OSPF GR 技术白皮书

Copyright © 2019 紫光恒越技术有限公司 版权所有，保留一切权利。
非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。
除紫光恒越技术有限公司的商标外，本手册中出现的其它公司的商标、产品标识及商品名称，由各自权利人拥有。本文中的内容为通用性技术信息，某些信息可能不适用于您所购买的产品。

目 录

1 概述 ····· 1
1.1 产生背景 ····· 1
1.2 技术优点 ····· 1
2 OSPF GR 技术详细介绍 ····· 1
2.1 概念介绍 ····· 1
2.2 IETF 标准 OSPF GR 实现机制 ····· 2
2.2.1 Grace LSA ····· 2
2.2.2 IETF 标准 OSPF GR 运行过程 ····· 3
2.3 非 IETF 标准 OSPF GR 实现机制 ····· 3
2.3.1 LLS 能力扩展 ····· 4
2.3.2 OOB 能力扩展 ····· 5
2.3.3 非 IETF 标准 OSPF GR 运行过程 ····· 5
3 典型组网应用 ····· 7
3.1 OSPF GR 典型组网应用 ····· 7
3.1.1 组网图 ····· 7
3.1.2 组网需求 ····· 7
4 参考文献 ····· 8

1 概述

GR（Graceful Restart，平滑重启）是一种在主备切换或协议重启时保证转发业务不中断的机制。其核心在于：设备进行协议重启时，能够通知其周边设备，使到该设备的邻居关系和路由在一定时间内保持稳定。在协议重启完毕后，周边设备协助其进行信息（包括支持 GR 的相关协议所维护的各种拓扑、路由和会话信息）同步，在尽量短的时间内恢复到重启前的状态。在协议重启过程中不会产生路由振荡，报文转发路径也没有任何改变，整个系统可以实现不间断运行。

OSPF GR 可以保证运行 OSPF 协议的路由器在进行主备切换或 OSPF 协议重启时，转发业务正常进行。

1.1 产生背景

一台路由器的 OSPF 协议重启之后，会发送 Hello 报文以发现邻居，而邻居路由器因为之前已经和该路由器建立有邻居关系，因此收到该 Hello 报文之后就会把该路由器从邻居列表中删除，断开与该路由器的邻居关系，并通知其他路由器。当该路由器与邻居路由器重新建立 OSPF 邻居关系后，会重新同步所有的路由信息数据，而周边路由器也需要重新进行路由计算，这样就会引起网络的路由振荡以及转发中断，对于一个大型网络，尤其是运营商网络，这些路由振荡和转发中断是不可容忍的。

如何能使协议重启而不引起网络振荡和转发中断呢？由于分布式设备的控制与转发是分开的，主控板负责整个设备的控制与管理，包括协议运行和路由计算，而接口板则负责数据转发。这样当发生主备倒换或协议重启时，数据转发仍然可以不中断进行。

同时，如果在重启期间周边设备能维持邻居关系不变并保持路由稳定，并且在设备重启后能协助设备进行路由信息同步，在尽量短的时间内使得本设备的路由信息恢复到重启前的状态，那么就可以保持网络拓扑稳定，不引起网络路由振荡。

为了实现这个目的，OSPF 路由协议进行了扩展，即 OSPF GR。通过 OSPF GR，可以避免网络的路由振荡及转发中断。

1.2 技术优点

- 保证协议重启或主备切换过程中转发业务不中断
- 减少协议重启或主备切换时路由振荡对全网的影响
- 减少单点故障，提高整个网络的可靠性

2 OSPF GR 技术详细介绍

2.1 概念介绍

OSPF GR 的基本术语：

- GR Restarter：发生协议重启事件且具有 GR 能力的设备。
- GR Helper：和 GR Restarter 具有邻居关系，协助完成 GR 流程的设备。

- GR Session：OSPF 邻居建立时进行关于 GR 能力的协商，一般把 GR 能力协商过程称为 GR Session。协商的内容包括双方是否都具备 GR 能力等。一旦 GR 能力协商通过，当协议重启时就可以进入 GR 流程。

说明

分布式设备可以充当 GR Restarter 和 GR helper；而集中式设备只能充当 GR Helper，协助 GR Restarter 完成 GR 流程。

