

**YD**

# 中华人民共和国通信行业标准

YD/T 1099—2001

---

## 千兆比以太网交换机设备技术规范

Technical Specification for Giga Bit Ethernet LAN Switch Equipment

2001-02-20 发布

2001-06-01 实施

---

中华人民共和国信息产业部 发布

# 目 次

前言 .....	III
<b>1 范围</b> .....	1
<b>2 引用标准</b> .....	1
<b>3 定义术语及缩写</b> .....	2
3.1 定义 .....	2
3.2 缩写 .....	3
<b>4 千兆比以太网交换机功能</b> .....	4
<b>5 千兆以太网接口</b> .....	4
5.1 千兆以太网接口类型 .....	5
5.2 物理传输媒体 .....	6
5.3 PMD (物理介质相关) 子层 .....	6
5.4 PMA (物理介质接入) 子层 .....	6
5.5 PCS (物理编码) 子层 .....	6
5.6 RS(协调子层)和 GMII(千兆位介质无关接口).....	6
5.7 自动协商 .....	7
5.8 逻辑链路层功能 .....	7
<b>6 数据帧的转发及过滤</b> .....	7
6.1 转发数据帧 .....	7
6.2 过滤数据帧 .....	7
6.3 支持转发/过滤数据帧的功能 .....	8
<b>7 维护决定数据帧转发/过滤的信息</b> .....	10
7.1 维护过滤/转发信息 .....	10
7.2 网桥协议数据单元 .....	10
7.3 GARP 组播注册协议 (GMRP) .....	12
7.4 一般属性注册协议 (General Attribute Registration Protocol GARP) .....	12
7.5 生成树算法及协议 (Spanning Tree Algorithm and Protocol) .....	14
7.6 VLAN (虚拟局域网) 功能 .....	14
7.7 远程桥接 .....	16
7.8 多链路聚合 .....	16
<b>8 流量控制</b> .....	16
8.1 流量控制性能衡量 .....	16
8.2 半双工下的流控 .....	17
8.3 全双工下流控 .....	17
8.4 流量控制策略 .....	18
<b>9 性能指标要求</b> .....	18

9.1	端口数量 .....	18
9.2	设备吞吐量 .....	18
9.3	突发长度 .....	18
9.4	突发间隔 .....	18
9.5	过负荷 .....	18
9.6	转发速率 .....	18
9.7	拥塞控制 .....	18
9.8	队头阻塞 .....	18
9.9	地址缓存能力 .....	18
9.10	地址学习能力 .....	18
9.11	时延 .....	19
9.12	时延抖动 .....	19
9.13	丢包率 .....	19
9.14	误码率 .....	19
9.15	可靠性 .....	19
<b>10</b>	<b>运行与维护 .....</b>	<b>19</b>
10.1	定义 .....	19
10.2	交换机初始化 .....	19
10.3	运行和维护具体规定 .....	20
10.4	安全性考虑 .....	20
<b>11</b>	<b>网络管理协议 .....</b>	<b>21</b>
11.1	简单网络管理协议-SNMP .....	21
11.2	团体表格 .....	21
11.3	标准 MIBS .....	21
11.4	厂商指定的 MIBS .....	22
<b>12</b>	<b>环境要求 .....</b>	<b>22</b>
12.1	环境要求 .....	22
12.2	防电磁干扰要求 .....	22
12.3	交换机抗电磁干扰的能力 .....	23
12.4	交换机防雷击能力 .....	23
<b>13</b>	<b>电源与接地 .....</b>	<b>23</b>
13.1	电源 .....	23
13.2	交换机接地要求 .....	24

## 前 言

本标准以 RFC 文档以及 IEEE 标准为依据，针对我国具体要求制定而成。

本标准规定了千兆以太网交换机的技术要求，包括功能、指标、通信接口、通信协议环境要求。本标准列举了千兆以太网交换机必须实现的协议与功能。对协议实现的完整要求在 IEEE 标准及 RFC 协议中指出。本标准主要依据 IEEE 标准与 RFC 文档。随着 IP 技术的不断发展，需要对本标准不断补充和完善。

本标准由信息产业部电信研究院提出并归口。

本标准起草单位：信息产业部电信传输研究所

本标准主要起草人：魏亮

# 千兆比以太网交换机设备技术规范

Technical Specification for Giga Bit Ethernet LAN  
Switch Equipment

YD/T 1099—2001

## 1 范围

本标准规定了千兆比以太网交换机的技术要求，包括功能指标、通信协议、通信接口、环境要求等。千兆比以太网交换机是指拥有千兆比以太网接口的以太网交换机。

本标准中所有未指明的交换机的规定均特指对千兆比以太网交换机的规定。

本标准只对 ISO 的 OSI7 层参考模型中第二层数据链路层上工作的千兆比以太网交换机作规范，第三层交换和更高层交换在本标准范围之外。

本标准适合千兆比以太网交换机的研制开发和技术引进。

在本标准中：

- 必须：表示该条目是本标准必须。违反这样的要求是原则性错误。
- 必须实现：表示该要求必须实现，但不要求缺省使能。
- 不允许（不可以）：标识该条目绝对禁止。
- 应当（建议）：表示在某些特定条件下存在忽视该条目的理由，但是忽视或违反该条目时必须仔细衡量。
  - 应当（建议）实现：与应当（建议）类似，实现时不必要缺省使能。
  - 不应当（不建议）：表示在某些特定条件下存在所描述行为可接受或有效的理由，但实现该行为时必须仔细衡量。
- 可以：标识该条目确实可选。某些厂商可能出于市场或其他原因实现该选项，另一厂商可能出于类似理由不实现该选项。

## 2 引用标准

下列标准包含的条文，通过在本标准中引用而构成为本标准的条文。在标准出版时，所示版本均为有效。所有标准都会被修订，使用本标准的各方应探讨使用下列标准最新版本的可能性。

ANSI/IEEE Std 802.1D (1998) Edition	媒体访问控制(MAC) 网桥
IEEE Std 802.1Q (1998)	虚拟桥接局域网
IEEE Std 802.1g (1995)	远程媒体访问控制桥接
IEEE Std 802.1ad (2000)	多链路段聚合
IEEE Std 802.2 (1998)	逻辑链路控制
IEEE Std 802.3ab (1999)	在 4 对 5 类平衡双绞线上传输千兆比以太网的物理层参数及规定 1000BASE-T
IEEE Std 802.3 (1998) Edition	带碰撞检测的载波监听多重访问方式及物理层定义
RFC1213 (1991)	基于 TCP/IP 的互联网网络管理的管理信息库：MIB-II
RFC2011 SNMPv2 (1996)	对使用 SMIv2 互联网协议的管理信息库

RFC2012 SNMPv2 (1996)	对使用 SMIV2 传输控制协议的管理信息库
RFC2013 SNMPv2 (1996)	对使用 SMIV2 用户数据包协议的管理信息库
RFC1493 (1993)	对网桥管理对象的定义
RFC2233 (1997)	使用 SMIV2 的接口组 MIB
RFC1643 (1994)	对以太网接口类型管理对象的定义
RFC1757 (1995)	远程网络监视管理信息库
RFC2021 (1997)	远程网络监视管理信息库版本 2
RFC2074 (1997)	远程网络监视 MIB 协议标识符
RFC2613 (1999)	对交换网络的远程网络监视管理 MIB 扩展 Version 1.0

### 3 定义术语及缩写

#### 3.1 定义

本标准采用了下列定义。

##### 1) 网桥 (Bridge)

网桥工作于 ISO 的 OSI7 层参考模型中第二层数据链路层的 MAC 子层, 通过转发 MAC 帧实现网络互联。网桥的实现应当符合 ANSI/IEEE Std802.1D, 1998。网桥可以连接同种或不同种 MAC 技术的网络, 利用包含在 MAC 帧中的目的地址和源地址信息作智能转发决定。在连接以太网时, 网桥不但可以扩展物理网络拓扑结构, 还可以将端口上的子网隔离成独立的冲突域。

