

第 2 篇 路由基础

第 3 章 路由控制与转发

第 4 章 路由协议基础

第 5 章 路由负载分担与备份

第 6 章 路由聚合与 CIDR

第3章 路由控制与转发

为了使路由转发效率更高，系统使用了路由控制平面和转发平面。路由控制平面负责路由计算和维护，而路由转发平面负责 IP 数据报文的转发。本章介绍控制平面和转发平面的关系，并对路由表、FIB 表的生成和作用进行了详解。

3.1 本章目标

课程目标

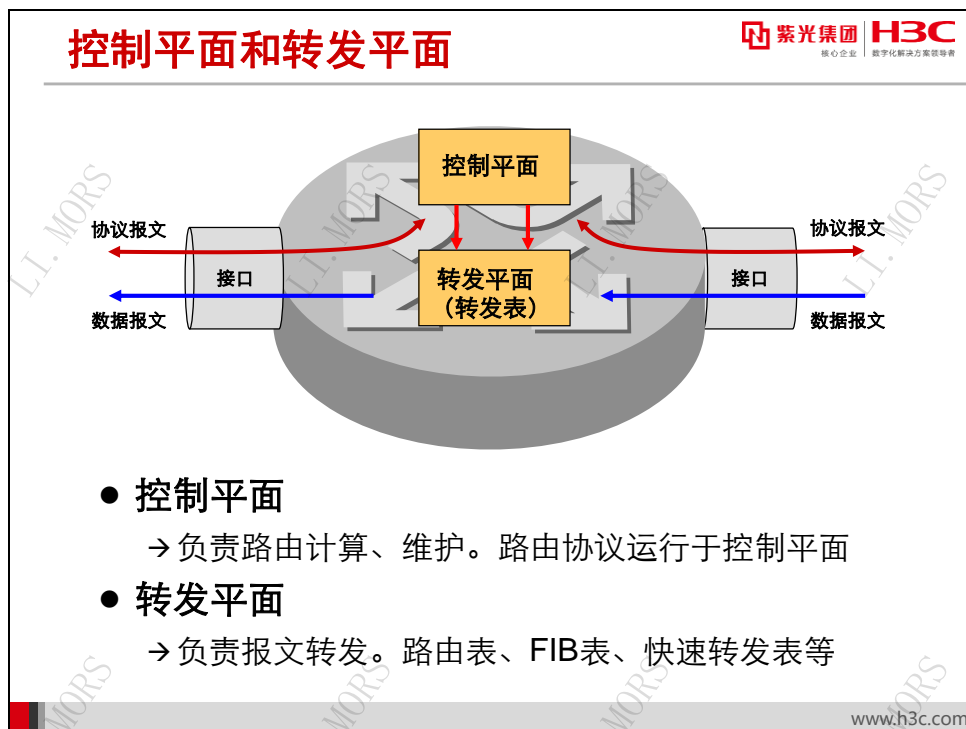
● 学习完本课程，您应该能够：

- 了解控制平面和转发平面的区别
- 了解FIB表项作用与生成过程
- 掌握快速转发工作原理
- 掌握快速转发的配置



www.h3c.com

3.2 路由的控制与转发平面



路由器、交换机承担着路由学习、MAC 地址学习、数据报文转发等重要的工作，其系统的稳定性是非常重要的。系统的设计者们尽力从架构上使系统工作稳定、可靠，其中很重要的一点就是控制平面和转发平面相对独立，以减少互相影响。本课程我们学习控制平面、转发平面的定义、作用，相关表项内容及配置。

● 控制平面

控制平面指系统中用来传送信令、计算表项的部分。诸如协议报文收发、协议表项计算、维护等都属于控制平面的范畴。例如在路由系统中，负责路由协议学习、路由表项维护的进程就属于控制平面；而交换系统中，负责 MAC 地址学习的进程则属于控制平面。

● 转发平面

转发平面指系统中用来进行数据报文的封装、转发的部分。诸如数据报文的接收、解封装、封装、转发等都属于转发平面的范畴。例如系统接收到 IP 报文后，需要进行解封装、查路由表、从出接口转发等，系统中负责以上行为的进程则属于转发平面。

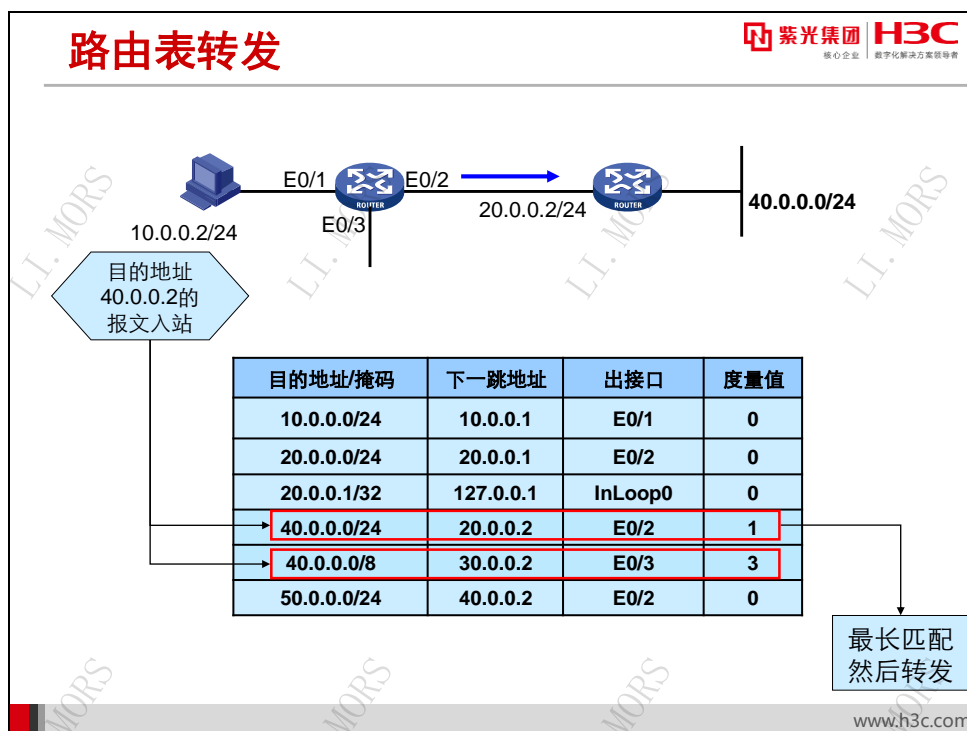
控制平面与转发平面相对独立又协同工作。系统的控制平面进行协议交互、路由计算后，生成若干表项，下发到转发平面，指导转发平面对报文进行转发。例如，路由器通过 OSPF 协议建立了路由表项，再进一步生成 FIB（Forwarding Information Base）表、快速转发表等，指导系统进行 IP 报文转发。

良好的系统设计应该是使控制平面与转发平面尽量分离，互不影响。当系统的控制平面暂时出现故障时，转发平面还可以继续工作，这样可以保证网络中原有的业务不受系统故障的影响，提高整个网络的可靠性。

控制平面与转发平面分离的另外好处是系统设计能够做到模块化，易于维护，便于扩展。例如原有系统仅能够支持 **RIP** 路由协议，伴随着网络设备数量的增长，**RIP** 协议不能够满足需求。系统开发者可以在不改变转发平面的基础上升级控制平面至支持 **OSPF** 协议，减少升级的代价。

控制平面与转发平面可以是物理分离，也可以是逻辑分离。高端设备如核心交换机、核心路由器，一般采用控制平面与转发平面物理分离。其主控板上的 **CPU** 不负责报文转发，而专注于系统的控制；而业务板则专注于数据报文转发。如果主控板损坏，业务板仍然能够转发报文。而对于入门级的网络设备，受限于成本，一般只做到逻辑分离。即设备启动后，系统将 **CPU** 和内存资源划分给不同的进程，有的进程负责路由学习，有的进程负责报文转发。

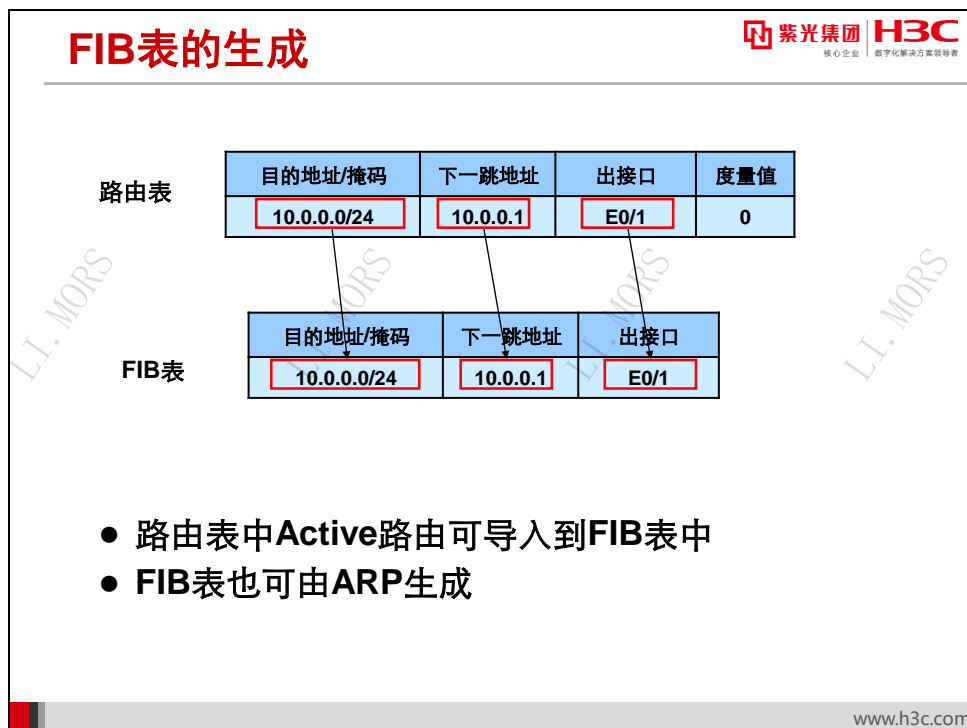
3.3 路由表和FIB表



当路由表中存在多个路由项可以同时匹配目的 IP 地址时，路由查找进程会选择其中掩码最长的路由项用于转发，此为最长匹配原则。

上图中，路由器接收到目的地址为 40.0.0.2 的数据包，经查找整个路由表，发现与路由 40.0.0.0/24 和 40.0.0.0/8 都能匹配。但根据最长匹配的原则，路由器会选择路由项 40.0.0.0/24，根据该路由项转发数据包。

由以上过程可知，路由表中路由项数量越多，所需查找及匹配的次数则越多，其转发效率就越低。



为了做到控制平面与转发平面完全分离，系统构建了另一张 FIB（Forwarding Information Base）表，又称为转发表，专注于数据报文的转发。FIB 表项来源于路由表项。

在计算路由信息的时候，不同路由协议所计算出的路径可能会不同。在这种情况下，路由器会选择较高路由优先级的路由协议发现的路由作为最优路由，并置为 **Active**（活跃）状态；而其它路由作为备份路由，置为 **Inactive**（非活跃）状态。此时，**Active**（活跃）状态的路由表项会由系统导入到 FIB 表中，作为系统转发的依据。另外，在某些系统中，FIB 表项也可能来源于 ARP 解析；即系统将通过 ARP 解析而得到的本地网段内的主机路由也添加到 FIB 表中。

FIB 表与路由表是同步更新的。系统的控制平面发现新的路由信息，根据路由信息更新自己的路由表，生成新的 **Active** 状态的路由表项，然后更新 FIB 表；如果原路由表中处于 **Active** 状态的路由表项失效，系统也会删除相关 FIB 表项。

由于 FIB 表中没有处于 **Inactive** 状态的冗余路由，所以通常 FIB 表项数量小于路由表项。所以，系统可以设计将 FIB 表项加载到硬件中，以大大加快数据转发速度。如某些高端交换机在启动后，FIB 表被系统加载到接口业务板的硬件中，数据报文通过硬件转发，不再需要通过 CPU 转发，可以做到没有转发时延。

FIB表摘要信息

[H3C]display fib

Destination count: 8 FIB entry count: 8

Flag:

U:Useable G:Gateway H:Host B:Blackhole D:Dynamic S:Static
R:Relay F:FRR

Destination/Mask	Nexthop	Flag	OutInterface/Token	Label
0.0.0.0/32	127.0.0.1	UH	InLoop0	Null
127.0.0.0/8	127.0.0.1	U	InLoop0	Null
127.0.0.0/32	127.0.0.1	UH	InLoop0	Null
127.0.0.1/32	127.0.0.1	UH	InLoop0	Null
127.255.255.255/32	127.0.0.1	UH	InLoop0	Null
224.0.0.0/4	0.0.0.0	UB	NULL0	Null
224.0.0.0/24	0.0.0.0	UB	NULL0	Null
255.255.255.255/32	127.0.0.1	UH	InLoop0	Null

