NETWORK SCALING WITH BGP LABELED UNICAST

Design and Configuration Guide

Table of Contents

Table of Figures

Introduction

As MPLS deployments expand beyond the service provider core and edge to the access and metropolitan networks, the number of edge-to-edge label-switched paths (LSPs) in many networks is increasing substantially. This continued growth can present scaling challenges—with some networks already reaching their limit—and can also slow end-to-end restoration.

Dividing the network into multiple regions can alleviate these issues by limiting the total number of end-to-end LSPs, and enabling failures to be contained and restored in a single region. These regions operate separate instances of interior gateway protocol (IGP), and use BGP Labeled Unicast (BGP-LU) to advertise route information between inter-region routers. By providing connectivity and communication between regions, BPG-LU enables service providers to massively scale the number of MPLS-enabled devices on their networks.

Scope

The scope of this document includes the implementation and configuration of BGP Labeled Unicast to scale an MPLS network. We provide sample configurations and explanations for implementing BGP-LU in multi-region networks, where regions consist of individual OSPF areas or IS-IS levels. The document assumes the reader has experience with service provider network design, routing protocols such as OSPF, IS-IS, and BGP, and knowledge of MPLS protocols like LDP and RSVP.

Design Considerations: Network Regions

Regions are an important concept because they address many of the challenges inherent in large routed networks. By dividing the network into regions, service providers can increase the scale of their networks and improve convergence times. Regions essentially partition the network into sections or zones, which can be OSPF areas or IS-IS levels within a single autonomous system (AS), or each region can be an AS using a separate IGP.

The characteristics of a multi-region network are quite similar to a multi-area OSPF network, multilevel IS-IS network, or BGP AS, but the regions don’t exchange routing information as would a typical area or level. No IGP routing information, LDP signaling, or RSVP signaling is exchanged between regions. Rather, regions are connected by and communicate with BGP labeled unicast.

Like these other concepts, the primary advantage of regions is reducing the number of entries in the routing and forwarding tables of individual routers. This simplifies the network, enabling greater scale and faster convergence.

LDP and RSVP label-switched paths are contained within a region, reducing the amount of LDP and RSVP state network-wide. Lowering the amount of resources required by each node prolongs the life span of each node as the network continues to grow.

Regions also simplify network integration and troubleshooting. Network integrations and expansions do not require compatible IGPs or compatible LDP/RSVP implementations between networks. The new network or region only needs BGP labeled unicast compatibility with the existing network. Troubleshooting a multi-region network is simplified because problems are more likely to be contained within a single region rather than spread across multiple regions.

In a multi-region network, BGP-LU is essential to enabling inter-region end-to-end routing, as it provides the communication and connectivity between multiple regions. Defined in RFC 3107[1], it enables BGP to advertise unicast routes with an MPLS label binding (a prefix and label). To accomplish this, BGP-LU leverages Multiprotocol Border Gateway Protocol (MP-BGP) and subsequent address family identifier (SAFI) 4 which indicates that the network layer reachability information (NLRI) contains a label mapping. BGP-LU has long been used for inter-AS VPN services such as “carrier’s carrier” and is now being applied to intra-AS in a similar way to achieve massive scaling.

---

[1] “Carrying Label Information in BGP-4”: www.rfc-editor.org/rfc/rfc3107.txt

Implementing Regions

For networks that are not already using regions, the first step is determining which network elements will be grouped together to form a region and its boundaries. Nodes that delineate the region boundary are referred to as regional border routers (RBRs).

In this paper, we will use a simplified example of network topology to illustrate region mapping within an AS. Each network has unique characteristics that might require a variation of the mapping method we describe, however the general concept described below is universally applicable. Consider the network topology in Figure 1. A typical service provider network consists of core, aggregation, and edge layers. Often this formation of nodes is aligned with a geographic location, and is contained within a point of presence (POP).

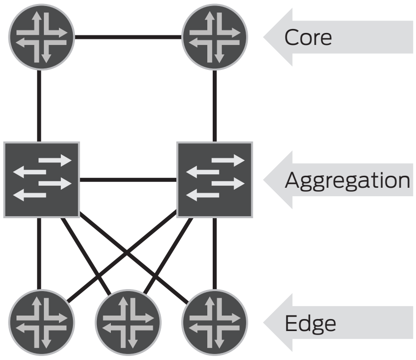

Figure 1: Typical network design

The core nodes in each POP commonly provide the WAN connectivity to other core nodes in remote POPs. Consider the network topology in Figure 2.

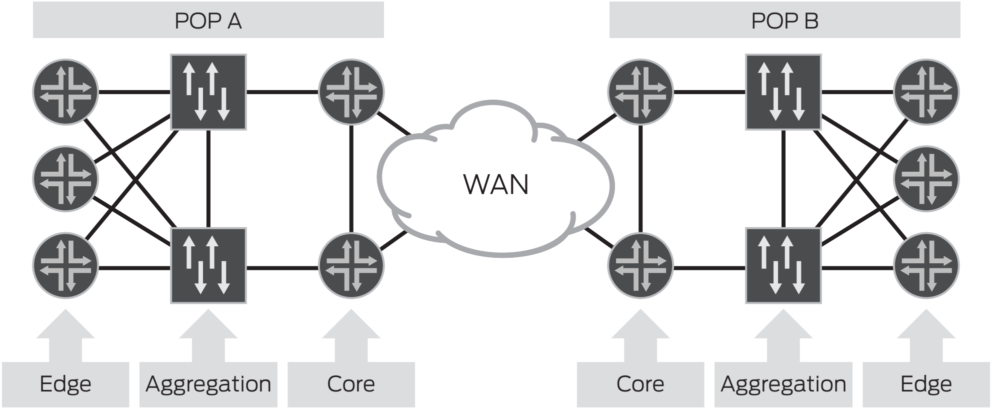

Figure 2: Typical POP hierarchy

The nodes at each layer of the POP hierarchy perform various functions depending on the design and protocols that are implemented. Below are some common characteristics that each layer of nodes might have.

