

修订记录

本页不打印

课程编码	适用产品	产品版本	课程版本

作者/工号	时间	审核人/工号	新开发/优化
陈俭元/cwx652820	2020/3/18		





BFD协议原理与配置



前言

- 随着网络应用的广泛部署，网络发生故障极大可能导致业务异常。为了减小链路、设备故障对业务的影响，提高网络的可靠性，网络设备需要尽快检测到与相邻设备间的通信故障，以便及时采取措施，保证业务正常进行。
- BFD（Bidirectional Forwarding Detection，双向转发检测）提供了一个通用的、标准化的、介质无关和协议无关的快速故障检测机制，用于快速检测、监控网络中链路或者IP路由的转发连通状态。
- 本章节主要介绍BFD工作原理以及常见的应用场景。



目标

- 学完本课程后，您将能够：
 - 阐明BFD的作用
 - 描述BFD工作原理
 - 实现BFD配置与应用



目录

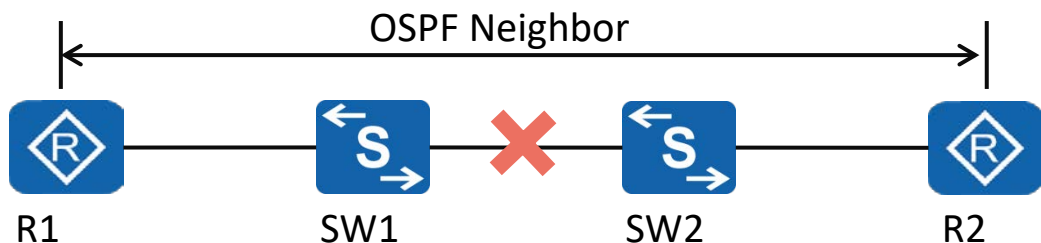
- 1. BFD概述**
2. BFD工作原理
3. BFD应用场景
4. BFD基本配置



网络故障检测遇到的问题

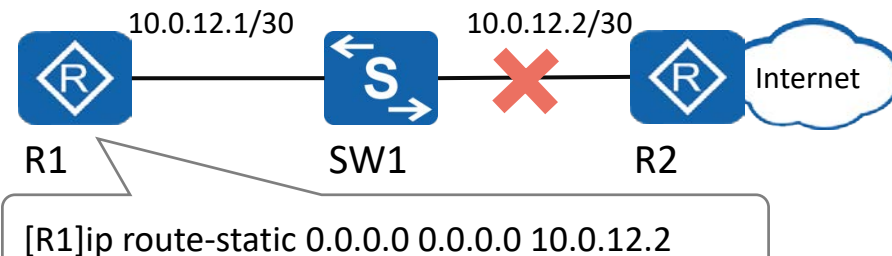
- 在无法通过硬件信号检测故障的系统中，应用通常采用上层协议本身的Hello报文机制检测网络故障。
- 常用路由协议的Hello报文机制检测时间较长，检测时间超过1秒钟。当应用在网络中传输的数据超过GB/s时，秒级的检测时间将会导致应用传输的数据大量丢失。
- 在三层网络中，静态路由本身没有故障检查机制。

动态路由故障检测问题



R1与R2通过二层交换机连接建立OSPF邻接关系，当SW1与SW2之间的链路出现故障时，R1与R2能快速感知吗？

静态路由故障检测问题

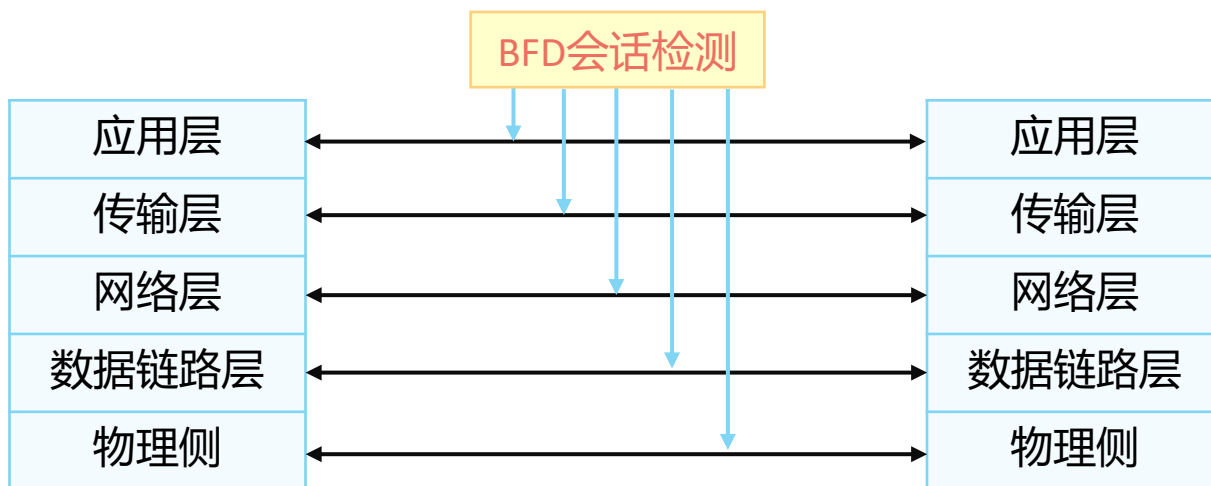


在R1路由器上配置到网关R2的默认路由，当SW1与R2之间的链路出现故障时R1能快速感知吗？



BFD概述

- BFD提供了一个通用的、标准化的、介质无关的、协议无关的快速故障检测机制，有以下两大优点：
 - 对相邻转发引擎之间的通道提供轻负荷、快速故障检测。
 - 用单一的机制对任何介质、任何协议层进行实时检测。
- BFD是一个简单的“Hello”协议。两个系统之间建立BFD会话通道，并周期性发送BFD检测报文，如果某个系统在规定的时间内没有收到对端的检测报文，则认为该通道的某个部分发生了故障。





