运行 CatOS 配置和管理的 Catalyst 4500/4000、5500/5000 和 6500/6000 系列交换机的最佳实践

目录

简介

本文档介绍 Cisco Catalyst 系列交换机（特别是 Catalyst 4500/4000、5500/5000 和 6500/6000 平台）在网络中的实施。介绍配置和命令时假定您运行的是 Catalyst OS (CatOS) General Deployment 软件 6.4(3) 或更高版本。尽管提供了一些设计注意事项，但本文档并不包括总体园区设计。

先决条件

要求

本文档假定读者熟悉 Catalyst 6500 系列命令参考 7.6。

除引用公共联机材料供您进一步阅读外，本文档还提供了下面一些基础和教育参考内容：

- Cisco ISP 基本要素 — 每个 ISP 应考察的基本 IOS 功能。
- Cisco 网络监控和事件相关指南
- 千兆园区网络设计 — 原理和体系结构
- Cisco SAFE：企业网络的安全蓝图

使用的组件

本文档不限于特定的软件和硬件版本。

规则

有关文档规则的详细信息，请参阅 Cisco 技术提示规则。

背景信息

这里提供的解决方案体现了 Cisco 工程师面对众多大型客户以及复杂网络的多年现场工作经验。因此，本文档重点介绍实现网络成功运行的实际配置。本文档提供以下解决方案：

- 在统计上覆盖面最广因而风险最低的解决方案。
- 通过牺牲一定程度的灵活性换取确定性结果的简单解决方案。
- 由网络运营团队配置的易于管理的解决方案。
- 促进高可用性和高稳定性的解决方案。

本文档分为以下四个部分：

- 基本配置 — 大部分网络使用的功能，如生成树协议 (STP) 和中继。
- 管理配置 — 设计注意事项以及使用简单网络管理协议 (SNMP)、远程监控 (RMON)、Syslog、Cisco 发现协议 (CDP) 和网络时间协议 (NTP) 的系统和事件监控。
- 安全配置 — 密码、端口安全、物理安全和使用 TACACS+ 的身份验证。
- 配置清单 — 建议配置模板的汇总信息。

基本配置

本部分讨论大多数 Catalyst 网络中部署的功能。

Catalyst 控制层面协议

本部分介绍正常操作情况下在交换机之间运行的协议。掌握这些协议的基础知识有助于理解各个部分。

Supervisor 数据流

Catalyst 网络中启用的大多数功能都需要两台或多台交换机进行协作，因此必须控制 keepalive 消息、配置参数和管理更改的交换。无论这些协议是 Cisco 专有的（如 CDP）还是基于标准的（如 IEEE 802.1d (STP)），在 Catalyst 系列中实施时都具有一些共同的元素。

在基本帧转发中，源自终端系统的用户数据帧及其源地址和目标地址在第 2 层 (L2) 交换域中不会更改。每个交换机 Supervisor 引擎上的内容可寻址存储器 (CAM) 查询表通过源地址识别过程进行填写，并指定哪个出口端口必须转发接收的每个帧。如果地址识别过程不完整（目标未知或帧发往广播或多播地址），则会在该 VLAN 中的所有端口向外转发（泛洪）该过程。

交换机还必须识别哪些帧将通过系统进行交换，以及哪些帧必须定向到交换机 CPU 本身（也称为网络管理处理器 [NMP]）。

Catalyst 控制层面是使用 CAM 表中的特殊条目（称为系统**条目**）创建的，用于接收数据流并将其定向到内部交换机端口上的 NMP。因此，通过将协议与熟知的目标 MAC 地址结合使用，可以将控制层面流与数据流分离。在交换机上发出 show CAM system 命令以确认这一点，如下所示：

```
>show cam system

* = Static Entry. + = Permanent Entry. # = System Entry. R = Router Entry.
X = Port Security Entry
VLAN  Dest MAC/Route Des    [CoS]  Destination Ports or VCs / [Protocol Type]
----  -----------------    -----  ------------------------------------------
1     00-d0-ff-88-cb-ff  #          1/3
!--- NMP internal port. 1 01-00-0c-cc-cc-cc # 1/3 !--- CDP and so on. 1 01-00-0c-cc-cc-cd # 1/3
!--- Cisco STP. 1 01-80-c2-00-00-00 # 1/3 !--- IEEE STP. 1 01-80-c2-00-00-01 # 1/3 !--- IEEE
flow control. 1 00-03-6b-51-e1-82 R# 15/1 !--- Multilayer Switch Feature Card (MSFC) router. ...
```

Cisco 具有保留的以太网 MAC 和协议地址范围，如下所示。本文档随后将逐一对其进行介绍。但为方便起见，下表提供了汇总信息。

功能	SNAP HDLC 协议类型	目标多播 MAC
端口聚合协议 (PAgP)	0x0104	01-00-0c-cc-cc-cc
生成树 PVSTP+	0x010b	01-00-0c-cc-cc-cd
VLAN 网桥	0x010c	01-00-0c-cd-cd-ce
单向链路检测 (UDLD)	0x0111	01-00-0c-cc-cc-cc
Cisco 发现协议	0x2000	01-00-0c-cc-cc-cc
动态中继 (DTP)	0x2004	01-00-0c-cc-cc-cc
STP Uplink Fast	0x200a	01-00-0c-cd-cd-cd
IEEE 生成树	N/A -	01-80-c2-00-00-00

802.1d	DSAP 42 SSAP 42	
交换机间链路 (ISL)	不适用	01-00-0c-00-00-00
VLAN 中继 (VTP)	0x2003	01-00-0c-cc-cc-cc
IEEE 暂停 802.3x	N/A - DSAP 81 SSAP 80	01-80-C2-00-00-00>0F

大多数 Cisco 控制协议采用 IEEE 802.3 SNAP 封装（包括 **LLC 0xAAAA03、OUI 0x00000C**），这在 LAN 分析程序跟踪文件中可以看到。这些协议的其他共有属性包括：

- 这些协议假定具有点对点连接。注意，有意使用多播目标地址将使两个 Catalyst 能够通过非 Cisco 交换机进行透明通信，因为无法识别和拦截帧的设备将只对其进行泛洪。但是，通过多供应商环境的单点对多点连接可能会产生不一致的行为，通常必须避免这种情况。
- 这些协议在第 3 层 (L3) 路由器终止，它们仅在一个交换机域内起作用。
- 这些协议通过入口专用集成电路 (ASIC) 处理和计划接收高于用户数据的优先级。

为保持内容完整，介绍了控制协议目标地址后，还必须介绍源地址。交换机协议使用从机箱上 EPROM 提供的可用地址组中获取的一个 MAC 地址。**发出 show module 命令以显示每个模块在发送数据流（如 STP 网桥协议数据单元 (BPDU) 或 ISL 帧）时可以使用的地址范围。**

```
>show module

...
Mod MAC-Address(es)                         Hw     Fw          Sw
--- --------------------------------------- ------ ----------  -----------------
1   00-01-c9-da-0c-1e to 00-01-c9-da-0c-1f 2.2    6.1(3)      6.1(1d)
    00-01-c9-da-0c-1c to 00-01-c9-da-0c-1
    00-d0-ff-88-c8-00 to 00-d0-ff-88-cb-ff
!--- MACs for sourcing traffic. ... VLAN 1
```

VLAN1

VLAN 1 在 Catalyst 网络中具有特殊意义。

Catalyst Supervisor 引擎在建立中继时，总是使用默认的 VLAN（即 VLAN 1）来标记许多控制和管理协议（如 CDP、VTP 和 PAgP）。包括内部 sc0 接口在内的所有端口均默认配置为 VLAN 1 的成员。所有中继均默认承载 VLAN 1，在低于 5.4 的 CatOS 软件版本中，无法阻止 VLAN 1 中的用户数据。

为帮助明确 Catalyst 网络连接中常用的一些术语，需要了解以下定义：

- 管理 VLAN 是 sc0 驻留的位置，可以更改此 VLAN。
- 本地 VLAN 定义为没有中继时端口返回的 VLAN，也是 802.1Q 中继上的未标记 VLAN。默认情况下，VLAN 1 为本地 VLAN。
- 要更改本地 VLAN，请发出 **set vlan** *vlan-id mod /port* 命令。**注意：** 必须先创建 VLAN，才能将其设置为中继的本地 VLAN。

下面是几个有力的理由，说明为何要调整网络并更改 VLAN 1 中端口的行为：

- 与所有其他 VLAN 一样，当 VLAN 1 的范围大得足以影响稳定性时（特别是从 STP 角度而言），便需要对其进行修剪。本文档的带内管理部分对此进行详细讨论。

- VLAN 1 的控制层面数据必须与用户数据分开，以简化故障排除并最大限度地使用 CPU 周期。
- 在没有 STP 的情况下设计多层园区网络时，必须避免 VLAN 1 中出现 L2 环路，如果存在多个 VLAN 和 IP 子网，接入层仍然需要中继。为此，请手动将 VLAN 1 从中继端口中清除。

总之，请注意有关中继的以下信息：

- **CDP、VTP 和 PAgP** 更新总是在带有 VLAN 1 标记的中继上转发。即使 VLAN 1 已从中继中清除且不是本地 VLAN，也是如此。如果 VLAN 1 的用户数据已清除，则对仍然使用 VLAN 1 发送的控制层面数据流没有任何影响。
- 在 ISL 中继上，DTP 数据包在 VLAN 1 中发送。即使 VLAN 1 已从中继中清除且不再是本地 VLAN，也是如此。在 802.1Q 中继上，DTP 数据包在本地 VLAN 中发送。即使本地 VLAN 已从中继中清除，也是如此。
- 在 PVST+ 中，除非 VLAN 1 已从中继中清除，否则会在通用生成树 VLAN 1 上以无标记形式转发 **802.1Q IEEE BPDU**，以便与其他供应商进行互操作。无论本地 VLAN 配置如何，都是如此。对于所有其他 VLAN，将发送 **Cisco PVST+ BPDU** 并对其进行标记。有关更多信息，请参阅本文档中的生成树协议部分。
- 802.1s 多生成树 (MST) BPDU 始终在 ISL 和 802.1Q 中继上的 VLAN 1 中发送。即使 VLAN 1 已从中继中清除，也是如此。
- 不要在 MST 网桥和 PVST+ 网桥之间的中继上清除或禁用 VLAN 1。但在已禁用 VLAN 1 的情况下，MST 网桥必须成为根，这样所有 VLAN 才能避免 MST 网桥的边界端口陷入根不一致状态。有关详细信息，请参阅了解多生成树协议 (802.1s)。

建议

为使 VLAN 在未连接客户端或主机的情况下保持**打开/打开状态**，需要在该 VLAN 中至少连接一台物理设备。否则，VLAN 将处于**打开/关闭状态**。目前，当交换机中没有用于该 VLAN 的活动端口时，没有用于将 VLAN 接口设置为**打开/打开**的命令。

如果不想连接设备，可以在任意端口为该 VLAN 连接一个回环插件。或者，也可以尝试使用交叉电缆连接该 VLAN 在同一交换机上的两个端口。此方法将强制端口打开。有关详细信息，请参阅 T1/56K 线路的回环测试的回环插件部分。

如果网络对于服务提供商是多宿主网络，该网络将充当两个服务提供商之间的中转网络。从一个服务提供商传递给另一个服务提供商时，如果数据包中接收的 VLAN 编号需要转换或更改，可以使用 QinQ 功能转换该 VLAN 编号。

VLAN 中继协议

在创建 VLAN 之前，请确定要在网络中使用的 VTP 模式。使用 VTP，可以在一台或多台交换机上集中进行 VLAN 配置更改。这些更改将自动传播到域中的所有其他交换机。

操作概述

VTP 是用于保持 VLAN 配置一致性的 L2 消息传递协议。VTP 按网络范围管理 VLAN 的添加、删除和重命名。VTP 能够将可能引起大量问题的错误配置和配置不一致问题减至最少，其中包括重复的 VLAN 名称、不正确的 VLAN 类型规格和安全违规。VLAN 数据库是一个二进制文件，与配置文件分开存储在 VTP 服务器的 NVRAM 中。

VTP 协议使用以太网目标多播 MAC 地址 (**01-00-0c-cc-cc-cc**) 和 SNAP HDLC 协议类型 Ox2003 在交换机之间进行通信。它不通过非中继端口工作（VTP 是 ISL 或 802.1Q 的有效负载），因此在

DTP 使中继联机之前，无法发送消息。

消息类型包括每五分钟发送一次的汇总通告、发生更改时的子集通告和请求通告，以及启用 VTP 修剪时的加入信息。服务器上每发生一次更改，VTP 配置修订号就加 1，随后在整个域中传播新表。

如果一个 VLAN 被删除，曾是该 VLAN 成员的端口将处于非活动状态。同样，如果处于客户端模式的交换机无法在启动时（从 VTP 服务器或另一个 VTP 客户端）接收 VTP VLAN 表，则会停用 VLAN 中除默认 VLAN 1 外的所有端口。

下表提供了各种 VTP 模式的功能比较汇总：

功能	服务器	客户端	透明	1
源 VTP 消息	是	是	否	否
监听 VTP 消息	是	是	否	否
转发 VTP 消息	是	是	是	否
创建 VLAN	是	否	是（仅在本地有意义）	是（仅在本地有意义）
记住 VLAN	是	否	是（仅在本地有意义）	是（仅在本地有意义）

在 VTP transparent 模式下，会忽略 VTP 更新（VTP 多播 MAC 地址会从通常用于拾取控制帧并将其定向到 Supervisor 引擎的 CAM 系统中删除）。因为协议使用了多播地址，transparent 模式中的交换机（或另一家供应商的交换机）会将帧泛洪到域中的其他 Cisco 交换机。

[1个]CatOS软件版本7.1引入选项禁用与使用的VTP `off`。在 VTP off 模式下，交换机的行为方式与 VTP transparent 模式非常类似，但 off 模式还会抑制转发 VTP 更新。

下表对初始配置进行概述：

功能	默认值
VTP DOMAIN NAME	空
VTP 模式	服务器
VTP 版本	版本 1 已启用
VTP 密码	无
VTP 修剪	已禁用

VTP 版本 2 (VTPv2) 具有以下功能上的灵活性。但是，它不能与 VTP 版本 1 (VTPv1) 进行互操作：

- 令牌环支持
- 未识别的 VTP 信息支持；交换机现在可传播其不能解析的值。

- 与版本相关的 transparent 模式；transparent 模式不再检查域名。这会启用在一个透明域中对多个域的支持。
- 版本号传播；如果所有交换机均支持 VTPv2，则可通过单个交换机的配置启用所有交换机。

有关详细信息，请参阅了解和配置 VLAN 中继协议 (VTP)。

VTP 版本 3

CatOS 软件版本 8.1 引入了对 VTP 版本 3 (VTPv3) 的支持。与现有版本相比，VTPv3 的功能有所增强。这些增强功能允许：

- 支持扩展的 VLAN
- 支持专用 VLAN 的创建和通告
- 支持 VLAN 实例和 MST 映射传播实例（在 CatOS 版本 8.3 中受支持）
- 改进后的服务器身份验证
- 防止将“错误的”数据库意外插入 VTP 域
- 与 VTPv1 和 VTPv2 进行交互
- 能够基于每个端口进行配置

VTPv3 实现与早期版本之间的一个主要差别是引入了 VTP 主服务器。理想情况下，如果 VTPv3 域未分区，则该域中必须仅有一个主服务器。对 VTP 域进行的所有更改只有在 VTP 主服务器上执行，才能传播到 VTP 域。VTPv3 域中可以有多个服务器，这些服务器也称为辅助服务器。将交换机配置为服务器时，默认情况下交换机会成为辅助服务器。辅助服务器可以存储域的配置，但不能修改该配置。辅助服务器通过从相应交换机进行成功接管后便可以成为主服务器。

运行 VTPv3 的交换机仅接受修订版号高于当前主服务器的 VTP 数据库。此过程与 VTPv1 和 VTPv2 明显不同，在 VTPv1 和 VTPv2 中交换机始终接受来自同一域的邻居的高级配置。VTPv3 的此更改可提供保护功能。网络中引入的具有较高 VTP 修订版号的新交换机不能覆盖整个域的 VLAN 配置。

VTPv3 还引入了针对 VTP 处理口令方式的增强功能。如果您使用隐藏的口令配置选项以将口令配置为“隐藏”，则会出现以下情形：

- 口令不会以明文形式显示在配置中。口令会以加密十六进制的格式保存在配置中。
- 如果您尝试将相应交换机配置为主服务器，则系统会提示您输入口令。如果您的口令与加密口令匹配，则该交换机会成为主服务器，这样您便可对域进行配置。

注意： 务必注意只有在您需要修改任何实例的 VTP 配置时才需要主服务器。VTP 域可在没有活动主服务器的情况下运行，这是因为辅助服务器可确保重新加载后配置始终存在。会因为以下原因而导致退出主服务器状态：

- 交换机重新加载
- 在活动 Supervisor 引擎和冗余 Supervisor 引擎之间进行高可用性切换
- 从另一个服务器进行接管
- 模式配置发生更改
- 任何 VTP 域配置更改，例如在以下方面发生更改：**version**域名域密码

VTPv3 还允许交换机参与 VTP 的多个实例。在这种情况下，同一个交换机可以是一个实例的 VTP 服务器和另一个实例的客户端，这是因为 VTP 模式特定于不同的 VTP 实例。例如，某台交换机可以针对 MST 实例在 transparent 模式下运行，还可针对 VLAN 实例配置为 server 模式。

在与 VTPv1 和 VTPv2 进行交互方面，所有版本 VTP 的默认行为是较早版本的 VTP 直接丢弃新版本的更新。除非 VTPv1 和 VTPv2 交换机处于 transparent 模式，否则会丢弃所有 VTPv3 更新。另一方面，VTPv3 交换机在中继上接收旧 VTPv1 或 VTPv2 帧后，VTPv3 交换机会将其精简版的数

据库更新传递给 VTPv1 和 VTPv2 交换机。但是，这种信息交换是单向的，因为 VTPv3 交换机不接受来自 VTPv1 和 VTPv2 交换机的任何更新。在中继连接上，VTPv3 交换机会持续发送精简版的更新以及完全版的 VTPv3 更新，以此满足跨中继端口的 VTPv2 和 VTPv3 邻居的需求。

为了针对扩展的 VLAN 提供 VTPv3 支持，VLAN 数据库的格式（即 VTP 为每个 VLAN 分配 70 字节）也进行了更改。此更改仅允许对非默认值进行编码，而不允许传送旧协议未修改的字段。由于此更改，4K VLAN 支持为生成的 VLAN 数据库的大小。

建议

对于使用 VTP client/server 模式，还是使用 VTP transparent 模式，没有具体的建议。尽管存在稍后说明的一些注意事项，部分客户仍倾向于使用易于管理的 VTP client/server 模式。建议在每个域中放置两个 server 模式交换机以实现冗余，一般为两个分布层交换机。域中其余的交换机必须设置为 client 模式。在您使用 VTPv2 实现 client/server 模式时，请注意在同一 VTP 域中始终接受较高的修订版号。如果将一台配置为 VTP client 或 server 模式的交换机引入 VTP 域，且该交换机的版本号高于现有的 VTP 服务器，这将覆盖 VTP 域中的 VLAN 数据库。**如果无意中进行了配置更改且 VLAN 已删除，则这种覆盖可能会导致网络发生严重中断。**为了确保 client 或 server 交换机始终具有低于服务器的配置修订版号，请将客户端 VTP 域名更改为不同于标准名称的其他名称。然后恢复为标准名称。此操作会将客户端上的配置修订版号设置为 0。

可以方便地在网络上更改 VTP，这一点有利也有弊。由于以下原因，很多企业倾向于使用 VTP transparent 模式这一较为谨慎的方法：

- 该模式有助于养成良好的更改控制习惯，因为必须以逐个交换机的形式考虑在交换机或中继端口上修改 VLAN 的要求。
- 该模式可降低发生影响整个域的管理员错误（例如意外删除 VLAN）的风险。
- 不存在因为在网络中引入具有较高 VTP 修订版号的新交换机而导致整个域的 VLAN 配置被覆盖的任何风险。
- 该模式有助于将 VLAN 从正在运行的中继修剪到在该 VLAN 中没有端口的交换机。这使帧泛洪能够更为高效地利用带宽。手工修剪也是很有益的，因为这样可以减小生成树直径（请参阅本文档中的 DTP 部分）。在端口信道中继上修剪未使用的 VLAN 前，请确保连接到 IP 电话的所有端口均已配置为语音 VLAN 的接入端口。
- CatOS 6.x 和 CatOS 7.x 中编号从 1025 到 4094 的扩展 VLAN 范围只能使用上述方式配置。有关详细信息，请参阅本文档中的扩展的 VLAN 和 MAC 地址缩减部分。
- 包含在 Cisco Works 2000 中的 Campus Manager 3.1 支持 VTP transparent 模式。在 VTP 域中至少需要一个服务器的旧限制已取消。

VTP 命令示例	备注
set vtp domain name pa	CDP 检查名称以帮助检查域之间是否有布线错误。简单的口令也是一项防止无意更改的有用预防措施。如果进行粘贴，请注意区分名称的大小写或空格。

password x	
set vtp mode transparent	
set vlan vlan number name name	每个在 VLAN 中具有端口的交换机。
set trunk mod/port vlan range	启用中继以根据需要传送 VLAN - 默认为所有 VLAN。
clear trunk mod/port vlan range	通过手动修剪（例如在从分布层到接入层的中继上进行修剪，该中继中不存在 VLAN）限制 STP 直径。

注意： 使用 **set** 命令指定 VLAN 只会添加 VLAN，而不会清除 VLAN。**例如，set trunk x/y 1-10 命令不会将允许的列表设置为仅 VLAN 1-10。需发出 clear trunk x/y 11-1005 命令才能实现所要的结果。**

尽管令牌环交换在本文档的讨论范围之外，但请注意对于 TR-ISL 网络不建议使用 VTP transparent 模式。令牌环交换的基础是整个域形成单个分布式多端口网桥，因此每个交换机必须具有相同的

VLAN 信息。

其他选项

VTPv2 是实现令牌环环境的一个要求，在该环境中强烈建议使用 client/server 模式。

VTPv3 能够实现更为严格的身份验证和配置修订控制。基本上而言，VTPv3 提供与 VTPv1/VTPv2 transparent 模式级别相同的功能，但是具有更高的安全性。此外，VTPv3 与旧 VTP 版本部分兼容。

在本文档中积极介绍了修剪 VLAN 以减少不必要的帧泛洪的优点。set vtp pruning enable命令修剪 VLAN自动地，终止帧的无效溢流他们不是需要的。不同于手动 VLAN 修剪，自动修剪不限制生成树直径。

从 CatOS 5.1 开始，Catalyst 交换机可以将大于 1000 的 802.1Q VLAN 编号映射到 ISL VLAN 编号。在 CatOS 6.x 中，根据 IEEE 802.1Q 标准 Catalyst 6500/6000 交换机支持 4096 个 VLAN。这些 VLAN 被组织为以下三个范围，只有部分 VLAN 会通过 VTP 传播到网络中的其他交换机：

- 普通范围的 VLAN：1 – 1001
- 扩展范围的 VLAN：1025 – 4094（只能由 VTPv3 传播）
- 保留范围的 VLAN：0， 1002 — 1024， 4095

IEEE 生成了一个基于标准的体系结构，从而实现与 VTP 类似的结果。作为 802.1Q 通用属性注册协议 (GARP) 的成员，通用 VLAN 注册协议 (GVRP) 可以在供应商之间实现 VLAN 管理互操作，但本文档未对此协议展开讨论。

注意： CatOS 7.x 引入了将 VTP 设置为 off 模式（一种与 transparent 模式非常类似的模式）的选项。但是，交换机不转发 VTP 帧。对于某些设计，这一点在中继到管理控制范围外的交换机时可能会很有用。

扩展的 VLAN 和 MAC 地址缩减

MAC 地址缩减功能可启用扩展范围的 VLAN 标识。启用 MAC 地址缩减会禁用用于 VLAN 生成树的 MAC 地址池并且会保留单个 MAC 地址。此 MAC 地址可标识交换机。CatOS 软件版本 6.1(1) 引入了对 Catalyst 6500/6000 和 Catalyst 4500/4000 交换机的 MAC 地址缩减支持，从而按照 IEEE 802.1Q 标准支持 4096 个 VLAN。

操作概述

交换机协议使用从机箱上的 EPROM 提供的可用地址组中获取的 MAC 地址，作为在 PVST+ 下运行的 VLAN 的网桥标识符的一部分。Catalyst 6500/6000 和 Catalyst 4500/4000 交换机支持 1024 个或 64 个 MAC 地址，具体取决于机箱类型。

在默认情况下具有 1024 个 MAC 地址的 Catalyst 交换机不启用 MAC 地址缩减。MAC 地址会按顺序进行分配。范围中的第一个 MAC 地址会分配到 VLAN 1。范围中的第二个 MAC 地址会分配到 VLAN 2，以此类推。这使交换机能够支持 1024 个 VLAN，且每个 VLAN 使用唯一的网桥标识符。

机箱类型	机箱地址

WS-C4003-S1、WS-C4006-S2	1024
WS-C4503、WS-C4506	64[1]
WS-C6509-E、WS-C6509、WS-C6509-NEB、WS-C6506-E、WS-C6506、WS-C6009、WS-C6006、OSR-7609-AC、OSR-7609-DC	1024
WS-C6513、WS-C6509-NEB-A、WS-C6504-E、WS-C6503-E、WS-C6503、CISCO7603、CISCO7606、CISCO7609、CISCO7613	64[1]

默认情况下[1] MAC地址减少为有64 MAC地址的交换机启用，并且功能不可能禁用。

对于具有 1024 个 MAC 地址的 Catalyst 系列交换机，启用 MAC 地址缩减可支持在 PVST+ 下运行的 4096 个 VLAN 或者 16 个多实例 STP (MISTP) 实例具有唯一标识符，且在交换机上需要的 MAC 地址的数量不会增加。MAC 地址缩减会将 STP 所需的 MAC 地址的数量从一个 VLAN 或一个 MISTP 实例一个缩减到一个交换机一个。

下图显示了未启用网桥标识符 MAC 地址缩减的示例。网桥标识符由一个 2 字节网桥优先级和一个 6 字节 MAC 地址组成：

MAC 地址缩减可修改 BPDU 的 STP 网桥标识符部分。原始的 2 字节优先级字段会拆分成两个字段。此拆分会导致一个 4 字节网桥优先级字段和一个 12 位系统 ID 扩展，而该扩展允许从 0 到 4095 的 VLAN 编号。

当您在 Catalyst 交换机上启用 MAC 地址缩减以便充分利用扩展范围的 VLAN 时，请在同一 STP 域内的所有交换机上启用 MAC 地址缩减。必须采用此步骤，才能使所有交换机上的 STP 根计算保持一致。在您启用 MAC 地址缩减之后，根网桥优先级变为 4096 的倍数与 VLAN ID 之和。因为未执行 MAC 地址缩减的交换机对网桥 ID 的选择更加精细，所以这些交换机可能会无意中声明根。

配置指南

当配置扩展的 VLAN 范围时，必须遵从某些准则。交换机可以出于内部目的从扩展范围中分配 VLAN 块。例如，交换机可以为路由端口或 Flex WAN 模块分配 VLAN。VLAN 块的分配始终从 VLAN 1006 开始并按升序进行分配。如果在该范围内存在 Flex WAN 模块必需的任何 VLAN，则不会对所有这些必需的 VLAN 进行分配，这是因为永远不会从用户 VLAN 区域分配 VLAN。**在交换机上发出 show vlan 命令或 show vlan summary 命令，以便同时显示用户分配的 VLAN 和内部 VLAN。**

```
>show vlan summary

Current Internal Vlan Allocation Policy - Ascending
```

```
Vlan status    Count  Vlans
-------------  -----  ----------------------------------------
VTP Active         7   1,17,174,1002-1005

Internal           7   1006-1011,1016
!--- These are internal VLANs. >show vlan
---- -------------------------------- --------- ------- --------

1    default                          active    7       4/1-48


!--- Output suppressed. 1006 Online Diagnostic Vlan1 active 0 internal 1007 Online Diagnostic
Vlan2 active 0 internal 1008 Online Diagnostic Vlan3 active 0 internal 1009 Voice Internal Vlan
active 0 internal 1010 Dtp Vlan active 0 internal 1011 Private Vlan Internal Vlan suspend 0
internal 1016 Online SP-RP Ping Vlan active 0 internal !--- These are internal VLANs.
```