路由的标志

转发接口

www.h3c.com

设备上可以使用 **display fib** 命令用来显示所有的 FIB 转发信息。如下所示：

[H3C]display fib

Destination count: 8 FIB entry count: 8

Flag:

U:Useable G:Gateway H:Host B:Blackhole D:Dynamic S:Static
R:Relay F:FRR

Destination/Mask	Nexthop	Flag	OutInterface/Token	Label
0.0.0.0/32	127.0.0.1	UH	InLoop0	Null
127.0.0.0/8	127.0.0.1	U	InLoop0	Null
127.0.0.0/32	127.0.0.1	UH	InLoop0	Null
127.0.0.1/32	127.0.0.1	UH	InLoop0	Null
127.255.255.255/32	127.0.0.1	UH	InLoop0	Null
224.0.0.0/4	0.0.0.0	UB	NULL0	Null
224.0.0.0/24	0.0.0.0	UB	NULL0	Null
255.255.255.255/32	127.0.0.1	UH	InLoop0	Null

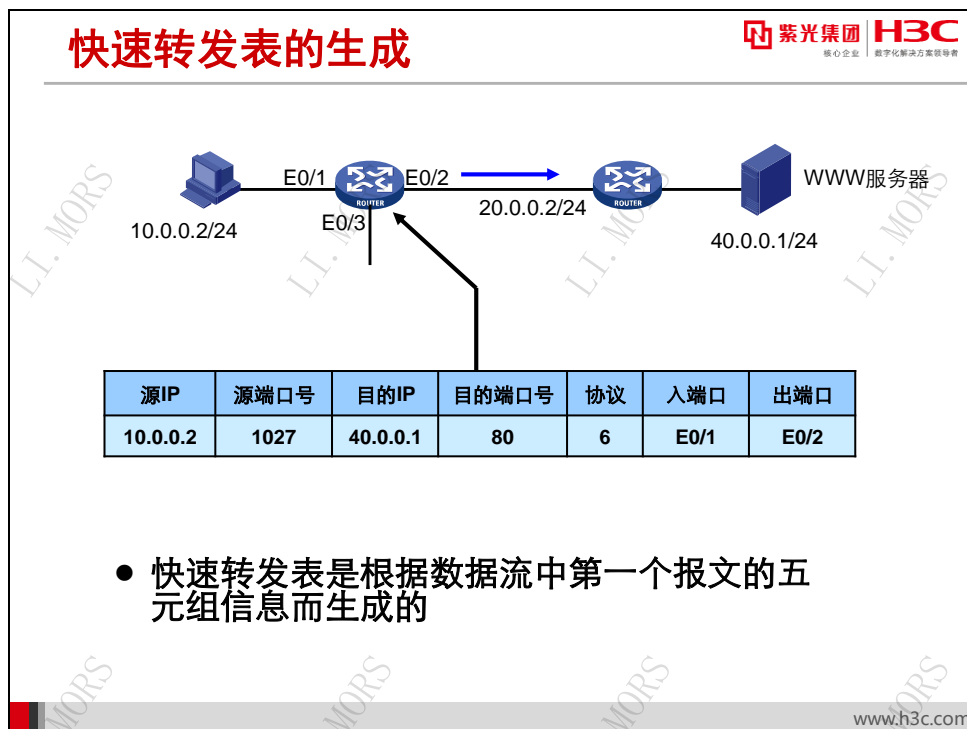
其中各参数含义如下表所示：

表3-1 display fib 命令显示信息描述表

字段	描述
Destination count	目的地址的个数
FIB entry count	FIB表项数目
Destination/Mask	目的地址/掩码长度
Nexthop	转发的下一跳地址

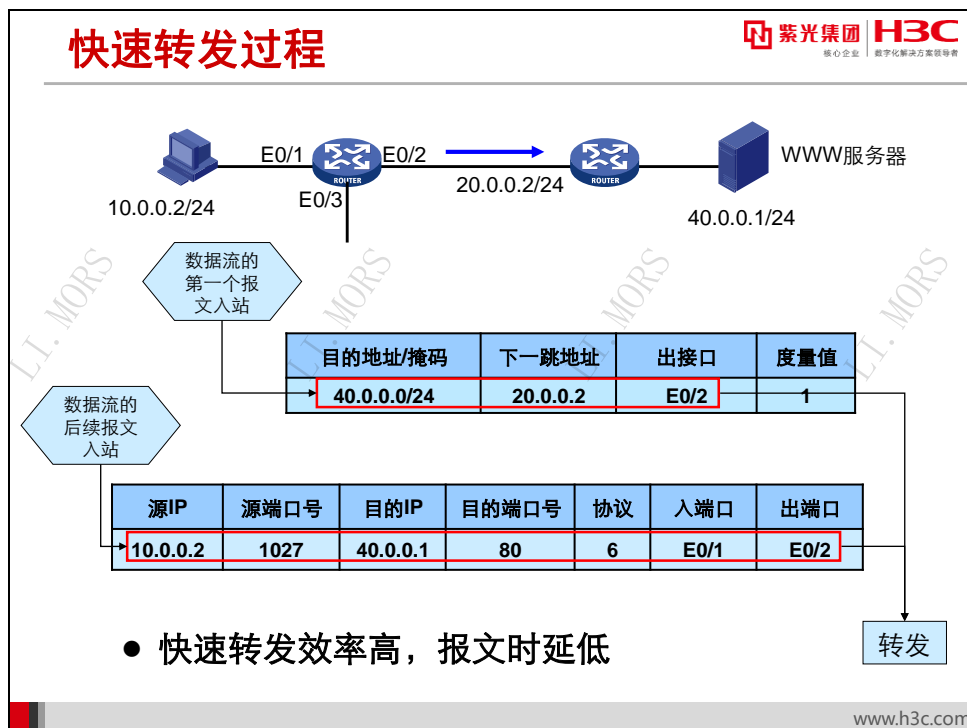
字段	描述
Flag	路由的标志： U：表示路由可用 G：表示网关路由 H：表示主机路由 B：表示黑洞路由 D：表示动态路由 S：表示静态路由 R：表示迭代路由 F：表示快速重路由
OutInterface/Token	转发接口/LSP索引号
Label	内层标签值

3.4 快速转发表



报文转发效率是衡量路由器性能的一项关键指标。按照常规流程，路由器收到一个报文后，将它从接口存储器拷贝至 CPU 中，CPU 根据报文的地址寻找 FIB 表中与之匹配的转发项，然后确定一条最佳的路径，同时还将报文按照数据链路层上使用的协议进行封装，最后，封装后的链路层帧通过 DMA（Direct Memory Access，直接内存访问）拷贝到输出队列中进行报文转发。这个过程两次经过系统总线，每一个报文的转发都要重复这个过程。

快速转发是采用高速缓存来处理报文，采用了基于数据流的技术。Internet 上的数据基本上都是基于数据流的，一个数据流的传输就是指在两台主机之间的一次特定的应用，比如访问 HTTP 服务的一次操作。我们一般用一个 5 元组来描述一个数据流：源 IP 地址、源端口号、目的 IP 地址、目的端口号、协议号。当一个数据流的第一个报文通过查找 FIB 表转发后，在高速缓存中生成相应的转发信息，该数据流后续报文的转发就可以通过直接查找高速缓存来实现。



在上图所示网络中，主机需要访问 WWW 服务器上的 HTTP 应用。它发出的第一个报文到达路由器的接口后，路由器查找快速转发表以期快速转发。但因为这个报文是第一个报文，快速转发表并没有这条数据流的转发信息高速缓存，所以系统并不能进行快速转发。系统只能把这个报文转交到普通的 FIB 转发流程，由 CPU 负责在 FIB 表中查找相关转发项，然后进行封装，从出接口转发出去。与此同时，系统记录报文中的 5 元组信息，在高速缓存中生成相应快速转发信息。

系统根据 5 元组信息生成相应的快速转发信息缓存的同时还记录了转发时的封装信息及接口信息。后续报文来到后，系统查看报文中的 5 元组，如果命中了快速转发缓存，则根据缓存中的封装信息直接进行二层数据帧的封装，然后在中断中直接送到出接口发送。此过程不需要上报 CPU 进行查表操作，也不需要进行内存访问操作，不占用系统总线资源。

快速转发技术大大缩减了 IP 报文的排队流程，减少报文的转发时间，提高 IP 报文的转发吞吐量。同时，由于高速缓存中的转发表已经做过优化，因此查找速度非常快。

快速转发配置和维护



- 显示快速转发表信息

```
[H3C]display ip fast-forwarding cache
```

- 清除快速转发缓冲区

```
<H3C>reset ip fast-forwarding cache
```

- 快速转发功能缺省情况下开启

www.h3c.com

由于快速转发具有增加报文转发效率的优点，所以缺省情况下此功能是开启的。

如果想查看系统目前的快速转发缓存信息，则命令如下：

```
display ip fast-forwarding cache
```

如果想清除快速转发缓冲区，则需要在用户模式下使用如下命令：

```
reset ip fast-forwarding cache
```

快速转发表摘要信息

紫光集团 H3C
核心企业 数字化解决方案领导者

```

[H3C]display ip fast-forwarding cache
Total number of fast-forwarding entries: 4
SIP      SPort  DIP      DPort  Pro  Input_If  Output_If  Flg
30.1.1.1 24321  2.2.2.2  0      1    GE0/0     GE0/1      1
30.1.1.1 24577  2.2.2.2  0      1    GE0/0     GE0/1      1
2.2.2.2  24321  30.1.1.1 2048   1    GE0/1     GE0/0      1
2.2.2.2  24577  30.1.1.1 2048   1    GE0/1     GE0/0      1
  
```

源IP 源端口 目的IP 目的端口 协议 入接口 出接口

www.h3c.com

在查看系统快速转发缓存时，有如下输出：

```

[H3C]display ip fast-forwarding cache
Total number of fast-forwarding entries: 4
SIP      SPort  DIP      DPort  Pro  Input_If  Output_If  Flg
30.1.1.1 24321  2.2.2.2  0      1    GE0/0     GE0/1      1
30.1.1.1 24577  2.2.2.2  0      1    GE0/0     GE0/1      1
2.2.2.2  24321  30.1.1.1 2048   1    GE0/1     GE0/0      1
2.2.2.2  24577  30.1.1.1 2048   1    GE0/1     GE0/0      1
  
```

表中各参数含义如下表：

表3-2 display ip fast-forwarding cache 命令显示信息描述表

字段	描述
SIP	源IP地址
SPort	源端口号
DIP	目的IP地址
DPort	目的端口号
Pro	协议号
Input_If	报文入接口类型和接口号
Output_If	报文出接口类型和接口号

3.5 本章总结

本章总结

- 控制平面和转发平面相分离
- 数据报文实际上是通过FIB表转发
- 系统通过构建快速转发缓存来提高报文转发效率
- 通过命令可查看与配置快速转发

www.h3c.com

3.6 习题和解答

3.6.1 习题

1. 关于控制平面和转发平面，下列说法中正确的是（ ）
 - A. 控制平面指系统中用来传送信令、计算表项的部分
 - B. MAC 地址表学习进程属于转发平面
 - C. 通过 FIB 进行报文转发属于转发平面的范畴
 - D. 控制平面和转发平面必须是物理分离的
2. 路由器查找路由表进行报文转发时，采用如下哪个原则？（ ）
 - A. 最长匹配
 - B. 精确匹配
 - C. 随机转发
 - D. 哈希算法
3. 描述数据流的 5 元组中所包含的元素为：（ ）
 - A. 源 MAC 地址、源 IP 地址、目的 MAC 地址、目的 IP 地址、协议号
 - B. 源 MAC 地址、源端口号、目的 MAC 地址、目的端口号、协议号
 - C. 源 IP 地址、源端口号、目的 IP 地址、目的端口号、接口号
 - D. 源 IP 地址、源端口号、目的 IP 地址、目的端口号、协议号
4. 在路由器上查看快速转发缓存信息的命令是？（ ）
 - A. [H3C] display fast-forwarding
 - B. [H3C] display fast-forwarding cache
 - C. [H3C] display ip fast-forwarding
 - D. [H3C] display ip fast-forwarding cache
5. 关于快速转发过程，下列说法中正确的是（ ）
 - A. 数据流中的第一个报文进行查找 FIB 表操作
 - B. 数据流中的后续报文进行查找 FIB 表操作
 - C. 数据流中的第一个报文进行查找快速转发表操作
 - D. 数据流中的后续报文进行查找快速转发表操作

3.6.2 习题答案

1. AC

2. A

3. D

4. D

5. AD

第4章 路由协议基础

路由器可以从三种方式获得网络中的路由信息，包括从链路层协议直接学习、人工配置静态路由、从动态路由学习。通过使用动态路由协议，路由器可以自动维护路由信息。根据算法的不同，动态路由协议又可以分成链路状态型和距离矢量型。本章介绍了这两种不同算法路由协议的基本原理，并比较了不同算法路由协议之间的异同，并对路由选择过程进行了深入分析。