Table 1: Node Characteristics

CORE NODES	AGGREGATION NODES	EDGE NODES
• OSPF area border router (ABR) • IS-IS L1/L2 router • MPLS transit label-switching router (LSR) • BGP route reflector	• L2/L3 edge aggregation • OSPF single area router • IS-IS single level router • BGP route reflector client • MPLS ingress, egress, and transit LSR	• Customer services • Peering • OSPF single area router • OSPF autonomous system border router (ASBR) • IS-IS single level router (IS-IS ASBR) • BGP route reflector client • MPLS provider edge (PE) ingress/egress LSR • VPN PE

The key task when mapping regions to a network is to determine which nodes will perform the role of RBR. After the RBRs are designated, the roles of other nodes logically fall into place. The network topology in Figure 2 suggests that core nodes would make ideal RBRs because, in many cases, they are already performing functions that provide hierarchy and separation of the network.

The network topology in Figure 3 depicts how a typical service provider network could be transformed into a multiregional network.

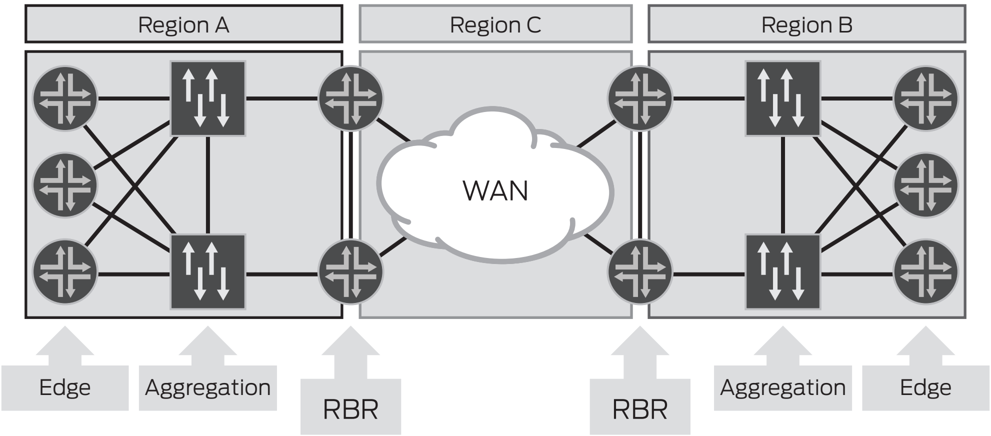

Figure 3: Core routers make logical choices for RBRs

In the diagram above, each region maintains separate IGP and LDP/RSVP signaling. Region C essentially functions as a transport network between the PE routers in regions A and B. The table below summarizes the characteristics of each region.

Table 2: Region Characteristics

	REGION A	REGION B	REGION C
IGP	Independent. No IGP route leaking toward region C.	Independent. No IGP route leaking toward region A or B.	Independent. No IGP route leaking toward region C.
Intraregional MPLS transport	No LDP/RSVP signaling to region C.	No LDP/RSVP signaling to region A or B.	No LDP/RSVP signaling to region C.
Interregional MPLS transport	RBR BGP-LU route reflector Edge BGP-LU client	RBR BGP-LU full mesh RBR BGP-LU route reflector to local region	RBR BGP-LU route reflector Edge BGP-LU client
L2/L3 aggregation	Transit LDP/RSVP LSR No BGP-LU	N/A	Transit LDP/RSVP LSR No BGP-LU

Protocol Operation: BGP-LU Route and Label Advertisement

Once the regions and roles for each router have been defined, the regions need a way to create end-to-end connectivity and enable inter-region communication. Regions do not share IGP routing information, so there needs to be another way for PEs to reach remote PEs in other regions.

BGP-LU provides reachability between regions by advertising PE loopbacks and label bindings to the RBR, which in turn advertises the loopbacks and label bindings to remote PEs in other regions. The diagram in Figure 4 outlines the operation of BGP-LU at a high level in each of the different types of routers in the regions.

