

# Cisco MDS 9000 系列交换机的高性能和可扩展性

## 简介

现在这个在强大的技术基础设施之上建设成功 IT 机构的时代，让企业可以提高他们的关键任务型应用的自动化水平和运行速度。企业对于技术基础设施的依赖性还将不断加强，从而使企业对于基础设施的性能和可扩展性的要求不断提高。这种要求的提高意味着供应商在设计应用环境（包括存储基础设施）时，需要越来越多地考虑到如何提高性能、可扩展性和效率。对于性能的更高要求促使供应商设计和部署速度更快、更加便于扩展的 IT 环境——从应用到存储子系统。

过去，一种公认的观点是，应用基础设施中速度最慢的环节是存储的访问。存储网络（SAN）在很大程度上消除了应用环境中的这个瓶颈。但是，人们目前对于存储网络的需求已经超过了今天的存储网络产品——即光纤通道交换机，所能达到的性能和可扩展性水平。可扩展性不仅指的是网络带宽的增长能力，还包括能否利用现有的光纤通道交换机产品的端口数量，有效地建设大规模的存储网络。随着存储网络需求的不断增长，用户迫切需要一种高性能的、能够通过扩展满足带宽和端口数量要求的、强大的存储网络平台。这样的平台必须能够提供投资保护，即它可以支持未来的需求、传输方式和端口密度，并能在短期内产生投资回报。

Cisco MDS 9000 系列多层光纤通道交换机平台是一种高度可扩展的模块化交换机，它采用了独特的设计，不仅可以超越目前的存储网络的性能和密度要求，还可以确保兼容性，能够支持未来的性能和可扩展性要求。Cisco MDS 9000 系列可以为未来提供出色的投资保护，并且可以满足用户目前的需要。

图 1 Cisco MDS 9500 系列控制器级交换机及 9216 光纤通道交换机



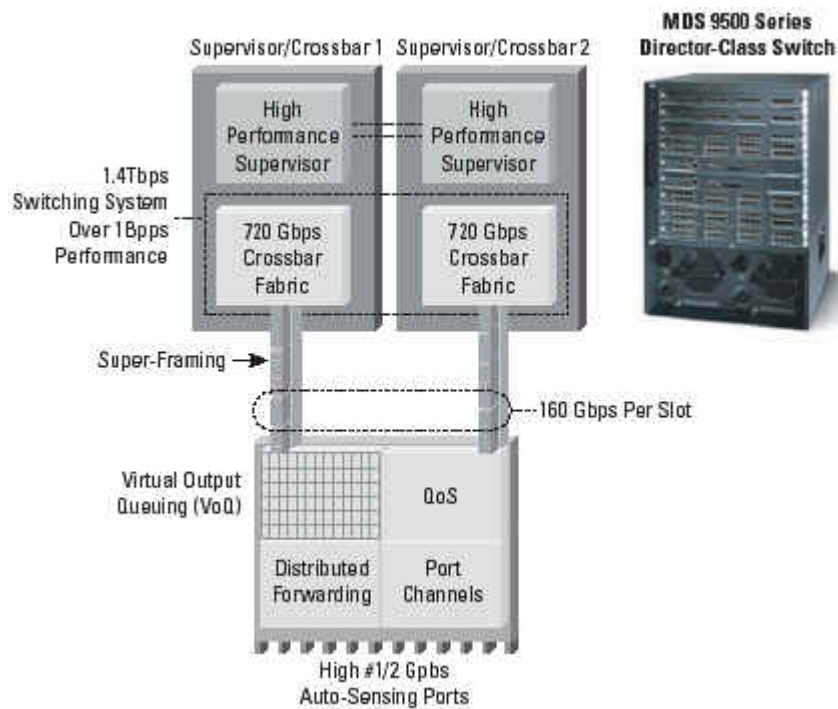
Cisco MDS 9000 系列光纤通道交换机是思科全新开发的一款产品，旨在提供出色的性能和较高的端口密度。但是，同样重要的是，它能够优化存储网络中的带宽资源的配置。Cisco MDS 9000 系列可以通过一个经过改进的架构和相关的功能集，为用户提供很多智能化的手段，帮助用户控制带宽的部署和使用方式。此外，该产品中的很多工具可以就存储网络中资源的使用情况提供精确的报告。以下将着重介绍 Cisco MDS 9000 系列交换机的性能和可扩展性，以及怎样利用这些功能建立高性能、高密度的存储网络和优化网络资源的配置。

## 有助于提升性能的架构

Cisco MDS 9500 系列控制器采用了一种高度可扩展的架构，并具有 Cisco IOS SAN 操作系统中的一组智能化功能，从而可以优化网络性能。核心硬件架构是这种高性能平台的基础。核