4.1 本章目标

课程目标

● 学习完本课程，您应该能够：


- 掌握路由的分类
- 掌握距离矢量型路由协议的工作原理
- 掌握链路状态型路由协议的工作原理
- 掌握路由选择过程
- 了解不同类型路由协议的异同



www.h3c.com

4.2 路由分类

路由分类



紫光集团 H3C
核心企业 | 数字化解决方案领导者

- **直连路由**
 - 无须配置及维护，由链路层协议自动发现
- **静态路由**
 - 人工配置及维护，不能自己适应网络拓扑变化
 - 无协议开销
- **动态路由**
 - 协议自动学习、计算，无需人工配置及维护，自动适应网络拓扑变化
 - 路由协议开销大

www.h3c.com

按照路由的来源，可以分为以下 3 种：

● 直连路由

直连路由是由链路层协议发现的。直连路由无须配置，在接口存在 IP 地址时，由路由进程自动生成。它的特点是开销小，配置简单，无需人工维护，但只能发现本接口所属网段的路由。

● 静态路由

由管理员手工配置而成的路由称之为静态路由。静态路由无开销，配置简单，适合简单拓扑结构的网络。静态路由的缺点是无法自动根据网络拓扑变化而改变。当一个网络故障发生后，静态路由不会自动修正，必须有管理员的介入。

● 动态路由

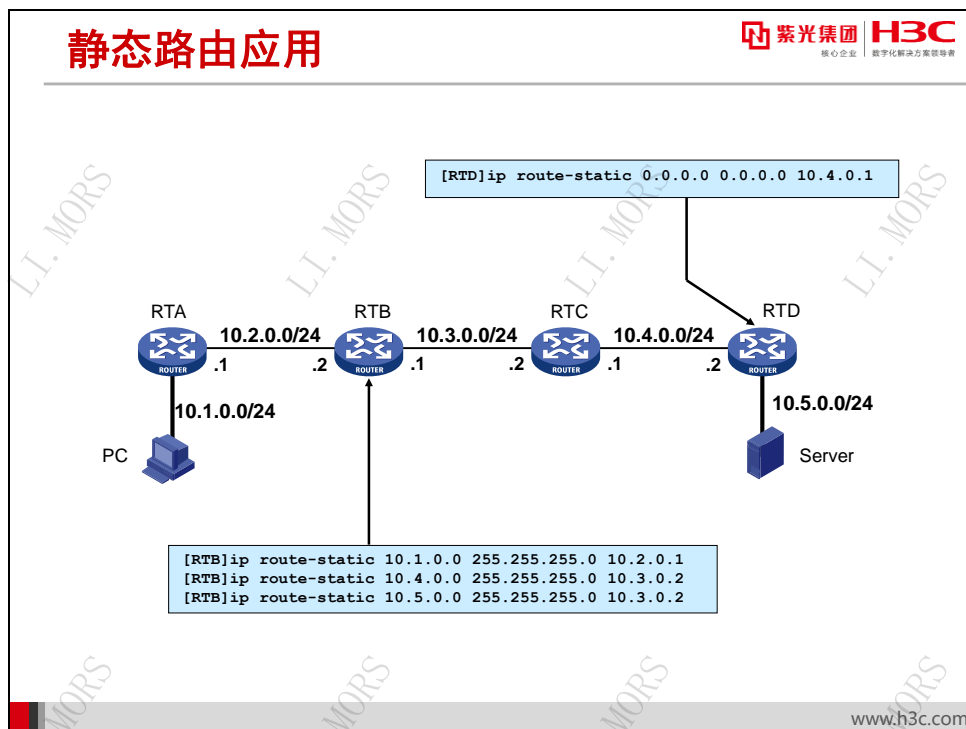
动态路由协议自动发现和维持的路由称为动态路由。动态路由的优点是无须人工配置具体路由表项，而由路由协议自动发现和计算。这样当网络拓扑结构复杂时，使用动态路由可减少管理员的配置工作，且减少配置错误。另外，动态路由协议支持路由备份，如果原有链路故障导致路由表项失效，协议可自动计算和使用另外的路径，无须人工维护。但系统启用动态路由协议后，系统之间交互协议报文，会占用一部分链路开销；并且动态路由协议配置复杂，需要管理员掌握一定的路由协议知识。

各类路由各有优缺点，可根据网络结构和实际需求来选择。

如果网络拓扑是星型，各节点之间没有冗余链路，则可以使用静态路由；如果网络中有冗余链路，如全互连或环型拓扑，则可以使用动态路由，以增强路由可靠性。

如果网络是分层的，则通常在接入层使用静态路由，以降低设备资源的消耗；而在汇聚或核心层使用动态路由，以增加可靠性。

4.3 静态路由应用



恰当地设置和使用静态路由可以改进网络的性能，并可为重要的网络应用保证带宽。而在路由器上合理配置缺省路由能够减少路由表中表项数量，节省路由表空间，加快路由匹配速度。

如上图，可以在 RTA 和 RTD 上配置缺省静态路由，而在 RTB 和 RTC 上配置静态路由。

配置 RTA:

```
[RTA]ip route-static 0.0.0 0.0.0.0 10.2.0.2
```

配置 RTB:

```
[RTB]ip route-static 10.1.0.0 255.255.255.0 10.2.0.1
[RTB]ip route-static 10.4.0.0 255.255.255.0 10.3.0.2
[RTB]ip route-static 10.5.0.0 255.255.255.0 10.3.0.2
```

配置 RTC:

```
[RTC]ip route-static 10.1.0.0 255.255.255.0 10.3.0.1
[RTC]ip route-static 10.2.0.0 255.255.255.0 10.3.0.1
[RTC]ip route-static 10.5.0.0 255.255.255.0 10.4.0.2
```

配置 RTD:

```
[RTD]ip route-static 0.0.0.0 0.0.0.0 10.4.0.1
```

配置静态路由时，要注意双向配置，避免出现单程路由。因为几乎所有的 Internet 应用如 HTTP、FTP 等都是双向传输，所以单程路由对用户的业务是没有意义的。

4.4 动态路由协议

路由协议概览		
		紫光集团 H3C 核心企业 数字化解决方案领导者
路由协议	协议算法	IGP/EGP
RIP	距离矢量	IGP
OSPF	链路状态	IGP
IS-IS	链路状态	IGP
BGP	路径矢量	EGP

根据作用的范围，路由协议可分为：

- 内部网关协议（Interior Gateway Protocol，简称 IGP）：在一个自治系统内部运行，常见的 IGP 协议包括 RIP、OSPF 和 IS-IS。
- 外部网关协议（Exterior Gateway Protocol，简称 EGP）：运行于不同自治系统之间，BGP 是目前最常用的 EGP。

根据使用的算法，路由协议可分为：

- 距离矢量协议（Distance-Vector）：包括 RIP 和 BGP。其中，BGP 也被称为路径矢量协议（Path-Vector）。
- 链路状态协议（Link-State）：包括 OSPF 和 IS-IS。

以上两种算法的主要区别在于发现和计算路由的方法不同。

说明：

自治系统（Autonomous System）是拥有同一选路策略，并在同一技术管理部门下运行的一组路由器。

动态路由协议工作原理

紫光集团 H3C
核心企业 数字化转型领导者

● 邻居发现

→ 路由器通过发送广播报文或发送给指定的路由器邻居以主动把自己介绍给网段内的其它路由器。

● 路由交换

→ 每台路由器将自己已知的路由相关信息发给相邻路由器。

● 路由计算

→ 每台路由器运行某种算法，计算出最终的路由来。

● 路由维护

→ 路由器之间通过周期性地发送协议报文来维护邻居信息。

www.h3c.com

各种动态路由协议所共同的目的是计算与维护路由。通常，各种动态路由协议的工作过程大致相同，都包含以下几个阶段：

● 邻居发现

运行了某种路由协议的路由器会主动把自己介绍给网段内的其它路由器。路由器通过发送广播报文或发送给指定的路由器邻居来做到这一点。

● 交换路由信息

发现邻居后，每台路由器将自己已知的路由相关信息发给相邻的路由器，相邻路由器又发送给下一台路由器。这样经过一段时间，最终每台路由器都会收到网络中所有的路由信息。

● 计算路由

每一台路由器都会运行某种算法，计算出最终的路由来。（实际上需要计算的是该条路由的下一跳和度量值）

● 维护路由

为了能够观察到某台路由器突然失效（路由器本身故障或连接线路中断）等异常情况，路由协议规定两台路由器之间的协议报文应该周期性地发送。如果路由器有一段时间收不到邻居发来的协议报文，则认为邻居失效了。

各个路由协议工作原理大体类似，但在实现细节上会有所不同。

距离矢量型路由协议特点

- 周期性、广播式发送路由更新
- 路由更新中携带全部路由表，接收方据此更新自己的路由
- 超过一定时间接收不到路由更新，则认为路由失效
- 以到目的地的距离（跳数）作为度量值
- 拓扑变化以逐跳的方式扩散
- 路由收敛速度慢
- 采用距离矢量算法，可能导致路由环路

距离矢量路由协议通常不维护邻居信息。在开始阶段，采用这种算法的路由器以广播或组播发送协议报文，请求邻居的路由信息；邻居路由器回应的协议报文中携带全部路由表，这样就完成路由表的初始化过程。

为了维护路由信息，路由器以一定的时间间隔向相邻的路由器发送路由更新，路由更新中携带本路由器的全部路由表。系统为路由表中的表项设定超时时间，如果超过一定时间接收不到路由更新，则系统认为原有的路由失效，会将其从路由表中删除。

距离矢量路由协议以到目的地的距离（跳数）作为度量值，距离越大，路由越差。但采用跳数作为度量值并不能完全反映链路带宽的实际状况，有时会造成协议选择了次优路径。

当网络拓扑发生变化时，距离矢量路由协议首先向邻居通告路由更新。邻居路由器根据收到的路由更新来更新自己的路由，然后再继续向外发送更新后的路由。这样，拓扑变化的信息会以逐跳的方式扩散到整个网络。

距离矢量路由协议基于贝尔曼-福特算法，也称 D-V 算法。这种算法的特点是计算路由时只考虑到目的网段的距离和方向。系统从邻居接收到路由更新后，将路由更新中的路由表项加入到自己的路由表中，其度量值在原来基础上加一，表示经过了一跳；并将路由表项的下一跳置为邻居路由器的地址，表示是经过邻居路由器学到的。距离矢量路由协议完全信任邻居路由器，它并不知道整个网络的拓扑环境，这样在环型拓扑网络中可能会产生路由环路。所以采用 D-V 算法的路由器采用了一些避免环路的机制，如水平分割、路由毒化、毒性逆转等。

RIP 协议是一种典型的距离矢量路由协议。它的优点是配置简单，算法占用较少的内存和 CPU 处理时间。它的缺点是算法本身不能完全杜绝路由自环，收敛相对较慢，周期性广播路由更新占用网络带宽较大，扩展性较差，最大跳数不能超过 16 跳。

链路状态型路由协议特点

- 通过Hello报文来发现邻居
- 建立邻接关系后，只发送链路状态公告（LSA）
- 根据自己链路状态信息库（LSDB）来计算路由
- 以到目的地的开销（cost）作为度量值
- 链路状态变化时，马上发送LSA到区域内所有路由器
- 路由收敛速度快
- 采用SPF算法，无路由自环

链路状态路由协议基于 Dijkstra 算法，有时被称为最短路径优先算法。

在开始阶段，采用这种算法的路由器以组播方式发送 Hello 报文，来发现邻居。收到 Hello 报文的邻居路由器会检查报文中所定义参数，如果双方一致就会形成邻居关系。有路由信息交换需求的邻居路由器会生成邻接关系，进而可以交换 LSA（Link State Advertisement，链路状态通告）。

链路状态路由协议用 LSA 来描述路由器周边的网络拓扑和链路状态。邻接关系建立后，路由器会将自己的 LSA 发送给区域内的所有邻接路由器，同时也从邻接路由器接收 LSA。每台路由器都会收集其它路由器通告的 LSA，所有的 LSA 放在一起便组成了 LSDB（Link State Database，链路状态数据库）。LSDB 是对整个自治系统的网络拓扑结构的描述。

路由器将 LSDB 转换成一张带权的有向图，这张图便是对整个网络拓扑结构的真实反映。各个路由器得到的有向图是完全相同的。每台路由器根据有向图，使用最短路径优先算法计算出一棵以自己为根的最短路径树，这棵树给出了到自治系统中各节点的路由。