另外，在您使用扩展范围的 VLAN 之前，必须删除所有现有的 802.1Q 到 ISL 的映射。此外，在 VTPv3 之前的版本中，必须使用 VTP 透明模式在每台交换机上静态配置扩展的 VLAN。有关详细信息，请参阅配置 VLAN 中的*扩展范围的 VLAN 配置指南* 部分。

注意： 在软件版本 8.1(1) 之前的软件中，不能为扩展范围的 VLAN 配置 VLAN 名称。此功能与任何 VTP 版本或模式无关。

建议

尽量在同一个 STP 域内保持一致的 MAC 地址缩减配置。但是，当具有 64 个 MAC 地址的新机箱引入 STP 域时，在所有网络设备上强制实施 MAC 地址缩减是不切实际的。默认情况下，会为具有 64 个 MAC 地址的交换机启用 MAC 地址缩减，且不能禁用该功能。请注意：当两个系统配置有同一生成树优先级时，未执行 MAC 地址缩减的系统有更高的生成树优先级。发出此命令以启用或禁用 MAC 地址缩减：

```
set spantree macreduction enable | disable
```

内部 VLAN 的分配按升序进行，从 VLAN 1006 开始分配。为了避免在用户 VLAN 和内部 VLAN 之间发生冲突，分配用户 VLAN 时尽可能接近 VLAN 4094。对于运行 Cisco IOS® 系统软件的 Catalyst 6500 交换机，可以按降序配置内部 VLAN 分配。CatOS 软件的等效命令行界面 (CLI) 尚未获得正式支持。

自动协商

以太网/快速以太网

自动协商是 IEEE 快速以太网 (FE) 标准 (802.3u) 的可选功能，它使设备能够通过链路自动交换有关**速度和双工**功能的信息。自动协商在第 1 层 (L1) 运行，目标是接入层端口，**临时用户**（如 PC）通过这些端口连接到网络。

操作概述

造成 10/100 Mbps 以太网链路性能问题的最常见原因是：链路上的一个端口在半双工状态运行，而另一个端口在全双工状态运行。如果对链路上的一个或两个端口进行重置，并且自动协商过程未导致两个链路伙伴使用相同配置，则会偶尔发生这种情况。这种情况还会在管理员重新配置链路的一端而忘记重新配置另一端时发生。此的典型症状增加帧校验序列、循环冗余冗余校验(CRC)，校准或者不全计数器在交换机。

在以下这些文档中将详细说明自动协商。这些文档介绍了自动协商的工作原理和配置选项。

- 以太网 10/100Mb 半双工/全双工自动协商的配置和故障排除
- 排除 Cisco Catalyst 交换机的 NIC 兼容性问题

对自动协商的一种常见误解是：可以手动配置 100 Mbps 全双工的一个链路伙伴，并进行自动协商以便与另一个链路伙伴进行全双工通信。实际上，尝试执行此操作会导致双工不匹配。这是因为一个链路伙伴进行自动协商，没有看到另一个链路伙伴的任何自动协商参数，于是默认设置为半双工形式。

大多数 Catalyst 以太网模块支持 10/100 Mbps 和半/全双工，不过将由 show port capabilities mod/port 命令对此进行确认。

FEFI

远端故障指示 (FEFI) 可保护 100BASE-FX（光纤）和千兆接口，而自动协商则可防止 100BASE-TX（铜缆）出现与物理层/信令相关的故障。

远端故障是一个站点可以检测到而另一个站点不能检测到的链路错误，例如断开 TX 线。在本示例中，发送站点仍可以接收有效数据，并通过链路完整性监视器检测到链路正常。它检测不到其传输并没有被另一个站点接收。检测到这样一个远程故障的 100BASE-FX 站点可以修改其传输的 IDLE 流，以发送特殊的位模式（称为 FEFI IDLE 模式），以便将该远程故障通告邻居。FEFI-IDLE 模式随后触发远程端口关闭 (errdisable)。有关故障保护的详细信息，请参阅本文的 UDLD 部分。

下面的硬件和模块支持 FEFI：

- Catalyst 5500/5000：WS-X5201R、WS-X5305、WS-X5236、WS-X5237、WS-U5538 和 WS-U5539
- Catalyst 6500/6000 和 4500/4000：所有 100BASE-FX 模块和 GE 模块

建议

无论是在 10/100 链路上配置自动协商，还是对速度和双工进行硬编码，最终都取决于链路伙伴的类型或您连接到 Catalyst 交换机端口的终端设备类型。终端设备和 Catalyst 交换机之间的自动协商一般都进行得很好，并且 Catalyst 交换机符合 IEEE 802.3u 规范。但是，当 NIC 或供应商交换机不完全符合该规范时，可能会发生问题。另外，由于特定于供应商的高级功能（例如自动极性或布线完整性）未在 IEEE 802.3u 的 10/100 Mbps 自动协商规范中予以说明，也可能导致存在硬件不兼容和其他问题。有关此问题的示例，请参阅 Field Notice：Intel Pro/1000T NIC 与 CAT4K/6K 连接时的性能问题。

可能会遇到一些需要设置主机、端口速度和双工的情况。一般来说，请遵从下列基本的故障排除步骤：

- 确保在链路两端都配置了自动协商，或者在两端都配置了硬编码。
- 查看 CatOS 发行版本注释中的常见问题说明。
- 验证您所运行的 NIC 驱动程序或操作系统的版本，这是因为通常需要最新的驱动程序或修补程序。

通常，首先尽量对任意类型的链路伙伴使用自动协商。为临时设备（如便携式计算机）配置自动协商有明显的好处。理想情况下，自动协商还适用于非临时设备（如服务器和固定的工作站），也适用于交换机到交换机以及交换机到路由器。由于所述的部分原因，可能出现协商问题。在这些情况下，请按照所提供的 TAC 链接中介绍的基本故障排除步骤操作。

如果 10/100 Mbps 以太网端口上的端口速度设置为 auto，将就速度和双工进行自动协商。发出此命

令可将端口设置为 auto 模式：

```
set port speed port range auto
!--- This is the default.
```

如果对端口进行硬编码，请发出以下这些配置命令：

```
set port speed port range 10 | 100 set port duplex port range full | half
```

在 CatOS 8.3 及更高版本中，Cisco 引入了可选的 **auto-10-100** 关键字。请在支持 10/100/1000 Mbps 速度但不需要对 1000 Mbps 使用自动协商的端口上使用 **auto-10-100** 关键字。使用 **auto-10-100** 关键字可以使端口的行为与速度设置为 auto 的 10/100-Mbps 端口相同。将仅对 10/100-Mbps 端口协商速度和双工，1000-Mbps 速度不参与协商。

```
set port speed port_range auto-10-100
```

其他选项

如果交换机之间不使用自动协商，针对某些问题的 L1 故障指示也可能丢失。使用 L2 协议（例如主动 UDLD）增强故障检测是很有帮助的。

千兆以太网

千兆以太网 (GE) 具有一个自动协商过程 (IEEE 802.3z)，该过程比 10/100 Mbps 以太网的自动协商过程更广泛，可用于交换流控制参数、远程故障信息和双工信息（即使 Catalyst 系列千兆以太网端口仅支持全双工模式）。

注意： 802.3z 已被 IEEE 802.3:2000 规范取代。有关详细信息，请参阅 IEEE LAN/MAN 标准在线订阅：归档文件。

操作概述

默认情况下会启用 GE 端口协商，GE 链路两端的端口必须具有相同设置。与 FE 不同，如果链路各端的端口上的自动协商设置不同，则不能建立 GE 链路。但是，若要使禁用自动协商的端口连通链路，所需的唯一条件是从远端发来一个有效的千兆位信号。此行为与远端的自动协商配置无关。例如，假设有两个设备 A 和 B。每个设备均可以启用或禁用自动协商。此表列出了可能的配置和相应的链路状态：

协商	B 启用	B 禁用
A 启用	两端均为 up	A 为 down，B 为 up
A 禁用	A 为 up，B 为 down	两端均为 up

在 GE 中，在通过使用特殊的保留链路代码字序列启动链路时，将执行同步和自动协商（如果它们已启用）。

注意： 提供了一个有效字字典，并非所有可能的字在 GE 中都有效。

可以用以下方式描述 GE 连接的生存期特征：

丢失同步操作意味着 MAC 检测到链路断开。不管是启用还是禁用自动协商，都可能丢失同步操作。在某些出现故障的情况下，例如连续收到三个无效字，将丢失同步操作。如果此情况持续 10 毫秒，则会断定发生了同步失败情况，并且链路更改为 link_down 状态。在丢失同步操作后，需要另外三个连续的有效空闲才能再次执行同步。其他灾难性事件（例如无法收到 (Rx) 信号）会导致链路断开事件。

自动协商是链路连通过程的一部分。当链路连通时，将结束自动协商。但是，交换机仍将监视链路的状态。如果端口上禁用自动协商，则“autoneg”阶段不再是选项。

GE 铜缆规范 (1000BASE-T) 支持通过“下页交换”进行自动协商。“下页交换”允许对铜缆端口上的 10/100/1000 Mbps 速度进行自动协商。

注意： GE 光纤规范只对双工、流控制和远程故障检测的协商做出了相关规定。GE 光纤端口不会对端口速度进行协商。有关自动协商的详细信息，请参阅 IEEE 802.3-2002 规范的第 28 和 37 部分。

同步重启延迟是一项软件功能，用来控制自动协商的总时间。如果自动协商在此时间内不成功，固件会重新启动自动协商，以防出现死锁。**set port sync-restart-delay 命令仅在自动协商设置为 enable 时有效。**

建议

在 GE 环境中启用自动协商比在 10/100 环境中启用自动协商更重要。实际上，只有在与不支持协商的设备连接的交换机端口上，或者因为互操作性问题导致连接问题时，才必须禁用自动协商。Cisco 建议在所有交换机到交换机的链路中以及通常所有 GE 设备中启用千兆协商（默认设置）。发出以下命令可以启用自动协商：

```
set port negotiation port range enable
!--- This is the default.
```

一个已知的例外情况是：当连接到运行 Cisco IOS 软件 12.0(10)S 版（该版本添加了流控制和自动协商功能）之前版本的千兆交换路由器 (GSR) 时。在此情况下，请关闭这两个功能，否则交换机端口报告 not connected，并且 GSR 会报告错误。下面是命令序列示例：

```
set port flowcontrol receive port range off set port flowcontrol send port range off set port
negotiation port range disable
```

必须根据具体情况分别查看交换机到服务器的连接。Cisco 客户在 Sun、HP 和 IBM 服务器上遇到过千兆协商问题。

其他选项

流控制是 802.3x 规范的可选部分；如果使用该功能，则必须对其进行协商。设备也许能够或者不能发送和/或响应 PAUSE 帧（**即大家所熟知的 MAC 01-80-C2-00-00-00 0F**）。而且，它们不能同意远端邻居的流控制请求。如果端口的输入缓冲区被填满，则该端口将向链路伙伴发送一个 PAUSE 帧，链路伙伴将停止传输，并将任何其他帧保存到链路伙伴输出缓冲区中。这不能解决任何持续存

在的超额订阅问题，但是可以在突发传输过程中有效地使输入缓冲区增大，增大幅度为伙伴的输出缓冲区的一部分。

此功能最适用于接入端口和终端主机之间的链路，在这些链路中，主机输出缓冲区与其虚拟内存的大小可能相同。在交换机到交换机的链路中使用此功能受益有限。

发出以下这些命令可在交换机端口上实现流控制：

```
set port flowcontrol mod/port receive | send off |on | desired
```

```
>show port flowcontrol

 Port  Send FlowControl    Receive FlowControl   RxPause TxPause
       admin    oper       admin    oper
-----  -------- --------   -------- --------     ------- -------
 6/1   off      off        on       on           0       0
 6/2   off      off        on       on           0       0
 6/3   off      off        on       on           0       0
```

注意： 如果进行协商，所有 Catalyst 模块都可以响应 PAUSE 帧。由于某些模块（例如 WS-X5410、WS-X4306）是无阻塞模块，因此它们永远不会发送 PAUSE 帧，即使它们协商要执行此操作。

动态中继协议 (DTP)

封装类型

中继通过临时识别和标记（本地链路）原始以太网帧来扩展设备之间的 VLAN，从而使它们能够在单个链路中实现多路复用。这也确保了能在交换机之间维护单独的 VLAN 广播域和安全域。CAM 表用于维护交换机范围内的帧到 VLAN 映射。

尽管此处仅介绍了以太网对中继的支持，但是包括 ATM LANE、FDDI 802.10 和以太网在内的几种 L2 媒体都支持中继功能。

ISL 操作概述

Cisco 专有的识别或标记方案 ISL 已使用多年。此外，还提供了 802.1Q IEEE 标准。

通过在两层标记方案中完全封装原始帧，ISL 实际上是一种隧道协议，并且还具有能够传送非以太网帧的优点。它向标准以太网帧添加了 26 字节的报头和 4 字节的 FCS，这样将生成更大的以太网帧，并由配置为中继的端口进行处理。ISL 支持 1024 个 VLAN。

ISL 帧格式

40 位	4 位	4 位	48 位	16 位	24 位	24 位	15 位	位	16 位	16 位	可变长度	32 位
目的。地址	类型	用户	SA	LEN	SNAP LLC	HAS	VLAN	BPDU	索引	保留	封装的帧	FCS
01-00-0c-00-					AAAA03	00000C						

00												

有关详细信息，请参阅交换机间链路和 IEEE 802.1Q 帧格式。

802.1Q 操作概述

IEEE 802.1Q 标准除规定封装类型外，还规定了许多其他内容，其中包括生成树增强功能、GARP（请参阅本文档中的 VTP 部分）以及 802.1p 服务质量 (QoS) 标记。

802.1Q 帧格式保留了原始以太网源地址和目标地址，但交换机现在要求必须接收到小巨人帧，即使在主机可使用标记表示 QoS 信令的 802.1p 用户优先级的接入端口上也是如此。该标记包含 4 个字节，因此 802.1Q 以太网 v2 帧包含 1522 个字节（IEEE 802.3ac 工作组取得的成果）。802.1Q 还支持对 4096 个 VLAN 空间进行编号。

传输和接收的所有数据帧都带有 802.1Q 标记，但本地 VLAN 上的数据帧除外（存在基于交换机输入端口配置的隐式标签）。本地 VLAN 上的帧始终以无标记方式传输，并且通常以无标记方式接收。但是，也可以以带标记方式接收这些帧。

有关详细信息，请参阅通过 IEEE 802.10 实现 VLAN 标准化和获取 IEEE 802。

802.1Q/801.1p 帧格式

		标记头						
		TPID	TCI					
48 位	48 位	16 位	3 位	1 位	12 位	16 位	可变长度	32 位
DA	SA	TPID	优先级	CFI	VLAN ID	长度/类型	数据和 PAD	FCS
		0x8100	0-7	0-1	0-4095			

建议

由于所有较新硬件均支持 802.1Q（某些硬件仅支持 802.1Q，例如 Catalyst 4500/4000 系列和 CSS 11000），Cisco 建议所有新的实施都应遵循 IEEE 802.1Q 标准，并且早期网络应逐渐从 ISL 迁移。

IEEE 标准允许供应商执行互操作。随着支持 802.1p 的新型主机 NIC 和设备的问世，这将有利于所有 Cisco 环境。虽然 ISL 和 802.1Q 的实施已经成熟，但 IEEE 标准最终将得到更广泛的应用，并将获得更多的第三方支持，如网络分析器支持。802.1Q 比 ISL 具有更低的封装开销是它受到青睐的次要原因。

由于会在多台使用 DTP 的交换机之间就封装类型进行协商，并且如果两端都支持 ISL，则会在默认情况下优先选择 ISL，因此必须发出以下命令以指定 dot1q：

```
set trunk mod/port mode dot1q
```

如果从中继中清除了 VLAN 1（如本文档中的带内管理部分所述），尽管不会传输或接收用户数据

，但 NMP 仍将在 VLAN 1 中继续传送控制协议（例如，CDP 和 VTP）。

此外，当建立中继时，将始终在 VLAN1 中发送 CDP、VTP 和 PAgP 数据包，如本文档中的 VLAN1 部分所述。如果采用 dot1q 封装，当交换机的本地 VLAN 发生改变时，这些控制帧将带有 VLAN1 标记。如果启用中继到路由器的 dot1q，并且已在交换机上更改了本地 VLAN，则需要使用 VLAN 1 中的子接口，以便接收带标记的 CDP 帧，并在路由器上提供 CDP 邻居可见性。

注意： 本地 VLAN 的隐式标记可能导致 dot1Q 出现安全问题，这是因为可能会将帧从一个 VLAN 发送到另外一个 VLAN，而不会经过路由器。请参阅 VLAN 实施是否存在漏洞？ 以获取更多详细信息。解决方法是对不用于最终用户访问的中继本地 VLAN 使用 VLAN ID。大多数 Cisco 客户只需将 VLAN 1 保留为中继上的本地 VLAN，并向除 VLAN 1 以外的 VLAN 分配接入端口即可实现此目标。

中继模式

DTP 是第二代动态 ISL (DISL)，开发 DTP 的目的是为了确保中继两端的交换机就发送 ISL 或 802.1Q 帧所涉及到的不同参数（例如，配置的封装类型、本地 VLAN 和硬件功能）达成一致。通过确保端口及其邻居处于一致状态，还可以防止无中继端口泛溢标记帧，从而避免了严重的安全风险。

操作概述

DTP 是用于在交换机端口及其邻居之间协商配置参数的第 2 层协议。它使用另一多播 MAC 地址 (01-00-0c-cc-cc-cc) 和 SNAP 协议类型 0x2004。下表汇总了多种配置模式：

模式	功能	是否传输 DTP 帧	最终状态（本地端口）
Auto	使端口愿意将链路转换为中继。如果邻接端口设置为 on 或 desirable 模式，那么该端口将变成中继端口。	是，定期。	建立中继
在	将端口置于永久 trunking 模式，并通过协商把链路转换成中继。该端口成为中继端口，即使其邻接端口不同意此更改。	是，定期。	中继，无条件。
Nonegotiate	将端口置于永久中继模式，但阻止端口生成 DTP 帧。必须将邻接端口手动配置为中继端口，以建立中继链路。这对不支持 DTP 的设备非常有用。	否	中继，无条件。
	使端口主动尝试将链路转换成中继链路。如果邻接端口设置为 on、desirable 或 auto 模式，那么该端口将变成中继端口。	是，定期。	仅当远程模式为

			on、auto 或 desirable 时，才会最终处于 trunking 状态。
	将端口置于永久非中继模式，然后通过协商将链路转换成非中继链路。该端口成为非中继端口，即使其邻接端口不同意此更改。	在稳定状态下不会传输，但从 on 模式更改至此模式后，传输会发出加快远程终端检测的通知。	非中继

以下是有关该协议的一些要点：

- DTP 采用点对点连接，而 Cisco 设备仅支持采用点对点连接的 802.1Q 中继端口。
- 在 DTP 协商时，端口不会加入 STP。仅在端口已成为三种 DTP 类型（接入、ISL 或 802.1Q）之一后，才会将该端口添加到 STP。否则，在端口加入 STP 之前，PAgP（如果已配置）将成为运行的下一个进程。
- 如果端口在ISL模式下中继，DTP数据包将在VLAN 1上发出，否则(对于802.1q中继或非中继端口)会在本地VLAN上发出。
- 在 desirable 模式下，DTP 数据包会传输 **VTP 域名**（域名必须匹配才会出现协商的中继），以及中继配置和管理状态。
- 协商期间将每秒发送一次消息；协商之后将每 30 秒发送一次消息。
- 确保了解 on、nonegotiate 和 off 模式都明确指定了端口的最终状态。不良配置可能导致一端是中继、另一端不是中继的危险状态或不一致状态。
- on、auto 或 desirable 模式下的端口会定期发送 DTP 帧。如果处于 auto 或 desirable 模式下的端口在五分钟之内未收到 DTP 数据包，则会设置为非中继。

有关 ISL 的更多详细信息，请参阅在 Catalyst 5500/5000 和 6500/6000 系列交换机上配置 ISL 中继。有关 802.1Q 的更多详细信息，请参阅使用 802.1Q 封装和 Cisco CatOS 系统软件在 Catalyst 4500/4000、5500/5000 和 6500/6000 系列交换机之间建立中继。

建议

Cisco 建议在两端配置 desirable 模式的显式中继。在此模式下，网络操作员可以信任 syslog 和命令行状态消息显示的端口正在运行和建立中继，这与 on 模式不同，后者即使在错误配置邻居的情况下仍会使端口看起来正在运行。另外，在链路的一端无法成为中继或丢弃中继状态的情况下，desirable 模式下的中继可提供稳定性。发出以下命令以设置 desirable 模式：

```
set trunk mod/port desirable ISL | dot1q
```

注意： 在所有非中继端口上将中继设置为 off。这有助于在启动主机端口时避免浪费协商时间。**此命令还会在使用 set port host 命令时执行；**有关详细信息，请参阅 STP 部分。发出以下命令可以在一个端口范围中禁用中继：

```
set trunk port range off
!--- Ports are not trunking; part of the set port host command.
```

其他选项

另一个常见的客户配置仅在分布层使用 desirable 模式，并在接入层使用最简单的默认配置（auto 模式）。

某些交换机（如 Catalyst 2900XL）、Cisco IOS 路由器或其他供应商设备目前不支持通过 DTP 进行中继协商。您可以在 Catalyst 4500/4000、5500/5000、6500/6000 交换机上采用 nonegotiate 模式设置端口，以便无条件地与这些设备建立中继，这有助于实现校园网中的通用设置标准化。此外，您可以实施 nonegotiate 模式以减少总体的链路初始化时间。

注意： 信道模式和 STP 配置等因素也可影响初始化时间。

发出以下命令以设置 nonegotiate 模式：

```
set trunk mod/port nonegotiate ISL | dot1q
```

当连接到 Cisco IOS 路由器时，Cisco 建议采用 nonegotiate 模式，这是因为当进行桥接时，从 on 模式接收的一些 DTP 帧可以返回到中继端口。收到 DTP 帧后，交换机端口将设法进行不必要的重新协商（或关闭和启动中继）。如果已启用 nonegotiate，交换机不会发送 DTP 帧。

生成树协议

基本考虑因素

生成树协议 (STP) 维护冗余交换网络和桥接网络中的无环路 L2 环境。如果不采用 STP，帧会无限循环和/或增加，高流量数据会持续中断广播域中的所有设备，进而导致网络崩溃。

虽然从某些方面来讲，STP 是一个成熟的协议，最初是为基于软件的缓慢网桥规范 (IEEE 802.1d) 开发的，但要在大型交换网络中顺利实施此协议可能比较复杂，这些网络包含多个 VLAN、在一个域中具有多台交换机、需要支持多家供应商以及具有较新的 IEEE 增强功能。

CatOS 6.x 继续采用了新的 STP 技术，如 MISTP、环路防护、根防护和 BPDU 到达时间迟滞检测，以供将来使用。此外，CatOS 7.x 中还提供了已进一步标准化的协议，例如，IEEE 802.1s 共享生成树和 IEEE 802.1w 快速收敛生成树。

操作概述

在每个 VLAN 中，选择具有最小根网桥标识符 (BID) 的交换机作为根网桥。BID 由网桥优先级和交换机 MAC 地址组合而成。

首先，从所有交换机发送 BPDU，这些 BPDU 包含每台交换机的 BID 和到达该交换机的路径成本。这可以确定根网桥和到根之间的最低成本路径。根发送的 BPDU 中传送的其他配置参数将覆盖本地配置的参数，以使整个网络采用一致的计时器。

然后，通过以下步骤来收敛拓扑：

1. 为整个生成树域选择一个根网桥。
2. 在每个非根网桥中选择一个根端口（面向根网桥）。
3. 选择一个指定端口，以便在每个网段中转发 BPDU。
4. 非指定端口将被阻塞。

有关详细信息，请参阅配置生成树。

基本计时器默认值（秒）	名称	功能
2	Hello	控制 BPDU 的发送。
15	(Fwddelay)	控制端口在 Listening 和 Learning 状态下所消耗的时长，同时影响拓扑更改进程（请参阅下一部分）。
20	Maxage	控制交换机在寻找其他路径之前保持当前拓扑的时间。经过 Maxage 秒之后，交换机会认为 BPDU 已过期，并从阻塞端口池中寻找新的根端口。如果没有任何阻塞端口可用，则将其自身声明为指定端口上的根。

端口状态	含义	到下一状态的默认计时
	以管理方式关闭。	不适用
	接收 BPDU 和停止用户数据。	监控 BPDU 接收。等待 20 秒（以使 Maxage 过期）或立即更改（如果检测到直接/本地链路故障）。
	发送或接收 BPDU，以检查是否需要返回到 Blocking 状态。	Fwddelay 计时器（等待 15 秒）
	生成拓扑/CAM 表。	Fwddelay 计时器（等待 15 秒）
	发送/接收数据。	
	基本拓扑更改时间总计：	**如果等待 Maxage 过期，则时间为 20 + 2 (15) = 50 秒，如果直接链路发生故障，则为 30 秒**

STP 中的两种 BPDU 类型包括配置 BPDU 和拓扑更改通知 (TCN) BPDU。

配置 BPDU 流

配置 BPDU 来源于根网桥上每个端口上的每个 hello 间隔，并在随后流至所有分支交换机，以便维持生成树的状态。在稳定状态中，BPDU 流是单向的：根端口和阻塞端口仅接收配置 BPDU，而指定端口仅发送配置 BPDU。

对于交换机从根接收的每个 BPDU，Catalyst 中央 NMP 都会处理和外发一个包含根信息的新 BPDU。换句话说，如果根网桥丢失，或者到根网桥的所有路径丢失，将会停止接收 BPDU（直到 Maxage 计时器开始重新选择）。

TCN BPDU 流

当在生成树中检测到拓扑更改时，将从分支交换机获取 TCN BPDU，并将其发送到根网桥。根端口仅发送 TCN，指定端口仅接收 TCN。

TCN BPDU 传输至根网桥，并会在每个步骤中对其进行确认，因此该机制非常可靠。一旦它到达根网桥，根网桥将通过获取带有 TCN 标志设置的配置 BPDU，向整个域通知已发生更改，时间为 **maxage + fwddelay** 时间（默认为 35 秒）。这会导致所有交换机将其正常 CAM 老化时间从五分钟（默认值）更改为 **fwddelay** 指定的间隔（默认情况下为 15 秒）。有关更多详细信息，请参阅了解生成树协议拓扑更改。

生成树模式

可以采用三种不同的方法将 VLAN 与生成树相关联：

- 一个针对所有 VLAN 的生成树，或单一生成树协议，例如 IEEE 802.1Q
- 基于每个 VLAN 的生成树，或共享生成树，例如 Cisco PVST
- 基于每组 VLAN 的生成树，或多生成树，例如 Cisco MISTP 和 IEEE 802.1s

针对所有 VLAN 的单一生成树仅允许一个活动拓扑，因此无法实现负载均衡。STP 阻塞端口将阻塞所有 VLAN，并且不能传送数据。

每个 VLAN 的生成树允许负载均衡，但随着 VLAN 数量的增加，需要进行更多的 BPDU CPU 处理工作。在每个交换机的生成树中，CATOS版本注释提供建议的逻辑端口数量的相关指导。例如，Catalyst 6500/6000 Supervisor 引擎 1 的公式如下：

