



# iStack 技术白皮书

文档版本 01  
发布日期 2012-08-15

华为技术有限公司



**版权所有 © 华为技术有限公司 2012。 保留一切权利。**

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

## 商标声明



HUAWEI 和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

## 注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

# 华为技术有限公司

地址：                  深圳市龙岗区坂田华为总部办公楼                  邮编：518129

网址：                  <http://www.huawei.com>

客户服务邮箱：      [support@huawei.com](mailto:support@huawei.com)

客户服务电话：      0755-28560000 4008302118

客户服务传真：      0755-28560111

---

# 目 录

---

<b>1 iStack .....</b>	<b>1</b>
1.1 介绍.....	1
1.2 原理描述.....	2
1.3 应用场景.....	5
1.4 故障处理案例.....	9
1.4.1 设备无法组建堆叠.....	9
1.5 FAQ.....	10
<b>A 术语与缩略语 .....</b>	<b>A-1</b>

# 1 iStack

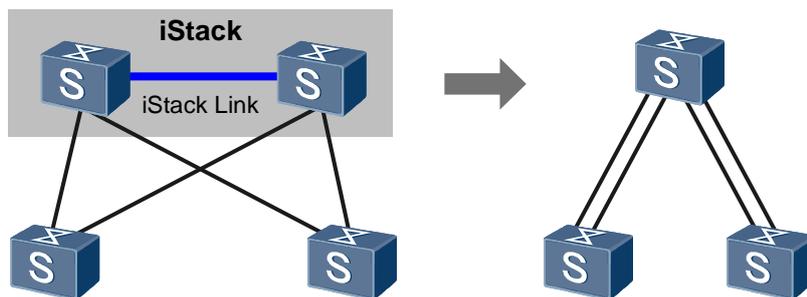
## 1.1 介绍

目前，网络中主要存在两种形态的通信设备：盒式设备和框式设备。

- 盒式设备成本低廉，但是没有高可用性支持，缺乏不中断的业务保护，无法应用于重要的场合（例如核心层、汇聚层、数据中心等）；在复杂的组网环境中，盒式设备扩展性差，用户不得不维护更多的网络设备，并且为了增加这些设备修改早期的组网结构。
- 框式设备具有高可用性、高性能、高端口密度的优点，因此经常被应用于一些重要场合（例如核心层、汇聚层、数据中心等）。但它相比盒式交换机也有一些缺点，比如首次投入成本高、单端口成本高等。

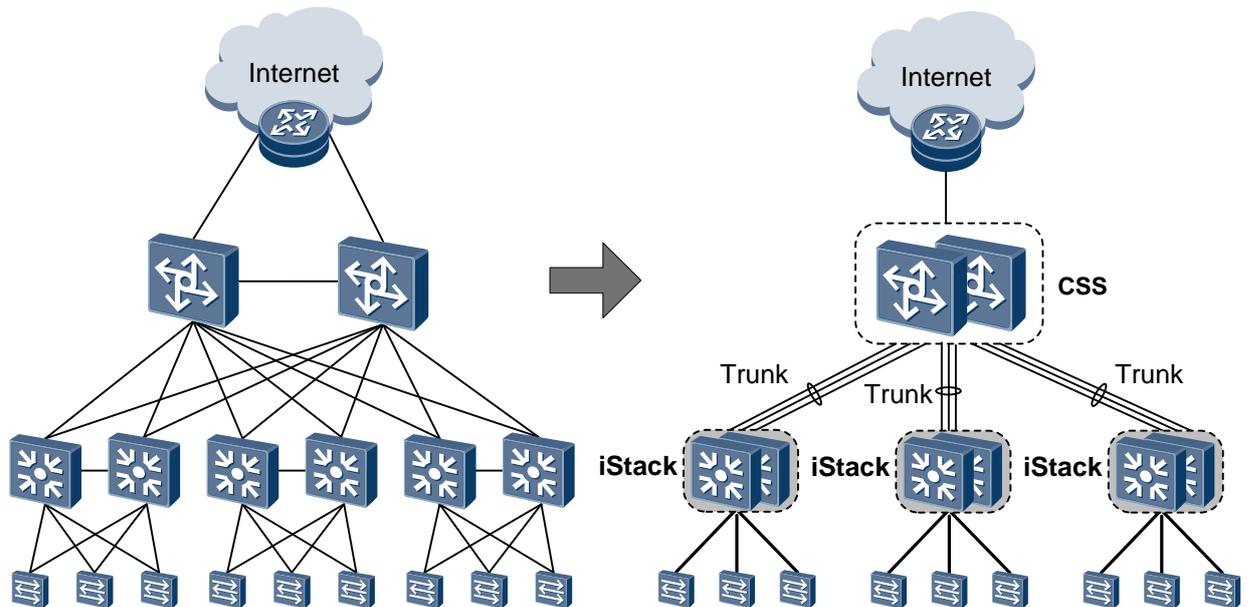
针对盒式设备与框式设备的这些特点，一种结合了两种设备优点的 iStack 堆叠技术应运而生。iStack 堆叠是指将多台支持堆叠特性的交换机设备组合在一起，从逻辑上组合成一台整体交换设备，如图 1-1 所示。这种虚拟设备既具有盒式设备的低成本优点，又具有框式设备的扩展性以及高可靠性优点。