心硬件架构建立在一种业界独一无二、强大的高性能骨干网的基础之上。硬件本身包含的很多改进措施可以进一步提高平台的性能，满足和超越用户目前的需求。

图 2 Cisco MDS 9500 系列性能架构



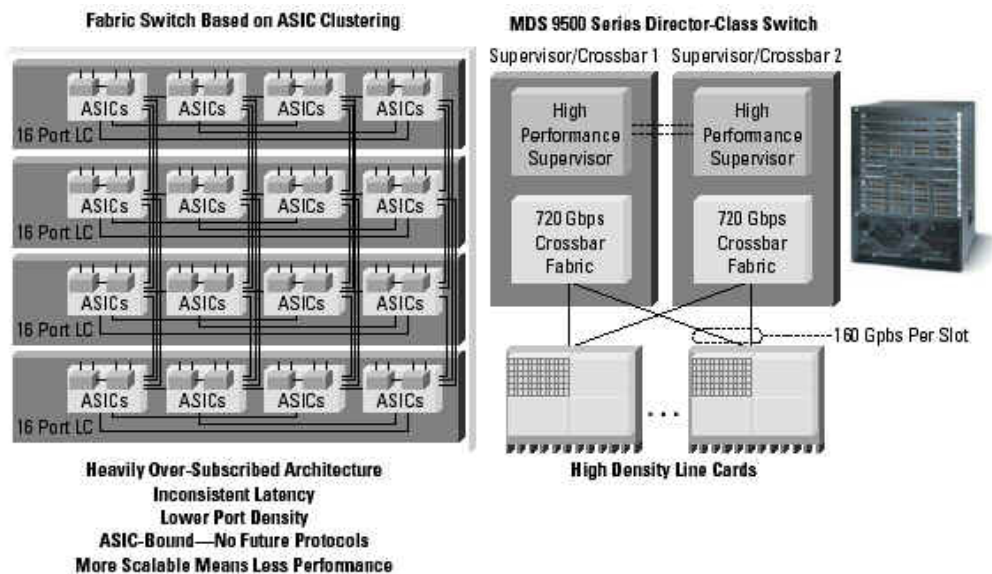
Cisco MDS 9216 光纤通道交换机由于只配置了单个处理器，因此不属于控制器级别，但在系统中的所有端口之间提供一个高达 80Gbps 带宽的交换路径阵列，所以系统能够以全速运行。

### 交叉交换矩阵(Crossbar)

用户对存储网络的性能和容量的需求一直在迅速增长。为了满足这种性能需求，存储网络交换机平台也必须不断地提升性能和容量。这种增长方式包括支持新的协议、新的线速或者新的传输方式。但是，存储网络平台的这种增长不应当以反复更换核心系统组件或者进行彻底升级为代价。目前部署在存储网络中的平台必须能够随着人们对于网络基础设施的需求的增长而增长。

目前有很多供应商都选择在较小的交换机特定用途集成电路（ASIC）的群集的基础上开发他们的交换机产品。尽管这种架构让很多供应商都能够以相对较快的速度，推出不同端口配置的低密度产品，但是人们对于更高性能和密度的要求很快就让这种架构毫无用武之地。当这种架构需要扩展规模时，它们内部的性能会变得无法预测，最终不可避免地会出现过载。它的另外一个缺点是不能迅速地支持新的协议、新的传输方式和新的线速，这是交换 ASIC 的固定特性造成的。

图 3 ASIC 群集交换机与基于交叉矩阵的控制器级交换机的比较



利用思科在开发高度可扩展的网络平台方面所积累的经验，Cisco MDS 9000 系列建立在一种高性能、强大的交叉交换矩阵架构的基础上。在像 Cisco MDS 9000 系列这样的系统中使用交叉交换矩阵，就可以通过一个灵活的机型提供很高的交换容量，而且这种机型可以根据未来的需要进行相应的改动。交叉矩阵的特性是无论所要转发的是什么协议和什么格式的数据帧，它都能够以很高的速率转发流量。因此，一个基于交叉矩阵的系统可以方便地满足光纤通道目前的高性能要求，并可以在将来方便地加以改进，从而可以在不需要更换任何设备的情况下支持新的协议、传输方式或者线速。

Cisco MDS 9500 系列具有 1.44Tbps 的原始交换性能。在一个九插槽的系统中，这相当于系统中的每个插槽具有 160Gbps 的纯带宽。要提供这样的性能，用户还必须考虑到高可用性。为了确保可用性不会因为性能的提高而降低，1.44Tbps 的交换容量实际上是通过两个 720Gbps 的交叉矩阵实现的。这两个冗余矩阵以一种活动—活动负载分担方式工作。Cisco MDS 9500 系列采用的架构是，如果某个交叉矩阵发生故障，另外一个交叉矩阵将会无缝地接管全部的负载，整个过程中系统的性能不会受到任何影响。因为刚开始时每个交换模块并不需要超过 80Gbps 的带宽，所以系统的冗余容量为 160Gbps。

### Supervisor 处理器模块

Supervisor 处理器模块是 Cisco MDS 9000 系列交换机的控制中心。它的控制引擎包含一个功能强大的处理器，它负责运行系统软件 and 所有相关功能。Supervisor 中的处理器还负责管理所有相关的交换阵列服务、状态表、数据库和系统中的很多其他功能。为了确保交换阵列的可扩展性，Cisco MDS 9000 系列交换机具有由一个高速的 Intel Pentium 3 处理器提供的、很高的 Supervisor 处理能力。这代表了目前市场中一个控制器级光纤通道交换机所能获得的最高处理能力。这样强大的处理能力使得 Cisco MDS 9000 系列能够支持更多的附加设备，具有端口密度更高、规模更大的网络，经过改进的智能化交换阵列服务，范围广泛的故障诊断和管理功能。

MDS 9216 交换机和 9500 控制器级使用的是同样的 Supervisor 处理器，只不过 9216 交换机的 Supervisor 处理器是固定在机箱中，而 9500 控制器的 Supervisor 处理器是以可热插拔的模块并且配备了双 Supervisor 处理器模块以支持在线软件升级功能。