端口数量 + （中继数量 * 中继上的 VLAN 数量）< 4000

Cisco MISTP 和新的 802.1s 标准只允许定义两个有效 STP 实例/拓扑，所有 VLAN 都将映射到这两个生成树中的一个。此技术允许 STP 扩展到数千个 VLAN，同时支持负载均衡。

BPDU 格式

为了支持 IEEE 802.1Q 标准，现有的 Cisco STP 实现通过在 IEEE 802.1Q 单一生成树区域中添加隧道技术支持而扩展成为 PVST+。因此，PVST+ 与 IEEE 802.1Q MST 和 Cisco PVST 协议兼容，无需额外的命令或配置。另外，PVST+ 添加了验证机制，以确保交换机之间端口中继和 VLAN ID 配置的一致性。

以下是 PVST+ 协议的一些操作要点：

- PVST+ 利用所谓的通用生成树 (CST) 通过 802.1Q 中继与 802.1Q 单一生成树交互操作。CST 始终位于 VLAN 1 上，因此需要在中继上启用此 VLAN，才能与其他供应商交互操作。CST BPDU（总是没有标记）被传送到IEEE标准bridge-group(MAC地址01-80-c2-00-00-00，DSAP 42，SSAP 42)。为保持描述的完整性，同时将一组对应的 BPDU 发送到 VLAN 1 的 Cisco 共享生成树 MAC 地址。
- PVST+ 通过隧道在 802.1Q VLAN 区域中以多播数据的形式传输 PVST BPDU。对于中继上的每个 VLAN，Cisco 共享生成树 BPDU 会被传送到 MAC 地址 01-00-0c-cc-cc-cd（SNAP

HDLC 协议类型 0x010b）。BPDU 在本地 VLAN 上未进行标记，而在所有其他 VLAN 上则进行了标记。
- PVST+ 检查端口与 VLAN 是否存在不一致。PVST+ 将阻塞那些接收不一致 BPDU 的端口，以防止转发环路。它还将有关任何配置不匹配的情况通过 Syslog 消息通知用户。
- PVST+ 与在 ISL 中继上运行 PVST 的现有 Cisco 交换机向后兼容。ISL 封装的 BPDU 仍通过 IEEE MAC 地址发送或接收。换句话说，每个 BPDU 类型都是本地链路；不存在转换问题。

建议

默认情况下，所有 Catalyst 交换机均启用 STP。即使选中的设计不包括 L2 环路从而未启用 STP，我们仍然推荐使用这种方法，这样，就可以使它积极维护阻塞端口。

```
set spantree enable all
!--- This is the default.
```

出于以下原因，Cisco 建议使 STP 保持启用状态：

- 如果存在环路（由于不匹配、电缆损坏等原因导致），STP 将防止多播数据和广播数据对网络造成负面影响。
- 防止 EtherChannel 中断。
- 多数网络配置有 STP，以获得最大覆盖面。更多暴露通常等同于稳定代码。
- 防止双连接 NIC 行为不当（或在服务器上启用桥接）。
- 许多协议（例如 PAgP、IGMP 监听和中继）的软件与 STP 紧密相关。在没有 STP 的情况下运行可能会导致不良结果。

请勿更改计时器，因为此操作可能对稳定性产生负面影响。大多数部署的网络未经过调整。可通过命令行访问的一些简单的 STP 计时器（如 hello 间隔和 Maxage）本身由一组复杂的其他假设的和固有计时器组成，因此调整计时器和考虑可能出现的所有后果是很困难的。此外，可能会破坏 UDLD 保护。

理想情况下，禁止在管理 VLAN 上传输用户流量。特别是使用更旧的Catalyst交换机处理器时，最好的办法是保持管理VLAN与用户数据分离，避免STP出现问题。行为不当的终端站可能会使 Supervisor 引擎处理器忙于处理广播数据包，以致可能漏掉一个或多个 BPDU。然而，较新的拥有更强大 CPU 和扼杀控件的交换机可以解除您的这一顾虑。有关详细信息，请参阅本文档的带内管理部分。

请勿过度设计冗余。这可能会使故障排除工作变得极其困难 - 过多的阻塞端口给长期稳定性造成负面影响。**保持总 SPT 直径不超过 7 跳。**尽可能地设计为 Cisco 多层模型，这样交换域更小、STP 三角和决定性的阻塞端口更少（详见千兆园区网络设计 - 原理和体系结构中的说明）。

影响并了解根功能和阻塞端口所在的位置，并在拓扑图上对它们进行记录。STP 故障排除应该从阻塞端口开始 - 问题根源分析的一个关键部分，通常就是找出阻塞状态转换到转发状态的原因。**选择分布和核心层作为根/辅根所在的位置**，因为它们被认为是网络最稳定的部分。检查用 L2 数据转发路径覆盖 L3 和 HSRP 的最佳方式。此命令是用于配置网桥优先级的宏；设置为根，将大大低于默认值 (32768)，设置为辅根则适当低于默认值：

```
set spantree root secondary vlan range
```

注意：此宏将根优先级设置为 8192（默认值）、当前根优先级减 1（如果已知另一个根网桥），或当前根优先级（如果其 MAC 地址低于当前根）。

从中继端口修剪不必要的 VLAN（双向实施）。这会限制不需要特定 VLAN 的网络部分的 STP 直径

和 NMP 处理开销。VTP 自动修剪功能不会从中继中删除 STP。有关详细信息，请参阅本文档的 VTP 部分。也可以使用 CatOS 5.4 及更高版本从中继中删除默认 VLAN 1。

有关其他信息，请参阅生成树协议问题及相关设计注意事项。

其他选项

Cisco 还有一个称为 VLAN 网桥的 STP。此协议使用目标 MAC 地址 **01-00-0c-cd-cd-ce** 运行，协议类型为 0x010c。

如果您需要桥接VLAN之间的不可路由或传统协议，而不干扰在那些VLAN上运行的IEEE生成树实例，这种方法是最有用的。如果非桥接数据流的 VLAN 接口阻塞 L2 数据流（当它们与 IP VLAN 加入同一个 STP 时，很可能发生这种情况），覆盖的 L3 数据流也会无意中被删除 - 这是一个我们不希望看到的负面影响。因此对于桥接协议，VLAN 网桥是一个单独的 STP 实例，它提供单独的拓朴，可以在不影响 IP 数据流的情况下进行操作。

如果在 Cisco 路由器（例如 MSFC）的 VLAN 之间需要网桥，Cisco 建议您运行 VLAN 网桥。

Portfast

PortFast 用于绕过接入端口上正常的生成树运行，以加快终端站与它们在链路初始化后需要连接的服务之间的连接。对于某些协议（如 IPX/SPX），应确保在链路接通后立即将接入端口转入转发模式以避免 GNS 问题，这一点非常重要。

有关详细信息，请参阅使用 PortFast 和其他命令解决工作站启动连接延迟问题。

操作概述

在得知链路即将运行后，PortFast 会将端口从阻塞模式直接转换为转发模式，以跳过 STP 的正常的 listening 和 learning 状态。如果未启用此功能，STP 将丢弃所有用户数据，直到确定端口已准备好转入 forwarding 模式。这可能会需要多达两倍于 ForwardDelay 时间的时间（默认情况下总共是 30 秒）。

PortFast 模式还可防止每次端口状态从 learning 更改为 forwarding 时生成 STP TCN。TCN 本身并不是问题，但是，如果大量 TCN 发送到根网桥（通常是在早上人们打开其 PC 时），则可能不必要地延长收敛时间。

STP PortFast 对于多播 CGMP 和 Catalyst 5500/5000 MLS 网络尤为重要。这些环境中的 TCN 可能导致静态 CGMP CAM 表条目过期（造成多播数据包丢失，直到下个 IGMP 报告），和/或清理 MLS 缓存条目（需要重新建立，根据缓存大小的不同，可能导致路由器 CPU 使用率达到高峰）。（Catalyst 6500/6000 MLS 实现和通过 IGMP 监听获知的多播条目不受影响。）

建议

Cisco 建议对所有活动主机端口启用 STP PortFast，并针对交换机之间的链路和未使用的端口禁用 STP PortFast。

此外，对于所有主机端口，必须禁止建立中继和开辟信道。默认情况下，会为建立中继和开辟信道而启用各个接入端口，但不会有意将相邻交换机放在主机端口上。如果这些协议需要协商，端口激活中的后续延迟可能会导致我们不愿意看到的情况，即使来自工作站的初始数据包（例如 DHCP 请求）未被转发。

CatOS 5.2 引入了宏命令 set port host port range，用于为接入端口实施此配置，并有助于大幅提高自动协商和连接性能：

```
set port host port range
!--- Macro command for these commands: set spantree portfast port range enable set trunk port
range off set port channel port range mode off
```

注意： PortFast 并不意味着生成树根本没有在那些端口上运行。仍然发送、接收和处理 BPDU。

其他选项

PortFast BPDU 防护通过在非中继端口上收到 BPDU 时，将该端口转换为 errdisable 状态，从而防止产生环路。

由于主机端口不能与交换机连接，因此在为 PortFast 配置的接入端口上永远不会收到 BPDU 数据包。如果观察到 BPDU，它表明这是一个无效的并且可能很危险的配置，需要执行管理操作。启用 BPDU 防护功能后，生成树将关闭配置了 PortFast 的负责接收 BPDU 的接口，而不是将它们转入 STP blocking 状态。

如下所示的命令基于每台交换机（而不是基于每个端口）运行：

```
set spantree portfast bpdu-guard enable
```

如果端口关闭，将通过 SNMP 陷阱或 syslog 消息通知网络管理员。也可以为处于 errdisable 状态的端口配置自动恢复时间。有关详细信息，请参阅本文档的 UDLD 部分。有关详细信息，请参阅生成树 PortFast BPDU 防护增强功能。

注意： 中继端口的 PortFast 是在 CatOS 7.x 中引入的，对更早期版本的中继端口没有任何影响。中继端口的 PortFast 旨在增加 L3 网络的收敛时间。为实现这项功能，CatOS 7.x 还引入了在每个端口上配置 PortFast BPDU 防护的潜在功能。

UplinkFast

UplinkFast 提供在网络接入层中发生直接链路故障以后进行快速 STP 收敛的功能。它不修改 STP，其目的是将特定环境下的收敛时间缩短到不足三秒，而不是通常的延迟 30 秒。有关详细信息，请参阅了解和配置 Cisco Uplink Fast 功能。

操作概述

在接入层上使用 Cisco 多层设计模型时，如果转发上行链路丢失，阻塞的上行链路将立即转换为 forwarding 状态，而不等待 listening 和 learning 状态。

上行链路组是一组可以被当作根端口和备用根端口的基于 VLAN 的端口。通常情况下，根端口可保证从接入层到根的连通性。如果此主根连接由于任何原因而失败，可以立即启用备用根链路，而无需经过通常为 30 秒的收敛延迟。

由于这能够有效地绕过正常 STP 拓扑更改处理流程（listening 和 learning），因此需要一个备选的拓扑更正机制，以便更新本地终端站能够通过备选路径到达的域中的交换机。运行 UplinkFast 的接入层交换机也为其 CAM 到多播 MAC 地址（01-00-0c-cd-cd-cd，HDLC 协议 0x200a）中的每个 MAC 地址生成帧，以使用新拓扑更新相应域中所有交换机中的 CAM 表。

建议

Cisco 建议对带有阻塞端口的交换机（通常在接入层）启用 UplinkFast。如果没有获得某备用根链路的隐含拓扑信息（通常是 Cisco 多层设计中的分布交换机和核心交换机），请勿在交换机上使用该备用根链路。可以添加该功能，而无需中断生产网络。发出此命令以启用 UplinkFast：

```
set spantree uplinkfast enable
```

此命令还会将**网桥优先级**设置为高，以便将变为根网桥的风险降到最低；并将端口优先级设置为高，以便将变为指定端口（这将中断此功能）的风险降到最低。如果要恢复已启用 UplinkFast 的交换机，则必须禁用该功能，用“clear uplink”命令清除上行链路数据库，并手动恢复网桥优先级。

注意： 启用协议过滤功能时，必须使用 UplinkFast 命令的 **all protocols** 关键字。因为在启用协议过滤后，CAM 将记录协议类型及 MAC 和 VLAN 信息，因此必须为每个 MAC 地址上的各个协议生成 UplinkFast 帧。**rate** 关键字指示 UplinkFast 拓扑更新帧的每秒数据包数。建议使用默认值。无需为快速生成树协议 (RSTP) 或 IEEE 802.1w 配置 BackboneFast，因为 RSTP 中自带该机制并会自动启用。

Backbonefast

BackboneFast 提供从间接链路故障中进行快速收敛的功能。通过 STP 的此项新增功能，收敛时间通常可以从默认值 50 秒缩短到 30 秒。

操作概述

当交换机上的根端口或阻塞端口收到来自其指定网桥的下级 BPDU 时，即启动此机制。当下游交换机已失去与根的连接，并开始发送它自己的 BPDU 来选择新根时，可能会出现这种情况。**下级 BPDU** 将交换机同时标识为根网桥和指定网桥。

在正常的生成树规则下，接收交换机将在配置的最大老化时间（默认为 20 秒）内忽略下级 BPDU。然而，使用 BackboneFast，交换机将下级 BPDU 当作拓朴结构可能已更改的信号，并尝试使用根链路查询 (RLQ) BPDU 确定它是否具有通往根网桥的备选路径。通过添加此协议，可允许交换机检查根是否仍然可用，在更短时间内将阻塞端口转入 forwarding 状态，以及通知发送下级 BPDU 的隔离交换机此根仍然在那个位置。

以下是有关该协议操作的一些要点：

- 交换机仅将 RLQ 数据包从根端口中传出（即，传向根网桥）。
- 如果接收 RLQ 的交换机是根交换机，或者它知道自己与根的连接已经断开，则此交换机可以进行回复。如果该交换机不知道这些事实，它必须从其根端口转发查询。
- 如果交换机失去了与根的连接，它必须以否定的方式回复此查询。
- 此回复只能从查询传入的端口发出。
- 根交换机必须始终使用肯定回复回应此查询。
- 如果在非根端口收到回复，该回复将被丢弃。

由于 maxage 没有期限限制，因此 STP 收敛时间最多可以减少 20 秒钟。

有关详细信息，请参阅了解和配置 Catalyst 交换机上的 Backbone Fast。

建议

Cisco 建议在运行 STP 的所有交换机上启用 BackboneFast。可以添加该功能，而无需中断生产网络。发出以下命令以启用 BackboneFast：

```
set spantree backbonefast enable
```
注意： 需要在域内的所有交换机上配置此全局级命令，因为它向 STP 协议中添加了所有交换机都需要了解的功能。

其他选项

2900XL 和 3500 系列不支持 BackboneFast。如果交换机域除了 Catalyst 4500/4000、5500/5000 和 6500/6000 交换机之外还包含这些交换机，则不能启用此功能。

无需为 RSTP 或 IEEE 802.1w 配置 BackboneFast，因为 RSTP 中自带该机制并会自动启用。

生成树环路防护

环路防护是 Cisco 针对 STP 的专有优化。环路防护保护 L2 网络，使其不会因以下原因而形成环路：

- 网络接口发生故障
- CPU 忙
- 阻止 BPDU 正常转发的任何原因

在冗余拓扑中，当阻塞端口错误地转换为转发状态时，则会产生 STP 环路。通常，发生此转换的原因是物理冗余拓扑中的其中一个端口（不一定是阻塞端口）停止接收 BPDU。

环路防护只是在交换机由点对点链路连接的交换网络中有用。多数现代园区网络和数据中心网络都是这种类型的网络。在点对点链路上，除非发送一个下级 BPDU 或关闭链路，否则指定的网桥是不会消失的。STP 环路防护功能是在适用于 Catalyst 4000 和 Catalyst 5000 平台的 CatOS 版本 6.2(1) 中，以及适用于 Catalyst 6000 平台的 CatOS 版本 6.2(2) 中引入的。

有关环路防护的详细信息，请参阅使用环路防护和 BPDU 迟滞检测功能的生成树协议增强功能。

操作概述

环路防护检查以确定根端口或替代/备用根端口是否接收 BPDU。如果端口不接收 BPDU，环路防护会将端口置于不一致状态（阻塞），直至端口重新开始接收 BPDU。处于不一致状态的端口不会传输 BPDU。如果此端口重新接收 BPDU，则该端口（和链路）将再次被视为可用。将从该端口删除环路不一致情况，STP 能确定端口状态是因为此类恢复是自动进行的。

环路防护可隔离故障，并让生成树汇聚成稳定的拓扑，而不发生链路或网桥故障。环路防护将以所用 STP 版本的速度防止出现 STP 环路。它不依赖于 STP 本身（802.1d 或 802.1w），也不受 STP 计时器调整的影响。由于以上原因，在依赖 STP 并且软件支持相应功能的拓朴结构中，请与 UDLD 一起实施环路防护。

当环路防护阻塞了一个不一致的端口时，会将以下消息记录到日志中：

```
set spantree backbonefast enable
```
一旦在处于环路不一致 STP 状态的端口收到 BPDU，该端口就会转换到其他 STP 状态。根据收到的 BPDU，恢复是自动进行的，无需进行干预。在恢复之后，会将以下消息记录到日志中。

```
set spantree backbonefast enable
```

与其他 STP 功能的交互作用

- **根防护**根防护始终强制指定某一端口。仅当端口为根端口或备用端口时，环路防护才有效。这些功能互相排斥。在一个端口上不能同时启用环路防护和根防护。
- **UplinkFast**环路防护与 UplinkFast 兼容。如果环路防护将一个根端口置于阻塞状态，则 UplinkFast 会将一个新的根端口置于转发状态。并且，UplinkFast 不会选择 loop-inconsistent 端口作为根端口。
- **Backbonefast**环路防护与 BackboneFast 兼容。接收到来自指定网桥的下级 BPDU 会触发 BackboneFast。由于是从此链路中接收到 BPDU，环路防护未激活，因此 BackboneFast 与环路防护兼容。
- **Portfast**一旦链接连通，PortFast 会立即将端口转换为转发指定状态。由于启用 PortFast 的端口不能是根端口或备用端口，因此环路防护和 PortFast 是互相排斥的。
- **Pagp**环路防护使用 STP 已知的端口。因此，环路防护可利用 PAgP 提供的逻辑端口抽象概念。但是，要形成信道，在信道中分组的所有物理端口均需具有兼容的配置。PAgP 会在所有物理端口上强制实施环路防护的统一配置以形成信道。**注意：** 以下是在 EtherChannel 上配置环路防护时需注意的问题：STP 始终选取信道中的第一个操作端口来发送 BPDU。如果此链路变为单向，那么即使信道中的其他链路正常工作，环路防护也会阻塞信道。如果已被环路防护阻塞的端口组合在一起以形成信道，则 STP 会丢失这些端口的所有状态信息。新信道端口可获得具有指定角色的转发状态。如果信道被环路防护阻塞且信道中断，则 STP 会丢失所有状态信息。即使形成信道的一个或多个链路为单向链路，各个物理端口也可获得具有指定角色的转发状态。在上面所列出的后两种情况下，可能会形成环路，直到 UDLD 检测到该故障为止。但是环路防护无法检测到该环路。

环路防护和 UDLD 功能比较

环路防护功能和 UDLD 功能部分重叠。两者均可防止单向链路导致的 STP 故障。但是这两个功能在问题解决方法以及功能方面有所不同。具体来说，存在某些 UDLD 无法检测到的单向故障，例如由不发送 BPDU 的 CPU 导致的故障。另外，使用主动 STP 计时器和 RSTP 模式可能会导致形成环路，直到 UDLD 可检测到该故障。

在共享链路上，或者在链路自连通后成为单向链路的情况下，环路防护不起作用。如果链路自连通后成为单向链路，端口将从不接收 BPDU 且会成为指定端口。此行为可能是正常的，因此该特定情况不在环路防护的范围之内。UDLD 可以防止出现这样的情况。

若要提供最高级别的防护，请同时启用 UDLD 和环路防护。有关环路防护和 UDLD 功能比较的信息，请参阅使用环路防护和 BPDU 迟滞检测功能的生成树协议增强功能中的*环路防护与单向链路检测 (UDLD)* 部分。

建议

Cisco 建议您在具有物理环路的交换机网络上全局启用环路防护。在 Catalyst 软件的 7.1(1) 版本及更高版本中，您可以在所有端口上全局启用环路防护。实际上，是在所有点对点链路上启用此功能。链路的双工状态可检测到点对点链路。如果双工是全双工，则认为链路是点对点链路。若要启用全局环路防护，请发出以下命令：

```
set spantree global-default loopguard enable
```

其他选项

对于不支持全局环路防护配置的交换机，请在所有各个端口（包括端口信道端口）上启用此功能。尽管在指定端口上启用环路防护没有任何优点，但启用此功能不会产生问题。另外，有效的生成树重新收敛实际上可将指定端口变为根端口，这可使此功能在该端口上变得有用。若要启用环路防护，请发出以下命令：

```
set spantree guard loop mod/port
```

如果意外形成环路，具有无环路拓扑的网络仍可从环路防护功能中受益。但是，在此类型的拓扑中启用环路防护可能会导致网络隔离问题。若要构建无环路的拓扑并避免网络隔离问题，请发出以下命令，从而全局或分别禁用环路防护。请不要在共享链路上启用环路防护。

- ```
  set spantree global-default loopguard disable
  !--- This is the global default.
  ```
  或
- ```
  set spantree guard none mod/port
  !--- This is the default port configuration.
  ```

生成树根防护

根防护功能提供了在网络中强制执行根网桥安置的方法。根防护可确保启用根防护的端口为指定端口。通常，除非根网桥的两个或多个端口连接在一起，否则根网桥端口全部为指定端口。如果网桥在启用了根防护的端口上收到高级 STP BPDU，则网桥会将此端口转换为根不一致 STP 状态。此根不一致状态实际上等效于监听状态。此时不会通过此端口转发任何流量。根防护以这种方式强制确定根网桥的位置。在针对 Catalyst 29xx、4500/4000、5500/5000 和 6500/6000 的 6.1.1 版及更高版本的 CatOS 中提供了根防护功能。

操作概述

根防护是 STP 内置机制。根防护自身没有计时器，它仅依赖于接收 BPDU。如果对端口应用根防护，则根防护不允许端口成为根端口。如果 BPDU 的接收触发可使指定端口成为根端口的生成树收敛，则该端口会置于根不一致状态。以下 syslog 消息显示该操作：

```
set spantree guard none mod/port
!--- This is the default port configuration.
```

端口停止发送高级 BPDU 后，该端口会再次解除阻塞。通过 STP，端口会从监听状态进入学习状态，并最终转换为转发状态。恢复是自动进行的，并且不需要人为干预。以下 syslog 消息提供一个示例：

```
set spantree guard none mod/port
!--- This is the default port configuration.
```

根防护会强制某端口成为指定端口，且只有在该端口为根端口或备用端口的情况下环路防护才有效。因此，这两个功能是互相排斥的。在一个端口上不能同时启用环路防护和根防护。

有关详细信息，请参阅生成树协议根防护增强功能。

建议

Cisco 建议您在与不受直接管理控制的网络设备连接的端口上启用根防护功能。若要配置根防护，请发出以下命令：

set spantree guard root *mod/port*

EtherChannel

EtherChannel 技术允许将多个信道（Catalyst 6500/6000 上最多有 8 个）逆向多路复用到单个逻辑链路中。虽然在实施中每个平台与下一个平台有所不同，但了解共同的要求很重要：

- 用于在多个信道上统计多路复用帧的算法
- 创建逻辑端口，以便可以运行单个 STP 实例
- 信道管理协议，例如 PAgP 或链路聚合控制协议 (LACP)

帧多路复用

EtherChannel 包括一种帧分配算法，该算法可以跨组件 10/100 或千兆链路有效地多路复用帧。每个平台算法上的区别在于做出分配决策时，每种硬件类型提取帧头信息的能力不同。

负载分配算法是可用于两种信道控制协议的全局选项。PAgP 和 LACP 使用帧分配算法，这是因为 IEEE 标准不要求使用任何特定的分配算法。但是，任何分配算法均可确保在接收帧时，算法不会导致属于任何给定对话的帧顺序混乱或帧重复。

注意： 必须考虑以下信息：

- Catalyst 6500/6000 具有比 Catalyst 5500/5000 更新的交换硬件，并且能够以线速读取 IP 第 4 层 (L4) 信息，从而做出比简单的 MAC L2 信息更明智的多路复用决策。
- Catalyst 5500/5000 功能取决于在模块上是否存在以太网捆绑芯片 (EBC)。**show port capabilities mod/port 命令可确认在每个端口上能实现哪些功能。**

请参阅下表，该表详细说明了针对各所列平台的帧分配算法：

平台	信道负载均衡算法
Catalyst 5500/5000/5000 系列	带有必要模块的 Catalyst 5500/5000 允许在每个 FEC1 中存在两到四条链路，尽管这些链路必须在同一模块上。源和目标 MAC 地址对可确定为帧转发选择的链路。将对源 MAC 地址和目标 MAC 地址的最低有效的两位执行 X-OR 运算。此运算会产生以下四种结果之一：(0 0)、(0 1)、(1 0) 或 (1 1)。其中每个值均指向 FEC 捆绑中的一个链路。对于双端口 Fast EtherChannel，在 X-OR 运算中只使用一个位。可能会出现源/目标地址对中的一个地址固定不变的情况。例如，目标可能是服务器，或者更有可能是路由器。在这种情况下，会看到统计负载均衡，因为源地址始终是不同的。
Catalyst	Catalyst 4500/4000 EtherChannel 根据每个帧的源和目标 MAC 地址的低位，向信道的各链路（位于单个模块上）分配帧。与 Catalyst 5500/5000 比较，该算法更加复杂，且使用 MAC DA（3、5、6 个字节）、SA（3、5、6 个字节）、输入端口和 VLAN ID 这些字

4500/4000系列	段的确定性散列。帧分配方法是不可配置的。
Catalyst 6500/6000系列	有两种可能的散列算法，具体取决于 Supervisor 引擎硬件。哈希是在，在所有的情况下，采取MAC地址、IP地址或者IP TCP/UDP2端口号并且运用算法创造三比特值的硬件方面实现的第十七个度多项式。将分别对源和目标地址执行该算法。然后，将对结果执行 XOR 运算以生成另一个三位值，该值用于确定使用信道中的哪个端口来转发数据包。Catalyst 6500/6000 上的信道可以在任何模块上的端口之间形成，并且最多可以达到 8 个端口。

[1] FEC =快速以太信道

[2] UDP =用户数据报协议

下表说明各个 Catalyst 6500/6000 Supervisor 引擎型号支持的分配方法及其默认行为。

硬件	说明	分配方法
WS-F6020（L2 引擎）	早期的 Supervisor 引擎 1	L2 MAC：SA;DA;SA & DA
WS-F6020A (L2引擎) WS-F6K-PFC (L3引擎)	较新的 Supervisor 引擎 1 和 Supervisor 引擎 1A/PFC1	L2 MAC：SA;DA;SA & DA L3 IP:SA;DA;SA 和 DA（默认）
WS-F6K-PFC2	Supervisor 引擎 2/PFC2（需要 CatOS 6.x）	L2 MAC：SA;DA;SA & DA L3 IP:SA;DA;SA & DA (默认) L4会话：S 端口；D 端口；S & D 端口（默认）
WS-F6K-	Supervisor引擎720/PFC3A	L2

PFC3BXL WS-F6K-PFC3B WS-F6K-PFC3A	(需要CatOS 8.1.x) Supervisor引擎720/Supervisor引擎32/PFC3B (需要CatOS 8.4.x) Supervisor引擎720/PFC3BXL (需要CatOS 8.3.x)	MAC：SA;DA;SA & DA L3 IP:SA;DA;SA & DA (默认) L4会话：S 端口；D 端口；S & D端口IP-VLAN-L4会话：SA & VLAN & S 端口；DA & VLAN & D 端口；SA & DA & VLAN & S 端口 & D 端口

