

1. IP 路由再认识

1.1 路由基础.....	2
1.2 路由器及路由基本原理.....	2
1.3 静态路由与动态路由.....	3
1.3.1 静态路由与动态路由的区别.....	3
1.3.2 动态路由的分类.....	3
1.4 路由表和 FIB 表.....	4
1.4.1 路由表	4
1.4.2 路由表中的内容.....	4
1.4.3 路由超限自动恢复.....	6
1.4.4 FIB 表的匹配.....	7
1.5 路由迭代.....	8
1.6 路由协议的优先级.....	9
1.7 路由的度量.....	11
1.8 负载分担与路由备份.....	12
1.8.1 负载分担	12
1.8.2 路由备份	13
1.9 IP FRR	14
1.9.1 IP FRR 的分类与实现.....	14
1.9.2 IP FRR 的典型应用.....	14
1.10 路由收敛.....	15
1.11 缺省路由.....	16
1.12 不同路由协议的互相引入.....	17
1.13 自治系统.....	17
1.14 下一跳分离.....	18

1.1 路由基础

路由是数据通信网络中最基本的要素。**路由信息就是指导报文发送的路径信息，路由的过程就是报文转发的过程。**

根据**路由目的地**的不同，路由可划分为：

- **网段路由**：目的地为网段，IPv4 地址子网掩码长度小于 32 位或 IPv6 地址前缀长度小于 128 位。
- **主机路由**：目的地为主机，IPv4 地址子网掩码长度为 32 位或 IPv6 地址前缀长度为 128 位。

根据**目的地与该路由器是否直接相连**，路由又可划分为：

- **直连路由**：目的地所在网络与路由器直接相连。
- **间接路由**：目的地所在网络与路由器非直接相连。

根据**目的地址类型**的不同，路由还可以分为：

- **单播路由**：表示将报文转发的目的地址是一个单播地址。
- **组播路由**：表示将报文转发的目的地址是一个组播地址。

1.2 路由器及路由基本原理

在因特网中，网络连接设备用来控制网络流量和保证网络数据传输质量。常见的网络连接设备有**集线器（Hub）、网桥（Bridge）、交换机（Switch）和路由器（Router）**。

路由器是一种典型的网络连接设备，用来进行**路由选择和报文转发**。路由器根据收到报文的地址选择一条合适的路径（包含一个或多个路由器的网络），然后将报文传送到**下一个路由器**，**路径终端的路由器负责将报文送交目的主机**。

路由就是报文从源端到目的端的路径。当报文从路由器到目的网段有多条路由可达时，路由器可以根据路由表中最佳路由进行转发。最佳路由的选取与发现此路由的路由协议的优先级、路由的度量有关。当多条路由的协议优先级与路由度量都相同时，可以实现负载分担，缓解网络压力；当多条路由的协议优先级与路由度量不同时，可以构成路由备份，提高网络的可靠性。

1.3 静态路由与动态路由

路由器不仅支持静态路由，同时也支持 RIP（Routing Information Protocol）、OSPF（Open Shortest Path First）、IS-IS（Intermedia System-Intermedia System）和 BGP（Border Gateway Protocol）等动态路由协议。

1.3.1 静态路由与动态路由的区别

路由协议是路由器之间维护路由表的规则，用于发现路由，生成路由表，并指导报文转发。依据来源的不同，路由可以分为三类：

- 通过**链路层协议**发现的路由称为**直连路由**。
- 通过网络管理员**手动配置**的路由称为**静态路由**。
- 通过**动态路由协议**发现的路由称为**动态路由**。

静态路由配置方便，对系统要求低，适用于拓扑结构简单并且稳定的小型网络。缺点是不能自动适应网络拓扑的变化，需要人工干预。

动态路由协议有自己的路由算法，能够自动适应网络拓扑的变化，适用于具有一定数量三层设备的网络。缺点是配置对用户要求比较高，对系统的要求高于静态路由，并将占用一定的网络资源和系统资源。

1.3.2 动态路由的分类

对动态路由协议的分类可以采用以下不同标准：

根据作用范围不同，路由协议可分为：

- **内部网关协议** IGP（Interior Gateway Protocol）：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** EGP（Exterior Gateway Protocol）：运行于不同自治系统之间。BGP 是目前最常用的 EGP 协议。

根据使用算法不同，路由协议可分为：

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

以上两种算法的主要区别在于发现路由和计算路由的方法不同。

1.4 路由表和 FIB 表

路由器转发数据包的关键是**路由表和 FIB 表**，每个路由器都至少保存着一张路由表和一张 FIB（Forwarding Information Base）表。**路由器通过路由表选择路由，通过 FIB 表指导报文进行转发。**

