

链路聚合

HUAWEI TECHNOLOGIES CO., LTD.





## 前言

随着网络规模不断扩大，用户对骨干链路的带宽和可靠性提出了越来越高的要求。在传统技术中，常用更换高速率的接口板或更换支持高速率接口板的设备的方式来增加带宽，但这种方案需要付出高额的费用，而且不够灵活。

采用链路聚合技术可以在不进行硬件升级的条件下，通过将多个物理接口捆绑为一个逻辑接口，来达到增加链路带宽的目的。在实现增大带宽目的的同时，链路聚合采用备份链路的机制，可以有效的提高设备之间链路的可靠性。

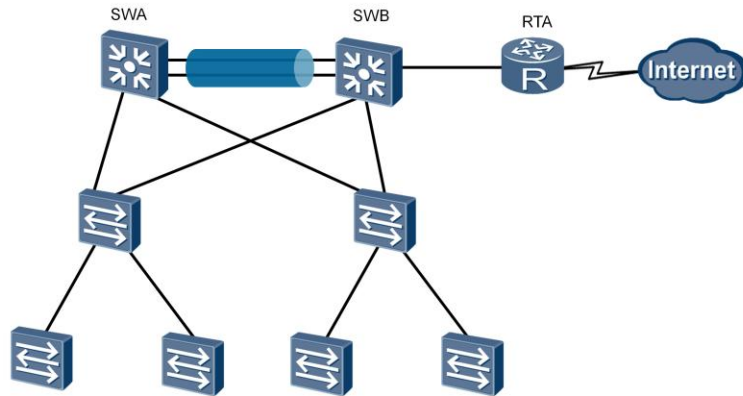


## 学习目标

学完本课程后，您应该能：

- 掌握链路聚合的原理
- 掌握链路聚合的配置

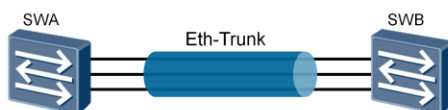
## 链路聚合的应用场景



- 链路聚合一般部署在核心结点，以便提升整个网络的数据吞吐量。

在企业网络中，所有设备的流量在转发到其他网络前都会汇聚到核心层，再由核心层设备转发到其他网络，或者转发到外网。因此，在核心层设备负责数据的高速交换时，容易发生拥塞。在核心层部署链路聚合，可以提升整个网络的数据吞吐量，解决拥塞问题。本示例中，两台核心交换机SWA和SWB之间通过两条成员链路互相连接，通过部署链路聚合，可以确保SWA和SWB之间的链路不会产生拥塞。

## 链路聚合



- 链路聚合能够提高链路带宽，增强网络可用性，支持负载分担。

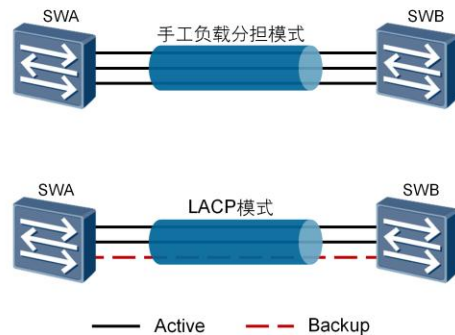
链路聚合是把两台设备之间的多条物理链路聚合在一起，当做一条逻辑链路来使用。这两台设备可以是一对路由器，一对交换机，或者是一台路由器和一台交换机。一条聚合链路可以包含多条成员链路，在ARG3系列路由器和X7系列交换机上默认最多为8条。

链路聚合能够提高链路带宽。理论上，通过聚合几条链路，一个聚合口的带宽可以扩展为所有成员口带宽的总和，这样就有效地增加了逻辑链路的带宽。

链路聚合为网络提供了高可靠性。配置了链路聚合之后，如果一个成员接口发生故障，该成员口的物理链路会把流量切换到另一条成员链路上。

链路聚合还可以在一个聚合口上实现负载均衡，一个聚合口可以把流量分散到多个不同的成员口上，通过成员链路把流量发送到同一个目的地，将网络产生拥塞的可能性降到最低。

## 链路聚合模式



- 手工负载分担模式下所有活动接口都参与数据的转发，分担负载流量。
- LACP模式支持链路备份。

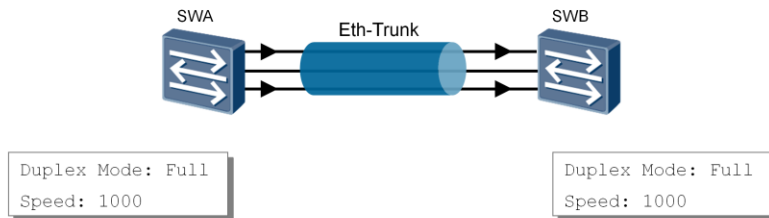
链路聚合包含两种模式：手动负载均衡模式和静态LACP（Link Aggregation Control Protocol）模式。