目录

1. BFD产生背景
- 2. BFD工作原理**
3. BFD应用场景
4. BFD基本配置



BFD报文结构

- BFD检测是通过维护在两个系统之间建立的BFD会话来实现的，系统通过发送BFD报文建立会话。
- BFD控制报文根据场景不同封装不同，报文结构由强制部分和可选的认证字段组成。

强制部分

Ver	Diag	Sta	P	F	C	A	D	M	Detect Mult	Length
My discriminator										
Your discriminator										
Desired Min TX Interval										
Required Min RX Interval										
Required Min Echo RX Interval										

可选部分（认证字段）

Auth-Type	Auth-Len	Authentication Data
-----------	----------	---------------------

1. Sta：BFD本地状态。
2. Detect Mult：检测超时倍数，用于检测方计算检测超时时间。
3. My Discriminator：BFD会话连接本地标识符（Local Discriminator）。发送系统产生的一个唯一的、非0鉴别值，用来区分一个系统的多个BFD会话。
4. Your Discriminator：BFD会话连接远端标识符（Remote Discriminator）。从远端系统接收到的鉴别值，这个域直接返回接收到的“My Discriminator”，如果不知道这个值就返回0。
5. Desired Min TX Interval：本地支持的最小BFD报文发送间隔。
6. Required Min RX Interval：本地支持的最小BFD报文接收间隔。
7. Required Min Echo RX Interval：本地支持的最小Echo报文接收间隔，单位为微秒（如果本地不支持Echo功能，则设置0）。

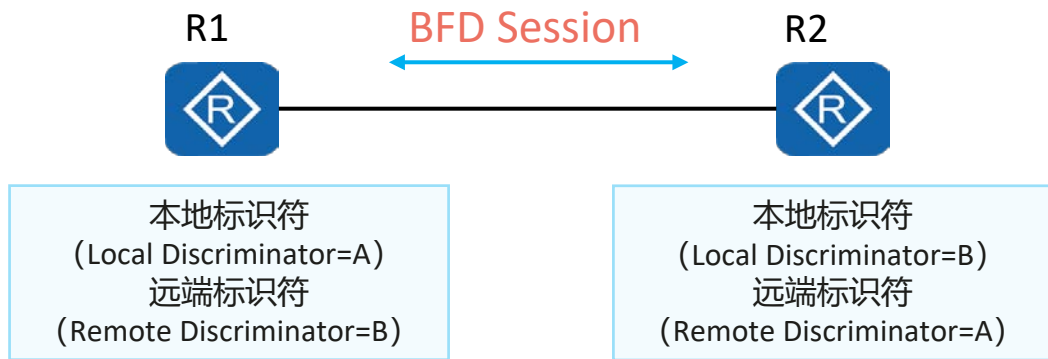


BFD会话建立

BFD会话的建立有两种方式，即静态建立BFD会话和动态建立BFD会话。BFD通过控制报文中的本地标识符和远端标识符区分不同的会话。静态和动态创建BFD会话的主要区别在于Local Discriminator和Remote Discriminator的配置方式不同。