1.4.1 路由表

每台路由器中都保存着一张**本地核心路由表**（即设备的**IP 路由表**），同时**各个路由协议也维护着自己的路由表**。

- **本地核心路由表**

路由器使用本地核心路由表用来保存决策优选路由，并负责把**优选路由下发到 FIB 表**，通过 FIB 表指导报文进行转发。这张路由表依据各种路由协议的优先级和度量值来选取路由。

说明：对于支持 L3VPN（Layer 3 Virtual Private Network）的路由器，每一个 VPN-Instance 拥有一个自己的管理路由表（本地核心路由表）。

- **协议路由表**

协议路由表中存放着该协议发现的路由信息。

路由协议可以引入并发布其他协议生成的路由。例如，在路由器上运行 OSPF 协议，需要使用 OSPF 协议通告直连路由、静态路由或者 IS-IS 路由时，要将这些路由引入到 OSPF 协议的路由表中。

1.4.2 路由表中的内容

在路由器中，执行命令 **display ip routing-table** 时，可以查看路由器的路由表概要信息，如下所示：

```
<Huawei> display ip routing-table
```

Proto: Protocol Pre: Preference 后面缩写释义

Route Flags: R - relay, D - download to fib, T - to vpn-instance

路由标记: R: 表示该路由是迭代路由 D: 表示该路由下发到 FIB 表

Routing Table: _public_ 此路由表是公网路由表

Destinations : 14 目的网络/主机的总数

Routes : 14 路由的总数

Destination/Mask	Proto	Pre	Cost	Flags	NextHop	Interface
0.0.0.0/0	Static	60	0	RD	10.137.216.1	Vlanif20
10.10.10.0/24	Direct	0	0	D	10.10.10.10	Vlanif20
10.10.10.10/32	Direct	0	0	D	127.0.0.1	InLoopBack0
10.10.10.255/32	Direct	0	0	D	127.0.0.1	InLoopBack0
10.10.11.0/24	Direct	0	0	D	10.10.11.1	LoopBack0
10.10.11.1/32	Direct	0	0	D	127.0.0.1	InLoopBack0
10.10.11.255/32	Direct	0	0	D	127.0.0.1	InLoopBack0
10.137.216.0/23	Direct	0	0	D	10.137.217.208	Vlanif20
10.137.217.208/32	Direct	0	0	D	127.0.0.1	InLoopBack0
10.137.217.255/32	Direct	0	0	D	127.0.0.1	InLoopBack0
127.0.0.0/8	Direct	0	0	D	127.0.0.1	InLoopBack0
127.0.0.1/32	Direct	0	0	D	127.0.0.1	InLoopBack0
127.255.255.255/32	Direct	0	0	D	127.0.0.1	InLoopBack0
255.255.255.255/32	Direct	0	0	D	127.0.0.1	InLoopBack0

路由表中包含了下列关键项：

Destination：表示此路由的目的地址。用来标识 IP 包的目的地址或目的网络。

Mask：表示此目的地址的子网掩码长度。与目的地址一起来标识目的主机或路由器所在的网段的地址。

将目的地址和子网掩码“逻辑与”后可得到目的主机或路由器所在网段的地址。例如：目的地址为 10.1.1.1，掩码为 255.255.255.0 的主机或路由器所在网段的地址为 10.1.1.0。

掩码由若干个连续“1”构成，既可以用点分十进制表示，也可以用掩码中连续“1”的个数来表示。例如掩码 255.255.255.0 长度为 24，即可以表示为 24。

Proto：表示学习此路由的路由协议。

Pre：表示此路由的路由协议优先级。针对同一目的地，可能存在不同下一跳、出接口等多条路由，这些不同的路由可能是由不同的路由协议发现的，也可以是手工配置的静态路由。优先级高（数值小）者将成为当前的最优路由。各协议路由优先级请参见**路由协议的优先级**。