图1-1 堆叠示意图



iStack 是一种虚拟化技术，在不改变网络物理拓扑连接结构条件下，将网络同一层的多台设备虚拟化成单台逻辑设备，达到简化网络结构、简化网络协议部署、提高网络可靠性和可管理性的目的，如图 1-2 所示。

图1-2 网络横向虚拟化



## 技术优点

堆叠主要具有以下优点：

- 高可靠性。堆叠系统多台成员设备之间 1:N 冗余备份；堆叠支持跨设备的链路聚合功能，实现跨设备的链路冗余备份。
- 强大的网络扩展能力。通过增加成员设备，可以轻松的扩展堆叠系统的端口数、带宽和处理能力。
- 简化网络结构和协议部署。堆叠技术可以将复杂的网络拓扑结构简化为层次分明、互联关系简单的网络结构，网络各层之间通过链路聚合，自然消除环路，不需要再部署 MSTP、VRRP 等协议。
- 简化配置和管理。堆叠形成后，多台物理设备虚拟成为一台设备，用户可以通过任何一台成员设备登录堆叠系统，对堆叠系统所有成员设备进行统一配置和管理。

## 1.2 原理描述

### 基本概念

#### 角色

堆叠中所有的单台设备都称为成员设备，按照功能不同，可以分为三种角色：

- 主交换机，即 Master，负责管理整个堆叠系统。堆叠系统中只有一台主交换机。

- 备交换机，即 **Standby**，是主交换机的备份交换机。当主交换机故障时，备交换机会接替原主交换机的所有业务。堆叠系统中只有一台备交换机。
- 从交换机，即 **Slave**，除主交换机外，堆叠中所有交换机都是从交换机，其中备交换机同时承担备交换机和从交换机两种角色。