链路状态路由协议以到目的地的开销（Cost）作为度量值。路由器根据该接口的带宽自动计算到达邻居的权值，带宽与权值成反比，带宽越高，权值越小，表示到邻居的路径越好。在使用最短路径优先算法计算最短路径树时，将自己到各节点的路径上权值相加，也就计算出了到达各节点的开销，将此开销作为路由度量值。


当网络拓扑发生变化时，路由器并不发送路由表，而只是发送含有链路变化信息的 LSA。LSA 在区域内扩散，所有路由器都收到，然后更新自己的 LSDB，再运行 SPF 算法，重新计算路由。这样的好处是带宽占用小，路由收敛速度快。

因为采用链路状态路由协议的路由器知道整个网络的拓扑，且采用 SPF 算法，从根本上避免了路由环路产生。

OSPF 和 IS-IS 是链路状态路由协议。它们能够完全杜绝协议内的路由自环，且采用增量更新方式来通告变化的 LSA，占用带宽少。OSPF 和 IS-IS 采用路由分组和区域划分等机制，所以能够支持大规模的网络，且扩展性较好。但相对 RIP 来讲，OSPF 和 IS-IS 的配置较复杂一些。

路径矢量型路由协议特点

- 仅在邻居刚建立时发送全部路由表
- 邻居建立后发送增量路由
- 如果邻居失效，则认为路由失效
- 丰富的路由属性作为度量值
- 拓扑变化以逐跳的方式扩散
- 采用机制防止路由环路



紫光集团 H3C
核心企业 数字化解决方案领导者

www.h3c.com

路径矢量路由协议结合了距离矢量路由协议和链路状态路由协议的优点。

路径矢量路由协议采用单播方式与相邻路由器建立邻居关系。邻居关系建立后，根据预先配置的策略，路由器将全部或部分带有路由属性的路由表发送给邻居。邻居收到路由表后，根据预先配置的策略将全部或部分路由信息加入到自己的路由表中。

当路由信息发生变化时，路径矢量路由协议只发送增量路由给邻居，减少带宽的消耗。

邻居关系是以单播方式，通过 TCP 三方握手形式建立；并且在建立后定时交换 Keepalive 报文，以维持邻居关系正常。如果邻居断开，则相关路由失效。

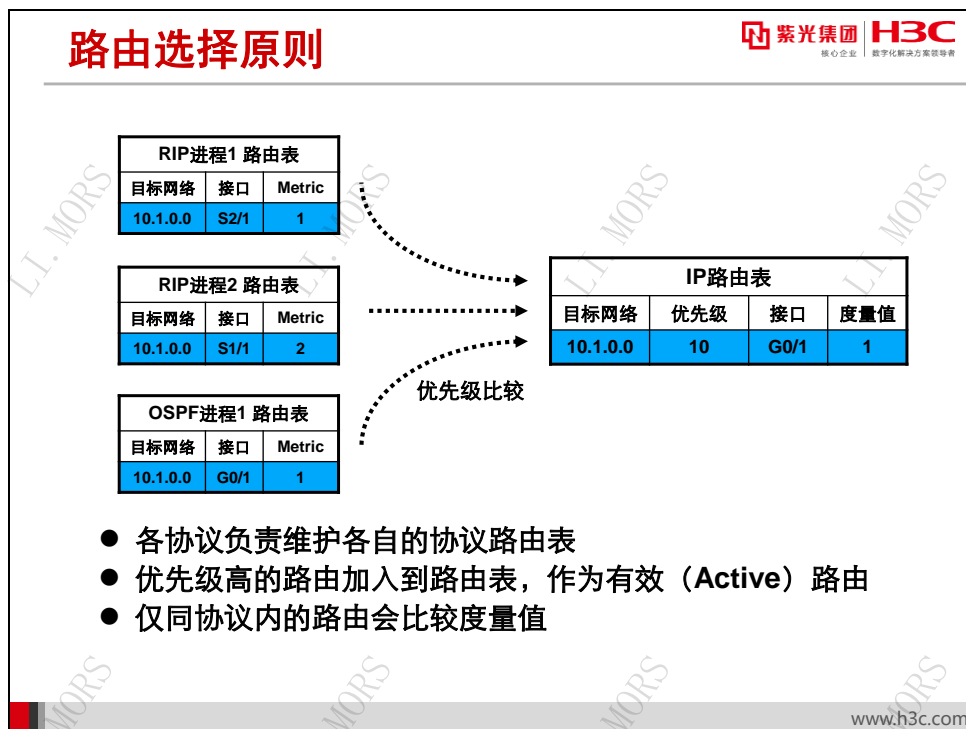
路径矢量路由协议采用丰富的路由属性作为路由度量值。属性包括路由的起源、到目的地的距离、本地优先级、MED 值等，且这些路由属性都可根据网络实际情况由管理员自己进行修改。

在拓扑发生变化时，路径矢量路由协议仅将变化的路由信息发送给邻居路由器，以逐跳的方式在全网络内扩散。但由于采用触发更新机制，变化的路由能够很快通知到整个网络。

BGP 协议是路径矢量路由协议。它采用一些方法能够防止路由环路。**BGP** 协议把 **AS** 间传递的路由都记录了经过的 **AS** 号码，这样路由器接收到路由时可以据此查看此条路由是不是自己发出的。在 **AS** 内，**BGP** 协议规定路由器不能把从邻居学到的路由再返回给邻居。

BGP 通过与邻居路由器建立对等体来交换路由信息，并采用增量更新机制来发送路由更新，只有当路由表变化时才发送路由更新信息，节省了相邻路由器之间的链路带宽。

4.5 路由选择原则



每个路由协议都维护了自己的路由表，称为协议路由表。协议路由表中只记录了本路由协议学习和计算的路由。

大多数路由协议都支持多进程。各个协议进程之间互不影响，相互独立。各个进程之间的交互相当于不同路由协议之间的路由交互。

各个路由协议的各个进程独立维护自己的路由表，然后统一汇总到 IP 路由表中。IP 路由表首先选择路由协议优先级高的路由使用。如果协议优先级一致，则再选择度量值最优的路由，作为 IP 路由表的有效（Active）路由，指导 IP 报文转发。其余的路由作为备份，如果有有效路由失效，再进行重新选择。

路由度量值只在同一种路由协议内有比较意义，不同的路由协议之间的路由度量值没有可比性，也不存在换算关系。

各类路由缺省优先级

路由类型	缺省优先级
直连路由 (Direct)	0
OSPF内部路由	10
静态路由 (Static)	60
RIP路由	100
OSPF外部路由	150
BGP路由	255

对于相同的目的地，不同的路由协议（包括静态路由）可能会发现不同的路由，但这些路由并不都是最优的。事实上，在某一时刻，到某一目的地的当前路由仅能由唯一的路由协议来决定。为了判断最优路由，各路由协议（包括静态路由）都被赋予了一个优先级，当存在多个路由信息源时，具有较高优先级的路由协议发现的路由将成为当前路由。各种路由协议及其发现路由的缺省优先级如下表所示。

表4-1 路由协议及缺省时的路由优先级

路由协议或路由种类	相应路由的优先级
DIRECT	0
OSPF	10
IS-IS	15
STATIC	60
RIP	100
OSPF ASE	150
OSPF NSSA	150
IBGP	255
EBGP	255
UNKNOWN	256

其中：**0**表示直接连接的路由，**256**表示任何来自不可信源端的路由。数值越小表明优先级越高。除直连路由（**DIRECT**）外，各种路由的优先级都可由用户手工进行配置。另外，每条静态路由的优先级都可以不相同。

4.6 路由协议比较

路由协议可靠性、安全性			
协议	协议端口	可靠性	安全性（是否支持验证）
RIP-1	UDP 520	低	否
RIP-2	UDP 520	低	是
OSPF	IP 89	高	是
IS-IS	基于链路层协议	高	是
BGP	TCP 179	高	是

目前常用的路由协议包括 RIP-1/2, OSPF, IS-IS, BGP 等。本节对其协议特点进行全面的比较。

RIP 协议是最早的路由协议，其设计思想是为小型网络中提供简单易用的动态路由，其算法简单，对 CPU 和内存资源要求低。RIP 采用广播（RIP-1）或组播（RIP-2）方式在邻居间传送协议报文，传输层采用 UDP 封装，端口号是 520。由于 UDP 是不可靠的传输层协议，所以 RIP 设计成为周期性的广播全部路由表，如果邻居超过 3 次无法收到路由更新，则认为路由失效。RIP-1 协议不支持验证，其安全性较低；RIP-2 对其进行了改进，从而能够支持验证，安全性提高了。

OSPF 是目前应用最广泛的 IGP 协议。OSPF 设计思想是为大中型网络提供分层次的、可划分区域的路由协议。其算法复杂，但能够保证无域内环路。OSPF 采用 IP 来进行承载，所有的协议报文都由 IP 封装后进行传输，端口号是 89。IP 是尽力而为的网络层协议，本身是不可靠的；所以为了保证协议报文传输的可靠性，OSPF 采用了确认机制：在邻居发现阶段，交互 LSA 的阶段，OSPF 都采用确认机制来保证传输可靠。OSPF 支持验证，使 OSPF 的安全性得到了保证。

IS-IS 是另外一种链路状态型的路由协议，其同样采用 SPF 算法，支持路由分组管理与划分区域，同样可应用在大中型网络中，可扩展性好。与 OSPF 不同的是，IS-IS 的运行直接基于链路层，其所有协议报文通过链路层协议来承载。所以 IS-IS 也可以运行在无 IP 的网络中，如 OSI 网络中。为了保证协议报文传输的可靠性，IS-IS 同样设计了确认机制来保证协议报文在传输过程中没有丢失。IS-IS 也支持验证，安全性得到了保证。

BGP 协议是唯一的 EGP 协议。与其它协议不同，BGP 采用 TCP 来保证协议传输的可靠性，TCP 端口号是 179。TCP 本身有三方握手的确认机制，运行 BGP 的路由器首先建立可靠的 TCP 连接，然后通过 TCP 连接来交互 BGP 协议报文。这样，BGP 协议不需要自己设计可靠传输机制，降低了协议报文的复杂度和开销。另外，BGP 的安全性也可以由 TCP 来保证，TCP 支持验证功能，通过验证的双方才能够建立 TCP 连接。

BGP 自己不学习路由，它的路由来源于 IGP 协议如 OSPF 等。管理员手工指定哪些 IGP 路由能够导入到 BGP 中，并手工指定 BGP 能够与哪些邻居建立对等体关系从而交换路由信息。

路由协议特性比较					
特性	RIP-1	RIP-2	OSPF	IS-IS	BGP
距离矢量算法	√	√	—	—	√
链路状态算法	—	—	√	√	—
支持VLSM	—	√	√	√	√
支持手工聚合	—	√	√	√	√
支持自动聚合	√	√	—	—	√
支持无类别	—	√	√	√	√
收敛速度	慢	慢	快	快	慢
度量值	跳数	跳数	开销	开销	路径属性

RIP 与 BGP 协议属于距离矢量型路由协议，其中 BGP 又属于路径矢量型路由协议。由于 RIP-1 是早期的路由协议，所以其不支持无类别（Classless）路由，只能支持按类自动聚合，不支持可变长子网掩码（VLSM），所以其应用有一定限制。

除 RIP-1 外，其它路由协议都能够支持 VLSM 和手工聚合，这样能够对网络进行很细致的子网划分和汇聚，从而节省 IP 地址，减少路由表数量。

由于 RIP 和 BGP 协议的路由更新需要以逐跳的方式进行传播，所以路由收敛速度慢。而链路状态型的路由协议采用 SPF 算法，根据自己的 LSDB 进行路由计算，所以收敛速度快。

RIP 是使用跳数作为度量值，而 OSPF 和 IS-IS 使用开销作为度量值。BGP 的度量值较为复杂，它包含了多个属性，并可手工修改属性值以控制路由。

路由协议定时器比较

紫光集团 H3C
核心企业 数字化转型领导者

协议	周期性发送全部路由	Hello定时器	保持定时器
RIP-1	30秒	—	180秒
RIP-2	30秒	—	180秒
OSPF	触发更新	广播链路上10秒	4倍Hello定时器
IS-IS	触发更新	10秒	4倍Hello定时器
BGP	触发更新	60秒	3倍Hello定时器

www.h3c.com

所有的路由协议都采用定时器来维护邻居关系和路由信息。

RIP 协议不需要建立邻居关系，直接交换路由信息。RIP 协议定义了 Update 定时器，表示发送路由更新的时间间隔，其缺省时间是 30 秒；同时定义了 Timeout 定时器，表示路由老化时间，缺省参数是 180 秒。如果在老化时间内没有收到关于某条路由的更新报文，则该条路由在路由表中的度量值将会被设置为 16，表示无效路由。

OSPF 和 IS-IS 需要首先建立邻居关系，然后在形成邻接关系的路由器之间交互 LSA。OSPF 定义了 Hello 定时器，表示接口向邻居发送 Hello 报文的时间间隔，其广播网络类型链路上的缺省时间是 10 秒；同时定义了邻居失效时间，广播网络类型链路上的缺省时间是 40 秒。在邻居失效时间内，如果接口还没有收到邻居发送的 Hello 报文，路由器就会宣告该邻居无效。