##### 2) 以太网交换机 (Ethernet Switch)

以太网交换机实质上是支持以太网接口的多端口网桥。交换机通常使用硬件实现过滤、学习和转发数据帧。

交换机必须实现网桥功能中相应功能。

##### 3) 千兆比以太网交换机 (Giga Bit Ethernet Switch)

千兆比以太网交换机即通信接口支持千兆比以太网的以太网交换机。千兆比以太网交换机具体功能要求在本标准第 4 章中详细描述。

##### 4) 存储转发 (storage-forward)

在过滤或转发处理之前, 整个帧必须已经完全接收。

##### 5) 直通转发 (cut-through)

在接受完整个网络帧之前, 转发已经开始的转发方式。

##### 6) 虚拟局域网 (Virtual Local Access Network)

VLAN 功能指通过桥接的局域网内活跃拓扑中工作站的划分, 各 VLAN 使用 VID (VLAN 标识符) 区分。各个 VLAN 是原桥接的局域网的一个子集。

##### 7) 远程桥接 (Remote MAC Bridgeing)

远程媒体访问控制桥接是指在互连的局域网间使用远程媒体访问控制桥的操作以及远程媒体访问控制桥通过非局域网通信设备按照生成树算法配置被桥接局域网的协议。

##### 8) 链路聚合 (Link Aggregation)

多链路聚合是指在逻辑上将多条独立的链路作为一条单独链路使用, 以此获得灵活的高带宽以及链路冗余。

##### 9) 1000 BASE-CX

运行在专门屏蔽平衡铜缆上的 1000 BASE-X (见 IEEE 802.3 39 子句)

##### 10) 1000 BASE-LX

在多模或单模光纤上使用长波长激光设备的 1000 BASE-X。(见 IEEE802.3 38 子句)

##### 11) 1000 BASE-SX

在多模光纤上使用短波长激光设备的 1000 BASE-X。(见 IEEE 802.3 38 子句)

## 12) 1000 BASE-T

IEEE 802.3 对使用 4 对平衡五类线的 1000 Mbit/s CSMA/CD 局域网的物理层规定。(见 IEEE 802.3 40 子句)

## 13) 1000 BASE-X

IEEE 802.3 对使用 ANSI X3.230-1994 (FC-PH) [B19] 9 得到的物理层的 1000 Mbit/s CSMA/CD 局域网的物理层规定。(见 IEEE 802.3 36 子句)

## 14) 8B/10B 传输编码

一种 dc-平衡, 基于字节的数据编码规定。

## 3.2 缩写

本标准使用下列缩写:

AFC	Asymmetric Flow Control	(不对称流量控制)
AUI	ATTACHMENT UNIT INTERFACE	(附加单元接口)
BPDU	Bridge Protocol Data Unit	(桥接协议数据单元)
CRC	Cyclic Redundancy Check	(循环冗余校验)
FCS	Frame Check Sequence	(帧检验序列)
E-ISS	Enhanced Internal Sublayer Service	(增强的内部子层服务)
FID	Filter Identifier	(过滤标识符)
GARP	General Attribute Registration Protocol	(一般属性注册协议)
GARP PDU	GARP Protocol Data Unit	(GARP 协议数据单元)
GID	GARP Information Declaration	(GARP 信息发布)
GIP	GARP Information Propagation	(GARP 信息广播)
GMII	Gigabit Media Independent Interface	(千兆比特媒体无关接口)
GMRP	GARP Multicast Registration Protocol	(GARP 组播注册协议)
GVRP	GARP VLAN Registration Protocol	(GARP VLAN 注册协议)
IETF	Internet Engineering Task Force	(互联网工程任务组)
IGMP	Internet Group Management Protocol	(互联网组管理协议)
ISS	Internal Sublayer Service	(内部子层服务)
IVL	Independent VLAN Learning	(独立的 VLAN 学习)
LAN	Local Area Network	(局域网)
LLC	Logical Link Control	(逻辑链路控制)
MAC	Media Access Control	(媒体控制访问)
MAU	Medium Attachment Unit	(媒体附加接口)
MDI	Media Dependent Interface	(媒体依赖接口)
MIB	Management Information Base	(管理信息库)
MII	Media Independent Interface	(媒体无关接口)
MSDU	MAC Service Data Unit	(MAC 服务数据单元)
NCFI	Non-Canonical Format Indication	(非规范的格式标识符)
PCS	Physical Coding Sublayer	(物理编码子层)
PICS	Protocol Implementation Conformance Statement	(协议实现一致性声明)
PHY	Physical Layer Device	(物理层设备)
PLS	Physical Layer Signaling	(物理层信令)
PDU	Protocol Data Unit	(协议数据单元)
PMA	Physical Medium Attachment	(物理介质接入)
PMD	Physical Medium Dependent	(物理媒体相关)

PVID	Port VID	(端口 VID)
RIF	Routing Information Field(ISO/IEC8802-5)	(路由信息域)
STPID	SNAP-encoded Tag Protocol Identifier SNAP	(编码标记协议标识符)
SVL	Shared VLAN Learning	(共享 VLAN 学习)
TCI	TAG Control Information	(标记控制信息)
TPID	TAG Protocol Identifier	(标记协议信息)
VID	Virtual LAN Identifier	(虚拟局域网标识符)
VLAN	Virtual LAN	(虚拟局域网)

#### 4 千兆以太网交换机功能

交换机必须实现：

##### 1) 千兆以太网接口功能

千兆以太网交换机必须至少拥有一个千兆以太网接口。千兆以太网接口应符合第 5 章规定。

##### 2) 逻辑链路层功能

千兆以太网交换机必须实现一类 LLC 支持类型 1 操作。对 LLC 的实现必须符合 ISO/IEC 8802-2。

##### 3) 数据帧转发功能

转发是指交换机在不同端口所连接的被桥接的 MAC 间交换 MAC 用户数据帧。交换机必须实现转发数据帧。交换机转发数据帧应当实现 IEEE802.1p 中规定的 QoS。数据帧的转发在 6.1 中规定。

##### 4) 数据帧过滤功能

过滤是指交换机为防止数据帧重复，对某些端口上数据帧不转发（丢弃）到其他接口的行为。交换机必须实现基本过滤服务。数据帧的过滤在 6.2 中规定。

##### 5) 维护决定数据帧转发及过滤的信息

交换机必须实现维护数据帧转发/过滤信息。对上述信息的维护在第 7 章中规定。

##### 6) 运行维护功能

交换机必须实现运行维护功能。运行维护功能在第 8 章中规定。

##### 7) 网络管理功能

交换机必须实现网络管理接口及协议。网络管理接口及协议在第 9 章中规定。

#### 5 千兆以太网接口

建议千兆以太网接口按照图 1 分层实现。

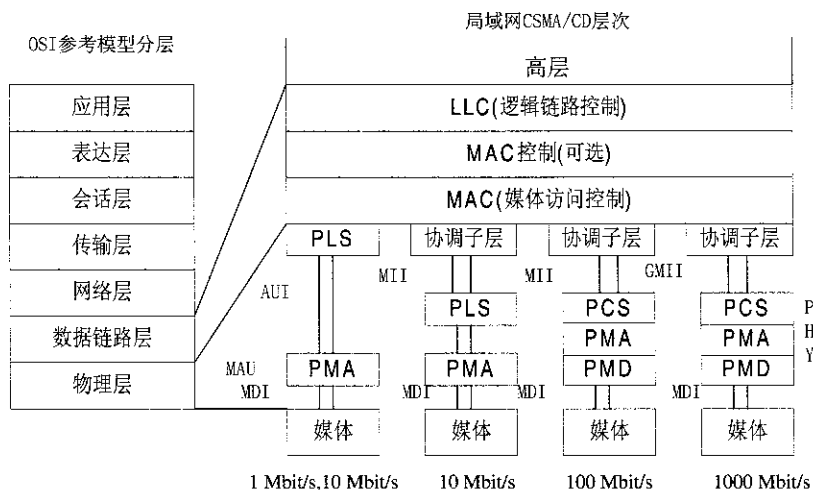


图 1 千兆以太网接口分层



## 5.1 千兆以太网接口类型

千兆以太网接口可以采用 1000 BASE-SX、1000 BASE-LX、1000 BASE-CX 及 1000 BASE-T。

### 5.1.1 1000 BASE-SX

符合 IEEE802.3 定义的 1000 BASE-SX 以太网标准，这种以太网标准表示在光缆上使用短波长激光发送器，基于 8B/10B 信号的 1000 Mbit/s 千兆位以太网，特性为：

- 端口类型为光端口；
- 传输介质为光纤，1000 BASE-SX 网段基于波长为 770~860 nm 的光纤激光传输器，它可以被耦合到多模光缆中，可以使用两种光纤，一种是纤芯为 62.5  $\mu\text{m}$  的多模光纤，最长为 275 m；另一种是纤芯为 50  $\mu\text{m}$  的多模光纤，最长为 550 m，标准将指定为最大可达到的距离；
- 信号编码方式为 8B/10B；
- 半双工下的介质访问控制协议为具有冲突检测的载波侦听多路访问协议（CSMA/CD），但是时间槽长度增加到 512 字节；
- 信号方式为基带方式，这种信号方式中，在任何时刻仅有一个信号在介质中传送；
- 支持全双工和半双工的数据通信方式，在全双工下，工作速率达到 2000 Mbit/s，在半双工下，工作速率为 1000 Mbit/s；
- 连接器：SC(female)（符合 IEC-61754-4 及 ISO/IEC11801/尺度和接口规范中定义）或者其他标准的连接器。

