地址。如果该地址在缓存中，将该数据包发送到相应

用缓存的链路层地址。

b) 如果地址不在缓存中, 则建立一个包含目标 IP 地址的 NHRP 请求包。将该请求发送到该接口上配置的 NHRP 服务器。这可能是一个独立的逻辑过程或路由气内一个实体。

c) NHRP 服务器将返回正确的链路层地址, 该地址用作传送数据包及后续数据包。当等待 NHRP 服务器响应, 系统可以将数据包传送到传统的“下一跳”路由器。

7.4.1.4.3 下一跳地址

路由器应用前一节中的算法决定 IP 目标地址是否相邻。如果相邻, 则下一跳地址与目的地 IP 地址相同, 否则该包只能通过下一路由器转发才能到达立即目标。

如果数据包包含 SSRR, 路由器应丢弃该包, 并响应一个 ICMP 源路由不正确差错消息。否则路由器在路由表中查找目标 IP 地址来决定正确的下一跳地址。

下一跳选择的目的是检查路由器的转发信息库 (FIB) 中的实体, 为 IP 包在 FIB 可用的路由中选择最佳路由 (如果存在)。

任何路由查找算法都在包含 FIB 实体内容的候选路由中查找。算法包含一系列步骤从集合中剔除路由, 这些步骤称为裁减规则。一般在集合中仅存在一条路由时算法结束。如果该集合变为空, IP 包因目的地不可达被丢弃。集合中也可能存在多条路由。这种情况下路由器决定选择一条路由, 或者采用最经常使用的路由来实现“负荷分担”。

路由器为 IP 包选择下一跳时应采用下列裁减规则 (规则 3 (轻量级 TOS) 例外), 如果路由器在决定下一跳时不考虑 TOS, 规则 3 应以下面所述顺序应用。这些规则应以出现的次序应用到 FIB。

1) 基本匹配

该规则丢弃所有除数据包 IP 目标地址外其他目的地的路由。

更精确地, 假设每条路由由目的地属性 (称为 route.dest), 相应的网络前缀长度称为 route.length, 用于指定 route.dest 中哪些比特有用。需转发的 IP 目的地址是 IP.dest。该规则丢弃所有除 route.dest 中 route.length 有用比特与 IP.dest 相同外的其他候选路由。

2) 最长匹配

最长匹配是上述基本匹配的改进。完成基本匹配后, 路由器查找最长的 route.length 值。丢弃所有其他路由。

3) 轻量级 TOS

每条路由有 TOS 属性 (称为 route.tos), 其值假设等于 IP 头中 TOS 域的值。路由协议分发 TOS 信息, 在加入 FIB 的路由信息中填入合适的 route.tos; 其他路由协议的路由的 TOS 作为缺省 TOS (0) 对待。IP 头中 TOS 域作为 IP.tos。

检查候选路由是否存在 route.tos = IP.tos, 如果存在则剔除所有其他路由。如果不存在, 在候选路由中剔除 route.tos 不为 0 的路由项。

4) 最佳度量

每条路由有度量属性 (称为 route.metric) 和路由域标识符 (称为 route.domain)。选路由相互比较, route.domain 相同的路由比较 route.metric, 剔除 route.metric 较差的路由。路由好坏的比较通常是简单的算数比较, 尽管某些协议使用需要复杂比较的结构表示 route.metric。

5) 厂商策略

厂商策略用于弥补上述规则的不足。厂商策略因厂商而异。

基本匹配和最长匹配裁减规则对一些特殊路由处理一般化。下面类型路由选择中，倾向由高到低递减：

1) 主机路由：到特定端系统的路由。

2) 分级网络前缀路由：到一个特定网络前缀的路由。在 FIB 中可能包含多个相互竞争的到达特定网络的路由（某前缀是另一网络前缀的前缀）。对此，按照网络前缀长度递减选择。

3) 缺省路由：到所有没有显式指定路由的网络的路由。是指定成前缀长度为 0 的路由。

如果使用上述规则后路由集为空（即没有发现路由），该 IP 包被丢弃，产生一个相应的 ICMP 差错（如果 IP 目的地址从源路由选项域中得到，则产生差错源路由 ICMP；否则按照 7.2.3.1 节描述产生目标主机不可达 ICMP 或者目标网络不可达（ICMP））。

7.4.1.4.4 管理倾向（Administrative Preference）

对厂商策略裁减规则的一个建议机制是使用管理倾向，该机制是简单优先级算法。管理倾向是手工为可选择的路由定义优先级。

每条路由关联一个基于不同属性的倾向性值（为倾向性值指定机制在下面建议）。倾向性值是一个 [0...255] 的整数，0 表示最倾向，254 表示最不倾向，255 是一个特殊值，指该路由不应使用。厂商裁减规则的第一步是丢弃所有最不可能的路由（倾向值为 255 的路由总是首先丢弃）。

由于误用该策略可能导致路由循环，该策略较不安全。由于没有协议来确保路由器配置的倾向性规则与相邻路由器配置的倾向性规则一致，因此在配置倾向性规则时，应考虑以下一些因素。

1) 地址匹配

应对所有路由（从同一路由域中学习到的路由）赋予一个相同的倾向性值，这些路由到一组匹配特定网络前缀的特定目的地。

2) 路由类

应在维护目的地的路由协议中，能对所有路由（从同一路由域中学习到的路由）赋予一个相同的倾向性值，这些路由拥有一特定的路由类（域中、域间、内部度量之外或外部度量之外）。

3) 接口

应对所有路由（从同一路由域中学习到的路由）赋予一个相同的倾向性值，这些路由能使包路由到一个特定逻辑接口（逻辑接口通常与路由器网络接口一一对应，除非某网络接口拥有多个 IP 地址，对应多个逻辑接口）。

4) 源路由器

应对所有路由（从同一路由域中学习到的路由）赋予一个相同的倾向性值，这些路由从一组路由器获得，该组路由器拥有符合一组特定网络前缀的源地址。

5) 始发 AS

应对提供信息的路由协议，能对所有路由（从同一路由域中学习到的路由）赋予一个相同的倾向性值，这些路由在另一个路由域中始发。对 BGP 路由，始发 AS 是列在路由的 AS_PATH 属性中的第一个 AS。对于 OSPF，如果标记的自动比特被设置，且标记的路径长度不等于 3，始发 AS 可以认为是路由的外部路由标记的低序的 16 比特。

6) 外部路由标记

应能对所有 OSPF 外部路由（从同一路由域中学习到的路由）赋予一个相同的倾向性值，这些路由的外部路由标记匹配一组指定值中的任何一个。

7) AS 路径

应能对所有 BGP 路由（从同一路由域中学习到的路由）赋予一个相同的倾向性值，这些路由 AS 路径匹配一组指定值中的任何一个。一个简单选择是匹配所有在路由的 AS_PATH 属性中显示（或者不显示）特定 AS 号的所有路由。一个更通用的替代方案是允许 AS 路径匹配一个指定的一般表达式。

7.4.1.4.5 负荷分担

在下一跳选择进程最后仍可能存在多条路由，这种情况下路由器有多种选择：它可以武断地丢弃一些路由；可以通过比较路由域中路由的度量来减少候选路由；也可以保留多个路由，使用负荷分担机制来分摊流量。由于在某个条件下非常有效的负荷分担可能在另一条件下无效，因此应提供允许管理员禁止负荷分担的手段。

7.4.1.5 未使用的 IP 头比特

在 IP 头的 TOS 字段和标志字段中包含几个未使用比特，不允许路由器只因为未使用比特非零值而丢弃 IP 包。

路由器应不理睬未使用比特的值，并原封不动转发。如果路由器将包分段，应将上述比特复制到每一个分段的相应位置。

7.4.1.6 分段及重组

路由器应支持 IP 分段。

在转发以前，路由器不能重组任何 IP 数据包。

7.4.1.7 互联网控制消息协议（ICMP）

本节描述由路由器发送的 ICMP 消息。

7.4.1.7.1 目的地不可达

ICMP 路由器不可达消息是路由器对因为目的地不可达或服务不可用不能转发的 IP 包的响应。例如消息的目标主机没有开机，不能响应 ARP 请求，消息的目的地网络前缀在路由器中没有有效的路由等。

路由器应有能力发送 ICMP 目的地不可达消息并且能选择一个与不可达原因最接近的编码。

在 IETF RFC792 和 IETF RFC1122 中规定的编码如下：

0 = 网络不可达：路由器在到目标网络的转发路径（路由）不可用时发送。

1 = 主机不可达：路由器在到直连网络上目标主机转发路径（路由）不可用（主机不应答 ARP）时发送。

2 = 协议不可用：最终目标的传输层不支持数据包中指定的传输层协议时发送。

3 = 端口不可用：当指定的传输层协议（例如 UDP）不能将包发给最终目标的传输层但没有相应的协议机制通知发送者时发送；

4 = 分段需求但 DF 设置：路由器需要将数据包分段，但因 DF 标志设置无法完成分段，路由器发送该消息。

5 = 源路由失败：当路由器不能按照源路由选项域将数据包转发到下一跳时发送。

6 = 目标网络不可知：由于该消息暗示目标网络不存在，该消息不应使用（应使用网络不可达消息 0 替代目标网络不可知消息 6）。

7 = 目标主机不可知：仅当路由器能判定（从链路层获知）目标主机不存在时发送。

11 = 网络因 TOS 不可达：在到达目标网络的转发路径（路由）上不存在指定的或者缺省的 TOS 时路由器发送。

12 = 主机因 TOS 不可达：由于到达目标的路由不匹配数据包中所要求的 TOS 或缺省 TOS，路由器不能转发包时发送。

附加编码规定如下：

13 = 通信管理禁止：路由器因管理者过滤器要求不能转发包时发送。

14 = 主机优先级冲突：第一跳路由器因下列原因通知主机不能得到指定的优先级时发送：特定的源/目标主机或网络组、上层协议、源/目标端口组合。

15 = 优先级禁止：网络管理员引入对操作的最低优先级要求，而数据包优先级低于该要求时发送。

注 1：路由器不应使用在标准 RFC1122 中定义的编码 8（源主机孤立消息），应使用编码 0（网络不可达）或编码 1（主机不可达）替代。

注 2：在标准 RFC1122 中定义了编码 9（通信的目标网络被管理员禁止）和编码 10（通信的目标主机被管理员禁止）。这些定义由端到端封装设备使用。路由器在相应情况下应使用新定义的编码 13。

路由器可以设置可配置的选项域来禁止编码 13（通信管理禁止）消息发送。该选项域使能时，因通信管理禁止而丢弃包时，路由器应不发送 ICMP 差错消息。

相似地，路由器可以设置可配置的选项域禁止编码 14（主机优先级冲突）以及编码 15（优先级禁止）消息发送。相应选项域使能时，由于上述原因丢弃包时，路由器不发送 ICMP 差错消息。

当网络上其他主机可能可达时，路由器应使用主机不可达或者目标主机不可知消息，否则源主机可能会错误地认为网络上所有的主机不可达。

IETF RFC1191 定义了对目的地不可达消息的改变以包含编码 4（分段需求但 DF 设置）。路由器发送编码 4（目的地不可达消息）时应使用上述改变。

7.4.1.7.2 重定向

ICMP 重定向消息是路由器通知本地主机某一类流量应使用另一个不同的下一跳路由器。

路由器应不产生 IETF RFC792 中规定的网络重定向消息（编码 0），或网络和 TOS 重定向消息（编码 2）。路由器应能产生主机重定向消息（编码 1），应能产生在 IETF RFC792 中指定的 TOS 和主机重定向消息。

路由器不能因 IETF RFC792 中指定的网络或 TOS 消息（编码 0，2）发送网络重定向消息。路由器应能产生主机重定向消息（编码 1），应能产生在 IETF RFC792 中规定的 TOS 和主机重定向消息（编码 3）。

当触发重定向的包具有一个目的地，其路径需要路由器根据所要求的 TOS 来选择时，则路由器可以发送编码 3（因主机和 TOS 重定向）消息。

能产生编码 3（主机和 TOS）重定向消息的路由器应设置可配置的选项域（缺省为允许）来允许发送编码 1（主机）重定向来替代编码 3 重定向。如果作相应配置，路由器应发送编码 3 重定向替代编码 1 的重定向。

除非满足下列所有条件，路由器不可以产生重定向消息：

- 1) 数据包正在转发到接收该包的同一物理接口，
- 2) 包中的源 IP 地址与下一跳 IP 地址在同一 IP 逻辑（子）网上，并且

3) 包中没有包含 IP 源路由选项域。

ICMP 重定向消息中源地址应与目标地址在同一 IP 逻辑(子)网上。

应用路由协议(除静态路由外)的路由器不允许在转发包时使用从 ICMP 重定向中学习到路由。如果路由器没有应用路由协议,路由器可以设置可配置的选项域允许在转发包时考虑从 ICMP 重定向中学习到路由。

7.4.1.7.3 超时

当由于 TTL 域超时丢弃一个 IP 包时,路由器应产生一个超时消息(编码 0)。路由器可以设置基于每个端口的选项域来禁止产生该消息,但是路由器在缺省条件下应允许产生该消息。

7.4.2 特定规定

7.4.2.1 生存时间(TTL)

IP 头中定义的 TTL 域作为限制数据包生存时间的定时器。TTL 是以秒为单位的 8 比特域。即使实际处理时间小于 1s,处理 IP 包的每个路由器(或者其他模块)应至少将 TTL 减 1。通常情况下,TTL 可以作为衡量数据包在互联网上能传输多少跳的限制。

当路由器转发数据包时,应将 TTL 至少减 1。如果处理包时间大于 1s,路由器可以将 TTL 每秒减 1。

如果 TTL 减到 0 或更少,数据包应被丢弃,并且,如果目的地不是组播地址,路由器应向 IP 包发送方发送 ICMP 超时消息编码 0。另外,路由器不可以因为预测到数据包最终目的地路径上其他路由器会丢弃该数据包而预先丢弃该 TTL 不为 0 的单播数据包或广播包。但是路由器可以对组播包作预丢弃,以更有效地实现 IP 组播的扩展环搜索算法(见 IETF RFC1112)。

7.4.2.2 服务类型(TOS)

IP 头中 TOS 字节分 3 部分:优先级字段(高 3bit),TOS 字段(下 4bit)以及保留比特(最低 1 个 bit)。对保留比特的规定见 7.3.1.2.3 节。对优先级比特的规定见 7.4.2.3 节。对 TOS 域的规定及使用见 IETF RFC1349。

IP 头中 TOS 字节也可以理解为 DSCP 域,对 DSCP 域的规定及使用见 IETF RFC2474。

当决定如何转发 IP 包时,路由器应考虑包 IP 头中 TOS 域。

路由器应在它的路由表中为每个路由维护 TOS 值。如果通过路由协议学习到的路由不支持 TOS,则应分配 0 值(缺省 TOS 值)。

为选择到目的地的路由,路由器应使用等效于下面描述的算法:

- 1) 在路由表中找到所有到目的地的路由。
- 2) 如果没有上述路由,路由器因目的地不可达丢弃该包。
- 3) 如果存在一条或多条 TOS 匹配包中 TOS 选项域的路由,路由器选择最佳路由。
- 4) 否则路由器重复上述步骤查找 TOS=0 的路由。

5) 如果没有选择内容,则路由器因目的地不可达丢弃该包。路由器返回 ICMP 目的地不可达差错包,指定编码:网络因 TOS 不可达(编码 11)或者主机因 TOS 不可达(编码 12)。

7.4.2.3 IP 优先级

本小节规定路由器中 IP 优先级字段正确处理的要求。优先级机制是基于不同业务流的相对重要性而在网络中分配相应的资源。IP 层规范规定了为不同类型流量而使用不同的特定值。

路由器中优先级处理的基本机制是基于资源的优先分配,包括基于优先级顺序的队列服务,基于优

优先级的拥塞控制以及链路层优先级属性的选择。路由器也为它产生的路由、管理、控制流量分配 IP 优先级。

在本节中讨论的优先级顺序队列服务包括但并不限于转发过程的排队和输出链路排队。路由器支持的优先级应同样包含何时对有限资源（例如包缓存或链路层连接）分配考虑优先级处理。

7.4.2.3.1 优先级顺序队列服务

路由器应实现优先级顺序队列服务。优先级顺序队列服务表示当选择数据包到输出链路时，先选择队列中最高优先级的数据包。实现优先级顺序队列服务的路由器应设置可配置的选项域在互联网层抑制优先级顺序队列服务。

路由器可以实现基于其他策略（不同于严格优先级队列）的吞吐量管理规程，但是应设置可配置的选项域来关闭该吞吐量管理规程。

在拥塞控制中，实现优先级顺序队列服务的路由器应在丢弃高优先级包之前先丢弃低优先级包。预占（包处理或传输的中断）不是互联网层的功能，其他层协议可以提供这种预占特性。

7.4.2.3.2 低层优先级映射

实现优先级顺序排队的路由器应提供低层优先级映射，建议其他路由器也提供低层优先级映射。

实现低层优先级映射的路由器应满足：

- 1) 应能将 IP 优先级映射到链路层优先级，如果该链路层具有规定的优先级特性；
- 2) 应实现一个可配置的选项域来配置对所有 IP 流量选择链路层缺省的优先级处理；
- 3) 建议在每个接口配置 IP 优先级值到链路层优先级值特定的非标准映射。

7.4.2.3.3 所有路由器的优先级处理

路由器（无论是否使用优先级顺序队列服务）应支持：

- 1) 正常情况下应接受并处理所有优先级的流量，除非管理员配置要求不这样做。
- 2) 可以实现对特殊流量源使用管理员规定的优先级过滤器。如果提供上述特性，该过滤器不允许过滤或截断下列类型 ICMP 差错消息：目的地不可达、重定向、超时和参数问题。如果提供该过滤器，该过滤器中同样需要根据地址的包过滤规程。

当数据包被过滤器丢弃时，应发送编码为 14 的 ICMP 目的地不可达消息，除非配置要求禁止发送该 ICMP 消息。

- 3) 可以实现截断功能，该功能允许路由器拒绝或丢弃低于某一优先级的流量。该功能可以由管理措施而激活，或者由相关依赖的一些机制来激活，但应提供一个可配置的选项域禁止非人工干预的激活。当包由截断功能丢弃时，路由器应发送编码 15 的 ICMP 目的地不可达消息，除非配置要求禁止发送。

不允许路由器因优先级截断功能拒绝转发优先级为 6（互联网络控制）或 7（网络控制）的数据包。一般情况下，主机流量优先级应该是 5（CRITIC/ECP）或者更低。