注意： 如果采用 L4 分配，第一个分段的数据包使用 L4 分配。所有后续数据包使用 L3 分配。

有关其他平台上的 EtherChannel 支持以及如何对其进行配置和故障排除的更多详细信息，可在以下文档中找到：

- 了解 Catalyst 交换机上的 EtherChannel 负载均衡和冗余
- 在运行 CatOS 系统软件的 Catalyst 4500/4000、5500/5000 和 6500/6000 交换机之间配置 EtherChannel
- 在 Catalyst 6500/6000 和 Catalyst 4500/4000 之间配置 LACP (802.3ad)
- 配置第 3 层和第 2 层 EtherChannel

建议

默认情况下 Catalyst 6500/6000 系列交换机按照 IP 地址执行负载均衡。建议在 CatOS 5.5 中这样做，且假定 IP 为主要协议。若要设置负载均衡，请发出以下命令：

```
set port channel all distribution ip both
!--- This is the default.
```

在大多数网络中，对于 Catalyst 4500/4000 和 5500/5000 系列可以按照 L2 MAC 地址进行帧分配。但是，如果只有两个主设备通过信道通信，则所有流量会使用同一链路（因为 SMAC 和 DMAC 固定不变）。通常是服务器备份和其他大文件传输问题或两个路由器之间的传输分段问题。

尽管逻辑聚合端口 (agport) 能够由 SNMP 作为单独的实例管理，并聚合收集的吞吐量统计，Cisco 仍建议您单独管理每一个物理接口，从而检查帧分配机制的工作方式以及是否实现了统计负载均衡。

与使用 CatOS 5.x 中的 show counters mod/port 命令或 show mac mod/port 命令检查各端口计数器相比，使用 CatOS 6.x 中的新命令 show channel traffic 命令更容易显示百分比分配统计。使用 CatOS 6.x 中的另一个新命令 show channel hash 命令，可以根据分配模式查看将选择哪个端口作为特定地址和/或端口号的传出端口。对于 LACP 信道，其等效命令是 show lacp-channel traffic 命令和 show lacp-channel hash 命令。

其他选项

如果因 Catalyst 4500/4000 或 Catalyst 5500/5000 基于 MAC 的算法受到相对限制而出现问题，且没有实现良好的统计负载均衡，则可以采取以下步骤：

- 点部署 Catalyst 6500/6000 交换机
- 增加带宽，但不通过交换建立信道，例如从几个 FE 端口到一个 GE 端口，或者从几个 GE 端口到一个 10 GE 端口
- 对具有大量数据流的终端站对进行重新编址
- 为高带宽设备提供专用链路/VLAN

EtherChannel 配置指南和限制

在将兼容端口聚合到单个逻辑端口之前，EtherChannel 会验证所有物理端口上的端口属性。配置指南和限制对不同交换机平台会有所不同。请遵照指南以避免出现捆绑问题。例如，如果已启用 QoS，则在捆绑具有不同 QoS 功能的 Catalyst 6500/6000 系列交换模块时，EtherChannel 不会形成。**在 Cisco IOS 软件中，您可以使用 no mls qos channel-consistency 端口信道接口命令禁止对 EtherChannel 捆绑执行 QoS 端口属性检查。**用于禁止 QoS 端口属性检查的等效命令在 CatOS 中不可用。**您可以发出 show port capability mod/port 命令来显示 QoS 端口功能并确定端口是否兼容。**

请遵照以下针对不同平台的指南以避免出现配置问题：

- 配置 EtherChannel 的 *EtherChannel 配置指南* 部分 (Catalyst 6500/6000)
- 配置 Fast EtherChannel 和 Gigabit EtherChannel 的 *EtherChannel 配置指南和限制* 部分 (Catalyst 4500/4000)
- 配置 Fast EtherChannel 和 Gigabit EtherChannel 的 *EtherChannel 配置指南和限制* 部分 (Catalyst 5000)

注意： Catalyst 4000 支持的最大端口信道数是 126。使用软件版本 6.2(1) 及更低版本时，6 插槽和 9 插槽 Catalyst 6500 系列交换机最多可支持 128 个 EtherChannel。在软件版本 6.2(2) 及更高版本中，生成树功能可处理端口 ID。因此，对于 6 插槽或 9 插槽机箱，可支持的最大 EtherChannel 数为 126；对于 13 插槽机箱，可支持的最大 EtherChannel 数为 63。

端口聚合协议 (PAgP)

PAgP 是一种管理协议，用于检查链路两端的参数的一致性，并协助信道适应链路故障或添加。请注意有关 PAgP 的以下事实：

- PAgP 要求信道中的所有端口均属于同一 VLAN 或均配置为中继端口。（由于动态 VLAN 可以强制将端口更改到不同的 VLAN，因此在 EtherChannel 参与中未包括这些动态 VLAN。）
- 如果已存在链路捆绑且修改了某个端口的配置（例如更改 VLAN 或中继模式），则将修改链路捆绑中的所有端口以与该配置匹配。
- PAgP 不会对以不同速度或端口双工运行的端口进行分组。如果在存在链路捆绑的情况下更改速度和双工，PAgP 会更改链路捆绑中所有端口的速度和双工。

操作概述

PAgP 端口控制要分组的每一个物理（或逻辑）端口。使用对 CDP 数据包使用的同一多播组 MAC 地址 **01-00-0c-cc-cc-cc** 发送 PAgP 数据包。协议值是 0x0104。以下是协议操作的汇总：

- 只要物理端口处于 up 状态，就会在检测期间每秒钟传输一次 PAgP 数据包，在稳定状态下每 30 秒钟传输一次。
- 协议监听 PAgP 数据包，这种数据包证明物理端口与另一个支持 PAgP 的设备具有双向连接。
- 如果收到数据包但没有收到 PAgP 数据包，则假设端口连接到不支持 PAgP 的设备。

- 只要在一组物理端口上接收到两个 PAgP 数据包，就会尝试形成聚合端口。
- 如果 PAgP 数据包停止一段时间，则 PAgP 状态将为 down。

正常处理

必须定义以下概念以帮助理解协议行为：

- **Agport** - 由同一聚合中的所有物理端口组成的逻辑端口，可以通过其自己的 SNMP ifIndex 确定。因此，agport 不包含非操作端口。
- **信道** - 满足形成条件的聚合；因此它可以包含非操作端口（agport 是信道的子集）。协议（包括 STP 和 VTP，但不包括 CDP 和 DTP）在 agport 的 PAgP 之上运行。在 PAgP 将这些协议的 agport 连接到一个或多个物理端口之前，这些协议都无法发送或接收数据包。
- **组功能** - 每个物理端口和 agport 都拥有一个称为组功能的配置参数。当且仅当两个物理端口具有相同组功能时，其中一个物理端口才能与另一个物理端口聚合。
- **聚合过程** - 当物理端口到达 UpData 或 UpPAgP 状态时，它会连接到适当的 agport 上。当该程序离开这两个状态中的一个而进入另一个状态时，它将从 agport 分离。

下表给出了状态和创建过程的定义：

状态	含义
UpData	未接收到任何 PAgP 数据包。PAgP 数据包已发送。物理端口是连接到其 agport 的唯一端口。非 PAgP 数据包在物理端口和 agport 之间传入和传出。
Bidir	恰好已接收到一个 PAgP 数据包，证明仅与一个相邻端口存在双向连接。物理端口未连接到任何 agport。PAgP 数据包已发送并且可以被接收到。
UpPAgP	此物理端口（可能与其他物理端口关联）已连接到 agport。PAgP 数据包在该物理端口上发送和接收。非 PAgP 数据包在物理端口和 agport 之间传入和传出。

两个连接的两端都必须就分组的目标达成一致，该目标定义为连接两端都许可的 agport 中的最大端口组。

当物理端口到达 UpPAgP 状态时，会将其分配给具有成员物理端口的 agport，这些成员物理端口与新物理端口的组功能匹配，且处于 BiDir 或 UpPAgP 状态。（同时会将任何此类 BiDir 端口转移到 UpPAgP 状态。）如果不存在其构成物理端口参数与新就绪的物理端口兼容的 agport，则会将其分配给具有适当参数（这些参数没有关联的物理端口）的 agport。

在该物理端口已知的上一个相邻端口上可能会出现 PAgP 超时。从 agport 中删除端口超时。同时，也将删除其计时器也已超时的同一 agport 上的所有物理端口。这将导致另一端已停止的 agport 突然关闭，而不是一次关闭一个物理端口。

出现故障的行为

如果现有信道中的链路失败（例如，拔掉端口，删除千兆接口转换器 [GBIC] 或光纤损坏），则将更新 agport，并在一秒钟内通过其余链路散列数据流。不需要在故障后重新散列的任何数据流（继续在同一链路上发送的数据流）不会遭受任何损失。恢复故障链路将触发再次对 agport 进行更新，并且会再次对数据流进行散列处理。

注意： 由于关闭电源或删除模块而导致信道中出现链路故障时的行为可能有所不同。根据定义，信道需要有两个物理端口。如果在双端口信道已经丢失一个端口，则关闭逻辑agport，并用相关 生成树重新 初始化原始物理端口。这意味着在 STP 允许数据再次可以使用端口之前，可能会丢弃数据流。

在 Catalyst 6500/6000 上，此规则有一个例外。在 CatOS 6.3 以前的版本中，如果信道仅由模块 1 和模块 2 上的端口组成，则在删除模块期间 agport 的状态不会为 down。

在计划进行网络维护时，两种失败模式中的此差别很重要，因为在执行联机删除或插入模块时需要考虑 STP TCN。如上所述，使用 NMS 管理信道中的每个物理链路非常重要，因为 agport 可能未受故障干扰。

执行以下建议的步骤可减少对 Catalyst 6500/6000 进行不必要的拓扑更改：

- 如果每个模块都使用单个端口形成信道，则必须使用三个或多个模块（总共三个或多个端口）。
- 如果信道跨越两个模块，则必须在每个模块上使用两个端口（总共四个端口）。
- 如果在两个卡之间需要双端口信道，则只使用 Supervisor 引擎端口。
- 升级到 CatOS 6.3，它可处理模块的删除而不会对模块间的信道拆分进行 STP 重新计算。

配置选项

可以在不同模式下配置 EtherChannel，如下表所总结：

模式	可配置选项
在	PAgP 未在运行。不论相邻端口的配置方式如何，都对端口建立信道。如果相邻端口模式为 on，则形成信道。
	不论相邻端口的配置方式如何，都不会对端口建立信道。
Auto	聚合处于 PAgP 协议的控制之下。将端口置于被动协商状态，且不在接口上发送 PAgP 数据包，直到至少接收到一个指示发送方正在所需模式下运行的 PAgP 数据包。
	聚合处于 PAgP 协议的控制之下。将端口置于活动 negotiating state，在该状态下，端口通过发送 PAgP 数据包来启动与其他端口的协商。将与另一个处于 desirable 或 auto 模式的端口组形成信道。
Non-Silent（Catalyst 5500/5000 光纤 FE 和	auto 或 desirable 模式关键字。如果没有在接口上接收到数据包，则接口始终不会链接到 agport，也不能用于传输数据。由于某些链路故障会导致信道分离，因此为特定 Catalyst 5500/5000 硬件提供了此双向检查。由于已

GE 端口上的默认值）	启用 non-silent 模式，因此永远不允许恢复的相邻端口重新启动并不必要地分离信道。默认情况下，在 Catalyst 4500/4000 和 6500/6000 系列硬件中提供了更灵活的捆绑和改进的双向检查。
Silent（Catalyst 6500/6000 和 4500/4000 端口以及 5500/5000 铜缆端口上的默认值）	auto 或 desirable 模式关键字。如果在 15 秒超时期过后接口没有接收到任何数据包，接口将自动连接到 agport 并因此可用于数据传输。如果伙伴可以是从不发送 PAgP 的分析器或服务器，Silent 模式还允许信道操作。

静音/非静音设置影响端口如何对导致单向数据流的情况起反应或如何实现故障切换。当某个端口无法传输（例如由于物理子层 [PHY] 故障或者光纤或电缆损坏所致）时，相邻端口可能仍处于操作状态。伙伴将继续传输数据，但数据丢失，因为无法收到返回数据流。由于链路具有单向性，因此也可形成生成树环路。

一些光纤端口具有所需的功能，即当端口丢失其接收信号时 (FEFI) 时，将端口置于非操作状态。这将导致相邻端口进入非操作状态，实际上会导致链路两端的端口关闭。

如果使用的设备传输数据（例如 BPDU）且无法检测单向条件，则必须使用 non-silent 模式以允许端口保持非操作状态，直到存在接收数据且验证链路是双向的。PAgP 检测单向链路所需要的时间大约为 3.5 * 30 秒 = 105 秒，其中 30 秒是发送两个连续 PAgP 消息之间的时间。建议使用 UDLD，因为它是一个速度更快的单向链路探测器。

如果使用的设备不传输任何数据，则必须使用 silent 模式。这会强制端口处于已连接和操作状态，而不论接收到的数据是否存在。此外，对于能够检测单向环境(例如使用L1 FEFI和UDLD的更新的平台)的存在的端口，默认使用安静模式。

验证

下表描述了两个直接连接的交换机（交换机 A 和交换机 B）之间的所有可能的 PAgP 信道建立模式方案的摘要。其中某些组合可能会导致 STP 将信道建立端的端口置于 errdisable 状态（即，某些组合会关闭信道端的端口）。

交换机 A 信道模式	**交换机 B 信道模式**	**信道状态**
在	在	信道（非 PAgP）
在		不是信道 (errdisable)
在		不是信道 (errdisable)
在		不是信道 (errdisable)
	在	不是信道 (errdisable)

	在	不是信道 (errdisable)
		PAgP Channel
	在	不是信道 (errdisable)
		PAgP Channel
		PAgP Channel

建议

Cisco 建议在所有交换机到交换机信道连接上启用 PAgP，以避免 on 模式。首选方法是在链路的两端设置 desirable 模式。其他建议是将 silent/non-silent 关键字保留为默认值，即在 Catalyst 6500/6000 和 4500/4000 交换机上为 silent，在 Catalyst 5500/5000 光纤端口上为 non-silent。

如本文档中所讨论，在所有其他端口上将信道建立显式配置为 off，对于快速转发数据非常有帮助。最多等待 15 秒，因为必须避免在不用于信道建立的端口上出现 PAgP 超时，尤其是因为然后会将端口移交给 STP，而 STP 本身可能需要 30 秒来允许转发数据，此外 DTP 可能需要 5 秒，总共为 50 秒。本文档的 STP 部分中更详细地讨论了 **set port host** 命令。

```
set port channel port range mode desirable

set port channel port range mode off
!--- Ports not channeled; part of the set port hostcommand.
```

此命令为信道分配一个管理组号码，使用 show channel group 命令可以看到该号码。然后可以根据需要，按管理组号码管理以下操作：将信道端口添加到同一个 agport 以及从中删除信道端口。

其他选项

对于在接入层具有最小管理模型的客户，另一个常用配置是：在分布或核心层将模式设置为 desirable，让接入层交换机使用默认的 auto 配置。

与不支持 PAgP 的设备建立信道时，信道需要将硬编码设置为 on。这适用于多种设备，例如服务器、本地定向器、内容交换机、路由器、具有较旧软件的交换机、Catalyst XL 交换机和 Catalyst 8540。发出以下命令：

```
set port channel port range mode on
```

由于 CatOS 7.x 中提供的新 802.3ad IEEE LACP 标准可带来跨平台和供应商互操作性的优势，因此，从长期来看该标准很可能会取代 PAgP。

链路聚合控制协议

LACP 协议允许具有类似特性的端口通过与相邻的交换机进行动态协商来形成信道。PAgP 是 Cisco 专有的协议，只能在 Cisco 交换机和许可供应商发布的交换机上运行。但是 IEEE 802.3ad 中定义的 LACP 允许 Cisco 交换机用符合 802.3ad 规范的设备管理以太网信道。CatOS 7.x 软件版本引入了 LACP 技术支持。

LACP 和 PAgP 在功能方面几乎没有区别。这两个协议在每条信道中都最多支持八个端口，并且在捆绑形成之前检查相同的端口属性。这些端口属性包括：

- 速度
- 双工
- 本地 VLAN
- 中继类型

LACP 和 PAgP 之间的显著差异包括：

- LACP 只能在全双工端口运行，并且 LACP 不支持半双工端口。
- LACP 支持热备用端口。LACP 总是尝试配置一个信道中的最大兼容端口数，最多为硬件允许的最大数量（八个端口）。如果 LACP 不能聚合兼容的所有端口，则信道中无法主动包括的所有端口将转入热备用状态，并且只有当其中一个已使用端口发生故障时，才使用这些端口。例如，如果远程系统有限制性更强的硬件限制，则 LACP 不能聚合所有兼容端口。

注意： 在 CatOS 中，同一个管理键可以分配的最大端口数是 8。在 Cisco IOS 软件中，LACP 尝试配置一个 EtherChannel 中的最大兼容端口数，最多为硬件允许的最大数量（八个端口）。另外八个端口可以配置为热备用端口。

操作概述

LACP 控制要捆绑的每个物理（或逻辑）端口。LACP 数据包使用多播组 MAC 地址 **01-80-c2-00-00-02** 发送。类型/字段值为 0x8809，子类型为 0x01。以下是协议操作的汇总：

- 协议依靠设备来通告它们的聚合功能和状态信息。传输将在每条“可聚合”链路上定期发送。
- 只要物理端口打开，检测时 LACP 数据包每秒传输一次，并在处于稳定状态时每 30 秒传输一次。
- “可聚合”链路上的伙伴监听协议内发送的信息，并决定要采取何种操作。
- 在信道中配置兼容端口，最多可达硬件允许的最大数量（八个端口）。
- 通过在链路伙伴之间定期、及时交换最新状态信息来维护聚合。如果配置发生更改（例如由于链路故障），协议伙伴超时并根据系统的新状态采取相应的措施。
- 除定期 LACP 数据单元 (LACPDU) 传输之外，如果更改了状态信息，协议将向伙伴传送事件驱动的 LACPDU。协议伙伴根据系统的新状态采取相应的措施。

LACP 参数

为了允许 LACP 确定一组链路是否连接到同一个系统，以及从聚合角度来看这些链路是否是兼容，建立以下参数的能力是必要的：

- 参加链路聚合的每个系统的全局唯一标识必须为运行 LACP 的每个系统指定优先级，优先级可以自动选择，也可以由管理员指定。默认系统优先级是 32768。系统优先级主要用于与系统的 MAC 地址一起形成系统标识符。
- 根据指定系统对与每个端口及每台汇聚路由器相关的功能集的了解，用于识别该功能集的方法系统中的每个端口都必须自动分配或由管理员指定优先级。默认值为 128。优先级与端口号一起形成端口标识符。
- 识别链路聚合组及其关联汇聚路由器的方法端口与其他端口的聚合能力由另一个严格大于 0 的简单 16 位整数参数来汇总。此参数称为“键”。每个键由多种因素确定，例如：端口物理特性，包括：数据传输速度Duplexity点对点或共享介质网络管理员建立的配置约束以下两个键与每个端口相关联：管理键 - 此键允许通过管理来操纵键值。用户可以选择此键。操作键 - 系统使

用此键以形成聚合。用户不能选择或直接更改此键。系统中共享同一操作键值的一组端口被认为是同一键组的成员。

如果有两个系统和使用同一管理键的一组端口，每个系统都会尝试聚合这些端口。每个系统从最高优先级系统中具有最高优先级的端口开始。因为每个系统知道自己的优先级（用户或系统已指定）及其伙伴的优先级（通过 LACP 数据包发现），此行为是可能的。

出现故障的行为

LACP 的故障行为与 PAgP 的行为相同。如果现有信道中的一条链路发生故障，将更新 agport，并在 1 秒内在剩余链路上对数据流进行散列处理。链路会因为以下原因和其他原因而出现故障：

- 拔掉端口
- GBIC 被移除
- 光纤断开
- 硬件故障（接口或模块）

不需要在故障后重新散列的任何数据流（继续在同一链路上发送的数据流）不会遭受任何损失。恢复故障链路将触发再次对 agport 进行更新，并且会再次对数据流进行散列处理。

配置选项

可以在不同模式下配置 LACP EtherChannel，如下表所总结：

模式	可配置选项
在	不进行任何 LACP 协商，强制形成链路聚合。交换机既不发送 LACP 数据包，也不处理任何传入 LACP 数据包。如果相邻端口模式为 on，则形成信道。
	不论相邻端口的配置方式如何，都不会对端口建立信道。
Passive	这类似于 PAgP 中的 auto 模式。交换机不启动信道，但可以识别传入的 LACP 数据包。对等体（在 active 状态）通过发送 LACP 数据包启动协商。交换机接收并回复数据包，最终与对等交换机形成汇聚信道。
	这类似于 PAgP 中的 desirable 模式。交换机启动协商以形成 aglink。如果另一端以 LACP active 或 passive 模式运行，则形成链路聚合。

验证（LACP 和 LACP）

本部分中的表描述了两台直接连接的交换机（交换机 A 和交换机 B）之间的所有可能的 LACP 信道建立模式方案的摘要。其中某些组合可能会导致 STP 将信道建立端的端口置于 errdisable 状态。这意味着某些组合会关闭信道建立端的端口。

交换机 A 信道模式	交换机 B 信道模式	交换机 A 信道状态	交换机 B 信道状态
在	在	Channel (non-LACP)	Channel (non-LACP)
在		不是信道 (errdisable)	
在		不是信道 (errdisable)	
在		不是信道 (errdisable)	
		LACP Channel	LACP Channel
		LACP Channel	LACP Channel

验证（LACP 和 PAgP）

本部分中的表描述了两台直接连接的交换机（交换机 A 和交换机 B）之间的所有可能的 LACP 对 PAgP 信道建立模式方案的摘要。其中某些组合可能会导致 STP 将信道建立端的端口置于 errdisable 状态。这意味着某些组合会关闭信道建立端的端口。

交换机 A 信道模式	交换机 B 信道模式	交换机 A 信道状态	交换机 B 信道状态
在	在	Channel (non-LACP)	信道（非 PAgP）
在		不是信道 (errdisable)	
在		不是信道 (errdisable)	
在		不是信道 (errdisable)	
	在		不是信道 (errdisable)
	在		不是信道 (errdisable)
	在		不是信道 (errdisable)

建议

Cisco 建议您对 Cisco 交换机之间的信道连接启用 PAgP。当您与不支持 PAgP 但支持 LACP 的设备之间建立信道时，通过在设备两端将 LACP 配置为 active 来启用 LACP。如果设备的任何一端都不支持 LACP 或 PAgP，则您需要将此信道硬编码为 on。

- **set channelprotocol lacp** *module*

  在运行 CatOS 的交换机上，Catalyst 4500/4000 和 Catalyst 6500/6000 上的所有端口默认情况下都使用 PAgP 信道协议，因此不运行 LACP。为了配置端口以使用 LACP，您需要将模块上的信道协议设置为 LACP。LACP 和 PAgP 不能在运行 CatOS 的交换机的同一个模块上运行。
- **set port lacp-channel** *port_range admin-key*
  admin key（管理密钥）参数在 LACP 数据包中进行交换。信道只在具有同一个管理密钥的端口之间形成。**set port lacp-channel port_range admin-key 命令用于为信道分配一个管理密钥号码。show lacp-channel group 命令用于显示该号码。set port lacp-channel port_range admin-key** 命令用于对端口范围内的所有端口分配同一个管理密钥。如果没有配置特定密钥，将随机分配管理密钥。这样，您便可以在需要时引用管理密钥，以便对同一 agport 的添加和删除信道端口操作进行管理。
- **set port lacp-channel** *port_range* **mode active**
  set port lacp-channel port_range mode active 命令用于对先前已获得同一管理密钥的一组端口将信道模式更改为 active。

另外，LACP 在 LACP EtherChannel 建立后使用一个 30 秒间隔计时器 (Slow_Periodic_Time)。在收到的 LACPDU 信息失效之前使用的长超时 (3 x Slow_Periodic_Time) 秒数是 90。使用 UDLD，这是一个更快速的单向链路探测器。为了在形成信道后保持该信道，您现在不能调整 LACP 计时器，并且不能将交换机配置为使用快速 PDU 传输（每秒）。

其他选项

如果在接入层您具有一个最小管理模型，则通常的配置是在分布层和核心层将模式设置为 active。使接入层交换机采用默认的 passive 配置。

单向链路检测 (UDLD)

UDLD 是 Cisco 专有的轻量级协议，为检测设备之间的单向通信实例而开发。虽然有其他方法（如 FEFI）可检测传输介质的双向状态，但仍有某些实例，其中的 L1 检测机制不足。这些情况可能导致下列任何一种情况发生：

- 不可预测的 STP 操作
- 数据包不正确或过度泛洪
- 流量黑洞

UDLD 功能用于解决光纤和铜以太网接口上的这些故障情况：

- 监控实际布线配置，并关闭任何布线有误的端口，使其处于 errdisable 状态。
- 防范单向链路。当由于介质或端口/接口出现故障而检测到单向链路时，受到影响的端口将关闭并处于 errdisable 状态，同时还将生成相应的 syslog 消息。
- 此外，UDLD 主动模式检查以前视为双向的链路在拥塞时不会失去连接及变得不可用。UDLD 在整个链路上执行持续的连接测试。UDLD 主动模式的主要目的在于避免在某些故障情况下出

现流量黑洞。

由于生成树具有状态稳定的单向 BPDU 流，因此深受这些故障的影响。很容易便会看到某个端口突然无法传输 BPDU，从而导致邻居上的 STP 状态从 blocking 变为 forwarding。由于端口仍能接收，因而该变化会产生环路。

操作概述

UDLD 是在 LLC 层以上运行的第 2 层协议（目标 MAC 01-00-0c-cc-cc-cc，SNAP HDLC 协议类型 0x0111）。当与 FEFI 和自动协商第 1 层机制一同运行 UDLD 时，可以验证链路的物理（第 1 层）和逻辑（第 2 层）完整性。

UDLD 提供了 FEFI 和自动协商不能执行的功能和保护（即邻居信息检测和缓存），能够关闭所有错误连接端口并检测非点对点链路（那些遍历介质转换器或集线器）上的逻辑接口/端口故障。

UDLD 使用了两个基本机制；它获得有关邻居的信息，并将最新信息保存在本地缓存中，而且只要当它检测到新的邻居或者邻居请求缓存重新同步时，它就会发送一系列 UDLD probe/echo (hello) 消息。

UDLD 经常对启用 UDLD 的所有端口发送探测消息。只要某个端口收到特定的“触发”UDLD 消息，便会开始检测阶段和验证过程。如果在此过程结束时满足所有有效条件，则端口状态不会更改。为了满足这些条件，端口必须为双向且正确地进行布线。否则，端口为 errdisable，并显示一条 syslog 消息。Syslog 消息类似于下列消息：

- `UDLD-3-DISABLE Unidirectional link detected on port [dec]/[dec].Port disabled`
- `UDLD-4-ONEWAYPATH A unidirectional link from port [dec]/[dec] to port[dec]/[dec] of device [chars] was detected`

有关设备发出的完整的系统消息（包括 UDLD 事件）列表，请参阅消息和恢复过程（Catalyst 系列交换机，7.6）。

建立链路并归类为双向后，UDLD 继续以默认的 15 秒间隔公布 probe/echo 消息。此表表示在 **show udld port** 命令的输出中报告的有效 UDLD 链路状态：

端口状态	注释
未确定	正在进行检测，或者已禁用一个相邻的 UDLD 实体或已阻止其传输。
不适用	UDLD 已禁用。
shutdown	已检测到单向链路并禁用此端口。
双向	已检测到双向链路。