### 5.1.2 1000 BASE-LX

符合 IEEE802.3 定义的 1000 BASE-LX 以太网标准，这种以太网标准表示在光缆上使用长波长激光发送器，基于 8B/10B 信号的 1000 Mbit/s 千兆位以太网，特性为：

- 端口类型为光端口；
- 传输介质为光纤，1000 BASE-LX 网段基于波长为 1270~1355 nm 的光纤激光传输器，它可以被耦合到单模或者多模光缆中，可以使用三种光纤，一种是纤芯为 62.5  $\mu\text{m}$  的多模光纤，最长为 550 m；一种是纤芯为 50  $\mu\text{m}$  的多模光纤，最长为 550 m；另一种是纤芯为 10  $\mu\text{m}$  的单模光纤，最长为 5000 m，标准将指定为最大可达到的距离；
- 信号编码方式为 8B/10B；
- 半双工下的介质访问控制协议为具有冲突检测的载波侦听多路访问协议（CSMA/CD），但是时间槽长度增加到 512 字节；
- 信号方式为基带方式，这种信号方式中，在任何时刻仅有一个信号在介质中传送；
- 支持全双工和半双工的数据通信方式，在全双工下，工作速率达到 2000 Mbit/s，在半双工下，工作速率为 1000 Mbit/s；
- 连接器：SC(female)（符合 IEC-61754-4 及 ISO/IEC11801/尺度和接口规范中定义）或者其他标准的连接器。

### 5.1.3 1000 BASE-CX

符合 IEEE802.3 定义的 1000 BASE-CX 以太网标准，这种以太网标准是基于短距离铜质跳线的以太网，特性为：

- 端口类型为电端口；
- 传输介质为跳线电缆（一端连续屏蔽的平衡电缆）。最长 25 m。
- 信号为差分正向 ECL（PECL）逻辑信号。通常需要在电缆或连接其中构造均衡器。
- 半双工下的介质访问控制协议为具有冲突检测的载波侦听多路访问协议（CSMA/CD），但是时间槽长度增加到 512 字节；
- 支持全双工和半双工的数据通信方式，在全双工下，工作速率达到 2000 Mbit/s，在半双工下，工作速率为 1000 Mbit/s；
- 连接器：1) D 微型连接器（符合 IEC-60807-3 中定义）；2) 8 针带屏蔽 ANSI 光纤通道类型

2 连接器（符合 IEC-61076-3-103 中定义）。

#### 5.1.4 1000 BASE-T

符合 IEEE802.3ab 定义的 1000 BASE-T 以太网标准，这种以太网标准表示基于双绞线电缆的 1000 Mbit/s 千兆位以太网，特性为：

- 端口类型为电端口；
- 传输介质可以是 5 类非屏蔽双绞线（UTP），最长为 100m，传输的数据线数为 4 对；
- 半双工下的介质访问控制协议为具有冲突检测的载波侦听多路访问协议（CSMA/CD），但是时间槽长度增加到 512 字节；
- 信号方式为基带方式，这种信号方式中，在任何时刻仅有一个信号在介质中传送；
- 支持全双工和半双工的数据通信方式，在全双工下，工作速率达到 2000 Mbit/s，在半双工下，工作速率为 1000 Mbit/s；
- 连接器：RJ45。

#### 5.2 物理传输媒体

千兆以太网接口类型中规定，千兆以太网接口物理传输媒体可以采用：

- 纤芯为 62.5  $\mu\text{m}$  的多模光纤；
- 纤芯为 50  $\mu\text{m}$  的多模光纤；
- 纤芯为 10  $\mu\text{m}$  的单模光纤；
- 非屏蔽 5 类双绞线；
- 短跳线电缆。

#### 5.3 PMD（物理介质相关）子层

PMD 子层的功能是支持在 PMA 子层和介质之间交换串行化的 8B/10B 符号代码位。PMD 子层将这些电信号转换成适合于在某种特定介质上传输的形式。

PMD 是物理层的最低子层，千兆以太网标准中规定物理层负责从介质上发送和接收信号。IEEE Std 802.3, 1998 Edition 以及 IEEE Std 802.3ab 共规定 4 种类型的介质。

PMD 子层在 IEEE Std 802.3, 1998 Edition 38, 39 子句以及 IEEE Std 802.3ab 中规定。实现千兆以太网接口 PMD 子层的交换机必须符合上述标准中相关章节规定。

#### 5.4 PMA（物理介质接入）子层

PMA 子层提供了 PCS 和 PMD 层之间的串行化服务接口。和 PCS 子层的连接称为 PMA 服务接口。在把 10 位符号传递给 PMD 之前 PMA 子层的发送部分把它们转化成串行位流。PMA 子层的接收部分在把接收到的串行位流传递给 PCS 之前把它们转化成 10 位符号。另外 PMA 子层还从接收位流中分离出用于对接收到的数据进行正确的符号对齐（定界）的符号定时时钟。

PMA 子层在 IEEE Std 802.3, 1998 Edition 36 子句中规定。实现千兆以太网接口 PMA 子层的交换机必须符合上述标准中相关章节规定。

#### 5.5 PCS（物理编码）子层

PCS 子层位于协调子层（通过 GMII）和物理介质接入层（PMA）子层之间。PCS 子层完成将经过完善定义的以太网 MAC 功能映射到现存的编码和物理层信号系统的功能上去。

PCS 子层和上层 RS/MAC 的接口由 GMII 提供，与下层 PMA 接口使用 PMA 服务接口。对于 GMII 没有定义连接器。虽然接口上信号强制规定，但所有接口上的信号不需要为实现兼容而物理外露。

PCS 子层在 IEEE Std 802.3, 1998 Edition 36 子句中规定。实现千兆以太网接口 PCS 子层的交换机必须符合上述标准中相关章节规定。

#### 5.6 RS(协调子层)和 GMII(千兆位介质无关接口)

协调子层的功能是将 GMII 的八位宽度通路数据和相关控制信号映射到原始 PLS 服务接口定义 (MAC/PLS)接口上。

GMII 接口提供了千兆位 MAC 和物理层间的逻辑接口。GMII 和协调子层使 GMAC 可以连接到不

同类型的物理介质上。GMII 中数据通路宽度 8 位。每时钟周期内通过一个八位组。

RS 子层与 GMII 在 IEEE Std 802.3, 1998 Edition 35 子句中规定。实现千兆以太网接口 RS 子层与 GMII 的交换机必须符合上述标准中相关章节规定。

### 5.7 自动协商

自动协商功能是指链路两端设备自动配置，不需要用户干预。自动协商可以协商出两端设备可以共同正常工作的最优模式。当设备间没有可共享的最优模式时，自动协商确保网络上其他设备能正常工作不受影响，并且通过网管功能报告。

自动协商功能实际上是 PCS 内部的一种功能，由于比较重要，在本节单独描述。

千兆以太网上的自动协商应当可以工作在使用各种连接器类型的光纤和铜介质上，且应使用普通 8B/10B 代码交换普通页和下一页信息。在千兆以太网上的 PAUSE 帧操作应当可以不对称。

千兆以太网自动协商在 IEEE Std 802.3, 1998 Edition 37 子句中规定。实现千兆以太网自动协商的交换机必须符合上述标准中相关章节规定。

### 5.8 逻辑链路层功能

为支持网桥协议实体，交换机端口必须支持 LLC 类型 1 过程运行。对 LLC 的实现必须符合 ISO/IEC 8802-2。

交换机端口可以支持其他类型 LLC 过程，该类型 LLC 可以由其他上层协议使用。

## 6 数据帧的转发及过滤

### 6.1 转发数据帧

交换机转发数据帧必须：

- 1) 符合寻址规定；
- 2) 提供：
  - 在不提供 48 比特通用管理地址时分配组 MAC 地址来标识网桥协议实体的途径；
  - 端口标识符在实现生成树算法及协议时标识交换机每一端口。

交换机转发数据帧可以：

- 1) 提供转发时控制优先级映射的能力；
- 2) 提供多种流量分类；
- 3) 对独立 MAC 地址的转发行为作规定；
- 4) 管理转发帧的优先级。

数据帧的转发可以基于存储转发或直通式转发。

交换机必须支持存储转发。

交换机可以支持直通式转发。实现直通式转发的交换机必须缺省设置为存储转发。

### 6.2 过滤数据帧

交换机过滤数据帧必须符合：

- 实现基本过滤服务，对每个端口关联单一流量类 (Traffic Class)；
- 对过滤数据库下列参数使用规定的值：
  - 过滤数据库大小，过滤数据库所能容纳的最大条目数；
  - 永久数据库大小，永久数据库所能容纳的最大条目数。