Cost：路由开销。当到达同一目的地的多条路由具有相同的路由优先级时，路由开销最小的将成为当前的最优路由。

说明：**Preference** 用于不同路由协议间路由优先级的比较，**Cost** 用于同一种路由协议内

部不同路由的优先级的比较。

NextHop：表示此路由的下一跳地址。指明数据转发的下一个设备。

Interface：表示此路由的出接口。指明数据将从本地路由器哪个接口转发出去。

如下图所示的网络中，路由器 A 与三个网络相连，因此有三个 IP 地址和三个出接口，其路由表如图所示。

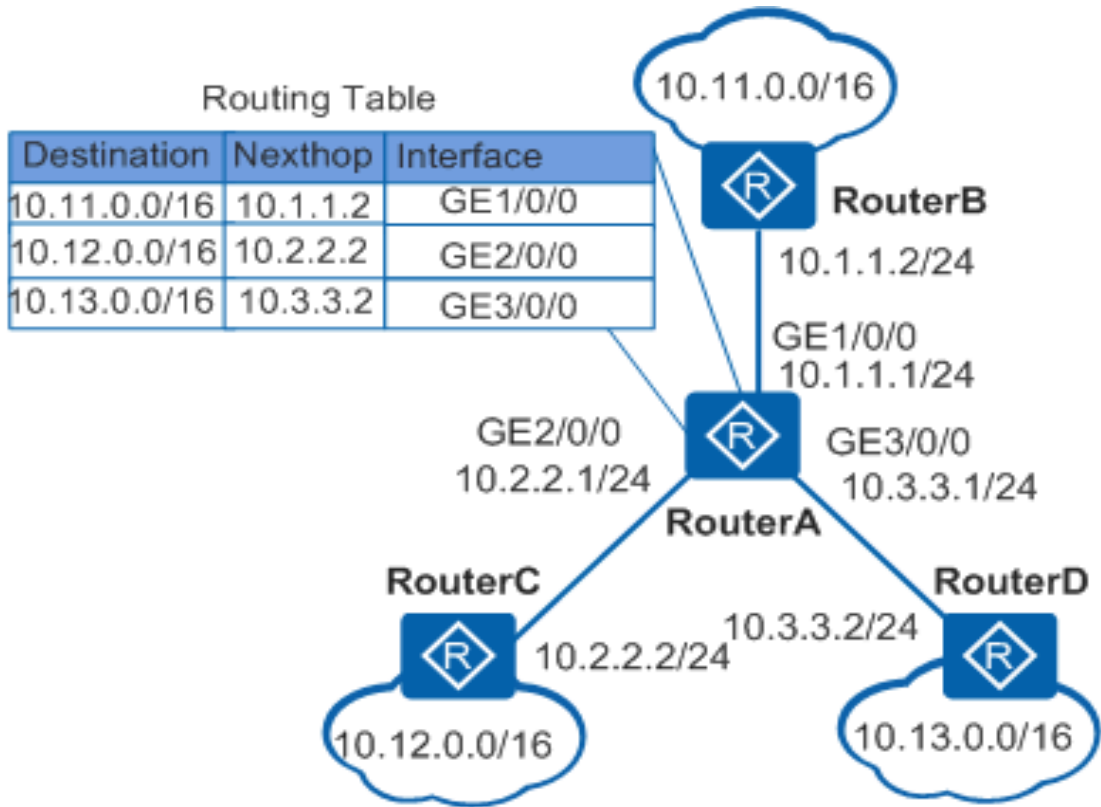


图 1 路由表示意图

1.4.3 路由超限自动恢复

本地核心路由表里保存着各路由协议的路由，如果本地核心路由表里的路由数量达到系统上限，协议路由表将无法向本地核心路由表添加路由。本地核心路由表有以下几种路由限制：

- 整机路由限制：指定所有路由条数的上限值。
- 整机路由前缀限制：指定所有路由的地址前缀范围。
- 组播 IGP 路由限制：指定组播 IGP 路由条数的上限值。
- 多拓扑路由限制：指定多拓扑路由条数的上限值。
- 所有私网路由限制：指定所有私网路由条数的上限值。

- VPN 路由限制：指定 VPN 路由条数的上限值。
- VPN 路由前缀限制：指定 VPN 路由的地址前缀范围。

如果协议由于某种路由限制而向本地核心路由表添加路由失败，系统会记录本次添加路由的协议和对应的路由表 Table ID。

当协议删除本地核心路由表里的路由释放了路由表的空间之后，路由超限解除，系统会通知所有向本地核心路由表添加路由失败的协议，重新向本地核心路由表添加路由，使得本地核心路由表中的路由能够得到最大程度的恢复。是否可以完全恢复，取决于释放的路由表空间的大小。

1.4.4 FIB 表的匹配

在路由表选择出路由后，路由表会将激活路由下发到 FIB 表中。当报文到达路由器时，会通过**查找 FIB 表进行转发**。

FIB 表中每条转发项都指明到达某网段或某主机的报文应通过路由器的哪个物理接口或逻辑接口发送，然后就可到达该路径的下一个路由器，或者不再经过别的路由器而传送到直接相连的网络中的目的主机。

FIB 表的匹配遵循**最长匹配原则**。查找 FIB 表时，报文的目地址和 FIB 中各表项的掩码进行按位“逻辑与”，得到的地址符合 FIB 表项中的网络地址则匹配。最终选择一个最长匹配的 FIB 表项转发报文。

例如，一台路由器上的路由表如下：

Routing Tables:

Destination/Mask	Proto	Pre	Cost	Flags	NextHop	Interface
0.0.0.0/0	Static	60	0	D	192.168.0.2	GigabitEthernet1/0/0
10.8.0.0/16	Static	60	3	D	192.168.0.2	GigabitEthernet1/0/0
10.9.0.0/16	Static	60	50	D	172.16.0.2	GigabitEthernet3/0/0
10.9.1.0/24	Static	60	4	D	192.168.0.2	GigabitEthernet2/0/0
10.20.0.0/16	Direct	0	0	D	172.16.0.1	GigabitEthernet4/0/0

