

目 录

以太网链路聚合	1
以太网链路聚合简介	1
基本概念	1
静态聚合模式	4
动态聚合模式	5
聚合负载分担类型	7

以太网链路聚合

以太网链路聚合简介

以太网链路聚合简称链路聚合，它通过将多条以太网物理链路捆绑在一起成为一条逻辑链路，从而实现增加链路带宽的目的。同时，这些捆绑在一起的链路通过相互间的动态备份，可以有效地提高链路的可靠性。

如 [图 1](#) 所示，Device A 与 Device B 之间通过三条以太网物理链路相连，将这三条链路捆绑在一起，就成为了一条逻辑链路 Link aggregation 1，这条逻辑链路的带宽等于原先三条以太网物理链路的带宽总和，从而达到了增加链路带宽的目的；同时，这三条以太网物理链路相互备份，有效地提高了链路的可靠性。

图 1 链路聚合示意图



基本概念

1. 聚合组、成员端口和聚合接口

将多个以太网接口捆绑在一起所形成的组合称为聚合组，而这些被捆绑在一起的以太网接口就称为该聚合组的成员端口。每个聚合组唯一对应着一个逻辑接口，我们称之为聚合接口。聚合组/聚合接口可以分为以下两种类型：

- 二层聚合组/二层聚合接口：二层聚合组的成员端口全部为二层以太网接口，其对应的聚合接口称为二层聚合接口（Bridge-aggregation Interface, BAGG）。
- 三层聚合组/三层聚合接口：三层聚合组的成员端口全部为三层以太网接口，其对应的聚合接口称为三层聚合接口（Route-aggregation Interface, RAGG）。



说明

- 不同型号的设备支持的聚合组/聚合接口的类型不同，请以设备的实际情况为准。
- 聚合组与聚合接口的编号是一一对应的，譬如聚合组 1 对应于聚合接口 1。
- 在创建了三层聚合接口之后，还可以继续创建该三层聚合接口的子接口（简称三层聚合子接口）。三层聚合子接口也是一种逻辑接口，工作在网络层，主要用来在三层聚合接口上支持收发携带 VLAN Tag 的报文。不同型号的设备对三层聚合子接口的支持情况不同，请以设备的实际情况为准。
- 聚合接口的速率和双工模式取决于对应聚合组内的选中端口（请参见“[2. 成员端口的状态](#)”）：聚合接口的速率等于所有选中端口的速率之和，聚合接口的双工模式则与选中端口的双工模式相同。

2. 成员端口的状态

聚合组内的成员端口具有以下两种状态：

- 选中（**Selected**）状态：此状态下的成员端口可以参与用户数据的转发，处于此状态的成员端口简称为“选中端口”。
- 非选中（**Unselected**）状态：此状态下的成员端口不能参与用户数据的转发，处于此状态的成员端口简称为“非选中端口”。

3. 操作 Key

操作 **Key** 是系统在进行链路聚合时用来表征成员端口聚合能力的一个数值，它是根据成员端口上的一些信息（包括该端口的速率、双工模式等）的组合自动计算生成的，这个信息组合中任何一项的变化都会引起操作 **Key** 的重新计算。在同一聚合组中，所有的选中端口都必须具有相同的操作 **Key**。

4. 配置分类

根据对成员端口状态的影响不同，我们可以将成员端口上的配置分为以下三类：

- (1) 端口属性类配置：包含速率、双工模式和链路状态（up/down）这三项配置内容，是成员端口上最基础的配置内容。
- (2) 第二类配置：包含的配置内容如 [表 1](#) 所示。在聚合组中，只有与对应聚合接口的第二类配置完全相同的成员端口才能够成为选中端口。

表 1 第二类配置的内容

配置项	内容
端口隔离	端口是否加入隔离组、端口所属的端口隔离组
QinQ 配置	端口的 QinQ 功能开启/关闭状态、VLAN Tag 的 TPID 值、添加的外层 VLAN Tag、内外层 VLAN 优先级映射关系、不同内层 VLAN ID 添加外层 VLAN Tag 的策略、内层 VLAN ID 替换关系
VLAN 配置	端口上允许通过的 VLAN、端口缺省 VLAN ID、端口的链路类型（即 Trunk、Hybrid、Access 类型）、基于 IP 子网的 VLAN 配置、基于协议的 VLAN 配置、VLAN 报文是否带 Tag 配置
MAC 地址学习配置	是否具有 MAC 地址学习功能、端口是否具有最大学习 MAC 地址个数的限制、MAC 地址表满后是否继续转发