**图 4 Cisco MDS 9500 系列 Supervisor 处理器模块**



为了确保可靠性，Cisco MDS 9500 系列控制器配备了两个全状态同步的 Supervisor 处理器模块，可以实现迅速、不间断的软件升级和故障恢复。

为了进一步提高 Cisco MDS 9500 平台的可扩展性，一个配置齐全的系统（包含两个交叉矩阵模块和两个 Supervisor 处理器模块）只需要占用两个插槽。这是通过使用一个能够安放一个 Supervisor 模块和一个集成化交叉矩阵的模块而实现的。通过将所占用的插槽总数限制到两个，可以为系统添加两个接口交换模块，从而可以增加最多 64 个 1/2Gbps 的光纤通道端口。

### **虚拟输出队列（VoQ）**

在光纤通道交换机中，通常每个端口都设有一个输入队列，用于存放从一个端口发往交换机中另外一个端口的数据帧。所有数据帧（无论要发往哪个目的地端口）都按照它们被接收的顺序排列。如果某个目的地端口发生阻塞，这种方式可能会导致末端阻塞。如果发生这种情况，输入队列中的所有其他帧都将被堵住，直到队列前端的某个帧被发送到发生阻塞的输出端口。在这种情况下，输入队列会被防止清空，而很多 I/O 交换都可能会出现不必要的延时。数据帧将会在队列中溢出，最终源设备将停止发送数据。

在图 5 所示的例子中，某个连接到交换机上“端口 1”的设备正在与另外三个分别连接到“端口 4、5 和 6”上的设备通信。如果某个输出端口发生阻塞，一个没有 VoQ 的交换机可能会发生末端阻塞，并且导致延时。尤其需要指出的是，如果不使用 VoQ，那么一旦“端口 4”上的设备因为无法处理流量负载而出现阻塞，那么所有发往端口 4 的数据帧都将被堵塞在“端口 1”队列中。而且，这种情况还会是使得从“端口 1”发出的所有其他帧都无法到达它们各自的目的地端口，例如“端口 5”和“端口 6”。因此，一个目的地端口上发生的阻塞可能会影响始发端口的所有其他流量。在存储网络等关键的环境中以这种低效率、不可预测的方式转发流量无疑是让人无法接受的。

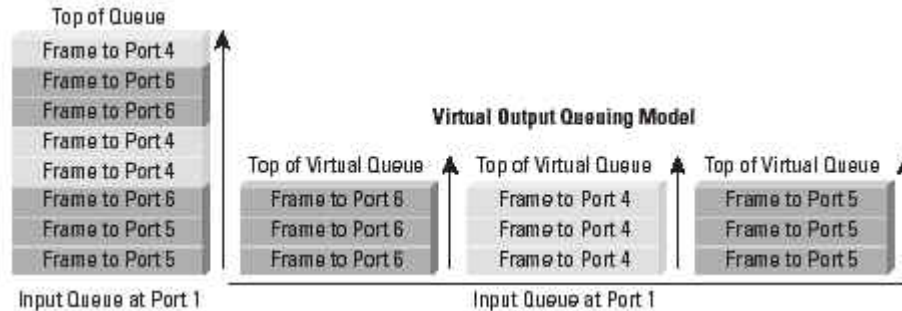
利用 VoQ 可以解决这种问题。如图 5 所示，交换机中的每个目的地端口都在输入端口拥有它自己单独的虚拟输出队列。因此，如果“端口 4”发生阻塞，它将不会对发往其他端口（例如“端口 5”和“端口 6”）的流量产生任何影响。这可以大大提高整个矩阵的效率。此外，VoQ 有助于在整个交换机中确保统一的性能，同时最大限度地降端口或者设备阻塞所产生的影响。

Cisco MDS 9000 系列交换机中的每个交换模块都采用了 VoQ。交换模块上的每个端口都为交换机的每个目的地端口设置了一个虚拟输出队列。此外，每个目的地虚拟输出队列实际上包含 4 个不同的队列，从而可以在系统的每个输出端口上，为各个等级的服务质量分别（QoS）



设置一个队列（每个端口 4 个等级）。这样做的结果是每个物理输入端口最多可以有 1024 个虚拟输出队列（256 个可能的目的地端口×4 个 QoS 等级）。VoQ 可以在网络发生拥塞时提供必要的保障，并可以提供一种有效的隔离手段，降低网络拥塞对于应用的整体影响。

图 5 虚拟输出队列模式可以解决线端阻塞问题



## 超级帧

Cisco MDS 9000 系列中的交叉矩阵可以在系统的各个交换模块之间转发数据帧。但是，为了获得对交叉矩阵的访问权限，需要转发数据帧的交换模块必须申请对交叉矩阵的访问权限。仲裁流程就是对访问交叉矩阵的访问权限进行集中的控制。因此，每个需要从系统的一个端口发送到另一端口的数据帧都必须申请对交叉矩阵的访问权限。申请方和交换模块之间使用一个基于信用点的系统，它负责授予通过交叉矩阵发送数据帧的权限。但是，在交换小型数据帧的情况下，这种方式的效率很低，因为在仲裁期间提供的每个信用点只能提供通过交叉矩阵发送少量字节的权限。为了优化交叉矩阵的资源利用率，思科在 Cisco MDS 9000 系列交换机中加入了一种名为“超级帧”的功能。超级帧技术是指将多个较小的数据帧整合成一个较大的超级帧，再在一个仲裁周期中通过交叉矩阵发送。超级帧可以提高交叉矩阵的工作效率，因为它不需要使用信用点，无需等待负责接收的交换模块认可并使用每个数据帧的信用点。