交换机过滤数据帧可以：

- 提供读取和更新过滤数据库和永久数据库的能力；
- 提供设置过滤数据库更新时间的能力。提供该能力的交换机应当实现本标准指定的所有可选值；
- 对独立 MAC 地址的过滤行为作规定。

### 6.3 支持转发/过滤数据帧的功能

交换机必须实现下列支持数据帧转发/过滤、提供 QoS 的功能：

- 帧接收；
- 丢弃所收到的错误帧；
- 丢弃 `frame_type` 不是 `user_data_frame` 或 `mac_action` 参数不是 `request_with_no_response` 的帧；
- 如果需要，重新产生用户优先级；
- 丢弃过滤信息应用指定需要丢弃的帧；
- 丢弃传输服务用户数据单元大小超过 ISO/IEC 15802-3, 6.3.8 中规定的帧；
- 发送所收到得到其他端口的帧；
- 根据过滤信息应用选择流量类；
- 根据流量类对帧排队；
- 丢弃超过最大网桥传输时延的帧；
- 在排队的帧中选择帧传输；
- 选择带外访问优先级 (ISO/IEC 15802-3, 6.3.9)；
- 如果需要，映射服务数据单元，重新计算帧检验序列；
- 帧发送。

#### 6.3.1 帧接收

关联在交换机端口上的 MAC 实体应当检查所连接局域网上所有被传输的帧。

所有正确的帧都被提交到 `M_UNITDATA` 标识原语，按照下面描述处理。

所有 `M_UNITDATA.indication` 原语中 `frame_type` 和 `mac_action` 参数分别是 `user_data_type` 和 `request_with_no_response` 的帧被提交到学习和转发进程。

所有 `frame_type` 和 `mac_action` 参数是其他值的帧不应提交到转发进程，但可以提交到学习进程。

所有 `frame_type` 是 `user_data_type`，目的地址指向交换机端口的帧应当提交到 LLC。这样的帧的目的地址域中应当携带交换机端口的独立地址或者关联在端口上的组地址。提交给 LLC 的帧同样可以按照上面描述提交给学习进程和/或转发进程。

将交换机端口作为终端的帧和从其他端口转发到端口的帧也应当提交给 LLC。

#### 6.3.2 重新生成用户优先级

所收到帧的 `user_priority` 由包含在帧中的优先级信息和接收端口的用户优先级重新生成表得到。对于每个接收端口，用户优先级重新生成表应当包含 8 个条目，对应于 `user_priority` 可能的 8 个值 (0~7)。每个条目指示给定优先级后，重新生成的用户优先级。

表 1 定义了对所收到的数据中指示的 8 种可能的用户优先级所重新生成的用户优先级的缺省值。

表 1 用户优先级重新生成

用户优先级	缺省得重新生成用户优先级	范围
0	0	0~7
1	1	0~7
2	2	0~7
3	3	0~7
4	4	0~7
5	5	0~7
6	6	0~7
7	7	0~7

交换机可选支持通过管理功能改变用户优先级重新生成值。如果支持该功能，交换机应当能对任意接收端口的任意到达优先级独立指定范围中任意值。

### 6.3.3 帧发送

关联在交换机端口上的每个 MAC 实体应当发送由 MAC 中继实体提交的帧。

转发进程提交被中继的帧用于发送。被发送帧所关联 M\_UNITDATA.request 原语使用所接收到相应 M\_UNITDATA.indication 原语中源和目的地址域。

LLC 协议数据单元由 LLC 作为交换机端口提供的 MAC 服务的使用者提交。用于传输上述协议数据单元所发送的帧将端口的独立 MAC 地址作为源地址。

每一个帧都发给 MAC 进程，供特定的 IEEE802 局域网技术观察。相应 M\_UNITDATA.request 原语中 frame\_type 和 mac\_action 参数分别应当为 user\_data\_type 和 request\_with\_no\_response。

由交换机端口所提供的 MAC 服务的 LLC 用户请求的帧发送应当提交给 MAC 中继实体。

### 6.3.4 执行拓扑限制

当且仅当下列条件满足时，交换机的端口才能作为可用的传输端口：

- 接收帧的端口处于转发状态；
- 需要发送帧的端口处于转发状态；
- 发送帧的端口不同于收到该帧的端口；
- 所需发送帧中 mac\_service\_data\_unit 大小不超过发送端口连接的局域网所支持的 mac\_service\_data\_unit 最大尺寸。

### 6.3.5 过滤帧

转发进程应基于下列条件作过滤决定：

- 收到的帧中 MAC 目的地址；
- 过滤数据库中关于 MAC 地址和接收端口的信息；
- 对可用发送端口的缺省组过滤行为。

### 6.3.6 帧排队

转发进程应当为排队的帧提供存储服务，等待时机将帧提交给关联在交换机端口上的 MAC 实体。

交换机可以在端口上提供多个队列。帧基于使用流量类的用户优先级决定所使用的存储队列，上述流量类是关联在每个端口上状态信息的一部分。对每个可能的用户优先级必须对流量类赋值。用户优先级可以由 0~7。队列应当一一对应到流量类。

出于管理考虑，交换机应最多支持 8 个流量类来支持将各个用户优先级的帧独立排队。

表 2 给出用户优先级到流量类映射的建议。

表 2 用户优先级到流量类映射的建议

		可用的流量类数量							
		1	2	3	4	5	6	7	8
用户 优 先 级	0 (缺省)	0	0	0	1	1	1	1	2
	1	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	1
	3	0	0	0	1	1	2	2	3
	4	0	1	1	2	2	3	3	4
	5	0	1	1	2	3	4	4	5
	6	0	1	2	3	4	5	5	6
	7	0	1	2	3	4	5	6	7

当帧提交到关联在端口上的 MAC 实体之后，应当从存储的队列中删除。即使发送失败，交换机转发进程不应试图重传。当缓存溢出时，交换机可以在队列中删除帧；交换机可以不按次序发送。

端口离开转发状态时，帧队列应当删除。

在某一特定端口队列中帧删除并不表示删除其他端口帧队列中该帧。

## 7 维护决定数据帧转发/过滤的信息

### 7.1 维护过滤/转发信息

交换机应当实现：

- 计算及配置被桥接的以太网的拓扑；
- 永久配置保留地址；
- 显式配置静态过滤信息；
- 通过察看被桥接局域网的流量中源地址自动学习对单目的地址的动态过滤信息；
- 对所学到的动态过滤信息设置定时器实现按时间老化；
- 通过 GMRP 协议自动添加/删除动态过滤信息；
- 显式配置关联到交换机端口上的流量类信息；
- 显式配置关联到交换机端口上的端口 VID (PVID)；
- 显式配置关联到交换机端口上的允许接收帧类型参数；
- 显式配置关联到交换机端口上的使能入口过滤参数；
- 通过使用 GVRP 自动配置动态 VLAN 注册实体；
- 显示配置通过静态 VLAN 注册实体方式关联 GVRP 操作的管理控制；
- 通过观察网络流量自动学习关联到 VLAN 的 MAC 地址；
- 显式配置每个端口需要的出口标记。

### 7.2 网桥协议数据单元

交换机应当实现网桥协议数据单元。网桥协议数据单元在 ANSI/IEEE Std 802.1D 1998 中定义。

#### 7.2.1 网桥协议数据单元结构

网桥协议数据单元用于实现生成树算法/协议的网桥协议实体。网桥协议数据单元分为两种：配置 BPDU 和拓扑变化通知 BPDU。

##### 1) 配置 BPDU

在图 2 中配置 BPDU 结构如下：

- 协议标识符位于第 1~2 字节。使用 0000 0000 0000 0000 标识生成树算法/协议。
- 协议版本标识符位于第 3 字节，使用 0000 0000。
- BPDU 类型位于第 4 字节，使用 0000 0000，标识配置 BPDU。
- 拓扑改变确认标志在第 5 字节第 8 比特。
- 拓扑变化标志位于 BPDU 第 5 字节第 1 比特。
- 根标识符位于 BPDU 第 6~13 字节。
- 根路径代价位于 BPDU 第 14~17 字节。
- 网桥标识符位于 BPDU 第 18~25 字节。
- 端口标识符位于 BPDU 第 26, 27 字节。
- 消息时限定时器值位于 BPDU 第 28, 29 字节。
- 最大时限定时器值位于 BPDU 第 30, 31 字节。
- Hello 时间定时器值位于 BPDU 第 32, 33 字节。
- 传递时延定时器值位于 BPDU 第 34, 35 字节。

消息时限值应小于最大时限值。

协议标识符	1
	2
协议版本标识符	3
BPDU 类型	4
标志	5
根标识符	6
	7
	8
	9
	10
	11
	12
	13
根路径代价	14
	15
	16
	17
网桥标识符	18
	19
	20
	21
	22
	23
	24
	25
端口标识符	26
	27
消息时限	28
	29
最大时限	30
	31
Hello 时间	32
	33
传递时延	34
	45

