

思科 StackWise 技术

本白皮书概括地介绍了思科 StackWise 技术及其通过连接多个固定配置交换机，创建一个统一的逻辑堆叠交换架构的特殊机制。本文着重介绍思科 StackWise 技术的下列关键部分：堆叠互联行为；堆叠的创建和改动；第二层和第三层转发；服务质量（QoS）机制。本文的目的是帮助读者了解思科 StackWise 技术怎样为语音、视频和千兆位以太网应用提供出色的性能。

技术概述

思科 StackWise 技术可以为统一地利用一组交换机的功能提供一种创新的方法。单个交换机可以智能化地结合到一起，创建一个单一的交换单元，并具有 32Gbps 的交换堆叠互联。同一堆叠中的所有交换机共享配置和路由信息，从而创建一个单一的交换单元。用户可以在不影响性能的情况下，从一个正在工作的堆叠中添加或者移除交换机。

交换机可以通过特殊的堆叠互联电缆，组织成一个单一的逻辑单元，从而创建一条双向的封闭环路。这条双向环路可以充当它所连接的所有交换机的交换矩阵。

图 1 利用 StackWise 技术建立的 Cisco Catalyst 3750 系列交换机堆叠



网络拓扑和路由信息通过堆叠互联不断更新。堆叠的所有成员都可以充分地使用堆叠互联带宽。堆叠作为一个单一的单元，由从堆叠成员交换机中选出一台主交换机负责管理。

堆叠中的每台交换机都能够充当管理层次中的主交换机或者从交换机。从堆叠成员中选出的主交换机将充当堆叠的控制中心。从交换机充当转发处理器。每台交换机都将被指定一个编号。一个堆叠中最多可以连接九台交换机。用户可以在不影响堆叠性能的情况下，在堆叠中添加或者移除交换机。

每个由 Cisco Catalyst 3750 系列交换机构成的堆叠都拥有一个单一的 IP 地址，并且作为一个统一的对象进行管理。单一 IP 管理适用于包括故障检测、虚拟 LAN（VLAN）创建和更改、安全和 QoS 控制等多种活动。每个堆叠只有一个配置文件，它将被分发给堆叠中的各个成员。这让堆叠中的每台交换机都可以共用相同的网络拓扑、介质访问控

思科系统公司

所有内容都归思科系统公司 1992 - 2003 年版权所有。重要声明和保密声明。

第 1 页，共 1 页

制 (MAC) 地址和路由信息。

堆叠互联功能

利用特殊的堆叠互联电缆和堆叠软件,思科 StackWise 技术最多可以将 9 台单独的 Cisco Catalyst 3750 交换机连接到一个统一的逻辑单元中。堆叠相当于一个单一的交换单元,由一个从成员交换机中选出的主交换机管理。主交换机可以自动地创建和升级所有的交换信息和可选的路由表。一个工作中的堆叠可以在不中断服务的情况下,添加新的成员或者移除旧的成员。

双向流

为了有效地平衡流量负载,分组被分配到两条逻辑计数循环路径上。每个计数循环路径都支持 16Gbps 的流量,总共支持 32Gbps。输出队列会计算路径的使用率,以确保流量负载的平衡分配。

无论数据帧在何时可以发送到路径上进行传输,系统都会计算路径的使用率,判断哪条路径拥有最多的带宽。确定之后,整个数据帧就会被复制到这条路径上。系统为流量提供服务的方式取决于它的服务等级 (CoS) 或者差分服务代码点 (DSCP) 值。低延时的流量会获得较高的优先级。

如果检测到电缆中存在断点,流量将会立即转到另外一条 16Gbps 路径上继续转发。

在线添加和移除堆叠中的交换机

用户可以在不影响堆叠性能的情况下,在一个工作中的堆叠中添加或者移除交换机。在添加了一个新的交换机以后,主交换机会自动地利用目前正在使用的 Cisco IOS 软件镜像和堆叠配置对该单元进行设置。堆叠将会搜集包括交换表在内的各种信息,并在获得新地址之后更新 MAC 地址。网络管理员不需要对交换机进行任何配置,就可以直接使用该交换机。同样,用户也可以在不对剩余交换机产生影响的情况下,从一个工作中的堆叠中移除交换机。当堆叠发现一系列端口不再可用时,它将在不影响转发或者路由的情况下更新相关信息。