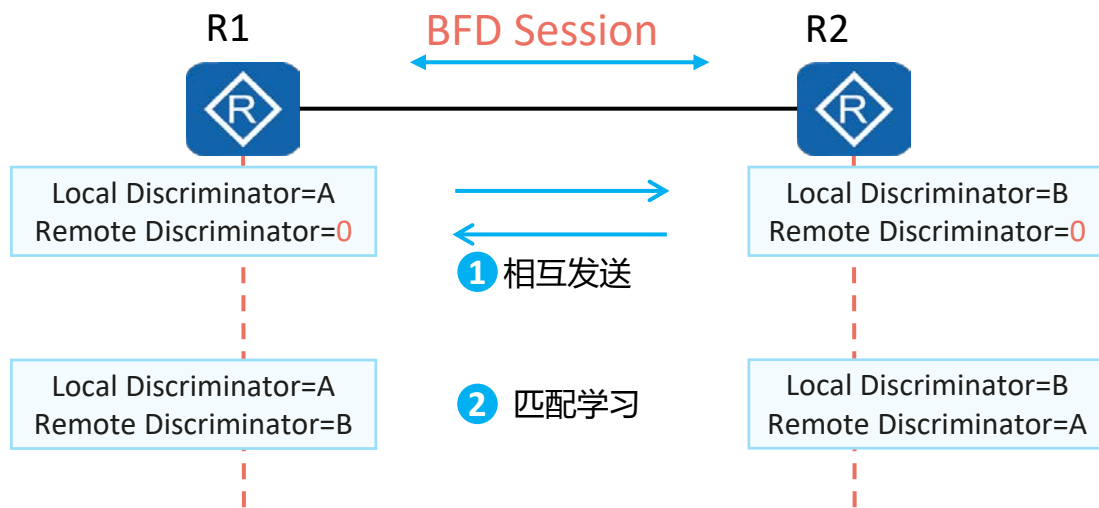
静态建立BFD会话

静态建立BFD会话是指通过命令行手工配置BFD会话参数，包括配置本地标识符和远端标识符等，然后手工下发BFD会话建立请求。



动态建立BFD会话

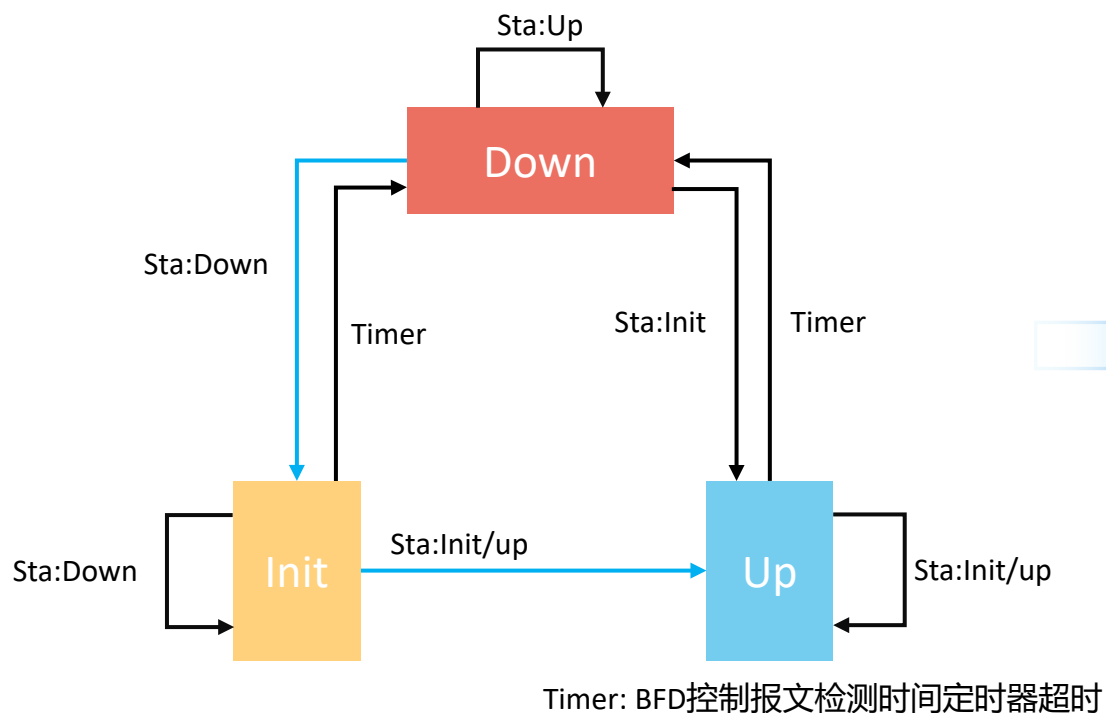
动态建立BFD会话的本地标识符由触发创建BFD会话的系统动态分配，远端标识符从收到对端BFD消息的Local Discriminator的值学习而来。



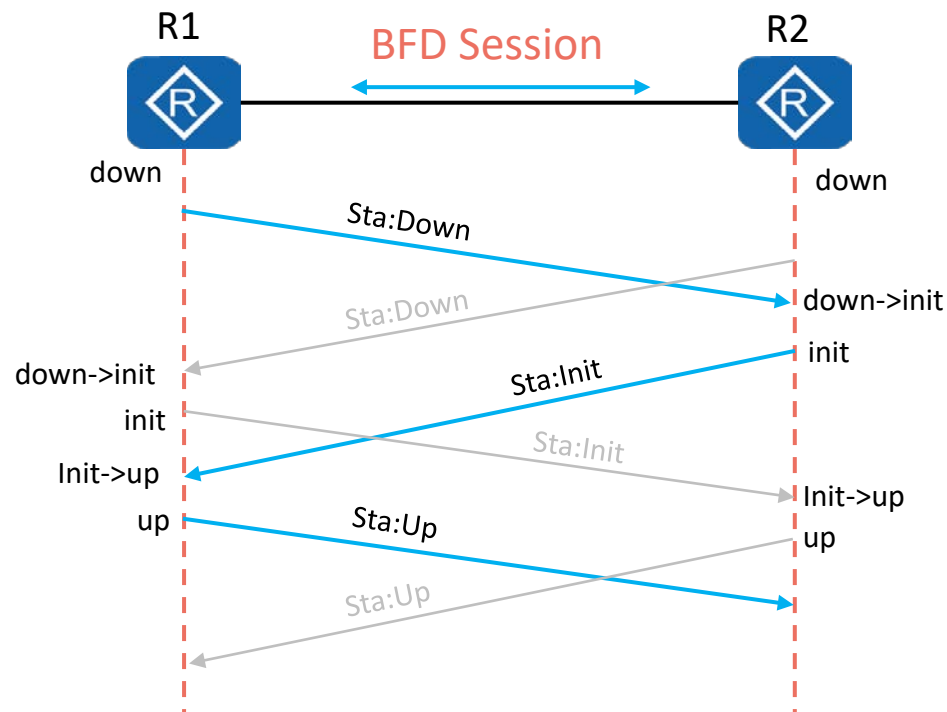


BFD会话状态

BFD会话有四种状态：Down、Init、Up和AdminDown。会话状态变化通过BFD报文的State字段传递，系统根据自己本地的会话状态和接收到的对端BFD报文驱动状态改变，如左下图所示。BFD状态机的建立和拆除都采用三次握手机制，如右下图所示，以确保两端系统都能知道状态的变化。



BFD会话状态转换图



BFD会话建立状态迁移流程

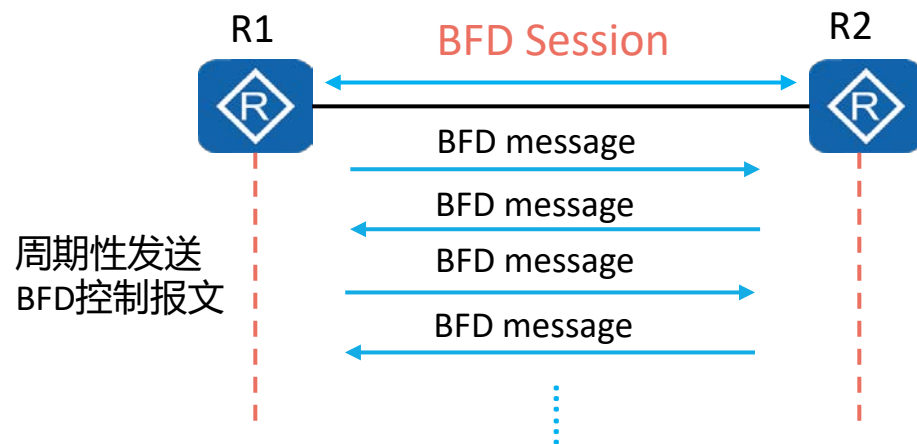


BFD检测模式

BFD的检测机制：两个系统建立BFD会话，并沿它们之间的路径周期性发送BFD控制报文，如果一方在既定的时间内没有收到BFD控制报文，则认为路径上发生了故障。BFD的检测模式有异步模式和查询模式两种。