说明

- 在聚合接口上所作的第二类配置，将被自动同步到对应聚合组内的所有成员端口上。当聚合接口被删除后，这些配置仍将保留在这些成员端口上。
- 由于成员端口上第二类配置的改变可能导致其选中/非选中状态发生变化，进而对业务产生影响，因此当在成员端口上进行此类配置时，系统将给出提示信息，由用户来决定是否继续执行该配置。

- (3) 第一类配置：是相对于第二类配置而言的，包含的配置内容有 GVRP、MSTP 等。在聚合组中，即使某成员端口与对应聚合接口的第一类配置存在不同，也不会影响该成员端口成为选中端口。



说明

在成员端口上所作出的第一类配置，只有当该成员端口退出聚合组后才能生效。

5. 参考端口

参考端口从成员端口中选出，其端口属性类配置和第二类配置将作为同一聚合组内的其它成员端口的参照，以确定这些成员端口的状态。

6. LACP 协议

基于 IEEE802.3ad 标准的 LACP（Link Aggregation Control Protocol，链路聚合控制协议）协议是一种实现链路动态聚合的协议，运行该协议的设备之间通过互发 LACPDU（Link Aggregation Control Protocol Data Unit，链路聚合控制协议数据单元）来交互链路聚合的相关信息。

(1) LACP 协议的功能

根据所使用的LACPDU字段的的不同，可将LACP协议的功能分为基本功能和扩展功能两大类，如 [表 2](#) 所示。

表 2 LACP 协议的功能分类

类别	说明
基本功能	<p>利用 LACPDU 的基本字段可以实现 LACP 协议的基本功能，基本字段包含以下信息：系统 LACP 优先级、系统 MAC 地址、端口聚合优先级、端口编号和操作 Key。</p> <p>动态聚合组内的成员端口会自动启用 LACP 协议，并通过发送 LACPDU 向对端通告本端的上述信息。当对端收到该 LACPDU 后，将其中的信息与本端其它成员端口收到的信息进行比较，以选择能够处于选中状态的成员端口，使双方可以对各自接口的选中/非选中状态达成一致，从而决定哪些链路可以加入聚合组以及某链路何时可以加入聚合组。</p>
扩展功能	<p>通过对 LACPDU 的字段进行扩展，可以实现对 LACP 协议的扩展。譬如，通过在扩展字段中定义一个新的 TLV（Type/Length/Value，类型/长度/值）数据域，可以实现 IRF（Intelligent Resilient Framework，智能弹性架构）中的 LACP MAD（Multi-Active Detection，多 Active 检测）机制。对于支持 LACP 协议扩展功能的设备来说：</p> <ul style="list-style-type: none"> ● 如果同时支持 IRF，则该设备可以作为成员设备或中间设备来参与 LACP MAD ● 如果不支持 IRF，则该设备只能作为中间设备来参与 LACP MAD

(2) LACP 优先级

根据作用的不同，可以将LACP优先级分为系统LACP优先级和端口聚合优先级两类，如 [表 3](#)所示。

表 3 LACP 优先级的分类

类别	说明	比较标准
系统 LACP 优先级	系统 LACP 优先级用于区分两端设备优先级的高低。要想使两端设备的选中端口一致，可以使一端具有较高的优先级，另一端则根据优先级较高的一端来选择本端的选中端口	优先级数值越小，优先级越高
端口聚合优先级	端口聚合优先级用于区分各成员端口成为选中端口的优先程度	

(3) LACP 超时时间

LACP 超时时间是指成员端口等待接收 LACPDU 的超时时间。在三倍 LACP 超时时间之后，如果本端成员端口仍未收到来自对端的 LACPDU，则认为对端成员端口已失效。LACP 超时时间只有短超时（1 秒）和长超时（30 秒）两种取值。