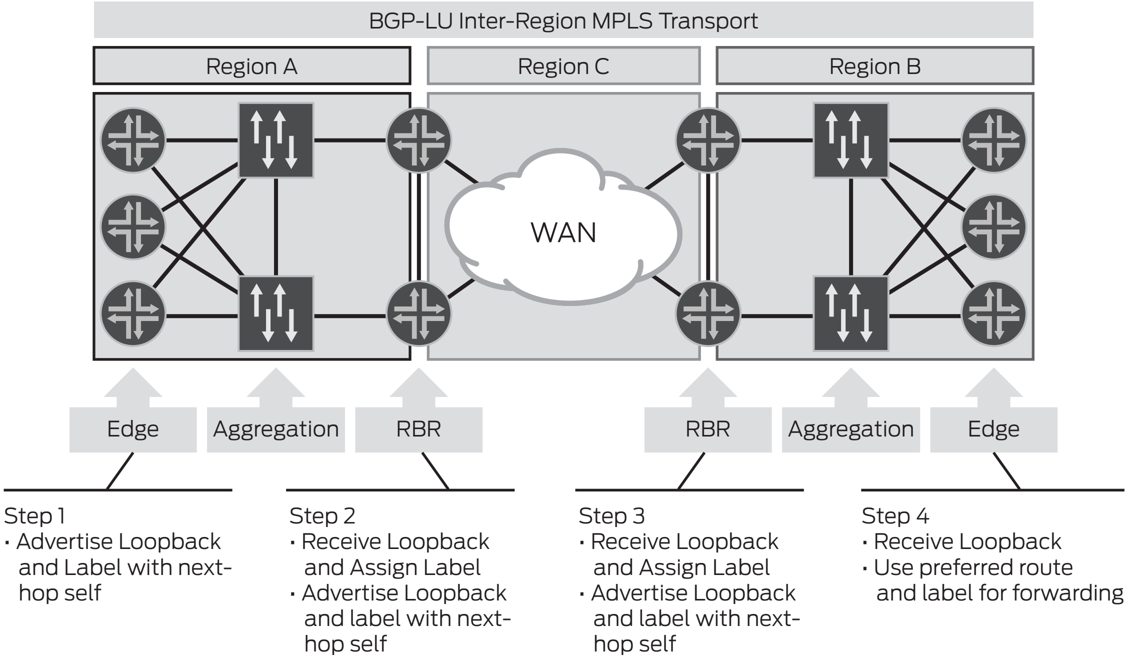

Figure 4: BGP-LU process throughout the network

The structure of the BGP-LU advertisements and resulting forwarding states create the inter-region MPLS transport plane. One of the benefits of BGP-LU is that the transport routers need not have any awareness of, or participation in, BGP-LU—only the PE and RBR routers require support. This makes BGP-LU relatively easy to implement, and also contributes to scalability.

Implementation and Configurations

Now that we have outlined the basics of regions and BGP-LU, we will provide specific configurations that can be used when implementing multi-region networks with BGP-LU. First, we will examine the configurations used for IGPs such as OSPF and IS-IS, followed by the means by which to enable the intra-region transport of BGP-LU. Finally, we will review the configurations that provide inter-region transport of BGP-LU.

IGP

As noted previously, a region can consist of an OSPF area or IS-IS level within an AS, or it can be an AS in and of itself. When using OSPF or IS-IS as the IGP, a configuration must be implemented that prevents the exchange of routing information outside the region—this means within an OSPF area, and it means between Levels 1 and 2 in IS-IS. The full configurations for OSPF and IS-IS are outside the scope of this document, but we will provide the configurations necessary to limit the exchange of routing information. The configurations shown below are based on the assumption that the network is a single AS, with the regions consisting of individual OSPF areas or IS-IS levels.

OSPF

In an OSPF network with multiple areas, the backbone area is referred to as area 0, and the site areas are given another identifying number (other than 0, which is reserved for the backbone area). Area border routers (ABRs) are used to connect the backbone area to non-backbone areas.

To prevent the exchange of routing information between OSPF areas, the non-backbone areas can be configured as stub areas. Stub areas do not receive summary link-state advertisements (LSAs) or a default route from the ABR. Below, we detail the configurations necessary to designate area 1 as a stub area, using the example outlined in Figure 5.

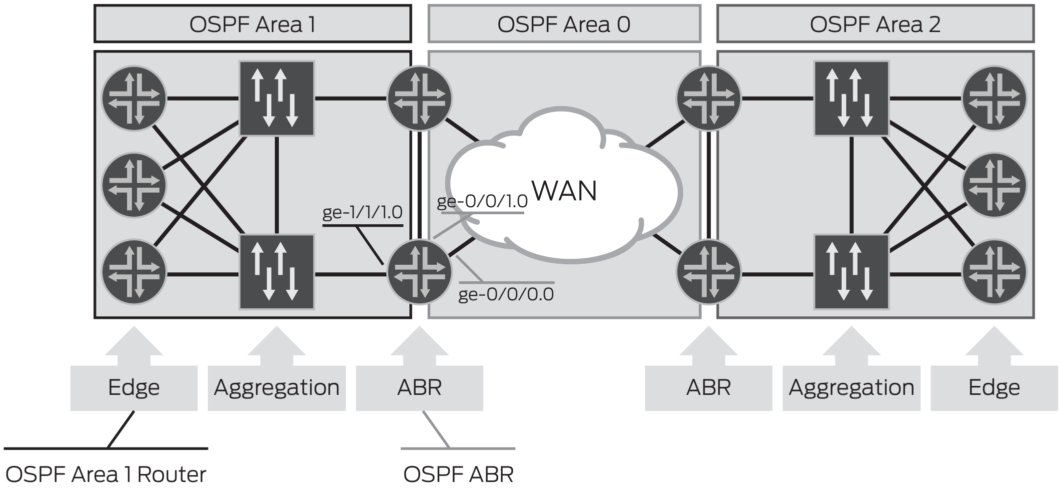

Figure 5: Non-backbone areas can be configured as stub areas

Area border routers and edge routers both need to be configured in this manner.

ABR configurations:

```
protocols {
    ospf {
        area 0.0.0.0 {
            interface lo0.0;
            interface ge-0/0/0.0;
            interface ge-0/0/1.0;
        }
        area 0.0.0.1 {
            stub no-summaries;
            interface ge-1/1/1.0;
        }
    }
}
```

Area 1 edge router configurations:

```
protocols {
    ospf {
        area 0.0.0.1 {
            stub;
            interface lo0.0;
            interface ge-1/1/1.0;
        }
    }
}
```

IS-IS

The scope of the IS-IS configurations in this document will focus on preventing the exchange of routing information between IS-IS level 1 (L1) and level 2 (L2). IS-IS L2 is loosely analogous to the OSPF backbone area 0 and typically provides access to the WAN or "backbone" of the network. Similarly, IS-IS L1 roughly corresponds with an OSPF non-backbone area.

In addition, with IS-IS the concept of areas and area-IDs can be used in conjunction with levels to enable scaling and the desired routing behavior. Because hierarchy can help enable scale, we recommend that each region be configured with its own unique IS-IS area-ID.

L1 routers use L1/L2 routers as the next hop for the default route 0.0.0.0/0. The L1/L2 routers do not advertise a default route—instead, the L1/L2 routers set the attached bit, which is advertised to the L1 router and causes the L1 router to install a default route with the L1/L2 router as the next hop. To prevent a default route from being installed, the attached bit must be ignored, and this can be accomplished by configuring the "ignore-attached-bit" command.