物理菊花链连接

如图 2 所示,交换机在物理上通过菊花链连接在一起。任何一条电缆发生中断,都将导致堆叠的带宽降低到它的总容量的一半。次秒级定时机制可以检测流量故障,及时地进行故障切换。这种机制可以在定时机制检测到电缆上的互动时重新恢复双路径传输。

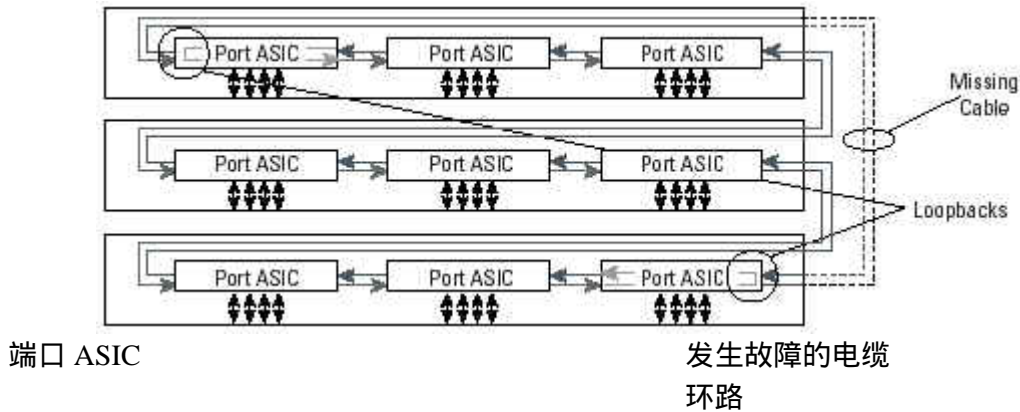
图 2 思科 StackWise 技术富有弹性的布线方式



次秒级故障切换

在路径的某个部分发生中断后的几个微秒内，所有数据都将被转移到双向路径的另外一条可以使用的路径上。（如图 3 所示）

图 3 在电缆发生中断后建立环路



交换机可以连续监控堆叠端口的活动，纠正数据传输中的错误。如果错误条件超过某个特定的阈值，或者电缆和它的端口之间缺乏足够的电磁接触，交换机将会检测到这种故障，并会发送一个消息给沿中断电缆方向上距离最近的交换机。随后这两台交换机将把所有流量都转移到另外一条没有故障的路径上。

单一管理 IP 地址

堆叠在进行初始设置时将获得一个 IP 地址。在创建了堆叠 IP 地址之后，连接到堆叠的物理交换机都将成为主交换机组的一部分。在连接到一个组时，每台交换机都将使用堆叠 IP 地址。在选出了一个新的主交换机之后，它将使用这个 IP 地址继续与网络进行交互。

堆叠的创建和改动

堆叠由各个交换机通过堆叠电缆相连而成。当堆叠端口检测到机电活动时，每个端口都开始传输关于它的交换机的信息。在获悉所有交换机的情况之后，堆叠将会选出一个成员作为主交换机，它负责维护和升级配置文件、路由信息和其他堆叠信息。整个堆叠将拥有一个单一的 IP 地址，该地址将供所有交换机使用。

思科系统公司

所有内容都归思科系统公司 1992 - 2003 年版权所有。重要声明和保密声明。

第 3 页，共 3 页

1:N 的主冗余

1:N 的主冗余让每个堆叠成员都可以充当主交换机，从而为数据的转发提供了最高的可靠性。堆叠中的每台交换机都可以充当一个主交换机，从而为网络控制创建了一种 1:N 的可用性机制。在某个单元发生故障时（尽管发生这种情况的可能性很小），所有其他单元都可以继续转发流量和保持正常运行。

主交换机的选择

堆叠相当于一个单一的交换单元，由一个从成员交换机中选出的主交换机管理。主交换机可以自动地创建和升级所有的交换信息和可选的路由表。堆叠中的任何成员都可以成为主交换机。在安装或者重启整个堆叠之后，堆叠中的交换机之间将进行一次选举。选举标准的层次结构如下：