目前有两种方式实现OSPF GR 技术：

- 一种是基于 IETF 标准，Restarter 通过向 Helper 发送一种称为 Grace LSA（Link state advertisement，链路状态信息公告）的 9 类 Opaque LSA 来控制 GR 的交互过程。
- 另外一种是非 IETF 标准，Restarter 与 Helper 之间是通过相互发送携带 LLS（Link local signaling，本地链路信令）与 OOB（Out-of-band LSDB Resynchronization，带外的链路状态信息库重新同步）扩展信息的 OSPF 报文来完成 GR 的交互过程。

说明

GR Restarter 上只能配置 IETF 标准与非 IETF 标准其中的一种能力，两种方式是互斥的。

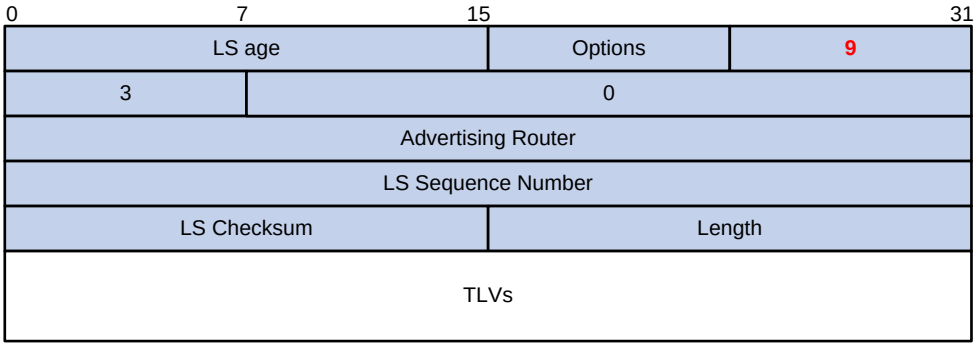
2.2 IETF标准OSPF GR实现机制

对 OSPF 协议做了如下扩展：增加一种 9 类 Opaque LSA——Grace LSA，用以在 GR Restarter 重启时，通知周边邻居进入 GR Helper 流程。

2.2.1 Grace LSA

Grace LSA 格式如图 1 所示。

图1 Grace LSA 格式



其中，几个主要的 TLV 的含义如下：

- Grace Period TLV：Type 取值为 1，长度为 4 字节，表示邻居设备进入 GR Helper 处理流程的最长保持时间。如果超过这段时间 GR Restarter 还没有完成 GR 处理流程，则周边邻居不再担任 GR Helper 角色。该 TLV 是 Grace LSA 必须携带的。
- Graceful Restart Reason TLV：Type 取值为 2，长度为 1 字节，告知邻居设备 GR Restarter 的重启原因。Value 取值为 0 表示原因未知，取值为 1 表示软件重启，取值为 2 表示软件重新加载（升级），取值为 3 表示 GR Restarter 进行主备倒换。Grace LSA 中必须携带该 TLV。
- IP Interface Address TLV：Type 取值为 3，长度为 4 字节，用来告知发送 Grace LSA 的接口的 IP 地址，在网络上需要用该 IP 地址来唯一标识一台重启设备。

2.2.2 IETF 标准 OSPF GR 运行过程

如图 2 所示，假设 Router A 和 Router B 原来有稳定的 OSPF 邻居关系，并且 Router A 使能了 GR 能力，此时 Router A 重启，将按照以下的流程进行路由信息交互：

(1) Router A 重启之后向 Router B 发送 Grace LSA。

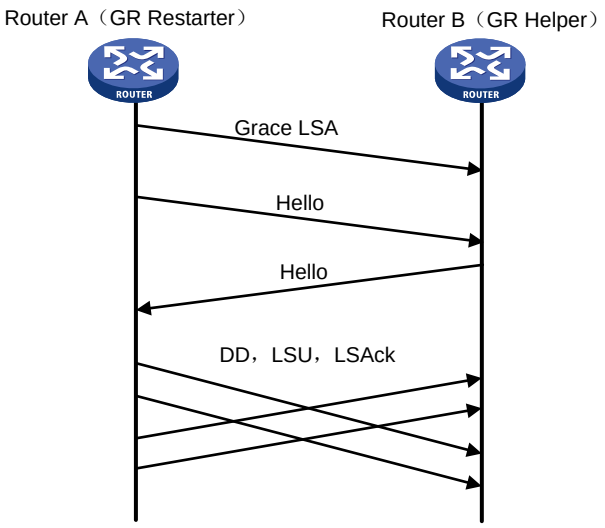
(2) Router B 收到 Router A 发出的 Grace LSA 后，会维持与 Router A 的邻居关系不变。