- **邻居缓存维护** — UDLD 为了维护 UDLD 邻居缓存的完整性，对每个活动接口都定期发送 hello probe/echo 数据包。无论何时收到 hello 消息，该消息都将存入缓存并在内存中保存定义为“保持时间”的一段最长期限。当保持时间过期时，相应的缓存条目便会老化。如果在保持时间期限内收到新的 hello 消息，此新消息便会取代旧条目并重置对应的生存时间计时器。
- 为了维护 UDLD 缓存的完整性，每当启用 UDLD 的接口被禁用或重置设备时，会清除受到配置更改影响的所有现有缓存条目，UDLD 传送至少一条消息，通知相应邻居刷新对应的缓存条目。
- **回声检测机制** — 回声机制构成了检测算法的基础。每当UDLD设备获得新邻居信息或者从失步邻居收到再次同步请求时，它会在连接端启动/重新启动检测窗口，同时发送响应信息进行回复。因为此行为在所有邻居中肯定都是相同的，所以回声发送人希望在应答中收到回声。如果检测窗口结束时没有收到有效的应答消息，则会将链路视为单向，并会触发链路重建或端口关闭进程。

收敛时间

为防止 STP 环路，CatOS 5.4(3) 将 UDLD 默认消息间隔从 60 秒降低到 15 秒，以便在阻塞端口可以转换为转发状态前关闭单向链路。

注意： 消息间隔值确定了邻居在联结或检测阶段之后发送 UDLD 探测的速率。消息间隔在链路的两端不需要匹配，虽然在可能的情况下最好配置一致。当 UDLD 邻居建立时，将发送已配置的消息间隔，并且将该对等体的超时间隔计算为 (3 * message_interval)。所以，在丢失三次连续的 hello（或探测）之后，对等关系便会超时。由于每一端的消息间隔不同，因此该超时值在每一端也是不同的。

UDLD 检测单向故障所需的时间大约为 (2.5 * message_interval + 4 秒)，或者大约为 41 秒（使用 15 秒的默认消息间隔）。这远低于 STP 重新收敛通常需要的 50 秒。如果 NMP CPU 有一些备用的循环并且您仔细监控其使用级别，则可以使消息间隔降低（甚至）至 7 秒的最小限度。此消息间隔帮助通过重要因素加快检测速度。

因此，UDLD 在默认生成树计时器上具有假定相关性。如果调整 STP 以便比使用 UDLD 更快收敛，请考虑采用替代机制，例如 CatOS 6.2 环路防护功能。另外，当您实现 RSTP (IEEE 802.1w) 时，请考虑采用替代机制，因为 RSTP 具有取决于拓扑的收敛特性（以毫秒为单位）。对于这些实例，请将环路防护与 UDLD 结合使用，这将提供大多数防护。环路防护防止 STP 环路具有正在使用的 STP 版本的速度，UDLD 检测各个 EtherChannel 链路上或 BPDU 不沿断开方向传输的情况下出现的单向连接。

注意： UDLD 并不捕获每种 STP 故障情况，例如超过 (2 * FwdDelay + Maxage) 时间不发送 BPDU 的 CPU 造成的故障。为此，Cisco 建议您在依赖 STP 的拓扑中与环路防护（在 CatOS 6.2 中引入）一起实现 UDLD。

警告： 对使用不可配置的 60 秒默认消息间隔的较早版本的 UDLD，请务必小心。这些版本易受生成树环路影响。

UDLD 主动模式

主动 UDLD 是专为解决那些（几种）需要对双向连接执行持续测试的情况而创建的。这样，主动模式功能在下列情况下便可针对危险的单向链路提供增强的防护：

- 当 UDLD PDU 对称丢失并且两端都超时时，则任何一个端口的状态都不为 errdisable。
- 链路的一端出现端口阻塞（传输 [Tx] 和 Rx）。
- 链路的一端保持接通状态，而另一端却已经关闭。
- 禁用自动协商或另一个第 1 层故障检测机制。
- 需要减少对第 1 层 FEFI 机制的信赖。
- 需要对点对点 FE/GE 链路提供针对单向链路故障的最大防护。特别是在两个邻居之间不允许存在故障的情况下，可将 UDLD 主动探测视为“检测信号”，该信号的存在可确保链路正常运行。

实施主动 UDLD 的最常见的情况是为了在自动协商或其他第 1 层故障检测机制禁用或不可用时对链路捆绑的某个成员执行连接检查。这对 EtherChannel 连接来说确实如此，因为即使启用了 PAgP/LACP，在稳定状态下也不使用非常低的 hello 计时器。在这种情况下，主动 UDLD 具有防止可能出现生成树环路的额外益处。

造成 UDLD 探测数据包对称丢失的情况更不易于区分。您必须了解正常 UDLD 会检查单向链路情况，即使在链路到达双向状态后也会如此。UDLD 的用途在于检测引起 STP 环路的第 2 层问题，并且那些问题通常是单向的，因为在稳定状态下 BPDU 仅沿一个方向流动。因此，将常规 UDLD 与自动协商和环路防护（用于依赖 STP 的网络）一起使用通常便足够了。然而，在两个方向都受到拥

塞同等影响从而导致在两个方向都失去 UDLD 探测的情况下，使用 UDLD 主动模式非常有益。例如，如果链路每一端的 CPU 使用率都提升，则会出现 UDLD 探测丢失的情况。双向连接丢失的其他示例包括下列这些设备之一出现的故障：

- 密集波分复用 (DWDM) 转发器
- 介质转换器
- 集线器
- 另一个第 1 层设备**注意：** 自动协商无法检测到故障。

主动 UDLD 错误禁用处于这些故障情况下的端口。当您对非点对点链路启用 UDLD 主动模式时，请仔细考虑后果。具有介质转换器、集线器或类似设备的链路是非点对点链路。中间设备可防止 UDLD 数据包转发以及在不必要时强制关闭链路。

在端口的所有邻居都老化之后，UDLD 主动模式（如果已启用）尝试重新启动联结顺序以与所有潜在不同步邻居重新同步。此尝试在公告或检测阶段发生。如果一串快速消息发出后（8 次失败的重新尝试），链路仍被视为“未确定”，则将端口置于 errdisable 状态。

注意： 某些交换机不支持主动 UDLD。目前，Catalyst 2900XL 和 Catalyst 3500XL 具有 60 秒的硬编码消息间隔。一般认为此间隔过长，不足以防止可能出现的 STP 环路（使用默认 STP 参数的情况下）。

路由链路上的 UDLD

为便于此讨论，路由链路为两种连接类型之一：

- 两个路由器节点之间点对点此链路通过 30 位子网掩码进行配置。
- 具有多个端口但仅支持路由连接的 VLAN以分割第 2 层核心拓扑为例。

每个内部网关路由协议 (IGRP) 都具有与其如何处理邻居关系和路由收敛相关的特性。当您将目前用到的两个较常见的路由协议（开放最短路径优先 (OSPF) 协议和增强 IGRP (EIGRP)）进行对比时便可发现本部分讨论的特性彼此相关。

首先，请注意在任何点对点路由网络上的第 1 层或第 2 层故障都会导致几乎立即拆卸第 3 层连接。由于在发生第 1 层/第 2 层故障时该 VLAN 中的唯一交换机端口转换为非连接状态，因此 MSFC Autostate 功能会在大约两秒的时间内同步第 2 层和第 3 层端口状态。此同步使第 3 层 VLAN 接口处于打开/关闭状态（在线路协议关闭的情况下）。

假定采用默认的计时器值。OSPF 每隔 10 秒发送 hello 消息且停顿间隔为 40 秒 (4 * hello)。这些计时器对 OSPF 点对点网络和广播网络是一致的。由于 OSPF 需要双向通信以便形成邻接关系，因此最坏情况下的故障切换时间是 40 秒。此故障切换情况如下，即使第 1 层/第 2 层故障不完全在点对点连接上发生，此情况也会造成第 3 层协议必须处理的半可操作情形。由于 UDLD 的检测时间非常类似于到期的 OSPF 停机计时器的时间（大约 40 秒），因此对 OSPF 第 3 层点对点链路配置 UDLD 正常模式的优势是有限的。

在许多情况下，EIGRP 比 OSPF 收敛的速度更快。然而，您必须记住，不必采用双向通信也可使邻居交换路由信息。在非常特定的半可操作故障情形下，EIGRP 易受持续不断的流量黑洞的攻击，直到某些其他事件通过使邻居“活动”建立路由为止。UDLD 正常模式能缓和本部分提到的这些情况。UDLD 正常模式检测到单向链路故障，并且错误使端口禁用。

对于使用任何路由协议的第 3 层路由连接，UDLD 正常模式仍可防止其不受初始链路激活时的问题的影响。此类问题包括布线错误或硬件故障。另外，UDLD 主动模式对第 3 层路由连接也提供这些优点：

- 防止不必要的流量黑洞**注意：** 在某些情况下需要最少的计时器。
- 将一个抖动链路置于 errdisable 状态
- 防止出现第 3 层 EtherChannel 配置导致的环路

UDLD 的默认行为

默认情况下，UDLD 全局禁用并准备在光纤端口上启用。由于 UDLD 是只在交换机之间才需要的基础结构协议，因此默认情况下在铜线端口上禁用 UDLD。铜线端口往往用于主机访问。

注意： 在邻居能够达到双向状态之前，必须全局启用 UDLD 并将其置于接口级别。在 CatOS 5.4(3) 及更高版本中，默认消息间隔为 15 秒并可在 7 到 90 秒之间进行配置。

默认情况下，全局禁用 Errdisable 恢复。在全局启用之后，如果端口进入 errdisable 状态，则在选定时间间隔后该端口将自动重新启用。默认时间为 300 秒，这是一个全局计时器且为交换机中的所有端口保持。如果将端口的 errdisable 超时设置为 disable，则可以手动阻止该端口重新启用。**发出 set port errdisable-timeout mod/port disable 命令。**

注意： 是否使用此命令取决于您的软件版本。

当您在不具备带外网络管理功能的情况下实现 UDLD 主动模式时，特别是在接入层或在发生 errdisable 情况时会从网络隔离的任何设备上，请考虑使用 errdisable 超时功能。

有关为处于 errdisable 状态的端口配置超时时间段的详细信息，请参阅配置以太网、快速以太网、千兆以太网和 10 千兆以太网交换。

建议

如果您正确使用 UDLD 正常模式并与适当的功能和协议结合使用，则 UDLD 正常模式可满足大多数情况的需要。这些功能/协议包括：

- FEFI
- 自动协商
- 环路防护

部署 UDLD 时，请考虑是否有必要对双向连接（主动模式）执行持续测试。一般情况下，如果启用了自动协商，则不必使用主动模式，因为自动协商可针对第 1 层的故障检测进行纠正。

对于 UDLD 消息间隔设置为默认值 15 秒的 Cisco 交换机之间的所有点对点 FE/GE 链路，Cisco 建议启用 UDLD 正常模式。此配置采用默认的 802.1d 生成树计时器。另外，在依赖 STP 实现冗余和收敛的网络中，请将 UDLD 与环路防护一起使用。此建议适用于拓扑中的一个或多个端口处于 STP 阻塞状态的网络。

发出以下命令以启用 UDLD：

```
set udld enable
!--- After global enablement, all FE and GE fiber !--- ports have UDLD enabled by default. set
udld enable port range
!--- This is for additional specific ports and copper media, if needed.
```

对于因为单向链路问题而被错误禁用的端口，必须手动启用。发出 **set port enable** 命令。

有关详细信息，请参阅了解和配置单向链路检测协议 (UDLD) 功能。

其他选项

为了尽量防止单向链路导致的问题，请配置主动模式 UDLD：

```
set udld aggressive-mode enable port_range
```

另外，可以在支持的情况下，在每一端将 UDLD 消息间隔值在 7 到 90 秒之间进行调整，以便加速收敛：

```
set udld interval time
```

对于在发生 errdisable 情况时会从网络隔离的任何设备，请考虑使用 errdisable 超时功能。对于接入层，当您在不具备带外网络管理功能的情况下实现 UDLD 主动模式时，这种情况通常适用。

如果将端口置于 errdisable 状态，则默认情况下该端口保持关闭。您可以发出以下命令，它可在超时间隔过后重新启用端口：

注意： 默认情况下超时间隔是 300 秒。

```
>set errdisable-timeout enable ?
bpdu-guard
!--- This is BPDU port-guard. channel-misconfig !--- This is a channel misconfiguration. duplex-
mismatch udld other !--- These are other reasons. all !--- Apply errdisable timeout to all
reasons.
```

如果合作伙伴设备（例如终端主机或路由器）不支持 UDLD，请不要运行此协议。发出以下命令：

```
set udld disable port_range
```

测试和监控 UDLD

如果在实验室中不存在确实有故障或者单向的组件（如发生故障的 GBIC），就不太容易测试 UDLD。此协议设计用来检测的故障情形要比实验室中经常出现的故障情形更为少见。例如，如果您想执行一个简单测试，并且拔掉一个光纤束以便观察所需的 errdisable 状态，则需要关闭第 1 层自动协商。否则，物理端口会进入 down 状态，这会重置 UDLD 消息通信。在 UDLD 正常模式下，远端会出现未确定状态。如果使用 UDLD 主动模式，远端会出现 errdisable 状态。

另外，还有一个用于模拟 UDLD 的邻居 PDU 丢失的测试方法。为了阻止 UDLD/CDP 硬件地址，但是允许其他地址通过，请使用 MAC 层过滤器。

为了监控 UDLD，请发出下列命令：

```
>show udld

UDLD              : enabled
Message Interval  : 15 seconds


>show udld port 3/1

UDLD              : enabled
Message Interval  : 15 seconds
Port      Admin Status  Aggressive Mode  Link State
--------  ------------  ---------------  ----------------
3/1       enabled       disabled         bidirectional
```

同样是从 enable 模式，您可以发出 show udld neighbor 隐藏命令来检查 UDLD 缓存内容（以与 CDP 相同的方式）。若要验证是否存在特定于协议的异常，将 UDLD 缓存与 CDP 缓存进行比较通

常是有用的。每当 CDP 也受到影响时，所有 PDU/BPDU 通常都会受到影响。因此，还需检查 STP。例如，请检查是否存在最近进行的根身份更改或根/指定端口放置更改。

```
>show udld neighbor 3/1

Port  Device Name            Device ID    Port-ID OperState
----- --------------------  ------------ ------- ----------
3/1   TSC07117119M(Switch)   000c86a50433 3/1     bidirectional
```

此外，您可以使用 Cisco UDLD SNMP MIB 变量来监控 UDLD 状态和配置一致性。

超巨型帧

对于所有以太网端口（包括 GE 和 10 GE）而言，默认的最大传输单元 (MTU) 帧大小为 1518 字节。超巨型帧功能使接口能够交换大于标准以太网帧大小的帧。若要优化服务器到服务器性能，以及支持可增加原始帧大小的应用程序（如多协议标签交换 (MPLS)、802.1Q 隧道和第 2 层隧道协议版本 3 (L2TPv3)），此功能非常有用。

操作概述

对于常规帧，IEEE 802.3 标准规范定义的最大以太网帧大小为 1518 字节，对于 802.1Q 封装帧，则为 1522 字节。802.1Q 封装帧有时称为“小巨型帧”。一般而言，当数据包超出特定以太网连接的指定以太网最大长度时，会将该数据包归类为巨型帧。巨型数据包也称为“超巨型帧”。

至于为什么某些帧的 MTU 大小可能超出 1518 字节，原因有很多。下面是一些示例：

- 供应商特定的要求 — 应用程序和某些 NIC 可能指定了超出标准的 1500 字节的 MTU 大小。之所以倾向于指定这种 MTU 大小，是因为有研究证明增加太网帧的大小可以增加平均吞吐量。
- 中继 — 为了在交换机或其他网络设备之间传递 VLAN ID 信息，使用中继来增大标准的以太网帧。目前，最为常见的两种中继形式是 Cisco 专有的 ISL 封装和 IEEE 802.1Q。
- MPLS —在MPLS在接口后启用，有可能性增添数据包的帧大小。这种增大取决于 MPLS 标记数据包的标签栈中的标签数量。一个标签的总大小为 4 字节。一个标签栈的总大小为 *n* x 4 字节。如果形成了标签堆栈，则帧数可能会超过 MTU。
- 802.1Q 隧道 — 802.1Q 隧道数据包包含两个 802.1Q 标记，通常每次仅有一个标记为硬件可见。因此，内部标记会对 MTU 值（有效负载大小）添加 4 字节。
- 通用传输接口 (UTI)/L2TPv3 — UTI/L2TPv3 可封装将通过 IP 网络转发的第 2 层数据。封装可将原始帧大小增大多达 50 字节。新的帧包括一个新的 IP 报头（20 字节）、一个 L2TPv3 报头（12 字节）和一个新的第 2 层报头。L2TPv3 有效负载包括完整的第 2 层帧，而第 2 层帧包括第 2 层报头。

不同 Catalyst 交换机支持各种帧大小的能力取决于多种因素，这些因素包括硬件和软件。即使是在相同的平台中，某些模块也可支持比其他模块更大的帧大小。

- 在运行 CatOS 6.1 版本时，Catalyst 5500/5000 交换机支持超巨型帧。在端口上启用超巨型帧功能后，MTU 大小可增加到 9216 字节。在基于 10/100 Mbps 非屏蔽双绞线 (UTP) 的板卡上，支持的最大帧大小仅为 8092 字节。此限制是由 ASIC 限制导致的。一般而言，在启用超巨型帧大小功能方面没有限制。您可以将此功能与中继/非中继和信道/非信道技术一起使用。
- 由于存在 ASIC 限制，Catalyst 4000 交换机（Supervisor 引擎 1 [WS-X4012] 和 Supervisor 引擎 2 [WS-X4013]）不支持超巨型帧。然而，802.1Q 中继例外。
- 运行 CatOS 版本 6.1(1) 及更高版本时，Catalyst 6500 系列平台可支持超巨型帧大小。但是，这种支持与所使用的板卡类型有关。一般而言，在启用超巨型帧大小功能方面没有限制。您可以将此功能与中继/非中继和信道/非信道技术一起使用。在单个端口上启用超巨型帧支持后

，默认的 MTU 大小为 9216 字节。默认的 MTU 不可使用 CatOS 配置。**不过，Cisco IOS 软件版本 12.1(13)E 引入了 system jumbomtu 命令来覆盖默认的 MTU。**
有关详细信息，请参阅 Catalyst 交换机上支持超巨型帧/巨型帧配置示例。

此表介绍了 Catalyst 6500/6000 系列交换机的不同板卡支持的 MTU 大小：

注意： MTU 大小或数据包大小仅指以太网有效负载。

线路卡	**MTU 大小**
默认	9216 字节
WS-X6248-RJ-45、WS-X6248A-RJ-45、WS-X6248-TEL、WS-X6248A-TEL、WS-X6348-RJ-45(V)、WS-X6348-RJ-21(V)	8092 字节（受 PHY 芯片限制）
WS-X6148-RJ-45(V)、WS-X6148-RJ-21(V)、WS-X6148-45AF、WS-X6148-21AF	9100个字节(@100Mbps) 9216个字节(@10Mbps)
WS-X6148A-RJ-45、WS-X6148A-45AF、WS-X6148-FE-SFP	9216 字节
WS-X6324-100FX-MM、-SM、WS-X6024-10FL-MT	9216

	字节
Supervisor 引擎 1、2、32 和 720 的 WS-X6548-RJ-45、WS-X6548-RJ-21、WS-X6524-100FX-MM、WS-X6148X2-RJ-45、WS-X6148X2-45AF、WS-X6196-RJ-21、WS-X6196-21AF、WS-X6408-GBIC、WS-X6316-GE-TX、WS-X6416-GBIC、WS-X6516-GBIC、WS-X6516A-GBIC、WS-X6816-GBIC 上行链路	9216 字节
WS-X6516-GE-TX	8092个字节(@100 Mbps) 9216个字节(@10或1000 Mbps)
WS-X6148-GE-TX、WS-X6148V-GE-TX、WS-X6148-GE-45AF、WS-X6548-GE-TX、WS-X6548V-GE-TX、WS-X6548-GE-45AF	1500字节（不支持超巨型帧）
WS-X6148A-GE-TX、WS-X6148A-GE-45AF、WS-X6502-10GE、WS-X67xx 系列	9216 字节
OSM ATM (OC12c)	9180 字节
OSM CHOC3、CHOC12、CHOC48、CT3	9216个

	字节(OCx和DS3) 7673个字节(T1/E1)
Flex WAN	7673个字节(CT3 T1/DS0) 9216个字节(OC3c POS) 7673个字节(T1)
CSM (WS-X6066-SLB-APC)	9216字节(自CSM 3.1(5)和3.2(1)起

	)
OSM POS OC3c、OC12c、OC48c；OSM DPT OC48c、OSM GE WAN	9216 字节

第 3 层超巨型帧支持

凭借运行于 Supervisor 引擎上的 CatOS 和运行于 MSFC 上的 Cisco IOS 软件，通过使用 PFC/MSFC2、PFC2/MSFC2 或更高版本硬件，在运行 Cisco IOS® 软件版本 12.1(2)E 和更高版本时，Catalyst 6500/6000 交换机还可支持第 3 层超巨型帧。如果入口和出口 VLAN 都针对超巨型帧进行了配置，则所有数据包都可由 PFC 以线速进行基于硬件的交换。如果入口 VLAN 针对超巨型帧进行了配置，而出口 VLAN 未进行配置，则会发生两种情形：

- 由终端主机在设置了不分段 (DF) 位（针对路径 MTU 发现）的情况下发送的超巨型帧 — 丢弃数据包，向该终端主机发送 Internet 控制消息协议 (ICMP) 不可到达的消息，消息代码为 fragment needed and DF set。
- 由终端主机在未设置 DF 位的情况下发送的超巨型帧 — 数据包传送到 MSFC2/MSFC3 以在软件中进行分段和交换。

此表概述了针对多种平台的第 3 层超巨型帧支持：

第 3 层交换机或模块	**第 3 层最大 MTU 大小**
Catalyst 2948G-L3/4908G-L3 系列	不支持超巨型帧。
Catalyst 5000 RSM[1]/RSFC[2]	不支持超巨型帧。
Catalyst 6500 MSFC1	不支持超巨型帧。
Catalyst 6500 MSFC2 及更高版本	Cisco IOS 软件版本 12.1(2)E：9216 字节

[1] RSM =路由交换模块

[2] RSFC =路由交换机功能卡

网络性能方面

现已对 WAN (Internet) 上的 TCP 性能进行了广泛的研究。此等式解释了 TCP 吞吐量的上限是如何由以下因素决定的：

- 最大数据段大小 (MSS)，即 MTU 长度减去 TCP/IP 报头长度
- 往返时间 (RTT)
- 数据包丢失率

$$Throughput \leq\sim 0.7 \times MSS / \left(RTT \times \sqrt{packet\_loss}\right)$$

根据此公式，可达到的最大 TCP 吞吐量与 MSS 成正比。基于常量 RTT 和丢包率，如果您将数据包大小增大一倍，则 TCP 吞吐量也会增大一倍。同样，在使用超巨型帧而不使用 1518 字节帧的情况下，数据包大小增加六倍有可能使以太网连接上的 TCP 吞吐量也增加六倍。

其次，因为服务器群的性能需求持续增长，要求采用更有效的手段来确保网络文件系统 (NFS) UDP 数据报实现更高的数据速率。NFS 是用来在基于 UNIX 的服务器之间传输文件的应用最为广泛的数据存储机制，它具有 8400 字节的数据报。假设以太网具有扩展的 9 KB MTU，则单个超巨型帧的大小足以传送一个 8 KB 的应用程序数据报（例如，NFS）以及数据包报头开销。因为不再需要软件来将 NFS 块分段到单独的 UDP 数据报中，所以此功能偶尔会允许在主机上进行更为高效的直接内存访问 (DMA) 传输。

建议

当您需要超巨型帧支持时，请仅对其内所有交换机模块（第 2 层）和接口（第 3 层）支持超巨型帧的网络区域使用超巨型帧。此配置可防止在路径上的任何位置进行分段。如果配置了大于路径中支持的帧长度的超巨型帧，则会导致因使用该功能而实现的所有改进都完全消失，原因是需要分段。正如此超巨型帧部分的表中所示，就支持的最大数据包大小而言，不同的平台和板卡各不相同。

对于支持超巨型帧的主机设备所在的整个第 2 层 VLAN，请将这些主机设备的 MTU 大小配置为网络硬件支持的最小公分母。若要对可支持超巨型帧的模块启用超巨型帧支持，请发出以下命令：

```
set port jumbo mod/port enable
```

另外，如果希望跨第 3 层边界实现超巨型帧支持，请在所有适用的 VLAN 接口上配置最大的可用 MTU 值，即 9216 字节。在 VLAN 接口下发出 **mtu** 命令：

```
interface vlan vlan# mtu 9216
```

此配置可确保这些模块支持的第 2 层超巨型帧 MTU 始终小于或等于为流量所经过的第 3 层接口配置的值。这样可以防止流量从 VLAN 跨第 3 层接口路由时进行分段。

管理配置

本部分讨论如何帮助控制和配置 Catalyst 网络以及对其进行故障排除的注意事项。

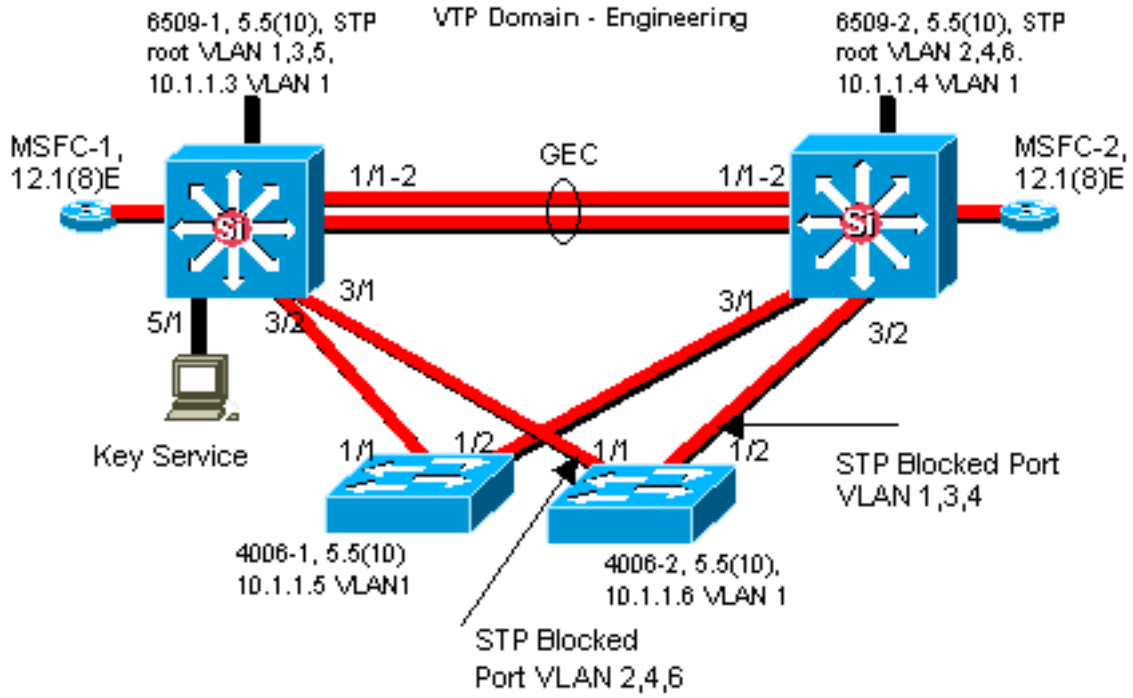
网络图

清晰的网络图是网络运行的一个基本部分。网络图在进行故障排除时非常重要，而且在由于发生中断而向供应商和合作伙伴上报时，网络图也是最为重要的信息交流工具。切不可低估网络图制备、其随时可用性以及可访问性的重要性。