- 1) 可以改变不是由本路由器产生的数据包的优先级设置。
- 2) 应能为支持的每个路由或管理协议单独配置优先级（某些协议除外，例如 OSPF，该协议指定的优先级值应使用）。
- 3) 可以不依赖对端地址而配置路由或管理流量优先级。
- 4) 应能正确响应提供的链路层优先级有关的差错指示。当链路因优先级有关条件不能接收包时，IP 包被丢弃，路由器应产生编码 15 的 ICMP 目的地不可达消息，除非配置要求禁止发送。

7.4.2.3.4 链路层广播转发

封装在大多数链路层协议中的 IP 包（PPP 除外）允许接收方只通过检查链路层协议的头（通常是链路层目的地址）来区别广播、组播或单播包。这一节中提到的链路层广播只适用于可以区分出广播包的链路层协议，同时提到的链路层组播包只是用于可以区分组播包的链路层协议。

路由器不允许转发作为链路层广播收到的任何包，除非目的地址指向一个 IP 组播地址。

路由器不允许转发作为链路层组播收到的包，除非数据包的目的地址是一个 IP 组播地址。

如果一个链路层广播帧既没有携带一个 IP 组播地址，也没有携带一个 IP 广播地址，则路由器应丢弃该帧。

当路由器作为链路层广播发送包时，IP 目的地址应是有效的 IP 组播或者 IP 广播地址。

7.4.2.4 互联网层广播转发

主要有两种 IP 广播地址类型：有限广播与定向广播。定向广播有 3 种子类型：对指定网络前缀的定向广播；向指定子网的定向广播；向指定网络所有子网的广播。对广播进行上述分类是根据路由器对目标网络子网结构的理解，同一广播地址对不同路由器可能得到不同的分类。

有限广播地址定义成全 1 地址：{ -1, -1 } 或者 255.255.255.255。

针对网络前缀广播包含网络前缀以及全 1 的主机地址即 { <网络前缀>, -1 }。例如 A 类广播地址：net.255.255.255，B 类广播地址：net.net.255.255，C 类广播地址：net.net.net.255。其中 net 表示网络地址字节。

针对所有子网广播没有在 CIDR 中定义。

当路由器遇到一些非标准 IP 广播地址：

- 1) 0.0.0.0 是废弃的有限广播地址形式。
- 2) { <网络前缀>, 0 } 是废弃的针对网络前缀广播地址。

按照本节描述，路由器收到上两种广播包后应丢弃，如果不这样处理，则应按照未废弃的广播地址处理。这些规则在下面几节中描述。

7.4.2.4.1 有限广播

有限广播报文不能转发。有限广播报文不能丢弃。在有限广播有效的情况下，应使用有限广播替代定向广播。

7.4.2.4.2 定向广播（可选）

路由器应将网络前缀定向广播作为有效，对远端网络或连接的无子网网络定向广播。

路由器可以设置选项域来允许接收针对网络前缀广播，可以设置一个选项域允许转发针对网络前缀的广播。这些选项域应缺省设置成禁止接收且禁止转发针对网络前缀的广播（见 IETF RFC2644）。

7.4.2.4.3 针对所有子网定向广播

路由器不应支持针对所有子网定向广播。

7.4.2.4.4 针对子网定向广播

在 IETF RFC1812 中规定了针对子网广播的算法。在 CIDR 路由域中，针对子网广播与针对网络广播没有区别，所以针对子网广播可以看作针对网络广播。

7.4.2.5 拥塞控制

网络中的拥塞可以不严格地定义成一种状态，在这种状态下对资源的需求（通常是带宽或 CPU 时间）

超出了路由器的能力范围。拥塞避免是路由器试图阻止请求过度的资源，拥塞恢复是路由器试图从拥塞状态中恢复正常运行。路由器可以综合使用上述两种机制。本标准建议使用 IETF RFC1154。

当主机使用合理的拥塞策略时，路由器处理峰值需求所需要的存储器总量与链路带宽和链路上延迟的乘积有关。所以当带宽乘延迟增加时，路由器存储器相应增大。存储能力与丢包率的关系尚属未知。

当路由器收到一个包超出其存储能力时，应丢弃该包或其他一个（些）包。一种选择是丢弃从最繁忙链路上收到的包，但需要考虑其他相关因素：包括流量的优先级，带宽预留以及选择需丢弃包的复杂度等。

路由器可以丢弃刚收到的包，这是最简单但不是最好的策略。路由器应选择一个严重滥用链路的流的包，假定它应用的 QoS 允许这样做。对数据包环境中使用 FIFO 队列的建议策略是在队列中随机选择一个包丢弃。路由器中使用公平排队的相似算法是在最长对列中丢弃或者使用虚时间最长中丢弃。路由器可以使用这些算法决定如何丢弃数据包。

如果路由器实现一种丢弃策略（例如随机丢弃），在候选丢弃的包中：

- 1) 如果实现且使能优先级顺序队列服务，路由器不允许丢弃一个 IP 优先级比未丢弃包 IP 优先级高的包。
- 2) 在不违背上一规则的情况下，路由器可以保护 IP 头中要求最高 TOS 可靠性的包。
- 3) 路由器可以保护分段的 IP 包，因为丢弃一个分段数据包可能引发该数据包所有分段重发，并进一步加重拥塞。
- 4) 路由器可以保护用作路由控制，链路控制或网络管理的包。专用路由器（不同于用作主机或终端服务器等的路由器）只要检查包的源或目标地址是否是路由器本身地址即可大致达到上述目的。

高级的拥塞管理需要考虑公平性，需要得到丢包惩罚的用户应是引起拥塞的用户。无论实现哪一种带宽拥塞控制技术，都要注意使 CPU 负荷足够小，不要因此引起 CPU 拥塞。

本标准建议不要向被丢弃包的发送者发送 ICMP 源抑制消息。ICMP 源抑制消息功能非常弱，路由器没有必要发送，主机软件不应该把它作为拥塞指示。

7.4.2.6 地址过滤

如果 IP 源地址是第 7.3.1.2.11 节中所指出地址而且不是一各单播地址，则该地址无效。

如果 IP 目的地址是第 7.3.1.3.1 节中规定地址或者是 E 类地址（255.255.255.255 除外）则该目的地址无效。

路由器不应转发源地址无效或源地址为 0（网络 0）的数据包。路由器不应转发源地址是 127 网络的数据包（除非是在环回接口上）。路由器可以设置一个开关允许管理员关闭这些检查。如果提供上述开关，上述开关的缺省状态应实施上述检查。

路由器不应转发目的地址无效或目的地址为 0（网络）的数据包。路由器不应转发目的地址是 127 网络的数据包（除非是在环回接口上）。路由器可以设置一个开关允许管理员关闭这些检查。如果提供上述开关，上述开关的缺省状态应实施上述检查。

如果路由器因上述规则而丢弃 IP 包，路由器应尽可能将下述内容写入日志：源 IP 地址、目的地 IP 地址。如果因源 IP 地址丢弃，还应将收到 IP 包的物理接口，发送该包的路由器或主机的链路层地址写入日志。

7.4.2.7 源地址检验

路由器应能实现按照 IP 包的源地址以及收到该 IP 包的逻辑端口的转发表过滤 IP 包。如果该过滤器被允许，当收到包的接口不是该包应被转发到的目的地包含的源地址时，该包被丢弃。简单说，如果路由器不愿意从某一特定接口路由包到某地址，该路由器不能相信从该接口上收到包内的源地址。

如果实现该功能，该功能缺省情况下应被关闭。

7.4.2.8 包过滤以及访问列表

为得到安全性及限制某一网络的流量，路由器应提供选择转发（过滤）的能力。如果提供该能力，路由器应能配置成转发所有包或者按照源地址或目的地址前缀有选择地转发或者按照其他消息属性过滤。对每个源或目的地址应能任意指定前缀长度。

如果指出上述能力，路由器应能被配置成以下两条中之一：

- 允许列表——指定一组能转发的消息定义；
- 拒绝列表——指定一组不能转发的消息定义。

其中消息定义是指源或目的地网络前缀，也可以包含其他标识信息例如 IP 协议类型、TCP 端口号码等。

路由器可以提供可配置的开关来制定允许或拒绝列表，或其他等效控制。

应允许一个匹配任何地址的值（例如关键字“任意（any）”或者全“0”掩码的地址或者全“0”的网络前缀）作为源或目的地址。

在地址对之外，路由器可以组合传输层和/或应用层协议指定的源和目的地端口来匹配。

路由器应允许丢弃 IP 包（即丢弃后不发送 ICMP 差错消息）。

路由器应允许当包丢弃时发送恰当的 ICMP 不可达消息。ICMP 消息应指定为通信管理禁止（编码 13）替代目的地不可达。

路由器应允许对可配置的允许指定的地址对、协议类型和端口发送 ICMP 目的地不可达消息（编码 13）。

路由器应能对丢弃的包计数，应能有选择性地将未转发的包写入日志。

7.4.2.9 组播路由

IP 路由器应能支持转发 IP 组播包，转发基于静态组播路由或者由组播路由协议动态决定的组播路由。转发 IP 组播包的路由器称为组播路由器。

7.4.2.10 转发控制

对每个物理接口，路由器应设置一个可配置的选项域来指定该接口上是否允许转发。当接口上禁止转发时路由器应：

- 1) 应丢弃所有该接口上收到的不是发给该路由器的包；
- 2) 除该路由器生成的包外，该接口不发送任何其他包；
- 3) 不通过任何路由协议宣布通过该接口的路经的可用性。

7.4.2.11 状态改变

在路由器运行期间，接口可能出现故障或被人为禁止，也可能重新可用。类似的，可能在接口或整个路由器上禁止转发或转发重新可用，路由器应能正确处理这种转变。

7.4.2.11.1 路由器停止转发时

当路由器停止转发时，路由器应停止广播路由（第三方路由除外）。该路由器可以继续接收并使用在

同一路由域其他路由器得到的路由。如果转发数据库仍存在，路由器不能停止转发数据库上的定时器。如果从其他路由器学习到的路由存在于缓存，路由器不能停止对路由的定时器。即使转发禁止，路由器仍应删除超时的转发路由。

路由器可以对不能转发的数据包的发送者发送 ICMP 目的地不可达（主机不可达）消息。路由器不应发送 ICMP 重定向消息。

7.4.2.11.2 路由器开始转发时

当路由器开始转发时，它应加速地向通常交换路由信息的其他路由器发送更新的路由信息。

7.4.2.11.3 当接口故障或禁止时

当接口故障或禁止时，路由器应删除并停止广播转发数据库中所有使用该接口的路由。该路由器应禁止所有使用该接口的静态路由。如果路由器得知或记得到相同目的地的其他路由时，路由器应使用一条最佳替代路由并加入到转发数据库中。路由器应发送 ICMP 目的地不可达或 ICMP 重定向消息，用作由于接口不可用而不能转发的数据包的响应。

7.4.2.11.4 当接口使能时

当不可用的接口变为可用时，路由器应使能所有使用该接口的静态路由。当路由器通过该接口学到路由时，路由器应权衡从其他接口学到的路由，决定将那条路由放到转发数据库。

路由器应加速地向通常交换路由信息的其他路由器发送更新的路由信息。

7.4.2.12 IP 选项域

一些选项域，例如路由记录和时间戳选项域，包含路由器转发路由时需要插入 IP 地址的位置。然而每一个类似的选项域都有有限个位置以供插入 IP，路由器可能发现已没有插入 IP 的位置。在第 7.4.1.5 节描述如何选择一个 IP 地址插入到选项域中。

7.4.2.12.1 转发中不可识别的选项域和不可识别的 IP 选项域

转发中，包含不可识别的选项域和不可识别的 IP 选项域的包通过后应毫无改变。

7.4.2.12.2 安全性选项域

某些环境需要在每个包中提供安全性选项域。路由器应实现 IETF RFC1112 中描述的安全性选项域。

7.4.2.12.3 流标识选项域

该选项域已废弃。如果路由器转发的包中包含流标识选项域，路由器应不理睬该选项域，且转发时不改变该选项域。

7.4.2.12.4 源路由选项域

转发包时，路由器应实现支持源路由选项域。路由器可以设置可配置的选项域，该选项域设置时，路由器丢弃所有源路由包。缺省情况下该选项域不能设置。

7.4.2.12.5 路由记录选项域

转发包时，路由器应支持路由记录选项域。路由器可以设置可配置的选项域，该选项域设置时，路由器不理睬所有的路由记录选项域（转发时不记录路由）。如果提供该选项域，缺省情况下不设置。该选项域不影响路由器处理收到的自身发出的带路由记录选项域的包（ICMP echo 请求中的路由记录选项域仍按照第 7.3.2.3.6 节中规定处理）。

7.4.2.12.6 时间戳选项域

转发包时，路由器应支持时间戳选项域。时间戳的值应按照 IETF RFC1122 中的规定来处理。

如果标志域=3（时间戳和预指定地址），且下一预指定地址匹配路由器中任一地址，路由器应写入时间戳。预指定地址不一定是收到包的接口地址或者转发包使用的接口地址。

路由器可以提供一可设置的选项域，设置该选项域后路由器转发包时即使标志域=0（时间戳）或=1（时间戳及注册 IP 地址）路由器也不理睬该选项域（转发时不改变）。如果提供上述选项域，该选项域应关闭（路由器不忽略时间戳）。该选项域不影响路由器处理收到的自身发出的有时间戳选项域的包，在路由器收到的数据包时间戳选项域中插入时间戳，ICMP echo 请求中的时间戳选项域仍按照 7.3.2.3.6 节中规定来处理。

7.5 传输层协议要求

路由器通常应该支持传输控制协议（TCP）和用户数据包协议（UDP）。

7.5.1 用户数据报协议-UDP

用户数据包协议在 IETF RFC768 中规定。

除下面两条外，路由器实现的 UDP 应符合，且无条件符合 IETF RFC1122 的要求：

- 1) 本标准不规定不同协议层间的接口；
- 2) 与 IETF RFC1122 相反，路由器应该产生 UDP 校验和。

7.5.2 传输控制协议-TCP

传输控制协议在 IETF RFC793 中规定。

除下面所述之外，路由器实现的 TCP 应符合，且无条件符合 IETF RFC1122 的要求：

- 1) 本标准不规定不同协议层间的接口。路由器不需要符合 IETF RFC1122 中的下列要求：

— “推动（PSH）”标志的使用（IETF RFC793 第 2.8 节）：将收到的“推动（PSH）”标志传递给应用层是任选的。

— 紧急指针（IETF RFC793 第 3.1 节）：当收到紧急指针且没有先前未处理的紧急数据时，或者紧急指针在数据流中推进时，TCP 应异步通知应用层。应用层应有方法获知连接中还有多少紧急数据，或者至少获知是否存在未处理的紧急数据。

— TCP 连接故障：对特定连接，应用应能为 R2 配置值，例如，交互应用可能将 R2 设置成“无限大”，这样可以使用户能控制何时断线。

— TCP 多归特性：当具有多归特性主机上的一个应用在打开 TCP 连接时没有指定使用的 IP 时，TCP 应在发送第一个 SYN 之前要求 IP 选择本地 IP 地址。

— IP 选项域：当打开 TCP 连接时，路由器应能指定一条源路由，并应在收到数据包中的源路由之前。

2) 关于 IETF RFC1122 中最大分段尺寸选项域的修改如下：实现 MTU 发现协议主机部分的路由器只有在路径 MTU 未知时使用 576 作为发送 MSS 的缺省值；如果路径 MTU 已知，发送 MSS 缺省值是路径 MTU-40。

3) 关于 IETF RFC1122 中最大分段尺寸选项域的修改如下：ICMP 目的地不可达消息编码 11 和 12 是附加软件差错状态，所以这些消息不应导致 TCP 放弃连接。

7.6 路由协议

7.6.1 概述

互联网路由系统包含两部分——内部路由与外部路由。自治域（AS）允许描述一组路由器从内部路

由到外部路由的转变。IP 数据包通常要穿过两个或多个 AS 的路由器才能到达目的地，AS 系统应相互提供拓扑信息才能允许这种转发。内部网关协议用作在 AS 内部分发路由信息（即 AS 内部路由）。外部网关协议用作在 AS 间交换路由信息（即 AS 间路由）。

7.6.1.1 路由安全性考虑

路由器应提供将路由信息源由最值得信赖到最不值得信赖排列的能力，并首先从最值得信赖的路由信息源接收路由信息。上述规定在使用外部路由网关 EGP 和其他内部路由协议的始发核心/末梢（core/stub）AS 系统中隐含使用。

路由器应提供一种机制来过滤过时无效的路由（例如 127 网络）。

缺省情况下，路由器不能分发不是该路由器使用、信任或认为有效的路由信息。有时路由器需要分发值得怀疑的路由信息，但由管理员直接人为干预。

对等实体之间（peer-to-peer）的认证涉及多种测试。对口令（password）消息的申请和可接收相邻路由器列表的使用有效地提高了路由数据库的鲁棒性。路由器应实现允许显式指定有效相邻路由器的管理控制。路由器对应对支持的路由协议实现对等实体之间（peer-to-peer）认证。

路由器应能基于源地址和接收数据包的端口来检验相邻路由器。路由器与直接相连的子网上的路由器应严格按照相连接口或者通过非编号接口进行通信。从其他接口上得到的信息应悄悄丢弃。

路由器应谨慎接收来自其他路由器的路由信息。

7.6.1.2 优先级

除非特殊路由协议的指定，路由器应将携带路由信息流量的IP数据包的优先级设置成6（互联网控制）。在MPLS帧中将EXP字段设置为6。

7.6.1.3 消息检验

对等实体之间（peer-to-peer）的认证涉及多种测试。对通行字（password）消息的申请和可接收相邻路由器列表的使用有效地提高了路由数据库的鲁棒性。路由器应实现允许显式指定有效相邻路由器的管理控制。路由器对应对支持的路由协议实现对等实体之间（peer-to-peer）认证。

路由器应能基于原地址和接收数据包的端口来检验相邻路由器。路由器与直接相连的子网上的路由器应严格按照相连接口或者通过非编号接口进行通信。从其他接口上得到的信息应丢弃。

7.6.2 内部网关协议

7.6.2.1 定义

内部网关协议（IGP）用作在特定 AS 内部路由器间分发路由信息。对特定 IGP 算法的实现相对独立，但应实现下列功能：

- 1) 应能迅速反映 AS 内部拓扑的改变；
- 2) 提供一种机制使电路振荡时不引起连续的路由更新；
- 3) 提供快速收敛成无环回（loop-free）路由；
- 4) 使用最少的带宽；
- 5) 提供等效路由以便负荷分担；
- 6) 提供一种认证的路由更新方法。

路由器除实现静态路由外，应实现 OSPFv2、IS-IS、RIP v2。

7.6.2.2 开放最短路径优先-OSPF

基于最短路径优先（SPF）的路由协议是一类基于链路状态算法的协议，它们基于 Dijkstra 的最短路径算法。在基于 SPF 的系统中，每个路由器通过称为洪泛（flooding）算法的过程得到完整的拓扑数据库。洪泛过程确保信息可靠传输。每一个运行 SPF 算法的路由器在数据路上建立 IP 路由表。