BGP 采用 TCP 来建立 BGP 对等体，然后交换 BGP 路由。当对等体间建立了 BGP 连接后，它们定时向对端发送存活（Keepalive）消息，以防止路由器认为 BGP 连接已中断。若路由器在设定的连接保持时间（Holdtime）内未收到对端的 Keepalive 消息或任何其它类型的报文，则认为此 BGP 连接已中断，从而断开此 BGP 连接。缺省情况下，BGP 的存活时间间隔为 60 秒，保持时间为 180 秒。

4.7 本章总结

本章总结

- 路由包括直连、静态、动态等
- 距离矢量型路由协议的工作原理
- 链路状态型路由协议的工作原理
- 系统通过优先级来进行不同协议间路由选择
- 距离矢量型路由协议与链路状态型路由协议的比较

www.h3c.com

4.8 习题和解答

4.8.1 习题

1. 以下哪些是动态路由协议的优点？（ ）
A. 无须人工维护路由表项 B. 协议本身占用链路带宽小
C. 由链路层协议发现，无须配置 D. 能够自动发现拓扑变化
2. 缺省情况下，静态路由的优先级是（ ）
A. 0 B. 10 C. 60 D. 100
3. 下列哪些路由协议能够支持手工聚合？（ ）
A. OSPF B. BGP
C. RIPv1 D. IS-IS
4. 下列哪些路由协议是基于 TCP 承载的？（ ）
A. OSPF B. BGP
C. RIPv1 D. IS-IS
5. 下列哪些是路径矢量型路由协议的特点？（ ）
A. 邻居建立后发送增量路由 B. 采用机制能够防止路由环路
C. 具有丰富的路由属性 D. 路由收敛速度快

4.8.2 习题答案

1. AD
2. C
3. ABD
4. B
5. ABC

第5章 路由负载分担与备份

通过在路由表中生成具有多个不同下一跳的等值路由，路由可以在多路径上实现负载分担。同时，合理的配置静态与动态路由协议，可以在网络中实现路由备份，提高路由可靠性。本章介绍静态等值路由的配置，如何通过浮动静态路由对动态路由实现备份，如何对拨号网络中的动态路由实现备份。

5.1 本章目标

课程目标

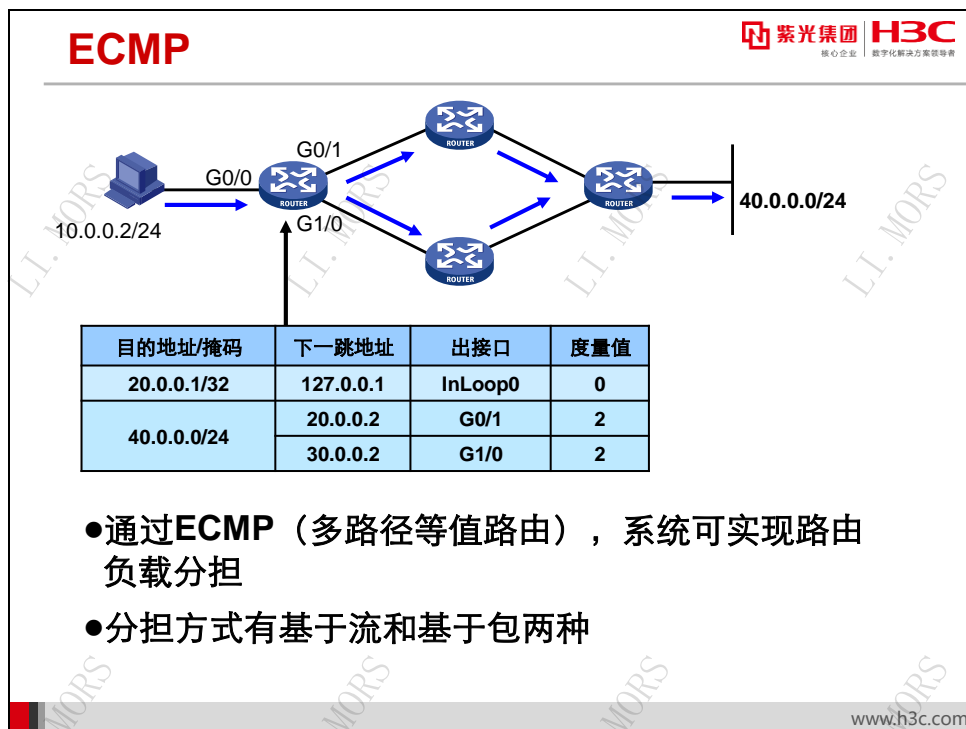
● 学习完本课程，您应该能够：

- 掌握路由负载分担原理
- 掌握路由备份的原理和应用
- 掌握浮动静态路由的原理及应用
- 掌握动态路由备份的原理及应用



www.h3c.com

5.2 路由负载分担



ECMP（Equal-cost multi-path routing，多路径等值路由），也称为等价路由，表示到达一个目的地有多条相同度量值的路由项。

对于同一目的地，路由协议可能会发现几条等值的路由，如果该路由协议在所有活跃的路由协议中优先级最高，那么这几条不同的路由都被看作当前有效的路由。或者，管理员可以手工配置到同一目的地的几条等值的路由作为有效路由。

路由器对数据报文进行转发时，如果发现到目的地有多条最优路径，会将数据按照一定的策略在多条路径上依次发送。通过 **ECMP**，在路由协议层面上实现了 **IP** 流量的负载分担。

负载分担方式有基于流和基于包 2 种。

- 基于流的负载分担

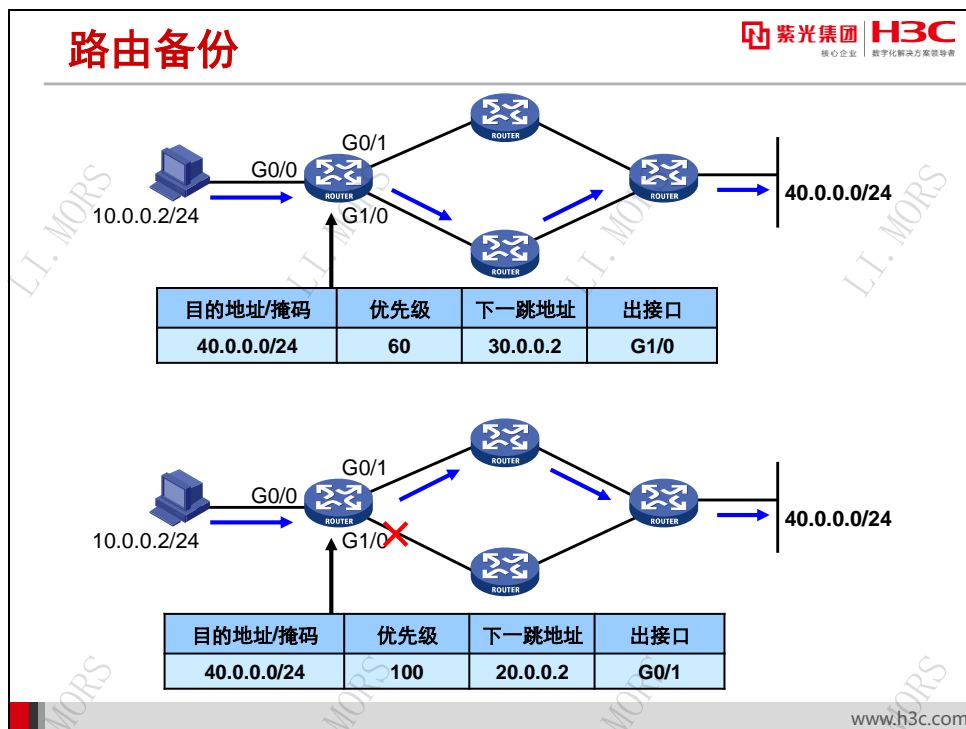
路由器根据 **IP** 报文中的 5 元组信息将数据分成不同的流。具有相同 5 元组信息的 **IP** 报文属于同一个流。转发数据时，路由器把不同的数据流根据算法从多个路径上依次发送出去。

- 基于包的负载分担

转发数据时，路由器把数据包从多个路径上依次发送出去。

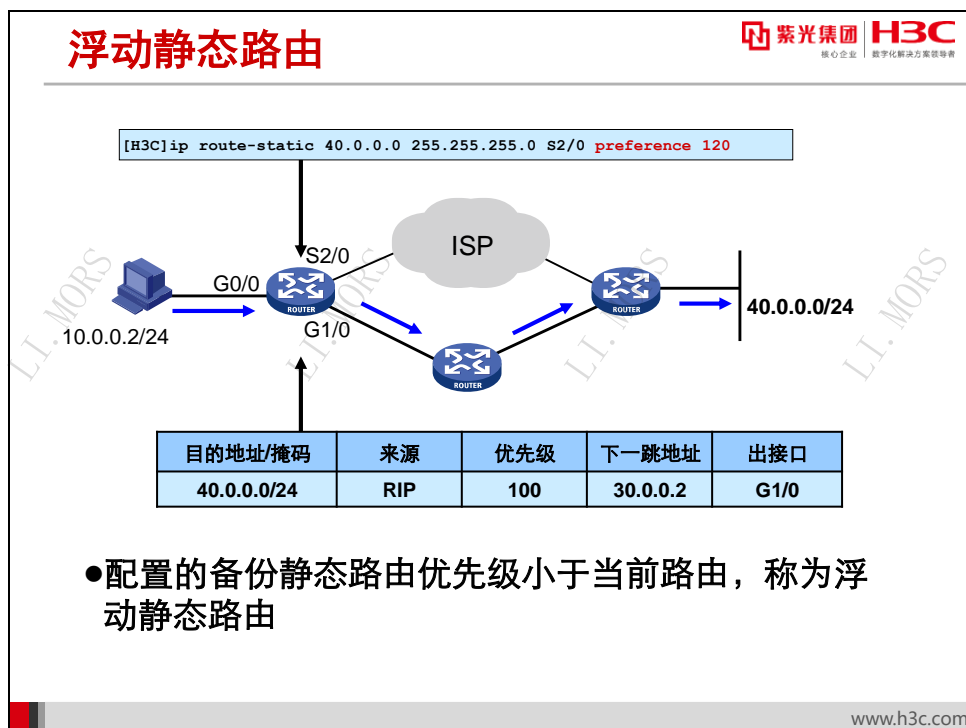
基于包转发能够做到更精确的负载分担。但是由于路由器要对每一个包都进行路由查表与转发操作，无法使用快速转发缓存来转发数据，所以转发效率降低了。另外，**Internet** 应用都是基于流的，如果路由器采用基于包的负载分担，一条流中的数据包会经过不同路径到达目的地，可能会造成接收方的乱序接收，影响应用程序的正常运行。

5.3 路由备份



使用路由备份可以提高网络的可靠性。用户可根据实际情况，配置到同一目的地的多条路由，其中优先级最高的一条路由作为主路由，其余优先级较低的路由作为备份路由。

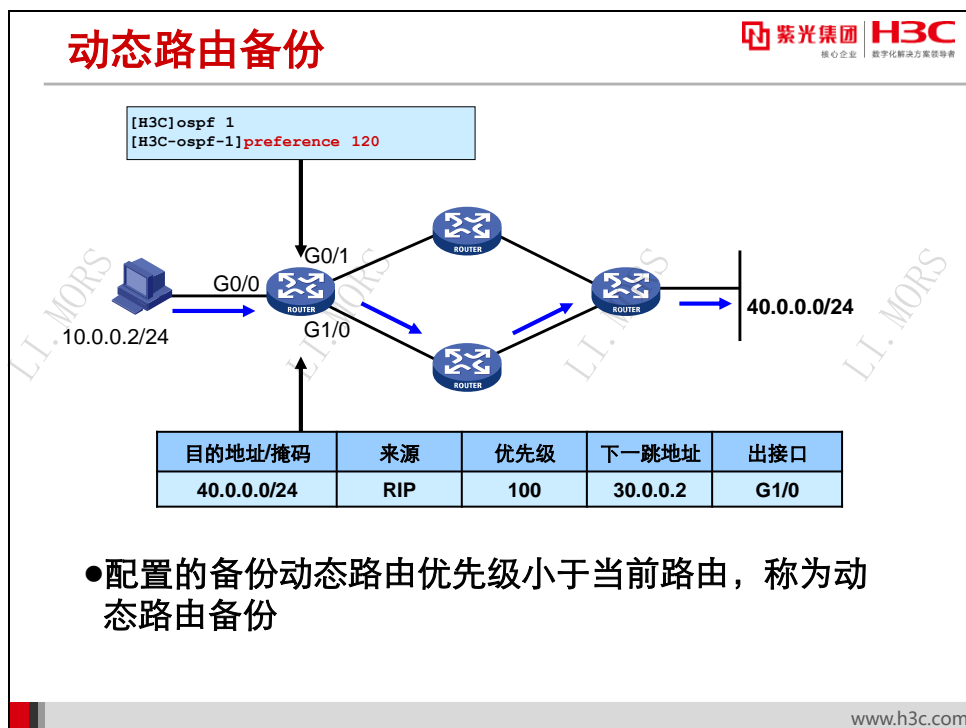
正常情况下，路由器采用主路由转发数据。当线路出现故障时，该路由变为非激活状态，路由器选择备份路由中优先级最高的转发数据。这样，也就实现了从主路由到备份路由的切换。当主路由恢复正常时，路由器也恢复相应的路由，并重新选择路由。由于该路由的优先级最高，路由器选择主路由来发送数据。这就是从备份路由到主路由的切换。