一个目的地址是 10.9.1.2 的报文进入路由器，查找对应的 FIB 表。

FIB Table:

Total number of Routes : 5

Destination/Mask	NextHop	Flag	TimeStamp	Interface
TunnelID				
0.0.0.0/0	192.168.0.2	SU	t[37]	GigabitEthernet1/0/0

0x0

10.8.0.0/16 0x0	192.168.0.2	DU	t[37]	GigabitEthernet1/0/0
10.9.0.0/16 0x0	172.16.0.2	DU	t[9992]	GigabitEthernet3/0/0
10.9.1.0/24 0x0	192.168.0.2	DU	t[9992]	GigabitEthernet2/0/0
10.20.0.0/16 0x0	172.16.0.1	U	t[9992]	GigabitEthernet4/0/0

首先，目的地址 10.9.1.2 与 FIB 表中各表项的掩码“0、16、24”作“逻辑与”运算，得到下面的网段地址：0.0.0.0/0、10.9.0.0/16、10.9.1.0/24。这三个结果可以匹配到 FIB 表中对应的三个表项。最终，路由器会选择**最长匹配 10.9.1.0/24 表项**，从接口 GE2/0/0 转发这条目的地址是 10.9.1.2 的报文。

路由器路由查找注意事项：

- **不同的路由前缀**（注意路由前缀包含网络号+掩码，缺一不可），**在路由表中属于不同的路由**。
- **相同的路由前缀**，通过不同的协议获取，**先比优先级**，优选优先级小的，**后比 cost**。（这是一般情况，当然有例外情况，这就要看特定的环境和特定的路由协议了。）
- 默认采用**最长匹配原则**，匹配，则转发；无匹配，则找默认路由，默认路由都没有，则丢弃。
- **路由器的行为是逐跳的**，到目标网络的沿路径每个路由器都必须有关于目的地的路由。
- **数据是双向的**，考虑流量的时候，要关注流量的**往返**。

1.5 路由迭代

路由必须有直连的下一跳才能够指导转发，但是路由生成时下一跳可能不是直连的，因此需要**计算出一个直连的下一跳和对应的出接口**，这个过程就叫做**路由迭代**。BGP 路由、静态路由和 UNR 路由的下一跳都有可能不是直连的，都需要进行路由迭代。

例如，BGP 路由的下一跳一般是非直连的对端 loopback 地址，不能指导转发，需要进行迭代。即根据以 BGP 学习到的下一跳为目的地址在 IP 路由表中查找，当找到一条具

有直连的下一跳、出接口信息的路由后（一般为一条 IGP 路由），将其下一跳、出接口信息填入这条 BGP 路由的 IP 路由表中并生成对应的 FIB 表项。

对于 BGP 私网路由，需要隧道进行转发，路由的下一跳一般是远端 PE 的 Loopback 地址，不能指导转发，也需要进行路由迭代，即在隧道列表中查找到达该 Loopback 地址的隧道，将该隧道信息填入路由表中并生成对应的 FIB 表项。

1.6 路由协议的优先级

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（取值较小）**的路由协议发现的路由将成为最优路由，并将最优路由放入本地路由表中。

路由器分别定义了**外部优先级**和**内部优先级**。外部优先级是指用户可以手工为各路由协议配置的优先级，缺省情况下如表 1 所示。

表 1 路由协议缺省时的外部优先级

路由协议的类型	路由协议的外部优先级
Direct	0
OSPF	10
IS-IS	15
Static	60
UNR (User Network Route)	- DHCP (Dynamic Host Configuration Protocol): 60 - AAA-Download: 60 - IP Pool: 61 - Frame: 62 - Host: 63 - NAT (Network Address Translation): 64 - IPSec (IP Security): 65 - NHRP (Next Hop Resolution Protocol): 65 - PPPoE (Point-to-Point Protocol over Ethernet): 65
RIP	100
OSPF ASE	150
OSPF NSSA	150
IBGP	255
EBGP	255

说明：其中，0 表示直接连接的路由，255 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，各种路由协议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。

路由协议的内部优先级则不能被用户手工修改，如表 2 所示。

表 2 路由协议内部优先级

路由协议的类型	路由协议的内部优先级
Direct	0
OSPF	10
IS-IS Level-1	15
IS-IS Level-2	18
Static	60
UNR	65
RIP	100
OSPF ASE	150
OSPF NSSA	150
IBGP	200
EBGP	20

选择路由时**先比较路由的外部优先级**，当不同的路由协议配置了相同的优先级后，系统会通过**内部优先级**决定哪个路由协议发现的路由将成为最优路由。例如，到达同一目的地 10.1.1.0/24 有两条路由可供选择，一条静态路由，另一条是 OSPF 路由，且这两条路由的外部优先级都被配置成 5。这时路由器系统将根据表 2 所示的内部优先级进行判断。因为 OSPF 协议的内部优先级是 10，高于静态路由的内部优先级 60。所以系统选择 OSPF 协议发现的路由作为最优路由。