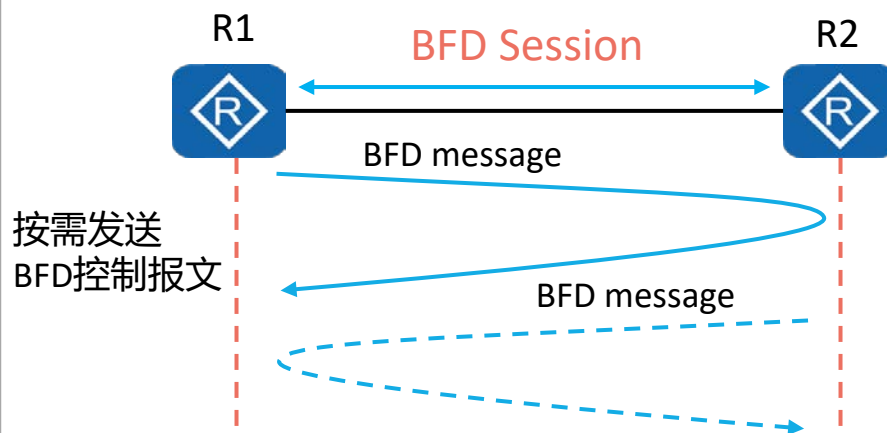
异步模式

系统之间相互周期性地发送BFD控制包，如果某个系统在检测时间内没有收到对端发来的BFD控制报文，就宣布会话为Down。



查询模式

在需要验证连接性的情况下，系统连续发送多个BFD控制包，如果在检测时间内没有收到返回的报文就宣布会话为Down。

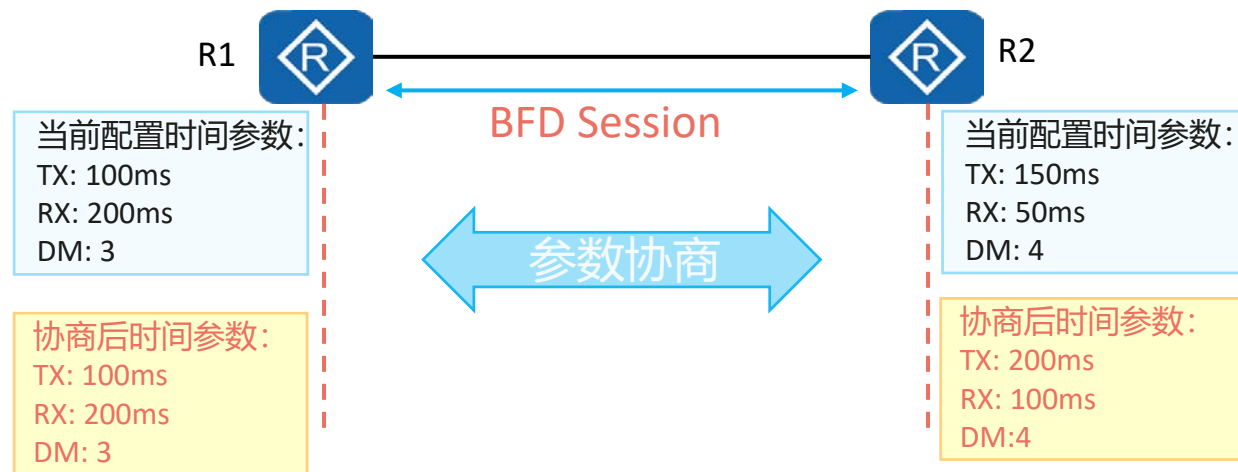




BFD检测时间

BFD会话检测时长由TX (Desired Min TX Interval) , RX (Required Min RX Interval) , DM (Detect Multi) 三个参数决定。BFD报文的实际发送时间间隔, 实际接受时间间隔由BFD会话协商决定。

- 本地BFD报文实际发送时间间隔 = $\text{MAX} \{ \text{本地配置的发送时间间隔}, \text{对端配置的接收时间间隔} \}$
- 本地BFD报文实际接收时间间隔 = $\text{MAX} \{ \text{对端配置的发送时间间隔}, \text{本地配置的接收时间间隔} \}$
- 本地BFD报文实际检测时间:
 - 异步模式: 本地BFD报文实际检测时间 = 本地BFD报文实际接收时间间隔 \times 对端配置的BFD检测倍数
 - 查询模式: 本地BFD报文实际检测时间 = 本地BFD报文实际接收时间间隔 \times 本端配置的BFD检测倍数



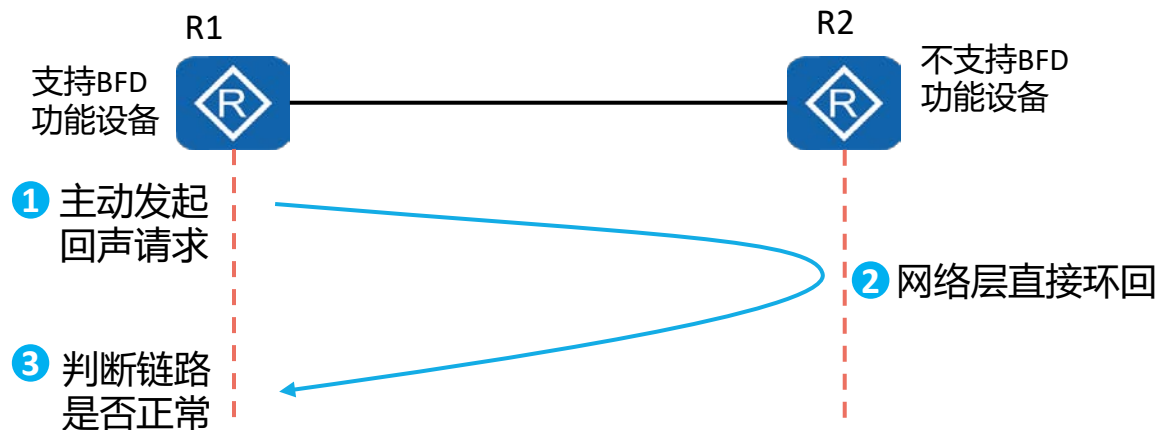
R1在异步模式下的实际检测时间 = $4 \times 200\text{ms}$