1. **用户优先级**——网络管理员可以选择一个交换机作为主交换机。
2. **硬件和软件优先级**——这种优先级将自动分配给具有最全面的功能集的设备。具有增强多层软件镜像（EMI）版本的 Catalyst 3750 交换机具有比标准多层软件镜像（SMI）版本高的优先级。
3. **缺省配置**——如果某个交换机拥有预设的配置信息，它将拥有比尚未设置的交换机高的优先级。
4. **正常运行时间**——运行时间最长的交换机将被选中。
5. **MAC 地址**——每个交换机将向它的所有相邻交换机报告它的 MAC 地址，以便比较。具有最低的 MAC 地址的交换机将被选中。

主交换机的活动

主交换机将充当 IP 功能（例如 Telnet 进程、ping、命令行界面（CLI）和路由信息交换）的主要联络点。主交换机负责将转发表下载到各个从交换机。组播和单播路由任务都由主交换机完成。QoS 和 ACL 配置信息也由主交换机分发到各个从交换机。当某个新的从交换机加入堆叠后，或者某个现有的交换机从堆叠中移除，主交换机将会针对这个事件发布一个通知，所有从交换机将按照该通知更新它们的路由表。

共享的网络拓扑信息

主交换机负责搜集和维护正确的路由和配置信息。它通过定期地向堆叠中的所有从交换机发送复本或者升级，保持这些信息的时效性。当某个新的主交换被选出来时，它将会重新使用以前的主交换机所保持的配置信息，以确保用户和网络的连续性。

从交换机的活动

每个交换机都拥有包含它自己的本地 MAC 地址的路由表和包含堆叠中其他 MAC 地址的路由表。主交换机将保存包含所有向堆叠报告的 MAC 地址的路由表。主交换机还会创建一个包含整个堆叠中所有 MAC 地址的地图，并将其发给所有从交换机。这样每个交换机都可以检测到堆叠的每个端口。这消除了重复的学习流程，为系统创建了一个更快、更有效的交换基础设施。

从交换机为它们所支持的每个 VLAN 保存它们自己的生成树。主交换机负责为堆叠中的每个 VLAN 保存所有生成树表的复本。当增加或者移除一个新的 VLAN 时，所有现

有的交换机都将获得该事件的通知，并将根据该通知升级它们的路由表。

从交换机需要等待从主交换机接收正在使用的配置的复本，并在收到最新的信息之后开始发送数据。这确保了所有交换机都将只使用最新的信息，而且在制定转发决策时只使用一个网络拓扑。

多种用于确保高可用性的机制

思科 StackWise 技术可以支持多种用于在堆叠中实现很高弹性的机制：

- **CrossStack EtherChannel 技术**——一个堆叠中的多台交换机可以创建一条 EtherChannel 连接。失去一台交换机将不会影响其他交换机的连接。
- **等成本路由**——交换机可以为不同的路由器提供双归属功能，以提高冗余性。
- **1:N 的主冗余**——堆叠中的每台交换机都可以充当主交换机。如果现有的主交换机发生故障，堆叠将会选出另外一个主交换机。
- **堆叠电缆弹性**——当双向环路中的某处发生中断时，交换机会自动地开始通过仍然可以使用的一半环路发送信息。如果整个 32Gbps 带宽都被占用，QoS 机制将会控制流量，以便首先传输对抖动和延时较为敏感的流量，将优先级较低的流量推后传输。
- **在线插拔**——用户可以在不影响堆叠性能的情况下，添加或者移除交换机。
- **分布式的第二层转发**——在主交换机发生故障时，各个交换机还将继续根据它们之前从主交换机处获得的路由表转发信息。
- **用于第三层弹性的 RPR +**——通过初始化，每台交换机都具有路由功能，可以在目前的主交换机发生故障时被选为主交换机。从交换机不需要进行重启，因此第二层转发可以继续。

第二层和第三层转发

思科 StackWise 技术可以为第二层和第三层转发的管理提供一种创新的方法。第二层转发是通过一种分散的方法实现的。第二层转发则通过一种集中方法实现。这为堆叠中的路由和交换活动提供了最大限度的弹性和效率。

切换主交换机时的转发弹性

在某个主交换机停止工作之后，堆叠将会选出一个新的主交换机，在此期间，堆叠将保持正常工作。第二层连接将不会受到任何影响。新的主交换机会利用它的热备用单播电缆，继续处理单播流量。组播电缆和路由电缆将会重新载入，以避免出现环路。