1.7 路由的度量

路由的度量标示出了这条路由到达指定的目的地址的代价，通常以下因素会影响到路由的度量。

- 路径长度

路径长度是最常见的影响路由度量的因素。**链路状态路由协议**可以为每一条链路设置一个**链路开销**来标示此链路的路径长度。在这种情况下，路径长度是指经过的所有链路的链路开销的**总和**。**距离矢量路由协议**使用**跳数**来标示路径长度。跳数是指数据从源端到目的端所经过的**设备数量**。例如，路由器到与它直接相连网络的跳数为 0，通过一台路由器可达的网络的跳数为 1，其余以此类推。

- **网络带宽**

网络带宽是一个链路实际的传输能力。例如，一个 10 千兆的链路要比 1 千兆的链路更优越。虽然带宽是指一个链路能达到的最大传输速率，但这不能说明在高带宽链路上路由要比低带宽链路上**更优越**。比如说，一个高带宽的链路正处于拥塞的状态下，那报文在这条链路上转发时将会花费更多的时间。

- **负载**

负载是一个网络资源的使用程度。计算负载方法包括 CPU 的利用率和它每秒处理数据包的数量。持续监测这些参数可以及时了解网络的使用情况。

- **通信开销**

通信开销衡量了一条链路的运营成本。尤其是只注重运营成本而不在乎网络性能的时候，通信开销则就成了一个重要的指标。

1.8 负载分担与路由备份

当多条路由的**路由优先级和路由度量都相同**时，这几条路由就称为**等价路由**，多条等价路由可以实现负载分担。当这几条路由为**非等价路由**时，就可以实现**路由备份**。

1.8.1 负载分担

路由器支持多路由模式，即允许配置多条目的地相同且优先级也相同的路由。当到达同一目的地存在同一路由协议发现的多条路由时，且这几条路由的开销值也相同，那么就满足负载分担的条件。

当实现负载分担时，路由器根据**五元组**（源地址、目的地址、源端口、目的端口、协议）进行转发，当五元组相同时，路由器总是选择与上一次相同的下一跳地址发送报文。当五元组不同时，路由器**会选取相对空闲的路径进行转发**。

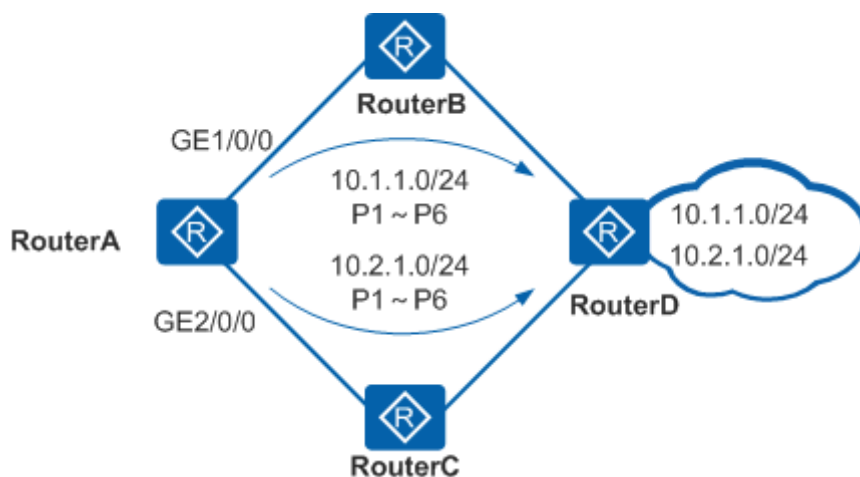


图 2 负载分担组网图

如图 2 所示，RouterA 已经通过接口 GE1/0/0 转发到目的地址 10.1.1.0/24 的第 1 个报文 P1，随后又需要分别转发报文到目的地址 10.1.1.0/24 和 10.2.1.0/24。其转发过程如下：

- 当转发到达 10.1.1.0/24 的第 2 个报文 P2 时，发现此报文与到达 10.1.1.0/24 的第 1 个报文 P1 的五元组一致。所以之后到达该目的地的报文都从 GE1/0/0 转发。
- 当转发到达 10.2.1.0/24 的第 1 个报文 P1 时，发现此报文与到达 10.1.1.0/24 的第 1 个报文 P1 的五元组不一致。所以选取从 GE2/0/0 转发，并且之后到达该目的地的报文都从 GE2/0/0 转发。

说明：系统所允许的负载分担的具体路由条数，与实际产品型号相关。

1.8.2 路由备份

路由备份功能，可以提高网络的可靠性。用户可以根据实际情况，配置到同一目的地的多条路由，其中一条路由的**优先级最高，作为主路由**，其余的路由**优先级较低，作为备份路由**。