R1在查询模式下的实际检测时间 = $3 \times 200\text{ms}$



BFD Echo功能

- BFD Echo功能也称为BFD回声功能，是由本地发送BFD Echo报文，远端系统将报文环回的一种检测机制。
- 在两台直接相连的设备中，其中一台设备支持BFD功能（R1）；另一台设备不支持BFD功能（R2），只支持基本的网络层转发。为了能够快速的检测这两台设备之间的故障，可以在支持BFD功能的设备上创建单臂回声功能的BFD会话。支持BFD功能的设备主动发起回声请求功能，不支持BFD功能的设备接收到该报文后直接将其环回，从而实现转发链路的连通性检测功能。





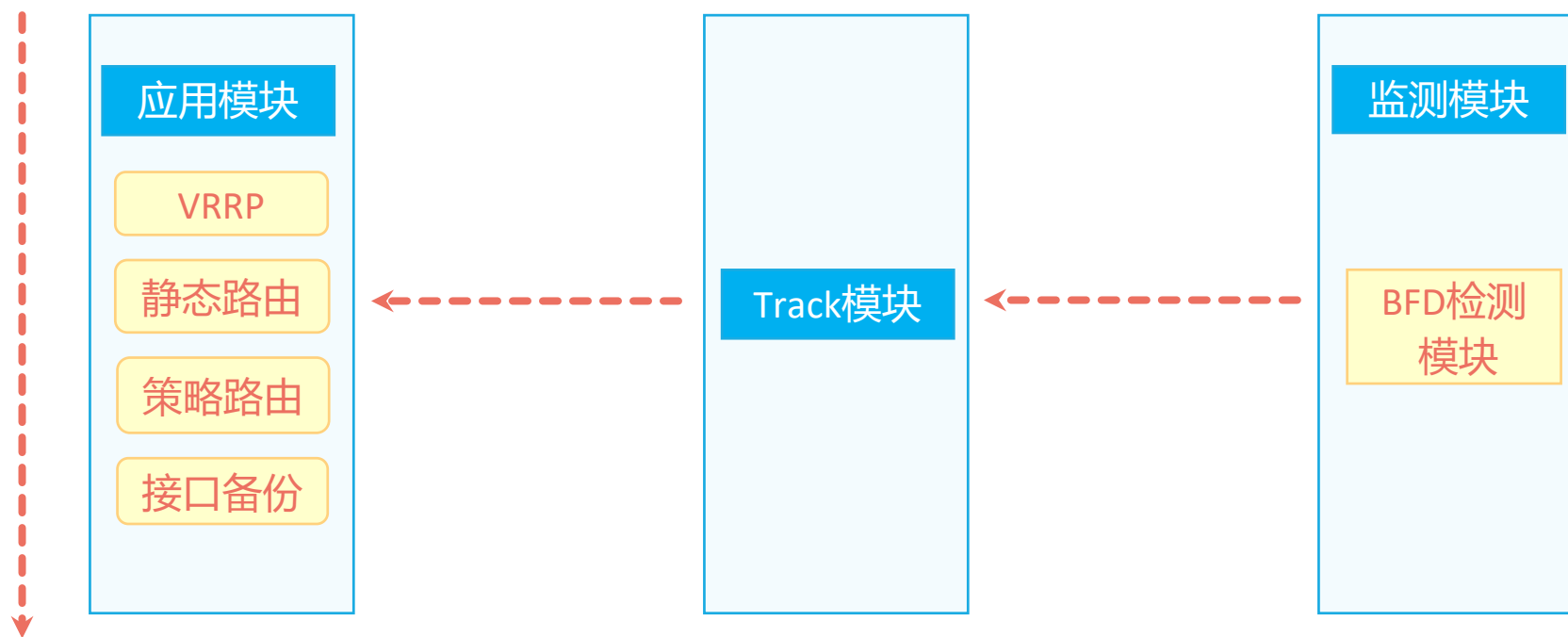
目录

1. BFD产生背景
2. BFD工作原理
- 3. BFD应用场景**
4. BFD基本配置



联动功能简介

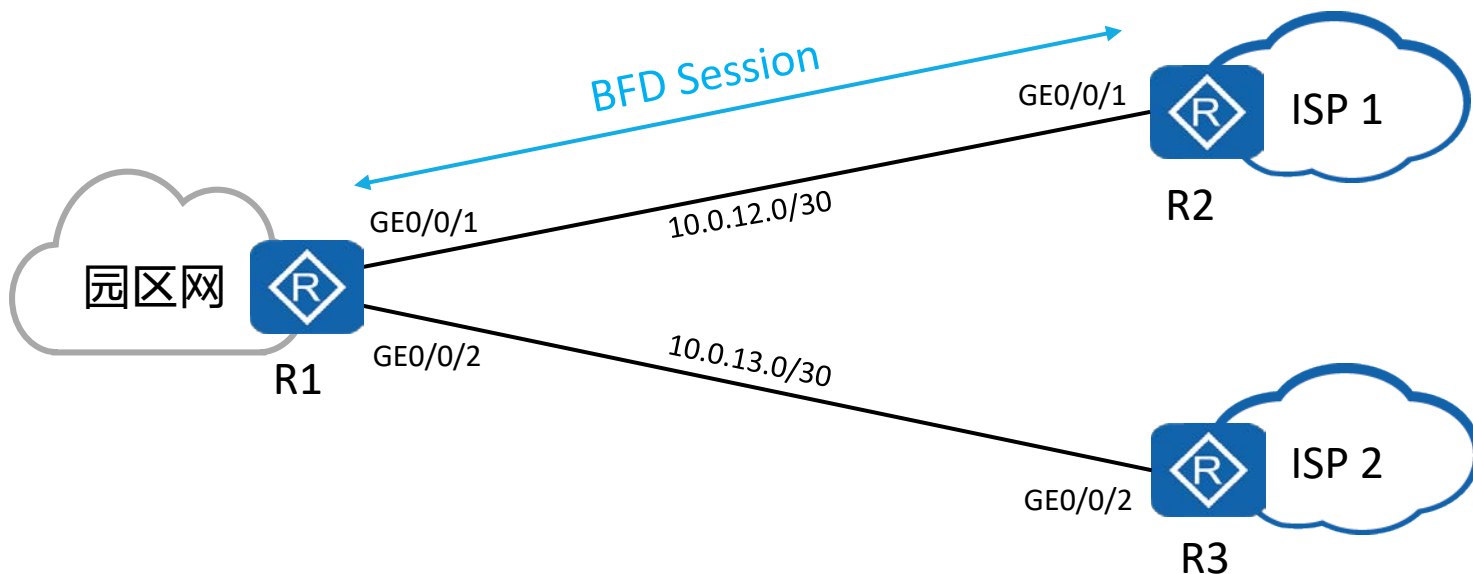
联动功能由检测模块、Track和应用模块三部分组成。





静态路由与BFD联动

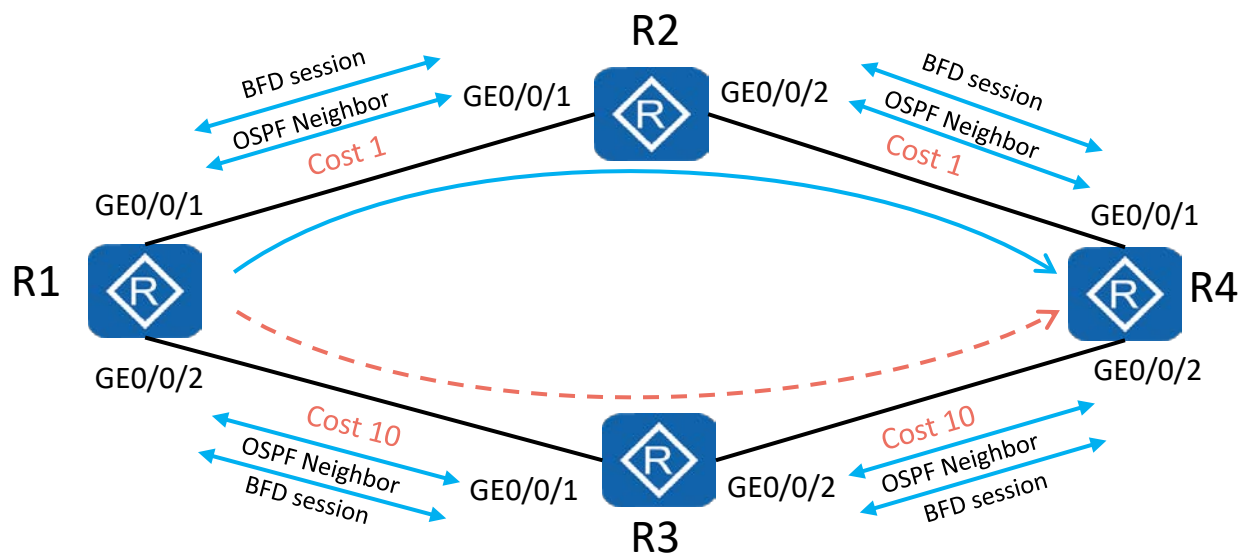