### 堆叠 ID

堆叠 ID，即成员编号（Member ID），用来标识和管理成员设备，堆叠系统中所有成员设备的堆叠 ID 都是唯一的。

### 堆叠优先级

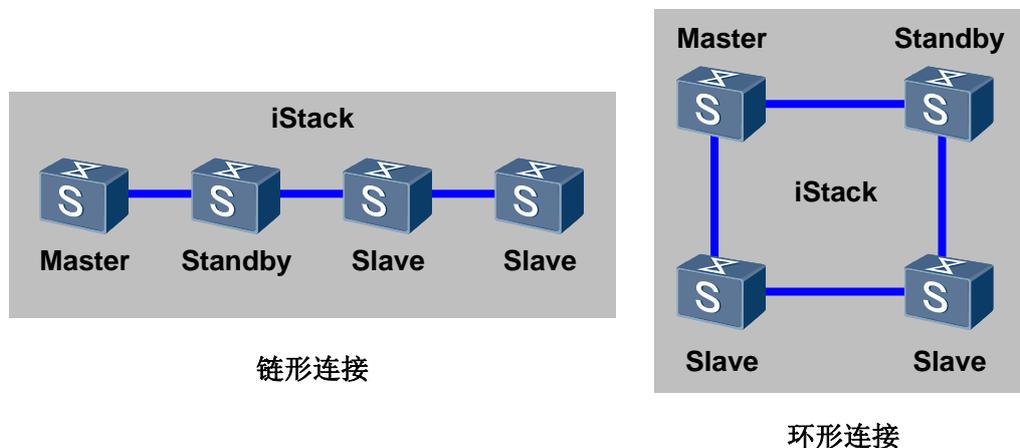
堆叠优先级是成员设备的一个属性，主要用于角色选举过程中确定成员设备的角色，优先级值越大表示优先级越高，优先级越高当选为主交换机的可能性越大。

## 堆叠物理连接

堆叠的连接拓扑有两种：

- 链形连接：使用堆叠电缆将一台设备的左口（右口）和另一台设备的右口（左口）连接起来，依次类推，第一台设备的右口（左口）和最后一台设备的左口（右口）没有连接堆叠电缆。
- 环形连接：将链形连接第一台设备的右口（左口）和最后一台设备的左口（右口）连接起来。

图1-3 堆叠连接拓扑示意图



环形连接比链形连接有更高的可靠性，因为当链形连接中出现链路故障时，会引起堆叠分裂，而环形连接中某条链路故障时，会形成链形连接，堆叠的业务不会受到影响。

## 堆叠竞争规则

堆叠系统由多台成员交换机组成，每台成员交换机具有一个确定的角色。堆叠建立时，成员交换机间相互发送堆叠竞争报文，选举出主、备、从交换机。主交换机收集成员信息并计算堆叠拓扑，将堆叠拓扑信息同步到所有的成员交换机。

主交换机的竞争规则如下（依次从第一条开始判断，直至找到最优的交换机才停止比较）：

1. 运行状态比较，已经正常运行的交换机优先处于启动状态的交换机竞争为主交换机。
2. 堆叠优先级比较，堆叠优先级高的交换机优先竞争为主交换机。
3. MAC 地址比较，MAC 地址小的交换机优先竞争为主交换机。

备交换机的选举规则与主交换机的选举规则相同。

堆叠系统建立之前，每台交换机都是单独的实体，每台交换机有自己独立的 IP 地址，用户需要独立的管理所有的交换机；堆叠建立后堆叠成员对外体现为一个统一的逻辑实体，用户使用一个 IP 地址对堆叠中的所有交换机进行管理和维护。

## 堆叠管理

堆叠建立后，所有的成员设备组成一台虚拟设备存在于网络中，所有成员设备的资源由主交换机统一管理。用户可以通过任意一台成员设备上的 Console 口或以 Telnet 方式登录堆叠系统，对整个堆叠系统进行管理和配置。

Master 设备作为 iStack 系统的管理中枢，负责响应用户的登录请求，即用户无论使用什么方式，通过哪台成员设备登录 iStack，最终都是对 Master 设备进行配置，Master 设备负责将用户的配置再下发给各个 Slave 设备。这种方式可以使 iStack 内各设备的配置保持高度统一。

堆叠系统使用堆叠 ID 来标识和管理成员设备，堆叠系统中所有成员设备的堆叠 ID 都是唯一的。堆叠 ID 被引入到端口编号中（堆叠 ID/子卡号/端口号），端口编号的第一维表示该端口的堆叠 ID。