正常情况下，路由器采用主路由转发数据。当主链路出现故障时，主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路由转发数据**。这样，也就实现了主路由到备份路由的切换。**当主链路恢复正常时，由于主路由的优先级最高，路由器重新选择主路由来发送数据。这样，就实现了从备份路由回切到主路由。**

1.9 IP FRR

快速重路由 FRR（Fast Reroute）是指当物理层或链路层检测到故障时将故障消息上报至上层路由系统，并立即启用一条备份链路转发报文。**IP FRR** 是一种快速实现路由备份的方式。

在传统的 IP 网络上，转发链路出现底层故障后，最为直观的表现是路由器上的物理接口状态变为 **Down** 状态。路由器检测到这种故障后，会通知上层路由系统进行相应更新，并重新计算路由。通常从链路故障发生到路由系统完成路由收敛（重新选择了一条可用的路由），要经历几秒钟的时间。

但是对于网络上某些对延时、丢包等非常敏感的业务来说，秒级的收敛时间是不能忍受的，因为这将导致当前业务的中断。比如 **VoIP 业务** 所能容忍的网络中断时间为**毫秒级**。

IP FRR 特性能够保证转发系统快速应对出现的链路故障，直接启用备份路由进行数据转发，尽快让业务恢复正常。

1.9.1 IP FRR 的分类与实现

IP FRR 针对 IP 网络路由而设计，分为公网 **IP FRR** 和私网 **IP FRR**：

- **公网 IP FRR**：用于保护公网路由器。
- **私网 IP FRR**：用于保护 CE（Customer Edge）。

IP FRR 的主要实现手段如下：

- 1、在主链路可用时，通过设置 **IP FRR** 策略，把备份路由的转发信息同时提供给转发引擎。
- 2、当转发引擎感知到主链路不可用时，能够在**控制平面路由**收敛前直接使用备份路径转发报文。

1.9.2 IP FRR 的典型应用

如图 3 所示，网络中通过部署 **IP FRR** 来增强可靠性。其中 **CE1** 双归到 **PE1** 和 **PE2**，并配置私网备份出接口和备份下一跳，使链路 **B** 为链路 **A** 的备份，链路 **A** 出现故障时可以快速切换到链路 **B** 上。

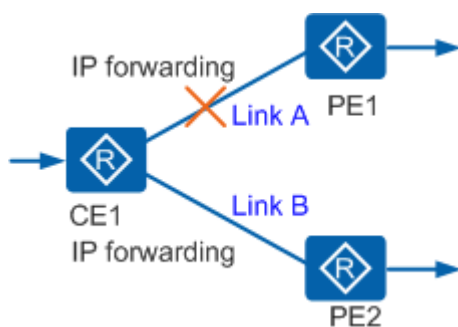


图 1 配置 IP FRR 功能

1.10 路由收敛

路由收敛是指网络拓扑变化引起的通过重新计算路由而发现替代路由的行为。随着网络的融合，区分服务的需求越来越强烈。某些路由可能指导关键业务的转发，如 VoIP，视频会议、组播等，这些关键的业务路由需要尽快收敛，而非关键路由可以相对慢一点收敛。因此，系统需要对不同路由按不同的收敛优先级处理，来提高网络可靠性。

按优先级收敛是指系统为路由设置不同的收敛优先级，分为 **critical**、**high**、**medium**、**low** 四种。系统根据这些路由的收敛优先级采用相对的优先收敛原则，即按照一定的调度比例进行路由收敛安装，指导业务的优先转发。

路由协议优先计算并下发高收敛优先级的路由给系统。缺省情况下的公网路由收敛优先级如表 3 所示。

表 3 缺省时的公网路由收敛优先级

路由协议或路由种类	收敛优先级
Direct	high
Static	medium
OSPF 和 IS-IS 的 32 位主机路由	medium
OSPF（除 32 位主机路由外）	low
IS-IS（除 32 位主机路由外）	low
RIP	low
BGP	low

说明：对于私网路由，除了 OSPF 和 IS-IS 的 **32 位主机路由**标识为 **medium**，其余路由统一标识为 **low**。

如图 4 所示，在图中的组播服务中，网络上运行 OSPF 和 IS-IS 协议，接收者在 RouterA 端，组播源服务器 10.10.10.10/32 在 RouterB 端。其中要求到组播服务器的路由优先于其他路由（例如 10.12.10.0/24）收敛。这时可以配置路由 10.10.10.10/32 的收敛优先级高于路由 10.12.10.0/24 的收敛优先级，这样当网络路由重新收敛时，就能确保到组播源的路由 10.10.10.10/32 优先收敛，保证组播业务的转发。