路由器应实现在 IETF RFC2328 中规定的 OSPFv2，支持可变长子网掩码（VLSM），支持广播网络，支持非广播多接入网络（NBMA），支持虚链路，支持末梢（stub）域和 NSSA，支持等开销多路径。

实现 OSPF 的路由器应实现 OSPF MIB（见 IETF RFC4750）。

7.6.2.3 中间系统到中间系统-IS-IS

IS-IS是基于链路状态（SPF）路由算法，拥有所有该类协议的优点。

路由器应实现 IS-IS。

IS-IS 在 IETF RFC1142 和 IETF RFC1195 中规定。

实现 IS-IS 的路由器应实现 IS-IS MIB。

7.6.2.4 路由信息协议

RIP应用极其广泛，是自治域内路由协议的事实标准之一。

路由器可以实现 RIPv1 和 RIPv2。支持 RIPv1 和 RIPv2 的路由器应实现 RIPv1 和 RIPv2 MIB。

7.6.3 外部网关协议

7.6.3.1 概述

外部网关协议在自治系统间使用，为特定自治系统内一组网络与相邻自治系统交换可达性信息。

路由器应实现 BGP 4。

7.6.3.2 边界网关协议-BGP4

7.6.3.2.1 定义

边界网关协议（BGP4）是自治域间路由协议，是在 BGP 运行者之间交换网络可达性信息的协议。网络信息包含流量到达某个网络所应经过的完整 AS 列表。该信息应确保路径内没有环路。该信息应足够丰富以用作构建 AS 互连图，在 AS 互连图中，应裁减路由环回，应被实施 AS 层的策略决定。

BGP4 在 IETF RFC4271 中规定。

实现 BGP4 的路由器应实现 BGP4 MIB（见 IETF RFC4273）。

建议实现 BGP 的路由器遵从 IETF RFC1772 第 6 章中的规定。

7.6.3.2.2 协议描述

BGP4 提供对非常复杂的路由策略的支持，但不要求所有对 BGP4 的实现都支持这样的策略。BGP4 至少要实现：

1) 应允许 AS 控制 BGP4 学到的路由是否广播到相邻 AS。BGP4 的实现应至少在单个网络范围内实现上述规定。BGP4 的实现同样应在自治系统范围内实现上述规定，上述自治系统可能是产生路由的自治系统，或者可能是将路由广播到本地系统的自治系统（相邻自治系统）。

2) 应允许 AS 当存在多条路径时倾向于使用某条特定路径。该功能应通过允许管理员向自治系统度量赋值来实现，使路由选择进程选择一条最低度量的路由（路由的度量定义为有关该路由 AS_PATH 路径属性所有 AS 的度量之和）。

3) 应允许 AS 忽略 AS_PATH 路径属性中包含某特定 AS 的路由。这样的功能可以使用 2) 中提到的技术实现：将某 AS 度量赋为无限大。路由选择进程应不理睬度量为无限大的路由。

- 4) 提供 BGP4 路由反射, 并符合 IETF RFC4456 的规定。
- 5) 提供 BGP4 区域属性, 并符合 IETF RFC1997 的规定。
- 6) 提供自治系统联合。
- 7) 支持 IETF RFC2439 规定的路由振荡抑制。
- 8) 支持 IETF RFC4760 规定的 BGP 4 的多协议扩展。

7.6.4 静态路由

静态路由提供一种途径来显示定义到一个特定目的地的下一跳路由器。路由器应提供一种途径来定义到特定目的地的静态路由, 其中目的地由网络前缀定义。该机制应允许对每一条静态路由指定度量 (metric)。一个支持动态路由协议的路由器应允许静态路由定义成任何路由协议使用的有效的度量。路由器应允许用户规定一组静态路由是否通过路由协议扩散。另外如果路由器支持使用下列信息的路由协议, 应在静态路由中支持这些附加信息。这些信息是:

- TOS;
- 子网掩码;
- 前缀长度;
- 对给定路由协议引入静态路由的特定度量。

7.6.5 路由信息的过滤

网络中每个路由器基于转发数据库中包含的信息作转发决定。在一个简单网络中转发数据库内的信息可以静态配置。当网络变得复杂时, 动态更新转发数据库对网络有效运行至关重要。

如果要求通过网络的数据流尽可能地高效转发, 则需要一种机制来控制哪些路由可以用作创建转发数据库的信息。这种控制任务可以通过哪一个路由信息源可以信任, 选择相信哪一条消息的形式来实现。转发数据库是可用的路由消息经过过滤后的结果。

除有效性之外, 控制路由消息传播可以通过阻止不正确或错误的路由信息的扩散而增加转发数据库的稳定性。

在某些情况下, 本地策略可能要求不能广泛传播整个路由信息。

这些过滤器要求只用于非 SPF 协议 (对不实现距离矢量协议的路由器没有影响)。

7.6.5.1 路由检验

当路由更新宣告中路由违反本标准的规定时, 路由器应作为差错写入日志, 除非接收的更新路由协议使用这些值编码那些特殊路由编码 (例如缺省路由)。

7.6.5.2 基本路由过滤

过滤路由信息允许对路由器用于转发报文的路径进行控制。路由器应可以配置从哪一个路由信息源接收路由消息, 哪一条路由可以信任, 因此路由器应指定:

- 路由信息可以从哪个逻辑接口接收, 从每个逻辑接口上可以接收哪些路由;
- 在一个逻辑接口上传播所有路由或者只传播缺省路由。

某些路由协议不能将逻辑接口作为路由信息源, 在这种情况下, 路由器应指定: 从哪一个相邻路由器可以接收路由信息。

7.6.5.3 高级路由过滤

当网络拓扑变得越复杂时, 越需要对复杂的路由进行过滤, 因此路由器应对每个路由协议分别提供:

- 1) 从哪一个逻辑端口或路由器可以接收路由信息，那些路由可以信任；
- 2) 哪些路由将通过哪个逻辑接口来发送；
- 3) 路由信息将发送到哪个路由器。

在许多环境下，需要将其他路由器上收到的路由信息赋予可信度。路由器可以指定对收到的每条路由赋予可信度或优先级。无论每条路由所关联的路由度量如何，赋予高可信度的路由将被优先选择。

如果路由器支持赋予优先级值，路由器不允许传播不作为第一方信息选择的路由。如果路由器使用的路由协议不支持区分第一方与第三方信息，路由器不能传播任何不能优先选择的路由。

如果路由器不使用某路由信息中的路由，则路由器不能该路由传播给其他路由器。

7.6.6 路由协议之间信息交换

如果一些独立的IP路由进程可以同时运行在同一路由器上，路由器应能在独立的IP内部路由协议之间交换路由信息。如果路由器配置成在独立内部路由协议间双向交换路由信息，则应提供某些机制来防止路由环回。路由器应提供优先级机制，在独立的路由进程中选择路由。当穿过管理边界时，路由器应提供IGP-IGP交换的管理控制。

路由器应提供某种机制来翻译或转变基于每个网络的度量。路由器（或路由协议）可以允许在IGP中引入外部路由的全局（程）优先级。

7.7 MPLS 协议

路由器应支持MPLS协议，支持MPLS的路由器称为标记交换路由器。

支持MPLS的路由器应支持LDP标记分发协议，可选支持RSVP、CR-LDP。

支持MPLS的路由器应能配置备份LSP，支持负荷分担的多路径LSP，支持标记压栈根据路由表的下一跳等参数将包路由至输出LSP。

有关MPLS协议具体见YD/T 1162.1-2005。

7.8 IPsec 协议

IPSec是在IP层提供通信安全而制定的一套协议族，它可有效地保护IP数据包的安全，它采取的具体保护形式包括：数据起源地验证；无连接数据的完整性验证；数据内容的机密性；抗重播保护以及有限的数据流机密性保证。

路由器可选支持IPsec协议，支持IPsec的路由器称为IPsec路由器。

IPsec路由器应支持AH协议，在实现AH协议时需要实现下列算法：

- 使用MD5的HMAC算法（必选）；
- 使用SHA-1的HMAC算法（可选）。

路由器可选支持ESP协议，支持ESP协议的路由器需要实现下列算法：

- 使用MD5的HMAC算法（必选）；
- 使用SHA-1的HMAC算法（可选）。

IPsec路由器应支持手工密钥管理，可选支持IKE。

IPsec路由器在进行路由协议包交换时，应支持使用AH头通过MD5进行加密认证的功能。

IPSec具体的规定见YD/T 1466-2006。

7.9 虚拟专用网

路由器应支持虚拟专用网功能，支持虚拟专用网功能的路由称为VPN路由器。

路由器应支持BGP/MPLS VPN功能。BGP/MPLS VPN具体规定见YD/T 1476-2006。

VPN路由器可选支持Martini、VPLS功能。

Martini VPN具体规定见IETF RFC4906。

VPLS VPN具体规定见IETF RFC4762。

7.10 区分业务协议（Diff-Serv）

路由器应支持区分业务协议。

区分业务是指在用户和业务网接口处给业务分级，业务的分级基于每个数据包的不同标识。同一级别的业务在该网络域中会被聚合统一传送，保证相应的延迟、传送速率、抖动等服务质量参数。区分业务并不提供从发送者到接收者的端到端服务质量保证，而是在域的范围范围内保证与业务分类相对应的服务质量，每个域之间对于不同类别业务的服务质量都应有一定的约定和包标识的翻译机制。

区分业务的设计弱化了对信令的依赖，全部处理都基于对互联网业务流的分级，包括流的会聚以及适用于不同业务类别的每一跳行为（PHB）等，因此它具有较大的可扩展性。由于它不用在每个路由器上为每个业务流保存“软状态”，在网络不断扩展过程中，不必过于担心网络负担的增加是否会给网络造成根本性的损害，因此区分业务把对业务服务质量的控制缩小到了每个域的范围。

有关区分业务协议具体参见附录 C。

7.11 排队策略和拥塞控制

7.11.1 排队策略

1) 支持公平排队算法（Fair Queuing 或 Round Robin）。这一算法将为不同的队列公平地分配带宽，确保没有哪个流能独占网络资源。使用带宽少的流可以与使用带宽多的流获得近似的吞吐量，从而获得较小的延迟。采用公平排队算法后，路由器可降低对网络资源具有过度侵占性的流的速度。

2) 支持加权公平排队算法（WFQ）。它是对公平排队算法的改进。该算法给每个队列一个权（weight），由它决定该队列可享用的链路带宽。这样实时业务可以确实得到所要求的性能，非弹性业务流可以与尽力而为（Best-effort）业务流相互隔离。但由于需要对流进行分类，转发速度可能会有所下降。

7.11.2 拥塞控制

1) 应支持 WFQ、随机早期探测 RED、加权的随机早期探测 WRED 等拥塞控制机制。RED 算法的原理是当队列开始不断加长（有发生拥塞的危险时）但还远没有发生拥塞时，对于每一队列收到的分组进行随机的丢弃，以便保持队列维持一个较小的平均长度。队列的平均队长越长，随机丢弃的概率就越大。这种丢弃将使得一些 TCP 源降低数据发送速率，从而避免即将发生的网络拥塞。这种算法较好地解决了全网同步的问题，通过该方法使弹性业务降速还可以提高非弹性业务性能。WRED 算法是对 RED 的扩展，它将优先级策略与 RED 算法结合在了一起。它可以为具有不同服务等级的业务规定不同的队列长度与丢包率的映射关系。这样，当网络出现拥塞的可能时，对于不同的业务类型，路由器将根据上述映射关系以不同的概率进行随机丢弃处理，提供不同的业务保证。

2) 应支持一种机制，由该机制可为不符合其业务级别 CIR/Burst 合同的流量标记一个高的丢弃优先级，该优先级应比满足合同的流量和尽力而为的流量的丢弃优先级高。

3) 在有可能存在输出队列争用的交换环境中，应提供有效的方法消除头部拥塞。

7.12 组播路由协议

路由器应实现组播路由协议，应支持互联网组管理协议 IGMP v2（见 IETF RFC2236）和协议无关组

播协议一稀疏模式（PIM-SM），可以任选实现距离矢量组播路由协议 DVMRP（见 IETF RFC1075）和组播源发现协议 MSDP。在实现 PIM-SM 协议时应考虑与距离矢量组播路由协议 DVMRP 的互通。可选支持跨域的组播路由协议 MBGP。

有关各组播路由协议具体见 YD/T 1177-2002。

8 安全性要求

路由器安全性要求见 YD/T 1358-2005 和 YD/T 1359-2005。

9 性能指标要求

9.1 设备容量

- 1) 系统双向交换容量至少达到 60Gbit/s。
- 2) 当系统支持 OC-192/STM-64 接口时，应无需对现有通用部件进行升级。

9.2 服务质量

9.2.1 丢包率（packet loss rate）

丢包率是指路由器在稳定的持续负荷下由于资源缺少在应该转发的数据包中不能转发的数据包所占比例。

丢包率通常用作衡量路由器在超负荷工作时路由器的性能。

由于路由器设计使用在不同目的和应用环境，对丢包率作范围限制没有意义。本标准对丢包率不作规范。只作为重要的性能指标供比较。

9.2.2 吞吐量（throughput）

吞吐量是路由器在不丢包情况下所能达到的最大转发速率。

吞吐量与路由器端口数量、端口速率、数据包长度、数据包类型、路由计算模式（分布或集中）、测试方法有关。一般泛指处理器处理数据包的能力。

由于路由器设计使用在不同目的和应用环境，对吞吐量作范围限制没有意义。本标准对吞吐量不作规范。只作为重要的性能指标供比较。

9.2.3 转发延迟

路由器延迟指需转发的数据包最后一比特进入路由器端口到该数据包第一比特出现在端口链路上的时间间隔。

该时间间隔是存储转发方式工作的路由器的处理时间。对于直通转发（cut through）方式工作的设备可能会得到负的延迟（该种设备在收到部分数据包后即开始转发）。

通常所测试的延迟是指测试仪表发出数据包到经过路由器转发后收到该数据包的时间间隔。上述延迟与测试数据包的长度及链路速率都相关。

延迟对网络性能影响较大，特作如下规范：

在最坏情况下，1518byte 长度及以下的 IP 包延迟均应小于 1ms。

9.2.4 标记交换丢包率

标记交换丢包率是指路由器在稳定的持续负荷下由于资源缺少在应该交换的 MPLS 帧中不能交换的帧比例。

标记交换丢包率通常用作衡量路由器在超负荷工作时路由器的性能。

由于路由器设计使用在不同目的和应用环境，对标记交换丢包率作范围限制没有意义。本标准对丢包率不作规范。只作为重要的性能指标供比较。

9.2.5 标记交换吞吐量

标记交换吞吐量是路由器的包背板交换容量。

标记交换吞吐量与路由器端口数量、端口速率、数据包长度、数据包类型、路由计算模式（分布或集中），测试方法有关。一般泛指处理器处理数据包的能力。

由于路由器设计使用在不同目的和应用环境，对标记交换吞吐量作范围限制没有意义。本标准对吞吐量不作规范。只作为重要的性能指标供比较。

9.2.6 标记交换转发延迟

标记交换延迟指需交换的标记帧最后一比特进入路由器端口到该数据包第一比特出现在端口链路上的时间间隔。

该时间间隔是存储转发方式工作的路由器的处理时间。对于直通转发（cut through）方式工作的设备可能会得到负的延迟。（该种设备在收到部分数据包后即开始转发）。

通常所测试的延迟是指测试仪表发出标记帧到经过路由器转发后收到该标记帧的时间间隔。上述延迟与测试标记帧的长度及链路速率都相关。

延迟对网络性能影响较大，特作如下规范：

在最坏情况下，1518byte 长度及以下的标记帧延迟均应小于 1ms。

9.2.7 错序比

错序比指路由器转发数据包时错序包所占总包数的比例。

本标准对错序比不作规范，只作为重要的性能指标供比较。

9.2.8 路由表容量

路由表容量指路由器运行中可以容纳的路由数量：

- 1) 系统应能够支持至少 250, 000 条路由，平均每个目的地址至少提供 2 条路径；
- 2) 系统应支持至少 500 个 BGP 对等；
- 3) 系统应支持至少 1000 个 IGP 邻居。

9.2.9 转发表

转发表容量指路由器运行中可以容纳的转发表条目数量。

由于路由器设计使用在不同目的和应用环境，对转发表容量作范围限制没有意义。本标准对转发表容量不作规范，只作为重要的性能指标供比较。

9.2.10 标记转发表

标记转发表容量指路由器运行中可以容纳的标记转发表条目数量。

由于路由器设计使用在不同目的和应用环境，对标记转发表容量作范围限制没有意义。本标准对标记转发表容量不作规范，只作为重要的性能指标供比较。

9.2.11 标记交换路径

LSP容量指路由器运行中可以建立并转发的标记交换路径数量。

由于路由器设计使用在不同目的和应用环境，对LSP容量作范围限制没有意义。本标准对LSP容量不作规范，只作为重要的性能指标供比较。

9.3 可靠性和可用性要求

1) 系统应达到或超过 99.999% 的可用性。

2) 无故障连续工作时间

系统的无故障工作时间: MTBF>40 万小时。

3) 控制和数据路径应分离。

4) 要求设备具有高可靠性和高稳定性。主处理器、主存、交换矩阵(如果存在)、电源、总线仲裁器和管理接口等系统主要部件应具有热备份冗余。各种线路卡要求提供远端测试诊断功能。当某个系统电源故障时应能保持连接的有效性。

5) 系统应支持热插拔功能。

10 定时和同步要求

10.1 同步方式

路由器应采用主从同步方式。

10.2 外定时方式

路由器应设置外定时源输入接口,从通信楼定时供给设备(BITS)获得定时。同步接口可以为 2048kbit/s 或 2048kHz,优选 2048kbit/s 接口。

10.2.1 2048kbit/s 接口

2048kbit/s 接口物理/电气参数特性应符合 GB/T 7611 的要求,帧结构应符合 ITU-T G.704 的要求。

10.2.2 2048kHz 接口

根据工程需要,路由器也可提供 2048kHz 接口,物理/电气参数特性应符合 GB/T 7611 的要求。

10.3 线路定时

路由器应具有从线路信号中恢复定时,并用于同步的功能。当不能使用外定时方式时,采用线路定时。

10.3.1 SDH 传输线路定时

SDH 传输线路定时信息从 STM-N 线路信号中获得,帧结构应符合 ITU-T G.707 的要求,STM-1 电接口的物理/电气参数特性应符合 YD/T 877-1996 的要求,STM-N 光接口物理参数特性应符合 GB/T 20185-2006 的要求。