## 堆叠维护

堆叠维护的主要功能是监控成员设备的加入和退出，并随时收集新的拓扑，维护现有拓扑。

### 成员设备加入

堆叠维护过程中，继续进行拓扑收集工作，当发现有新的成员设备加入时会根据新加入设备的状态采取不同的处理：

- 新加入的设备本身未形成堆叠（比如，新加入的交换机配置了堆叠功能，之后断电，再使用堆叠线缆连接到已有堆叠系统，上电启动），则新加入的设备会被选为 Slave，堆叠系统中原有主、备角色不变。
- 加入的设备本身已经形成了堆叠（比如，新加入的设备配置了堆叠功能，之后使用堆叠电缆连接到已有堆叠），此时相当于两个堆叠合并。在这种情况下，两个堆叠系统的主交换机将自动选出一个更优的作为新堆叠系统的主交换机，其中一个堆叠系统将保持不变，业务也不会受到影响；而另外一个堆叠系统的所有交换机

将重新启动后加入新堆叠，并将同步主交换机的配置，该堆叠的原有业务也将中断。

### 成员设备退出

堆叠成员退出是指成员交换机从堆叠系统中离开。退出成员的角色不同对堆叠系统的影响有所不同。

- 主交换机退出：备交换机升级为主交换机，更新堆叠拓扑并指定一个新的备交换机。
- 备交换机退出：主交换机更新堆叠拓扑并指定一个新的备交换机。
- 从交换机退出：主交换机更新堆叠拓扑。

### 堆叠分裂

堆叠分裂是指稳态运行的堆叠系统中带电移出部分成员或者堆叠线缆多点故障，导致一个堆叠系统变成多个堆叠系统，如图 1-4 所示。堆叠系统分裂后，可能产生多个有相同配置的堆叠系统，导致网络中 IP 地址和 MAC 地址的冲突，引起网络故障。

图1-4 堆叠分裂示意图



## 成员设备软件自动升级

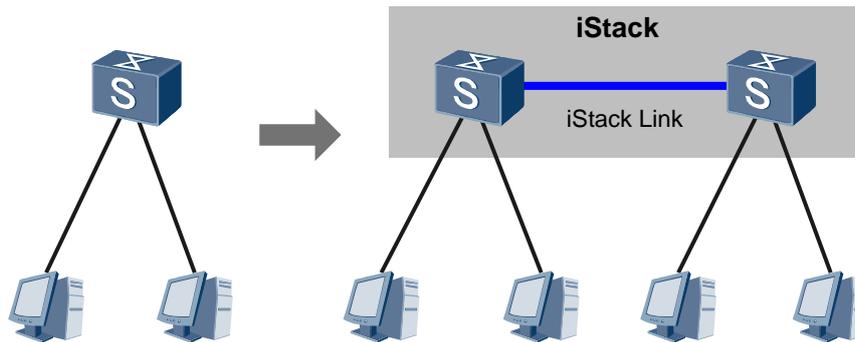
iStack 具有自动加载功能。在进行堆叠扩展增加新成员设备时，并不需要新加入的成员设备与原有虚拟设备具有相同软件版本，只要具有兼容的版本既可。新设备加入堆叠时，会与 Master 设备的软件版本号进行比较，如果不一致，则自动从 Master 设备下载系统启动文件，然后使用新的启动文件重启，重新加入堆叠。如果产品不支持该功能，则需要用户手工配置确保新加入的成员设备与原有虚拟设备版本一致后，新设备才能加入堆叠。

