

# 堆叠技术白皮书

## 摘要

堆叠技术，是非机箱式交换机所提供的技术，使得交换机可以通过堆叠模块或堆叠线缆，以一定的连接方式堆叠。堆叠交换机通过统一的带外通道(串口)或者管理 IP，进行配置和管理。锐捷交换机支持的堆叠数量为 2-8 台，在通过堆叠设备增加端口数量的同时，不增加配置的复杂度。

本文将介绍现在流行的堆叠技术，如菊花链式堆叠、星形堆叠等；通过对堆叠协议的简单介绍，让读者了解堆叠的建立过程；之后，将介绍锐捷不同系列交换机的堆叠使用的技术；最后，介绍各种锐捷交换机系列产品在堆叠时注意的使用限制，以及堆叠的相关配置和信息查看。

本技术白皮书仅针对 S20 系列、S21 系列、S23 系列、S26 系列、S32 系列、S3750 系列和 S5750 系列交换机的堆叠技术介绍，不包含 S29 系列和 S3760 系列产品，S29 系列和 S3760 系列产品请查阅其他相关文档。

## 关键词

堆叠 堆叠技术 堆叠协议

# 目 录

|                                    |    |
|------------------------------------|----|
| 摘要 .....                           | 1  |
| 关键词 .....                          | 1  |
| 1 概述 .....                         | 2  |
| 1.1 堆叠技术的提出 .....                  | 2  |
| 1.2 业界流行的堆叠连接方式 .....              | 3  |
| 1.2.1 菊花链式堆叠 .....                 | 3  |
| 1.2.2 星形堆叠 .....                   | 5  |
| 1.3 堆叠系统的术语概述 .....                | 6  |
| 1.3.1 堆叠主机和堆叠从机 .....              | 6  |
| 1.3.2 混合堆叠 .....                   | 7  |
| 1.4 产品支持 .....                     | 7  |
| 2 技术介绍 .....                       | 7  |
| 2.1 堆叠协议的运作 .....                  | 7  |
| 2.1.1 本机堆叠数据库的初始化 .....            | 8  |
| 2.1.2 堆叠拓扑发现 .....                 | 9  |
| 2.1.3 堆叠拓扑的分析 .....                | 11 |
| 2.1.4 堆叠完成的系统初始化 .....             | 11 |
| 3 锐捷堆叠技术特点 .....                   | 11 |
| 3.1 各款支持堆叠的交换机的参数比较 .....          | 11 |
| 3.2 锐捷交换机的不同堆叠方式 .....             | 13 |
| 3.2.1 虚拟堆叠 .....                   | 13 |
| 3.2.2 千兆堆叠 .....                   | 13 |
| 3.2.3 增强型堆叠 .....                  | 13 |
| 3.3 交换机堆叠时的功能限制 .....              | 13 |
| 3.3.1 虚拟堆叠的限制 .....                | 13 |
| 3.3.2 千兆堆叠的限制 .....                | 14 |
| 3.3.3 增强型堆叠的限制 .....               | 14 |
| 3.3.4 堆叠共同的限制 .....                | 14 |
| 3.4 易用的堆叠管理 .....                  | 15 |
| 3.5 其他注意事项 .....                   | 19 |
| 3.5.1 S3750 系列交换机的特殊配置 .....       | 19 |
| 3.5.2 S21 和 S26 交换机堆叠线缆的特殊连接 ..... | 20 |
| 3.5.3 设备优先级的配置 .....               | 21 |
| 3.5.4 堆叠设备别名的配置 .....              | 21 |
| 3.5.5 堆叠参数的保存 .....                | 22 |
| 3.5.6 堆叠起机时间 .....                 | 22 |
| 3.5.7 环形堆叠和线形堆叠的区别 .....           | 22 |
| 3.5.8 堆叠设备加入移除造成的影响 .....          | 23 |
| 4 结束语 .....                        | 23 |

# 1 概述

## 1.1 堆叠技术的提出

对于盒式交换机，由于其成本较低，经常用在网络的汇聚层和接入层。但是由于网络的规模不断提高，此时盒式交换机的缺点就暴露出来，端口数量的相对固定，不像机箱式的交换机，在端口出现不足的时候，能简单的通过增加线卡的方式添加端口容量，也就是说，盒式交换机的端口数量相对固定，如 24 个 10/100/1000M 端口和 48 个 10/100/1000M 端口的交换机，如图 1，就是锐捷的一款盒式交换机产品 S3250-24(上方)和 S3250-48(下方)：