If an L1/L2 router does not have an L2 adjacency to a different area than its own, the L1/L2 router will not set the attach bit, resulting in the L1 routers not installing a default route. As mentioned above, the default route installation can also be prevented by using a single IS-IS area-ID that causes the L1/L2 router to not set the attach-bit. Should the need arise for a default route, it's much easier to simply delete the ignore-attach-bit command than it is to renumber the IS-IS area-IDs or create a routing policy that explicitly advertises a default route. For these reasons, we recommend using unique IS-IS area-IDs to identify each region in conjunction with the ignore-attach-bit command.

By default, IS-IS L1 internal routes are installed into the L2 database; which is an exchange of routing information between levels that needs to be prevented in multi-region networks. To prevent L1 internal routes from being installed into the L2 database, an IS-IS export policy must be configured on the L1/L2 routers to reject L1 routes. L1 external routes are not installed into the L2 database by default, and IS-IS L2 internal routes are not installed into the L1 database. In the configurations below, consider the example network provided in Figure 6.

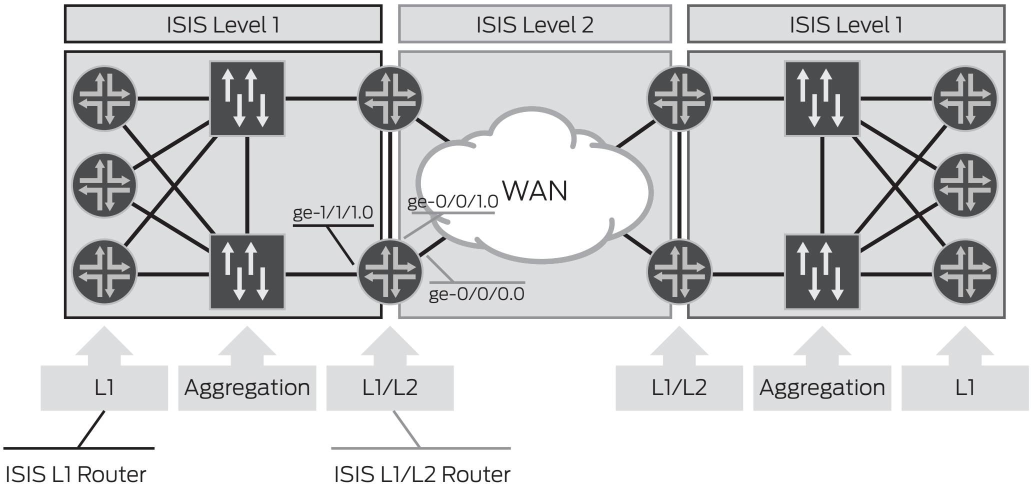

Figure 6: IS-IS configuration example

Similar configurations are required on all L1 and L1/L2 routers, so the configurations listed below would be the same for the routers in all regions.

IS-IS L1/L2 router configurations:

```
protocols {
    isis {
        export reject-l1-into-l2;
        interface ge-0/0/0.0 {
            point-to-point;
            level 2 disable;
        }
        interface fe-0/3/1.0 {
            point-to-point;
            level 1 disable;
        }
        interface lo0.0 {
            level 1 passive;
            level 2 passive;
        }
    }
}
policy-options {
    policy-statement reject-l1-into-l2 {
        term 1 {
            from {
                protocol isis;
                level 1;
            }
            then reject;
        }
    }
}
```

IS-IS L1 router configurations:

```
protocols {
    isis {
        ignore-attached-bit;
        interface ge-2/1/0.0 {
            point-to-point;
            level 2 disable;
        }
        interface lo0.0 {
            level 2 disable;
            level 1 passive;
        }
    }
}
```

Flat IGP Caveats

Using a “flat” IGP—defined as a single OSPF area or single IS-IS level—is supported and may be more suitable depending on the unique characteristics of the network. Using a flat IGP softens the region boundaries, which might make identifying regions and their routes slightly more difficult, and it requires additional configurations.

In a flat IGP, a PE will have two routes to reach remote PEs in other regions. One route is the IGP route (OSPF or IS-IS), and the second route is a BGP-LU route. To determine the preferred route, the router will compare the route-preference values and select the lowest. The Juniper Networks® Junos® operating system default route preferences are:

- RSVP: 7
- LDP: 9
- OSPF Internal: 10
- IS-IS L1 Internal: 15
- IS-IS L2 Internal: 18
- BGP: 170

Because the default BGP route preference is greater than the OSPF and IS-IS values, it will not be selected by default. This results in the BGP-LU label binding not being used for forwarding; which defeats the purpose of having BGP-LU deployed.

To make the BGP-LU route preferred, we must configure a BGP import policy to adjust the BGP-LU route’s route-preference value to a value less than that of the OSPF and IS-IS routes yet still greater than LDP and RSVP. For networks using IS-IS, a value less than 15 and greater than 9 would be suitable. For networks using OSPF, a value less than 10 is required, but because LDP has a default value of 9, this isn’t possible without also changing the LDP default preference to 8 and BGP to 9. In the example below, IS-IS was used at the IGP in conjunction with RSVP, so a value of 9 was sufficient. See example configuration snippets below.

PE BGP import policy configuration:

```
policy-statement L-BGP-Pref-14 {
                from {
                    protocol bgp;
                    rib inet.3;
                    route-filter 0.0.0.0/0 prefix-length-range /32-/32;
                    route-type internal;
                }
                then {
                    preference 9;
                    accept;
                }
            }
```

This policy changes the BGP route preference for /32 (loopback) prefixes to a value of nine, causing the BGP-LU route to be preferred over the IS-IS or OSPF route. This allows the BGP-LU label binding to be used for forwarding.

Intra-Region Transport

To provide transport for the BGP-LU LSP within an individual region, each region must have its own LSP mesh. When travelling within a region, the BGP-LU LSP will ride on top of these intraregional LSPs. To maintain consistency with the concept of regions, the intraregional LSP mesh will not extend beyond the regional boundaries.