(3) Router A 与 Router B 进行 Hello 报文与 DD（Database Description，链路状态信息库描述）报文交互和 LSDB 同步。由于 GR 过程中不能生成 LSA，所以在 LSDB 同步过程中，如果 Router A 从 Router B 收到自己产生的 LSA，直接存储下来并置上 Stale 标志。

(4) 完成 LSDB 的同步之后，Router A 发送 Grace LSA（Grace Period TLV 的 Type 值为 0）通知 Router B 结束 GR 流程。Router A 进入正常 OSPF 流程，然后 Router A 重新生成 LSA，并删除那些置了 Stale 标志而又没有被重新生成的 LSA。

(5) Router A 在恢复所有路由信息后重新进行路由计算，重新刷新 FIB 表。

图2 IETF 标准 OSPF GR 运行示意图



2.3 非IETF标准OSPF GR实现机制

与 IETF 标准的 OSPF GR 相比，非 IETF 标准 GR 主要是通过下面两个能力扩展来支持 OSPF GR：

- Link-Local Signaling：OSPF 本地链路通告扩展，用来标识当前路由器和其他路由器进行一些可选信息的通信。
- Out-of-band LSDB Resynchronization：带外的 LSDB 重新同步，完成无邻接关系变化的 LSDB 重新同步。

2.3.1 LLS 能力扩展

在 OSPF 的 Hello 报文和 DD 报文的尾部增加了 LLS Data。同时对 Hello 报文和 DD 报文的 Option 字段进行扩展，以标识是否携带 LLS。扩展后的 Option 字段如图 3 所示，如果 L 位置 1 表示携带了 LLS Data。

图3 Option 字段

*	*	DC	L	N/P	MC	E	*

经过扩展后的报文格式如图 4 所示。

图4 LLS 扩展后的 OSPF 报文格式图

IP Header
OSPF Header
OSPF Data
Authentication Data
LLS Data

LLS Data 的具体的格式如图 5 所示。

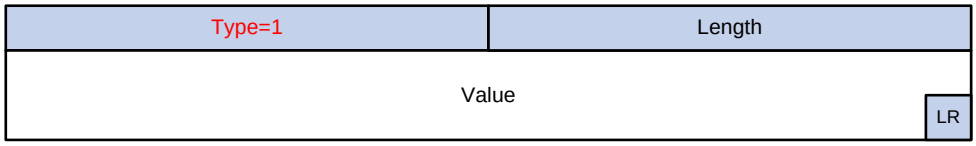
图5 LLS Data 字段

Checksum	LLS Data Length
TLVs	

LLS Data 中的 TLV 可以用于进行不同的 OSPF 扩展。例如为支持 GR，使用了类型为 1 的 TLV，即 Extended Option TLV（EO TLV）。下面我们讲一下 EO TLV 里两个重要字段：

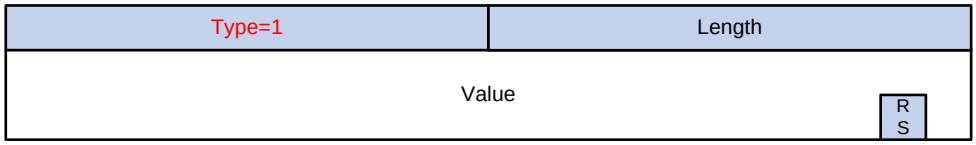
- LR 字段：用以实现 OOB 能力的协商，LR 在 EO TLV 中的位置如图 6 所示。如果当前路由器具备 OOB 能力，则在发送 Hello 和 DD 报文时将 LR（LSDB Resynchronization）进行置位；否则不对 LR 进行置位。关于 OOB 能力的介绍请参考 2.3.2 OOB 能力扩展。

图6 EO TLV LR 位说明图



- RS 字段：用以通知邻居当前路由器进入 GR 流程，RS 在 EO TLV 中的位置如图 7 所示。当路由器需要短暂离开网络时，可以通过将 Hello 报文中的 RS（Restart Signal）置位来向周边邻居通知自己需要进入 GR 流程，这样周边邻居就会保持与此路由器的邻居关系不变；否则不对 RS 进行置位。