利用 RPR + 建立面向路由弹性的高可用性架构

用于在切换主交换机时确保路由的高可用性的机制被称为路由处理器冗余 + (RPR +)。目前 Cisco 12000 和 7500 系列路由器和 Catalyst 6500 系列交换机产品利用它来实现很高的可用性。每个具有路由功能的从交换机都经过了初始化，可以在主交换机发生故障时接替路由功能。每个从交换机都经过了全面的初始化，并与主交换机建有连接。从交换机拥有相同的接口地址、封装类型、接口协议和服务。从交换机将不断地接收和集成由目前的主交换机发出的同步配置信息，并通过不断地执行自检，监控它们的工作准备情况。由于不需要初始化路由接口，所以采用这种技术的设备重建路由和连接的速度要比

思科系统公司

所有内容都归思科系统公司 1992 - 2003 年版权所有。重要声明和保密声明。

第 5 页，共 5 页

普通的第三层设备快得多。

添加新的成员

当交换堆叠确定一个主交换机之后，此后添加的任何新交换机都会自动成为一个从交换机。所有现有的路由和地址信息都将被下载到从交换机，以便让它可以立即开始传输流量。它的端口的 IP 地址变得与主交换机相同。包括 QoS 设置在内的全局信息将会被下载到新的从交换机中。

Cisco IOS 镜像必须相同

思科 StackWise 技术需要堆叠中的所有单元都使用同一版本的 Cisco IOS 软件。但是当堆叠首次建立时，各个单元可以拥有不同版本的 SMI 和 EMI。这将自动导致采用 EMI 版本的单元成为主交换机，继而让堆叠可以执行路由功能。但是，在您首次升级 Cisco IOS 软件的版本时，所有单元都必须运行与主交换机相同的 SMI 或者 EMI 版本。

通过主交换机进行自动的 Cisco IOS 软件升级/降级

当一个新的交换机加入一个现有的堆叠时，主交换机将会与该交换机交换信息，以判断它的 Cisco IOS 镜像是否与堆叠的镜像相同。如果相同，主交换机将会把堆叠配置发给设备，端口也可以立即上线。如果不同，将会发生下面三种情况中的一种：

1. 如果堆叠运行的 Cisco IOS 镜像可以支持新交换机的硬件，主交换机将在缺省情况下将主交换机闪存中的 Cisco IOS 镜像下载到新交换机中，发送堆叠配置，并将该交换机上线。
2. 如果堆叠运行的 Cisco IOS 镜像不能支持新交换机的硬件，而且用户为 Cisco IOS 镜像的下载设置了一个普通文件传输协议（TFTP）服务器，主交换将会自动地将 Cisco IOS 镜像通过 TFTP 服务器下载到新的交换机中，对它进行设置，继而将其上线。
3. 如果堆叠运行的 Cisco IOS 镜像不能支持新交换机的硬件，主交换机将行交换机设于某种待命状态，就版本的不兼容性通知用户，等待用户将主交换机的 Cisco IOS 镜像升级到某种可以同时支持这两种硬件的版本。主交换机随后将把堆叠的其他部分（包括新加入的交换机）升级到这个版本，并将堆叠上线。

对堆叠中的所有设备进行升级

因为交换机堆叠相当于一个统一的单元，因而可以一次性地对堆叠的所有成员进行升级。这意味着如果某个堆叠刚开始时在不同的交换机上包含 EMI 和 SMI 功能，那么在进行首次 Cisco IOS 软件升级时，堆叠中的所有单元都将具有所使用的镜像的特性。尽管这可以更加有效地向堆叠添加功能，但是确保在将单元从 SMI 升级到 EMI 功能之前，购齐所有相关的升级许可。否则，这些单元将触犯 Cisco IOS 软件策略。

智能单播和组播——一个分组，多个目的地

思科 StackWise 技术使用了一种非常有效的机制来传输单播和组播流量。每个数据分组都需在堆叠互联上发送一次，其中包括组播分组。每个数据分组都具有 24 个字节的报头和可以针对该分组和 QoS 指定者的活动列表。这个活动列表可以指定端口目的地或者目的地，以及需要对该分组采取的处理。在进行组播时，主交换机将会确定哪些端口应当收取分组的一个复本，并为这些端口添加一个目的地索引。分组的一个复本随后会