图 2 配置 BPDU

## 2) 拓扑变化通知 BPDU

图 3BPDU 结构如下：

- 协议标识符位于 BPDU 第 1~2 字节。使用 0000 0000 0000 0000，标识生成树算法/协议。
- 协议版本标识符位于 BPDU 第 3 字节，使用 0000 0000。

—— BPDU 类型位于 BPDU 第 4 字节，使用 1000 0000。

协议标识符	1
	2
协议版本标识符	3
BPDU 类型	4

图 3 拓扑变化通知 BPDU

### 7.2.2 对 BPDU 的验证

当且仅当 BPDU 长度大于 4 字节且满足下面条件之一时，网桥协议实体应当根据生成树算法/协议中规定处理收到的 BPDU。

- BPDU 类型为配置 BPDU 且 BPDU 包含 35 字节，BPDU 消息时限参数小于最大时限参数；
- BPDU 4 字节且类型为拓扑变化通知 BPDU。

### 7.3 GARP 组播注册协议 (GMRP)

交换机应当实现 GMRP。GMRP 在 ANSI/IEEE Std 802.1D 1998 中定义。

GMRP 提供一种机制允许交换机和终端系统动态注册组成员信息；该信息在所有支持扩展过滤服务的交换机/网桥间传播。GMRP 利用 GARP 提供服务。

需要通过 GMRP 注册、消除注册或传播的信息有：

- 组成员信息。该信息指示存在一个或多个 GMRP 参与者是某特定组成员，并且携带相关该特定组的组 MAC 地址。对组成源信息的交换可能是因为创建或更新过滤数据库中的组注册实体来指示组成员在哪个端口上被注册。
- 组服务要求信息。该信息只是一个或多个 GMRP 参与者要求将所有组或为注册组应用缺省组过滤行为。

交换机实现 GMRP 的一致性要求。实现 GMRP 的交换机应当符合：

- 实现 GARP 应用及注册状态机以及 LeaveAll generation 机制。
- 按照状态机要求交换 GARP PDU，该 PDU 携带应用特定信息。
- 广播注册信息应符合基本生成树的 GIP 操作。
- GMRP 实现符合 ANSI/IEEE Std 802.1D 1998 中 10.3 字句。
- 对携带 GARP 应用地址的数据帧实行转发、过滤及丢弃。

### 7.4 一般属性注册协议 (General Attribute Registration Protocol GARP)

交换机应当实现 GARP。GARP 在 ANSI/IEEE Std 802.1D 1998 中定义。

#### 7.4.1 GARP 定义

GARP 提供一种一般属性分发能力，该能力由 GARP 应用在桥接的局域网上与其他 GARP 参与者注册及消除注册属性值。属性的类型、属性的值和属性值相关的语义由 GARP 应用决定。

#### 7.4.2 GARP 实现要求

GARP 及相关算法用作在桥接的网络中建立、维护和解析属性注册以及在连接到局域网上的 GARP 参与者间分发注册信息。交换机必须实现 GARP。实现 GARP 的交换机必须满足下面要求。下面要求包括应用及注册机要求，失败条件下的错误恢复，性能，可扩展性，与非 GARP 设备反向兼容以及对交换机、终端和网络的负荷。

- 1) 实现 GARP 必须允许 GARP 参与者连接到桥接的局域网上发布关联于 GARP 应用的属性值声明。
- 2) 实现 GARP 必须允许 GARP 参与者连接到桥接的局域网上撤销关联于 GARP 应用的属性值声明。
- 3) 实现 GARP 必须允许交换机将收到的声明向通过该交换机所连接的局域网上能够访问的 GARP 参



与者广播。

4) 实现 GARP 必须允许 GARP 参与者维护指示当前声明状态和参与设备每个端口属性注册的状态信息。

5) 实现 GARP 必须允许 GARP 参与者消除关于部分或所有桥接网络的属性的状态信息。

6) 发布/撤销属性注册信息以及属性注册信息在桥接网络中传播的时延应当保持足够小，且按照网络直径线性增长。

7) 当一个 GARP 失败时，GARP 应保持可用。

8) 丢失一个数据帧时，GARP 应保持可用。

9) GARP 应当能正常运行在：

—— 同构网络中：即桥接的网络中所有交换机/网桥都支持基本过滤服务及扩展的过滤服务。

—— 异构网络中：即某些交换机/网桥只支持基本过滤服务，某些交换机/网桥支持基本过滤服务和扩展过滤服务。

10) GARP 需要的带宽应当足够小，且与桥接网络上的流量无关。该带宽应当只与局域网上维持成员关系的组的规模有关。

#### 7.4.3 GARP 参与者间互操作性要求：

1) 对每个定义的 GARP 应用必须在协议交换帧的目的地址使用唯一的组 MAC 地址，称为 GARP 应用地址。GARP 应用地址中 01-80-C2-00-00-20 用作 GMRP 地址，01-80-C2-00-00-21 到 01-80-C2-00-00-2F 保留供将来使用。实现已定义 GARP 应用地址的交换机不需要将目的地址是 GARP 地址的帧转发；未实现已定义 GARP 应用地址的交换机需要将目的地址是 GARP 地址的帧向所有活跃拓扑终端端口转发。

表 3 GARP 应用地址

分配	值
GMRP 地址	01-80-c2-00-00-20
保留	01-80-c2-00-00-21
保留	01-80-c2-00-00-22
保留	01-80-c2-00-00-23
保留	01-80-c2-00-00-24
保留	01-80-c2-00-00-25
保留	01-80-c2-00-00-26
保留	01-80-c2-00-00-27
保留	01-80-c2-00-00-28
保留	01-80-c2-00-00-29
保留	01-80-c2-00-00-2a
保留	01-80-c2-00-00-2b
保留	01-80-c2-00-00-2c
保留	01-80-c2-00-00-2d
保留	01-80-c2-00-00-2e
保留	01-80-c2-00-00-2f

2) GARP 参与者间 GARP PDU 的收发应采用所考虑 GARP 应用的格式，使用一般 PDU 格式，应通过 LLC 类型 1 过程取得。源和目的 LLC 地址应采用分配给生成树算法的标准 LLC 地址。

- 3) GARP 参与者收到格式不正确的 PDU 时应丢弃。
- 4) GARP 参与者的协议行为应符合协议中状态机的描述。

## 7.5 生成树算法及协议 (Spanning Tree Algorithm and Protocol)

生成树算法及协议将通过交换机或网桥连接的局域网的拓扑简化成一棵生成树。

交换机应当实现生成树算法及协议。对生成树算法及协议的实现必须符合 ANSI/IEEE Std 802.1D (1998) 中第 8 子句。

生成树算法的实现必须符合下列要求:

- 该算法必须能将一个被桥接的任意拓扑局域网中活跃的拓扑结构配制成一棵生成树。该生成树消除环路, 从而在任意两个节点间最多存在一条路径。
- 当由于交换机或网桥失败或瘫痪时, 该算法必须能在剩下局域网拓扑中自动重新配置生成树拓扑, 达到冗余目的。当存在交换机或网桥端口加入到桥接的局域网中时, 该算法必须能自动调节适应新的拓扑结构, 避免产生环路。
- 在任意规模桥接的局域网中, 构成生成树的活跃的拓扑必须稳定。该算法必须在大多数情况下在可预知的较短间隔内收敛, 减少网络不可用时间。
- 该算法的结果应当可预测, 可重复。结果应当可以由对算法参数的管理选择, 从而通过配置管理与流量分析达到性能管理的目的。
- 该算法应当对终端系统透明, 终端系统在使用 MAC 服务时应当无法得知是连接在单一局域网或桥接的局域网上。
- 由交换机或网桥使用用于建立或维护生成树的带宽必须只占链路可用带宽一小部分, 并且不随桥接网络的规模增加而增加。

另外为降低交换机及其配置的复杂性, 生成树算法的实现必须:

- 该算法关联交换机每端口的内存应当独立于桥接网络规模。
- 交换机在连入桥接网络之前除必要的 MAC 地址外不需作任何附加的配置便可正常工作。

为实现生成树算法, 交换机必须:

- 配置一个唯一的 MAC 组地址, 由桥接的局域网上所有其它交换机或网桥识别, 该地址用于标识连接在局域网上交换机或网桥桥接协议实体。
- 交换机的标识在桥接的局域网范围内唯一。
- 交换机的每个独立端口必须配置单独的端口标识, 该标识的分配独立于其他交换机或网桥的端口标识。