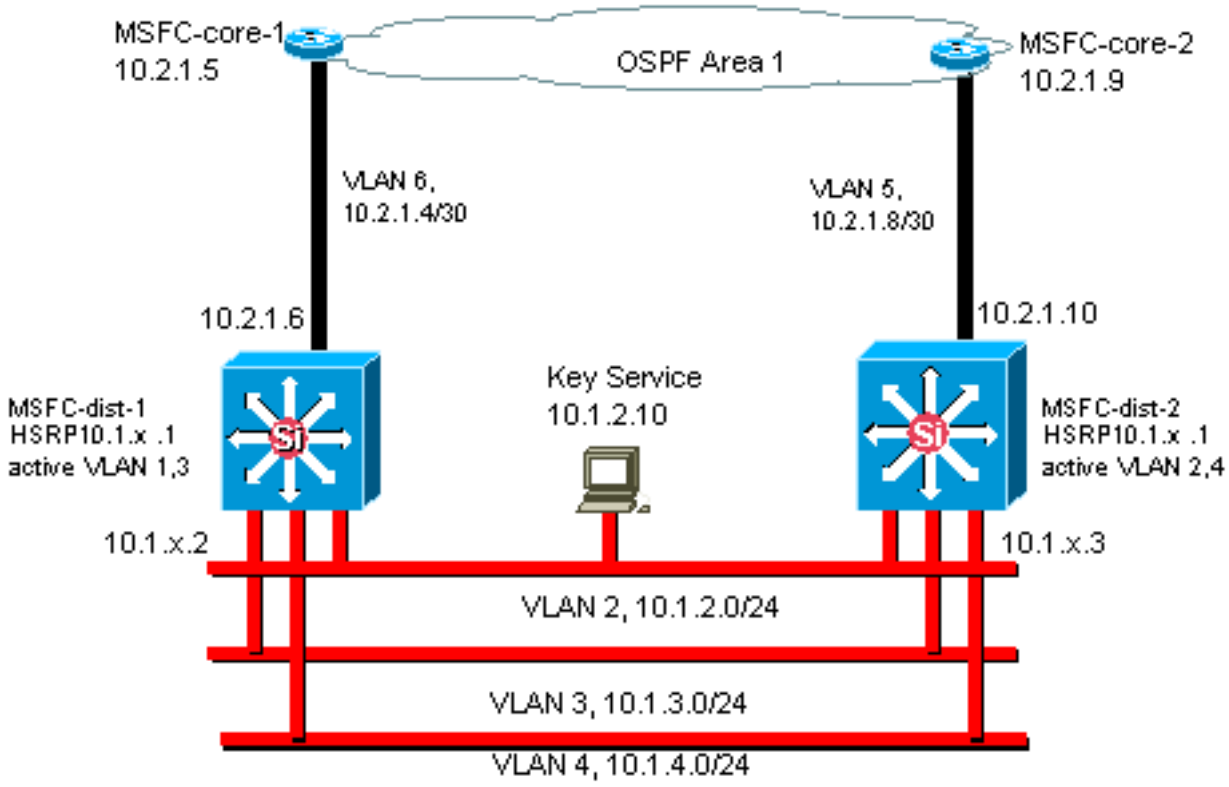
建议

Cisco 建议您创建以下三个图表：

- **整体图表** — 即使对于最大型的网络，显示端到端的物理和逻辑连接性的图表也非常重要。对于已经实施分层设计来分别说明每个层的企业来说，此图可能较为常见。但是在规划和解决问题时，它通常能够提供有关域的链接方式的有用信息，这些信息非常重要。
- **物理图表** — 显示所有交换机和路由器硬件以及布线。每个 VLAN 都必须标明中继、链路、速度、信道组、端口号、插槽、机箱类型、软件、VTP 域、根网桥、备用根网桥优先级、MAC 地址和阻塞端口。如将内部设备（例如 Catalyst 6500/6000 MSFC）描述为通过中继连接的枝干上的路由器，此图表通常会显得更为清晰。



- **逻辑图表** — 仅显示第 3 层功能（路由器充当对象，VLAN 充当以太网段）。必须标明 IP 地址、子网、辅助编址、活动和备用 HSRP、接入核心分布层和路由信息。



带内管理

根据配置的不同，交换机带内（内部）管理接口（称为 sc0）可能必须处理以下数据：

- 交换机管理协议，如 SNMP、Telnet、Secure Shell协议 (SSH) 和 syslog
- 用户数据，如广播和多播
- 交换机控制协议，如 STP BPDU、VTP、DTP、CDP 等

Cisco 多层设计的常见做法是配置跨交换域并且包含所有 sc0 接口的管理 VLAN。这有助于将管理流量与用户流量分开，并提高交换机管理接口的安全性。本部分介绍了使用默认 VLAN 1 和在同一

VLAN 中将通往交换机的管理流量作为用户流量运行的重要性和潜在问题。

操作概述

在使用VLAN1获得用户数据时，我们最关心的问题是，Supervisor Engine NMP不需要被终端站产生的许多组播和广播流量中断。尽管这一原则适用于所有 Supervisor 引擎，但是对于旧的 Catalyst 5500/5000 硬件，尤其是 Supervisor 引擎 I 和 Supervisor 引擎 II，处理此流量的资源有限。如果 Supervisor 引擎的 CPU、缓冲区或通向背板的带内信道被完全占用，用来监听不必要的流量，则可能会错过控制帧。在最坏的情况下，这可能会导致生成树环路或 EtherChannel 故障。

如果在 Catalyst 上发出 show interface 和 show ip stats 命令，则这些命令可以提供一些关于广播与单播流量的比例、IP 与非 IP 流量的比例的指示信息，而在管理 VLAN 中通常不会看到这些信息。

对旧版 Catalyst 5500/5000 硬件运行状况的进一步检查是针对资源错误 (RscrcErrors) 检查 **show inband/biga**（隐藏命令）的输出，类似于路由器中的缓冲区丢弃。如果这些资源错误不断增加，内存将不可用于接收系统数据包，原因可能是管理 VLAN 中的广播数据流过大。一个资源错误可能表示 Supervisor 引擎无法处理某个数据包（如 BPDU），但它可能迅速变为问题，因为协议（如生成树）不会重新发送丢失的 BPDU。

建议

如本文档 Cat 控制部分强调的那样，VLAN 1 是用于标记和处理大部分控制层面数据流的特殊的 VLAN。默认情况下，在所有中继上启用 VLAN 1。对于较大的园区网络，应注意 VLAN 1 **STP 域**的直径；网络某一部分的不稳定性可能会影响 VLAN 1，从而影响控制层面稳定性，以及所有其他 VLAN 的 STP 稳定性。在 CatOS 5.4 和更高版本中，可以使用以下命令限制 VLAN 1 传送用户数据和运行 STP：

```
clear trunk mod/port vlan 1
```

从网络分析程序的角度来看，此命令不会停止对从 VLAN 1 中一个交换机发送到另一个交换机的数据包的控制。但不通过此链路转发任何数据，也不运行 STP。因此，这种技术可用于将 VLAN 1 分为较小的故障域。

注意： 目前不能在 3500 和 2900XL 上清除 VLAN 1 中继。

在园区设计过程中，即使小心谨慎地将用户 VLAN 限制在较小的交换机域和相对较小的故障/L3 边界内，仍然有一些客户希望采用不同的方法处理管理 VLAN，并试图使用单个管理子网覆盖整个网络。没有技术原因要求中央 NMS 应用程序必须是邻近其所管理设备的第 2 层，这也不是合格的安全参数。Cisco 建议将管理 VLAN 的直径限制为与用户 VLAN 相同的路由域结构，并考虑采用带外管理和/或 CatOS 6.x SSH 支持作为提高网络管理安全性的方式。

其他选项

但在某些拓扑中，上述 Cisco 建议有一些设计注意事项。例如，理想的通用 Cisco 多层设计是避免使用活动生成树的设计。这要求将每个 IP 子网/VLAN 限制在单个接入层交换机或交换机集群中。在这些设计中，不能有向下配置到接入层的中继。

是否创建单独的管理 VLAN 并启用中继以便在 L2 接入层和 L3 分布层之间对其进行传送，这个问题很难回答。与 Cisco 工程师进行设计审核时，有以下两个选项：

- **选项 1**：将两个或三个唯一的 VLAN 从分布层向下中继到每台接入层交换机。例如，这样可以

提供数据 VLAN、语音 VLAN 和管理 VLAN，而且仍然具有 STP 处于非活动状态的优势。（请注意，如果 VLAN 1 已从中继中清除，则需要额外的配置步骤。）在此解决方案中，为了在故障恢复期间暂时避免路由数据流产生黑洞，在设计中还需要注意几点：用于中继的 STP PortFast（CatOS 7.x 和更高版本）或带有 STP 转发功能的 VLAN Autostate 同步（高于 CatOS 5.5[9] 的版本）。

- **选项 2：**可以使用单个 VLAN 作为数据 VLAN 和管理 VLAN。事实上对许多用户而言，借助新型交换机硬件（如更强大的 CPU 和控制层面速率限制控件），以及多层设计倡导使用的相对较小的广播域，将 sc0 接口和用户数据中分开不像以前那样困难了。作出最终决定的最佳方式是：检查该 VLAN 的广播流量配置文件，并与您的 Cisco 工程师讨论交换机硬件的功能。如果管理 VLAN 确实包括该接入层交换机上的所有用户，则强烈建议使用 IP 输入过滤器以确保交换机免受用户的侵害，如本文档的安全配置部分所述。

带外管理

进一步延续上一部分的论述，可以在生产网络周围构建单独的管理基础设施，使得无论发生怎样的数据流驱动事件或控制层面事件，都能随时远程访问设备，从而提高网络管理的可用性。下面是两个典型方法：

- 使用独有的 LAN 进行带外管理
- 使用终端服务器进行带外管理

操作概述

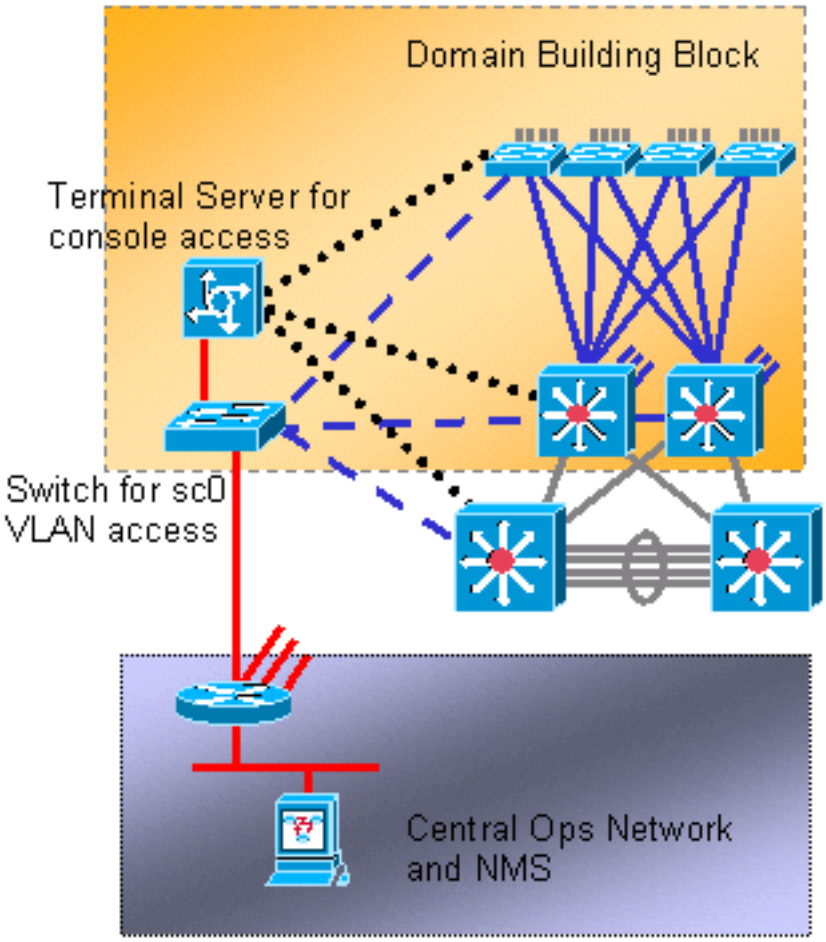
可以在管理 VLAN 上为网络中的每台路由器和交换机提供带外以太网管理接口。每个设备上的一个以太网端口在管理 VLAN 中进行配置，并通过 sc0 接口在生产网络外部连接到单独的交换式管理网络。请注意，Catalyst 4500/4000 交换机在 Supervisor 引擎上有一个特殊的 me1 接口，该接口仅用于带外管理，而不用作交换机端口。

此外，还可以通过以下配置实现终端服务器连接：使用 RJ-45 串行电缆将 Cisco 2600 或 3600 接入布局中的每台路由器和交换机的控制台端口。使用终端服务器还使您无需配置备份方案，例如在每个设备的辅助端口上配置调制解调器。可以在终端服务器的辅助端口上配置单个调制解调器，以便在发生网络连接故障期间为其他设备提供拨号服务。

建议

通过这种安排，除大量带内路径之外，还可为每台交换机和路由器提供两条带外路径，从而实现高可用性网络管理。带外负责以下事项：

- 带外将管理数据量从用户数据中分离出来。
- 带外的管理 IP 地址位于单独的子网、VLAN 和交换机中，从而实现较高的安全性。
- 带外提供在网络故障期间传递管理数据的更有效保证。
- 带外在管理 VLAN 中没有活动的生成树。冗余无关紧要。



系统测试

启动诊断

在系统启动期间，会执行一定数量的进程以确保有可靠的操作平台可以使用，从而使发生故障的硬件不会中断网络。Catalyst 启动诊断分为通电自检 (POST) 和联机诊断。

操作概述

将卡热插拔到机箱中时，会根据平台和硬件配置执行不同的启动诊断。诊断级别越高，可以检测出的问题越多，但也会产生更长的引导周期。可以选择以下三个级别的 POST 诊断（所有测试都会检查 DRAM、RAM、缓存是否存在以及大小，并对它们进行初始化）：

操作概述			
旁路	不适用	3	在使用 CatOS 5.5 或更低版本的 4500/4000 系列上不可用。
最小	仅对 DRAM 的首个 MB 进行模式写入测试。	30	在 5500/5000 和 6500/6000 系列上默认提供；在 4500/4000 系列上不可用。
完整	对所有内存进行模式写入测试。	60	在 4500/4000 系列上默认提供。

联机诊断

这些测试检查交换机内部的数据包路径。因此请注意，联机诊断是系统范围的测试，而不仅仅是端口测试，这一点很重要。在 Catalyst 5500/5000 和 6500/6000 交换机上，首先从备用 Supervisor 引擎开始执行测试，然后再从主 Supervisor 引擎执行测试。诊断的时长取决于系统配置（插槽、模块、端口的数量）。有三种测试类别：

- 回环测试 — 将来自 Supervisor 引擎 NMP 的数据包发送到每个端口，然后返回到 NMP，并检查错误。
- 捆绑测试 — 创建端口的信道（最多八个端口），并对 agport 执行回环测试以验证特定链路的散列（有关详细信息，请参阅本文档的 EtherChannel 部分）。
- 增强地址识别逻辑 (EARL) 测试 — 对中央 Supervisor 引擎和内联以太网模块 L3 重写引擎进行测试。先创建硬件转发条目和路由端口，然后通过每个模块上的交换硬件从 NMP 发送示例数据包（针对各种协议封装类型），再发送回 NMP。此测试适用于 Catalyst 6500/6000 PFC 模块和更新模块。

完整联机诊断用时约两分钟。最小诊断不对 Supervisor 引擎以外的模块执行捆绑或重写测试，大约需要 90 秒。

在内存测试过程中，如果发现读回模式与写入模式之间存在差别，端口状态将变为 faulty。如果发出 **show test** 命令后跟所检查的模块号，可以查看这些测试的结果：

```
>show test 9

Diagnostic mode: complete (mode at next reset: complete)
!--- Configuration setting. Module 9 : 4-port Multilayer Switch Line Card Status for Module 9 :
PASS Port Status : Ports 1 2 3 4 ----------------- . . . . Line Card Diag Status for Module 9 (.
= Pass, F = Fail, N = N/A) Loopback Status [Reported by Module 1] : Ports 1 2 3 4 -------------
--- . . F . !--- Faulty. Channel Status : Ports 1 2 3 4 ----------------- . . . .
```

建议

Cisco 建议将所有交换机设置为使用完整诊断，以提供最大范围的故障检测并防止正常操作期间发生中断。

注意： 在下次引导设备之前，此更改不会生效。发出以下命令可设置完整诊断：

```
set test diaglevel complete
```

其他选项

在某些情况下，与等待运行完整诊断相比，较快的启动速度可能更可取。系统的启动涉及到其他一些因素和时间，但总体而言，POST 和联机诊断会增加三分之一左右的时间。在使用带有单个 Supervisor Engine完全插满九插槽的单个Catalyst 6509机箱进行测试时，如果进行完全诊断，总启动时间约为380秒。如果进行最小诊断，启动时间约为300秒，如果绕过诊断，则启动时间仅为250秒。发出以下命令可配置绕过诊断：

```
set test diaglevel bypass
```

注意： Catalyst 4500/4000 接受配置为进行最小诊断，但这样仍会导致执行完整测试。此平台将来会支持最小模式。

运行时诊断

系统进入可操作状态后，交换机 Supervisor 引擎便会对其他模块执行各种监控。如果无法通过管理消息（在带外管理总线上运行的串行控制协议 [SCP]）到达某个模块，Supervisor 引擎将尝试重新启动该卡或采取其他适当行动。

操作概述

Supervisor 引擎自动执行各种监控；这不需要任何配置。对于 Catalyst 5500/5000 和 6500/6000，将监控交换机的下列组件：

- 通过监视器的 NMP
- 增强 EARL 芯片错误
- 从 Supervisor 引擎到背板的带内信道
- 通过经过带外信道的 keepalive 的模块 (Catalyst 6500/6000)
- 活动 Supervisor 引擎的状态由备用 Supervisor 引擎进行监控 (Catalyst 6500/6000)

系统和硬件错误检测

操作概述

在 CatOS 6.2 和更高版本中增加了更多功能，用于监控关键系统和硬件级别组件。支持以下三个硬件组件：

- 带内
- 端口计数器
- 内存

如果功能已启用并且发现错误情况，交换机将生成一个 syslog 消息。该消息在性能明显降低之前通知管理员存在问题。在 CatOS 版本 6.4(16)、7.6(12)、8.4(2) 和更高版本中，全部三个组件的默认模式从禁用变为启用。

带内

如果检测到带内错误，syslog 消息将在性能明显降低之前通知您存在问题。错误显示发生的带内故障的类型。下面是一些示例：

- 带内阻塞
- 资源错误
- 启动期间的带内故障

在检测到带内 ping 故障时，该功能还报告一条附加的 syslog 消息，包含带内连接当前 Tx 和 Rx 速率、CPU 和交换机背板负载的快照。通过此消息，您可以正确确定带内是否出现阻塞（无 Tx/Rx）或过载（Tx/Rx 过多）。此附加信息可帮助您确定带内 ping 故障的原因。

端口计数器

启用此功能时，它会创建并启动一个进程以调试端口计数器。端口计数器定期监控所选的内部端口错误计数器。线路卡的体系结构（具体来讲就是模块上的 ASIC）确定了功能查询哪些计数器。Cisco 技术支持或开发工程部门随后可使用此信息以排除问题故障。此功能不轮询与链路伙伴连通性直接关联的错误计数器，如 FCS、CRC、校准和残帧计数器。要纳入此功能，请参阅本文档的 EtherChannel/链路错误处理部分。

轮询每 30 分钟执行一次，并以所选错误计数器为背景运行。如果同一端口的两次连续轮询之间出现计数增加，syslog 消息将报告该事件并提供模块/端口和错误计数器的详细信息。

Catalyst 4500/4000 平台不支持端口计数器选项。

内存

启用此功能可对 DRAM 损坏情况执行背景监控和检测。此类内存损坏情况包括：

- 分配
- 释放
- 超出范围
- 校准错误

建议

在支持的位置启用所有错误检测功能，包括带内、端口计数器和内存。启用这些功能可改进 Catalyst 交换机平台的前瞻性系统和硬件警告诊断功能。发出以下命令以启用全部三个错误检测功能：

```
set errordetection inband enable
!--- This is the default in CatOS 6.4(16), 7.6(12), 8.4(2), and later. set errordetection
portcounters enable
!--- This is the default in CatOS 6.4(16), 7.6(12), 8.4(2), and later. set errordetection memory
enable
!--- This is the default in CatOS 6.4(16), 7.6(12), 8.4(2), and later.
```

发出以下命令以确认错误检测已启用：

```
>show errordetection

Inband error detection:                  enabled
Memory error detection:                  enabled
Packet buffer error detection:           errdisable
Port counter error detection:            enabled
Port link-errors detection:              disabled
Port link-errors action:                 port-failover
Port link-errors interval:               30 seconds
```

EtherChannel/链路错误处理

操作概述

CatOS 8.4 和更高版本中引入了一项新功能，用于提供数据流从 EtherChannel 中一个端口到同一 EtherChannel 中另一个端口的自动故障切换。当信道中某个端口在指定的时间间隔内超出可配置错误阈值时，将发生端口故障切换。只有 EtherChannel 中仍有可操作端口时，才会发生端口故障切换。如果发生故障的端口是 EtherChannel 中的最后一个端口，该端口不会进入 port-failover 状态。此端口继续传递数据流，而不考虑收到错误的类型。单个非信道端口不会进入 port-failover 状态。如果在指定时间间隔内超过错误阈值，这些端口将进入 errdisable 状态。

只有启用 **set errordetection portcounters** 时，此功能才有效。要监控的链路错误以下面三个计数器为基准：

- InErrors

- RxCRCs (CRCAlignErrors)
- TxCRCs

在交换机上发出 show counters 命令以显示错误计数器的数量。示例如下：

```
>show counters 4/48

.......

32 bit counters

0  rxCRCAlignErrors           =          0
.......

6  ifInErrors                 =          0
.......

12 txCRC                      =          0
```

下表列出了可能的配置参数及其默认配置：

参数	默认
全局	已禁用
RxCRC 的端口监控程序	已禁用
InErrors 的端口监控程序	已禁用
TxCRC 的端口监控程序	已禁用
操作	端口故障切换
间隔	30 秒
示例计数	连续 3 次
阈值下限	1000
阈值上限	1001

如果功能已启用，并且某个端口的错误计数在指定示例计数期间内达到可配置阈值的最高值，那么可配置操作为错误禁用或端口故障切换。错误禁用操作会使端口进入 errdisable 状态。如果配置端口故障切换操作，将考虑端口信道状态。只有当端口在信道，但是该端口不是信道中的最后一个可操作的端口时，才对端口执行错误禁用。另外，如果配置的动作是端口故障切换，并且端口是单个端口或非信道端口，则当端口错误计数达到阈值的高值时，端口将置于 errdisable 状态。

时间间隔是一个读取端口错误计数器的计时器常数。链路错误间隔的默认值为 30 秒。允许的范围在 30 到 1800 秒之间。

存在由于意外的一次性事件而发生端口意外错误禁用的风险。为了最大程度地降低风险，只有该情况在此连续采样次数期间内持续存在时才对端口采取操作。默认采样值为 3，允许的范围为 1 到 255。

阈值是一个根据链路错误间隔进行检查的绝对数。默认链路错误下限阈值为 1000，允许的范围是 1 到 65,535。默认链路错误上限阈值为 1001。当连续采样次数达到下限阈值时，发送 syslog。如果连续采样次数达到上限阈值，发送 syslog，并触发错误禁用或端口故障切换操作。

注意： 请对信道中的所有端口使用相同的端口错误检测配置。有关详细信息，请参阅“Catalyst 6500 系列软件配置指南”的以下部分：

- 检查状态和连通性的*配置 EtherChannel/链路错误处理* 部分
- 配置以太网、快速以太网、千兆以太网和 10 千兆以太网交换机的*配置端口错误检测* 部分

建议

由于此功能使用 SCP 消息以记录和比较数据，如果存在大量活动端口，可能会占用很多 CPU 资源。当阈值间隔设置为非常小的值时，此方案会占用更多 CPU 资源。对于被指定为关键链路并承载敏感应用程序数据流的端口，请谨慎启用此功能。发出此命令以全局启用链路错误检测：

```
set errordetection link-errors enable
```

并且，请从默认阈值、间隔和采样参数着手。并使用默认操作：端口故障切换。

发出以下命令以将全局链路错误参数应用于各个端口：

```
set port errordetection mod/port inerrors enable

set port errordetection mod/port rxcrc enable

set port errordetection mod/port txcrc enable
```

您可以发出以下命令以验证链路错误配置：

```
show errordetection

show port errordetection {mod | mod/port}
```

Catalyst 6500/6000 数据包缓冲诊断

在 CatOS 版本 6.4(7)、7.6(5) 和 8.2(1) 中，引入了 Catalyst 6500/6000 数据包缓冲诊断。默认情况下被启用的数据包缓冲诊断可检测由瞬时静态 RAM (SRAM) 故障造成的数据包缓冲故障。检测是在以下 48 端口 10/100 Mbps 线路模块上进行的：

- WS-X6248-RJ45
- WS-X6248-RJ21
- WS-X6348-RJ45
- WS-X6348-RJ21
- WS-X6148-RJ45
- WS-X6148-RJ21

当故障情况发生时，48 个 10/100 Mbps 端口中的 12 个继续保持连接，并可能遇到随机的连接性问题。从此情况恢复的唯一方法是对线路模块重新通电。

操作概述

数据包缓冲诊断检查存储在数据包缓冲区的特定部分中的数据，以确定该数据是否因为瞬时 SRAM 故障而损坏。如果此过程回读的内容与其写入的内容不同，则此过程会执行以下两个可能的可配置恢复选项：

1. 默认操作是对受缓冲故障影响的板卡端口执行错误禁用。
2. 第二个选项是对板卡重新通电。

添加了两个 syslog 消息。此消息为由于数据包缓冲错误造成的端口错误禁用或模块的重新通电提供警告：

```
show errordetection
```

```
show port errordetection {mod | mod/port}
```

在早于 8.3 和 8.4 的 CatOS 版本中，板卡重新通电时间在 30 和 40 秒之间。快速引导功能已在 CatOS 版本 8.3 和 8.4 中引入。该功能在初始引导过程中将固件自动下载到安装的板卡上，以便最大程度缩短启动时间。快速引导功能使重新通电时间降低到大约 10 秒。

建议

Cisco 推荐使用默认选项 errdisable。在生产期间，此操作对网络服务的影响最小。若可能，请将受因错误而禁用的端口影响的连接切换到其他可用的交换机端口以恢复服务。在维护窗口期间，安排板卡的手动重新通电。**发出 reset module mod 命令以从损坏的数据包缓冲区情况完全恢复。**

注意： 如果错误继续，在模块重置后，请设法再置模块。

发出此命令以启用 errdisable 选项：

```
set errordetection packet-buffer errdisable
!--- This is the default.
```

其他选项

由于板卡必须重新通电才能完全恢复遇到 SRAM 故障的所有端口，一项备选恢复操作就是配置重新通电选项。在可以接受网络服务中断持续 30 到 40 秒之间的情况下，此选项非常有用。此时间长度是为了线路模块完全重新通电并自行恢复服务而无需快速引导功能所必需的时间。通过重新通电选项，快速引导功能可以使网络服务的中断时间缩短到 10 秒。发出此命令以启用重新通电选项：

```
set errordetection packet-buffer power-cycle
```

数据包缓冲诊断

此测试仅适用于 Catalyst 5500/5000 交换机。此测试用来查找使用带有特定硬件的以太网模块的 Catalyst 5500/5000 交换机上的故障硬件，这些特定硬件在用户端口和交换机底板之间提供 10/100 Mbps 连接。由于它们无法对中继帧进行 CRC 检查，如果端口数据包缓冲区在运行时出现问题，数据包可能会被破坏并导致 CRC 错误。不幸的是，这可能导致坏帧进一步传播到 Catalyst 5500/5000 ISL 网络中，潜在地引发控制层面中断，并在最严重的情况下可能导致广播风暴。

更新的 Catalyst 5500/5000 模块和其他平台已经更新了内置的硬件错误检查，不需要数据包缓冲测试，因此没有配置它的选项。

需要数据包缓冲诊断的链路模块有 WS-X5010、WS-X5011、WS-X5013、WS-X5020、WS-X5111、WS-X5113、WS-X5114、WS-X5201、WS-X5203、WS-X5213/a、WS-X5223、WS-X5224、WS-X5506、WS-X5509、WS-U5531、WS-U5533 和 WS-U5535。