- 静态路由自身没有检测机制，如果静态路由存在冗余路径，通过静态路由与BFD联动，当主用路径故障时，实现静态路由的快速切换。
- 静态路由与BFD联动应用广泛，如下图中R1是园区网的出口路由器，R1通过两条链路分别连接ISP1和ISP2，正常情况下默认路由经过的链路为指向ISP1的链路，当通往ISP1的链路出现故障的时候，BFD会话能够快速感知，并通知路由器将流量切换到指向ISP2的链路。





OSPF与BFD联动 (1)

- OSPF在未绑定BFD的情况下，链路故障检测时间由协议Hello机制决定，通常是秒级。通过绑定BFD，可以实现毫秒级故障检测。
- BFD与OSPF联动就是将BFD和OSPF协议关联起来，BFD将链路故障的快速检测结果告知OSPF协议。

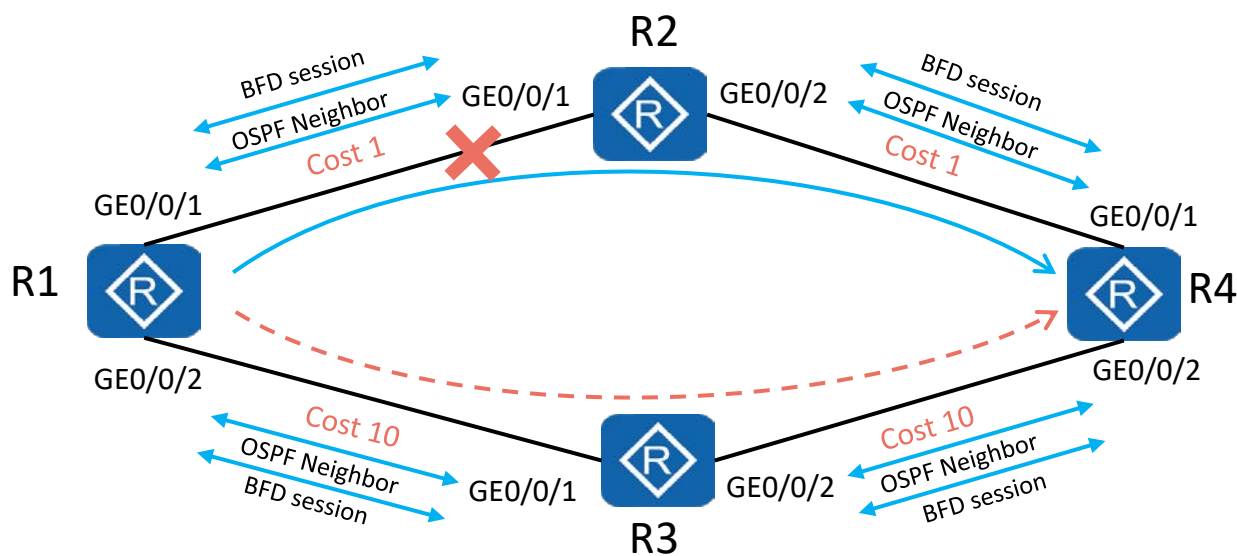


1. OSPF通过自己的Hello机制发现邻居并建立连接。
2. OSPF在建立了新的邻居关系后，将邻居信息（包括目的地址和源地址等）通告给BFD。
3. BFD根据收到的邻居信息建立会话，会话建立以后，BFD开始检测链路故障。
4. 正常情况下，R1根据OSPF路径开销大小选择经过R2到达R4。