For this intraregional LSP mesh, you can use LDP, RSVP, or a combination of both within a particular region. When choosing a protocol, you should consider factors such as ease of configuration, traffic management capabilities, network requirements, and personal preference. LDP, for example, typically requires fewer configurations but doesn’t offer the traffic engineering (TE) capabilities that RSVP offers. RSVP, on the other hand, requires more configurations and offers more traffic management tools.

Once the appropriate protocol has been selected, BGP-LU does not require any changes to the default behavior of LDP or RSVP. The configuration examples below are the minimum required to enable LDP and RSVP, and assume the network layout diagramed in Figure 7.

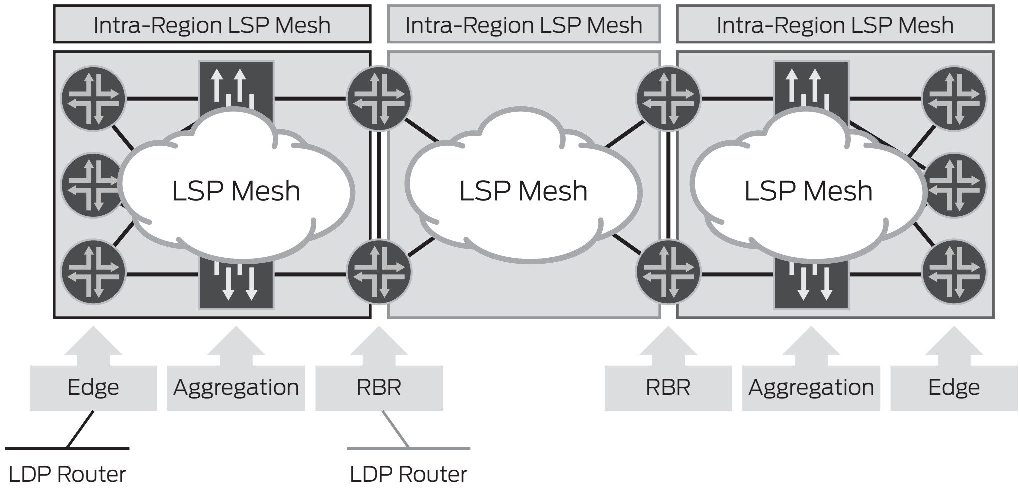

Figure 7: Sample network configuration for LDP and RSVP

In an LDP-based network, the configurations are the same for both the edge router and the RBR.

LDP configuration (both router edge and RBR):

```
Protocols {
ldp {
        interface ge-1/1/1.0;
    }
}
```

RSVP requires a slightly different configuration on the RBR and edge router.

RSVP configuration:

EDGE ROUTER

```
protocols {
    rsvp {
        interface ge-1/1/1.0;
    }
   mpls {
        label-switched-path Edge-to-RBR
{
            to 10.200.15.10;
        }
        interface ge-1/1/1.0;
    }
}
```

REGIONAL BORDER ROUTER (RBR)

```
protocols {
    rsvp {
        interface ge-1/1/1.0;
    }
   mpls {
        label-switched-path RBR-to-Edge
{
            to 10.200.15.11;
        }
        interface ge-1/1/1.0;
    }
}
```

Inter-Region Transport

BGP-LU

As described above, BGP-LU is the label signaling and routing protocol that provides the edge-to-edge or PE-to-PE reachability. The "hops" of the BGP-LU LSP do not need to be adjacent, just like two neighbors that form a unicast internal BGP (IBGP) session do not need to be adjacent. Rather, the BGP-LU LSP hops are those routers that participate in BGP-LU and are in the forwarding path. In this example, the BGP-LU hops are the region A edge, region A RBR, region B RBR, and region B edge. As noted above, the transit routers within each region do not require any awareness of, or participation in, BGP-LU—one of the reasons BGP-LU scales so well. Below we will provide configurations for the edge and RBR routers in the blue region A, as illustrated in Figure 8 The configuration would be similar in the red region so it has been omitted for brevity.

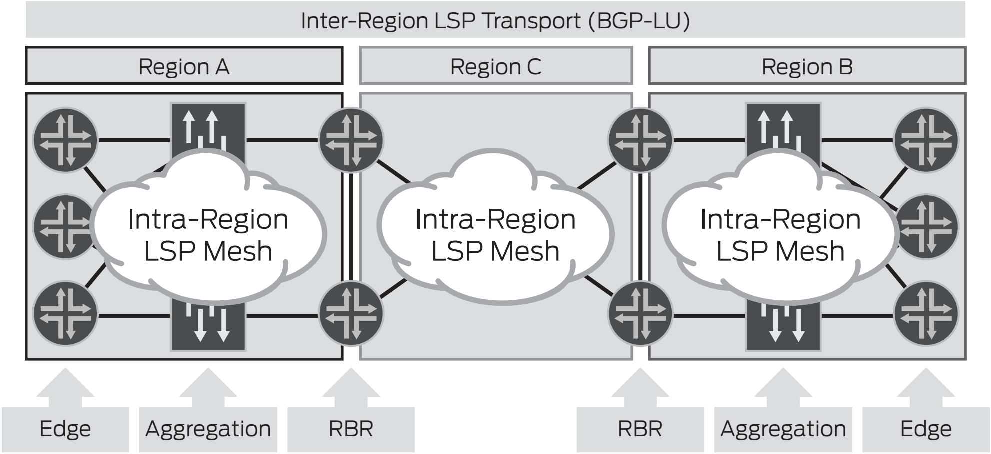

Figure 8: BGP-LU configuration

Region A Edge Router Configurations

Edge routers need configuration under three command-line interface (CLI) hierarchies: Routing-Options, Policy-Options, and Protocols BGP.