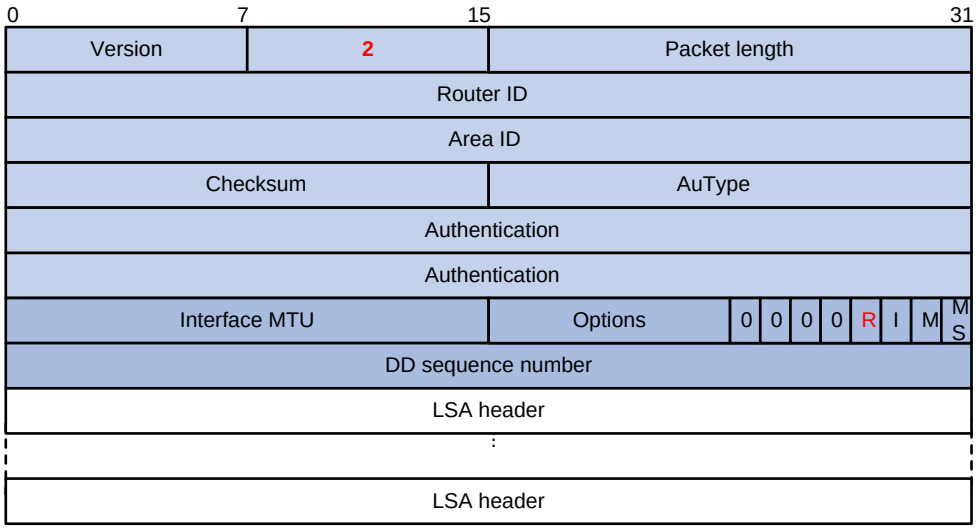
图7 EO TLV RS 位说明图



2.3.2 OOB 能力扩展

当路由器通过 LLS 完成 OOB 能力协商，确认双方都具备 OOB 能力后，则表示可以在邻居关系和网络拓扑稳定时进行 LSDB 同步，即进入 OOB 流程。这时需要在向邻居发送的 DD 报文中将 Option 字段中的 R 进行置位，R 在 Option 字段中的位置如图 8 所示。

图8 DD 报文



2.3.3 非 IETF 标准 OSPF GR 运行过程

如图 9 所示，假设 Router A 和 Router B 原来有稳定的 OSPF 邻居关系，并且 Router A 使能了 GR 能力，此时 Router A 重启，将按照以下的流程进行路由信息交互：

(1) Router A 重启后向 Router B 发送 LR 和 RS 都置位的 Hello 报文，向 Router B 通告自己只是暂时断开马上就会恢复正常，并且自己具备 OOB 能力；

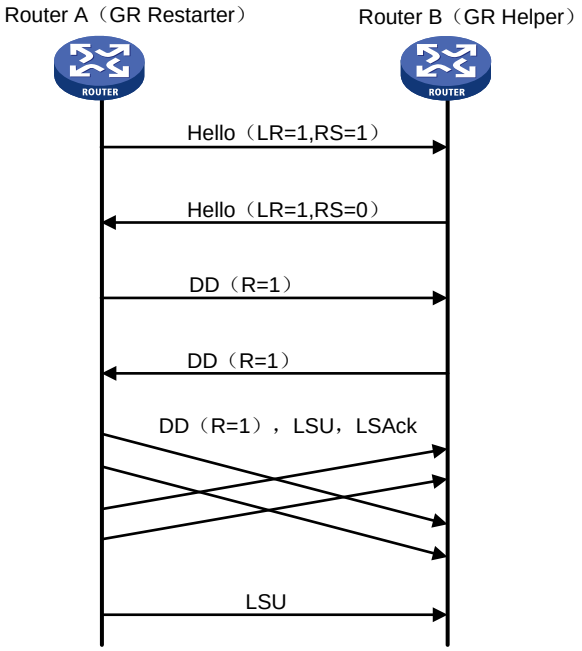
(1) Router B 收到 Hello 报文后，也会向 Router A 回复一个 LR 仍保持置位但是 RS 位已经清除的 Hello 报文，向 Router A 通告自己已经知道它是要短暂离开，并且自己也具备 OOB 能力；

(2) Router A 向 Router B 发送 R 置位的 DD 报文，向 Router B 发起 LSDB 同步请求，在 LSDB 同步期间，Router B 不会将 Router A 从自己的邻居列表中删除，在自己生成的 Router LSA（Network LSA）中，与 Router A 的邻居关系仍然为 Full。由于 GR 过程中不能生成 LSA，所以在 LSDB 同步过程中，如果 Router A 从 Router B 收到自己产生的 LSA，直接存储下来并置上 Stale 标志。