OSPF与BFD联动 (2)

BFD会话建立后会周期性地快速发送BFD报文，如果在检测时间内没有收到BFD报文则认为该双向转发路径发生了故障，通知被服务的上层应用进行相应的处理。



1. 当R1和R2之间链路出现故障，BFD首先快速检测到链路故障，BFD会话状态变为Down并通知R1。
2. R1处理邻居Down事件，通知本地OSPF进程邻居不可达，重新进行路由计算，选择通过R3到达R4。



目录

1. BFD产生背景
2. BFD工作原理
3. BFD应用场景
- 4. BFD基本配置**



BFD配置命令介绍 (1)

1. 创建BFD会话绑定信息，并进入BFD会话视图。

```
[Huawei] bfd session-name bind peer-ip ip-address [ vpn-instance vpn-name ] interface interface-type interface-number  
[ source-ip ip-address ]
```

缺省情况下，未创建BFD会话。在第一次创建单跳BFD会话时，必须绑定对端IP地址和本端相应接口，且创建后不可修改。如果需要修改，则只能删除后重新创建。

2. 创建使用组播地址作为对端地址的BFD会话，并进入BFD会话视图。

```
[Huawei] bfd session-name bind peer-ip default-ip interface interface-type interface-number [ source-ip ip-address ]
```

3. 创建BFD for IPv6的绑定信息，并进入BFD会话视图。

```
[Huawei] bfd session-name bind peer-ipv6 ip-address [ vpn-instance vpn-name ] interface interface-type interface-number  
[ source-ipv6 ip-address ]
```

在第一次创建单跳BFD6会话时，必须绑定对端IPv6地址和本端相应接口，且创建后不可修改。



BFD配置命令介绍 (2)

4. 创建静态标识符自协商BFD会话

```
[Huawei] bfd session-name bind peer-ip ip-address [ vpn-instance vpn-name ] interface interface-type interface-number  
[ source-ip ip-address ] auto
```

5. 创建单臂Echo功能的BFD会话

```
[Huawei] bfd session-name bind peer-ip ip-address [ vpn-instance vpn-name ] interface interface-type interface-number  
[ source-ip ip-address ] one-arm-echo
```

6. 配置BFD会话的本地标识符

```
[Huawei-bfd-session-test] discriminator local discr-value
```

此处假设BFD Session名称是test。

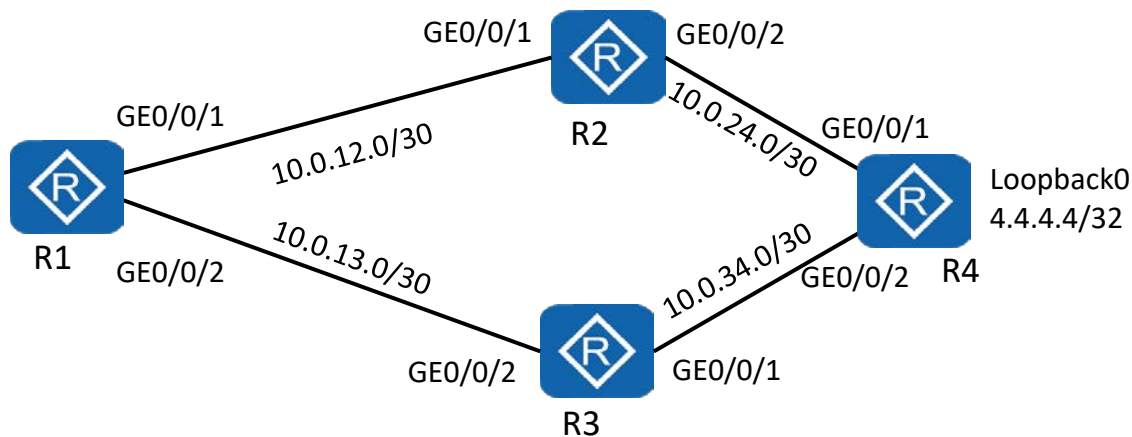
7. 配置BFD会话的远端标识符

```
[Huawei-bfd-session-test] discriminator remote discr-value
```

配置标识符时，本端的本地标识符与对端的远端标识符必需相同，否则BFD会话无法正确建立。并且，本地标识符和远端标识符配置成功后不可修改。



静态路由与BFD联动配置



实验要求：

- 如上图组网所示，在R1上配置到达R4的Loopback0：4.4.4.4/32网段的浮动静态路由，正常情况下通过R2访问R4，当R2故障时，自动选路通过R3访问R4的Loopback0；
- 在R1与R2之间建立BFD会话，并与静态路由绑定，实现故障快速检测和路径快速收敛。

在R1与R2之间建立静态BFD会话：

```
[R1]bfd
[R1]bfd 12 bind peer 10.0.12.2 interface GigabitEthernet 0/0/1
[R1-bfd-session-12]discriminator local 10
[R1-bfd-session-12]discriminator remote 20
[R1-bfd-session-12]commit
[R2]bfd
[R2]bfd 21 bind peer 10.0.12.1 interface GigabitEthernet 0/0/1
[R2-bfd-session-21]discriminator local 20
[R2-bfd-session-21]discriminator remote 10
[R2-bfd-session-21]commit
```

在R1上配置静态路由并绑定BFD会话：

```
[R1] ip route-static 4.4.4.4 32 10.0.12.2 track bfd-session 12
[R1] ip route-static 4.4.4.4 32 10.0.13.2 preference 100
```