交换机必须为上述参数分配, 实现分配机制。当交换机使用 48 比特全球统一管理地址时, 该唯一的 MAC 地址作为交换机组地址标识桥接协议实体。

另外为使生成树算法结果可配置, 交换机必须实现:

- 一种途径用作给予交换机在桥接的局域网中相对其他路由器或网桥的相对优先级。
- 一种途径用作给予某端口在该交换机中相对其他端口的相对优先级。
- 一种途径用作给予端口上路径的代价。

## 7.6 VLAN (虚拟局域网) 功能

交换机应当实现 VLAN。VLAN 在 ANSI/IEEE Std 802.1Q 1998 中定义。

### 7.6.1 VLAN 定义

VLAN 功能指通过桥接的局域网内活跃拓扑中工作站的划分, 各 VLAN 使用 VID (VLAN 标识符) 区分。各个 VLAN 是原桥接的局域网的一个子集。VLAN 具有如下特性:

- 可以在所有的 IEEE802 局域网 MAC 层协议上实现, 可以在共享媒体或点到点的局域网上实现。
- 对网上的工作站划分逻辑组, 组内成员通信正如同在同一局域网上通信。并且对逻辑组成员的增删改查提供方便的机制。

—— 管制 VLAN 间的流量。交换机只在同一 VLAN 中转发数据包，广播包和多播包。

### 7.6.2 实现 VLAN 交换机的功能要求

实现 VLAN 功能的以太网交换机除应符合上文中对交换机的功能要求外，还应符合：

- 1) 按照下面方法转发、过滤数据帧：
- 2) 在每个端口上对可接受的帧类型参数支持至少一个可允许的值。
- 3) 对每个支持非标记和优先级标记帧的端口上支持：
  - 端口 VLAN 标记值 (PVID)；
  - 能够配置至少一个 VLAN，非标记的帧的端口集包含该端口。
  - 通过管理操作配置 PVID。
  - 通过管理操作配置静态过滤实体。

4) 支持在转发的帧中插入标记头、更改标记头或删除标记头。这些增删改行为由每个端口上可接受帧类型参数以及每个端口是否被配制成传输 VLAN 标记的帧或/和非标记帧决定。一对端口间的帧转发要求概括在表 4 中。

表 4 帧转发要求

		接受端口接受 (不丢弃)		
		VLAN 标记	优先级标记	非标记
发送 端口 发送	非标记	应支持删除标记头	应支持删除标记头	N/A
	优先级标记	如果目标 MAC 所要求的格式与所接受的格式不同，则应当支持改变被标记的帧的格式。	如果目标 MAC 所要求的格式与所接受的格式不同，则应当支持在标记头中插入或改变 non-null 的虚拟局域网标识符 (VID)。	应当支持插入适合目标 MAC 格式的标记头，携带 non-null 的 VID。

5) 在所有端口上支持通过一般属性注册协议 (GARP) 和 VLAN 注册协议 (GVRP) 来自动配置和管理 VLAN 拓扑信息。

6) 通过静态或动态 VLAN 注册实体来支持在过滤数据库中包含至少对一个 VLAN 的动态或静态配置信息。

- 7) 支持至少一个过滤标识符 (FID)
- 8) 支持对每一个所支持的 FID 分配至少一个 VID。

支持 VLAN 的交换机可选支持：

1) 支持扩展过滤模式 (Extended Filtering Mode) 和 GARP 组播注册协议 (GMRP) 中的操作。

2) 通过动态或静态 VLAN 注册实体，支持在过滤数据库中包含超过一个最多达 4094 个 VLAN 的动态或静态配置信息。

3) 在每个端口上对可接受帧类型参数支持所有允许的值。当支持所有允许值时，通过管理配置参数值。

- 4) 支持使能或禁止入口过滤。
- 5) 支持配置多个 VLAN，这些 VLAN 非标记端口集包含某一端口。
- 6) 支持下列管理功能
  - 配置管理；
  - 错误管理；
  - 性能管理；
  - 安全管理；
  - 账号管理。

- 7) 支持多个 FID。
- 8) 支持对每个支持的 FID 分配多个 VID。
- 9) 支持通过管理配置 VLAN 强制学习。
- 10) 支持通过管理对 FID 配置固定的 VID。

### 7.6.3 VLAN 结构概述

VLAN 框架结构基于 3 层结构。包括：

- 配置；
- 对配置信息的分发；
- 转发。

VLAN 结构框架如图 4 所示。

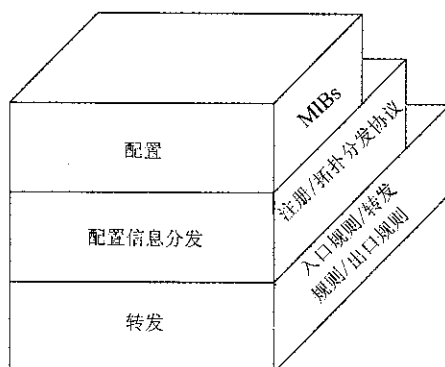


图 4 VLAN 结构框架

### 7.7 远程桥接

远程媒体访问控制桥接是指在互连的局域网间使用远程媒体访问控制桥的操作以及远程媒体访问控制桥通过非局域网通信设备按照生成树算法配置被桥接局域网的协议。

建议千兆以太网交换机实现远程桥接，对远程桥接的实现必须符合 IEEE Std 802.1g。

### 7.8 多链路聚合

多链路聚合是指在逻辑上将多条独立的链路作为一条单独链路使用，以此获得灵活的高带宽以及链路冗余。

建议千兆以太网交换机实现多链路聚合。由于规范链路聚合的协议 IEEE Std 802.1ad 1997 年开始制订，2000 年完成。所以对链路聚合的实现建议符合 IEEE Std 802.1ad。

## 8 流量控制

当流量超过交换机的最大传输能力时，交换机吞吐量可能下降。本标准规定的交换机不允许因背板能力不足引起拥塞。由于可能存在多个端口向某一端口发流量的情况，交换机应当实现流量控制。

流量控制是一种被交换机或拥塞实体用于限制网络访问的机制。流量控制通过对缓存器设置上限、修改发送速率或将发送源关闭一段时间实现。对触发流控机制的参数可以根据流量情况来静态或动态调整。

### 8.1 流量控制性能衡量

流量控制性能衡量需要将不同的流量控制策略进行比较。此外还可以对流量控制策略的参数进行调节优化性能。三个常用的性能衡量指标是吞吐量、包延时和丢包率。

包吞吐量代表设备处理网络负载的能力。在低负载情况下，吞吐量等于网络负载。在负载较高情况下，由于对资源的竞争吞吐量会下降。

包延时是吞吐量的函数。一种特定的流控可能针对高负载下时延特性优化，对低负载下时延不优化。

丢包率是用来衡量由于重传超时，过多的冲突，缓冲区溢出造成数据包丢失。随着丢包率的增加，吞吐量会降低，时延会增加。

## 8.2 半双工下的流控

### 8.2.1 背压流控

背压流控指交换机拥塞时，某个帧到达输入端口时在帧上加一个强制的冲突，迫使远端 DTE 放弃发送。远端 DTE 经后退间隔后重传。如果此时拥塞尚未解除，可以继续使用背压流控机制。对背压流控端口的选择在本标准范围之外。

背压式流控可以采用如下两种方式：

- 1) 背压流控可以简单地拥塞所有源端口，直到拥塞解除。这样做会影响网络中其他接点的性能。
- 2) 背压流控也可以检查帧中目的地地址，如果属于拥塞端口则产生强制冲突。

交换机可选实现背压流控。实现背压式流控的交换机建议采用方式 2)。

### 8.2.2 载波扩展流控

载波扩展流控是指当交换机拥塞时，交换机产生载波侦听信号。在发送载波侦听信号期间所有的 DTE 被抑制。在拥塞期间，交换机发送载波侦听信号必须发送有效数据位。

交换机可选实现载波扩展流控。对发送载波侦听信号端口的选择在本标准范围之外。

## 8.3 全双工下流控

### 8.3.1 PAUSE 控制

全双工下交换机流控可以采用 IEEE802.3 流量控制（又称 PAUSE 控制）。PAUSE 控制在 IEEE std802.3, 1998 版附件 31B 中规定。

PAUSE 控制用作禁止对端在一定时间内发送除 MAC 控制帧之外的帧。是否允许发送 PAUSE 控制在自动协商中决定。IEEE802.3x 定义了两个方向上都支持 PAUSE 功能的 PAUSE 帧格式配置。通过对 PAUSE 位置位，链路两端的设备可以发送并接受 PAUSE 帧。