手工负载分担模式下，Eth-Trunk的建立、成员接口的加入由手工配置，没有链路聚合控制协议的参与。该模式下所有活动链路都参与数据的转发，平均分担流量，因此称为负载分担模式。如果某条活动链路故障，链路聚合组自动在剩余的活动中链路中平均分担流量。当需要在两个直连设备间提供一个较大的链路带宽而设备又不支持LACP协议时，可以使用手工负载分担模式。ARG3系列路由器和X7系列交换机可以基于目的MAC地址，源MAC地址，或者基于源MAC地址和目的MAC地址，源IP地址，目的IP地址，或者基于源IP地址和目的IP地址进行负载均衡。

在静态LACP模式中，链路两端的设备相互发送LACP报文，协商聚合参数。协商完成后，两台设备确定活动接口和非活动接口。在静态LACP模式中，需要手动创建一个Eth-Trunk口，并添加成员口。LACP协商选举活动接口和非活动接口。静态LACP模式也叫M:N模式。M代表活动成员链路，用于在负载均衡模式中转发数据。N代表非活动链路，用于冗余备份。如果一条活动链路发生故障，该链路传输的数据被切换到一条优先级最高的备份链路上，这条备份链路转变为活动状态。

两种链路聚合模式的主要区别是：在静态LACP模式中，一些链路充当备份链路。在手动负载均衡模式中，所有的成员口都处于转发状态。

## 数据流控制



- Eth-Trunk链路两端相连的物理接口的数量、速率、双工方式、流控方式必须一致。

在一个聚合口中，聚合链路两端的物理口（即成员口）的所有参数必须一致，包括物理口的数量，传输速率，双工模式和流量控制模式。成员口可以是二层接口或三层接口。

数据流在聚合链路上传输，数据顺序必须保持不变。一个数据流可以看做是一组MAC地址和IP地址相同的帧。例如，两台设备间的SSH或SFTP连接可以看做一个数据流。如果未配置链路聚合，只是用一条物理链路来传输数据，那么一个数据流中的帧总是能按正确的顺序到达目的地。配置了链路聚合后，多条物理链路被绑定成一条聚合链路，一个数据流中的帧通过不同的物理链路传输。如果第一个帧通过一条物理链路传输，第二个帧通过另外一条物理链路传输，这样一来同一数据流的第二个数据帧就有可能比第一个数据帧先到达对端设备，从而产生接收数据包乱序的情况。

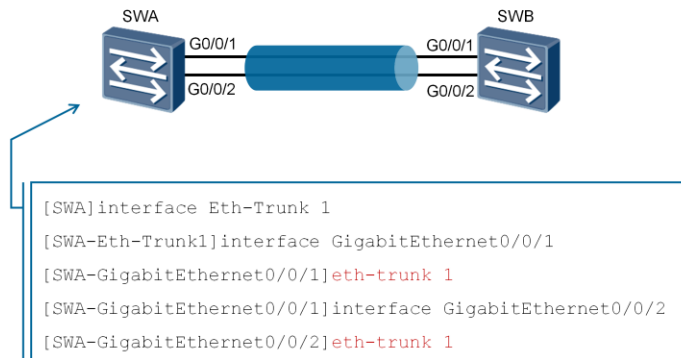
为了避免这种情况的发生，Eth-Trunk采用逐流负载分担的机制，这种机制把数据帧中的地址通过HASH算法生成HASH-KEY值，然后根据这个数值在Eth-Trunk转发表中寻找对应的出接口，不同的MAC或IP地址HASH得出的HASH-KEY值不同，从而出接口也就不同，这样既保证了同一数据流的帧在同一条物理链路转发，又实现了流量在聚合组内各物理链路上的负载分担，即逐流的负载分担。逐流负载分担能保证包的顺序，但不能保证带宽利用率。

负载分担的类型主要包括以下几种，用户可以根据具体应用选择不同的负载分担类型。

1. 根据报文的源MAC地址进行负载分担；
2. 根据报文的目的MAC地址进行负载分担；
3. 根据报文的源IP地址进行负载分担；
4. 根据报文的目的IP地址进行负载分担；
5. 根据报文的源MAC地址和目的MAC地址进行负载分担；
6. 根据报文的源IP地址和目的IP地址进行负载分担；
7. 根据报文的VLAN、源物理端口等对L2、IPv4、IPv6和MPLS报文进行增强型负载分担。



## 二层链路聚合配置



本例中，通过执行**interface Eth-trunk <trunk-id>**命令配置链路聚合。这条命令创建了一个Eth-Trunk口，并且进入该Eth-Trunk口视图。*trunk-id*用来唯一标识一个Eth-Trunk口，该参数的取值可以是0到63之间的任何一个整数。如果指定的Eth-Trunk口已经存在，执行**interface eth-trunk**命令会直接进入该Eth-Trunk口视图。