10.3.2 同步状态信息(SSM)

根据工程要求,路由器应能支持 STM-N 接口同步状态信息字节 SSMB(S1)功能,其要求应符合 ITU-T G.707 的规定。

10.4 路由器内部时钟要求

10.4.1 路由器配备的内部时钟等级

路由器内部时钟应采用二或三级时钟设备。

10.4.2 路由器内部时钟性能

路由器内部二级时钟主要性能要求如下:

1) 时钟单元可采用有保持功能的高稳晶体时钟。

2) 保持稳定度: B 类 $<1 \times 10^{-9}/\text{天}$; A 类 $<5 \times 10^{-10}/\text{天}$ 。

3) 自由运行频率准确度: $\pm 4 \times 10^{-7}$ 。

- 4) 牵引范围: $\pm 4 \times 10^{-7}$ 。

路由器内部三级时钟主要性能要求如下:

- 1) 时钟单元可采用有保持功能的高稳晶体钟。
- 2) 保持稳定度: 优于 $\pm 1 \times 10^{-8}$ /天
- 3) 自由运行频率准确度: $\pm 4.6 \times 10^{-6}$
- 4) 牵引范围: $\pm 4.6 \times 10^{-6}$

10.4.3 时钟工作方式

- 1) 快捕: 开机后首先进入快捕工作方式。
- 2) 正常(跟踪): 由快捕工作方式自动转入正常工作方式。
- 3) 保持: 路由器在失去全部输入频率基准合, 时钟自动进入保持工作方式。
- 4) 自由运行: 路由器应能人工选择自由运行工作方式, 用于时钟的自检、频率调整以及时钟的测试等。

10.4.4 时钟输出端的相位稳定性

- 1) 相位不连续性

对时钟进行的不经常的内部操作, 应满足:

- a) 在 2^{11} UI 内的任何时间, 相位变化应不超过 $1/8$ UI。
- b) 对大于或等于 2^{11} UI 时间, 每个 2^{11} UI 的间隔内的相位变化应不超过 $1/8$ UI, 并且漂移总量不超过 $1 \mu\text{s}$ 。

- 2) 长期相位变化

- a) 理想工作状态: 在输入频率基准无损伤的条件下, 对任何 $\geq 100\text{s}$ 的周期内, 时钟输出端的最大相对时间间隔误差 (MRTIE) 应不超过 $1 \mu\text{s}$ 。

- b) 实际工作状态: 待定

- c) 保持工作状态: 在保持(记忆)工作的情况下, 时钟的输出在任何 S 秒周期内的 MRTIE 不应超过下列限值:

$$\text{当 } S \geq 100, \text{MRTIE}(S) = [as + (1/2)bs^2 + C] \text{ ns}$$

其中, 参数 a 取值为 10.0, 即相当于初始频率偏差 1×10^{-8} 参数 b 取值为 2.3×10^{-4} 即相当于频率偏移为 2×10^{-8} /天, 参数 c 取值为 1000。

10.4.5 时钟可靠性

- 1) 平均故障间隔时间

路由器时钟的平均故障间隔时间: $\text{MTBF} > 10$ 年

- 2) 时钟的备用冗余度

路由器应设置两个性能相同的独立的时钟(主用和备用), 当一个时钟发生故障时, 另一个时钟应能立即正常工作。

10.4.6 时钟的可维护性

路由器时钟应具有频率粗调、微调功能, 生产厂家应提供维护方法, 应尽可能在现场调节频率。

10.5 频率基准的保护倒换

路由器应至少有两条同步链路的输入口, 即主用和备用。当失去频率基准后应自动倒换到备用频率

基准，倒换过程中不应产生滑动。如果备用频率基准也发生故障，则应自动转入保持工作状态。

10.6 同步性能的监测、告警和控制

路由器与数字链路相连，除接受输入的频率基准外，它应发出用于维护的告警和工作状态的显示，并能接受控制信息，其大中部分功能由同步设备来完成。

10.6.1 告警

对下述情况应能从控制中心和本设备自动检测并发出告警。

- 1) 对任何输入的 2048kbit/s 数字信号，每 24h 发生 4 次滑动，产生一般性告警
- 2) 锁相环路频率调节范围的临界告警。由于进钟晶体的老化而导致固有的时钟频率偏离锁相环的控制范围（控制信号超出时钟调节范围的 3/4）时发出一般性告警。
- 3) 路由器失去输入频率基准 10min 或连续错帧 10min 产生一般性告警。
- 4) 路由器若频率基准之一发生故障或降质（ $(\Delta f/f) \geq 2 \times 10^{-8}$ ）应产生一般性告警。如全部输入频率基准发生故障或降质（ $(\Delta f/f) \geq 2 \times 10^{-8}$ ）应该产生严重告警。
- 5) 对任何输入的 2048kbit/s 数字信号，每 24h 发生滑动次数等于或多于 255 次产生严重告警。
- 6) 路由器失去频率基准 24h 或连续错帧 24h 产生严重告警。
- 7) 时钟本身发生故障，例如恒温槽故障、时钟停止工作等发出严重告警。

10.6.2 工作状态显示

为了解同步设备的运行情况，以便正确维护交换设备，应对下列项目进行监测并给出可见的显示信号：

- 1) 时钟的工作方式，即快捕、跟踪、保持和自由运行；
- 2) 在使用的频率基准；
- 3) 在使用的时钟；
- 4) 上一次频率基准的倒换时间；
- 5) 输入频率基准的错帧率（错帧次数/小时或分钟）；
- 6) 人为强制状态应给予显示；
- 7) 相位变化达到或超过规定限值应计数。

10.6.3 控制

- 1) 在本地或控制中心，可实施下列人工控制功能：
 - a) 选择时钟的工作状态（快捕、跟踪和保持）；
 - b) 倒换时钟；
 - c) 倒换频率基准；
 - d) 切断频率基准。

- 2) 路由器设备的输入频率基准的功能。

路由器设备除在本地设自动倒换外，还需要在控制中心设人工倒换功能。

- 3) 同步设备应有自检、诊断和适用于维护的功能。

- 4) 对第 9.6.1 节中的 1)、2)、5) 项可根据需要修改软件。

11 操作管理维护要求

11.1 概述

本章描述路由器在设备管理、网络管理、设备的运行、操作、维护方面的要求以及在安全性方面应提供的功能。

11.2 网络管理协议

11.2.1 简单网络管理协议-SNMP

路由器应支持 IETF RFC3416、RFC3417 和 RFC3418 中规定的 SNMP v2。

SNMP 应使用 UDP/IP 作为传输层/网络层协议。也可以使用其他协议（例如 IETF RFC1418 和 IETF RFC1089）。

SNMP 管理请求向路由器任何一个接口的 IP 发出时，该操作应生效。实际的管理动作应由路由器或路由器的代理完成。

支持 SNMP v2 协议的路由器应实现 SNMP v2 MIB（见 IETF RFC3418）。

路由器应实现所有的 SNMP 操作。

路由器应提供一种机制来限制 SNMP 自陷（trap）消息的产生速率。路由器可以通过 IETF RFC1224 中描述的异步告警管理算法来实现上述机制。

11.2.2 团体（community）表格

为本标准描述方便，假设路由器中存在一个抽象的团体表格，该表格包含多个条目，每个条目给一个特定区域，包含完全定义该区域属性需要的参数。对抽象团体表格的实现方法在本标准范围之外，由实现者决定。

路由器的团体表格建议至少包含两个条目。

路由器应允许用户手工（即不使用 SNMP）检查、增、删、改 SNMP 团体表格中的条目。用户应能够设置区域名，或者构造 MIB 视图。用户应能以只读（即不允许 SET）或者读写（允许 SET）的方式配置区域。

用户应能定义至少一个 IP 地址，当使用自陷（trap）时，对每个捕获或 MIB 视图的通知将送到该 IP 地址。这些 IP 地址应定义在区域或 MIB 视图库内。允许或不允许在区域或 MIB 视图库上发通知应是可配置的。

路由器应提供为特定区域提供有效管理员列表的能力。如果提供上述列表，路由器应检验 SNMP 数据包源地址的有效性，如果该地址没有在上述列表中出现则应丢弃该数据包。如果数据包被丢弃，路由器应采取 SNMP 认证失败时的相应措施。

区域表应存储在非易失性存储器内。

区域表的初始状态应包含一个条目，其中区域名串为 Public，访问权限为只读。该条目的缺省状态为不允许发送自陷（trap）。如果实现，该条目应保存在区域表中，直到管理员改变或者删除。

11.2.3 标准 MIB

路由器应实现所有关于路由器配置的 MIB 有：

支持 SNMP v2 协议的路由器应实现 SNMP v2 MIB（见 IETF RFC3418）。

路由器应实现 MIB-II（见 IETF RFC1213）中的系统、接口、IP、ICMP 和 UDP 组。

路由器应实现接口扩展 MIB（见 IETF RFC2863）。

如果路由器实现 TCP（例如，远程登录），应实现 MIB-II（见 IETF RFC1213）中的 TCP 组。

如果路由器实现 OSPF v2，应实现 OSPF MIB（见 IETF RFC4750）。

如果路由器实现 BGP4，应实现 BGP4 MIB（见 IETF RFC4273）。

如果路由器 FDDI 接口实现 ANSI SMT 7.3，应实现 FDDI MIB（见 IETF RFC1512）。

如果路由器 FDDI 接口实现 ANSI SMT 6.2，应实现 FDDI MIB（见 IETF RFC1285）。

如果路由器有 E3 接口，应实现 E3 MIB（见 IETF RFC3896）。

如果路由器在任何接口上支持 PPP 及 IP NCP，应实现 MIB（IETF RFC1471、IETF RFC1472 和 IETF RFC1473）。

如果路由器支持 RIPv2，应实现 RIPv2 MIB（见 IETF RFC1724）。

如果路由器有 ATM 接口，应实现 ATM MIB（见 IETF RFC2515）。

如果路由器支持 ATM 上支持传统 IP 及 ARP，应实现 MIB（见 IETF RFC2320）。

如果路由器有吉比特以太网，应实现吉比特以太网 MIB（见 IETF RFC3635）。

如果路由器有 SDH 接口，应实现 SDH MIB（见 IETF RFC3592）。

如果路由器支持 IP/ICMP，应实现 MIB（见 IETF RFC4293）。

如果路由器支持 TCP，应实现 MIB（见 IETF RFC4022）。

如果路由器支持 UDP，应实现 MIB（见 IETF RFC4113）。

如果路由器支持 SNMP v2，应实现 SNMP v2MIB（见 IETF RFC1907）。

如果路由器支持接口协议，应实现 MIB（见 IETF RFC2863）。

如果路由器支持接口转发表，应实现 MIB（见 IETF RFC4292）。

如果路由器支持 VRRP，应实现 VRRP MIB（见 IETF RFC2778）。

如果路由器支持用于网管接口的 RS-232 接口，应实现 MIB（见 IETF RFC1659）。

11.2.4 RMON MIB

路由器应支持 RMON MIB（见 IETF RFC2819/RFC4502）。其中，路由器应支持 RMON 第 1 组（以太网统计数据组），第 2 组（历史记录控制组），第 3 组（以太网历史记录组），第 4 组（告警组），和第 10 组（事件组）。可以选择支持第 5 组（主机组），第 6 组（前 N 个主机组），第 7 组（矩阵组），第 8 组（筛选组）和第 9 组（包捕获组）。

11.2.5 厂商特定的 MIB

互联网标准和根据实验的 MIB 不能完全覆盖网络单元统计、状态、配置和控制信息。路由器厂商可以自己开发覆盖上述信息的 MIB 扩展，这些 MIB 扩展称为厂商特定的 MIB。

由于这些信息不能由标准或实验得到的 MIB 得到，厂商特定的 MIB 应提供存取这些统计、状态、配置和控制信息的方法，而且这些信息能用于监视和控制操作。

厂商应根据 IETF RFC1155 的规定使所有厂商特定的 MIB 变量可用，并以 IETF RFC1212 规定的方式来描述。

11.2.6 保存改变

通过 SNMP 调整的参数可以存储在非易失性存储器中。

11.3 自举协议（BOOTP）

11.3.1 概述

自举协议 (BOOTP) 是基于 UDP/IP 的协议, 该协议允许启动的主机自我动态配置, 无需用户干预。BOOTP 协议提供通知主机有关被赋予的 IP 地址、BOOTP 服务器主机的 IP 地址以及将要下载到内存并执行的、文件的、名字的方法。其他配置信息例如本地前缀长度或子网掩码, 本地时间偏移, 缺省路由器地址以及各种互联网服务器的地址同样通过 BOOTP 通知主机 (见 IETF RFC2132)。

11.3.2 BOOTP 中继代理

有时, BOOTP 客户以及相关联的 BOOTP 服务器不在同一 IP (子) 网中。在这种情况下, 需要一个第三方的代理在客户与服务器之间转发 BOOTP 消息。为了与路由器 IP 转发功能区别, 这个第三方的代理称为 BOOTP 中继代理。

中继代理的功能通常在路由器上实现。

路由器可以提供 BOOTP 中继代理功能。如果实现, 应符合 IETF RFC1542 中的规定。

在第 7.4.1.3 节中讨论了对数据包本地分发给路由器的环境。所有目标端口号为 BOOTP (67) 的 UDP 消息应由路由器的逻辑 BOOTP 中继代理处理。

为了支持 BOOTP 中继代理, 路由器应接收 IP 源地址是 0.0.0.0 的 BOOT REQUEST 消息, 并本地分发给中继代理。

11.4 运行维护要求

11.4.1 定义

11.4.1.1 有关设备性能

路由器的 O&M 中应包含以下措施:

- 设备资源利用率;
- 网络接口带宽利用率;
- 丢包率;
- 设备软件运行情况;
- 核心路由器的配置情况;
- 开关机配置;
- 安装或升级新硬件;
- 安装或升级新软件;
- 监视路由器及相连网络的状态及性能;
- 流量统计的收集;
- 以及上述措施的协调等。

11.4.1.2 有关设备告警

路由器的 O&M 中应包含以下措施:

- 诊断路由器的处理器、网络接口、相连的网络或通信链路的硬件问题;
- 故障定位、记录和存储;
- 故障告警;
- 在宕机后重新启动或重新引导路由器;
- 配置 (重新配置) 路由器;
- 发现及诊断互联网问题例如拥塞、路由环回、差错 IP 地址、黑洞、包雪崩、主机的错误行为;

- 暂时的或者永久的网络拓扑改变；
- 运行软件的故障检查及记录等。

路由器以及相连的通信链路通常作为一个系统由集中的 O&M 组织来维护运行。该组织可能通过一个网络运行中心 (NOC) 来执行 O&M 功能。由于路由器可能与 NOC 连接在不同网上，路由器支持 NOC 从互联网远程监视及控制非常重要。由于网络故障通常会终止网络访问，NOC 要求路由器应支持通过一条备用途径，通常是接在路由器配置口上的调制解调器来实现网络管理。

因为在互联网中传输的 IP 包通常会使用多于一个 NOC 控制下的路由器，互联网故障诊断将牵涉到多个 NOC 的合作。在某些情况下，路由器需要超过一个 NOC 来监视，但由于过多的监视会损害路由器性能，因此只有在必要时才可以这样做。

11.4.2 路由器初始化

11.4.2.1 最少的路由器配置

在路由器能转发包以前，存在一个最少的路由器配置条件：

- 1) 路由器知道该物理接口上相关联的至少一个逻辑接口的 IP 地址和子网掩码或网络前缀长度；
- 2) 路由器知道该接口是非编号接口，并且路由器知道其路由器 ID。
- 3) 路由器不允许使用厂商配置的缺省 IP 地址、前缀长度和路由器 ID。
- 4) 路由器不允许假设一个没有配置的接口是一个非编号接口。

11.4.2.2 地址及前缀初始化

路由器应允许静态配置 IP 地址，地址掩码或前缀长度，并存储在非易失性存储器中。

路由器可以在系统初始化过程中动态得到 IP 地址和相应掩码。

如果提供动态方式，在特定路由器中是否使用动态方式应可配置。

如上文所述，路由器的 IP 地址中主机地址部分和网络前缀部分不允许是全 0 或全 1。所以路由器应不允许将 IP 地址设置成上述形式。

路由器应对设置的掩码实施下述检查：

- 掩码既不是全 0 也不是全 1（前缀长度不为 0 或 32）。
- 相应与地址网络前缀部分的比特为全 1。
- 对应于网络前缀部分的比特是连续的。

11.4.2.3 使用 BOOTP 或 TFTP 的网络引导

路由器可以使用 TFTP 从网络引导，也可以使用 BOOTP 从服务器下载启动映像文件。

BOOTP 是引导端系统的协议，供路由器使用时需要进一步的扩展。如果路由器使用 BOOTP 寻找当前启动主机，它应使用第一个接口的硬件地址发送 BOOTP REQUEST，或者该路由器已预先配置，使用另一个接口的硬件地址，或者在 BOOTP 包的硬件地址字段中写入另一个数。这允许路由器无需硬件地址（例如只有同步口的路由器）就可以使用 BOOTP 用来启动发现。TFTP 可随后用作检索在 BOOTP REPLY 中找到的映像。如果没有预先配置使用的接口以及接口数量，路由器可以循环使用各接口硬件地址直到 BOOTP 服务器发现一个匹配。

路由器应具有将从 BOOTP 得到的参数存储到非易失性存储器的能力。路由器可以具有将从网络得到的系统映像存储到本地稳定存储器的能力。

路由器可以有一种机制允许远端用户要求该路由器得到一个新的引导映像。路由器应区分从下面 3 个地方获得的新引导映像：包含在请求中的引导映像；从上一个启动映像服务器得到的引导映像和通过

BOOTP 定位一个服务器得到的引导映象。

11.4.3 运行和维护具体规定

11.4.3.1 定义

在路由器上实施 O&M 功能有多个可用的模型：一个是仅在本地模型，该模型要求 O&M 功能只能在本地执行（例如，接在路由器上的终端）；一个是完全远程管理，在本地只允许作最少的操作（例如，强迫引导），大多数 O&M 从远端由 NOC 执行；另一个是中间模型，例如 NOC 人员可以登录到路由器上作为一个主机，使用 Telenet 协议执行本地也能申请的功能。仅在本地模型一般在路由器安装时使用，路由器通常需由 NOC 远端操作，所以路由器应实现远端操作。

远端 O&M 功能可以通过控制代理（程序）实现。在直接应用中，O&M 功能直接由 NOC 通过标准互联网协议实现（例如，SNMP、UDP、TCP）。在间接应用中，控制代理支持这些协议并控制路由器使用恰当的协议。建议使用直接应用的方式。