在创建超级帧时，需要遵循两个主要的原则：

- 数据帧必须来自于同一个源光纤通道 ID (FC\_ID)。
- 数据帧必须发往同一个目的地光纤通道 ID (FC\_ID)。

当某个数据帧离开虚拟输出队列 (VoQ) 时，系统将检查队列中是否有其他符合超级帧条件的数据帧。如果存在符合超级帧条件的数据帧，就会再创建一个超级帧，并在获得一个信用点之后立即发送。这样，只需要为整个超级帧申请一次仲裁，而不需要为每个数据帧申请仲裁。对于穿越交叉矩阵的小型数据帧，这可以最大限度地提高仲裁效率。利用超级帧技术，Cisco MDS 9000 系列的交换容量可以达到超过每秒 10 亿个光纤通道数据帧。

## 高密度系统

Cisco MDS 9000 系列的设计目的不仅是提供很高的性能，而且还是一种能够优化机架单元使用率的机型，提供很高的端口密度。这可以满足用户对于建设规模大于市场现有系统的大型存储系统的需求；Cisco MDS 9000 系列支持一系列交换模块，从而可以为用户提供多种性能和端口密度选择方案。无论端口密度如何，每个交换模块都支持多种性能优化功能，包括虚拟输出队列和超级帧。

此外，Cisco MDS 9000 系列交换模块还支持分布式转发技术(Distributed Forwarding)。分布式转发技术让系统可以在交换模块级别来制定帧转发策略，而不需要通过一个集中式的处理引擎制定这些策略。每个交换模块都拥有它自己的转发处理引擎，从而让它可以在本地制定关于将数据帧发送到哪里的策略。一旦策略制定完毕，数据帧就会被立即纳入队列，并通过交叉矩阵发送到目的地交换模块和端口。为了向交换模块提供用于制定转发策略的信息，系统 Supervisor 处理器将参与和管理所有控制协议，例如交换阵列注册和交换阵列最短路径优先 (PSPF) 升级。当这些协议发生变化时，升级信息将会从 Supervisor 处理器发送到在本地存储这些协议的交换模块。这些信息让交换模块可以非常迅速地制定它们自己的转发策略，而不需要求助于系统 Supervisor 处理器，也不需要通过交叉矩阵连接到某个集中转发处理引擎。

分布式转发的另外一个优势是用户可以方便地把其他的协议和传输方式集成到 Cisco MDS 9000 系列交换机中。例如，要在 Cisco MDS 9000 系统中加入一个新协议，用户需要设置一个具有本地转发智能的新交换模块，由它将数据帧发送到系统中其他现有的、支持同一种协议的交换模块和端口。控制协议仍然在 Supervisor 处理器中运行，转发信息的一个散列的、同步的副本将会被发送到新的交换模块。这种系统的优点在于在加入新的协议之后，仍然可以使用与原来相同的交换矩阵、机箱、电源和 Supervisor 处理器。此外，不同协议的交换模块可以共存于同一个系统，而不需要改动现有的接口交换模块。只有基于交叉矩阵的系统才能做到这一点。

## 模块化设计

Cisco MDS 9000 系列交换机建立在一个模块化的系统之上，它提供了一个非常便于扩展的产品系列。随着存储网络需求的增长，系统中可以加入更多的交换模块，以提高端口总数。利用 Cisco MDS 9500 系列中的三种型号，存储网络可以经济有效地从 16 个端口拓展到每交换阵列数千个端口。模块化设计还可以在 Cisco MDS 9000 系列的不同型号之间实现设备和配件的共用。Supervisor 处理引擎、交叉矩阵和接口交换模块都采用了独特的设计，可以在 Cisco MDS 9500 系列的所有三种型号之间互换使用。例如，一个中型交换阵列可以利用 Cisco MDS 9509 系列控制器级交换机构建，每个交换机包含三个或者四个交换模块，从而为将来的升级提供了足够的空间。如果密度需求迅速增长，Cisco MDS 9513 可以利用与 Cisco MDS 9509 相同的模块进行部署，但是前者可以提供更高的端口密度。先前的 Cisco MDS 9509 可以重新部署为边缘控制器级交换机，或者被用于扩建另外一个数据中心。无论采取什么方式，Cisco MDS 9500 系列控制器级交换机都堪称灵活的机箱和交换模块产品线，为用户提供最大限度的设计灵活性、最低限度的配件需求，减少需要管理的设备。

Cisco MDS 9216 的系统架构与 Cisco MDS 9500 系列相同，它将多层的智能和一个模块化的机型紧密地结合在一起，从而成为了业界最智能、最灵活的光纤通道交换机。随着存储网络的进一步发展，用户可以将 Cisco MDS 9000 系列模块从 Cisco MDS 9216 光纤通道交换机中卸载下来，移植到 Cisco MDS 9500 系列多层控制器中，从而获得平稳的移植、通用的部件和出色的投资保护。