常见的一种路由备份方法是使用静态路由来备份动态路由。

路由器在主链路上配置动态路由协议，与邻居交换路由信息；同时配置静态路由，指定优先级低于主路由，下一跳指向备份链路。由于主路由优先级高，所以正常情况下路由器采用主路由转发数据。如果主链路出现故障，主路由的下一跳不可达，变为非激活状态；此时静态路由“浮出”来，被激活，路由器选择备份链路来转发数据。当主链路恢复时，原主路由由恢复激活状态，路由器重新选择主路由来发送数据，转发路径切换回主链路上。

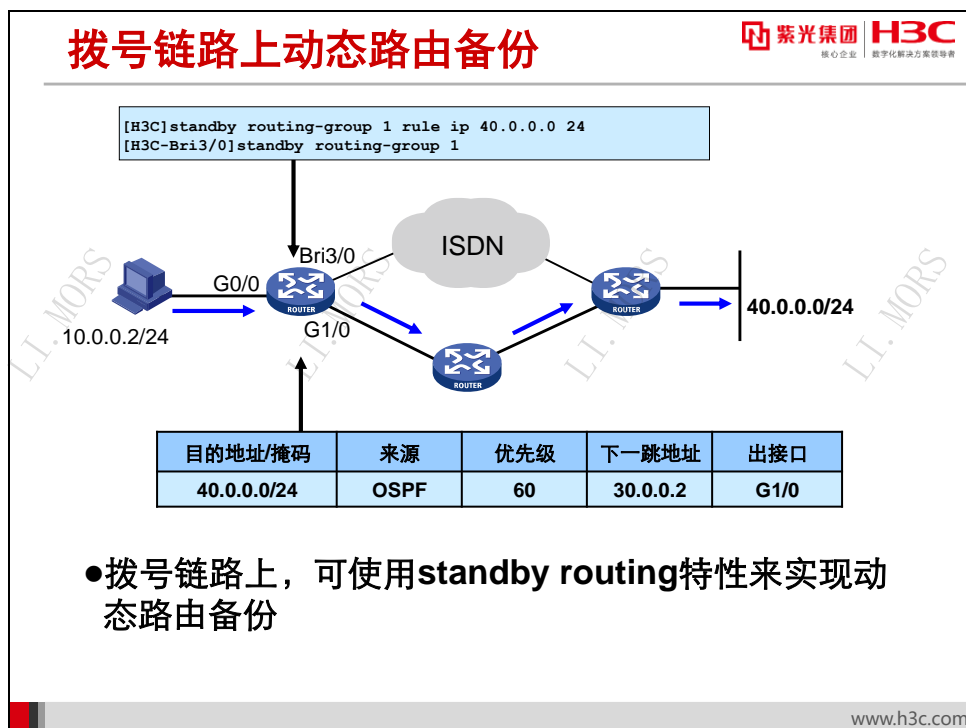
因为浮动静态路由没有链路带宽开销，所以通常适合于备份链路是低带宽链路的场合。



也可以采用一种动态路由来备份另一种动态路由。

路由器同时运行 2 种动态路由协议，并通过 2 种协议学习到了同一目的地的路由。使用动态路由备份时，需要管理员手工配置备份动态路由优先级小于主动态路由。

由于路由器同时运行 2 种动态路由协议，对路由器的 CPU 和内存要求较高；另外，动态路由协议开销较大，只有在备份链路对带宽消耗不敏感的情况下，才可以使用动态路由备份。



动态路由协议需要与邻居路由器交换路由信息。如果在拨号链路上运行了动态路由协议，协议会发周期性的发送协议报文，从而导致拨号接口周期性的向外拨号，产生不必要的费用。

如果要想在拨号链路上做路由备份，其中一个解决办法是配置浮动静态路由，另外一个解决办法是配置动态路由备份（**standby routing**）。

要实现动态路由备份功能，须要在路由器配置动态路由备份规则（**standby routing-rule**），规则中指定监控网段；同时在拨号接口上配置动态路由备份组（**standby routing-group**），以当被监控网段不可达时，路由器可以在拨号接口上发出拨号请求，激活拨号链路。拨号链路激活后，路由器通过拨号链路与邻居交换路由信息，从而学习到监控网段的路由，此时路由表项的出接口为本地拨号接口，数据通过拨号接口转发。当原来的主链路恢复后，路由器断开拨号链路，路由表项中的出接口切换回主链路接口，数据通过主接口转发。

如图中所示，路由器要实现动态路由备份功能，须首先在系统视图下创建动态备份路由规则：

```
[H3C]standby routing-group 1 rule ip 40.0.0.0 24
```

然后将规则应用到拨号接口上：

```
[H3C-Bri3/0]standby routing-group 1
```

5.4 本章总结

本章总结

- 通过等值路由可实现负载分担
- 通过浮动静态路由可对动态路由实行备份
- 动态路由间可实现互相备份
- 拨号链路上可使用**standby routing**特性实现动态路由备份

www.h3c.com

5.5 习题和解答

5.5.1 习题

- 关于路由的负载分担，以下哪些说法是正确的？（ ）
 - 基于包转发能够做到更精确的负载分担
 - 基于流转发效率比基于包要低
 - 路由器通过 ECMP 实现路由层面的负载分担
 - 为了防止应用程序接收方的乱序接收，应该采用基于流的负载分担方式
- 某公司分部使用 2 条链路连接到总部。为提高路由可靠性，使用低速链路（64K 的 DDN 线路）备份高速链路（2M 的 DDN）。主链路上运行了 OSPF 协议，则在备份链路上运行哪些路由协议比较合适？（ ）
 - 浮动静态路由
 - RIP 动态路由
 - standby routing 特性动态路由
 - OSPF 动态路由
- 某公司分部使用 2 条链路连接到总部。为提高路由可靠性，使用低速链路（56K 的 PSTN 线路）备份高速链路（2M 的 DDN）。主链路上运行了 OSPF 协议，则在备份链路上运行哪些路由协议比较合适？（ ）
 - 浮动静态路由
 - RIP 动态路由
 - standby routing 特性动态路由
 - OSPF 动态路由
- 某公司分部使用 2 条高速链路（100M 以太网线路）连接到总部。以下哪种路由方案能够同时提高路由可靠性和增加链路带宽？（ ）
 - 主链路运行 OSPF，备份链路运行浮动静态路由
 - 主链路和备份链路同时运行 OSPF，实现 ECMP
 - 主链路和备份链路同时运行 OSPF，调整链路开销以实现路由备份
 - 主链路运行 OSPF，备份链路运行静态路由，调整路由优先级以实现路由备份

份

5.5.2 习题答案

1. ACD 2. A 3. AC 4. B

第6章 路由聚合与 CIDR

在大规模网络中，数量众多的路由使得设备转发效率低下。通过应用路由聚合和 CIDR 技术，路由数量能够得到一定程度的限制。本章介绍如何在静态和 RIP 协议中配置路由聚合，并介绍聚合可能产生的问题和解决方法。最后，对 CIDR 技术的优点进行了总结。

6.1 本章目标

课程目标

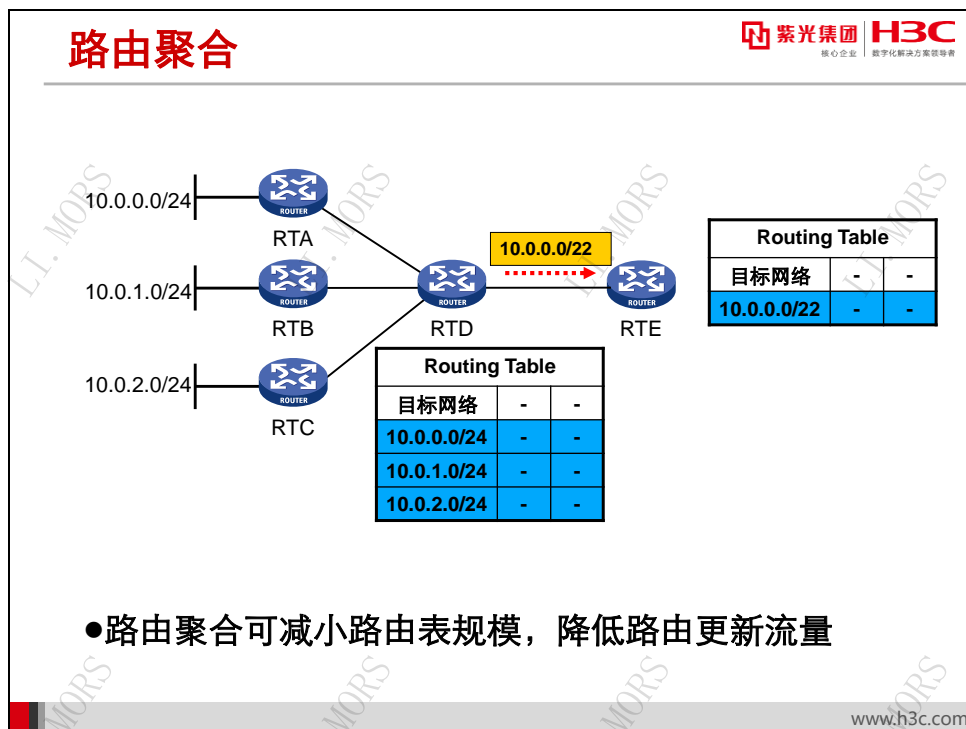
● 学习完本课程，您应该能够：

- 掌握路由聚合的种类和特点
- 掌握RIP协议中的聚合配置
- 了解聚合产生环路及解决方法
- 掌握CIDR的原理及应用



www.h3c.com

6.2 路由聚合

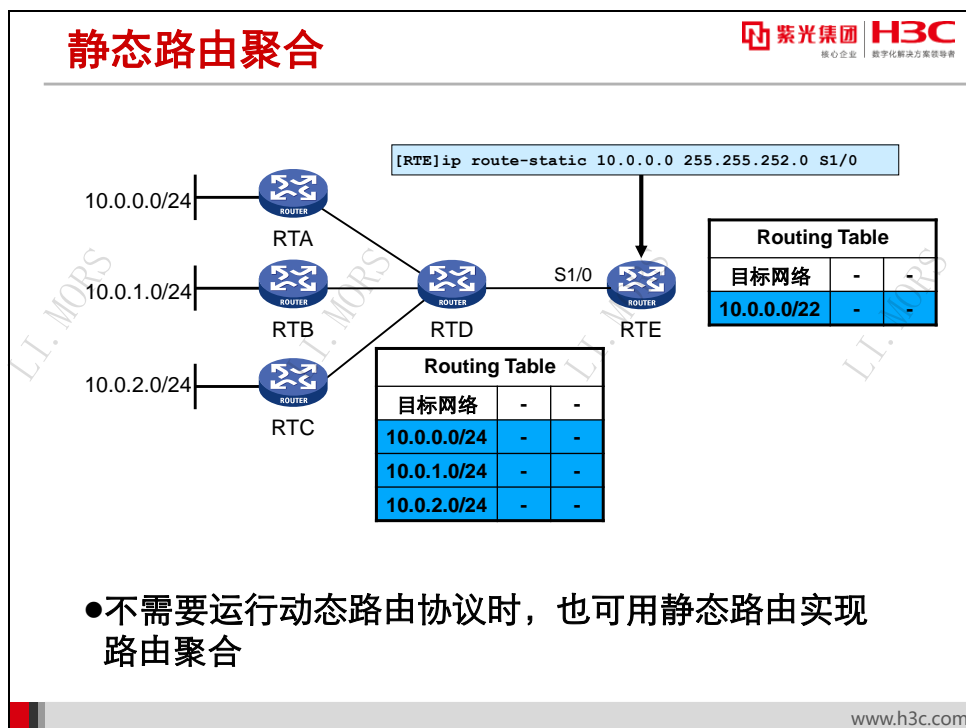


路由器在转发数据报文时，要进行路由表的查表操作，找出其中掩码最长的路由项用于转发。路由表中路由项数量越多，所需查找及匹配的次数则越多，所消耗的 CPU 及内存资源也越大。