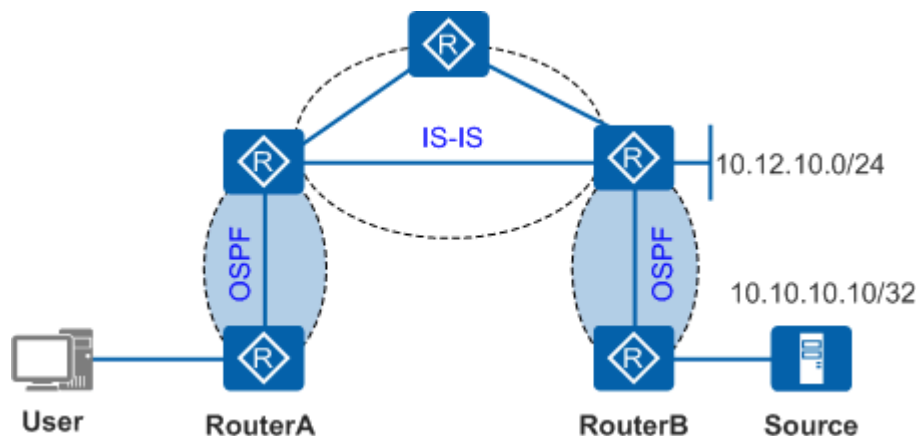


图 4 路由按优先级收敛应用组网图

1.11 缺省路由

缺省路由是另外一种特殊的路由。简单来说，缺省路由是没有在路由表中找到匹配的
路由表项时才使用的路由。如果报文的目的地址不能与路由表的任何目的地址相匹配，那
么该报文将选取缺省路由进行转发。如果没有缺省路由且报文的目的地址不在路由表中，
那么该报文将被**丢弃**，并向源端返回一个 ICMP（Internet Control Message Protocol）报
文，**报告该目的地址或网络不可达**。

在路由表中，缺省路由以到网络 0.0.0.0（掩码也为 0.0.0.0）的路由形式出现。可通
过命令 **display ip routing-table** 查看当前是否设置了缺省路由。通常情况下，管理员可以
通过手工方式配置缺省静态路由；但有些时候，**也可以使动态路由协议生成缺省路由**，如
OSPF 和 IS-IS。

1.12 不同路由协议的互相引入

由于采用的算法不同，不同的路由协议可以发现不同的路由。当网络规模比较大，使用多种路由协议时，不同的路由协议间通常需要发布其他路由协议发现的路由。

各路由协议都可以引入其它路由协议的路由、直连路由和静态路由，每种路由协议都有相应的路由引入机制，具体内容请参见各路由协议模块有关引入外部路由的描述。

1.13 自治系统

在互联网中，自治系统 AS（Autonomous System）是指在一个（有时是多个）实体管辖下的所有 IP 网络和路由器的网络，它们对互联网执行共同的路由策略。

每一个 AS 可以支持多个内部网关路由协议。一个 AS 内的所有网络都被分配同一个 AS 号，属于一个行政单位管辖。AS 号分为 2 字节 AS 号和 4 字节 AS 号。其中 2 字节 AS 号的范围为 1 至 65535。随着时间推进，可分配的 2 字节 AS 号已经濒临枯竭，需要将 AS 号的范围从之前的 2 字节扩展为 4 字节，其中 4 字节 AS 号的取值范围为 1 至 4294967295。4 字节 AS 号还可以用 X.Y 的形式表示，其中 X 的取值范围为 1 至 65535，Y 的取值范围为 0 至 65535。

AS 号根据在不同的网络上使用可以分为两类，如表 4 所示：

表 4 AS 号的分类及其取值范围

AS 号的类别	2 字节 AS 号	4 字节 AS 号
公有 AS 号	1~64511	1~64511,65536~4294967295
私有 AS 号	64512~65535	64512~65535

1.14 下一跳分离

下一跳分离是一种提高路由收敛性能的技术，它把路由前缀与下一跳转发信息的直接关系变为间接关系。下一跳分离支持下一跳信息的单独刷新，而不必逐条刷新大量前缀，从而获得较快的收敛速度。

在路由**需要迭代的场景中**，当 IGP 路由或隧道发生切换时，快速刷新 FIB 转发路径，实现流量的快速收敛，降低对业务的影响。

下一跳分离的基础是路由前缀与下一跳的映射法则。为了满足不同场景下迭代路由和迭代隧道的需求，下一跳信息由**地址族、原始下一跳地址、隧道策略等要素**构成。系统给每个下一跳信息分配一个索引，进行迭代并将迭代结果通知路由协议并下发 FIB。

按需迭代是当依赖路由变化时，只对相关的下一跳进行重新迭代。如果路由的目的地址是某个下一跳信息的原始下一跳地址，或是其网段地址，则路由变化时会影响这个下一跳信息的迭代结果，否则路由的变化对下一跳没有影响。所以，当一条路由发生变化时，通过判断目的地址可以实现只对相关的下一跳进行重新迭代。