First, interface routes are copied to the inet.3 routing table, so that the loopback address can be exported with a label.

Routing-Options configuration:

```
routing-options {
    interface-routes {
        rib-group inet ifrg-inet0-to-inet3;
    }
    rib-groups {
        ifrg-inet0-to-inet3 {
            import-rib [ inet.0 inet.3 ];
        }
    }
    router-id 10.10.10.199;
    autonomous-system 100;
}
```

Next, policy is created to export only the loopback addresses.

Policy-Options configuration:

```
policy-options {
    policy-statement advertise.lo0 {
        term 1 {
            from {
```

```
                protocol direct;
                route-filter 0.0.0.0/0 prefix-length-range /32-/32;
            }
            then accept;
        }
        term 2 {
            then reject;
        }
    }
}
```

Finally, we add the BGP-LU configuration.

```
Protocols BGP configuration:
bgp {
    group Edge.RegionA {
        type internal;
        local-address 10.10.10.199;
        family inet {
            labeled-unicast {
                rib {
                    inet.3;
                }
            }
        }
        export advertise.lo0;
        neighbor 10.10.10.138;
    }
}
```

Region A Regional Border Router Configuration

Likewise, RBRs also need configurations under Routing-Options, Policy-Options, and Protocols BGP.

The Routing-Options configuration is the same as the edge router configuration above, so it has been omitted for brevity. Please refer to Routing-Options configuration above.

RBR Policy-Options configuration:

```
policy-options {
    policy-statement advertise.lo0 {
        term self {
            from {
                protocol direct;
                route-filter 0.0.0.0/0 prefix-length-range /32-/32;
            }
            then accept;
        }
    }
    policy-statement nhs-all {
        term 1 {
            then {
                next-hop self;
            }
        }
    }
}
```

The BGP configuration is as follows. Edge routers are configured as Route-Reflector clients to the RBRs. All RBRs are then fully meshed. To avoid a full mesh in region C, it is possible to use a route reflector in this region as well. The configuration below does not account for a route reflector in region C.

BGP configuration:

```
bgp {
    group Edge.RegionA.clients {
        type internal;
        local-address 10.10.10.138;
        family inet {
            labeled-unicast {
                rib {
                    inet.3;
                }
            }
        }
        export nhs-all;
        cluster 10.10.10.138;
        neighbor 10.10.10.199;
    }
    group RegionC {
        type internal;
        local-address 10.10.10.138;
        family inet {
            labeled-unicast {
                rib {
                    inet.3;
                }
            }
        }
        export [ advertise.lo0 nhs-all ];
        neighbor 10.10.10.152;
        neighbor 10.10.10.154;
    }
}
```

MPLS PE

MPLS PE routers participate in the intra-region LSP mesh, which, as described earlier, can either be LDP, RSVP, or a combination of both. MPLS PE routers typically participate in unicast BGP, learning prefixes via External BGP (E-BGP) and announcing them to the local unicast route reflectors, which in most cases will also be the RBRs (see following section on route reflection for more detail). The local RBR advertises the unicast BGP route to the remote RBRs, which in turn “reflect” or advertise the unicast BGP prefixes to its local route reflector clients—the local PEs. The RBR-to-RBR unicast BGP sessions can be fully meshed or a layer of route reflector hierarchy can be introduced to improve scaling, if required.

Figure 9 demonstrates the BGP configurations of the edge and RBR routers in each region. The edge router in region A has an E-BGP session to 11.1.199.2 in AS 199, and the edge router in region B has an E-BGP session to 11.1.200.2 in the AS 200. An MPLS PE is the ingress for the intra-region LSP as well as for the BGP-LU LSP. The MPLS PE has a double push label operation. The inner or first label is derived from BGP-LU and identifies the remote PE loopback. The outer or second label is derived from LDP or RSVP and identifies the local RBR.

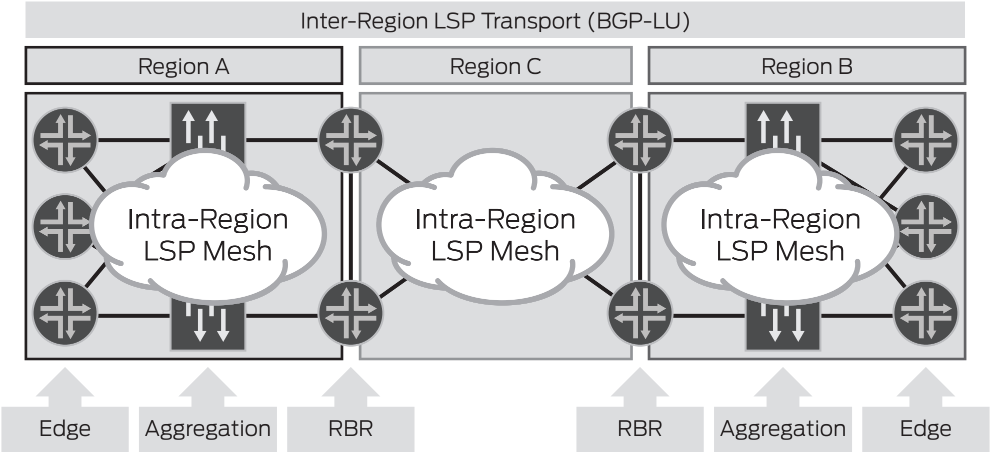

Figure 9: MPLS PE configuration

MPLS PE region A configuration:

```
bgp {
    group E-BGP {
        peer-as 199;
        neighbor 11.1.199.2;
    }
    group edge-A-to-RBR-A {
        type internal;
        local-address 10.10.10.199;
        export nhs-all;
        neighbor 10.10.10.197;
    }
}
```

RBR region A configuration:

```
bgp {
    group RBR-A-to-RBR-B {
        type internal;
        local-address 10.10.10.197;
        neighbor 10.10.10.198;
    }
    group RBR-A-to-edge-A {
        type internal;
        local-address 10.10.10.197;
        cluster 10.10.10.197;
```

```
        neighbor 10.10.10.199;
    }
}
```