图 1 锐捷 S3250 系列产品

对于图 1 所示的 S3250 系列产品，S3250-24 产品有 24 个百兆口，S3250-48 产品有 48 个百兆口，通过堆叠，将 1 台 S3250-24 和 1 台 S3250-48 混合堆叠，这样堆叠系统中就有了  $24+48=72$  个百兆口了。再举个例子，如果 8 台 S3250-48 堆叠，那么端口就可以扩展为  $48*8=384$  个百兆口了。在扩展了端口后，不需要对每台交换机进行分别管理，而只需要将串口接在主机上，或者通过配置管理 IP 进行 telnet 管理，在统一的界面下，即可以对整个堆叠系统中的所有端口进行统一的配置，同时，整个堆叠系统将共享同一个 MAC 地址表和路由表，也就是说，整个堆叠系统共享二层和三层的转发策略。

## 1.2 业界流行的堆叠连接方式

### 1.2.1 菊花链式堆叠

菊花链式堆叠是交换机级联连接方式的一种扩展，其连接方式示意如下：

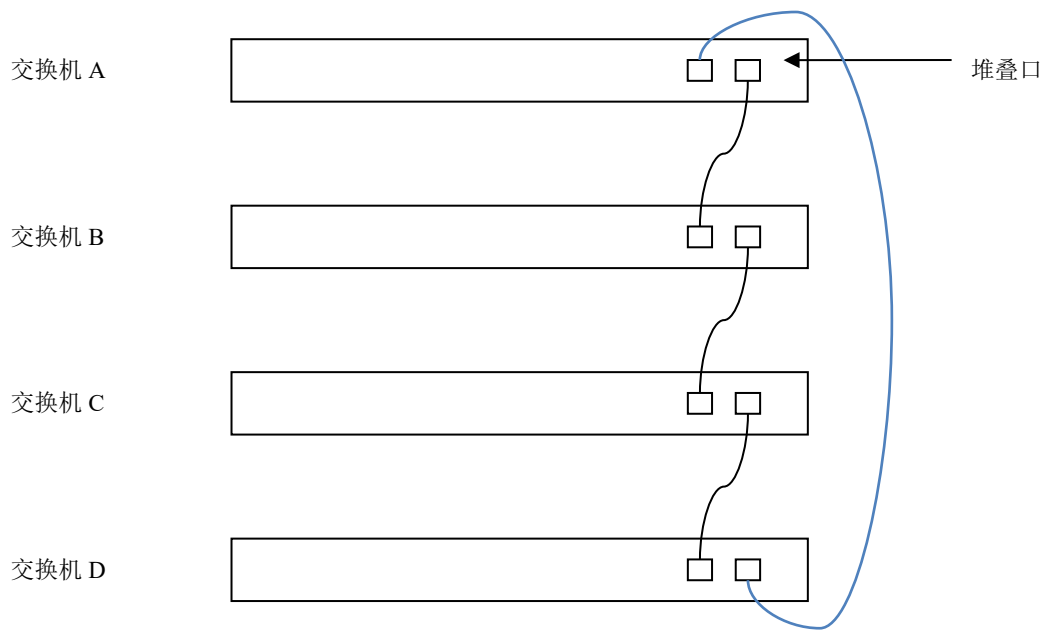


图 2 菊花链式堆叠连接示意图

如图 2 所示，菊花链式堆叠就是使用堆叠线缆将要堆叠的交换机通过彼此相连的方式连接起来，交换机 D 到交换机 A 的连接的链路作为多余链路被裁剪，在实际运作过程中，是不参与交换的。

#### 1.2.1.1 环形堆叠和线形堆叠的区别

如图 2 所示，蓝色标注的堆叠线为冗余链路，此时交换机堆叠处于环形堆叠状态；而如果不接那条蓝色冗余链路，则此时交换机处于线性堆叠状态，如图 3 所示：

图 3

环形堆叠相对与线形堆叠的好处在于，环形堆叠的冗余链路将会在其他正常链路的某一条出故障时，承担故障链路的工作，从而能保证系统的高稳定性。

### 1.2.1.2 单工堆叠和双工堆叠

菊花链式堆叠可分为单工堆叠和双工堆叠(也可以称为半双工堆叠和全双工堆叠),二者区别在于,堆叠口支持双工的,仅需要 1 条堆叠线缆将两个堆叠端口互联的(同时包括 TX 和 RX 方向),为双工堆叠,如图 2 和图 3 所示;而单工堆叠,堆叠口是以堆叠模块上两个端口(实际上是千兆端口的输出和输入方向)为一组的,包括了 TX 和 RX 两个方向,需要使用 2 条堆叠线缆进行堆叠口的连接,同时,需要交叉连接。