## 经过优化的交换模块

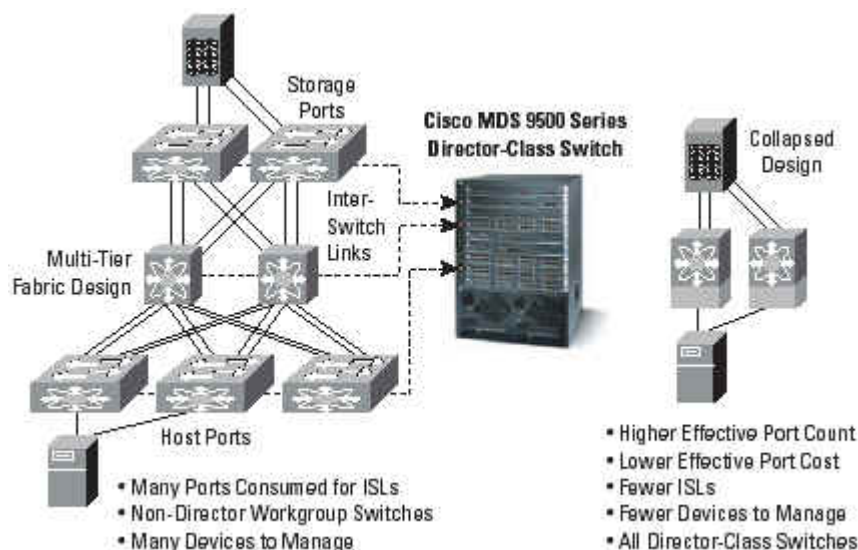
Cisco MDS 9000 系列交换机所提供的交换模块可以满足存储网络中不同的性能和端口密度要求。用市场中现有的存储网络产品设计一个规模较大的网络通常需要使用一种多层架构。在这种的配置中，常需要用高速的核心层和边缘层来提高最终网络的整体端口密度。图 6 显示

了这种设计的一个例子。依照这种设计，这种核心—边缘式部署在边缘层提供的整体带宽（相对于从边缘层到核心层的带宽）方面存在过载瓶颈问题。这种过载瓶颈问题是无法避免的，因为如果部署一个不存在过载的存储网络，网络中的每个可用端口（那些不用于交换机间连接（ISL）的端口）的成本将会迅速增加。但是，由于很多原因，这种发生在边缘的过载是可以接受的。首先，目前的存储网络主要工作在 1Gbps，很少有主机、磁盘或者磁带光纤通道连接的性能可以完全发挥出来。通常，大部分连接到存储网络的主机的最高利用率介于 20% 到 30% 之间（针对 1Gbps 的连接速率）。对于磁带阵列，因为磁带机制本身的技术限制，这个比例并不会显著提高。存储网络中用户连接带宽最高的部分无疑是磁盘子系统。但是，因为很多结构因素，这些阵列连接的突发利用率通常介于 50% 到 75% 之间，一般性使用的平均利用率则更低。在 2Gbps 光纤通道环境中，这种利用率不足的情况更加明显。此外，一些多宿主设备（包括主机和磁盘）以一种主/备用配置部署。在这种配置中，与交换阵列的连接将始终处于待命状态，直到发生某个故障为止。因此，大部分终端设备都不需要全部 1Gbps 或者 2Gbps 带宽容量。一般而言，光纤通道存储网络中利用率最高的连接通常是那些被用于互联交换机的连接，因为这种连接需要整合大量的流量。考虑到这些利用率因素，过载的边缘层不仅是接受的，而且对于成本优化来说是必不可少的。过载范围通常最低可达 3:1（成本昂贵），最高可达 16:1（成本优化），具体取决于应用服务器、磁盘和磁带的性能需求，以及每种设备的数量。

随着用户对于存储网络密度要求的不断提高，利用低密度交换机的核心—边缘式设计满足需求的能力正在不断减弱。建设大型存储网络的用户非常担心管理越来越多的低密度交换机所带来的管理负担。要达到所需要的总存储网络端口密度，就需要部署更多的交换机。但是，设备的数量也多，所需要的管理成本越高，管理的复杂性也越高。思科的用户明确地表达了在拓展存储网络的同时最大限度地减少所需交换机的需求，而 Cisco MDS 9000 系列交换机的交换模块可以满足他们的这种需求。

Cisco MDS 9000 系列交换机提供很高的端口密度（每个控制器最多 256 个端口，每个机架最多 768 个端口），让用户可以利用较少的高密度控制器建设大规模的存储网络。但是，根据用户的带宽需求和核心—边缘式设计的特性，核心层针对线速的交换机间连接进行了优化，这种连接通常需要很高的性能。边缘层通过高达 16:1 的过载系数，针对降低成本和提高端口密度进行了优化。出于性能和成本方面的考虑，思科采用了这种架构，而 Cisco MDS 9000 系列可以提供两种最主要的交换模块。