有些路由协议如 RIP，在与邻居路由器交换路由信息时，需要发送全部路由表。如果路由表项数量众多，则会占用有限的链路带宽资源。

路由聚合是指将同一网段内的不同子网的路由聚合成 1 条路由向外（其它网段）发送，目的是为了减小路由表的规模，从而减少网络上的流量。

在上图所示网络中，路由器将接收到的 3 条具体路由 10.0.0.0/24、10.0.1.0/24 和 10.0.2.0/24 聚合成 1 条路由向邻居发送，邻居收到后，路由表中就只有 1 条路由表项，从而减小了路由表的规模。

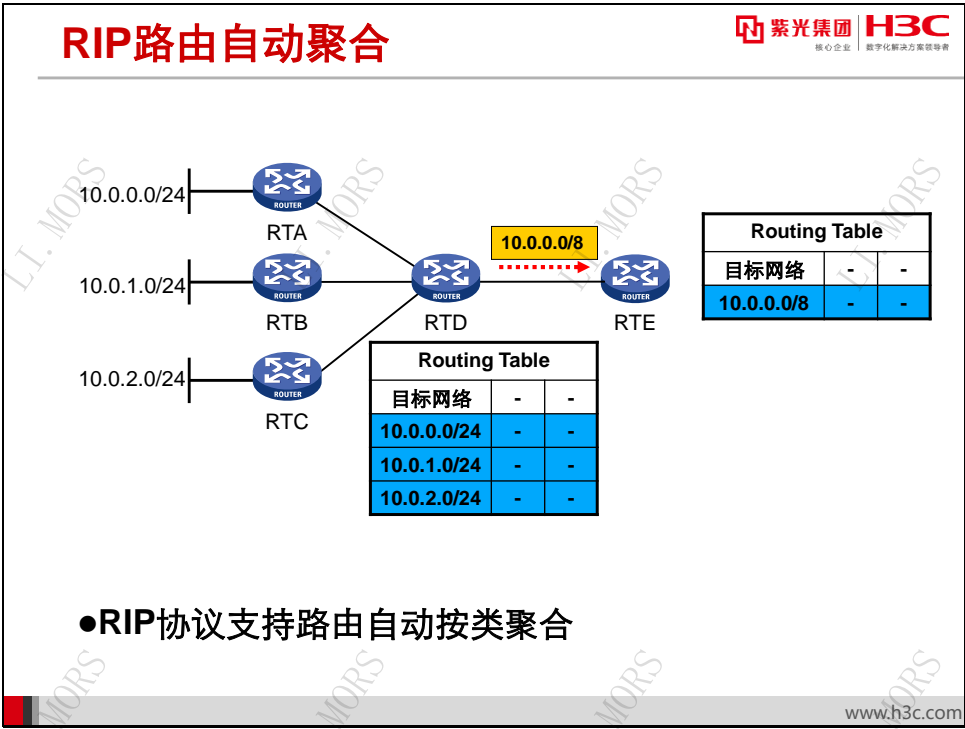


在路由器上不需要运行动态路由协议时，可以配置静态路由聚合。

如上图所示，RTD 上有 3 条具体路由 10.0.0.0/24、10.0.1.0/24 和 10.0.2.0/24。RTE 上没有运行动态路由，所以没有学习到这 3 条路由。但是 RTE 上可以配置 1 条静态路由，其目的地址为 10.0.0.0/22，也就是包含了这 3 条路由，下一跳指向 RTD。

事实上，可以认为缺省静态路由 0.0.0.0/0 就是把所有的路由都包含的聚合路由。

6.3 RIP协议中的聚合



RIP 协议支持路由自动按类聚合。路由器运行 RIP-1 或 RIP-2 后，会自动将子网路由聚合成自然掩码的路由向外发送。

上图中，如果启用 RIP-1，RTD 会只发送不带掩码的路由项 10.0.0.0 给 RTE；如果启用 RIP-2，则 RTD 会将 3 条具体路由 10.0.0.0/24、10.0.1.0/24 和 10.0.2.0/24 聚合成自然掩码的 10.0.0.0/8 向邻居 RTE 发送。

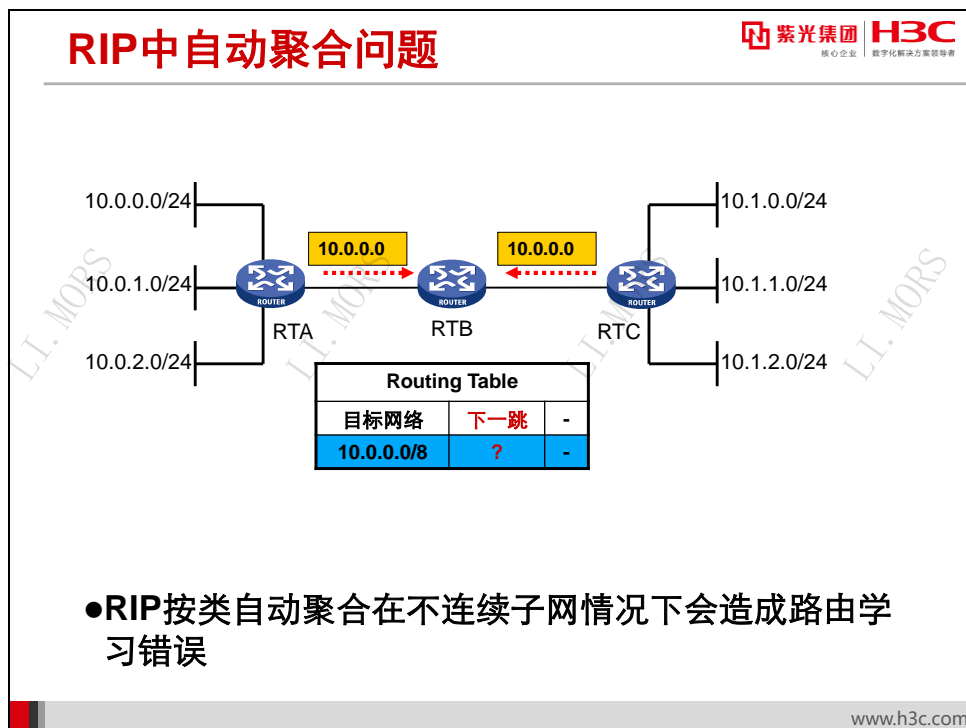
缺省情况下，RIP-1 和 RIP-2 的自动路由聚合功能是开启的。可以在 RIP 视图下配置关闭 RIP-2 自动路由聚合功能，命令如下：

undo summary

关闭自动聚合后，RIP-2 将发送具体路由到邻居路由器。

说明：

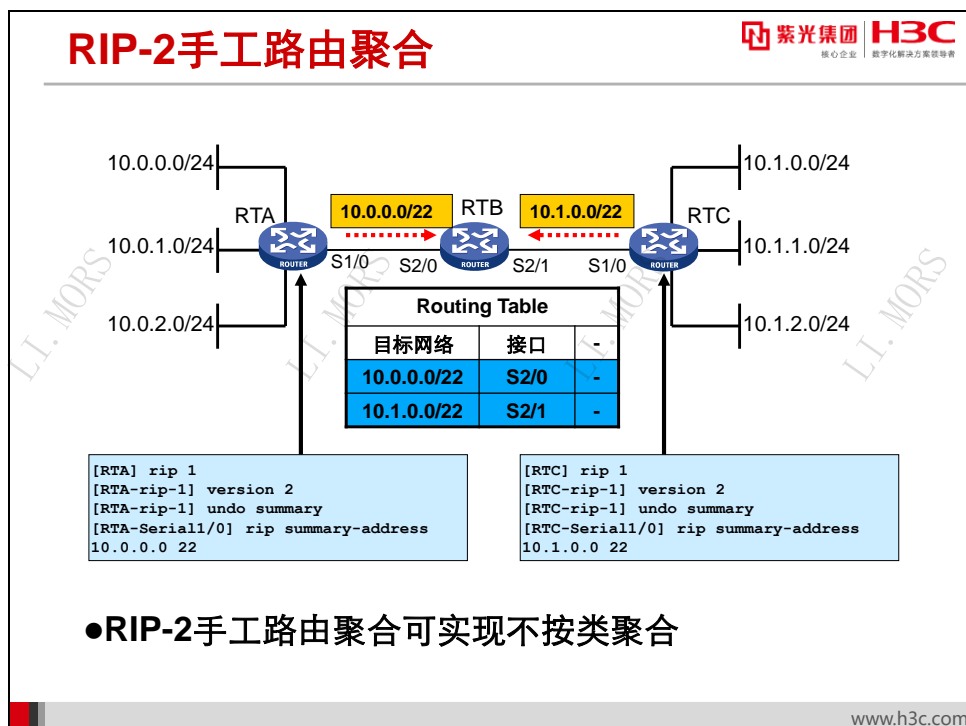
RIP-1 无法支持 VLSM，所以无法关闭自动聚合功能。



在不连续子网情况下，RIP 自动路由聚合功能会导致路由学习错误。

上图中，所有路由器运行 RIP-1。RTA 将发送不带掩码的路由项 10.0.0.0 给 RTB；同理，RTC 与 RTB 会发送不带掩码的路由项 10.0.0.0 给 RTB。这样，RTB 会从 2 个接口接收到同一目的地的路由信息，于是提示错误，无法正确学习路由。

要解决以上问题，需要网络设计者在规划阶段予以合理规划，避免在网络中出现不连续子网的情况。或者，可以使用 RIP-2 协议并关闭自动路由聚合功能。



RIP-2 协议能够支持手工路由聚合，以实现不按类，任意子网掩码的聚合。

可以在接口视图下配置手工路由聚合，命令如下：

rip summary-address *ip-address* { *mask* | *mask-length* }

如上图所示，在 RTA 及 RTC 上配置手工路由聚合，将聚合后的路由发布给 RTB。

RTA 配置如下：

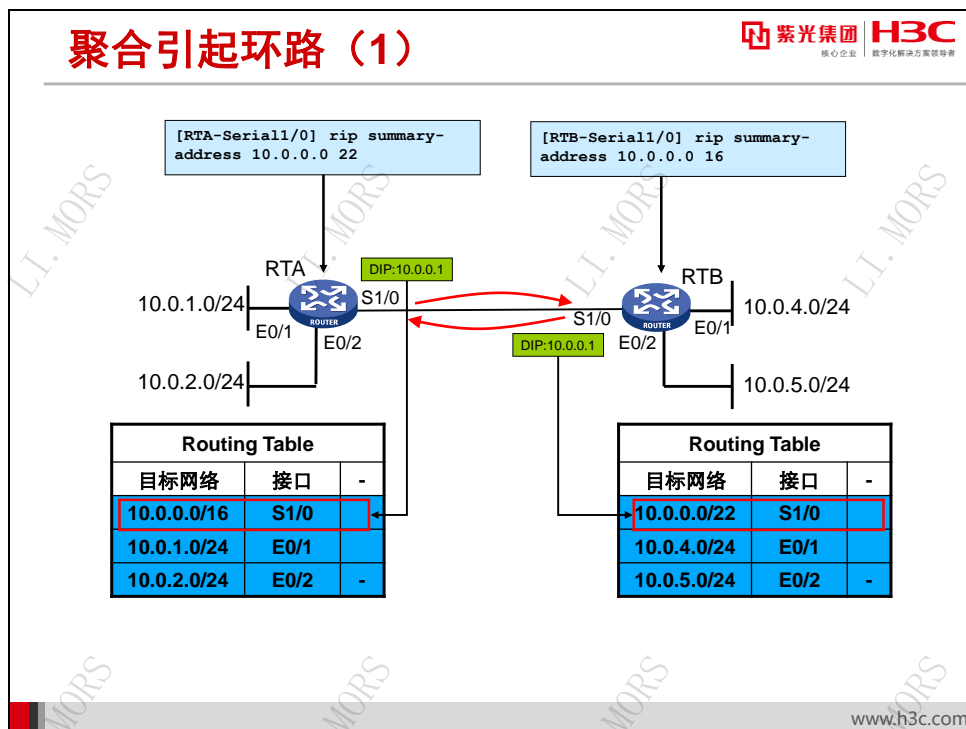
```
[RTA]rip 1
[RTA-rip-1]version 2
[RTA-rip-1]undo summary
[RTA-Serial1/0]rip summary-address 10.0.0.0 22
```

RTB 配置如下：

```
[RTC]rip 1
[RTC-rip-1]version 2
[RTC-rip-1]undo summary
[RTC-Serial1/0]rip summary-address 10.1.0.0 22
```