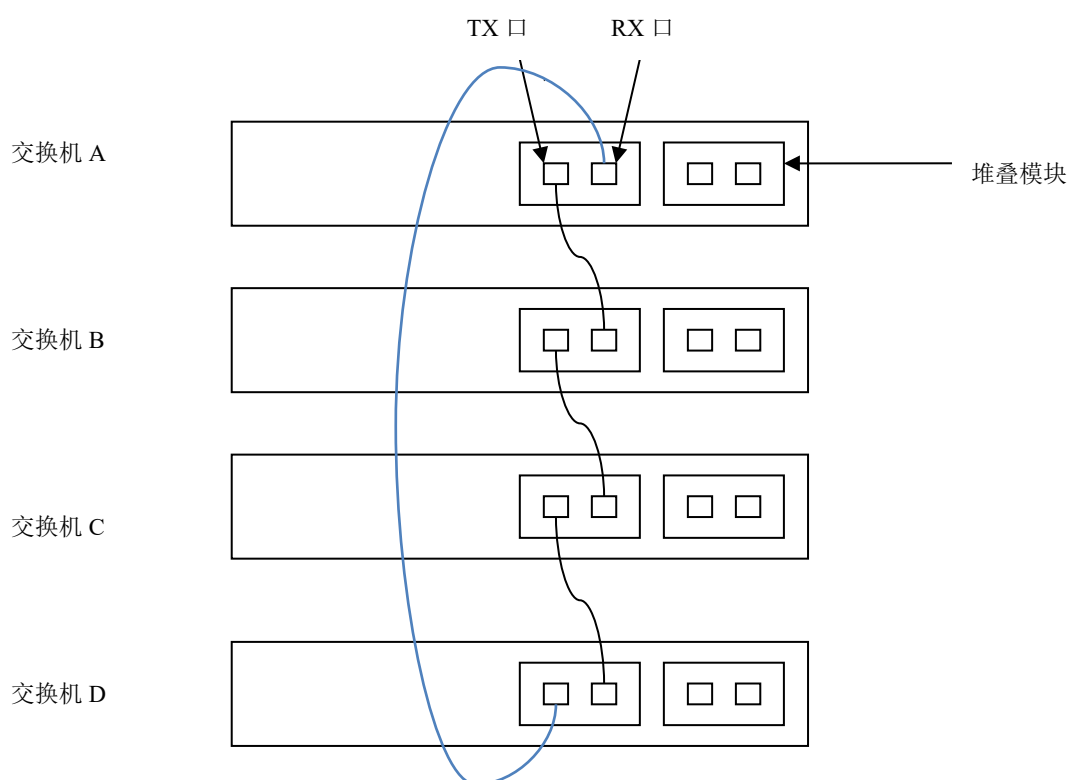


图 4 半双工堆叠

如图 4 所示,一组堆叠口含有一个 TX 端口和一个 RX 端口,连接时,需要将本交换机的 TX 端口连接到另一台交换机的 RX 端口,直到连成环状,这样的堆叠方式只要有一条堆叠链路出问题,就会导致整个堆叠系统出现问题,故单工堆叠不支持线形堆叠。

### 1.2.2 星形堆叠

星形堆叠是另一种较少应用的堆叠技术,通过将多台交换机连接到同一个交换中心上(称为 CORE/MATRIX),所有的堆叠主机都通过高速的堆叠接口,上联到统一的堆叠中心。