## 图 6 基于 Cisco MDS 9500 控制器级高密度交换机的折叠式设计



如前所述，对端口密度要求最高的部分是实际带宽最低的部分，即应用主机。为了满足这种需求，Cisco MDS 9000 系列提供了一种专门针对这些主机连接进行优化的交换模块。DS-X9032 是一种 32 端口的 1/2Gbps 自动检测交换模块，它的设计目的是，以一种类似于核心—边缘式设计中的边缘光纤通道交换机的方式，优化主机连接的端口密度和每端口成本。而且，DS-X9032 的配置使得每个由 4 个端口组成的群组可以共享与交叉矩阵的 2.5Gbps 带宽，平均的过载率为 3.2:1；明显优于大部分子系统供应商推荐的核心边缘设计和主机—磁盘端口连接方式。此外，DS-X9032 的一项独特优势是，它是一个功能强大的控制器级系统的组成部分。与边缘交换机不同，DS-X9032 位于一个具有冗余电源、高性能交叉矩阵和不间断软件升级等功能的控制器级系统中。DS-X9032 的配置可以为大部分常用的主机连接提供足够的带宽，让 4 端口群组中的每个连接立即增长到完全线速，并优化控制器光纤通道交换机端口的成本。DS-X9032 支持所有标准的光纤通道模式和成本，以及 Cisco MDS 9500 系列控制器级交换机的一些高级智能服务。

图 7 DS-X9032 32 端口 1/2Gbps 交换模块



对于那些需要很高持续吞吐量的光纤连接，Cisco MDS 9000 系列支持一个线速的 16 端口交换模块。DS-X9016 是一个 16 端口的 1/2Gbps 自动检测交换模块，可以在所有端口上，以 1Gbps 或者 2Gbps 的速度提供线速性能。尽管 DS-X9016 可以支持所有与 DS-X9032 交换模块相同的标准和高级智能服务，但是 DS-X9016 交换模块主要是针对矩阵中的存储子系统和 ISL 连接而设计。

图 8 DS-X9016 16 端口 1/2Gbps 交换模块





Cisco MDS 9000 系列控制器可以提供一个能够利用较少设备建设高密度交换阵列的平台。此外，高密度控制器的使用可以降低用户对于过多 ISL 连接的需求，增加有效端口（非 ISL），以及通过降低端口成本和管理成本，降低整个存储网络解决方案的成本。

### 针对网络可用性而设计

用户能够以 Cisco MDS 9000 系列交换机为平台，建立灵活、可扩展的存储网络。Cisco MDS 9000 系列的每个成员（还包括 Cisco MDS 9200 系列光纤通道交换机）都包含很多针对光纤通道的新型服务，可以提高存储网络部署的可扩展性。不仅平台本身具有很高的可扩展性，内嵌的智能协议也可将存储网络的可用性提高到今天的 SAN 交换机产品以上。

下面将概括地介绍 Cisco MDS 9000 系列交换机中包含的一些能够将可扩展性提高到现有产品的限值之上的高级功能。

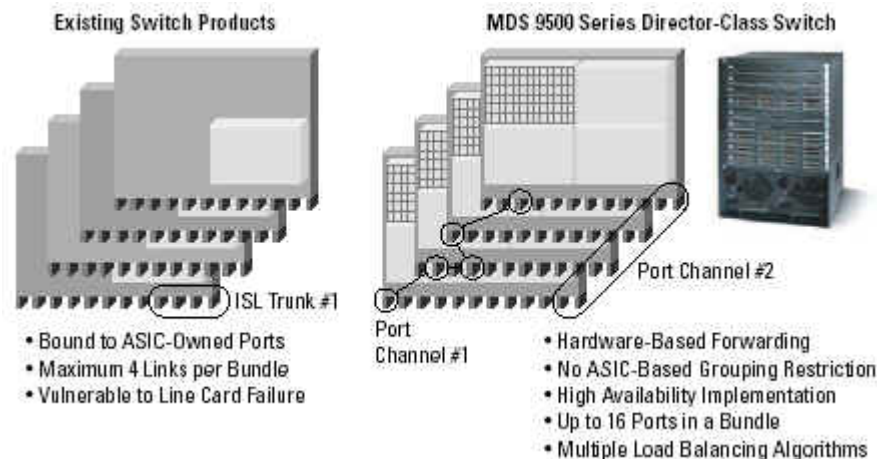
### 端口通道

随着用户不断地扩建存储网络，他们必须用更多的带宽来互联交换阵列中的交换机，以保证带宽水平。ISL 被用于实现交换机间的连接，并且通常是矩阵中利用率最高的连接，因为它们是流量的汇聚点。因为涉及到大量的带宽和 ISL 的重要性，这些连接必须采用冗余设计。利用思科的端口通道(PortChannel)技术，用户可以将最多 16 个独立的物理连接整合成一个逻辑 ISL。这不仅可以提供一个非常灵活的逻辑连接，还能在交换机之间提供高达 32Gbps 的带宽。如图 9 所示，相对于其他供应商的 ISL 端口汇聚方法，思科的端口通道技术的一个重要优势就是，捆绑的物理连接可以位于交换机的任何一个交换模块的任何一个端口上，而不会影响连接的线速性能。将物理连接分散到多个交换模块上，不仅可以防止连接故障、电缆中断和光纤故障，还可以防止交换模块发生故障。思科的端口通道技术的另外一项优势涉及到流量在逻辑 ISL 中的分发方式。目前，用户可以利用两种不同的算法，在逻辑 ISL 中的连接之间分配流量。两种算法可以通过让所有流量都与同一个物理连接上的两个终端之间的某种信息交换相关联，确保按序传输。这两种算法都完全基于硬件，从而可以确保性能不会受到任何影响。第一种算法可以根据散列数据帧的源 FCID 和目的地 FCID 来分配流量。因此，某对指定终端之间的所有流量都将通过捆绑中的同一个物理连接进行传输。第二种算法通过根据源/目的地 FCID，以及 RX\_ID/OX\_ID 标识符分发流量，提供了先进的负载均衡技术。这样，同样两个终端之间的信息交换可以通过捆绑中的多个连接转发，从而在不影响性能的情况下提高了捆绑的效率。