PAUSE 帧是一个特殊编码的通用 MAC 控制帧。该帧是一个具有最小合法长度的 802.3 以太网帧。该帧具有以下特点：

- 1) 目的地址域是为 MAC 控制 PAUSE 帧单独保留的多目地址 (01-80-c2-00-00-01)。
- 2) 源地址域是源/发送站点的 48 bit 地址。
- 3) 两字节的长度/类型域包含 16 进制值 88-08。表示 802.3 局域网的 MAC 控制帧。
- 4) MAC 控制操作码使用 00-01。
- 5) MAC 控制参数包含 2 字节的 PAUSE 定时器值。该定时器值 16 比特，以 LSB 在先方式传送。

PAUSE 时间单位是 512 位时间+1。

特别指出：

- 1) 当发送器暂停时不禁止发送 PAUSE 帧。
- 2) 0 是在 PAUSE 帧中传输的有效暂停定时器值。一个被暂停的站点在得到下一个 PAUSE 帧后可以重载暂停定时器值。
- 3) 由于为 PAUSE 帧保留的多目地址会被 802.1D（生成树）网桥特殊对待，无论 802.1D 网桥端口状态如何，是否实现 MAC 控制子层，网桥不传播 PAUSE 帧。

建议交换机支持 PAUSE 控制。

### 8.3.2 不对称流量控制 (AFC)

PAUSE 控制提供了对称的流量控制。交换机可以支持不对称流量控制 (AFC)。不对称流量控制是阻止源头的流量，使交换机不需要使用更多的缓冲区。在 AFC 配置下，交换机具有对端站流控的能力，端站不能对交换机流控。

互联的交换机间不应使用 AFC。

交换机对 AFC 的实现可选。

## 8.4 流量控制策略

启动流量控制通常使用的策略有基于水位的流量控制、基于信用的流量控制和基于速率的流量控制。

对启动流量控制策略的选择在本标准范围之外。

本标准定义了用于流量控制的机制，何时启动流量控制，启动多少时间在本标准范围之外。

## 9 性能指标要求

### 9.1 端口数量

设备拥有的端口的数量。要求 1000 Mbit/s 端口数量应当 $\geq 1$ 。

### 9.2 设备吞吐量

设备吞吐量指设备所有端口同时收发数据速率能力的总和。本标准建议吞吐量= $\Sigma$  端口速率（半双工），吞吐量= $\Sigma$  端口速率 $\times 2$ （全双工）。

### 9.3 突发长度

突发（burst）：以最小合法帧间隔发送的一组帧。

突发长度（burst size）：一定数量的突发帧。

突发长度可以从 1 到无限。在全双工接口上无论是单向/双向流量，理论上突发长度没有限制。在半双工接口上双向流量的突发长度有限，因为接口发送序列可能被接收帧中断。

建议全双工端口上突发长度无限。半双工端口上单向流量的突发长度无限。半双工端口上双向流量突发长度不作规范。

### 9.4 突发间隔

突发之间的时间间隔。

建议突发间隔=最小帧间隔。

### 9.5 过负荷

过负荷：试图使被测设备以超出端口媒体限制的速度传输。

过负荷可以通过缓存或拥塞控制方式实现。本标准建议实现过负荷。

### 9.6 转发速率

在一定负荷下，被测设备可以观察到正确转发帧的速率。

本标准建议设备端口采用线速转发数据帧。

### 9.7 拥塞控制

任何用作为避免帧丢失，请求外部数据源降低发送速率以免拥塞端口的机制。本标准对此不作规范，由生产厂家任选实现。

### 9.8 队头阻塞

队头阻塞是指由于输入端口试图向某一拥塞端口发送数据帧而导致该输入端口上目的地为不拥塞端口帧的丢失或附加时延。

本标准不强制实现避免队头阻塞，但建议生产厂家实现。

### 9.9 地址缓存能力

每个端口/模块/设备上能够缓存的 MAC 地址的能力。由于缓存的 MAC 地址才能使到达的帧不被丢弃或广播。

本标准建议 1000 Mbit/s 端口 MAC 地址缓存能力不低于 2048 个。

### 9.10 地址学习能力

交换机可以学习新的 MAC 地址（不用广播或丢弃数据帧）的速度。

该指标用作衡量网络重启后地址表建立速度。

标准建议 1000 Mbit/s 端口地址学习能力 $> 1000$  个/s。

### 9.11 时延

对于存储转发设备，时延为被测设备收到最后一比特到发出第一比特的时间间隔。对于按比特转发设备，时延为被测设备收到第一比特到发出第一比特的时间间隔。

本标准定义的时延为测试设备发出时戳的测试帧到收到该帧的时间间隔。本标准建议 64 Byte 长的数据帧时延不超过 100  $\mu$ s。

### 9.12 时延抖动

指时延变化。由于交换机时延较小，因此本标准对时延抖动不作规范。

### 9.13 丢包率

丢包率是指交换机因资源不足引起的包丢失率。

本标准建议交换机丢包率 < 0.01% (暂定)。

### 9.14 误码率

指端口满负荷传输时比特差错率。本标准对误码率不作规范。

### 9.15 可靠性

用在核心网时，建议系统的无故障工作时间 MTBF > 17520h。建议系统故障恢复时间 < 1h。建议主要部件热备份冗余。

## 10 运行与维护

### 10.1 定义

下面行为应包含在交换机的 O&M 中：

- 提供设备资源利用率；
- 提供网络接口带宽利用率；
- 诊断交换机的处理器，网络接口，相连的网络，调制解调器的硬件问题；
- 安装新硬件；
- 安装新软件；
- 在崩溃后重新启动或重新引导交换机；
- 配置（重新配置）交换机；
- 发现及诊断互联网问题例如拥塞，环路，错误 MAC 地址等错误行为；
- 改变网络拓扑，暂时（例如绕过有问题的通信链路）或者永久；
- 监视交换机及相连网络的状态及性能；
- 为网络设计收集流量统计；
- 在恰当的厂商及电信规范中协调上述行为。

交换机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。由于交换机可能与 NOC 连接在不同网上，交换机支持 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 要求交换机可选支持通过一条备用途径，通常是接在交换机配置口上的调制解调器来实现网络管理。

### 10.2 交换机初始化

#### 10.2.1 最少交换机配置

交换机应当能在不配置任何参数条件下正确转发数据帧。

#### 10.2.2 地址及前缀初始化

交换机可以允许静态配置 IP 地址、地址掩码或前缀长度，并存储在非—不稳定存储器中，用作管理。

如上文所述，交换机的 IP 地址中主机地址部分和网络前缀部分不允许是 0 或—1。所以交换机应当不允许将 IP 地址设置成上述形式。

交换机应当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。

### 10.3 运行和维护具体规定

#### 10.3.1 定义

在交换机上实施 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，接在交换机上的终端）；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 人员可以登录到交换机上作为一个主机，使用 Telenet 协议执行本地也能执行的功能。仅在本地模型一般在交换机安装时使用，交换机通常需要由 NOC 远端操作，所以交换机应实现远端操作。

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交换机使用恰当的协议。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它 NOC 软件应像在标准操作系统中编程一样。

交换机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交换机资源的有效控制，例如交换机监视时必须不过分占用 CPU 资源；另一方面，O&M 功能必须具有相对高的优先级，因为交换机堵塞的时候通常是最需要 O&M 操作的时候。

#### 10.3.2 带外访问

交换机应当提供带外（Out-Of-Band OOB）访问。OOB 访问应当提供所有带内访问的功能。带外访问应当实现访问控制，防止非授权访问。

#### 10.3.3 交换机 O&M 功能

##### 1) 维护—硬件诊断

在本地硬件维护时，每个交换机应当像一个独立设备一样被操作，在交换机端应当提供运行诊断程序需要的方法。交换机应当能在出错时运行诊断程序。

##### 2) 控制—配置交换机

每台交换机都可能有需要配置的参数。交换机参数更新后应当不需要重新启动；最坏情况下需要重新运行。可能存在某些情况，改变参数后必须重启交换机（例如改变 IP 地址）。这些情况下，必须小心将对交换机和周边网络的影响减少到最小。

##### 3) 对错误配置的检查与反应

必须实现一种机制检查错误配置并做出反应。如果命令不正确运行，交换机应当给出错误消息。交换机不应接受错误格式的命令，即使该命令本身是正确的。

另一种错误是对交换机连接网络的错误配置。交换机可以实现检测网络的误配置。交换机可以将发现的错误记录到日志或者网络上其他交换机或主机，管理员能看到可能存在的问题。

### 10.4 安全性考虑

#### 10.4.1 审计与审计记录

##### 1) 配置改变

交换机应当提供一种方法来记录配置的改变，指示记录操作人员改变配置的时间。

##### 2) 安全性审计

交换机必须提供一种机制审计与安全性相关的失败与冲突：

- 授权失败：错误口令，无效的 SNMP 通信，非法的授权令牌；
- 对控制策略的违反：被过滤掉的目的地址；
- 授权通过：正确口令，远程登录带内访问，配置口访问。