7. 聚合模式



说明

不同型号的设备支持的聚合模式不同，请以设备的实际情况为准。

根据成员端口上是否启用了 LACP 协议，可以将链路聚合分为静态聚合和动态聚合两种模式，它们各自的特点如表 4 所示。

表 4 不同聚合模式的特点

聚合模式	成员端口是否开启 LACP 协议	优点	缺点
静态聚合模式	否	一旦配置好后，端口的选中/非选中状态就不会受网络环境的影响，比较稳定	不能根据对端的状态调整端口的选中/非选中状态，不够灵活
动态聚合模式	是	能够根据对端和本端的信息调整端口的选中/非选中状态，比较灵活	端口的选中/非选中状态容易受网络环境的影响，不够稳定

处于静态聚合模式和动态聚合模式下的聚合组分别称为静态聚合组和动态聚合组，动态聚合组内的选中端口以及处于 up 状态、与对应聚合接口的第二类配置相同的非选中端口均可以收发 LACPDU。

静态聚合模式

在静态聚合模式下，聚合组内的成员端口上不启用 LACP 协议，其端口状态通过手工进行维护。静态聚合模式的工作机制如下：

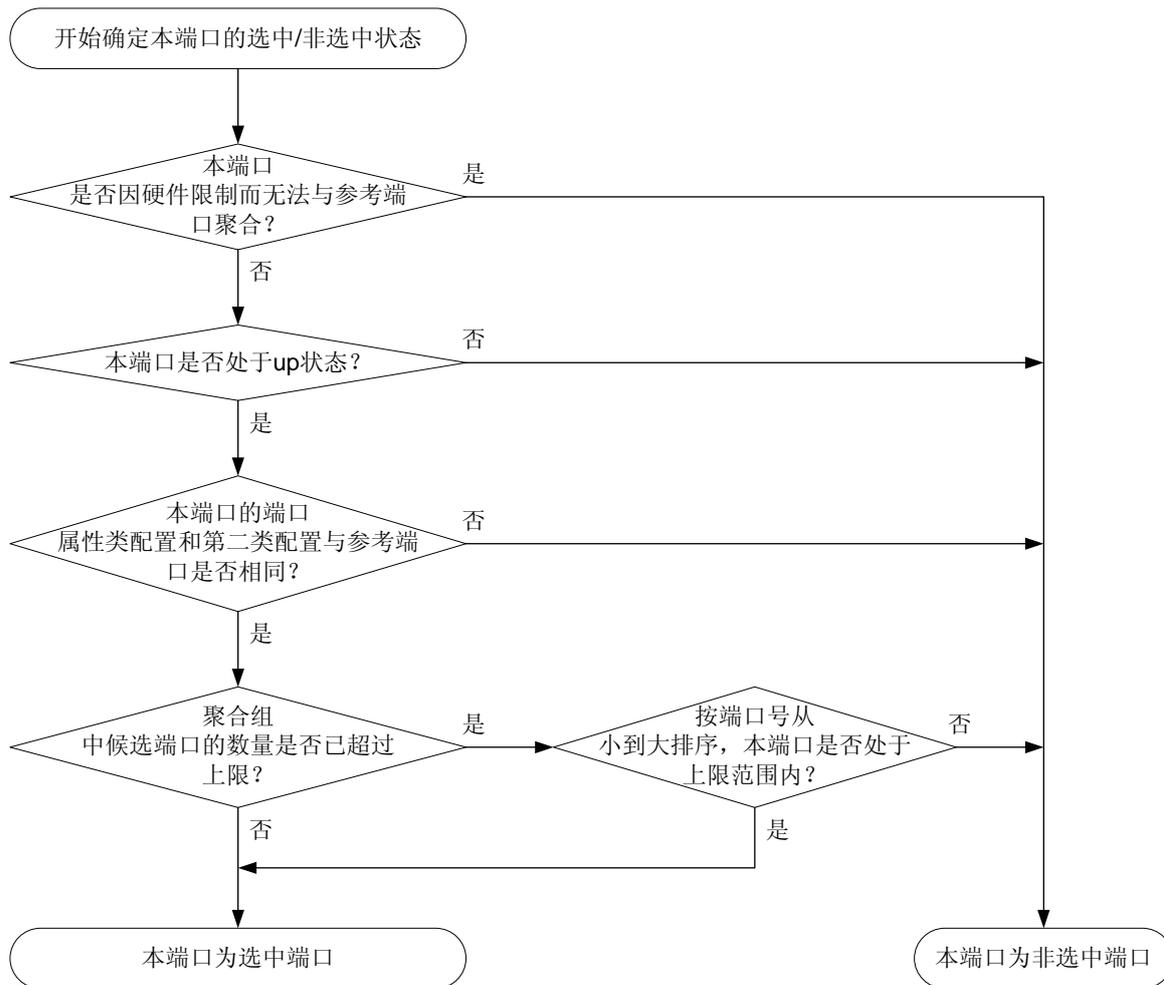
1. 选择参考端口

当聚合组内有处于 up 状态的端口时，先比较端口的聚合优先级，优先级数值最小的端口作为参考端口；如果优先级相同，再按照端口的全双工/高速率->全双工/低速率->半双工/高速率->半双工/低速率的优先次序，选择优先次序最高、且第二类配置与对应聚合接口相同的端口作为该组的参考端口；如果优先次序也相同，则选择端口号最小的端口作为参考端口。

2. 确定成员端口的状态

静态聚合组内成员端口状态的确定流程如图 2 所示。

图 2 静态聚合组内成员端口状态的确定流程



说明

- 静态聚合组内选中端口的最大数量与设备的型号有关，请以设备的实际情况为准。
- 当一个成员端口的端口属性类配置或第二类配置改变时，其所在静态聚合组内各成员端口的选中/非选中状态可能会发生改变。
- 当静态聚合组内选中端口的数量已达到上限时，后加入的成员端口即使满足成为选中端口的所有条件，也不会立刻成为选中端口。这样能够尽量维持当前选中端口上的流量不中断，但是由于设备重启时会重新计算选中端口，因此可能导致设备重启前、后各成员端口的选中/非选中状态不一致。