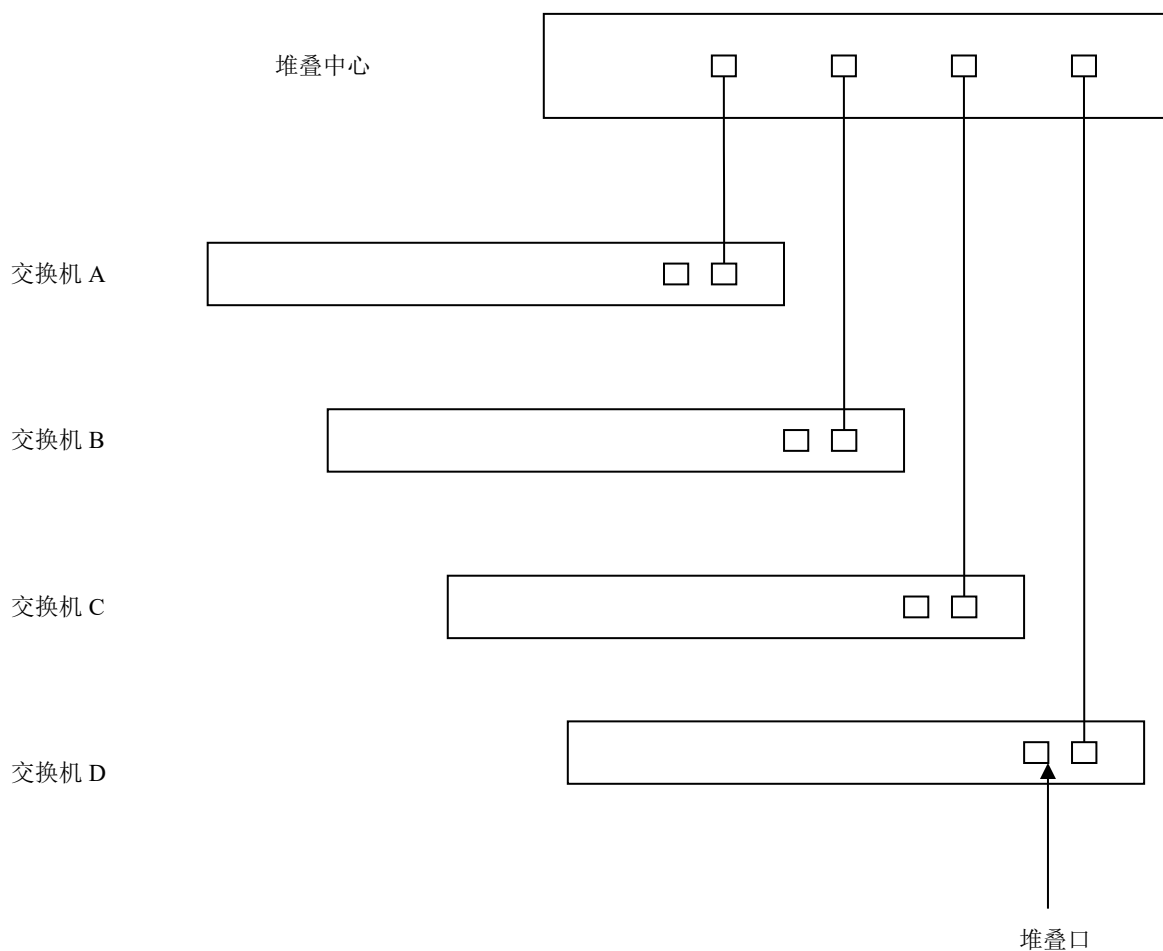


图 5 星形堆叠

如图 5，星形堆叠的连接，使用堆叠中心将 4 台交换机连接到一起，由堆叠中心保证交换的带宽。其中，堆叠中心和交换机 A 是可以集成在一起的，同时扮演交换机和堆叠中心的角色。从菊花链式堆叠和星形堆叠的示意图比较可以发现，星形堆叠的优势在于一台交换机到另一台交换机之间的转发路径要比菊花链式堆叠短，不会出现跨多台交换机的情况。但是，星形堆叠的成本较高，需要专用的堆叠中心支持，也与机箱式交换机的角色重复。

## 1.3 堆叠系统的术语概述

### 1.3.1 堆叠主机和堆叠从机

堆叠系统需要确立一台设备作为堆叠的主机，堆叠主机的串口可以提供用户访问进行管理，同时，在堆叠的主机上进行管理 ip 的配置，使得可以通过 telnet/web/snmp 各种方式对交换机堆叠系统进行统一的管理。

堆叠从机是堆叠系统中除堆叠主机外的交换机，无法使用从机的串口对堆叠从机进行管理，堆叠从机的串口仅供技术人员在定位故障时使用的调试界面。

### 1.3.2 混合堆叠

一般来说，只有相同系列交换机才能堆叠成功，也就是说，不同系列的交换机是无法堆叠成功的。相同系列不同型号的交换机之间可以堆叠成功，即可以混合堆叠。以 S3250 系列的交换机来说，S3250-48 和 S3250-24 是两个型号交换机，端口数目不同，但是可以混合堆叠成功。