厂商应提供这样一种环境：用户使用控制代理或其他 NOC 软件应象在标准操作系统中编程一样。使用标准互联网协议 TCP 和 UDP 应能帮助实现上述要求。

路由器远程监视和远程控制存在重要的访问控制问题：一方面应确保应用这些功能时路由器资源的有效控制，例如路由器监视时应不过分占用 CPU 资源；另一方面，O&M 功能应具有相对高的优先级，因为路由器拥塞的时候通常是最需要 O&M 操作的时候。

11.4.3.2 带外访问

路由器应提供带外（OOB）访问。OOB 访问应提供所有带内访问的功能。带外访问应实现访问控制，防止非法访问。

11.4.3.3 路由器 O&M 功能

11.4.3.3.1 硬件诊断维护

在本地硬件维护时，每个路由器应作为一个独立设备来操作，在路由器处应提供运行诊断程序的工具和方法。路由器应能在故障情况时运行诊断程序。

11.4.3.3.2 下载内存和重新引导的控制

路由器应同时提供带内和带外的机制来使网络管理员重新装载、停止、重新启动路由器。路由器应提供一种机制（例如 watchdog 定时器），当路由器因为软硬件差错挂起一定时间后，自动重新启动。

路由器应实现一种机制将路由器的内存（和/或路由器宕机后所有对厂家调试有用的状态信息）保存到本地稳定存储设备或者通过在线转储机制（例如 TFTP）保存到另一台主机中（见 IETF RFC1195、IETF RFC1123）。

11.4.3.3.3 配置路由器的控制

每台路由器都有需要配置的参数。路由器参数更新后应不需要重新引导，最坏情况下需要重新启动。可能存在某些情况，改变参数后应重启路由器（例如改变某接口的 IP 地址）。这些情况下，应小心将对路由器和周边网络的影响减少到最小。

应存在一种方法自动或人工地从网络配置路由器。路由器应能从另一台路由器或主机下载/上载配置参数。路由器应提供一种方法，无论作为应用程序方式或者路由器功能方式，能相互转换配置参数格式和人工可编辑格式。路由器应具有某种稳定的存储器存储配置。路由器不应信任例如 RARP，ICMP 地址掩码响应等协议，也可以不信任 BOOTP 协议。

11.4.3.3.4 系统软件网络引导

路由器应将系统软件保存在非易失性存储器中，例如 PROM、EPROM 或者磁盘中。路由器可以通过

网络从其他主机或路由器下载系统软件。

能将系统软件保存在本地非易失性存储器中的路由器可以实现配置成从网络引导系统软件。实现上述功能的路由器应配置成在无法从网络引导系统时可以从本地存储器引导系统。

路由器可以给予不同系统软件区分不同配置。如果不同版本软件的配置命令有所改变，路由器应能兼容上一版本的配置命令。

11.4.3.3.5 对差错配置的检测和响应

路由器应实现一种机制检测差错配置并做出响应。如果命令不正确运行，路由器应给出差错消息。路由器不应接受差错格式的命令，即使该命令本身是正确的。

另一种差错是对路由器连接网络的差错配置。路由器可以实现检测网络的误配置。路由器可以将发现的差错记录到日志或者网络上其他路由器或主机，管理员可以看到可能存在的问题。

11.4.3.3.6 最少干扰

对路由器配置的改变应最小程度地影响网络。当在路由器上作很小的改动时，路由表不应没有必要地刷新。如果路由器上运行多个路由协议，停止一个路由协议不应干扰其他路由协议，除非某网络需要通过多个路由协议获得路由。

11.4.3.3.7 管理方式

路由器应提供以下管理接入方式：Telnet、SNMP（MIB II）、RS232 接口，为基于 SNMP 的管理提供的带外分组接口（如以太网接口）。

路由器应支持以下管理方式：通过接入平台提供基于 SNMP 的配置和监控，通过 Telnet 或 RS232 提供 CLI 接口，配置应可下载和上载，应支持脚本语言以使复杂的管理任务自动化。

11.5 计费信息统计功能

路由器应提供包数、字节数、端口和业务类型等信息统计功能。

11.6 ATM 操作和维护（OAM）功能要求

本节所规定的 OAM 功能适用于所有类型的 ATM 连接。目前仅要求点到点的 ATM 连接的 OAM 功能。

11.6.1 物理层 OAM 功能

1) 基于 SDH 传输系统的物理层的 OAM 功能

与 SDH 有关的 OAM 功能应符合 ITU-T G.832/G.804 和 ITU G.782/G.783/G.784 相关要求。

ATM 特定的 OAM 功能包括信头差错监控、信元定界丢失（LCD）检测和路径 RDI 检测、空闲信元的插入/抑制缺陷的检测以及差错性能的监控/报告。

2) 基于信元的传输系统的物理层的 OAM 功能

基于信元的传输系统的物理层 OAM 功能包括信头差错监控、性能监控和报告、故障设备的定位以及对以下缺陷的检测：

- 信号丢失（LOS）或 F1 物理层 OAM 信元识别丢失的检测；
- 不可接受的差错性能的检测；
- 空闲信元的插入/抑制的缺陷的检测；
- F3 物理层 OAM 信元的识别的丢失的检测；
- 信元定界丢失缺陷的检测。

基于信元的传输系统的物理层的 OAM 功能应符合 YDN 052-1998 和 ITU-T I.432。

11.6.2 ATM 层的 OAM 功能

要求提供以下的 ATM 层 OAM 功能：

1) AIS 和 RDI 功能

在 VP 和 VC 两级都支持 AIS 和 RDI 故障管理功能。只要求提供端到端 AIS 和 RDI 信元，段 AIS 和 RDI 信元暂不要求支持。

2) 连续性检验功能

连续性检验功能是任选功能。

连续性检验信元在连续性检验功能激活后，以标称值每秒一个周期地产生且下行发送。

3) 环回能力

环回信元的在下行的一个点上的回送，或可遵循来自系统管理的指令进行，或可遵循包含在它的信息字段中的信息进行。

环回信元可以在连接点以及段和连接端点上插入。始发环回信元的 VP 或 VC 连接点，在对该环回信元的相关标记和源点标识符进行匹配后，应消除该环回信元。

段环回信元可以在连接点以及段和连接端点回送。VP 或 VC 连接点，在回送段环回 OAM 信元后，应消除此段环回 OAM 信元。在连接点环回情况，也采用环回位置标识符。

除了用来自网管的命令启动一个环回外，还应有其他手段来启动一个环回，例如，用户可以启动一个端到端环回。但这种情况仍须向 TMN 报告环回结果。

环回点可改变环回信元净荷中的环回指示字段。

端到端环回信元不应在中间连接点上回送。

4) 性能管理功能

可任选地提供性能管理功能。

性能监控应通过用户信元的监控块来进行，块大小 N 的标称值暂定为 128、256、512 和 1024，其他值可选。

性能监控可以在每一接口上所挑选的一定数量的 VPC 或 VCC 上同时执行。

性能监控可以在连接建立期间或连接建立后的任何时间激活。

在激活性能监控后，所接收到的第一个性能监控信元仅用于初始化，而不用来更新性能参数。

5) 激活/去活程序

激活/去活功能根据性能监视和连续性检验功能的提供而提供。

性能监控和连续性检验可以在连接/段的建立期间或连接/段建立以后的任何时间上激活。此类激活（以及相关的去活）由网管系统或由终端用户启动。在网管系统或终端用户请求性能监控或连续性检验的激活/去活以后，在连接（或连接段）的两个端点间需要一个启动程序来正确启动 OAM 过程。

启动程序的执行有两种方法：

- 用激活/去激活 OAM 信元；
- 完全通过网管系统。

第一种方法可任选支持，第二种方法要求支持，有的情况，如对于性能监控或连续性检验建立在一个连接或连接段（它的端点包含在单个管理范围内）的这种情况，OAM 功能激活/去活也可完全由网管系统来执行。

12 硬件要求

12.1 硬件系统基本要求

- 1) 应采用模块式的硬件结构, 便于扩充, 并能容纳新业务和新技术。
- 2) 提供的设备, 应全部采用经过老化测试和严格筛选的优质元器件, 组装过程应有严格的质量控制, 确保长期使用的高稳定性, 高可靠性。
- 3) 系统构成应具有冗余和容错等安全措施。
- 4) 当软件升级时, 应不影响硬件结构。
- 5) 如果系统包含交换矩阵, 交换矩阵应有一定的安全措施, 以保证系统的可靠性。
- 6) 应具有网络故障和硬件故障告警功能。

12.2 对处理机的要求

- 1) 处理机系统要有冗余度, 遇到处理机软硬件故障时, 具有倒机、分级再启动及系统再生等性能, 以保证其安全可靠。
- 2) 处理机系统具有故障脱机自动诊断功能。
- 3) 处理机系统应具有过负荷控制措施。
- 4) 处理机系统应具有软、硬件故障告警信号。

12.3 对输入、输出设备的基本要求

- 1) 人机命令尽可能采用菜单方式, 用作人机命令输入的设备应具有冗余度。
- 2) 显示器、各类告警信号除由打印机打印外, 还应在显示屏上显示, 且能用不同彩色显示出各类故障的严重程度。

13 软件要求

13.1 基本要求

- 1) 要求软件采用模块化结构, 模块之间的通信应按接口进行。任何一层的任何一个模块的维护和更新以及新模块的追加都不应影响其他模块。
- 2) 配置数据与处理程序应有相对的独立性, 配置数据的任何变更都不应引起运行版本程序的变更, 处理程序应与任何局的配置数据相适应。
- 3) 软件应有容错能力, 一般小的软件故障不应引起各类严重的系统再启动。
- 4) 软件设计应有防护性能, 某一软件模块内的软件差错应限制在本模块内, 而不应造成其他软件模块的差错。
- 5) 应具有软件运行故障的监视功能, 一旦软件出现死循环等重大故障时, 应能自动再启动, 并作出即时故障报告信息。
- 6) 在未达到设备的最终容量时, 增加或减少设备容量时, 只需变更配置数据, 并仅需使用一般的人机命令即可, 不应影响正常通信。
- 7) 同种型号的路由器应采用同一种软件版本, 同一型号路由器的不同时间的软件版本应能兼容。

13.2 软件功能要求

- 1) 要求有完善的实时操作系统。
- 2) 要求有完善的各类协议处理功能和路由处理功能。
- 3) 要求具有网管子系统及处理相应业务的功能, 要求具有路由变更控制功能和输入业务量、输出业

务量控制功能。

4) 要求具有完善的系统结构控制功能,可以灵活地组合路由器中完好的设备,构成运行系统。

5) 要求具有对种硬件设备测试的功能。

6) 要求具有对软件、硬件运行故障的监视功能,有完善的故障告警及障碍后处理功能。

7) 要求具有完善的、方便的人机通信控制功能。

8) 要求具有完善的维护管理功能,具有配置的维护管理,业务量观察管理、软件维护管理、设备维护管理、计费管理等功能。

9) 要求具有故障诊断和故障定位功能。故障中断定位后应能显示或打印,报告故障设备的物理位置等有关信息。

13.3 软件维护管理功能要求

1) 要求具有在不中断通信的情况下,完成程序打补丁的功能。补丁区应集中专用,如果总补丁数超过 100 个时,要求厂家无偿地提供新版本。

2) 要求对于全部局数据和用户数据都可以在不影响呼叫接续处理的情况,用人机通信方式进行下述操作:

a) 数据查询;

b) 数据修改变更;

c) 数据追加;

d) 由磁带或其他媒介进行批量数据的引入运行;

e) 原运行数据的暂存,重新运行,使用删除。

3) 如对修改后软件不满意或将修改后软件引入系统后,对系统有副作用或发现新版本有问题,应能方便而迅速地(在 1min 内)恢复到原来的程序。

4) 系统软件能在线升级,不需重启动。

5) 故障诊断软件的诊断精度:

——要求故障诊断软件能对硬件故障进行诊断和定位,故障诊断定位后应能显示或打印,报告故障设备的物理位置等有关信息。

——对硬件故障诊断定位的精度要求如下:对于各公共部件,如处理机、交换矩阵、线卡、存储器、输出/输入设备等的硬件故障应能达到:100%的故障能自动定位至 1 块板。

14 机械结构与工艺要求

14.1 概述

设备的总体机械结构,应充分考虑安装、维护的方便和扩充容量或调整设备数量的灵活性,实现硬件模块化。应具有足够的机械强度和刚度,设备的安装固定方式应具有防振抗震能力,应保证设备经过常规的运输、储存和安装后,不产生破损、变形。

厂方应提供设备的机械结构、品种规格及安装规程等方面的详细说明。

14.2 机架要求

1) 设备在预防意外撞击部位、可接触至布线的部位和危险电压的部位,均应提供罩盖,对高压等危险部位应有特殊标志。

2) 每一个机架在前方或背面应有清楚的标志。

3) 插入模块应有导向。

4) 厂家应提供为安装该系统所必需的铁架、支撑件、电缆支架、电缆走道、底座、底盘等。

14.3 接插件

1) 接插件应接触完全可靠，结构坚实，借助手或简单工具易于插入或拔出，并有定位和锁定装置。

2) 厂家应提供安装用应的端子板、连接条等。

14.4 布线及连接

1) 机架之间、机架内各机框之间采用接插件实现光/电连接。

2) 线缆在机架内排放的位置应设计合理，不得妨碍或影响日常维护、测试工作的进行。

3) 设备内的所有焊点不得有虚焊、假焊、漏焊和混线。厂方应保证不使用具有腐蚀性的助焊剂。

4) 厂家应提供与设备有关的全部布线及电缆，电缆两端应有编号标志。应提供布线及电缆的详细说明及有关的规范。

14.5 机械加工工艺

1) 零部件的形状尺寸，表面光洁度等技术参数应符合设计文件的规定。

2) 活动部分（如门及指示、控制面板等）应动作灵活、位置准确。

14.6 表面涂复处理

1) 设备的表面涂复，应满足安装地区的环境、气候所需的防腐、防蛀的要求。

2) 所有喷漆（塑）零件的表面应光滑平整，色泽一致，不允许有划痕、斑疵、流挂、脱落和破损。电镀零件的表面应有金属光泽，不允许有裂纹、锈点、毛刺和缺陷。

3) 机架（盘）、机台和外观应色彩协调，色调柔和，色泽一致。

14.7 印刷电路板

1) 所有印刷电路板，均应有防霉喷涂层，如采用深色覆盖涂层，需要在涂层外加印清楚的电路连接线条。

2) 印刷电路板应有插错保护功能。

3) 印刷板板面应平整，其翘曲的程度应以不影响印刷插件的顺利插拔或不造成插拔困难为限。

4) 每一印刷电路板均应标出名称或代号。安装在印刷板上的部（器）件，应有明显的与图纸一致的标志。其标志应方便维护人员查看，并应将所有部（器）件列表说明。

5) 各种印刷电路板均不允许有飞线。

6) 印刷电路板上应有插拔及锁定位置。

7) 同一品种的印刷电路板应具有完全的互换性。

14.8 可闻噪声及震动

可闻噪声及震动作出说明，以便于设备的使用维护部门采取相应的措施。

14.9 冷却、通风

设备的冷却应采用自然通风散热方式。厂家应对设备的散热要求提出详细说明。

15 过流过压保护要求

15.1 安全要求

设备的交流电源处应安装过压、过流保护器。出现以下两种情况时，过压、过流保护器能发挥保护作用：

- 雷电冲击线路设备；
- 通信线路与高压线过近而产生感应电压。

15.1.1 雷电冲击线路设备

设备在承受峰值电压为 1000V 的电压脉冲的冲击后，其部件性能应该不受任何影响。

15.1.2 通信线路与高压线过近而产生感应电压

设备应能经受住来自通信导线上的纵电势为 650V，0.5s 内的感应电压而不降低任何部件的性能。

15.2 过压自动恢复和过压告警

- 1) 当路由器遭遇上述各种过压冲击时，应能立即产生告警信号。
- 2) 保护设备在经受雷电冲击和高压线感应等冲击后，应能自动恢复，无需维护人员干预。

15.3 防电涌破坏

设备应带有防电涌器件，有效防止电涌对设备的损坏。

15.4 绝缘电阻

正常情况下，设备的外壳与电源间的绝缘电阻应不小于 50MΩ。

16 环境要求

16.1 环境温、湿度要求

路由器在以下温、湿度条件下的机房中应能正常工作，见表 20。

表 20 环境温、湿度要求

设备名称	温 度	(℃)	相对湿	度 (%)
	长期工作条件①	短期工作条件②	长期工作条件	短期工作条件
路由器	15℃~30℃	0℃~45℃	40%~65%	20%~90%

注 1: 机房内工作环境温、湿度的测量点，指在机架前后没有保护板时测量，距地板以上 1.5m 和距交换机架前方 0.4m 处测量的数值。

注 2: 短期工作条件指连续不超过 48h 和每年累计不超过 15 天。

注 3: 极端恶劣工作环境，一般指机房空调系统出现故障时可能出现的环境温度和湿度值。

每次不应超过 5h 能恢复正常工作范围

16.2 机房地面要求

要求机房地面具有良好的防静电性能。地板绝缘电阻应满足表 21 要求。

表 21 地板绝缘电阻

阻值要求分档	每档绝缘电阻值	说 明
最小绝缘电阻	$25 \times 10^3 \Omega$	
最大绝缘电阻	$1 \times 10^6 \Omega$	对新地板要求
最大绝缘电阻	$1 \times 10^{10} \Omega$	地板寿命终了时

机架上、下应留空间，满足通风、防静电及布缆要求。当机房处在相对湿度较低的地区环境时，特别当相对湿度处在 20% 以下的时间里，应加强其抗静电措施。

16.3 机房的防尘和对有害气体浓度的要求

16.3.1 对防尘的要求

- 1) 机房中应无爆炸、导电、导磁性及腐蚀性尘埃。
- 2) 灰尘粒子直径大于 5μm 的浓度，应 $\leq 3 \times 10^4$ 粒/米³ 要求。

16.3.2 对有害气体浓度的要求

机房中应无腐蚀金属的和破坏绝缘的气体。

16.4 抗电磁干扰的能力

机房具有抗外界电磁干扰的屏蔽效应

路由器本身在 0.01~10000MHz 频率范围内, 受到电场强度为 140dB (V/m 的外界电磁波干扰时, 应不出现故障和性能的下降。

在交流、直流电源线对和信号线对受到表 22 所示的 0.01~100MHz 频率范围的外界电磁干扰感应电流时, 路由器应不出现故障和性能的下降。

表 22 外界电磁干扰感应电流

频率范围 (MHz)	最大线路感应电流 (dB μ A)
0.01~0.8	$-21.05\lg f + 67.9$
0.8~100	70

16.5 路由器本身产生的电磁干扰要求

路由器本身产生的电磁干扰应满足以下各限值:

1) 由路由器发射出的无线电电磁波干扰强度应满足表 23 限值。

表 23 无线电电磁波干扰强度

频率范围 (MHz)	干扰电磁强度 (dB μ A/m)
0.01~ 0.024	$148.6 - 60\lg d$
0.024~0.8	$116.2 - 60\lg d - 20\lg f$
0.8~1.59	$118.2 - 60\lg d$
$1.59 \sim 47.7/d$	$126.2 - 60\lg d - 40\lg f$
$47.7/d \sim 88$	$59.1 - 20\lg d$
88~216	$63.6 - 20\lg d$
216~10000	$66.6 - 20\lg d$

其中:

- (1) d 为测试天线与靠近被测物间水平距离, 单位为米, d 限于 30m 内。
- (2) f 为频率, 以 MHz 为单位。
- (3) dB μ V 表示以 dB μ V 为单位的绝对电压电平值

2) 由路由器系统进入交流馈电线的干扰电流限值, 应符合表 24 规定。

表 24 交流馈电线干扰电流

频率范围 (MHz)	最大线路干扰电流 (dB μ A)
0.000061~0.001	$I - 20\lg f - 84.4$
0.001~0.01	$(124.4 - I) \lg f + 348.8 - 2I$
0.01~0.8	$-21.05\lg f + 57.9$
0.8~100	60

其中:

- (1) f 为频率, 单位为 MHz。
- (2) I 为接入到交流电源处的输入线路电流电平。
- (3) dB μ A 为以微安 (μ A) 为参考单元的分贝数, 即绝对的电流电平值

3) 由路由器进入到直流馈电和信号线上的干扰电流限值, 应符合表 25 要求。

表 25 导线干扰电流

频率范围 (MHz)	最大导线干扰电流 (dBμA)
0.01~0.8	$-21.05\lg f + 57.9$
0.8~100	60

16.6 抗地震措施

路由器机架及设备应进行抗震加固，应能达到抗里氏 7 级（美氏 9 级）地震的能力。

16.7 运输和仓储要求

路由器设备应能适应不同的运输环境条件如防水、防震等，并应能在无空调条件下运输和仓储，而不影响装机开通之后的正常运行。

17 电源与接地要求

17.1 电源要求

17.1.1 直流电压及其波动范围要求：

额定电压：为-48V 的直流电源

电压波动范围：在直流输入端子处测量-48V 电压允许变动范围为-57~-40V。路由器在此范围内应工作正常。

17.1.2 杂音电压指标：

在直流配电盘输出端子处测量的限值如下：

- 300Hz<频率≤3400Hz杂音电压≤2mV有效值；
- 0Hz≤频率≤300Hz峰-峰值杂音电压≤400mV有效值；
- 3.4kHz<频率≤150kHz宽带杂音电压≤100mV有效值；
- 150kHz<频率≤30MHz宽带杂音电压≤30mV有效值。

17.1.3 离散频率杂音电压

- 3.4kHz≤频率≤150kHz，≤5mV有效值；
- 150kHz<频率≤200kHz，≤3mV有效值；
- 200kHz<频率≤500kHz，≤2mV有效值；
- 500kHz<频率≤2mHz，≤1mV有效值。

17.1.4 交流电压及其波动范围要求

单相 220V±10%，频率 50Hz±5%。

线电压波形畸变率小于 5%

17.2 接地要求

17.2.1 接地方式

路由器所在机房应采取各类通信设备的工作地、保护地以及建筑防雷接地共同合用一组接地体的集中接地方式，即为联合接地方式。

17.2.2 接地要求

1) 由联合接地体的垂直接地总汇集线上所接的水平接地分汇集线引入机房，路由器的各个机架设备的接地线就近引入水平接地分汇集线上。

2) 路由器各机架上的直流电源工作地应从接地汇集线上引入。

3) 各机架设备做工作接地，机壳和机架应作保护接地。

17.2.3 接地线截面积

接地线（指各种需接地的机架、地线等设备与水平接地分汇集线之间的连线），其截面积应根据可能通过的最大电流负荷确定。接地线应采用良导体（铜）导线，并且不准使用裸导线布放。

17.2.4 接地电阻值

路由器所在机房的联合接地的接地电阻值要求 $<5\Omega$ 。

附 录 A
(规范性附录)
SDH 上传送 IP 的技术要求

A.1 概述

本附录提供 IETF RFC 1619/RFC2615 规定的 SDH 上传送 IP 的技术要求。

A.2 SDH上传送IP的定义

SDH 上传送 IP (IP Over SDH) 以 SDH 网络作为 IP 数据网络的物理传输网络, 它使用链路适配及成帧协议对 IP 数据包进行封装, 然后按字节同步的方式把封装后的 IP 数据包影射到 SDH 的同步净荷封装 (SPE) 中, SDH 上传送 IP (IP Over SDH) 的定义如图 A.1 所示。

IP
适配与成帧协议
SONET/SDH 路径
SONET/SDH 复用段
SONET/SDH 再生段

图 A.1 SDH 上传送 IP (IP Over SDH) 的定义

目前广泛使用 PPP 协议对 IP 数据包进行封装, 并采用 HDLC 的帧格式, 即 IP/PPP/HDLC/ SDH, PPP 协议提供多协议封装、差错控制和链路初始化控制等功能, 而 HDLC 帧格式负责同步传输链路上的 PPP 封装的 IP 数据帧的定界, IP/PPP/HDLC/SDH 的协议堆栈如图 A.2 所示。

IP	— 客户数据包: IP v4、IPv6。
PPP	— IP 多协议封装; — 差错检验; — 链路初始化控制。
HDLC	— PPP 分组定界。
SONET/SDH 路径	— 路径复用和带宽管理 (提供 1.5Gbit/s 到 10Gbit/s 的带宽); — 连接检验; — 差错检验。
SONET/SDH 复用段	— 高速复用; — 线路故障分段和保护切换; — 其他传送网维护功能。
SONET/SDH 再生段	— 高速传输 (例如成帧、扰码); — 再生器故障分段; — 其他传送网维护功能。

图 A2 IP/PPP/HDLC/SDH 的协议堆栈和功能

PPP 协议由 IETF RFC 1661 规定, 具有 HDLC 帧格式的 PPP 帧由 IETF RFC 1662 文件规定, SDH 上传送 IP (IP Over SDH) 的技术要求由 IETF RFC 1619/RFC2615 文件规定, SDH 的技术要求在 ITU-T G.707 中规定, 如图 A.3 所示。

IP	
PPP	IETF RFC 1619/RFC2615
HDLC	
SONET/SDH 路径	
SONET/SDH 复用段	ITU-T G.707
SONET/SDH 再生段	

图 A3 IP/PPP/HDLC/SDH 的协议

A.3 SDH上传送IP的技术要求

A.3.1 PPP协议族

SDH 上传送 IP (IP Over SDH) 实际上是把 IP 数据包封装在 PPP 协议中, 然后在把 PPP 帧放入 SDH 的净荷字段中。本节介绍 PPP 协议的基本概念。

点到点协议 (PPP) 实际上是一个协议族, 它包括下列协议:

- * PPP 协议点到点协议;
- * PPP 多链路协议点到点多链路协议;
- * LCP—链路控制协议;
- * LQR—链路质量报告;
- * PAP—通行字认证协议;
- * CHAP—握手认证协议;
- * IPCP—IP 控制协议;
- * IPXCP—IPX 控制协议;
- * ATCP—AppleTalk 控制协议;
- * BAP—带宽分配协议;
- * BACP—带宽分配控制协议;
- * BCP—桥接控制协议;
- * PPP-BPDU—PPP 桥接协议数据单元;
- * CCP—压缩控制协议;
- * IPv6CP—IPv6 控制协议;
- * SNACP—SNA PPP 控制协议;
- * BVCP—PPP Banyan Vines 控制协议;
- * NBFCP—PPP NetBios 帧控制协议;
- * DNCP—PPP DECnet Phase IV 控制协议;
- * L2F—第 2 层发送协议;
- * L2TP—第 2 层隧道协议;
- * ECP—保密控制协议;
- * OSINLCP—OSI 网络层控制协议;
- * PPTP—点到点隧道协议;
- * SDCP—串行数据控制协议。

图 A.4 表示了 PPP 协议族在 OSI 七层模型中的位置。

PPP 协议以前主要用于专线 (租用线路) 上, 它提供了通过点到点链路传输各种协议的数据分组的方法。最初 PPP 只用于运行高级数据链路控制 (HDLC) 协议的同步串行链路, 如图 A.4 所示, 现在, PPP 已经扩展为也能在异步连接上运行。

PPP 设计提供下列功能:

- * 用于建立、配置和监视串行连接的链路控制协议 (LCP);
- * 封装格式;

* 允许若干个网络层协议通过同一个串行路径复用一族网络层控制协议（NCP）。

PPP 协议定义了一个串行链路、点到点封装协议，封装协议在其信息(I)字段承载或封装网络层 PDU，并使用帧中的另一个字段来标识哪个网络层 PDU 位于信息字段。

PPP 封装串行通信链路上的网络层数据。该协议允许在点到点通信通路上进行通信的两台主机或路由器协商通信期间需要使用的特定网络层协议类型（如 IP 协议等）。协商完成后，在 HDLC 帧的信息字段封装承载网络层 PDU。PPP 协议既支持面向比特的同步传输，也支持异步（启动/停止）传输，它还可以用于交换链路或拨号链路，但需要全双工功能。

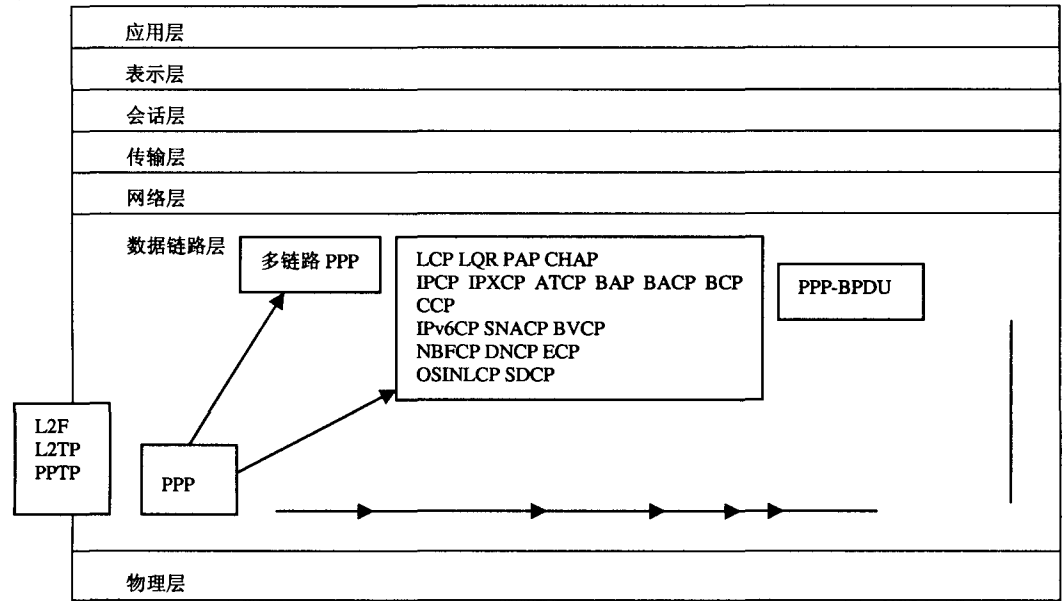


图 A.4 PPP 协议族与 OSI 模型的关系

PPP 分为 3 个主要部分，第一部分处理 HDLC 帧以及如何用它来把数据分组封装到帧的信息（I）字段。PPP 的第 2 个主要部分是链路控制协议（LCP），它用于建立链路、测试链路的各种服务质量特性，对它进行配置以及释放该链路。第 3 个主要部分是一族网络控制协议（NCP），以便确定哪个网络层协议要用于连接。

PPP 支持同时使用多个网络协议，例如该协议允许两个用户协商 IP、Digital 的 DECnet IV 网络层、IPX 和 XNS 的同时使用。

PPP 的连接建立包括 3 个阶段。

- (1) 交换 LCP 分组，建立链接并保证它的质量。链路质量报告（LQR）协议用于这个目的。
- (2) 如果建立了链接，则发送为每个配置协议准备的 NCP 分组，以建立这个协议的逻辑连接。例如，如果路由器运行 IP 和 IPX，为每个协议发送不同的 NCP 分组。如果 NCP 协商成功，则把该路由插入到该协议的路由表中。
- (3) 使用 PPP 封装通过连接发送协议的数据分组。一个 LAN 帧映射到一个 PPP 分组中，除非 LAN 的最大传输单元（MTU）超过串行线路的 MTU。不同协议的 LAN 分组不应被混合到同一个 PPP 帧中。

PPP 帧中的协议字段以及 NCP 协议，意味着能够使用 PPP 通过相同的物理链路多路复用不同的网络层协议。通过 PPP，甚至能发送桥接通信。

A.3.1.1 PPP 协议

点到点协议（PPP）用于在两个对等实体间传送分组的简单链路上，这些链路同时提供双方向的全双

工操作，并按先后次序传送分组。PPP 为各种主机、桥接设备和路由器之间提供了一种简便的连接方式。

PPP 协议信头的格式如图 A.5 所示。

标志	地址	控制	协议	信息	FCS	标志
1 字节	1 字节	1 字节	2 字节	可变长度	2 字节	1 字节

图 A.5 PPP 协议的信头结构

(1) 标志字段

标志字段是标准的 HDLC 标识，其值为 01111110。在实际通信中，前一个帧的结束和下一个帧的开始使用该标识来进行区分。由于在数据信息字段中也可能出现“01111110”，这时需要使用 HDLC “0”插入技术来区分数据与标志字段，即发送端在信息字段中连续 5 个 1 后插入一位 0。接收端则对输入比特流进行监视，当检测到连续 5 个 1 时，如果第 6 位是 0，则删除该 0，如果第 6 位是 1，则把该字符解释为一个标志。

(2) 地址字段

地址字段长度为 1byte，PPP 信头中的地址是 HDLC 广播地址，其值为 FFh，PPP 不分配单个端站地址。

(3) 控制字段

控制字段长度为 1byte，其值为 03h，表示该帧是轮询/结束（Poll/Final）置为 0 的 HDLC 无编号信息（UI）帧命令。该字段编码为其他值的帧将被丢弃。

(4) 协议字段

协议字段长度为 2byte，用来标识在 PPP 帧的信息字段封装的网络协议类型。协议字段还用来标识 PPP 控制协议的类型，即 LCP 或 NCP（如 IPC 等），其中“Cxxxh”范围的值保留给 LCP 使用，而“8xxxh”范围的值保留给 NCP 使用。

PPP 端站能够协商把协议字段的大小减少到 1byte。

(5) 信息字段

信息字段是可变长度的，在该字段放入高层协议数据。

(6) FCS 字段

帧校验序列（FCS）字段长度为 1byte，其值为 PPP 帧的 16 位循环冗余校验和。

A.3.1.2 链路控制协议

为了使 PPP 协议能适用于各种环境，PPP 提供链路控制协议（LCP），用于建立、配置和测试数据链路连接。使用 LCP 可以自动协商封装格式选择、处理各种分组长度限制、检测环回链路和其他配置差错、结束链路等。LCP 还包括其他一些可选设施，如对链路对端的对等端站身份的认证以及确定链路功能是否正常、链路是否出现故障等。

LCP 分组的格式如图 A.6 所示，LCP 分组各字段的意义描述如下。

编码	标识符	长度	数据
1byte	1byte	2byte	可变长度

图 A.6 LCP 分组的格式

(1) 编码字段

编码字段长度为 1byte，编码值指出 LCP 分组的类型。

编码 LCP 分组类型

- 1 配置请求
- 2 配置确认
- 3 配置未确认
- 4 配置拒绝
- 5 结束请求
- 6 结束确认
- 7 编码拒绝
- 8 协议拒绝
- 9 回显请求
- 10 回显响应
- 11 丢弃请求
- 12 链路质量报告

(2) 标识符字段

标识符字段长度为 1byte，用于匹配请求和响应。

(3) 长度字段

长度字段为 2byte 长，用来指示 LCP 分组的长度，包括编码、标识符、长度和数据字段。

(4) 数据字段

数据字段长度可变，它包含一个或多个配置选项域。LCP 配置选项域的格式如图 A.7 所示。

类型	长度	数据
----	----	----

图 A.7 LCP 配置选项域的格式

● 类型字段

类型字段长度为 1byte，其编码值用来指示配置选项域的类型。

● 长度字段

长度字段指出配置选项域的长度，包括类型、长度和数据字段。

● 数据字段

数据字段的值。

LCP 支持连接的建立，并允许协商某些配置选项域。协议还维护，并提供终止过程。为完成这些功能，LCP 要按下述 4 个阶段组织：

- 阶段 1：链路建立和配置协商；
- 阶段 2：链路质量确定；
- 阶段 3：网络层协议配置协商；
- 阶段 4：链路撤消。

PPP 要求执行 LCP 以便在交换任何网络层业务流之前，打开两个端站之间的连接，它由一系列配置分组的交换来完成。在交换了这些分组并且配置确认分组已在站点之间发送和接收之后，则认为连接已处于打开状态，可以开始网络层分组的交换，但 LCP 本声只限于链路层操作。LCP 并不理解如何协商网络层协议的实现。实际上，它并不关心和网络协议有关的上层协商。

A.3.1.3 IP控制协议

IP 控制协议 (IPCP) 负责在点到点链路两端的端站中配置 IP 协议参数。IPCP 使用与链路控制协议 (LCP) 相同的分组交换机制, 只有当 PPP 进入到网络层协议阶段后, 才能交换 IPCP 分组, 任何在此之前接收到的 IPCP 分组都将被丢弃。

IPCP 分组的格式如图 A.8 所示。

编码	标识符	长度	数据
1byte	1byte	2byte	可变长度

图 A.8 IPCP 分组的结构

各字段意义描述如下:

(1) 编码字段

编码字段长度为 1byte, 其编码指出 IPCP 分组的类型:

编码 IPCP 分组类型

- | | |
|---|-------|
| 1 | 配置请求 |
| 2 | 配置确认 |
| 3 | 配置未确认 |
| 4 | 配置拒绝 |
| 5 | 结束请求 |
| 6 | 结束确认 |
| 7 | 编码拒绝 |

(2) 标识符字段

标识符字段长度为 1byte, 其编码用来匹配请求和响应。

(3) 长度字段

长度字段为 2byte, 用来指出 IPCP 分组的长度, 包括编码、标识符、长度和数据字段。

(4) 数据字段

数据字段长度可变, 它包含一个或多个配置选项域, IPCP 配置选项域的格式如图 A.9 所示。

类型	长度	数据
----	----	----

图 A9 IPCP 配置选项域的格式

● 类型字段

类型字段长度为 1byte, 其编码值用来指示配置选项域的类型:

类型字段编码 配置选项域类型

- | | |
|---|---------|
| 1 | IP 地址 |
| 2 | IP 压缩协议 |
| 3 | IP 地址 |

当类型字段编码为 3 时, 表示在链路的本地端协商 IP 地址的方式。优选使用类型编码 3, 不建议使用类型编码 1。

● 长度字段

长度字段指出配置选项域的长度, 包括类型、长度和数据字段。

● 数据字段

数据字段的值。

A.3.1.4 链路质量报告协议

链路质量报告（LQR）协议规定了 PPP 中链路质量监视的机制。在核心路由器中，无需链路质量报告（LQR）协议。

A.3.2 对SDH接口的要求

应符合 YDN 099-1998 和 ITU-T G.707 的要求。

附录 B (规范性附录)

在 ATM 上支持传统 IP 及地址解析 (ARP) 的技术要求

B.1 概述

本附录规定了在 ATM AAL5 上传输 IP 数据包、ATM 地址解析协议 (ATMARP) 请求和应答的方法和协议。

B.2 总则

本附录描述了在“传统的”IP 网络中使用 ATM 来代替局域网 (ethernet) 和代替互联管理区域内或管理区域之间的路由器的 IP 链路的方法。这里“传统的”模式是指把 ATM 主机适配器当作运行在 LAN 情况下的 IP 协议栈网络接口来处理。

传统模型的特性包括:

- 1) 一个 LIS 内所有的 VC 使用相同的最大传输单元 (MTU), 见本附录 B5 节的规定;
- 2) 在缺省的情况下, 使用 LLC/SNAP 对 IP 分组进行封装;
- 3) 端到端 IP 路由结构与以前相同;

4) 在 LIS 内使用 ATMARP 从 IP 地址中解析出 ATM 地址, 在客户机看来, ATMARP 结构与 IETF RFC 826 定义的 ARP 模型完全相同;

- 5) 一个 IP 子网由许多主机和路由器组成, 同一个 LIS 中的两个 IP 成员由一条 VC 直接相连接。

在传统模型中, ATM 地址可以使用 ATM Forum NSAP 端点地址或 E.164 UNI 地址来, 在 ATMARP 协议中, 把 ATM 地址当作“硬件地址”。

B.3 IP子网配置

在逻辑 IP 子网 (LIS) 方案中, 每个独立的管理实体在一个封闭的逻辑 IP 子网内配置它的主机和路由器。每个 LIS 的运行和通信独立于同一个 ATM 网络中的其他 LIS。与 ATM 网络相连接的主机直接与同一 LIS 内的其他主机进行通信。与本地 LIS 以外主机的通信通过 IP 路由器进行。该路由器是一个与 ATM 网络相连接的 ATM 端点, 并被配置成一个或多个 LIS 的成员。该配置导致在同一个 ATM 网络上运行多个分离的 LIS。属于不同 IP 子网的主机之间应通过一个中间 IP 路由器进行通信, 即使在 ATM 网络上, 这两个 IP 成员间可以建立一条直接的 VC 连接。

在一个 ATM LIS 配置中运行的 IP 成员 (主机, 路由器) 的要求如下:

- 1) 所有成员具有同样的 IP 网络/子网号码和地址模;
- 2) 一个 LIS 内的所有成员都直接连接到 ATM 网络;
- 3) LIS 以外的所有成员通过路由器进行访问;

4) 当使用 SVC 的时候, 一个 LIS 内的所有成员应能完成 ATMARP 功能, 从目的 IP 地址中解析出目的 ATM 地址。一个 LIS 内的所有成员还应能完成 InATMARP 功能, 在 ARP 服务员中进行地址登记, 见第 B.6 章的规定;

5) 当使用 PVC 的时候, 一个 LIS 的所有成员应能完成 InATMARP, 从与之相连接的 VC 中解析出所有与该成员相连接的所有其他成员的 IP 地址, 见第 B.6 章的规定;

6) LIS 内的所有成员应能够通过 ATM 与该 LIS 内的其他所有成员进行通信, 即连接该子网成员的底层虚连接拓扑结构是全网状的。

下面列出了每个与 ATM 网络相连接的 IP 端站中应实施的一组特定的 ATM 参数:

1) ATM 硬件地址 (atm\$ha): 每个 IP 端站的 ATM 地址。

2) ATMARP 请求地址 (atm\$arp-req): atm\$arp-req 是每个 LIS 内的 ATMARP 服务器的 ATM 地址。在 SVC 环境下, 向该地址发送 ATMARP 请求, 用来从目的协议地址中解析出的 ATM 地址。该服务器应有权负责解析该 LIS 内所有成员的 ATMARP 请求。如果 LIS 仅进行 PVC 操作, 那么该参数被置为空并且 IP 端站不需要送 ATMARP 请求给 ATMARP 服务器。

在 ATM 网络上提供 LIS 功能的路由器, 也可以互联多个 LIS, 能互联多个不同 LIS 的路由器应支持多组参数 (每个 LIS 一组), 并且能够把每组参数与一个特定的 IP 网络/子网号码关联起来。此外, 互联多个 LIS 的具有一条 ATM 物理接口的路由器可以具有一个或多个不同的 ATM 端点地址。当使用 NSAP 地址时, 这并不意味着使用不同的端系统标志符 (ESI), NSAP 地址的最后一个字节是 NSAP 地址选择器 (SEL) 字段, 该字段能够被用来区分使用同一个 ESI 的 256 个不同的 LIS。

B.4 封装格式

所有实施方法应支持 IEEE 802.2 LLC/SNAP 封装, 如 RFC 2684 所述。LLC/SNAP 封装是用于 IP 数据包的缺省的分组格式。

B.5 MTU长度

在 ATM 网络上运行的 IP 成员的缺省 MTU 长度应该是 9180byte。LLC/SNAP 信头是 8byte, 因此缺省的 ATM AAL5 协议数据单元的长度是 9188byte。在传统的 IP 子网中, 如果且仅仅如果在 LIS 的所有成员已经被配置为使用非缺省值的情况下, 才能使用非缺省值。

B.6 地址解析

在 ATM 逻辑子网中的地址解析应该使用 ATM 地址解析协议 (ATMARP) 和反向 ATM 地址解析协议 (InATMARP)。ATMARP 以 IETF RFC 826 中定义的 ARP 协议为基础, 但扩展了必要的功能, 用以在单传送服务器 ATM 环境下支持 ARP。InATMARP 与 IETF RFC 1293 中定义的协议完全相同。

ARP 和 InARP 分组的格式在第 B.6.6 节描述。

B.6.1 永久虚连接环境下的地址登记

在使用 ATM PVC 连接的 LIS 中, 利用网管来配置和建立 LIS 内相关 IP 成员间的 PVC 连接。LIS 内的所有 IP 端站都应使用反向 ATM 地址解析协议 (InATMARP), 用以确定与该端站相连接的所有 IP 端站的 IP 地址。发送 IP 端站应在与其相连接的所有 PVC 上发送含有其 IP 地址的 InARP_REQUEST 消息, 所有收到该请求消息的 IP 端站都应用含有其本身 IP 地址的 InARP_REPLY 消息进行响应。图 B1 表示了 PVC 上 InATMARP 操作的举例。

InATMARP 分组的格式在第 B.6.6 节规定, 当 ATM 源和/或目的地址未知时, 在 InATMARP 分组中相应的 ATM 地址长度字段应被置为零, 指示 ATM 源和/或目的地址字段为空。

当请求端站收到 InARP 应答后, 它将使用在应答中所提供的地址信息填写其高速缓存区中的 ATMARP 表的相应条目。与 ATMARP 一样, 通过 InATMARP 获得的信息可能由于超时而被丢弃或在某些环境下变成无效。作为定时处理的一部分, 每个 IP 端站应负责使 ATMARP 表中的相应条目重新有效,

参见第 B.6.5 节的规定。

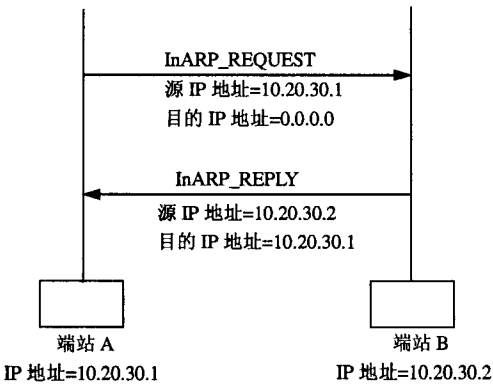


图 B.1 PVC 上 InATMARP 操作举例

B.6.2 交换虚连接（SVC）环境下的地址登记

ATM 网络是非广播、非多目广播的通信网络，在这种网络中利用 SVC 来支持 IP 协议需要使用 ATMARP 和 InATMARP 技术，在本标准中规定，在一个 LIS 内使用 ATMARP 服务器来负责完成 ATMARP 和 InATMARP 功能，ATMARP 服务器应能解析该 LIS 内的所有成员的 ATMARP 请求并能使用 InATMARP 来完成该 LIS 内所有成员的地址登记。

ATMARP 服务器具有一个 IP 地址，ATMARP 服务器还具有一个公共的 ATM 地址，即在第 B.3 章规定的 atm\$arp-req，该服务器服务的 LIS 内的所有客户机都配置有 ATMARP 服务器的 ATM 地址。ATMARP 服务器自己不能发起建立一条 VC 连接来进行地址登记，而是根据 LIS 内的客户机的动作来开始 ATMARP 登记程序。LIS 内的客户机根据其自身配置的 ATMARP 服务器的 ATM 地址，建立与 ATMARP 服务器相连接的点到点 VC。一旦这条使用 LLC/SNAP 封装的 VC 呼叫/建立完成后，ATMARP 服务器将发送一个 InARP_REQUEST，以确定客户机的 IP 地址。客户机使用 InARP_REPLY 来进行应答，该应答中包含了足够的信息，ATMARP 服务器可以据此建立/更新其 ATMARP 表中客户机的 IP-ATM 地址对条目。InATMARP 分组的格式见第 B.6.6 节，地址登记过程如图 B.2 所示。

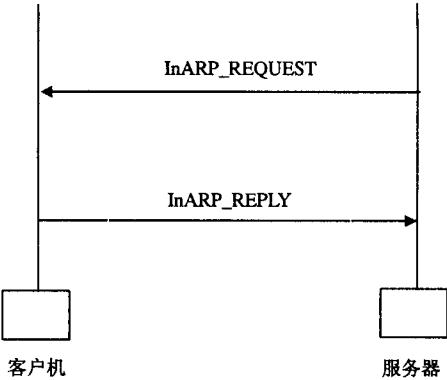


图 B.2 SVC InATMARP 地址登记过程

在一个 LIS 中可以只有一个 ATMARP 服务器，但为了安全和可靠起见，一个 LIS 内也可以有几个 ATMARP 服务器，但这些 ATMARP 服务器的数据库应同步，并具有与单个服务器一样的客户机—服务器接口。

B.6.3 ATMARP服务器工作要求

ATMARP 服务器接受发自它所服务的 LIS 内其他 ATM 端点的 ATM 呼叫/连接。当呼叫建立后，如

果该 VC 支持 LLC/SNAP 封装, 那么 ATMARP 服务器将向源端站发送一个 InATMARF 请求 (InARP_REQUEST)。在收到 InATMARF 应答 (InARP_REPLY) 后, 服务器将检查应答中的 IP 地址和 ATM 地址。服务器将增加或更新其 ATMARP 地址表中的<ATM 地址, IP 地址>映射条目和时间标签。如果 InATMARF 应答中的 IP 地址与服务器中 ATMARP 地址表某个条目中的 IP 地址相同, InATMARF 应答中的 ATM 地址与 ATMARP 地址表中对应条目的 ATM 地址不匹配并且存在一条开放的 VC 与 ATMARP 地址表中的该条目相关联, 则丢弃 InATMARF 信息并且不对服务器的 ATMARP 地址表进行修改, ATMARP 地址表条目保存到超时或无效才进行修改。VC 呼叫的清除不会删除 ATMARP 地址表的条目。

当 ATMARP 服务器接收到一个 ATMARP 请求 (ARP_REQUEST) 时, 如果在它的 ATMARP 地址表中存在对应的<ATM 地址, IP 地址>映射项目, 那么它将产生相应的 ATMARP 回答 (ARP_REPLY), 其中包含客户机所需解析的目的端站的 ATM 地址。如果在服务器的 ATMARP 地址表中不存在对应的<ATM 地址, IP 地址>映射项目, 则它将产生一个否定的 ATMARP 应答 (ARP_NAK)。ARP_NAK 应答是 ATMARP 协议专有的一个特征, 用来提高 ATMARP 服务器机制的可靠性。使用 ARP_NAK 应答, 客户机能够区分地址解析失败是由于服务器的严重故障还是 ATMARP 地址表查阅失败。ARP_NAK 分组格式与所接收到的 ARP_REQUEST 分组格式相同, 只是把操作码置为 ARP_NAK, 即产生 ARP_NAK 分组时, 把 ARP_REQUEST 分组数据复制下来, 而仅仅把 ARP_REQUEST 分组中的操作码字段置为 ARP_NAK, ARP_NAK 分组格式见第 B.6.6 节的描述。

当服务器在一条 VC 上收到一个 ATMARP 请求, 该请求中的源 IP 地址和 ATM 地址与服务器的 ATMARP 地址表中已存在的条目相匹配并且其中的 ATM 地址与该 VC 相关联, 服务器将复位源 ATMARP 地址表相应条目的时间标签: 即如果客户机在用来登记它的 ATMARP 地址表条目的相同 VC 上向服务器发送 ATMARP 请求, 服务器应检查该 ATMARP 请求, 并复位对应于客户机的 ATMARP 地址表条目的时间标签, 重新开始计时, 这样标记该客户机仍是“活动的”。

为了增加使用 ATMARP 的地址解析机制的可靠性: 当服务器在一条 VC 上收到一个 ARP_REQUEST 时, 它应检查源信息。如果 ATMARP 地址表中没有与该 VC 相关联的 IP 地址并且如果源 IP 地址与任何其他 VC 连接都不关联, 则服务器应在 ATMARP 地址表中增加<ATM 地址, IP 地址>映射条目和时间标记并且使这个条目与该 VC 相关联。

B.6.4 ATMARP 客户机工作要求

ATMARF 客户机负责向 ATMARP 服务器登记它自己的 ATMARP 信息并负责获取和刷新它自己的有关其他 IP 成员的 ATMARP 地址条目/信息。这意味着, ATMARP 客户机应配置有 ATMARP 服务器的 ATM 地址。

ATMARF 客户机应完成下列功能。

- 1) 建立与 ATMARP 服务器相连接的 VC, 用以发送和接收 ATMARP 和 InATMARF 分组。
 - 2) 正确地响应在任何 VC 上接收到的 ARP_REQUEST 和 InARP_REQUEST 分组 (见 IETF RFC 1293)。
 - 3) 产生 ARP_REQUEST 分组, 向 ATMARP 服务器发送 ARP_REQUEST 分组, 并且正确地处理从服务器接收到的 ARP_REPLY 和 ARP_NAK 分组。ARP_REPLY 分组用来建立/刷新客户机自己的 ATMARP 地址表条目。
 - 4) 在需要时, 产生和发送 InARP_REQUEST 分组并且正确地处理 InARP_REPLY 分组。
- InARP_REPLY 分组用来建立/刷新客户机自己的 ATMARP 地址表条目 (见 IETF RFC1293)。

5) 提供 ATMARP 地址表计时功能, 在一个时间周期后, 删除客户机自己的旧的 ATMARP 地址表条目。

如果客户机没有保持一条与服务器相连接的 VC, 那么客户机应在服务器中刷新它自己的 ATMARP 地址表信息, 并且至少 20s 一次。这通过建立一条与服务器相连接的 VC 并且交换 InATMARF 分组来进行。

B.6.5 ATMARP 地址表计时

ATMARF 客户机或服务器应知道下列信息: 与它相连接的所有 VC (永久的或交换的)、它们在 ATMARP 地址表中对应的条目、支持 LLC/SNAP 封装的 VC。

服务器 ATMARP 地址表中的条目有效的最长时间是 20min, 即地址表中对应的条目将无效, 在与服务器相连接的 VC 每 20min 应发送一次 InATMARF 请求, 使 ATMARP 地址表中的对应条目保持有效, 如果某条与服务器相连接的 VC 被拆除 20min 后, 在服务器的 ATMARP 客户机 ATMARP 地址表中的条目有效的最长时间是 15min。

在计时期满要删除一个 ATMARP 表条目前, ATMARP 服务器应在任何与该条目相关联的已建立 VC 上发送一个 InARP_REQUEST。如果收到一个 InARP_REPLY, 该表项目被更新并且不被删除。

如果不存在与 ATMARP 地址表中的某个条目相关联的已建立的 VC, 则该条目将被删除。

当 ATMARP 地址表条目超时, ATMARP 客户机必使该条目无效。如果没有与该无效条目相关联的已建立的 VC, 那么该条目将被删除。在一个无效条目对应一条已建立 VC 的情况下, ATMARP 客户机应在在该 VC 上发送任何非地址解析业务之前使该条目重新有效。在 PVC 的情况下, 客户机将发送一个 InARP_REQUEST, 当接收到 InARP_REPLY 时更新该条目, 使该条目重新有效。在 SVC 的情况下, 客户机将向 ATMARP 服务器一个发送 ARP_REQUEST, 当接收到 InARP_REPLY 时更新该条目, 使该条目重新有效。如果与一个无效的 ATMARP 地址表条目相关联的 VC 被拆除了, 则将删除该条目。

B.6.6 ATMARP和InATMARF分组格式

互联网地址的分配独立于 ATM 地址的分配。每个主机实施方法应知道它自己的 IP 和 ATM 地址并且应能正确地响应地址解析请求。IP 成员应也在需要使用 ATMARP 和 InATMARF 把 IP 地址解析到 ATM 地址。

ATMARF 和 InATMARF 分组格式如图 B.3 所示, ATMARP 和 InATMARF 协议使用与 IETF RFC 826 规定的 ARP 和 IETF RFC 1293 规定的 InARP 协议相同的硬件类型 (ar\$hrd)、协议类型 (ar\$pro) 和操作码 (ar\$op) 等字段格式。这些字段在 ATMARP 分组中的位置与在 ARP 和 InARP 分组中的比特位置相同。ATMARF 被分配给了一个惟一的硬件类型值, 此外, ATMARP 使用一个用于 ARP_NAK 的操作码。ATMARF/InATMARF 分组格式的剩余部分与 ARP/InARP 分组格式不同。