动态聚合模式

说明

本特性的支持情况与设备的型号有关，请以设备的实际情况为准。

在动态聚合模式下，聚合组内的成员端口上均启用 LACP 协议，其端口状态通过该协议自动进行维护。动态聚合模式的工作机制如下：

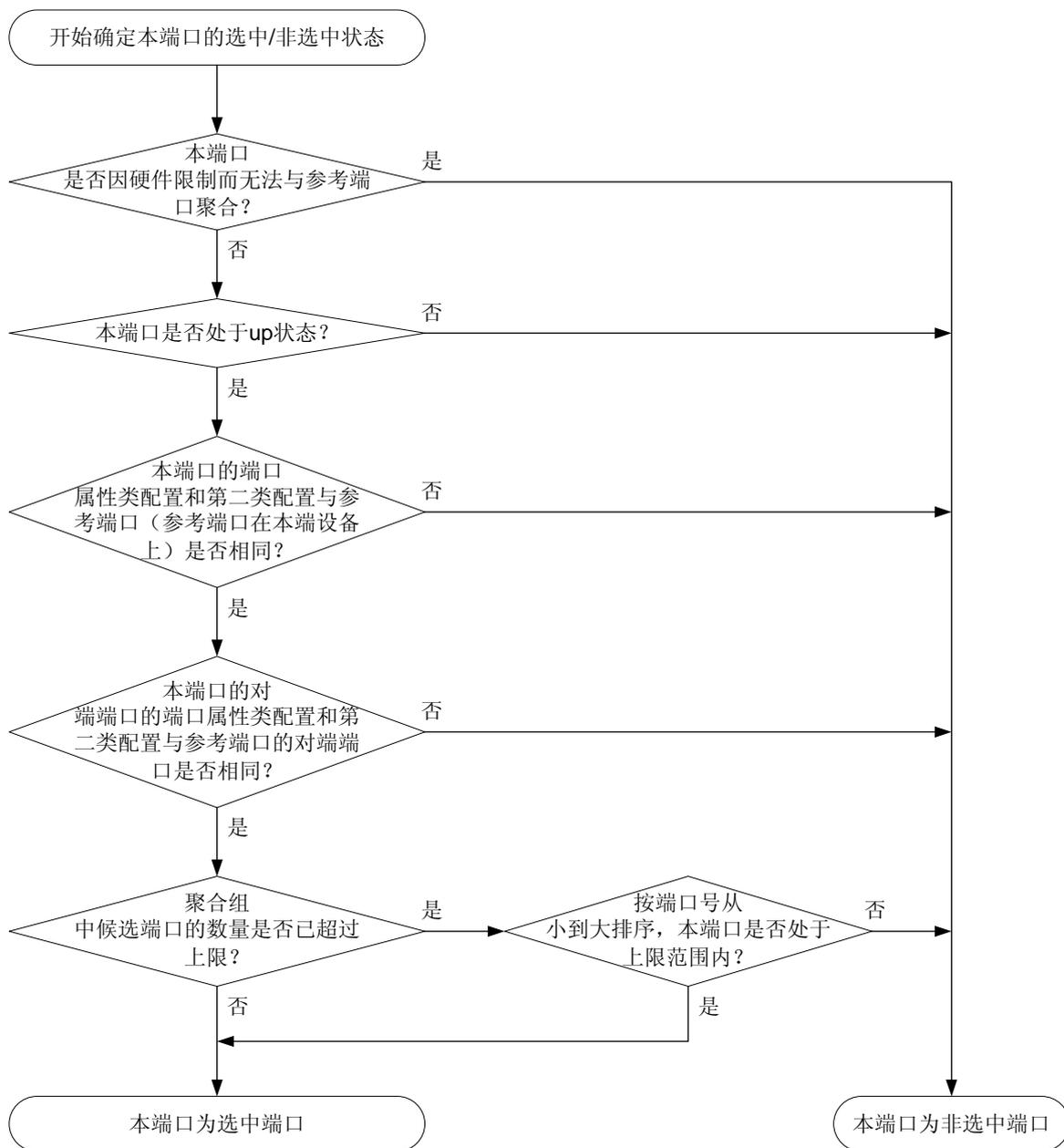
1. 选择参考端口

- (1) 首先，从聚合链路的两端选出设备 ID（由系统的 LACP 优先级和系统的 MAC 地址共同构成）较小的一端：先比较两端的系统 LACP 优先级，优先级数值越小其设备 ID 越小；如果优先级相同再比较其系统 MAC 地址，MAC 地址越小其设备 ID 越小。
- (2) 其次，对于设备 ID 较小的一端，再比较其聚合组内各成员端口的端口 ID（由端口的聚合优先级和端口的编号共同构成）：先比较端口的聚合优先级，优先级数值越小其端口 ID 越小；如果优先级相同再比较其端口号，端口号越小其端口 ID 越小。端口 ID 最小的端口作为参考端口。

2. 确定成员端口的状态

在设备ID较小的一端，动态聚合组内成员端口状态的确定流程如 [图 3](#)所示。

图 3 动态聚合组内成员端口状态的确定流程



与此同时，设备 ID 较大的一端也会随着对端成员端口状态的变化，随时调整本端各成员端口的状态，以确保聚合链路两端成员端口状态的一致。

 说明

- 动态聚合组内选中端口的最大数量与设备的型号有关，请以设备的实际情况为准。