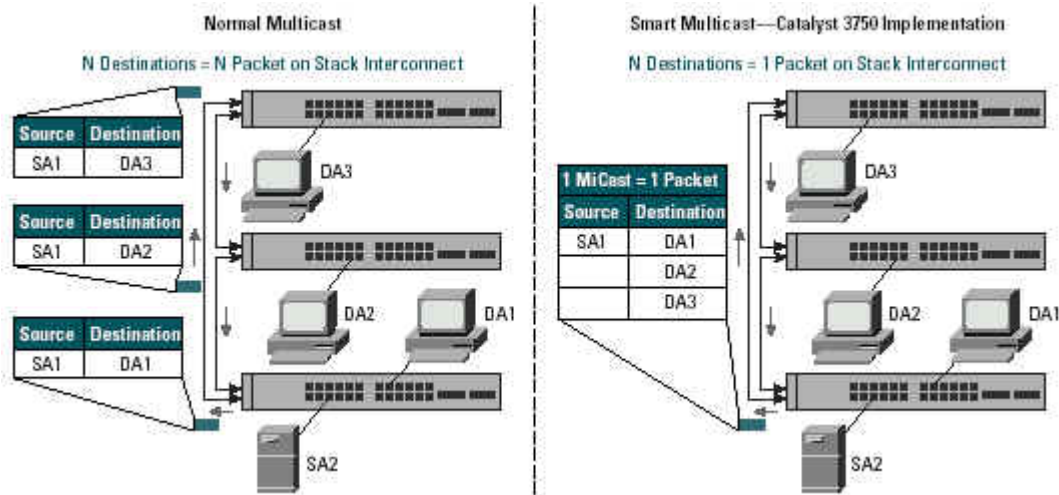
思科系统公司

所有内容都归思科系统公司 1992 - 2003 年版权所有。重要声明和保密声明。

第 6 页，共 6 页

被放在堆叠互联上。每个拥有一个目的地索引地址的交换机端口将会复制该分组。这为堆叠接收和管理组播信息提供了一种更加有效的机制（如图 4 所示）。

图 4 可堆叠交换机的普通组播和使用思科 StackWise 技术的 Cisco Catalyst 3750 系列交换机的智能组播之间的比较



普通组播	智能组播——Catalyst 3750 部署
N 个目的地 = 在堆叠互联上传输 N 个分组	N 个目的地 = 在堆叠互联上传输 1 个分组
源 目的地	1 个组播 = 1 个分组 源 目的地

QoS 机制

QoS 可以在用户接入网络的边缘提供精确的控制。这对于那些需要移植到融合应用的网络特别重要，这种网络需要对不同信息进行不同的处理。QoS 对于移植到千兆位以太网速度也非常重要，它可以有效地避免拥塞。

在网络边缘采用 QoS

思科 StackWise 技术支持四个整型的或者共享的循环输入队列。任何堆叠都可以整型或者共享。总输出也可以整型。网络管理员可以控制所有四个队列，设置各个队列的优先级，并对每个队列所能使用的带宽设置一个比例。缺省状态是共享所有队列。

当队列被设置为共享时，它们将被定义一个百分比。例如：队列一 50%，队列二 30%，队列三 20%，队列四 10%。

当队列设置为整型时，它们将被定义一定的带宽。例如：队列一 10Mb，队列二 40Mb，队列三 30Mb，队列四 20Mb。如果某个队列没有用完它所分配到的全部带宽，其他队列也可以使用该队列的资源——如果它们超出了限值的话。总带宽不能超过端口的最大线速。