## 1.3 应用场景

### 扩展端口数量

如图 1-5 所示，当接入的用户数增加到原交换机端口密度不能满足接入需求时，可以通过在原有的堆叠系统中增加新的交换机而得到满足。

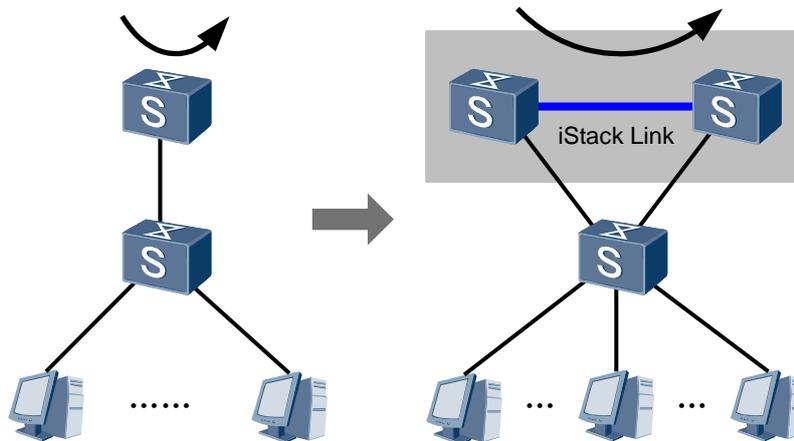
图1-5 扩展端口示意图



### 扩展系统处理能力

使用 iStack 扩展系统处理能力如图 1-6 所示。当交换机转发能力不能满足需求时，可以增加新交换机与原交换机组成堆叠系统来实现。若一台交换机转发能力为 64Mpps，则通过增加一台交换机进行扩展后，整个堆叠设备的转发能力为 128Mpps。需要强调的是，是整个堆叠设备的转发能力整体提高，而不是单个交换机的转发能力提高。