RBR region B configuration:

```
bgp {
   group RBR-B-to-RBR-A {
       type internal;
       local-address 10.10.10.198;
       neighbor 10.10.10.197;
   }
   group RBR-B-to-edge-B {
        type internal;
        local-address 10.10.10.198;
        cluster 10.10.10.198;
        neighbor 10.10.10.200;
    }
}




MPLS PE region B configuration:
bgp {
    group E-BGP {
        peer-as 200;
        neighbor 11.1.200.2;
    }
    group edge-B-to-RBR-B {
        type internal;
        local-address 10.10.10.200;
        export nhs-all;
        neighbor 10.10.10.198;
    }
}
```

IP-Only PE

If the PE router is IP-only and does not support MPLS, it cannot support BGP-LU. In this case, the unicast BGP configuration for the IP PE is the same as the MPLS PE, but the loopbacks of IP PEs will be advertised into BGP-LU by the local RBR. Without the ability to learn BGP-LU routes, the IP-only PE will, by default, be unable to reach remote PEs. This can be resolved by configuring the RBRs to advertise a default route to the IP PE with the RBR as the next hop. This ensures that the IP PE can reach destinations outside of its local region and still benefit from the BGP-LU infrastructure.

The configuration below illustrates the policy that allows the RBR to identify and advertise the IP PE loopbacks into BGP-LU. The IP PE loopback needs to be copied into inet.3 from the IGP. For this configuration, we have used IS-IS as the IGP.

RBR Configuration with IP only PE

IGP configuration:

```
isis {
      rib-group inet ifrg-inet0-to-inet3;
}
```

Routing-Options configuration:

```
routing-options {
   interface-routes {
                rib-group inet ifrg-inet0-to-inet3;
   }
   rib-groups {
          ifrg-inet0-to-inet3 {
                 import-rib [ inet.0 inet.3 ];
           }
   }
   router-id 10.10.10.197;
   autonomous-system 100;
}
```

Policy-Options configuration:

```
policy-options {
    policy-statement advertise.lo0 {
        term self {
            from {
                protocol [direct isis];
                route-filter  10.10.10.199/32 exact;
                route-filter 0.0.0.0/0 prefix-length-range /32-/32;
            }
            then accept;
        }
    }
    policy-statement nhs-all {
        term 1 {
            then {
                next-hop self;
            }
        }
    }
}
```

BGP-LU configuration:

```
bgp {
    group RegionC {
        type internal;
        local-address 10.10.10.197;
        family inet {
            labeled-unicast {
                rib {
                    inet.3;
                }
            }
        }
        export [ advertise.lo0 nhs-all ];
        neighbor 10.10.10.198;
        neighbor 10.10.10.154;
    }
}
```

VPN PE

In this section, we will review the configurations necessary for a PE router providing either L3VPN or L2VPN services to its customers. As with the other PE types, we recommend that the VPN PEs exchange routes via a route reflector hierarchy. The VPN route reflector hierarchy should align with the unicast and BGP-LU route reflector hierarchy—that is, the RBRs should perform route reflection for the VPN PEs. The reasons are the same; this eliminates the need for edge-to-edge BGP sessions and contributes to greater scalability.

A VPN PE performs three label push operations. The first label identifies the VPN and is derived from VPN BGP. The second label identifies the remote VPN PE loopback and is derived from BGP-LU (a different protocol family than VPN BGP). And the third label identifies the local RBR and is derived from LDP or RSVP. The configurations for RBR-A, region C, and region B have been omitted for brevity. Those additional configurations would follow the same route reflector design as described in the MPLS PE section.

VPN PE in Region A

BGP configuration:

```
bgp {
    group l3vpn-BGP {
        type internal;
        local-address 10.10.10.199;
        family inet-vpn {
            unicast;
        }
        neighbor 10.10.10.197;
    }
}
```

Routing-Instance configuration:

```
routing-instances {
    R197-vrf-1 {
        instance-type vrf;
        interface ge-5/1/2.1999;
        route-distinguisher 10.10.10.199:1;
        vrf-target target:199:1;
        vrf-table-label;
        protocols {
            bgp {
                group eBGP {
                    type external;
                    peer-as 2199;
                    neighbor 11.2.199.2;
                }
            }
        }
    }
}
```

Non-VPN IPv6 PE (6PE)

The configuration shown here is for a PE router that provides non-VPN IPv6 services to the customer. In this case, the customer facing interfaces are configured with IPv6 addresses, but the core facing interface is configured with an IPv4 address. The BGP configuration here is similar to the MPLS PE—the IPv6 traffic is tunneled through IPv4 intra-region MPLS LSPs and hence needs a knob under the MPLS protocol hierarchy.

Below we provide the necessary steps to configure IPv6 on the 6PE router in region A. Configuration would be similar for all other 6PE routers in other regions in the network.