 - 当动态聚合组内同时存在全双工端口和半双工端口时，全双工端口将优先成为选中端口；只有当所有全双工端口都无法成为选中端口，或动态聚合组内只有半双工端口时，才允许从半双工端口中选出一个成为选中端口，且只有一个半双工端口可成为选中端口。
 - 当一个成员端口的端口属性类配置或第二类配置改变时，其所在动态聚合组内各成员端口的选中/非选中状态可能会发生改变。
 - 当本端端口的选中/非选中状态发生改变时，其对端端口的选中/非选中状态也将随之改变。
 - 当动态聚合组内选中端口的数量已达到上限时，后加入的成员端口一旦满足成为选中端口的所有条件，就会立刻取代已不满足条件的端口成为选中端口。
-

聚合负载分担类型

通过采用不同的聚合负载分担类型及其组合，可以灵活地实现对聚合组内流量的负载分担。聚合负载分担的类型包括以下几种：

- 根据报文的 MAC 地址进行聚合负载分担
- 根据报文的 VLAN 标签进行聚合负载分担
- 根据报文的服务器端口号进行聚合负载分担
- 根据报文的入端口进行聚合负载分担
- 根据报文的 IP 地址进行聚合负载分担
- 根据报文的 IP 协议类型进行聚合负载分担
- 根据报文的 MPLS 标签进行聚合负载分担

用户可以指定系统按照上述聚合负载分担类型的其中之一或其组合来进行负载分担，此外用户也可以指定系统按照报文类型（如二层、IPv4、IPv6、MPLS 等）自动选择聚合负载分担的类型，还可以指定系统对每个报文逐包进行聚合负载分担。

 说明

不同型号的设备支持的聚合负载分担类型不同，请以设备的实际情况为准。