## 1.4 产品支持

目前锐捷的交换机中，支持堆叠的系列包括 S20 系列、S21 系列、S23 系列、S26 系列、S32 系列、S3750 系列和 S5750 系列，支持的版本有 RGNOS 1.69/1.7 以及 RGNOS 10.1(4)及以后版本。

目前锐捷的交换机均只支持菊花链式堆叠。

## 2 技术介绍

本章介绍菊花链式堆叠下，堆叠协议的运作方式。在交换机通过堆叠线缆连接好上电后，通过堆叠发现、设备检测和拓扑分析，完成堆叠的过程，从而建立起堆叠系统，确立主从机地位，通过配置分发使得系统得以最终运行起来。

### 2.1 堆叠协议的运作

堆叠系统中的交换机，运行着一个堆叠的协议，随着堆叠进程的不断运行，堆叠状态机从一个状态跃迁到另一个状态，状态跃迁如图 6 所示：

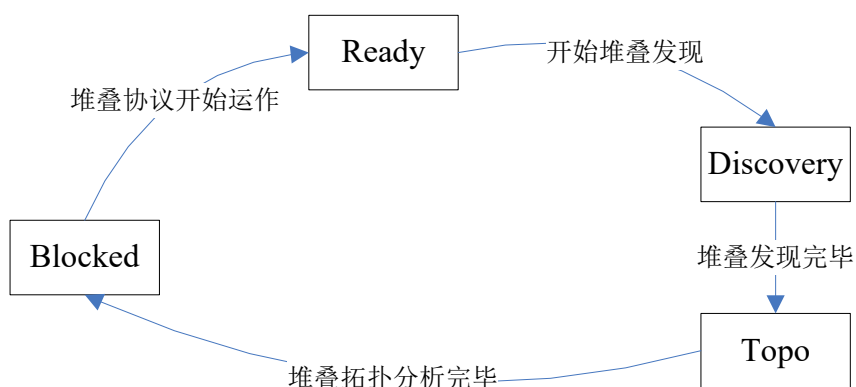


图 6

这里的状态包括：

- **Blocked:** 堆叠检测状态机挂起，此时设备可能刚刚上电、工作在单机模式下、完成堆叠检测的过程；
- **Ready:** 开始进行堆叠检测；
- **Discovery:** 正在搜集堆叠成员信息；

- **Topo:** 主机进行堆叠拓扑的配置。

从堆叠系统中的交换机上电，到整个系统堆叠完毕，需要经历以下基本过程：

- 本机堆叠数据库的初始化
- 堆叠拓扑发现
- 堆叠拓扑分析
- 堆叠完成的系统初始化

下面的章节中，就将分别介绍基本过程。

### 2.1.1 本机堆叠数据库的初始化

对于堆叠系统中的某台交换机，在堆叠系统开始运作之前，需要初始化本地和堆叠相关的数据库，用于在堆叠协议运行时，拓扑检测用到的与堆叠系统中其他交换机交换的堆叠信息。

堆叠数据库记录了一台交换用于堆叠的相关信息，包括的基本信息如下：

- 本机的序列号
- 本机的 MAC 地址
- 配置的堆叠优先级，未配置默认为 1
- 堆叠口数量
- 堆叠口的端口编号
- 本机的版本信息，包括 **BOOT** 版本、**CTRL** 版本和主程序版本，还有硬件版本

在堆叠拓扑检测的过程中，本机的堆叠数据库信息会打包到报文中，并发送到堆叠系统中的每台交换机上，到堆叠检测结束，每台交换机上都有所有交换机的堆叠数据库信息。

### 2.1.2 堆叠拓扑发现

在本机堆叠数据库初始化完毕后，堆叠任务就会开始拓扑发现(或者成为拓扑检测)。堆叠发现分为 3 个阶段：

- **Probe:** 使用 **probe** 报文，用于决定堆叠口的连接，搜集对端的堆叠口信息；
- **Routing:** 使用 **routing** 报文，将本地搜集到的堆叠口信息传播到所有邻居；
- **Config:** 使用 **config** 报文，用于各交换机间同步堆叠检测的结束状态。

### 2.1.2.1 Probe 阶段

为了支持单工/双工两种模式的堆叠，Probe 报文可分别用于记录堆叠口的发送/接收(TX/RX)方向的连接信息。本设备发出的 probe 报文记录着本机已知的堆叠口信息，当其他设备接收到 probe 报文后，将本机的堆叠口信息增加到该报文中，并转发到本地所有的堆叠口，最终这个 probe 报文回到发送者时就已经搜集到其他设备上已知的堆叠口信息，这些信息被更新到堆叠数据库中。