此实验其它配置此处省略



BFD会话配置验证

```
[R1]display bfd session all verbose
```

```
Session MIndex : 256 (One Hop) State : Up Name : 12
```

```
Local Discriminator : 10 Remote Discriminator : 20
```

```
Session Detect Mode : Asynchronous Mode Without Echo Function
```

```
BFD Bind Type : Interface(Vlanif10)
```

```
Bind Session Type : Static
```

```
Bind Peer IP Address : 10.0.12.2
```

```
NextHop Ip Address : 10.0.12.2
```

```
Bind Interface : GigabitEthernet0/0/1
```

```
FSM Board Id : 0 TOS-EXP : 7
```

```
Min Tx Interval (ms) : 1000 Min Rx Interval (ms) : 1000
```

```
Actual Tx Interval (ms): 1000 Actual Rx Interval (ms): 1000
```

```
Local Detect Multi : 3 Detect Interval (ms) : 3000
```

```
Echo Passive : Disable Acl Number : -
```

```
Destination Port : 3784 TTL : 255
```

```
----more----
```

BFD会话状态为UP

BFD会话类型为静态BFD

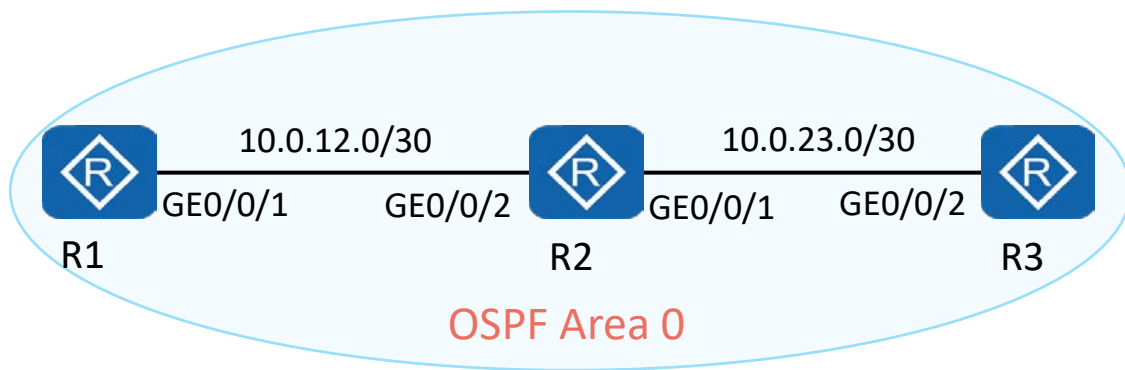
系统配置BFD控制报文最小接收间隔和最小发送间隔

系统协商后的BFD控制报文实际最小接受间隔和最小发送间隔

系统检测次数以及故障检测间隔



OSPF与BFD联动配置



实验要求:

- R1、R2、R3运行OSPF协议，且都属于Area 0；
- 配置OSPF与BFD联动，通过设置所有OSPF接口的BFD会话参数进一步提高链路状态变化时OSPF的收敛速度；
- 将BFD会话的最大发送间隔和最大接受间隔都设置为100ms,检测次数默认不变。

R1配置如下:

```
[R1]bfd
[R1]interface GigabitEthernet 0/0/1
[R1-GigabitEthernet0/0/1]ip address 10.0.12.1 30
[R1]ospf 1
[R1-ospf-1]area 0
[R1-ospf-1-area-0.0.0.0]network 10.0.12.0 0.0.0.3
[R1-ospf-1-area-0.0.0.0]quit
[R1-ospf-1]bfd all-interfaces enable
[R1-ospf-1]bfd all-interfaces min-tx-interval 100 min-rx-interval 100 detect-multiplier 3
```

R2和R3的配置与R1类似，此处省略。



BFD检测配置验证

```
[R1]display bfd session all verbose
```

Session Mince : **256** **(One Hop) State : Up** Name : dyn_8192

Local Discriminator : 8192 Remote Discriminator : 8192

Session Detect Mode : Asynchronous Mode Without Echo Function

BFD Bind Type : Interface(GigabitEthernet0/0/0)

Bind Session Type : **Dynamic**

Bind Peer IP Address : 10.0.12.2

NextHop Ip Address : 10.0.12.2

Bind Interface : GigabitEthernet0/0/0

FSM Board Id : 0 TOS-EXP : 7

Min Tx Interval (ms) : 100 Min Rx Interval (ms) : 100

Actual Tx Interval (ms): 100 Actual Rx Interval (ms): 100

Local Detect Multi : 3 Detect Interval (ms) : 300

Echo Passive : Disable Acl Number : -

```
[R1]display ospf 1 bfd session all
```

OSPF Process 1 with Router ID 10.0.12.1

Area 0.0.0.0 interface 10.0.12.1(GigabitEthernet0/0/0)'s BFD Sessions

NeighborId:10.0.12.2

AreaId:0.0.0.0

Interface:GigabitEthernet0/0/0

BFDState:up rx :100 tx :100

Multiplier:3

BFD Local Dis:8192

LocalIpAdd: 10.0.12.1

RemoteIpAdd:10.0.12.2

Diagnostic Info:No diagnostic information



思考题

1. （多选题） BFD会话建立过程中有以下哪几种状态？
 - A. Down
 - B. Init
 - C. Up
 - D. Establish
2. （多选题） BFD检测模式有哪些？
 - A. 异步模式
 - B. 同步模式
 - C. 查询模式
 - D. 回声模式



本章总结

- 本章详细介绍了BFD工作原理，内容包括会话建立的过程、检测模式、检测时间以及检测流程。
- BFD会话建立中经过三个不同的状态，包括down、init和up，会话状态建立根据收到报文中的sta字段进行转换。BFD检测模式分为异步检测和查询模式，它们的主要区别不仅是发送检测报文的方式和检测的位置不同，同时检测时间与检测流程也不同。
- BFD能够快速检测链路故障，但是只有被应用模块调用，例如OSPF、静态路由，才能够实现流量路径的快速切换，因此BFD在线网当中应用非常广泛。

The background of the slide features a blue-tinted image of several business professionals in a modern office environment. They are standing on a highly reflective floor, and their silhouettes are clearly visible. The overall aesthetic is professional and corporate.

谢谢

www.huawei.com