操作概述

此诊断校验存储在数据包缓冲区的特定部分的数据是否未被故障硬件意外损坏。如果过程回读的内容与写入的内容不同，该过程会在 failed 模式下关闭端口，因为此端口可能破坏数据。无需设置错误阈值。发生故障的端口不能重新启用，直到模块被重置（或替换）。

有两个模式可用于数据包缓冲测试：预定和按需。测试开始时 syslog 消息，以指示测试预期时间（计算到分钟）和测试已经开始。确切的测试长度根据端口类型、缓冲区容量大小和运行的测试类型而变化。

按需测试是积极型，以便在数分钟内完成。由于这些测试主动干扰数据包内存，因此测试之前，必须通过管理方式关闭端口。发出此命令以关闭端口：

```
> (enable) test packetbuffer 4/1
Warning: only disabled ports may be tested on demand - 4/1 will be skipped.
> (enable) set port disable 4/1
> (enable) test packetbuffer 4/1
Packet buffer test started. Estimated test time: 1 minute.
%SYS-5-PKTTESTSTART:Packet buffer test started
%SYS-5-PKTTESTDONE:Packet buffer test done. Use 'show test' to see test results
```

预定测试比按需测试被动得多，预定测试在后台执行。在多个模块上执行平行测试，但一次只对一个模块的一个端口执行。该测试可以保留、写入和读取小部分的数据包缓冲内存，然后恢复用户数据包缓冲数据，从而避免产生错误。然而，由于测试正在写入缓冲内存，它将阻拦传入数据包几毫秒，进而导致繁忙链路上的若干损失。在默认情况下，每个缓冲区写测试之间有 8 秒的暂停，可最大限度地减少数据包丢失，但这意味着，如果系统带有很多需要数据包缓冲测试的模块，可能需要 24 小时以上才能完成测试。在 CatOS 5.4 及更高版本上，默认每周星期日 03:30 开始进行预定测试，测试状态可通过以下命令确认：

```
>show test packetbuffer status

!--- When test is running, the command returns !--- this information: Current packet buffer test
details Test Type : scheduled Test Started : 03:30:08 Jul 20 2001 Test Status : 26% of ports
tested Ports under test : 10/5,11/2 Estimated time left : 11 minutes !--- When test is not
running, !--- the command returns this information: Last packet buffer test details Test Type :
scheduled Test Started : 03:30:08 Jul 20 2001 Test Finished : 06:48:57 Jul 21 2001
```

建议

Cisco 建议使用 Catalyst 5500/5000 系统的定期数据包缓冲测试功能，其优点是能够发现模块上的问题，足以抵消存在少量数据包丢失风险的缺点。

然后应该根据需要，通过允许用户从故障端口或 RMA 模块更改链路来在网络上制定标准的每周时间。由于此测试可能导致若干数据包丢失（取决于网络负载），它必须在网络使用较少的时间进行，如在星期天上午 3:30 AM（默认值）。发出此命令以设置测试时间：

```
set test packetbuffer Sunday 3:30
!--- This is the default.
```

一旦启用（当 CatOS 第一次被升级到 5.4 及更高版本），就有机会暴露以前隐藏的内存/硬件问题，结果会自动关闭端口。您可能发现此消息：

```
set test packetbuffer Sunday 3:30
!--- This is the default.
```

其他选项

如果不能接受每星期每个端口冒低水平丢包率的风险，那么建议在计划中断期间使用按需功能。发出以下命令，逐范围手动启动此功能（虽然必须首先通过管理方式禁用端口）：

```
test packetbuffer port range
```

系统日志记录

Syslog 消息是 Cisco 特定的，并是主动式故障管理的关键部分。与标准化的 SNMP 报告相比，使用 syslog 可以报告范围更广的网络和协议状况。管理平台（如 Cisco Resource Manager

Essentials (RME) 和网络分析工具包 (NATkit)）可以执行以下任务，从而充分利用 syslog 信息：

- 按严重性、消息、设备等呈现分析
- 启用对传入消息过滤以进行分析
- 触发器警告（如寻呼机）或按需收集库存和配置更改

建议

一个需要重点注意的问题是：什么级别的日志信息将在本地生成，并且存储在交换机缓冲区中，而不是哪些信息被发送到 syslog 服务器（使用 set logging server severity value 命令）。一些组织会集中记录高级信息，而其他一些组织会在交换机中查看更加详细的事件记录，或者只在故障排除期间启用更高级别的 syslog 获取。

CatOS 平台上的调试与 Cisco IOS 软件有所不同，它可以通过 set logging session enable 为每个会话启用详细的系统日志记录，而不会改变默认记录的日志。

Cisco 一般建议将生成树和系统 syslog 设备提高到第 6 层，这些是进行跟踪的关键的稳定功能。另外，对于多播环境，我们建议将 mcast 设备的日志记录级别提高为 4，以便在路由器端口被删除的情况下生成 syslog 消息。遗憾地是，在 CatOS 5.5(5) 之前，这有可能导致 IGMP 的加入和离开被记录在 syslog 消息中，这太过于繁杂以至难以监控。最后，如果使用 IP 输入列表，建议使用 4 级作为最低日志记录级别来捕获未授权的登录尝试。发出以下命令以设置这些选项：

```
set logging buffer 500
!--- This is the default. set logging server syslog server IP address set logging server enable
!--- This is the default. set logging timestamp enable
set logging level spantree 6 default
!--- Increase default STP syslog level. set logging level sys 6 default
!--- Increase default system syslog level. set logging server severity 4
!--- This is the default; !--- it limits messages exported to syslog server. set logging console
disable
```

关闭控制台消息，避免交换机在高消息流量时段因等待缓慢的终端响应或不存在的终端响应而产生的挂起现象。控制台日志记录在 CatOS 下拥有高优先级，主要用于故障排除或交换机崩溃时捕捉本地的最终消息。

此表为 Catalyst 6500/6000 提供单独的日志记录设备、默认级别和推荐的更改。根据支持的功能，每个平台的设施略有不同。

设备	默认级别	建议操作
ACL	5	保持原状。
cdp	4	保持原状。
警察	3	保持原状。
dtp	8	保持原状。
EARL	2	保持原状。
ethc[1]	5	保持原状。
filesys	2	保持原状。
gvrp	2	保持原状。
ip	2	**如果使用 IP 输入列表，请更改到 4。**
内核	2	保持原状。

1d	3	保持原状。
mcast	2	**如果使用多播（CatOS 5.5[5]及更高版本，请更改到 4）。**
mgmt	5	保持原状。
MLS	5	保持原状。
pagp	5	保持原状。
protfilt	2	保持原状。
修剪	2	保持原状。
Privatevlan	3	保持原状。
qos	3	保持原状。
radius	2	保持原状。
rsvp	3	保持原状。
安全	2	保持原状。
snmp	2	保持原状。
spantree	2	**更改到 6。**
sys	5	**更改到 6。**
TAC	2	保持原状。
tcp	2	保持原状。
telnet	2	保持原状。
Tftp	2	保持原状。
UDLD	4	保持原状。
VMPS	2	保持原状。
VTP	2	保持原状。

[1]在CatOS 7.x和以后， ethc设备代码替换pagp设备代码为了反射LACP支持。

注意： 目前，Catalyst 交换机会记录所执行的每个 **set** 或 clear 命令的配置更改 Syslog 第 6 级消息，这与 Cisco IOS 软件有所不同，IOS 软件仅在您退出配置模式后才触发该消息。如果您需要RME，在触发后实时备份配置，则需要将这些消息发送到RME系统日志服务器。对于大多数客户而言，Catalyst 交换机的定期配置备份已经足够用，不需要对默认服务器日志记录严重性进行更改。

如果您调整您的 NMS 警报，请参见系统消息指南。

简单网络管理协议 (SNMP)

SNMP 用于检索存储在网络设备管理信息库 (MIB) 中的统计数据、计数器以及表格。NMS（如 HP OpenView）可使用收集的信息来生成实时警报，测量可用性，生成容量规划信息，以及帮助执行配置和故障排除检查。

操作概述

利用一些安全机制，网络管理站能够通过 SNMP 协议的 get 和 get next 请求从 MIB 中检索信息，以及通过 **set** 命令更改参数。另外，可以将网络设备配置为针对 NMS 生成陷阱消息，从而实时发出警报。SNMP 轮询使用 IP UDP 端口 161，而 SNMP 陷阱使用端口 162。

Cisco 支持下列版本的 SNMP：

- SNMPv1 ：RFC 1157 Internet 标准，使用明文团队字符串安全性。IP 地址访问控制列表和口令可定义能够访问代理 MIB 的管理器团队。
- SNMPv2C ：SNMPv2和SNMPv2C的组合。SNMPv2是RFC 1902-1907定义的互联网标准草案，SNMPv2C是用于SNMPv2（在RFC 1901定义的试用草案）的基于社区的管理框架。优点包括：它具有批量检索机制，该机制支持检索表格和大量信息，可最大程度地减少所需的往返次数，并可改进错误处理。
- SNMPv3 ：RFC 2570 提议草案通过在网络上结合使用身份验证和数据包加密，提供对设备的安全访问。SNMPv3 中提供的安全功能包括：消息完整性：确保数据包在传输期间不会被篡改验证：确定消息的来源是否有效加密：对数据包的内容进行编码，以防止未经授权的来源轻易查看它

此表识别安全模式的组合：

模式级别	验证	加密	结果
v1	noAuthNoPriv，团队字符串	否	使用团体字符串匹配进行身份验证。
v2c	noAuthNoPriv，团队字符串	否	使用团体字符串匹配进行身份验证。
v3	noAuthNoPriv，用户名	否	使用用户名匹配进行身份验证。
v3	authNoPriv，MD5 或 SHA	Np	提供基于 HMAC-MD5 或 HMAC-SHA 算法的身份验证。
v3	authPriv，MD5 或 SHA	DES	提供基于 HMAC-MD5 或 HMAC-SHA 算法的身份验证。除了提供基于 CBC-DES (DES-56) 标准的身份验证之外，还提供 DES 56 位加密。

注意： 请记住以下有关 SNMPv3 对象的信息：

- 每个用户都属于一个组。
- 组定义一组用户的访问策略。
- 访问策略定义可以访问的 SNMP 对象以执行读取、写入和创建操作。
- 组确定其用户可以接收的通知的列表。
- 组还定义其用户的安全模式和安全等级。

SNMP 陷阱建议

SNMP 是所有网络管理的基础并且在所有网络上启用和使用。必须将交换机上的 SNMP 代理设置为使用管理站支持的 SNMP 版本。由于代理可与多个管理器通信，因此可以将软件配置为支持使用 SNMPv1 协议与一个管理站进行通信，并支持使用 SNMPv2 协议与另一个管理站进行通信。

大多数 NMS 工作站目前通过以下配置使用 SNMPv2c：

```
set snmp community read-only string
!--- Allow viewing of variables only. set snmp community read-write string
!--- Allow setting of variables. set snmp community read-write-all string<string>
```

```
!--- Include setting of SNMP strings.
```

Cisco 建议为所有正在使用的功能启用 SNMP 陷阱（如果需要，可以禁用未使用的功能）。**一旦启用陷阱，便可使用 test snmp 命令和在 NMS 上设置的相应处理对其进行测试，以查找错误（如寻呼警报或弹出项）。**

所有陷阱默认设置为禁用状态，需要单独或通过使用 **all** 参数添加到配置中，如下所示：

```
set snmp trap enable all
set snmp trap server address read-only community string
```

CatOS 5.5 中的可用陷阱包括：

陷阱	说明
验证	验证
网桥	网桥
机箱	机箱
设置	配置
实体	实体
ippermit	Ip 允许
模块	模块
中继器	中继器
stpx	生成树扩展
Syslog	Syslog 通知
VMPS	VLAN 成员策略服务器
VTP	VLAN 中继协议

注意： Syslog 陷阱还将交换机生成的所有 Syslog 消息作为 SNMP 陷阱发送到 NMS。如果分析器（例如，Cisco Works 2000 RME）已经发出了 Syslog 警报，则两次接收此信息并不一定有用。

与 Cisco IOS 软件不同的是，端口级 SNMP 陷阱默认设置为禁用状态，原因是交换机可能有数百个活动接口。因此，Cisco 建议关键端口（如基础设施到路由器、交换机和主服务器的链路）启用端口级 SNMP 陷阱。不需要其他端口（如用户主机端口）来帮助简化网络管理。

```
set port trap port range enable
!--- Enable on key ports only.
```

SNMP 轮询建议

建议执行网络管理审核以便详细地讨论具体需求。但是，下面列出了一些用于管理大型网络的基本 Cisco 准则：

- 简单易行，成效出众。
- 减少由于数据轮询、集合、工具和手动分析过量而导致的员工超负荷的情况。
- 用户只需使用几个工具，即可进行网络管理，例如：HP OpenView用于NMS、Cisco RME用于配置工具、syslog、库存、软件管理器、Microsoft Excel作为NMS数据分析程序、CGI作为发布到Web的方式。
- 通过向 Web 发布报表，用户（如高级管理人员和分析人员）可以自助获取一些信息，且不需向操作人员提出许多特殊要求。

- 弄清楚在网络上效果不错的项并将其保持原状。关注什么不适用。

NMS 实施的第一个阶段必须是为网络硬件确立基准。我们可以通过以下信息，推断出关于设备和协议状态的更多信息：路由器上的CPU、内存和缓冲区的利用率；交换机上的NMP CPU、内存和背板的利用率。仅当硬件基准执行 L2 和 L3 数据流加载之后，峰值和平均基准才完全有意义。根据公司的业务周期，通常需要用数月时间来建立基准以便了解每天、每周和每季度的趋势。

许多网络都遭遇了因轮询过度而导致的 NMS 性能下降及容量问题。因此，建议在建立基准之后在设备本身上设置警报和事件 RMON 阈值以针对异常更改向 NMS 发出警报，从而删除轮询。这使网络能够在不正常情况出现时告知操作员，而不是不断轮询，查看一切情况是否正常。可以根据各种规则设置阈值（如最大值以及百分比或与平均值的标准偏差），这些阈值不在本文档的讨论范围之内。

NMS 实施的第二个阶段是使用 SNMP 更加详细地轮询网络的特定区域。这包括可疑区域、更改之前的区域或被认定运行正常的区域。使用 NMS 系统作为探照灯来详细扫描网络并照亮热点（不要尝试照亮整个网络）。

Cisco 网络管理咨询组建议分析或在园区网络中监控这些关键故障 MIB。有关详细信息（例如，关于要轮询的性能 MIB），请参阅 Cisco 网络监控和事件相关准则。

对象名称	对象说明	OID	轮询间隔	阈值
MIB-II				
sysUpTime	系统正常运行时间（以 1/100 秒为单位）	1.3.6.1.2.1.1.3	5 分钟	< 30000

对象名称	对象说明	OID	轮询间隔	阈值
CISCO-PROCESS-MIB				
cpmCPUTotal5min	过去 5 分钟内的总 CPU 使用率	1.3.6.1.4.1.9.9.109.1.1.1.1.5	10 分钟	基准

对象名称	对象说明	OID	轮询间隔	阈值
CISCO-STACK-MIB				
sysEnableChassisTraps	指示是否必须生成此 MIB 中的 chassisAlarmOn 和 chassisAlarmOff 陷阱。	1.3.6.1.4.1.9.5.1.1.24	24 小时	1
sysEnableModuleTraps	指示是否必须生成此 MIB 中的 moduleUp 和 moduleDown 陷阱。	1.3.6.1.4.1.9.5.1.1.25	24 小时	1
sysEnableBridgeTraps	指示是否必须生成 BRIDGE-MIB (RFC 1493) 中的	1.3.6.1.4.1.9.5.1.1.26	24 小时	1

	newRoot 和 topologyChange 陷阱。			
sysEnableRepeaterTraps	指示是否必须生成 REPEATER-MIB (RFC1516) 中的陷阱。	1.3.6.1.4.1.9.5.1.1.29	24 小时	1
sysEnableIpPermitTraps	指示是否必须生成此 MIB 中的 Ip permit 陷阱。	1.3.6.1.4.1.9.5.1.1.31	24 小时	1
sysEnableVmpsTraps	指示是否必须生成在 CISCO-VLAN-MEMBERSHIP-MIB 中定义的 vmVmpsChange 陷阱。	1.3.6.1.4.1.9.5.1.1.33	24 小时	1
sysEnableConfigTraps	指示是否必须生成此 MIB 中的 sysConfigChange 陷阱。	1.3.6.1.4.1.9.5.1.1.35	24 小时	1
sysEnableStpxTrap	指示是否必须生成 CISCO-STP-EXTENSIONS-MIB 中的 stpxInconsistencyUpdate 陷阱。	1.3.6.1.4.1.9.5.1.1.40	24 小时	1
chassisPs1status	电源状态 1。	1.3.6.1.4.1.9.5.1.2.4	10 分钟	2
chassisPs1TestResult	有关电源状态 1 的详细信息。	1.3.6.1.4.1.9.5.1.2.5	必要时。	
chassisPs2Status	电源状态 2。	1.3.6.1.4.1.9.5.1.2.7	10 分钟	2
chassisPs2TestResult	有关电源状态 2 的详细信息	1.3.6.1.4.1.9.5.1.2.8	必要时。	
chassisFanStatus	机箱风扇的状态。	1.3.6.1.4.1.9.5.1.2.9	10 分钟	2
chassisFanTestResult	有关机箱风扇状态的详细信息。	1.3.6.1.4.1.9.5.1.2.10	必要时。	
chassisMinorAlarm	机箱次要警报状态。	1.3.6.1.4.1.9.5.1.2.11	10 分钟	1
chassis MajorAlarm	机箱主要警报状态	1.3.6.1.4.1.9.5.1.2.12	10 分钟	1
chassisTempAlarm	机箱温度警报状态。	1.3.6.1.4.1.9.5.1.2.13	10 分钟	1
moduleStatus	模块的操作状态。	1.3.6.1.4.1.9.5.1.3.1.1.10	30 分钟	2
moduleTestResult	有关模块状态的详细信息。	1.3.6.1.4.1.9.5.7.3.1.1.11	必要时。	

moduleStandbyStatus	冗余模块的状态。	1.3.6.1.4.1.9.5.7.3.1.1.21	30 分钟	= 1 或 = 4

对象名称	对象说明	OID	轮询间隔	阈值
CISCO-MEMORY-POOL-MIB				
dot1dStpTimeSinceTopologyChange	自上次实体检测到拓扑更改以来经过的时间（以 1/100 秒为单位）。	1.3.6.1.2.1.17.2.3	5 分钟	< 30000
dot1dStpTopChanges	自上次重置或初始化管理实体以来此网桥检测到的拓扑更改总次数。	1.3.6.1.2.1.17.2.4	必要时。	
dot1dStpPortState [1]	端口的当前状态（通过应用生成树协议而定义）。返回值可以是下列值之一：已禁用 (1)、阻塞 (2)、监听 (3)、识别 (4)、转发 (5)、或者中断 (6)。	1.3.6.1.2.1.17.2.15.1.3	必要时。	

对象名称	对象说明	OID	轮询间隔	阈值
CISCO-MEMORY-POOL-MIB				
ciscoMemoryPoolUsed	指示受管设备上的应用程序当前正在使用的内存池字节数。	1.3.6.1.4.1.9.9.48.1.1.1.5	30 分钟	基准
ciscoMemoryPoolFree	指示受管设备上当前未使用的内存池字节数。 **注意**：ciscoMemoryPoolUsed 和 ciscoMemoryPoolFree 的总和是池中的总内存量。	1.3.6.1.4.1.9.9.48.1.1.1.6	30 分钟	基准
ciscoMemoryPoolLargestFree	指示受管设备上当前未使用的内存池的最大连续字节数。	1.3.6.1.4.1.9.9.48.1.1.1.7	30 分钟	基准

有关 Cisco MIB 支持的详细信息，请参阅 Cisco 网络管理工具包 - MIB。

注意：一些标准的 MIB 假定一个特定的 SNMP 实体只包含一个 MIB 实例。因此，标准 MIB 的索引不允许用户直接访问 MIB 的特定实例。在这些情况下，提供团体字符串索引来访问标准 MIB 的

每个实例。语法是 [团体字符串] @ [实例编号]，其中实例通常是 VLAN 编号。

其他选项

SNMPv3 的安全方面意味着它的使用有望及时超越 SNMPv2。Cisco 建议客户准备将这个新协议作为他们的 NMS 策略的一部分。优点是，我们可从SNMP设备安全地收集该数据，无需惧怕被窜改或损坏。保密信息（如更改交换机配置的 SNMP **set** 命令数据包）可以进行加密，以防止其内容在网络中泄露。另外，不同的用户组可以具有不同的特权。

注意： SNMPv3 的配置与 SNMPv2 命令行有很大不同，并且预计 Supervisor 引擎上的 CPU 负载会增加。

远程监控

RMON 允许网络设备自身对 MIB 数据进行预处理，为网络管理器普遍使用或应用该信息做准备，例如，网络管理器使用该信息进行历史基准确定和阈值分析。

根据 RFC 1757中的规定，NMS 将 RMON 处理结果存储在 RMON MIB 中，以便于以后收集。

操作概述

Catalyst 交换机在每个端口上支持以硬件形式存在的微型 RMON，该微型 RMON 包括四个基本的 RMON-1 组：统计数据（组 1）、历史记录（组 2）、警报（组 3）和事件（组 9）。

RMON-1 最强大的部分是警报和事件组提供的“阈值机制”。正如已经谈及的，RMON 阈值的配置允许交换机在发生异常状况时发送 SNMP 陷阱。一旦识别了关键端口，便可使用 SNMP 轮询计数器或 RMON 历史记录组，并为那些端口创建记录正常数据流活动的基准。接下来，可以设置 RMON 上限阈值和下限阈值，并配置为在偏离基准达到定义的差异时发出警报。

因为在“警报”和“事件”表中成功创建参数行是一项繁琐的工作，所以最好是使用 RMON 管理程序包执行阈值配置。商业 RMON NMS 数据包（如 Cisco Traffic Director，它是 Cisco Works 2000 的一部分）纳入了 GUI，从而使 RMON 阈值的设置变得更简单。

为了确立基准，etherStats 组提供了一系列有用的 L2 数据流统计数据。此表中的对象可用于获得有关单播、多播和广播数据流以及各种 L2 错误的统计数据。也可以将交换机上的 RMON 代理配置为将这些示例值存储在历史记录组中。这一机制能够让轮询数量减少，而不降低采样率。RMON 历史记录能产生准确基准，且不会产生大量的轮询开销。然而，收集的历史记录越多，使用的交换机资源越多。

当交换机只提供 RMON-1 的四个基本组时，重要的是不要忘记其余的 RMON-1 和 RMON-2。所有组均是在 RFC 2021 中定义的，其中包括 UsrHistory（组 18）和 Probeconfig（组 19）。可以使用 SPAN 端口或 VLAN ACL 重定向功能从交换机中检索 L3 及更高版本的信息，利用这些重定向功能，可以将数据流复制到外部 RMON SwitchProbe 或内部网络分析模块 (NAM) 上。

NAM 支持所有 RMON 组，甚至可以检查**应用层数据**，包括当 MLS 处于启用状态时从 Catalyst 导出的 NetFlow 数据。运行 MLS 意味着路由器不会在流中交换所有数据包，因此，只有网络流数据导出（而不是接口计数器）才可以生成可靠的 VLAN 记账。

您可以使用 SPAN 端口和交换机探测器捕获特定端口、中继或 VLAN 的数据包流并上载要通过 RMON 管理程序包进行解码的数据包。SPAN 端口可通过 CISCO-STACK-MIB 中的 SPAN 组控制 SNMP，因此很容易使此过程实现自动化。Traffic Director 将这些功能与其流动代理程序功能配合

使用。

有一些警示跨整个 VLAN。即使您使用 1Gbps 探测器，来自一个 VLAN 或者来自一个甚至 1Gbps 的全双工端口的整个数据包流可能会超出 SPAN 端口的带宽。如果 SPAN 端口继续使用完整带宽运行，则数据可能会丢失。有关详细信息，请参阅配置 Catalyst 交换端口分析器 (SPAN) 功能。

建议

Cisco 建议部署 RMON 阈值和警报，以通过比单独使用 SNMP 轮询更智能的方式来帮助进行网络管理。这将减少网络管理数据流开销，并使网络可以在基准发生某种变化时发出智能警报。RMON 需要由一个外部代理（如 Traffic Director）驱动；不提供 CLI 支持。发出以下命令可以启用 RMON：

```
set snmp rmon enable
set snmp extendedrmon netflow enable mod
!--- For use with NAM module only.
```

重要的是记住交换机的主要功能是转发帧，而不是充当大型多端口 RMON 探测器。所以，当您在多个端口上针对多种状况设置历史记录和阈值时，请记住，这样会耗费资源。如果要扩展 RMON，请考虑使用 NAM 模块。并且请切记重要端口规则：在规划阶段，仅在被标识为“重要端口”的端口上轮询和设置阈值。

内存要求

RMON 内存使用率在所有交换机平台间是恒定的，与统计数据、历史记录、警报和事件有关。RMON 在 RMON 代理（在此示例中为交换机）上使用桶来存储历史记录和统计数据。在 RMON 探测器 (Switch Probe) 或 RMON 应用程序 (Traffic Director) 上定义桶大小，然后将其发送到将要设置的交换机。通常情况下，只有在 DRAM 少于 32MB 的旧 Supervisor 引擎上才考虑内存限制。请参阅下列准则：

- 为了支持微型 RMON，需在 NMP 映像中添加大约 450K 的代码空间，微型 RMON 包括以下四组 RMON：统计数据、历史记录、警报和事件。因为 RMON 依赖于运行时配置，所以它的动态内存需求是可变的。下面提供了每个微型 RMON 组的运行时 RMON 内存使用率信息：以太网统计数据组 - 获取每个交换式以太网/FE 接口的 800 个字节。历史记录组--对于以太网接口，每个已配置的带50个桶的历史记录控制项使用近3.6KB的内存空间，其他额外的每个桶将占用56个字节。警报和事件组 - 每个已配置的警报及其对应的事件项占用 2.6KB 的内存空间。
- 如果系统总 NVRAM 大小为 256K 或更多，保存与 RMON 相关的配置将占用约 20K 的 NVRAM 空间，如果总 NVRAM 大小为 128K，则保存与 RMON 相关的配置将占用 10K 的 NVRAM 空间。

网络时间协议 (NTP)

NTP (RFC 1305) 同步一组分布式时间服务器和客户端间的计时，并且允许在创建系统日志或发生其他特定于时间的事件时将事件关联起来。

NTP 可以确保客户端时间的准确性，在局域网上通常可以精确到毫秒，在广域网上最多可以精确到几十毫秒，这与同步到协调世界时 (UTC) 的主服务器相关。典型的 NTP 配置使用多台冗余服务器和不同的网络路径来实现高准确性和可靠性。一些配置包括加密身份验证，目的是防止偶然或恶意协议攻击。

操作概述

NTP 最早是在 RFC 958 中记录的，随后通过 RFC 1119（NTP 版本 2）演进，目前已经是第三个版本（如 RFC 1305 中所定义）。 它在 UDP 端口 123 上运行。所有 NTP 通信都使用 UTC，该时间与格林尼治标准时间相同。

访问公共时间服务器

NTP 子网目前包括 50 多个公共主服务器，这些服务器通过无线电、卫星或者调制解调器直接与 UTC 同步。通常，客户端数量相对较少的客户端工作站和服务器无法与主服务器同步。大约有 100 个公共辅助服务器与主服务器保持同步，主服务器可为 Internet 上的 100,000 多个客户端和服务器提供同步功能。当前列表在“公共 NTP 服务器列表”页（该页定期更新）上进行维护。此外，有许多专用的主服务器和辅助服务器通常不可公共使用。有关公共 NTP 服务器列表及公共 NTP 服务器使用方法的信息，请访问特拉华大学（University of Delaware）的时间同步服务器网站。