### 2.1.2.2 Routing 阶段

Routing 报文用于邻接成员之间交换堆叠信息(也就是本交换机堆叠口直连的对端交换机)，Routing 报文只发送到邻居节点，报文中包含当前所有已知的堆叠信息，包括经过的每个节点的堆叠数据库信息，通过这样的方式,堆叠的信息迅速在系统中传播。

### 2.1.2.3 Config 阶段

当本地堆叠口的接收/发送方向的信息都已搜集到后，本地交换机的堆叠状态标记为本地完成，表示本地设备已经完成堆叠发现工作。当堆叠数据库里的所有其他已知设备的堆叠状态都标记为本地完成，则本地交换机的堆叠状态就标记为全部完成，标识拓扑中的所有交换机设备都完成了堆叠发现工作。随后在每个设备的堆叠数据库基础上开始进行主从选举。

主从选举的决定因素有：

- 用户配置的优先级
- 交换机的 MAC 地址

其中用户优先级从 1 到 10 逐渐升高，MAC 地址较小的优先，总的来说，选择配置的用户优先级越大、MAC 地址越小的，则会被选举为主机。

举几个两台堆叠的例子来具体说明，3-8 台堆叠的情况可以通过两两比较的方式得到：

- 配置的优先级不同，MAC 地址小的配置优先级高：

堆叠设备 1，MAC 地址为 00d0f8000001，配置的优先级为 10

堆叠设备 2，MAC 地址为 00d0f8000002，配置的优先级为 9

那么堆叠设备 1 将会被竞选为主机。

- 配置的优先级不同，MAC 地址小的配置优先级低：

堆叠设备 1，MAC 地址为 00d0f8000001，配置的优先级为 9

堆叠设备 2，MAC 地址为 00d0f8000002，配置的优先级为 10

那么堆叠设备 2 将会被竞选为主机。

- 配置的优先级相同：

堆叠设备 1，MAC 地址为 00d0f8000001，配置的优先级为 10

堆叠设备 2，MAC 地址为 00d0f8000002，配置的优先级为 10

那么堆叠设备 1 将会被竞选为主机。

某个设备被选举为主机后，开始向系统中的其他成员发出 Config 报文，，当所有的成员都发送回应后，系统完成堆叠发现的过程，如果在一定的时间间隔内没有回应，则认为堆叠检测失败，给出错误提示。

### 2.1.3 堆叠拓扑的分析

完成堆叠拓扑的发现过程后，堆叠的主机与从机之间的堆叠数据库已经完成同步，主机的选举也已完成，堆叠系统进入拓扑分析过程。拓扑分析的目的是：允许堆叠系统的端口之间能够互相通讯；防止堆叠系统形成环路，也就是在环形堆叠的环境中，要切除冗余的那条堆叠链路；保证两个堆叠口之间的通路为最短路径。

### 2.1.4 堆叠完成的系统初始化

由于在堆叠协议运行的过程中，要防止数据报文影响堆叠，在堆叠协议运作开始前，要将所有的交换端口都设置为 shutdown 状态，在堆叠拓扑分析完成后，系统将恢复正常交换口的交换，即设置为 no shutdown 状态。

## 3 锐捷堆叠技术特点

在锐捷的盒式交换机中，使用以上堆叠协议的交换机包括 S20 系列、S21 系列、S23 系列、S26 系列、S32 系列、S3750 系列和 S5750 系列，对于 S29 系列、S3760 系列和 S5760 系列不在本文的介绍范围内。

### 3.1 各款支持堆叠的交换机的参数比较

| 比较项       | S21 系列            | S20 系列            | S23 系列            | S26 系列            | S32 系列                              | S3750 系列                            | S5750 系列                                                                                              |
|-----------|-------------------|-------------------|-------------------|-------------------|-------------------------------------|-------------------------------------|-------------------------------------------------------------------------------------------------------|
| 涵盖的<br>产品 | S2126G、<br>S2150G | S2026G、<br>S2052G | S2328G、<br>S2352G | S2628G、<br>S2652G | S3250-24、<br>S3250P-24、<br>S3250-48 | S3750-48、<br>S3750-24、<br>S3750E-24 | S5750-24SFP/12GT、<br>S5750-24GT/12SFP、<br>S5750-48GT/4SFP、<br>S5750S-24GT/12SFP、<br>S5750S-48GT/4SFP、 |

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/806235043012010211>