交换机必须提供一种机制来限制或禁止这样的审计，但缺省情况下审计应当存在。审计可能的方法包括在如果存在的控制口列出冲突，计数或者写入日志，通过 SNMP trap 机制送到远程安全服务器，



或者使用 UNIX 的日志机制。交换机必须实现至少一种上述方法，但可以实现多种方法。

#### 10.4.2 配置控制

在为交换机生产软件/固件时，厂商应负责良好的配置控制。

如果厂商提供用户远程改变交换机配置的能力，例如通过远程登录；这种能力应当是可配置的，缺省情况应当是不允许远程配置。在允许远程配置前，交换机应当要求有效的授权。这种授权不应当在网络上传输明文。例如：如果实现远程登录，厂商应当实现 Kerberos, S-Key, 或者其他类似授权机制。

交换机不允许存在未记载于文档的访问后门或通用密码。厂商必须确保这样的用于调试或者开发的访问途径在产品分销到客户手中之前已删除。

### 11 网络管理协议

#### 11.1 简单网络管理协议-SNMP

##### 11.1.1 SNMP 协议元素

交换机必须支持 RFC1902 至 RFC1906 中规定的 SNMP v2。

SNMP 必须使用 UDP/IP 作为传输层/网络层协议。也可以使用其它协议（例如，RFC1418, RFC1089）。

SNMP 管理请求向交换机任何一个接口发出时，该操作必须生效。实际的管理动作应由交换机或交换机的代理完成。

支持 SNMP v2 协议的交换机必须实现 SNMP v2 MIB RFC1907。

交换机必须实现所有的 SNMP 操作。

交换机必须提供一种机制来限制 SNMP 陷阱 (trap) 消息的产生速率。交换机可以通过 RFC1224 中描述的异步告警管理算法来实现上述机制。

##### 11.2 团体表格

为本标准描述方便，假设交换机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。

交换机的团体表格必须至少包含一个条目，建议至少包含两个条目。

交换机必须允许用户手工（即不使用 SNMP）检查、增、删、改 SNMP 团体表格中的条目。用户必须能够设置区域名，或者构造 MIB 视图。用户必须能以只读（即不允许 SET）或者读写（允许 SET）的方式配置区域。

用户必须能定义至少一个 IP 地址，当使用 trap 时，对每个区域或 MIB 视图的通知将送到该 IP 地址。这些 IP 地址应当被定义在区域或 MIB 视图库内。允许或不允许在区域或 MIB 视图库上发通知应当是可配置的。

交换机应当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，交换机必须验证 SNMP 数据报源地址的有效性；如果该地址没有在上述列表中出现则必须丢弃该数据报。如果数据报被丢弃，交换机必须作 SNMP 授权失败时相应动作。

团体表必须存储在非-不稳定的存储器内。

团体表的初始状态应当包含一个条目，其中区域名串为 Public，访问权限为只读。该条目的缺省状态不允许发送 trap。如果实现，该条目必须保存在团体表中，直到管理员改变或者删除。

##### 11.3 标准 MIBS

所有关于交换机配置的 MIB 都应实现：

- MIB-II RFC1213 中的系统、接口、IP、ICMP 和 UDP 组必须实现。
- 接口扩展 MIB RFC1229 必须实现。
- 如果交换机实现 TCP（例如，远程登录），MIB-II，RFC1213 中的 TCP 组必须实现。

- 以太网-链路 MIB RFC1643 必须实现。
- 网桥 MIB RFC1493 必须实现。
- 远程网络监视 MIB RFC1757 必须实现。
- 远程网络监视 MIB 对交换网络的扩展 V1, RFC2613 建议支持。

#### 11.4 厂商指定的 MIBS

互联网标准和根据实验的 MIB 不能完全覆盖网络元素统计、状态、配置和控制信息。交换机（或其他网络设备）的厂商通常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。

交换机上厂商特定的 MIB 必须提供一种方法来存取所有实现的统计、状态、配置和控制信息，这些信息不能由标准或实验得到的 MIB 得到。这些信息必须能被控制和监视操作使用。

厂商应当使所有厂商特定的 MIB 变量可用。这些指定必须符合 RFC1155，并且必须以 RFC1212 指定的方式描述。

## 12 环境要求

### 12.1 环境要求

#### 12.1.1 交换机正常工作的温度、湿度条件

- 长期工作条件：温度保持 15~30℃、相对湿度保持 40%~65%
- 短期工作条件：温度保持 0~40℃、相对湿度保持 20%~90%

注：

- 1) 交换机的正常工作温度和相对湿度的测量点指在地板以上 2 m 和交换机前方 0.4 m 处测量值。
- 2) 短期工作条件是指连续不超过 48 h 和每年累计不超过 15 d。
- 3) 相对湿度低于 20% 的环境应采用防静电地面。

#### 12.1.2 交换机的防尘要求

机房内灰尘粒子应是非导电，导磁和腐蚀性的。

### 12.2 防电磁干扰要求

交换机产生的电磁干扰应满足以下要求：

- 由交换机射出的无线电电磁干扰应符合表 5 的规定。
- 由交换机进入交流馈电线的无线电电磁干扰应符合表 6 的规定。
- 由交换机进入直流馈线和信号线的无线电电磁干扰应符合表 7 的规定。

表 5 交换机射出的无线电电磁干扰限值

频率, MHz	电磁强度, dB( $\mu$ V/m)	频率, MHz	电磁强度, dB ( $\mu$ V/m)
0.01~0.024	148.6-60lgd	47.7/d-88	59.1-20lgd
0.024~0.8	116.2-60lgd-20lgf	88~216	63.6-20lgd
0.8~1.59	118.2-60lgd	2160~10000	66.6-20lgd
1.59~47.7/d	120.2-60lgd-40lgf		

注：

- 1  $d$  为测试天线与靠近被测物间水平距离;单位为 m,  $d$  限于 30 m 内。
- 2  $f$  为频率, 以 MHz 为单位。
- 3 dB ( $\mu$ V/m) 表示以微伏 ( $\mu$ V/m) 为参考单元的分贝数。

表 6 进入交流馈电线的无线电电磁干扰限值

频率, MHz	最大线路电流, dB $\mu$ 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

注:

- 1  $f$  为频率, 以 MHz 为单位。
- 2  $I$  为接入到交流电源处的输入线路电流电平。
- 3 dB ( $\mu$ V/m) 表示以微伏 ( $\mu$ V/m) 为参考单元的分贝数。

表 7 进入直流馈线和信号线的无线电电磁干扰限值

频率, MHz	最大线路电流, dB $\mu$ A
0.01~0.8	$-21.05\lg f+57.9$
0.8~100	60

### 12.3 交换机抗电磁干扰的能力

交换机在受到 0.01~1000 MHz 频率范围内电场强度为 140 dB $\mu$ V/m 的外界电磁干扰时应不出现故障和性能下降。

在直流或交流电源线受到表 8 所示的 0.01~100 MHz 频率范围的外界电磁干扰电流时应不出现故障和性能下降。

表 8 进入电源线的电磁干扰限值

频率, MHz	最大线路电流, dB $\mu$ A
0.01~0.8	$-21.05\lg f+67.9$
0.8~100	70

### 12.4 交换机防雷击能力

交换机设备防雷击能力应当符合 GB3483-1983 《电子设备雷击实验》。

## 13 电源与接地

### 13.1 电源

#### a) 直流电压及其波动范围要求

额定电压: 为 -48 V 的直流电源。

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

#### b) 杂音电压指标

在直流配电盘输出端子处测量的限值如下:

300~3400 Hz 杂音电压  $\leq 2$  mV;

0~300 Hz 峰峰值杂音电压  $\leq 400$  mV;

3.4~15 kHz 宽带杂音电压  $\leq 100$  mV 有效值;

150 kHz~30 MHz 宽带杂音电压 $\leq 30$  mV 有效值。

c) 离散频率杂音电压

3.4~15 kHz,  $\leq 5$  mV 有效值;

150~200 kHz,  $\leq 3$  mV 有效值;

200~500 kHz,  $\leq 2$  mV 有效值;

500 kHz~2 MHz,  $\leq 1$  mV 有效值。

d) 交流电压及其波动范围要求

单相 220 V $\pm 10\%$ , 频率 50 Hz $\pm 5\%$ 。

线电压波形畸变率 $> 5\%$ 。

13.2 交换机接地要求

a) 接地方式应符合工作地、保护地和建筑防雷接地公用一组接地体的联合接地方式。

b) 接地线截面积

接地线截面积根据可能通过的最大电流负荷确定。应采用良导体导线, 不能使用裸导线布放。

接地电阻值: 联合接地的电阻值应 $< 5 \Omega$ 。

---