高可用性是用户通过采用端口通道技术所能获得的一项重要优势。如果捆绑中的某个连接发生故障，端口通道功能可以实现极为迅速地无中断恢复。逻辑 ISL 并不依靠 FSPF 进行捆绑间流量分发，并对快速检测和故障恢复进行了专门的优化。从 FSPF 的角度来看，逻辑 ISL 相当于一个单一的物理连接。

思科的光纤通道端口通道技术可以提供最大的逻辑 ISL 带宽、极快的恢复、最高的可用性，并能提高总带宽的有效利用率。

图 9 思科的端口通道(PortChannel)功能可以提供最大限度的可用性和带宽



### 转发拥塞控制(Forward Congestion Control)

随着存储网络的不断扩展，流量拥塞正在日益成为一个严重的问题。规模扩大的多层网络会带来更多的流量汇聚点，从而提高了流量拥塞的风险。即使只有一个连接发生拥塞，也会对整个网络产生严重的影响。转发拥塞控制（FCC）是 Cisco MDS 9000 系列特有的一种功能。FCC 让交换阵列中的一个发生拥塞的端口能够检测到拥塞源，进而限制这些特定端口所使用的带宽。一旦检测到拥塞源，发生拥塞的端口就会通知导致拥塞的端口降低输入流量。当源端口接收到该信息时，它们将通过降低发放信用点的速度，限制获得缓存—缓存信用点 (Buffer-Buffer Credit) 的节点设备。这个减缓发放信用点的动作可以降低节点设备所能使用的有效带宽，从而缓解拥塞。在拥塞得到缓解或者消除之后，源端口就会把向节点设备发放信用点的速度恢复到正常速度。在一个包含多个交换机的大规模交换阵列中，FCC 可以在不造成流量拥塞的情况下，提供必要的平衡和控制，从而平等地为所有应用服务器提供服务。

### VSAN

目前，很多 SAN 设计人员都出于各种原因建设独立的存储网络。在这种情况下，一个独立的存储网络指的是一个在物理上完全隔离的、用于将主机连接到存储的交换机或者交换机群组。下面列出了一些常见的原因：

1. **高可用性**——一种常见的做法是建设多个并行的交换阵列，并在并行的、物理上隔离的交换阵列中加入“多宿主”主机和磁盘。一般而言，采用这种隔离方式的主要原因是确保各个交换阵列的服务（例如名称服务）彼此隔离。如果某个交换阵列服务发生故障，它不会影响其他的并行交换阵列。因此，并行交换阵列可以提供从主机到磁盘的隔离路径。
2. **应用和备份交换阵列**——很多用户为他们的存储网络环境建设至少两个物理上独立的交换阵列。他们的主要目的是将一个交换阵列专门用作应用主机，另外一个专门用于备份环境。利用这种方式，备份流量在物理上与主应用流量完全隔离。
3. **部门交换阵列**——很多用户选择为部门应用扩展独立的存储网络环境。在这种情况下，用户会为每个部门应用建设一个独立的、规模较小的交换阵列。
4. **多种 OS 的交换阵列**——有些用户的做法是为使用不同操作系统的主机建设独立的交换阵列。由于一些操作系统的特性以及它们发现和使用存储的方法不同，很多用户会用不同的交换阵列隔离环境。典型例子之一是一个 Sun Solaris 交换阵列和一个 Windows

NT/2K 交换阵列。

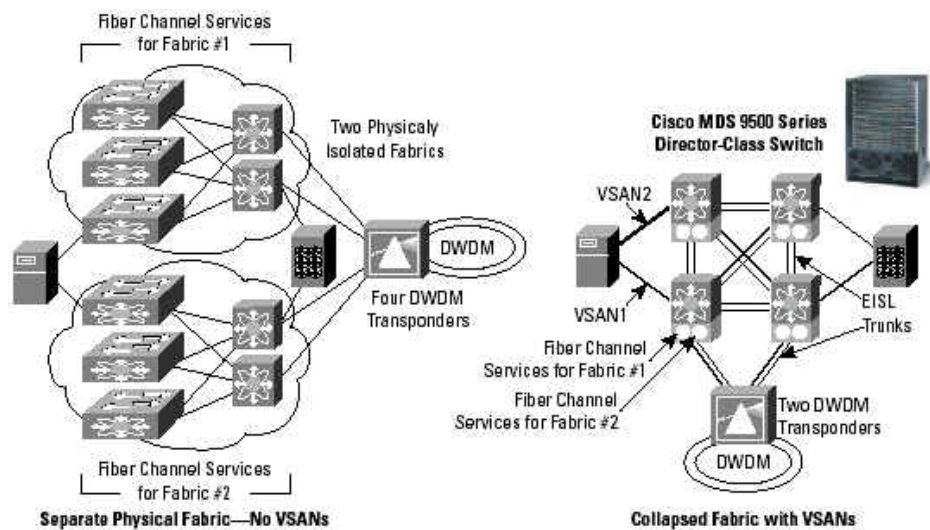
尽管每个原因都为建设独立的交换阵列提供了有力的依据，但是这种做法的确相当浪费。添加独立的交换阵列意味着购买更多的硬件、投入更多的资金，而硬件通常无法得到充分的利用。