配置完成后，RTB 路由表中出现了聚合后的路由 10.0.0.0/22 和 10.1.0.0/22。

6.4 路由聚合环路产生与避免



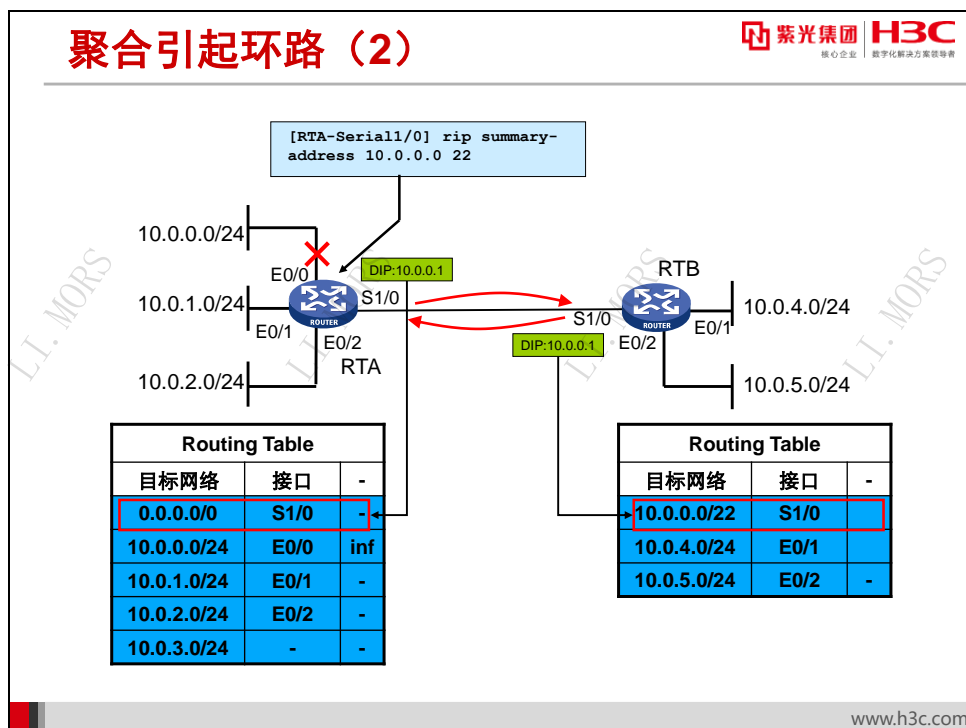
路由聚合相当于将原有的路由信息转换成新的路由信息，原有的信息丢失了。这样在某些情况下会产生路由环路。

如上图所示。RTA 连接有子网 10.0.1.0/24 和 10.0.2.0/24，配置路由聚合后，向 RTB 发送路由 10.0.0.0/22；RTB 连接有子网 10.0.4.0/24 和 10.0.5.0/24，配置路由聚合后，向 RTA 发送路由 10.0.0.0/16。这样，在 RTA 路由表中有路由表项 10.0.0.0/16，下一跳指向 RTB；RTB 路由表中有路由表项 10.0.0.0/22，下一跳指向 RTA。

如果此时 RTA 接收到目的地址为 10.0.0.1 的 IP 报文，经过查表，匹配了表项 10.0.0.0/16，转发到 RTB；RTB 经过查路由表，发现匹配表项 10.0.0.0/22，于是又转发回 RTA。路由环路就形成了。

造成上述路由环路的原因是聚合配置不当。在 RTA 和 RTB 上所配置的聚合路由，并没有完全包含相应具体路由。目的地址为 10.0.0.1 的 IP 报文不能够匹配路由表中的具体路由，而只能匹配聚合后路由，而聚合后路由是从对方学习到的，所以就互相转发给对方，造成环路。

所以在配置聚合路由时，要尽量使发布的聚合后路由恰好包含聚合前所有具体路由，以避免可能产生的环路。

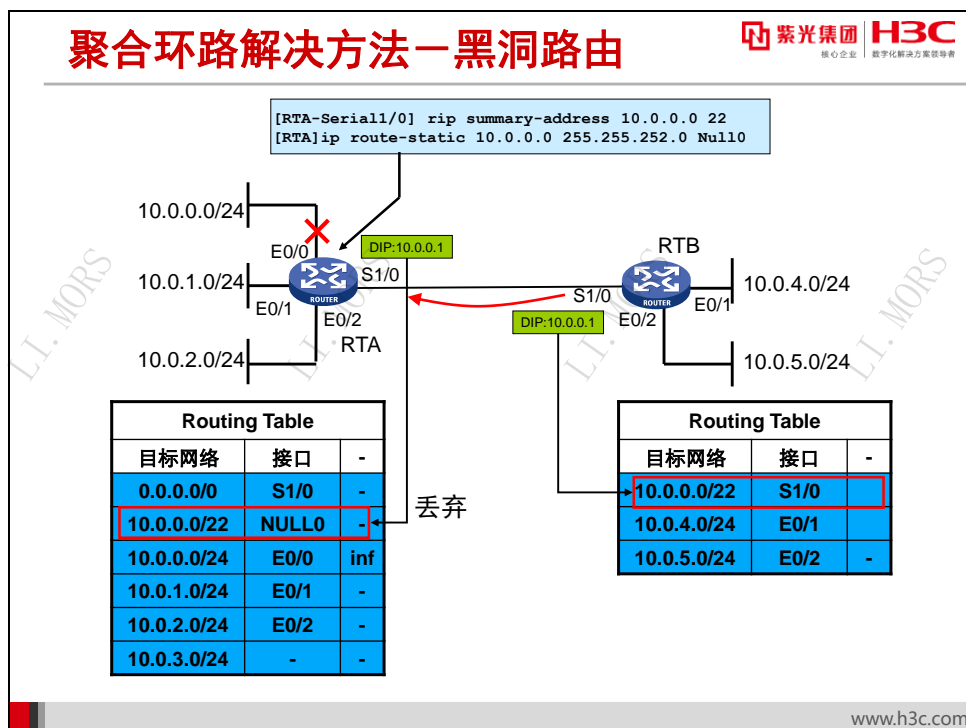


在某些情况下，尽管聚合路由已经完全包含了聚合前所有具体路由，但仍然会产生环路。

如上图所示，RTA 将 4 条直连路由 10.0.0.0/24、10.0.1.0/24、10.0.2.0/24 和 10.0.3.0/24 聚合成 1 条路由 10.0.0.0/22 发送给 RTB。同时，RTA 上配置了缺省静态路由 0.0.0.0/0，下一跳指向 RTB。

正常情况下，报文转发正常。但如果直连路由 10.0.0.0/24 的链路产生故障，路由器会将此路由置为失效状态。此时如果路由器接收目的地址是 10.0.0.1 的报文，经过查表操作，发现只能匹配缺省路由，于是将报文而转发给 RTB；而 RTB 经过查表，发现只能匹配从 RTA 学来的聚合后路由 10.0.0.0/22，于是将报文返回到 RTA，路由环路就产生了。

配置黑洞路由是解决聚合路由环路的较好方法。



如上图所示，在路由器 RTA 上配置 1 条黑洞路由 10.0.0.0/22，下一跳指向 Null0 接口。

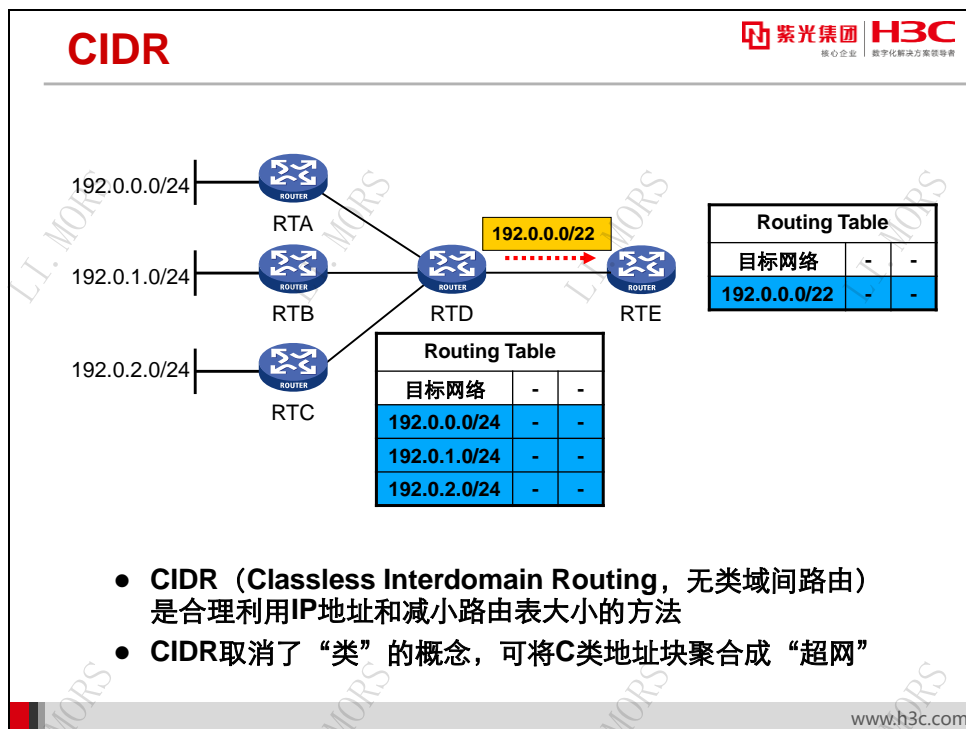
```
[RTA]ip route-static 10.0.0.0 255.255.252.0 Null0
```

Null0 接口是系统中一个虚拟的特殊接口，如果报文被转发到 Null0 接口，意味着报文实际上被丢弃。

配置完成后，在路由 10.0.0.0/24 失效的情况下，去往 10.0.0.1 的报文会匹配黑洞路由而被 RTA 丢弃。而如果路由 10.0.0.0/24 恢复正常，则报文会匹配具体路由 10.0.0.0/24（因其有较长的掩码）而转发到正确的目的地。

在配置黑洞路由时，注意所配置的网段大小要与聚合路由一致，这样才能完全防止聚合环路。

6.5 CIDR



CIDR (Classless Interdomain Routing, 无类域间路由) 是合理利用 IP 地址和减小路由表的方法。

在 Internet 发展早期, 所有地址块都是按类分配给 ISP 和各类组织。随着 Internet 用户快速发展, 这种地址分配方法的弊端显示出来。1 个 C 类地址块有 256 个 IP 地址, 1 个 B 类地址块有 65536 个 IP 地址, 而 1 个 A 类地址块有 16777216 个 IP 地址! 而按类分配意味着用户只能分配到整个地址块, 无法按照实际地址需求来分配, 从而导致很大一部分 IP 地址浪费。另外, Internet 骨干路由器上路由表的规模也在急剧增大。

CIDR 的出现解决了上述问题。通过应用 CIDR, IP 地址不再有“类”的限制。大的地址块如 A 类地址块可以被分割成多个小的地址块, 称为子网; 而小的多个地址块如 C 类地址块也可以被聚合成 1 个大的地址块, 称为超网。

在上图所示的网络中, RTD 把从 RTA、RTB 和 RTC 学来的 3 条 C 类网段路由聚合成 1 条 192.0.0.0/22 路由向 RTE 发布。这样可以减小 RTE 上路由表的规模。

在下一代 Internet 协议 IPv6 中, 地址已经完全取消了类的概念, 用前缀长度来指明地址属于哪一个网络。

在使用 CIDR 进行网络规划时, 要注意地址块要连续, 分层次, 以利于路由聚合。

6.6 本章总结

本章总结

- 使用聚合路由可减小路由表大小
- RIP协议可支持自动聚合和手工聚合
- 配置黑洞路由是避免聚合环路的方法
- CIDR取消了IP地址中“类”的概念，有利于合理利用地址和聚合路由

6.7 习题和解答

6.7.1 习题

1. 路由聚合的优点是 ()
A. 减小路由表规模, 加快路由匹配速度 B. 降低路由更新流量
C. 降低管理员对路由协议的配置工作 D. 减少路由环路产生
2. 以下哪些路由协议支持自动按类聚合功能? ()
A. RIP-1 B. RIP-2 C. OSPF D. BGP
3. 以下哪些路由协议支持手工聚合功能? ()
A. RIP-1 B. RIP-2 C. OSPF D. BGP
4. 在路由器上配置 RIP 手工聚合的命令是 ()
A. [RTA] rip summary-address 10.0.0.0
B. [RTA] rip summary-address 10.0.0.0 22
C. [RTA-Serial1/0] rip summary-address 10.0.0.0
D. [RTA-Serial1/0] rip summary-address 10.0.0.0 22
5. 关于 CIDR, 下列说法中正确的是 ()
A. CIDR 取消了类的概念
B. 使用 CIDR 进行 IP 地址规划, 可以合理利用 IP 地址
C. 使用 CIDR 进行 IP 地址规划, 有利于路由聚合
D. 大的 IP 地址块可以被划分成多个小的 IP 地址块, 称为子网划分

6.7.2 习题答案

1. AB 2. AB 3. BCD 4. D 5. ABCD