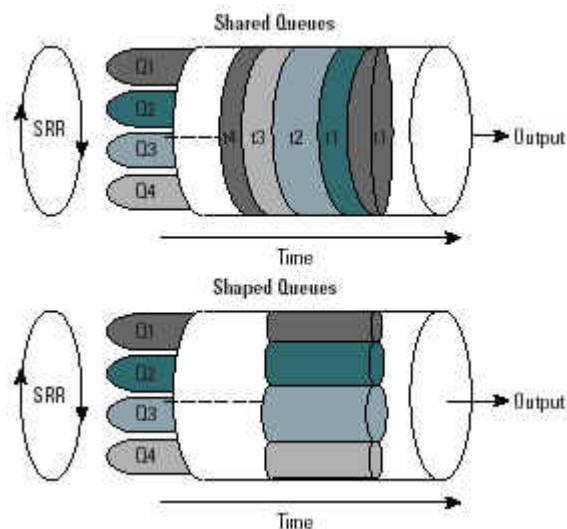
交换机的流量控制需要使用四个可用队列中的一个。网络管理员可以设置用哪个队列来

管理这些流量。

各个队列可以设置为支持帧或者字节数限制。

图 5 显示了整型的和共享的队列。

图 5 整型队列和共享队列的比较



共享队列
SRR 输出
 时间
整型队列
SRR 输出
 时间

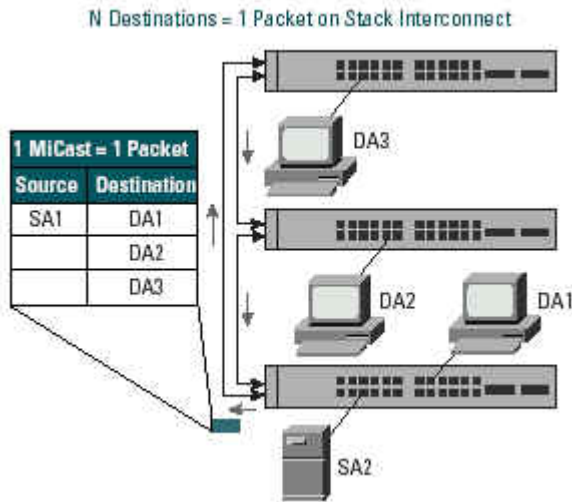
大型帧支持

思科 StackWise 技术可以在 10/100/1000 铜缆端口上支持最大 9KB 的大型帧。

智能 VLAN

VLAN 的工作方式与组播操作一样。如果主交换机检测到需要发往多个 VLAN 的信息，它就会创建一个具有很多目的地地址的分组副本。这将可以最大限度地利用堆叠互联（如图 6 所示）。

图 6 智能 VLAN 操作



N 个目的地 = 堆叠互联上的 1 个分组

1 组播 = 1 分组

源 目的地

覆盖整个堆叠的 EtherChannel 连接

因为堆叠上的所有端口都表现为一个统一的逻辑单元，所以 EtherChannel 技术可以覆盖堆叠中的多个物理设备。Cisco IOS 软件可以将堆叠中的任何交换机的最多八个单独的物理端口整合成一个逻辑通道上行连接。一个堆叠最多可以支持 12 个 EtherChannel 组。

管理

使用思科 StackWise 技术的产品可以通过 CLI 或者网络管理工具包进行管理。思科集群管理套件 (CMS) 软件专门针对思科可堆叠交换机的管理而开发。Cisco CMS 软件中针对堆叠单元的特殊向导让网络管理员可以对堆叠中的所有端口应用同一套配置。用于数据、语音、视频、组播、安全和 VLAN 间路由功能的预设向导让网络管理员可以一次性设置所有端口的配置。

思科 StackWise 技术也可以由 CiscoWorks 进行管理。

总结

思科 StackWise 技术让您可以提高您网络边缘的弹性和多样性，可以帮助您逐步采用高速、融合的应用。

北京

北京市东城区东长安街一号东方广场东一办公楼 19-21 层

邮政编码：100738

电话：(8610) 65267777

传真：(8610) 85181881

思科系统公司

所有内容都归思科系统公司 1992 - 2003 年版权所有。重要声明和保密声明。

第 9 页，共 9 页

广州

广州市天河北路 233 号中信广场 43 楼

邮政编码：510620

电话：(8620) 87007000

传真：(8620) 38770077

上海

上海市淮海中路 222 号力宝广场 32-33 层

邮政编码：200021

电话：(8621) 33104777

传真：(8621) 53966750

成都

成都市顺城大街 308 号冠城广场 23 层

邮政编码：610017

电话：(8628) 86758000

传真：(8628) 6528999

思科系统公司

所有内容都归思科系统公司 1992 - 2003 年版权所有。重要声明和保密声明。

第 10 页，共 10 页