配置Eth-Trunk口和成员口，需要注意以下规则：

1. 只能删除不包含任何成员口的Eth-Trunk口。
2. 把接口加入Eth-Trunk口时，二层Eth-Trunk口的成员口必须是二层接口，三层Eth-Trunk口的成员口必须是三层接口。
3. 一个Eth-Trunk口最多可以加入8个成员口。
4. 加入Eth-Trunk口的接口必须是hybrid接口（默认的接口类型）。
5. 一个Eth-Trunk口不能充当其他Eth-Trunk口的成员口。
6. 一个以太网接口只能加入一个Eth-Trunk口。如果把一个以太网接口加入另一个Eth-Trunk口，必须先把该以太网接口从当前所属的Eth-Trunk口中删除。
7. 一个Eth-Trunk口的成员口类型必须相同。例如，一个快速以太网口（FE口）和一个千兆以太网口（GE口）不能加入同一个Eth-Trunk。
8. 位于不同接口板（LPU）上的以太网接口可以加入同一个Eth-Trunk口。如果一个对端接口直接和本端Eth-Trunk口的一个成员口相连，该对端接口也必须加入一个Eth-Trunk口。否则两端无法通信。

9. 如果成员口的速率不同，速率较低的接口可能会拥塞，报文可能会被丢弃。
10. 接口加入Eth-Trunk口后，Eth-Trunk口学习MAC地址，成员口不再学习。

## 查看链路聚合信息

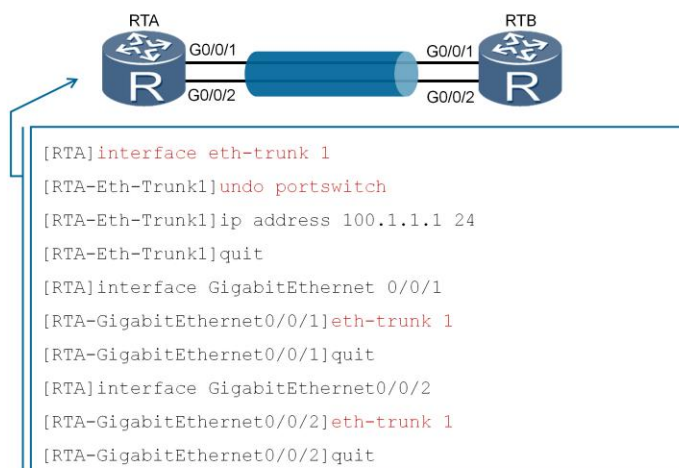
```
[SWA]display interface eth-trunk 1
Eth-Trunk1 current state : UP
Line protocol current state : UP
.....
-----
PortName                Status    Weight
-----
GigabitEthernet0/0/1    UP        1
GigabitEthernet0/0/2    UP        1
-----

The Number of Ports in Trunk : 2
The Number of UP Ports in Trunk : 2
```

- 两个成员接口已经被绑定到Eth-trunk 1。

执行**display interface eth-trunk <trunk-id>**命令，可以确认两台设备间是否已经成功实现链路聚合。也可以使用这条命令收集流量统计数据，定位接口故障。如果Eth-Trunk口处于UP状态，表明接口正常运行。如果接口处于Down状态，表明所有成员口物理层发生故障。如果管理员手动关闭端口，接口处于Administratively DOWN状态。可以通过接口状态的改变发现接口故障，所有接口正常情况下都应处于Up状态。

## 三层链路聚合配置



如果要在路由器上配置三层链路聚合，需要首先创建Eth-Trunk接口，然后在Eth-Trunk逻辑口上执行**undo portswitch**命令，把聚合链路从二层转为三层链路。执行**undo portswitch**命令后，可以为Eth-Trunk逻辑口分配一个IP地址。

## 查看链路聚合信息

```
[RTA]display interface eth-trunk 1
Eth-Trunk1 current state : UP
Line protocol current state : UP
.....
-----
PortName                Status      Weight
-----
GigabitEthernet0/0/1    UP          1
GigabitEthernet0/0/2    UP          1
-----
The Number of Ports in Trunk : 2
The Number of UP Ports in Trunk : 2
```

- 两个成员接口已经被绑定到Eth-trunk 1。

执行**display interface eth-trunk <trunk-id>**命令，可以确认两台设备间是否已经成功实现链路聚合。也可以使用这条命令收集流量统计数据，定位接口故障。如果Eth-Trunk口处于UP状态，表明接口正常运行。如果接口处于Down状态，表明所有成员口物理层发生故障。如果管理员手动关闭端口，接口处于Administratively DOWN状态。可以通过接口状态的改变发现接口故障，所有接口正常情况下都应处于Up状态。



## 总结

- 如果一个管理员希望将千兆以太网口和百兆以太网口加入同一个Eth-trunk, 会发生什么?
- 哪种链路聚合方法可以使用链路备份?

1. 一个快速以太网口（FE口）和一个千兆以太网口（GE口）不能加入同一个Eth-Trunk。如果将两个不同类型的接口加入到同一个Eth-Trunk口, 设备会提示发生错误。
2. 只有LACP模式支持备份成员链路。如需建立备份链路, 应使用LACP模式的链路聚合。



谢谢

[www.huawei.com](http://www.huawei.com)