ATMARF 和 InATMARF 协议有帧格式和值如图 B.3 所示的几个字段。

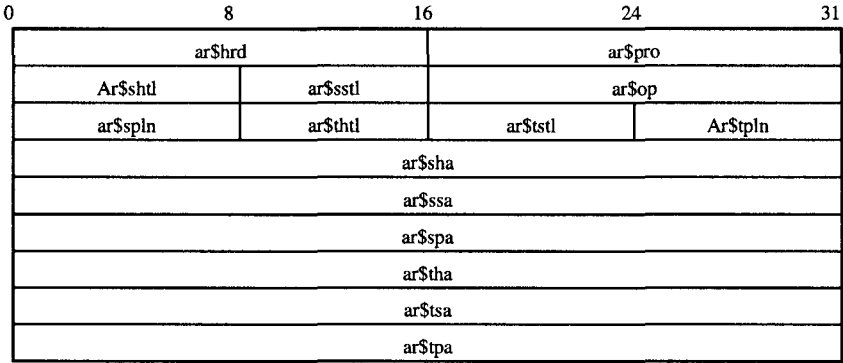


图 B.3 ATMARP 和 InATMARF 分组格式

图 B.3 中:

ar\$hrd	16 比特	硬件类型
ar\$pro	16 比特	协议类型
ar\$shtl	8 比特	源 ATM 号码 (q) 的类型和长度
ar\$sstl	8 比特	源 ATM 子地址 (r) 的类型和长度
ar\$op	16 比特	操作码 (请求, 响应或 NAK)
ar\$spln	8 比特	源协议地址 (s) 的长度
ar\$thtl	8 比特	目的 ATM 地址 (x) 的类型和长度
ar\$stsl	8 比特	目的 ATM 子地址 (y) 的类型和长度
ar\$tpln	8 比特	目的协议地址 (z) 的长度
ar\$sha	q 字节	源 ATM 号码
ar\$ssa	r 字节	源 ATM 子地址
ar\$spa	s 字节	源协议地址
ar\$tha	x 字节	目的 ATM 号码
ar\$tsa	y 字节	目的 ATM 子地址
ar\$tpa	z 字节	目的协议地址

其中

ar\$hrd — 分配给该字段的编码为 (0x0013) (参见 RFC 1430)。

ar\$pro — 使用 ATMARP 的协议类型号码 (IP 协议是 0x0800)。

ar\$op — 操作类型值 (十进制):

ARP_REQUEST = 1

ARP_REPLY = 2

InARP_REQUEST = 8

InARP_REPLY = 9

ARP_NAK = 10

ar\$spln — 源协议地址的长度 (以字节为单位, 下同), 对于 IP 协议, ar\$spln 为 4。

ar\$tpln — 目的协议地址的长度, 对于 IP 协议, ar\$tpln 为 4。

ar\$sha — 源 ATM 号码 (E.164 或 NSAP)

ar\$ssa — 源 ATM 子地址 (NSAP)

ar\$spa — 源协议地址

ar\$tha — 目的 ATM 号码 (E.164 或 NSAP)

ar\$tsa — 目的 ATM 子地址 (ATM NSAP)

ar\$tpa — 目的协议地址

ar\$shtl、ar\$sstl、ar\$thtl 和 ar\$stsl 的编码如图 B.4 所示。

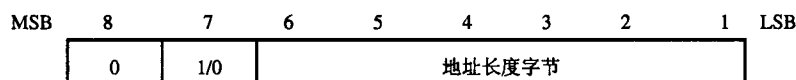


图 B.4 ar\$shtl、ar\$sstl、ar\$thtl 和 ar\$stsl 的编码

其中：

- 比特 8（保留） = 0（留待将来使用）
- 比特 7（类型） = 0 NSAP 格式
 = 1 E.164 格式
- 比特 6-1（长度） = 地址长度
 （MSB = 比特 6，LSB = 比特 1）

在 ITU-T Q.2931 和 ATM Forum UNI 3.1/4.0 规范中定义的 ATM 地址包括“主叫号码信息单元”和“主叫子地址信息单元”。这些信息单元(IE)应该分别映射到 ATMARP/InATMARF 的源 ATM 号码和源 ATM 子地址字段。在 ITU-T Q.2931 和 ATM Forum UNI 3.1/4.0 规范中定义的“被叫号码信息单元”和“被叫号码信息单元”应分别映射到 ATMARP/InATMARF 的目的 ATM 号码和目的 ATM 子地址字段。

目前使用 3 种号码和子地址的组合结构，如表 B.1 所示。

表 B.1 目前的使用 3 种号码和子地址的组合结构

	号码	子地址
结构 1	NSAP	空
结构 2	E.164	空
结构 3	E.164	NSAP

IP 成员应使用与它们的 ATM 网络连接对应的 ATM 地址结构，向它们的 ATMARP 服务器登记它们的 ATM 端点地址，例如遵循 ATM Forum UNI 3.1/4.0 的 ATM LAN 上实施的 LIS 可以使用结构 1 进行登记；在 E.164 “公共” ATM 网络上实施的 LIS 可以使用结构 2 进行登记。在 ATM LAN 和公共 ATM 网络组合环境中实施的 LIS 可以使用结构 3 注册。所有的实施应支持所有的 3 种 ATM 地址结构。

用于结构 1 和 2 的 ATMARP 和 InATMARF 请求和响应应指示一个空的 ATM 子地址；即 ar\$sttl.type = 1 且 ar\$sttl.length = 0、ar\$sttl.type = 1 且 ar\$sttl.length = 0。当 ar\$sttl.length 和 ar\$sttl.length = 0 时，没有 ar\$tsa 和 and ar\$ssa 字段。

B.6.7 ATMARP/InATMARF分组的封装

ATMARF 和 InATMARF 分组使用 LLC/SNAP 格式封装编码到 AAL5 PDU，用于封装 ATMARP/InATMARF PDU 的 AAL5 CPCS-SDU 净荷字段的帧格式如图 B.5 所示。

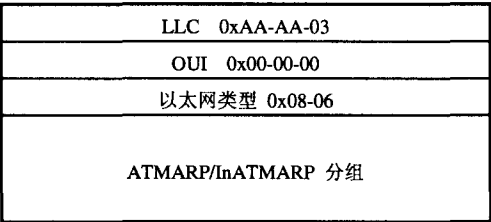


图 B.5 用于封装 ATMARP/InATMARF PDU 的 AAL5 净荷的格式

LLC 值 0xAA-AA-03（3 字节）指示了 SNAP 报头的存在。

OUI 值 0x00-00-00（3 字节）指示了下面两个字节为以太网类型。

以太网类型值 0x08-06（2 字节）指示是 ARP 类型，见 IETF RFC 1340。

LLC/SNAP 报头的总长度固定为 8 个八位组。这样就使 ATMARP 分组在 64bit 边界上与相关的 AAL5 CPCS-SDU 开端对齐。

这里所提出的对 ATMARP/InATMARF 进行 LLC/SNAP 封装与 IETF RFC 2684 中所说明的 ATM

AAL5 上 IP 多协议封装处理是相一致的，在格式上与在 IETF RFC 1042 中说明的 IEEE 802 网络上的 ATMARP 也是一致的。

在 LIS 中地址解析请求是向 LIS 内所有直接相连的 IP 成员进行广播。

B.7 IP广播地址

ATM 不支持广播寻址，因此不存在 IP 广播地址至 ATM 广播设备的有效映射。这种缺乏的映射并不限制网络成员发送或接收指定了 IETF RFC 1122 中所述的任意 4 种标准 IP 广播地址格式的 IP 数据包。网络成员一接收到发往其 LIS 的 IP 广播 IP 子网广播，应处理这些分组。

B.8 IP组播地址

ATM 不支持组播地址设备，因此不存在 IP 组播地址至 ATM 组播设备的有效映射。

附录 C

(资料性附录)

区分业务

C.1 概述

区分业务 (Diff-serv) 的基本机制是在网络的边缘路由器上根据某一业务的服务质量要求将该业务映射到一定的业务类别之中, 随后利用 IP 分组中的 DS 字段惟一地标志这一业务所需的服务类别, 网络中的各个节点将依据该字段对各种业务类别采取预先设定好的服务策略, 保证相应的延迟、传送速率、抖动等服务质量参数。这样, 对于一次会话中特定的数据流, 在每次连接的过程中, 将无须传递各种 QoS 信息, 从而避免了 RSVP 中高昂的建立成本。同时, 也使得这种技术具有较好的反应灵敏度, 特别适合于互联网中大量存在的短时间的连接。

Diff-serv 的基本原理见图 C.1 所示。

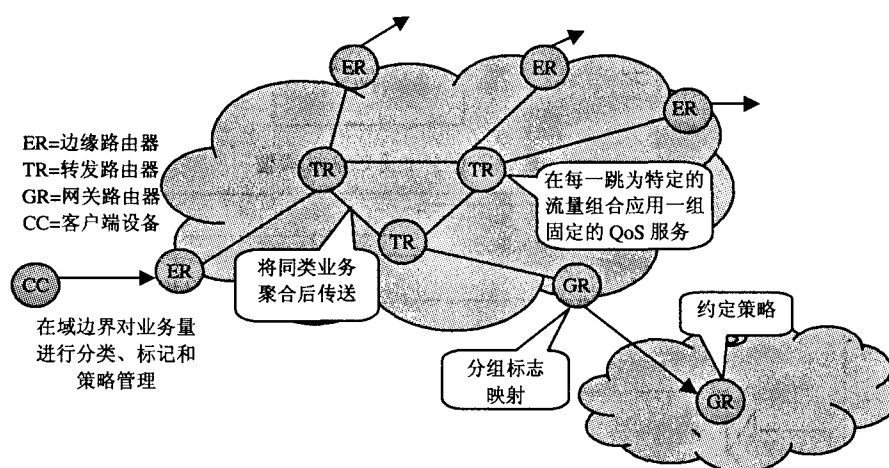


图 C.1 区分业务的原理

C.2 区分业务的结构

在区分业务域中, 路由器大致分为两类: 边缘路由器和转发路由器。

区分业务模型功能模型如图 C.2 所示, 区分业务的具体工作流程如下所述。

1. 在区分业务域的边缘路由器中将包含一系列的功能单元。边缘路由器将对来自于用户或其他网络的非区分业务的业务流进行分类, 并为每一个 IP 分组填入新的 DSCP 字段, 同时, 在网络的边缘路由器中建立起并开始应用与每一业务相对应的服务等级约定 (SLA) 以及每一跳行为 (PHB); 而对于来自用户或其他网络的区分业务业务流, 则依据分组中的 DSCP 字段, 为相应的业务选择特定的 PHB。

2. 在整个 DS 域的范围对相应的业务进行资源的分配, 使 DS 区域中对某一业务的服务质量达到一致。这一过程称为业务提供。

3. 随后, 将开始对业务分组的转发。

4. 在转发的过程中, 边缘路由器的各个策略单元将根据网络之间或网络与用户之间的 SLA, 对收到的业务流进行测量, 监视用户是否遵守业务等级协定 (SLA), 并将测量结果输入业务流策略单元, 对业务流进行整形, 丢弃, 标记 (实际上就是进行编码点的改写) 等工作。这一过程称为业务量调整 (Traffic Conditioning) 或业务量策略 (Traffic Policing)。

5. 此后，边缘路由器将对经过了上述步骤处理之后的业务流的分组进行 DSCP 字段的检查，并依据 DS 字段为业务流选择特定的 PHB，根据 PHB 所指定的排队策略将属于不同业务类别的业务量导入不同的队列加以处理，并按事先设定的带宽、缓冲处理输出队列，最后按照 PHB 所指定的丢弃策略对分组实施必要的丢弃。

6. 中间路由器在一般情况下，将只关心分组的 DS 字段，依据 DS 字段为业务流选择特定的 PHB，根据 PHB 所指定的排队策略将属于不同业务类别的业务量导入不同的队列加以处理，并按事先设定的带宽、缓冲处理输出队列，最后按照 PHB 所指定的丢弃策略对分组实施必要的丢弃。

可见，边缘路由器的功能是最为复杂的，它包含了能够实现整个区分业务模型中所有功能的各种功能单元。业务量调整单元见图 C.3 所示。

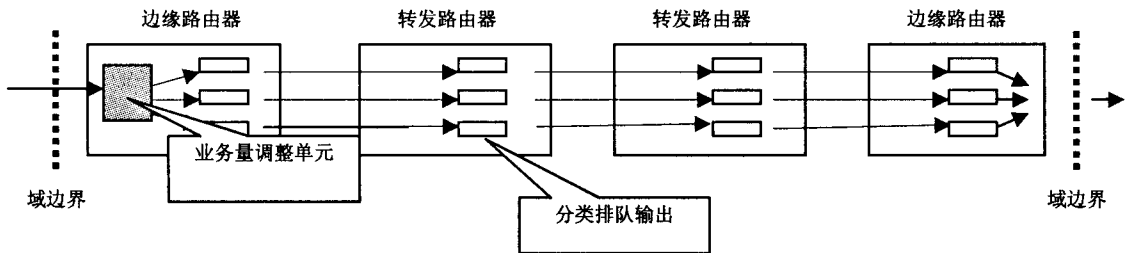


图 C.2 区分业务的功能模型

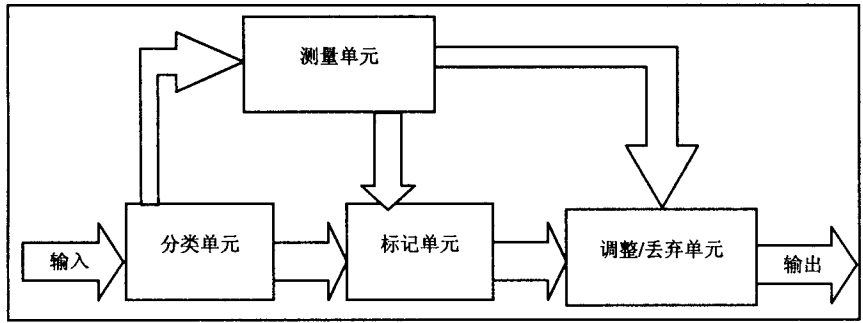


图 C.3 业务量调整单元的结构

分类部分包括分类与标记两个单元。

分类单元：在边缘路由器上，这一单元将检查分组的源 IP 地址、目的 IP 地址以及 TCP 或 UDP 的源/目的端口号等信息，籍此对分组所属的数据流进行辨别，这一过程实际上就是一个分类过程。对于 DS 字段的辨别也是由这一单元完成的。

分类单元将经过分类的业务流送至测量单元以便加以业务认证。同时，还将把业务流送至标记单元，以便对分组的标记进行必要的处理。

标记单元：在分类单元对业务流进行分类的基础上，标记单元将对没有填写 DS 字段的 IP 分组进行 DS 字段的填写。另外，如果来自测量单元的信息表明某一业务流超过了 SLA，则将由标记单元对业务流中的 IP 分组的 DS 域进行必要的改写，以便对相应的业务流进行服务质量的降级处理。

策略部分由测量单元与整形/丢弃单元组成。

测量单元：对某一业务流进行分类之后，还应对其速率，突发长度等参数进行测量，以便确定该业务流的资源消耗在持续时间内是否超出 SLA 的规定。除了业务认证功能之外，测量单元对于业务量的统计与计费功能也十分有用。在区分业务网络中，由于对享用不同服务质量的业务流的收费是不同的，所

以路由器中应有专门用于进行区分业务的业务测量的功能。

整形/丢弃单元：整形是指当业务流中存在突发的业务流时，通过一定的机制使路由器输出的业务流变得较为平稳。对于突发业务流，主要有三种整形策略：一是如果突发的业务流在一定的限度之内的话，则不予理会，继续正常转发；二是通过一定的机制（如漏桶算法）来对业务流的突发性进行削减；三是当业务流的突发超过一定的程度时，丢弃业务流中的一部分分组以便达到整形的目的。

当业务流的突发超过一定的程度或者业务流不符合 SLA 的约定时，边缘路由器将通过这一单元对业务流中的分组进行丢弃。具体的实施有许多不同的丢弃算法。

队列单元：在每一个路由器中，对应于每一个服务质量等级，将有一个队列单元。路由器将把属于不同服务质量等级的分组送入不同的队列单元中进行排队。排队单元是保证端到端服务质量的重要机构。

队列单元是实现 PHB 的关键部分。它的功能主要是两个方面：一方面是排队（scheduling）处理；另一方面是丢弃（Dropping）处理。所谓的 PHB 实际上就是一个对于不同的分组如何使用不同的队列（带宽不同，优先级不同）来进行处理，以及在拥塞发生时，如何对不同的分组采取不同的丢弃策略（丢弃机率不同）的问题，这两点也是决定分组获得服务质量好坏的关键。这两方面的功能将由一系列的排队（加权公平排队算法 WFQ，公平排队算法 RR（Round Robin））与丢弃算法（如随机早期探测 RED，加权的随机早期探测 WRED 等）来实现。

C.3 边缘与核心路由器

在集成业务（Int-serv）中，业务流所经过的每一个路由器都要按照特定的资源预留要求对所有的分组进行分类与策略。区分业务技术的一个重要目标就是将对业务流的分类与策略工作推到网络的边缘来完成，中间节点只需对 DS 字段进行检查并采取相应的交换策略即可；另一方面，区分业务技术对业务流的分类与策略功能将与网络单元的具体转发功能独立，这样做的目的是便于各种服务策略具体实现上的多样性，有利于网络功能的升级。可以看到，这一点与 MPLS 的设计思想是一致的。

在一个网络域中，也有可能有一些主机不支持区分业务，此时，与这些主机相邻的路由器将具有边缘路由器的功能，它们将对来自于这些主机的业务流进行分类与策略工作。

可以看到，核心路由器的工作实际上类似于 ATM 或 MPLS 的交换工作。这样，依据收到的分组的 DS 字段或标记，转发路由器就可以为不同的业务实现服务分类了。例如，如果有一个 DS 字段或者是标记表明相应的业务流是实时业务，则收到具有这一字段内容的路由器将把该业务流的分组优先于其他非延迟敏感型业务的分组转发。

另外，对于每一业务组合中的业务量也应有一定的限制，否则将无法达到理想的业务分类效果。道理很简单，如果大部分的业务都标志为高优先级的话，则高优先级将无从实现，最终在各个业务看来所获得的仍然是过去的尽力而为型服务质量。所以，边缘路由器应通过一定的业务认证机制对标有特定 DS 字段的分组进行策略，以便确认该分组能够享用相应的服务质量。

另外，对于分组的分类和标记工作也可以由用户端的主机来完成，因为这里是最了解各种业务的服务质量需求的地方。这时，Diff-serv网络的边缘实际上就移到了用户的主机之中。