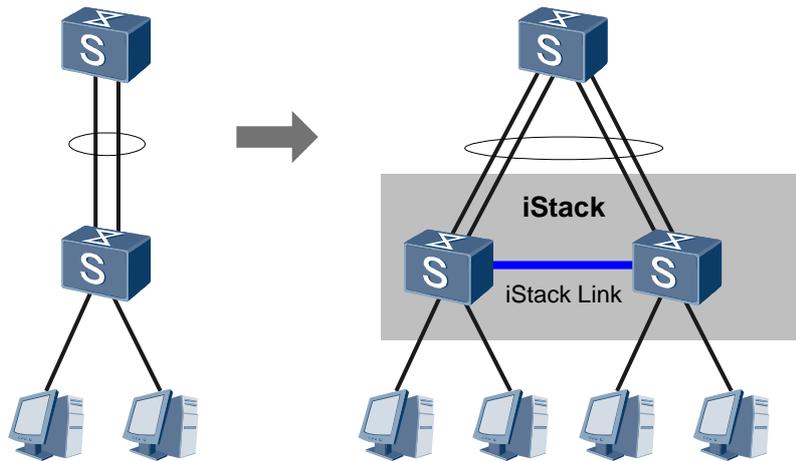
图1-6 扩展系统处理能力组网图



### 扩展带宽

如图 1-7 所示，当交换机上行带宽增加时，可以增加新交换机与原交换机组成堆叠系统，将成员交换机的多条物理链路配置成一个聚合组，提高交换机的上行带宽。

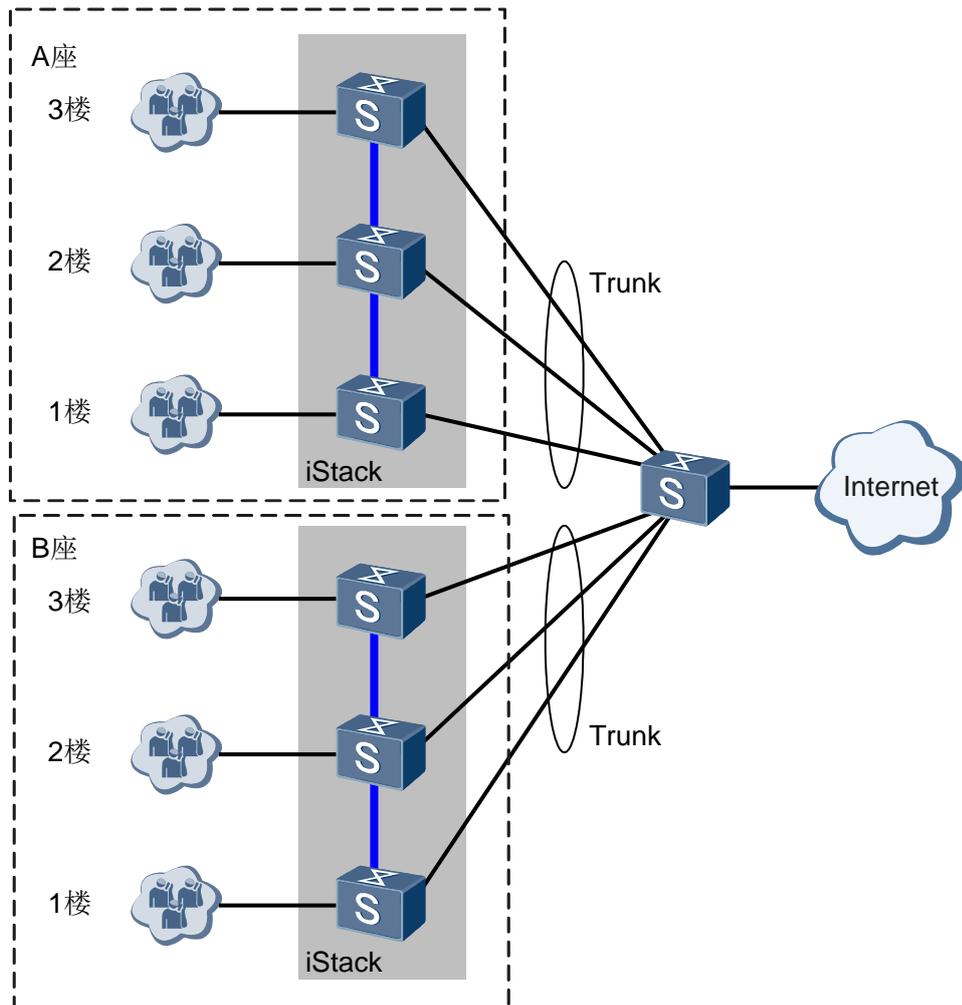
图1-7 扩展带宽示意图



## 跨越空间使用 iStack

iStack 可以将相距遥远的设备连接形成堆叠设备，如图 1-8 所示，每个楼层的用户通过楼道交换机接入外部网络，现将各楼道交换机连接起来形成一个堆叠设备，这样，相当于每个楼只有一个接入设备，网络结构变得更加简单。每个楼层有多条链路到达核心网络，网络变得更加健壮、可靠。对多台楼道交换机的配置简化成对堆叠系统的配置，降低了管理和维护的成本。

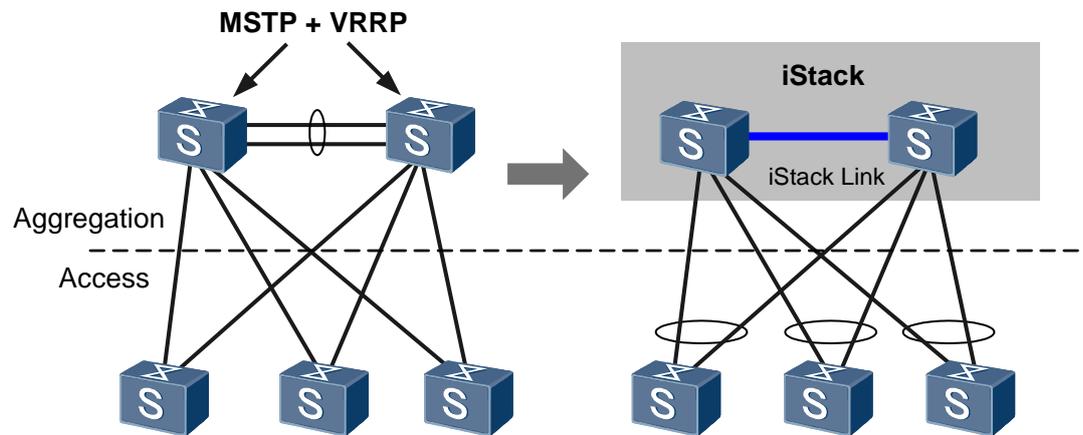
图1-8 跨越空间使用 iStack 组网图



### 简化组网

如图 1-9 所示，网络中汇聚层的多台设备组成堆叠，虚拟成单一的逻辑设备。简化后的组网不再需要使用 MSTP、VRRP 等协议，简化了网络配置，同时依靠跨设备的链路聚合，实现快速收敛，提高了可靠性。