对于迭代隧道的情况，当一条隧道的状态（UP/DOWN）改变时，只需要对原始下一跳地址与隧道目的地址相同的下一跳信息进行重新迭代。

迭代策略用于控制下一跳的迭代结果，以满足不同应用场景下的需求。迭代路由时不需要通过策略进行控制，只遵循最长匹配原则。只有私网路由迭代隧道时需要应用隧道策略。

缺省情况下，系统为 VPN 选择 LSP 隧道且不进行负载分担。如果要进行负载分担或选择其它类型的隧道，需要配置隧道策略并应用该隧道策略。VPN 应用隧道策略后，对其下一跳进行迭代时，选择隧道策略中绑定的隧道，或根据隧道策略中为不同类型的隧道指定的优先级，选择使用的隧道。

从转发层面看，对于公网路由，只需要下一跳和出接口就能指导转发，**对于私网路由，还需要指定公网隧道**，在采用下一跳分离下刷之前，转发信息（下一跳、出接口、隧

道) 需要由前缀直接带到 FIB，导致总的收敛时间与前缀数量相关，采用下一跳分离下刷后，大量的前缀与同一个下一跳对应，只需要通过这个下一跳将转发信息带到 FIB 中，相关前缀的流量就能同时切换。

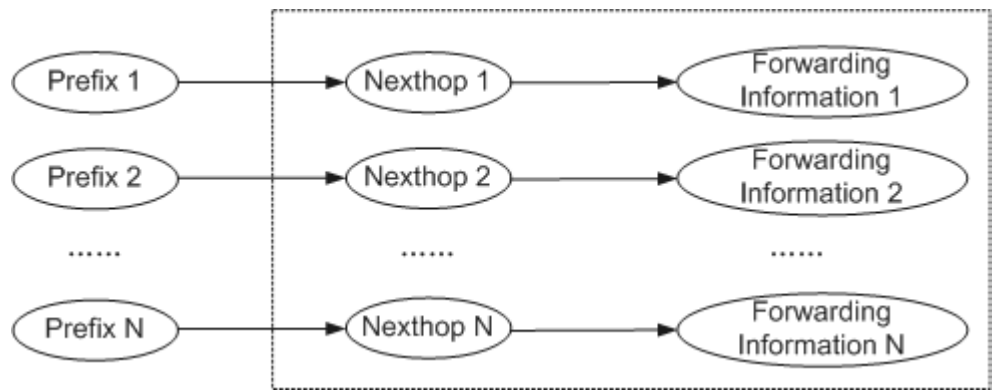


图 5 没有实现下一跳分离时路由的组织形式

如图 5 所示，没有实现下一跳分离时，前缀之间完全独立，分别对应自己的下一跳和迭代转发信息。当依赖路径变化时，分别对每个前缀对应的下一跳进行迭代，并以前缀为单位下刷，收敛速度与前缀数量相关。

实际上，如果前缀来自同一个 BGP 邻居，则它们的下一跳和迭代转发信息也一定相同，下刷的迭代转发信息是重复的。

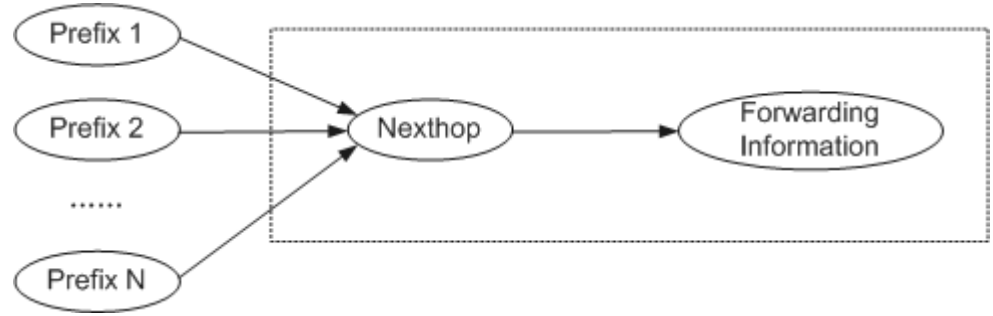


图 6 实现下一跳分离时路由的组织形式

如图 6 所示，实现下一跳分离后，来自同一个邻居的前缀共享同一个下一跳，依赖路径变化时，只对这个共同的下一跳进行迭代，并以下一跳为单位刷新其对应的迭代转发信息。这样所有前缀的流量就可以同时收敛，收敛速度与前缀数量无关。

迭代路由与迭代隧道的不同如下表所示。

表 5 迭代路由与迭代隧道的比较

迭代类型	特点
迭代路由	• 对 BGP 公网路由进行迭代。 • 由路由变化触发迭代。 • 支持按策略迭代下一跳。
迭代隧道	• 对 BGP 私网路由进行迭代。 • 由隧道或隧道策略变化触发迭代。 • 可通过隧道策略对迭代行为进行控制，满足不同应用场景下的需求。