因为既不能保证这些公共 Internet NTP 服务器的可用性，也不能保证它们将生成正确的时间，因此强烈建议考虑其他方法。这可能包括使用直接连接到若干路由器的各种独立全球定位服务 (GPS) 设备。

另一个可能的方法是使用配置为第一层主设备的各种路由器，但我们不建议使用这种方法。

层

每个 NTP 服务器均采用一个层，指示服务器距离外部时间源多远。第一层服务器有权访问某种类型的外部时间源，例如无线电时钟。第二层服务器从指定的一组第一层服务器获取时间详细信息，而第三层服务器从第二层服务器获取时间详细信息，依此类推。

服务器对等关系

- 服务器会响应客户端请求，但不会设法合并从客户端时间源获得的任何日期信息。
- 对等体会响应客户端请求，但会设法将客户端请求用作更好时间源的潜在候选者，并协助保持时钟频率稳定性。
- 为了成为真正的对等体，连接的两端必须形成对等体关系，而不是一端用户为对等体，而另一端用户为服务器。还建议对等体交换密钥，以便只有受信任的主机才能作为对等体彼此通话。
- 在客户端对服务器的请求中，服务器会应答客户端，并会忘记客户端曾经询问过问题；在客户端对对等体的请求中，服务器会应答客户端，并保留有关客户端的状态信息，以跟踪客户端在走时方面的情况及它在哪一层服务器上运行。**注意：** CatOS 只能充当 NTP 客户端。

NTP服务器处理几千个客户端没有任何问题。然而，处理数百个对等体即会对内存产生影响，且状态维护将消耗机箱上更多的 CPU 资源及带宽。

轮询

NTP 协议允许客户端根据需要随时查询服务器。实际上，在 Cisco 设备上首次配置 NTP 时，它会快速连续发出八次查询，间隔为 NTP_MINPOLL（24 = 16 秒）。NTP_MAXPOLL 为 214 秒（即 16,384 秒，或 4 小时 33 分钟 4 秒），这是 NTP 再次为回应轮询之前所经历的最长时间。目前，Cisco 还无法手动强制由用户设置轮询时间。

NTP轮询计数开始在2^6 (64)秒钟和由电源两增加(当两个服务器彼此同步)，到2^{10}。即您能盼望同步消息传送在间隔每配置的服务器或对等体64， 128， 256， 512或者1024秒。根据发送和接收数据包的锁相环路的不同，时间在 64 秒和 1024 秒（均为 2 的幂次方）之间变化。如果在相应时间内有很多抖动，则会更频繁地轮询。如果参考时钟精确，并且网络连接保持一致，您会在轮询之间的 1024 秒上看到轮询时间收敛。

实际上，这意味着 NTP 轮询间隔会随着客户端与服务器之间连接的变化而变化。连接越稳定，轮询间隔越长，这意味着 NTP 客户端已经收到了针对其前八个请求的八个响应（轮询间隔随后会加倍）。错过一个响应可导致轮询间隔减半。轮询间隔从 64 秒开始，最大可达到 1024 秒。在最好的情况下，轮询间隔从 64 秒变为 1024 秒将需要两个多小时的时间。

广播

决不转发 NTP 广播。**ntp broadcast** 命令可导致路由器在配置它的接口上生成 NTP 广播。broadcastclient命令的ntp导致路由器或听的交换机NTP在配置的接口广播。

NTP 流量级别

NTP 使用的带宽是最小的，因为在对端之间交换的轮询消息的间隔通常缩短到每 17 分钟（1024 秒）不超过一条消息。通过仔细规划，这可以在广域网链路的路由器网络之内得以维护。NTP 客户端必须与本地 NTP 服务器对等，而不是跨越广域网一直到将成为第二层服务器的中心站点核心路由器。

收敛的 NTP 客户端针对每台服务器每秒钟约使用 0.6 位。

建议

现在，很多客户在其 CatOS 平台上以客户端模式配置了 NTP，以便从 Internet 或无线电时钟的多个可靠的源实现同步。然而，在运行大量交换机的情况下，一种更加简单的服务器替代模式是：在广播客户端模式下，在交换域的管理 VLAN 上启用 NTP。此机制允许整个 Catalyst 域从单个广播消息中接收时钟。然而，由于信息流是单向的，因此计时精度稍有下降。

使用环回地址作为更新的来源，也有助于提高一致性。可以通过以下两种方式解决安全问题：

- 过滤服务器更新
- 验证

事件的时间相关性在以下两种情况下极其重要：故障排除和安全审核。应注意保护时间源和数据，并推荐使用加密，以防止有意或无意地清除关键事件。

Cisco 推荐进行以下配置：

Catalyst 配置

```
set ntp broadcastclient enable
set ntp authentication enable
set ntp key key
!--- This is a Message Digest 5 (MD5) hash. set ntp
timezone <zone name>
set ntp summertime <date change details>
```

备用 Catalyst 配置

```
!--- This more traditional configuration creates !---
more configuration work and NTP peerings. set ntp client
enable
set ntp server IP address of time server set timezone
zone name set summertime date change details
```

路由器配置

```
!--- This is a sample router configuration to distribute
!--- NTP broadcast information to the Catalyst broadcast
clients. ntp source loopback0
ntp server IP address of time server ntp update-calendar
clock timezone zone name clock summer-time date change
details ntp authentication key key ntp access-group
access-list
!--- To filter updates to allow only trusted sources of
NTP information. Interface to campus/management VLAN
containing switch sc0 ntp broadcast
```

Cisco 发现协议

CDP 通过数据链路层在相邻设备之间交换信息，它在确定逻辑或 IP 层以外的网络拓扑和物理配置方面非常有用。支持的设备主要包括交换机、路由器和 IP 电话。此部分着重介绍了 CDP 版本 2 相对于版本 1 的一些增强功能。

操作概述

CDP 将 SNAP 封装与类型代码 2000 一起使用。在以太网、ATM 和 FDDI 上，使用目标多播地址 **01-00-0c-cc-cc-cc, HDLC protocol type 0x2000**。在令牌环上，使用功能地址 c000.0800.0000。默认情况下，每分钟会定期发送 CDP 帧。

CDP 消息包含一个或多个子消息，允许目标设备收集并存储有关每个相邻设备的信息。

CDP 版本 1 支持下列参数：

参数	类型	说明
1	设备 ID	ASCII 格式的设备主机名称或硬件序列号。
2	地址	已发送更新的接口的 L3 地址。
3	端口 ID	已在其上发送 CDP 更新的端口。
4	功能	介绍设备的功能：路由器：0x01 TB 网桥：0x02 SR 网桥：0x04 交换机：0x08（提供 L2 和 L3 交换）主机：0x10 IGMP 条件过滤：0x20 网桥或交换机不会在非路由器端口上转发 IGMP 报告数据包。中继器：0x40
5	version	包含软件版本的字符串（与 **show version** 中的字符串相同）。
6	平台	硬件平台，例如 WS-C5000、WS-C6009 或 Cisco RSP。

在 CDP 版本 2 中，已引入其他协议字段。CDP 版本 2 支持所有字段，但所列出的字段在交换环境中特别有用，并且可用于 CatOS。

注意： 当交换机运行 CDPv1 时，它会丢弃 v2 帧。运行 CDPv2 的交换机在接口上接收 CDPv1 帧

时，除了从该接口发出 CDPv2 帧以外，还将从中开始发送 CDPv1 帧。

参数	类型	说明
9	VTP 域	VTP 域（如果是在设备上配置的）。
10	本地 VLAN	在 dot1q 中，这是未标记的 VLAN。
11	全双工/半双工	此字段包含发送端口的双工设置。

建议

默认情况下会启用 CDP，且 CDP 对于发现相邻设备并进行故障排除至关重要。网络管理应用程序也使用它来构建 L2 拓扑图。发出下列命令以设置 CDP：

```
set cdp enable
!--- This is the default. set cdp version v2
!--- This is the default.
```

在需要达到很高安全水平的网络的某些区域（例如面向 Internet 的 DMZ）中，必须通过以下方法关闭 CDP：

```
set cdp disable port range
```

show cdp neighbors 命令会显示本地 CDP 表。标有星号 (*) 的条目指示 VLAN 不匹配；标有 # 的条目指示双工不匹配。这可能非常有助于故障排除工作。

```
>show cdp neighbors

* - indicates vlan mismatch.
# - indicates duplex mismatch.
Port  Device-ID          Port-ID Platform
----- ------------------ ------- ------------
 3/1  TBA04060103(swi-2) 3/1     WS-C6506
 3/8  TBA03300081(swi-3) 1/1     WS-C6506
15/1  rtr-1-msfc         VLAN 1  cisco    Cat6k-MSFC
16/1  MSFC1b             Vlan2   cisco    Cat6k-MSFC
```

其他选项

某些交换机（如 Catalyst 6500/6000）能够通过 IP 电话的 UTP 电缆供电。通过 CDP 获取的信息有助于交换机上的电源管理。

因为可以将 PC 连接到 IP 电话，且两个设备均连接到 Catalyst 上的同一端口，所以交换机能够在独立的辅助 VLAN 中部署 VoIP 电话。这使得交换机能够轻松针对 VoIP 流量应用不同的服务质量 (QoS)。

另外，如果修改辅助 VLAN（例如，为了强制电话使用特定的 VLAN 或特定标记方法），此信息将通过 CDP 发送到电话。

参数	类型	说明
14	设备 ID	允许 VoIP 流量通过单独的 VLAN ID（辅助 VLAN）与其他流量区别开来。
16	功耗	VoIP 电话消耗的电量（以毫瓦为单位）。

注意： Catalyst 2900 和 3500XL 交换机目前不支持 CDPv2。

安全配置

理想情况下，用户应当已建立安全策略，来帮助他们确定 Cisco 的哪些工具和技术是合格的。

注意： 有很多文档（如 Cisco ISP 基本要素）对 Cisco IOS 软件安全性（与 CatOS 相对）进行了论述。

基本安全功能

密码

配置一个用户级口令（登录名）。在 CatOS 5.x 及更高版本中，口令区分大小写，长度可以是 0 到 30 个字符（包括空格）。设置启用口令：

set password *password* **set enablepass** *password*

所有口令都必须满足最小长度标准（例如，口令至少包含六个字符，且由字母和数字、大小写字母混合组成）才能进行登录，并在使用时激活口令。这些口令均使用 MD5 散列算法进行了加密。

为了更灵活地管理口令安全性和设备访问，Cisco 建议使用 TACACS+ 服务器。有关详细信息，请参阅本文档的 TACACS+ 部分。

Secure Shell (SSH)

使用 SSH 加密，可确保 Telnet 会话和到交换机的其他远程连接的安全性。SSH 加密仅在远程登录到交换机时才受支持。您无法对从交换机启动的 Telnet 会话进行加密。CatOS 6.1 支持 SSH 版本 1，并且在 CatOS 8.3 中增加了对 SSH 版本 2 的支持。SSH 版本 1 支持数据加密标准 (DES) 和三重 DES (3-DES) 加密方法，SSH 版本 2 支持 3-DES 和高级加密标准 (AES) 加密方法。您可以将 SSH 加密用于 RADIUS 和 TACACS+ 身份验证。SSH (k9) 图像支持此功能。有关详细信息，请参阅如何在运行 CatOS 的 Catalyst 交换机上配置 SSH。

set crypto key rsa 1024

为了禁用版本 1 回退和接受版本 2 连接，请发出以下命令：

set ssh mode v2

IP 允许过滤器

这些过滤器可确保通过 Telnet 和其他协议访问管理 sc0 接口。如果用于管理的 VLAN 还包含用户，则这些过滤器特别重要。发出以下命令，以便启用 IP 地址和端口过滤：

set ip permit enable
set ip permit *IP address mask Telnet|ssh|snmp|all*

然而，如果使用此命令无法访问 Telnet，则只能通过一些受信任的终端站来实现对 CatOS 设备的访问。此设置对故障排除来说可能是一大妨碍。请记住，它有可能会欺骗 IP 地址和愚弄过滤访问

，因此这只是保护的第一层。

端口安全性

考虑使用端口安全功能，以仅允许一个或几个已知 MAC 地址在特定端口上传送数据（举例来说，这可防止静态终端站在没有变更控制的情况下被新站取代）。这对静态 MAC 地址是可行的。

```
set port security mod/port enable MAC address
```

通过动态了解受限制的 MAC 地址也可以实现这一目的。

```
set port security port range enable
```

可以配置以下选项：

- set port security mod/port age time value — 指定在获取新地址之前，端口上的地址受到保护的持续时间。有效时间（以分钟为单位）为 10-1440。默认值是无老化时间。
- set port securitymod/port maximum value — 用于指定端口上要保护的 MAC 地址的最大数目的关键字。有效值为 1（默认值）至 1025。
- set port security mod/port violation shutdown — 如果发生冲突，会关闭端口（默认），并发送 Syslog 消息（默认）及丢弃流量。
- set port security mod/port shutdown time value — 端口保持禁用状态的持续时间。有效值为 10 至 1440 分钟。默认为永久关闭

在 CatOS 6.x 及更高版本中，Cisco 引入了 802.1x 身份验证，该功能允许客户端在可针对数据启用端口之前，对中央服务器进行身份验证。此功能出现在诸如 Windows XP 这样的平台上的早期技术支持中，但可能会被很多企业视为一个战略方向。有关如何在运行 Cisco IOS 软件的交换机上配置端口安全性的信息，请参阅配置端口安全。

登录 标识

创建适当的设备标语，以专门声明针对未经授权的访问所要执行的操作。请不要通告可能向未授权用户提供信息的站点名或网络数据。如果设备受损，并且攻击者被捉住，那么这些标语可提供追索权。

```
# set banner motd ^C
*** Unauthorized Access Prohibited ***
***  All transactions are logged   ***
------------- Notice Board -------------
----Contact Joe Cisco at 1 800 go cisco for access problems----
^C
```

物理安全

未经适当授权，不得接近设备，因此，必须将设备放置在一个受控制的（上锁的）场地。为了确保网络处于可操作状态，并且不会受到环境因素的不良影响，所有设备必须拥有适当的 UPS（尽可能提供足够的资源）并进行适当的温度控制（空调）。请记住，如果恶意攻击者突破了物理访问，则他们更有可能通过口令恢复或其他手段进行干扰。

终端访问控制器访问控制系统

默认情况下，无权和有权模式口令是全局性的，适用于访问交换机或路由器的每个用户，这些用户通过网络上的控制台端口或 Telnet 会话进行访问。其在网络设备上的实施过程很耗时，且非集中化

。使用易于引起配置错误的访问列表实施访问限制也很困难。

现已推出三种可帮助您控制和管理对网络设备的访问的安全系统。这些系统使用客户端/服务器体系结构在一个中央数据库中放置所有安全信息。这三个安全系统是：

- TACACS+
- RADIUS
- Kerberos

TACACS+ 是 Cisco 网络中的一种常见部署，同时也是本章论述的重点。它提供了下列功能：

- 身份验证 - 用户的识别和验证过程。认证用户的方法可能有几种，但最常见的方法包括同时使用用户名和密码。
- 授权 - 验证用户的身份后可授予其执行各种不同命令的权限。
- 记账 - 记录用户正在设备上执行的操作或已完成的操作。

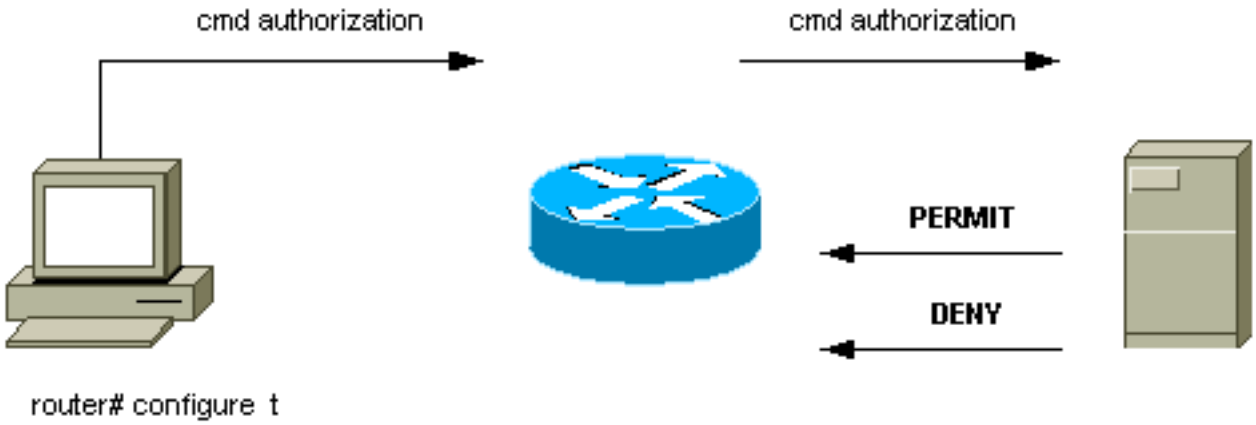
有关详细信息，请参阅在 Cisco Catalyst 交换机上配置 TACACS+、RADIUS 和 Kerberos。

操作概述

TACACS+ 协议向中央服务器转发用户名和口令，在此过程中会通过网络使用 **MD5** 单向哈希算法对用户名和口令加密 (RFC 1321)。 它使用 TCP 端口 49 作为传输协议；这与 UDP（由 RADIUS 使用）相比有下列优点：

- 面向连接的传输
- 已收到请求的单独确认 (TCP ACK) 与后台身份验证机制的当前加载方式无关
- 立即指示服务器崩溃情况（RST 数据包）

会话期间，如果需要进行额外的授权检查，那么交换机会与 TACACS+ 进行核对，确定是否授予了用户使用特定命令的权限。这样可以更好地控制从身份验证机制中分离时能够在交换机上执行的命令。使用命令记帐，可以审核连接到特定网络设备的特定用户发出的命令。



用户尝试简单的 ASCII 登录而使用 TACACS+ 进行身份验证以访问网络设备时，通常会发生以下过程：

- 建立连接后，交换机与 TACACS+ 后台程序通信以获得用户名提示，随后会向用户显示该提示。用户输入用户名，交换机便与 TACACS+ 后台程序通信以获得口令提示。交换机为用户显示密码提示，然后用户输入密码，并发送给TACACS+后台程序。
- 最后，网络设备从 TACACS+ 后台程序收到下列响应之一：ACCEPT - 用户已通过身份验证且可以开始服务。如果网络设备配置为需要授权，则此时会开始授权。REJECT - 用户未能通过身份验证。根据 TACACS+ 后台程序，系统可能会拒绝用户进行进一步访问或者会提示用户重

试登录序列。ERROR - 在身份验证期间的某一时间出错。这可能发生在后台程序中，也可能发生在后台程序和交换机之间的网络连接中。如果收到 ERROR 响应，网络设备通常会尝试使用另一种方法对用户进行身份验证。CONTINUE - 系统提示用户提供其他身份验证信息。
- 用户必须首先成功完成 TACACS+ 身份验证才能进行 TACACS+ 授权。
- 如果需要进行 TACACS+ 授权，那么交换机会再次与 TACACS+ 后台程序进行通信，并从其收到 ACCEPT 或 REJECT 授权响应。如果收到 ACCEPT 响应，则其中包含属性形式的数据，用于指导该用户的 EXEC 或 NETWORK 会话，并确定该用户可以访问的命令。

建议

Cisco 建议使用 TACACS+，因为可以使用 CiscoSecure ACS for NT、Unix 或其他第三方软件轻松实现 TACACS+。TACACS+ 功能包括详细的统计，可提供以下方面的记账数据：命令使用情况、系统使用情况、MD5 加密算法、身份验证和授权过程的管理控制。

在本示例中，登录并启用使用 TACACS+ 服务器进行身份验证的模式，这样，如果服务器不可用，便可以回退到本地身份验证。这是大部分网络中留下的一个重大后门。发出下列命令以设置 TACACS+：

```
set tacacs server server IP primary set tacacs server server IP
!--- Redundant servers are possible. set tacacs attempts 3
!--- This is the default. set tacacs key key
!--- MD5 encryption key. set tacacs timeout 15
!--- Longer server timeout (5 is default). set authentication login tacacs enable
set authentication enable tacacs enable
set authentication login local enable
set authentication enable local enable
!--- The last two commands are the default; they allow fallback !--- to local if no TACACS+
server available.
```

其他选项

可使用 TACACS+ 授权控制每个用户或用户组可以在交换机上执行的命令，但提供此方面的建议比较困难，因为所有客户对此都有各自的要求。有关详细信息，请参阅使用身份验证、授权和记账控制对交换机的访问。

最后，accounting 命令提供了有关每个用户所键入和所配置内容的审计线索。以下是在命令结束时接收审计信息所用的普遍做法的示例：

```
set accounting connect enable start-stop tacacs+
set accounting exec enable start-stop tacacs+
set accounting system enable start-stop tacacs+
set accounting commands enable all start-stop tacacs+
set accounting update periodic 1
```

此配置具有下列功能：

- 利用 **connect** 命令可以统计交换机上如 Telnet 等出站连接事件。
- 利用 **exec** 命令可以统计交换机上如操作人员等登录会话。
- 利用 **system** 命令可以统计交换机上如重新加载或重置等系统事件。
- **commands** 命令支持对输入到交换机中内容的记帐，适用于show和configuration命令。
- 每分钟定期更新服务器有助于记录用户是否仍然处于登录状态。

配置清单

本部分概述了所提供的建议配置，不包括安全详细信息。

这对标记所有端口十分有用。发出下面的命令可标记端口：

```
set port description descriptive name
```

以下各命令表中使用了下面的约定：

密钥：
粗体文本 - 建议的更改
常规文本 - 默认的建议设置

全局配置命令

命令	注释
set vtp domain name passwordx	避免受新交换机中未授权的 VTP 更新的危害。
set vtp mode transparent	选择本文档中提及的 VTP 模式。有关详细信息，请参阅本文档中的 VLAN 中继协议部分。
set spantree enable all	确保在所有 VLAN 上启用 STP。
set spantree root VLAN	建议根据 VLAN 安置根（和辅助根）网桥。
set spantree backbonefast enable	允许从间接故障中快速进行 STP 收敛（仅当域中的所有交换机都支持此功能时）。
set spantree uplinkfast enable	允许从直接故障中快速进行 STP 收敛（仅适用于接入层交换机）。
set spantree portfast bpdu-guard enable	允许端口在存在未经授权的生成树扩展的情况下自动关闭。
set udld enable	启用单向链路检测 (UDLD)（还需要进行端口级配置）。
set test diaglevel complete	允许在启动时进行全面诊断（Catalyst 4500/4000 上的默认设置）。
set test packetbuffer sun 3:30	启用端口缓冲区错误检查（仅适用于 Catalyst 5500/5000）。
set logging buffer 500	保持最大内部 Syslog 缓冲区。
set logging server IP 地址	为外部系统消息日志记录配置目标 Syslog 服务器。
set logging server enable	允许外部日志记录服务器。
set logging timestamp enable	在日志中启用消息的时间戳。
set logging level	增加默认的 STP Syslog 级别。

spantree 6 default	
set logging level sys 6 default	增加默认的系统 Syslog 级别。
set logging server severity 4	只允许导出更高严重级别的 Syslog。
set logging console disable	除非进行故障排除，否则禁用控制台。
set snmp community read-only 字符串	配置允许进行远程数据收集的口令。
set snmp community read-write 字符串	配置允许进行远程配置的口令。
set snmp community read-write-all 字符串	配置允许进行包括口令在内的远程配置的口令。
set snmp trap enable all	对 NMS 服务器启用故障和事件警报的 SNMP 陷阱。
set snmp trap 服务器地址字符串	配置 NMS 陷阱接收器的地址。
set snmp rmon enable	启用 RMON 以收集本地统计数据。有关详细信息，请参阅本文档中的远程监视部分。
set ntp broadcastclient enable	允许从上游路由器接收精确的系统时钟。
set ntp timezone 时区名称	设置设备的本地时区。
set ntp summertime 日期更改详细信息	如果适合时区，则配置夏令时。
set ntp authentication enable	出于安全考虑，配置加密的时间信息。
set ntp key 密钥	配置加密密钥。
set cdp enable	确保启用邻居发现（默认情况下在端口上处于启用状态）。
set tacacs server IP 地址 primary	配置 AAA 服务器的地址。
set tacacs server IP 地址	如果可以，则为冗余 AAA 服务器。
set tacacs attempts 3	允许对 AAA 用户帐户进行 3 次口令尝试。
set tacacs key 密钥	设置 AAA MD5 加密密钥。
set tacacs timeout 15	允许更长的服务器超时（默认值为 5 秒）。
set authentication login tacacs enable	使用 AAA 对登录进行身份验证。
set authentication enable tacacs enable	使用 AAA 对启用模式进行身份验证。
set authentication login local enable	默认;如果无可用 AAA 服务器，则允许回退到本地。
set authentication	默认;如果无可用 AAA 服务器，则

enable local enable	允许回退到本地。

主机端口配置命令

命令	注释
set port host 端口范围	删除不必要的端口处理。此宏将 spantree PortFast 设置为 enable、channel off、trunk off。
set udld disable *端口范围*	删除不必要的端口处理（默认情况下在铜线端口上处于禁用状态）。
set port speed *端口范围* auto	对最新主机 NIC 驱动程序使用自动协商。
set port trap *端口范围* disable	对一般用户来说无需 SNMP 陷阱；仅跟踪关键端口。

服务器配置命令

命令	注释
set port host 端口范围	删除不必要的端口处理。此宏将 spantree PortFast 设置为 enable、channel off、trunk off。
set udld disable *端口范围*	删除不必要的端口处理（默认情况下在铜线端口上处于禁用状态）。
set port speed 端口范围 10/*100*	通常配置静态/服务器端口；否则使用自动协商。
set port duplex 端口范围 full/*半*	通常配置静态/服务器端口；否则使用自动协商。
set port trap 端口范围 enable	关键服务端口必须向 NMS 发送陷阱。

未使用的端口配置命令

命令	注释
set spantree portfast *端口范围* disable	为 STP 启用必要的端口处理和保护。
set port disable 端口范围	禁用未使用的端口。
set vlan 未使用的虚拟 VLAN 端口范围	如果已启用端口，则将未经授权的流量定向到未使用的 VLAN。
set trunk 端口范围 off	禁用端口中继，直到管理为止。
set port channel 端口范围 mode off	禁用端口信道，直到管理为止。

基础架构端口（交换机-交换机，交换机-路由器）

命令	注释

set udld enable *端口范围*	启用单向链路检测 (UDLD)（并非铜线端口上的默认值）。
set udld aggressive-mode enable 端口范围	启用主动模式（对于支持该模式的设备）。
set port negotiation *端口范围* enable	允许链路参数的默认 GE 自动协商。
set port trap 端口范围 enable	允许这些关键端口的 SNMP 陷阱。
set trunk 端口范围 off	如果不使用中继，则禁用功能。
set trunk mod/port desirable ISL/*dot1q*/*协商*	如果使用中继，则首选 dot1q。
clear trunk mod/port vlan range	通过对不需要 VLAN 的中继中的 VLAN 进行修剪来限制 STP 的直径。
set port channel 端口范围 mode off	如果不使用信道，则禁用功能。
set port channel 端口范围 mode desirable	如果使用信道，则启用 PAgP。
set port channel all distribution ip both	如果使用信道，则允许 L3 源/目标负载均衡（Catalyst 6500/6000 上的默认值）。
set trunk mod/port nonegotiate ISL/*dot1q*	如果中继至路由器、Catalyst 2900XL、3500 或其他供应商，则禁用 DTP。
set port negotiation mod/port disable	协商可能会与某些旧的 GE 设备不兼容。

相关信息

- Catalyst 4500/4000 系列交换机上常见的 CatOS 错误消息
- Catalyst 5000/5500 系列交换机上常见的 CatOS 错误消息
- Catalyst 6500/6000 系列交换机上常见的 CatOS 错误消息
- 交换机产品支持
- LAN 交换技术支持
- 技术支持和文档 - Cisco Systems