为了实现同样隔离的环境，同时避免为了建设物理上隔离的交换阵列而花费大量资金，思科在 Cisco MDS 9000 系列中推出了虚拟 SAN（即 VSAN）功能。VSAN 可以提供在共享同一套物理基础设施上建设独立虚拟交换阵列的能力。利用 ISL 链路上的一种基于硬件的帧标签机制，各个虚拟交换阵列都彼此隔离。EISL 链路是一种增强的 ISL 链路，包括用于每个帧的附加标签信息，并可以支持互联任何 Cisco MDS 9000 系列交换机产品的链路。VSAN 的设置建立在物理端口的基础上，没有哪个物理端口可以属于多个 VSAN。因此，连接到一个物理端口的任何节点都将成为该端口的 VSAN 的成员。

VSAN 为用户提供了巨大的灵活性。例如，Cisco MDS 9000 系列产品在每个物理基础设施中支持 1024 个 VSAN。用户可以选择地在一个 EISL（增强 ISL）链路上添加或者去除 VSAN，以控制 VSAN 的范围。此外，它还可以提供了特殊的流量计数器，以跟踪每个 VSAN 的统计数据。

VSAN 最受欢迎的特性很可能是它的高可用性。VSAN 不仅可以提供严格的硬件隔离，还可以为每个新的 VSAN 创建一套完整的交换阵列服务。因此，当用户创建某个新的 VSAN 时，一套完全独立的服务（包括名称服务器、分区服务器、域控制器、别名服务器和注册服务器）将会在支持这个新的 VSAN 的交换机上创建并启用。这种复制的服务能在同一个物理基础设施上建设独立的环境，消除人们在 HA 方面的担忧。例如，在 VSAN1 中安装一个主分区集不会对 VSAN2 中的交换阵列造成任何影响。

图 10 利用 Cisco MDS 9000 系列的 VSAN 进行整合



VSAN 还可以通过一个远程基础设施，连接远程数据中心中的独立交换阵列。因为帧标签功能由硬件完成，并且包含在每个 EISL 帧之中，所以它可以通过密集波分复用（DWDM）或者稀疏波分复用（CWDM）等传输方式进行传输。因此，来自于多个 VSAN 的流量可以复合到一对光纤上，并可以在更长的距离上传输，同时仍然保持完全的隔离。VSAN 让用户可以利用一个通用的冗余物理基础设施来建设灵活、隔离的交换阵列，以实现 HA 目标，从而将可扩展性提高到了一个新的等级。

### 远距离传输性能

今天的业务连续性必须能够远程传输光纤通道。尽管一般用户通常只会将他们的存储网络拓展 5-50 公里，但是也有一些用户希望将他们的数据中心拓展到另外一个海岸或者大陆。为了在如此远的距离上传输光纤通道协议，任何一端的扩展连接上的光纤通道端口必须支持高数量等级的缓存—缓存信用点(Buffer-Buffer Credit)。B2B 信用点让发送端有权发送一个数据帧。当数据帧到达远端时，发送端会再发送一个数据帧。但是，如果距离或者相关延时过长，用户可能获得有限的有效带宽，这是因为损失了等待确认的时间。

因此，在较长的距离上传输光纤通道的关键是为发送端提供大量的 B2B 信用点。这种方法让发送端可以在等待返回确认信息的过程中，不断地在发送通道中装满大量的数据帧。根据光速计算，通常情况下每两公里距离就需要一个 B2B 信用点，以避免限制带宽。

对于市场中的大部分交换机，通过购买一个额外的使用许可，最多可以支持 60 个信用点。如果距离不超过 30 公里，用户只能建立一个 1Gbps 的传输通道。但是，利用 Cisco MDS 9000 系列，每个端口最多可以支持 255 个 B2B 信用点。此外，Cisco MDS 9000 支持一种可以提供 FC over IP 的 IP 存储服务模块。利用 Cisco MDS 9000 系列交换机的增强功能，用户能够以线速带宽，在扩展的连接上将光纤通道拓展到大约 128 公里。利用 Cisco MDS 9000 系列，用户可以在 TCP/IP、DWDM/CDWM 或者裸光纤的基础上，灵活地拓展光纤通道。

### 结论

今天的用户正处于实施存储网络的不同阶段。刚刚开始扩展 SAN 的用户面临着一个严峻的挑战，即如何选择能够满足未来需要的设计和 product。这些未来的需求通常很难估计，并且可能



在没有什么警示的情况下出现指数式的增长。而那些在建设更大规模的存储网络方面已经取得很大进展的用户则可能已经遇到了如何利用性能和端口密度都较低的交换机扩展一个存储网络基础设施的问题。所有用户都非常关心如何优化他们所购买和部署的存储网络资源的利用率。

利用 Cisco MDS 9000 系列交换机产品，思科提高了存储网络产品的性能和可扩展性标准。这个系列是业界端口密度较高的存储网络产品。它的四个型号（即 2、6、9、13 个插槽）让用户可以扩建他们的各种规模的网络，满足目前的需求。用户还可以确信的是，他们目前部署的网络将能够扩展到非常大的规模，满足未来出现的任何需求。

思科在设计 Cisco MDS 9000 系列交换机的所有能够提高可扩展性和性能的功能时，都考虑到了最高的可用性。每种可扩展性功能都具有迅速融合能力和可预测性，从而可以确存储网络的未来的发展——且这不会以牺牲可用性为代价。