(3) 完成 LSDB 的同步之后，Router A 结束 GR 流程，进入正常 OSPF 流程，然后 Router A 重新生成 LSA，并删除那些置了 Stale 标志而又没有被重新生成的 LSA。Router B 在和 Router A 的邻居关系重新达到 Full 状态后退出 GR 流程，进入正常 OSPF 流程。

(4) Router A 在恢复所有路由信息后重新进行路由计算，重新刷新 FIB 表。

图9 非 IETR 标准 GR 同步过程

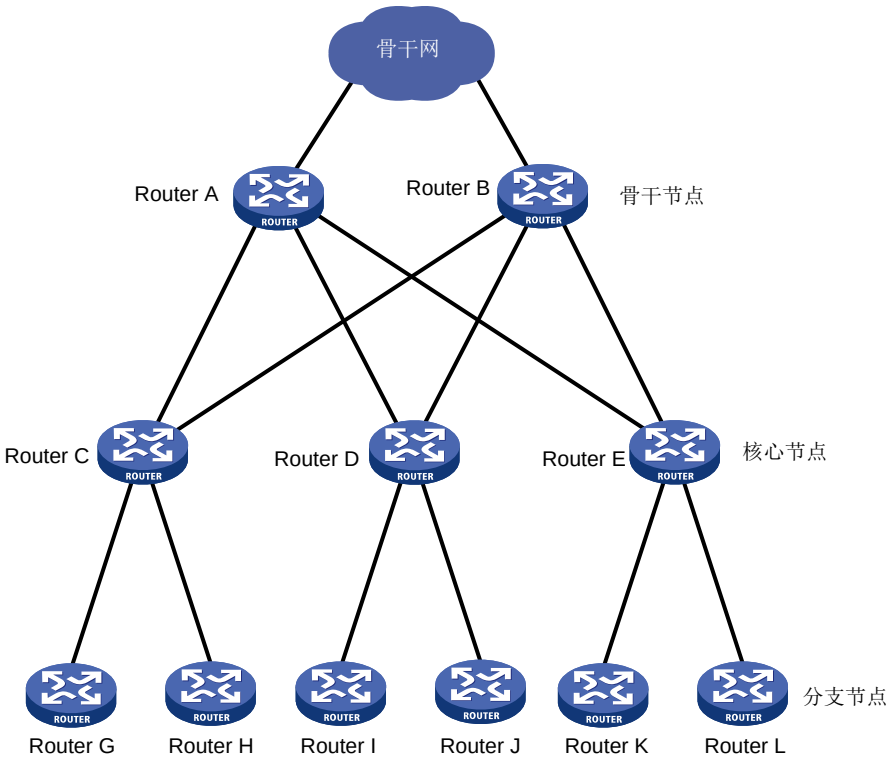


3 典型组网应用

3.1 OSPF GR典型组网应用

3.1.1 组网图

图10 OSPF GR 配置组网图



3.1.2 组网需求

- Router A、Router B、Router C、Router D、Router E、Router F、Router G、Router H、Router I、Router J、Router K、Router L 运行 OSPF 协议。
- Router A、Router B 与骨干网相连，属于骨干节点。
- Router G、Router H、Router I、Router J、Router K、Router L 是分支节点，通过核心节点 Router C、Router D、Router E 与骨干节点相连，通过骨干节点连接到骨干网上。
- 使用 GR 保证网络中的骨干节点和核心节点在出现协议重启时的转发业务不中断，避免出现不必要的路由振荡。
- 骨干节点和核心节点作为 GR Restarter（同时缺省也作为 GR Helper），分支节点作为 GR Helper。这样当骨干节点发生主备切换或重启 OSPF 进程时，核心节点可以作为 GR Helper 协助其进行 LSDB 重同步，并且保持转发不中断；当核心节点发生主备切换或重启 OSPF 进程时，骨干节点和分支节点都可以作为 GR Helper 协助其进行 LSDB 重同步，并且保持转发不中断。

4 参考文献

- RFC 2328：OSPF Version 2
- RFC 3623：Graceful OSPF Restart
- RFC 4811：OSPF Out-of-Band LSDB Resynchronization
- RFC 4812：OSPF Restart Signaling
- RFC 4813：OSPF Link-Local Signaling