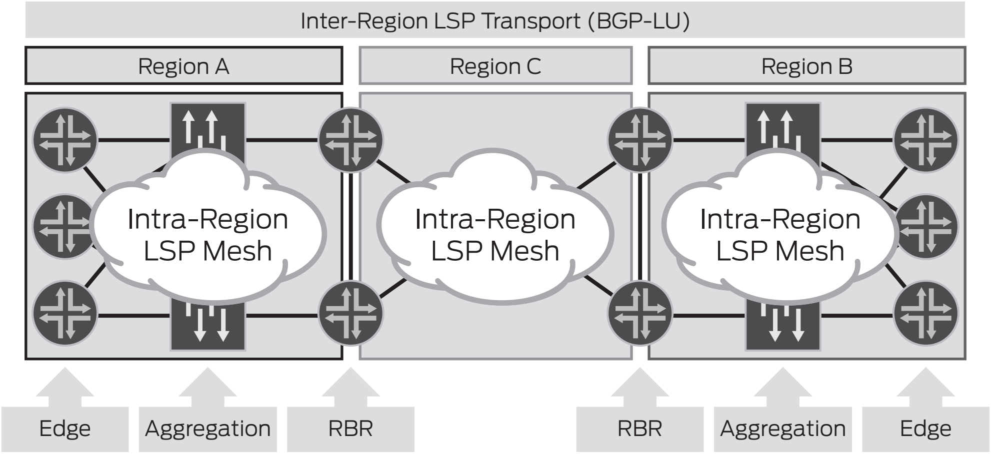

Figure 10: Topology considered for 6PE configuration below

6PE interface configuration:

```
interfaces {
    xe-0/0/0 {
    description “Core Facing Interface”
        unit 49 {
            vlan-id 49;
            family inet {
                address 10.1.1.21/30;
            }
            family iso;
            family inet6;
            family mpls;
        }
    }
    ge-5/1/2 {
    description “Customer Facing Interface”
        unit 49 {
            vlan-id 49;
            family inet6 {
                address 2001:0100:0049::1/126;
            }
        }
    }
    lo0 {
        unit 0 {
            family inet {
                address 10.10.10.49/32;
            }
            family iso {
                address 49.0001.0100.1001.0049.00;
```

```
            }
            family inet6 {
                address 2001:0100:0:1:0010:0010:0010:0049/128;
            }
        }
    }
}
```

Routing options:

```
routing-options {
    interface-routes {
        rib-group inet ifrg;
    }
    rib-groups {
        ifrg{
            import-rib [ inet.0 inet.3 inet6.3];
        }
    }
}
```

Policy options:

```
policy-options {
    policy-statement advertise.lo0 {
        term 1 {
            from {
                family inet6;
                route-filter 0::0/0 prefix-length-range /128-/128;
            }
            then accept;
        }
    }
}
```

MPLS configurations:

```
mpls {
    ipv6-tunneling;
    label-switched-path Edge-to-RBR {
        to 10.10.10.138;
    }
    interface xe-0/0/0.49;
}
```

BGP configurations:

```
bgp {
    group Edge-to-Customer {
        local-address 2001:0100:0049::1;
        family inet6 {
            unicast;
        }
        peer-as 49;
        neighbor 2001:0100:0049::2;
    }
    group Edge-to-RBR{
        type internal;
        local-address 10.10.10.49;
        family inet6 {
            labeled-unicast {
                 rib {
                inet6.3;
            }
        }
        export advertise.lo0;
        neighbor 10.10.10.138;
    }
    group Edge-to-Edge {
        type internal;
        local-address 10.10.10.49;
        family inet6 {
            unicast;
        }
        neighbor 10.10.10.50;
    }
}
```

Route Reflection

Using route reflection for BGP-LU prefix advertisements is a requirement if you want to achieve the scale and performance advantages of multiregional networks. Route reflectors eliminate the need for edge-to-edge IGP routing and LDP/RSVP LSP state, and therefore increase scale and performance.

In most cases, it is preferable for unicast BGP advertisements to use the same route reflection design as BGP-LU advertisements to avoid the need for interregional IGP routing. Typically, this means that RBRs should perform route reflection duties for both BGP-LU and unicast BGP.

However, there are some cases where it might not be possible or preferred to have the RBR perform the role of unicast route reflector. One such example is in a network that leverages a consolidated unicast route reflector design that uses a centralized route reflector pair serving multiple BGP-LU regions (rather than region-specific located route reflectors). Using a consolidated route reflection design can improve scale and performance by reducing the total number of route reflectors in a network (which also reduces the number of path advertisements).

In networks that leverage a consolidated route reflector design, the unicast route reflector resides outside of a BGP-LU region, so there must be a way to provide connectivity from that region to the remote unicast route reflector. Using a BGP-LU route to establish a unicast BGP session between two routers may be possible with extensive policy in some cases, but potentially introduces complexity that outweighs the benefit.

Instead, we recommend “leaking” the consolidated route reflector loopback prefixes between regions via the IGP. This allows a unicast BGP session to be established without interfering with the BGP-LU advertisements and route selection.

Conclusion

As the use of MPLS becomes increasingly widespread, implementing a multiregional network using BGP-LU can help improve network scale, manageability, and convergence and restoration times. The techniques and technologies outlined in this document are a part of Juniper's overall commitment to rapid convergence and increased scale in IP/MPLS networks, and the configurations outlined above will help you optimize your Juniper routers in multiregional network deployments.

Additional Resources

Seamless MPLS white paper: **www.juniper.net/us/en/local/pdf/whitepapers/2000316-en.pdf**

About Juniper Networks

Juniper Networks, Inc. is the leader in high-performance networking. Juniper offers a high-performance network infrastructure that creates a responsive and trusted environment for accelerating the deployment of services and applications over a single network. This fuels high-performance businesses. Additional information can be found at **www.juniper.net**.

Corporate and Sales Headquarters
Juniper Networks, Inc.
1194 North Mathilda Avenue
Sunnyvale, CA 94089 USA
Phone: 888.JUNIPER (888.586.4737)
or 408.745.2000
Fax: 408.745.2100
www.juniper.net

APAC Headquarters
Juniper Networks (Hong Kong)
26/F, Cityplaza One
1111 King's Road
Taikoo Shing, Hong Kong
Phone: 852.2332.3636
Fax: 852.2574.7803

EMEA Headquarters
Juniper Networks Ireland
Airside Business Park
Swords, County Dublin, Ireland
Phone: 35.31.8903.600
EMEA Sales: 00800.4586.4737
Fax: 35.31.8903.601

To purchase Juniper Networks solutions, please contact your Juniper Networks representative at 1-866-298-6428 or authorized reseller.

8020013-001-EN Jan 2010

♻ Printed on recycled paper