图1-9 简化组网示意图



## 1.4 故障处理案例

### 1.4.1 设备无法组建堆叠

#### 1.4.1 设备无法组建堆叠

##### 故障现象

多台交换机设备无法组建堆叠。

##### 操作步骤

1. 检查设备是否使能堆叠。

登录设备执行 **display stack** 命令，查看堆叠状态。

- 如果显示以下信息，表明设备没有使能堆叠功能。

```
<Quidway> display stack
Error: The stack function is not enabled.
```

在系统视图下执行 **stack enable** 命令，使能堆叠功能。执行此命令后，需要重启设备。

- 如果显示以下信息，表明设备已经使能堆叠功能，请执行下一步骤。

```
<Quidway> display stack
Stack topology type: Link
Stack system MAC: 0200-0001-0000
MAC switch delay time: never
Stack reserved vlanid : 4093
Slot#    role      Mac address      Priority  Device type
-----
1       Master   0200-0001-0000  100     S5700-28P-LI-AC
```

2. 检查设备的具体形态，不同系列之间的设备不能堆叠。

执行命令 **display elabel**，根据电子标签的 **Description** 字段，查看设备的具体形态。

```
<Quidway> display elabel
[Board Properties]
BoardType=CX22EMGEB
BarCode=21023518320123456789
Item=02351832
Description=Quidway S5700-28P-LI-AC,LS5ZC48CM, S5700-28P-LI-AC Mainframe
Manufactured=2009-02-05
VendorName=Huawei
IssueNumber=
CLEICode=
BOM=
```

如果堆叠系统中各交换机不是同种形态设备，则更换成同一形态设备；如果设备都属于同一形态，请执行下一步骤。

3. 检查设备的电子标签和堆叠卡的电子标签是否加载正确。
4. 检查堆叠卡或堆叠线缆是否正常。

执行命令 **display stack port all**，查看堆叠口的状态是否正常。

- 如果显示信息如下，表示堆叠口状态为 Up，堆叠卡与堆叠线缆正常。

```
<Quidway> display stack port all
Show stack port info:
Slot 0:
STACK 1,      status: UP,      peer: 1
STACK 2,      status: UP,      peer: 1
Slot 1:
STACK 1,      status: UP,      peer: 0
STACK 2,      status: UP,      peer: 0
```

- 如果显示信息如下，表示堆叠口状态为 Down，堆叠卡或堆叠线缆异常

```
<Quidway> display stack port all
Show stack port info:
Slot 0:
STACK 1,      status: DOWN,     peer: NONE
STACK 2,      status: DOWN,     peer: NONE
```

查看与之相连的对端设备是否下电或正在重启。如果是，等设备重启后再查看堆叠口状态是否 Up；如果不是，请更换本端的堆叠卡和堆叠线缆。

## 1.5 FAQ

### 能否从备、从设备的 Console 口及管理口登录设备？

可以从任意一台成员交换机的 Console 口或管理口登录堆叠系统，对堆叠系统进行管理。

## 不登录堆叠系统，如何从外观上识别哪个是主交换机？

设备上有一个模式切换按钮，用来设定系统的模式。在堆叠情况下，按这个模式按钮到 stack 模式，处于 master 地位的设备的接口指示灯会处于闪烁状态。

# A 术语与缩略语

术语与缩略语	英文全名	中文解释
iStack	Intelligent Stack	智能弹性架构，又称为堆叠
Eth-Trunk		又称链路聚合（Link Aggregation），是将一组物理接口捆绑在一起作为一个逻辑接口来增加带宽的一种方法