BNG Geo Redundancy

This chapter provides information about support of geographical redundancy through subscriber redundancy groups (SRGs) and session redundancy groups (SERGs).

Table 1: Feature History for Establishing Geo Redundancy

Release	Modification
Release 5.2.2	Introduced BNG geo redundancy.
Release 5.3.1	Geo redundancy support for PPPoE sessions was added.
Release 5.3.3	Peer route disable functionality was added.

This chapter covers these topics:

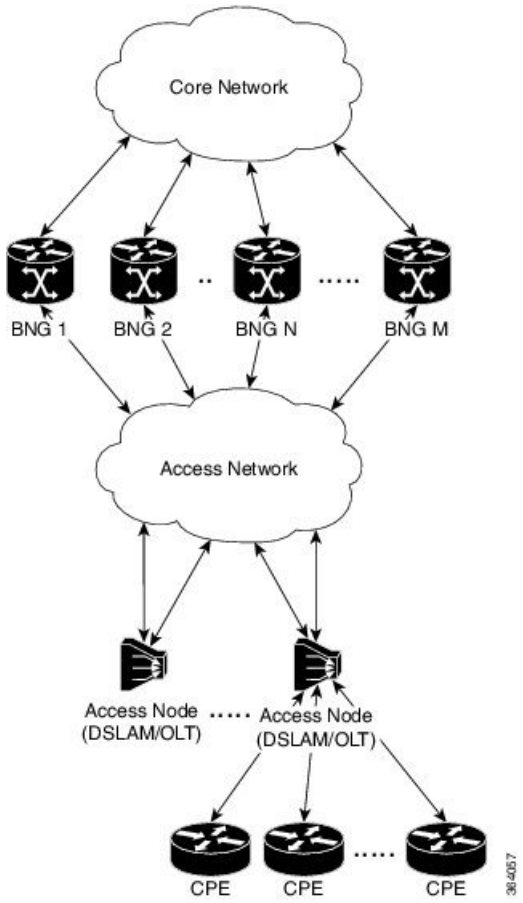
Geo Redundancy Overview

To provide redundancy for the subscriber sessions, BNG supports Geographical Redundancy across multiple BNGs, without having any L1 or L 2 connectivity between them. The BNG routers may be located in multiple geographical locations, and they have L3 connectivity over a shared core network through IP or MPLS routing.

Geo redundancy feature is supported for IPoE DHCP-triggered (IPv4, IPv6 and dual-stack) sessions and PPPoE (PTA and LAC) sessions.

This figure depicts a BNG geo redundancy deployment network model:

Figure 1: BNG Geo Redundancy Deployment Network Model



The redundancy pairing between BNG routers work by synchronizing the state from the master (active) to the slave (backup).

Geo redundancy works in conjunction with any of the access technologies. The CPEs are agnostic to redundancy; they see only one BNG or gateway. The access nodes are dual or multi-homed for redundancy using a variety of technologies based on the service provider network design and choices. Multi-chassis Link Aggregation (MC-LAG), dual-homed (Multiple Spanning Tree - Access Gateway or MST-AG), Ring (MST-AG or G.8032), xSTP and seamless MPLS (pseudowires) are a few such access networks.

Subscriber Redundancy Group (SRG)

Geo redundancy for subscribers is delivered by transferring the relevant session state from master BNG to slave BNG which can then help in failover (FO) or planned switchover (SO) of sessions from one BNG to another. Subscriber Redundancy Group (SRG) which is a set of access-interface (or a single access-interface) is introduced in BNG, and all subscribers in an SRG would FO or SO as a group.

The SRG has two modes of operation:

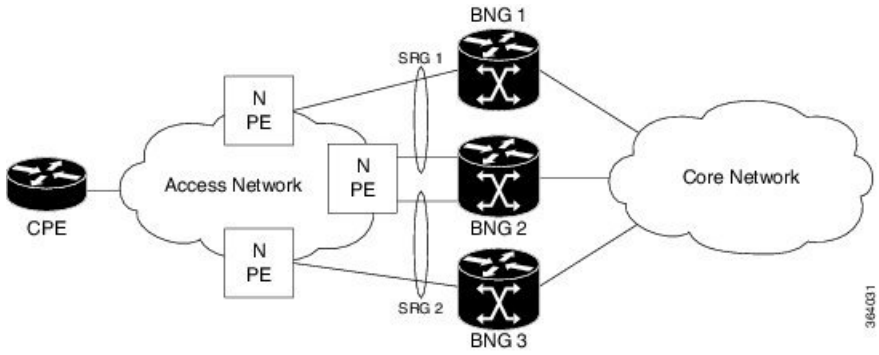
- Hot-standby
- Warm-standby

Currently BNG geo redundancy supports only the hot-standby slave mode. This is achieved by a 1:1 mirroring of subscriber session state from the master to the slave where the entire provisioning is done before the FO or SO. The sessions provisioned on slave is in sync with the set up on the master. Because the data plane is already set up for sub-second traffic impact, there is minimal action on switchover in the case of hot-standby mode and therefore, it is suitable for subscribers requiring high service level agreement (SLA). With appropriate capacity planning, the sessions can also be distributed across multiple BNGs to achieve an M: N model. The master-slave terminology is always in the context of a specific SRG; not for the BNG device as a whole.

Note Even after the Subscriber Redundancy Group (SRG) configuration is removed from the slave node, the CPE continues to receive ARP replies from both the master node and the slave node. This results in the network functioning in an uncertain manner. In order to avoid this uncertainty, shut down the access interface (that which corresponds to the slave node from which the configuration is removed) before removing the SRG configuration from the slave.

This figure depicts a typical BNG subscriber redundancy group (SRG):

Figure 2: BNG Subscriber Redundancy Group



SRG Virtual MAC

For seamless switchover between two BNGs, the L2-connected CPE devices must not detect change in gateway MAC and IPv4 or IPv6 addresses. The access technology like MC-LAG uses the same MAC address on both BNGs with active-standby roles, providing seamless switchover. Where MAC sharing is not provided by the access technology or protocol (like MST-AG, G.8032), the BNG SRG virtual MAC (vMAC) must be used. vMAC is configured as global MAC prefix or per SRG. This is integrated with BNG's dynamic master or slave role negotiation; additional protocols like VRRP or HSRP is not needed. vMAC (and its derived IPv6 link-local address) is used for control protocol exchanges (for example, ARP, ND, DHCP, PPPOE and so on) and data traffic for subscriber sessions or services only. It allows real port MAC to be used for Ethernet protocols (like E-OAM, xSTP, G.8032 and so on) that are leveraged by the SRG for doing failure detection, recovery and MAC Flush.

Session Distribution Across SRG

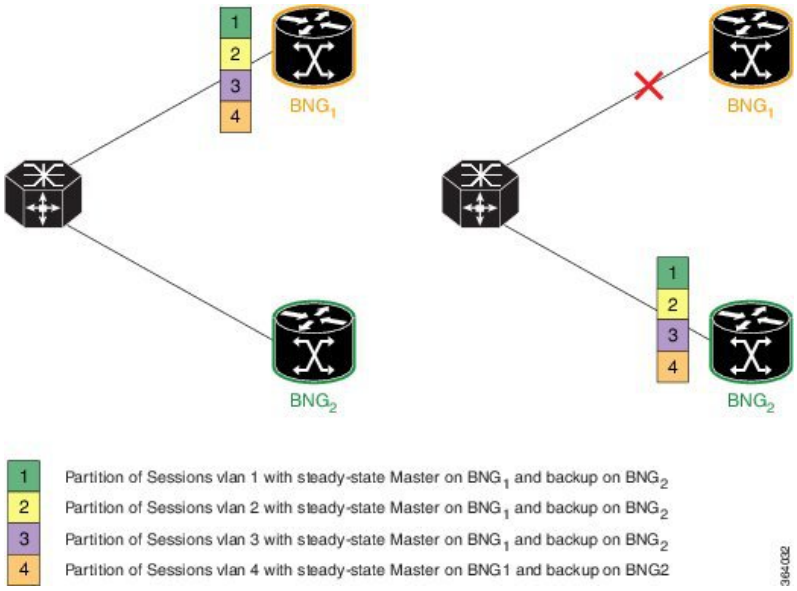
The session distribution across SRGs can be in either of these modes:

- Active-standby mode:

  In this mode, a dedicated backup BNG can be a slave for multiple SRGs from different active BNGs which are masters for those respective SRGs.

This figure shows an active-standby mode of session distribution across SRGs:

Figure 3: Active-standby Mode of Session Distribution



In figure a:

- Sessions are associated with partitions (VLAN 1, 2, 3 and 4) on BNG1, with each VLAN mapped to separate SRG configured as master role.
- BNG2 acts as backup for all VLANs.
- Each VLAN has 8K sessions terminated on it.

In figure b:

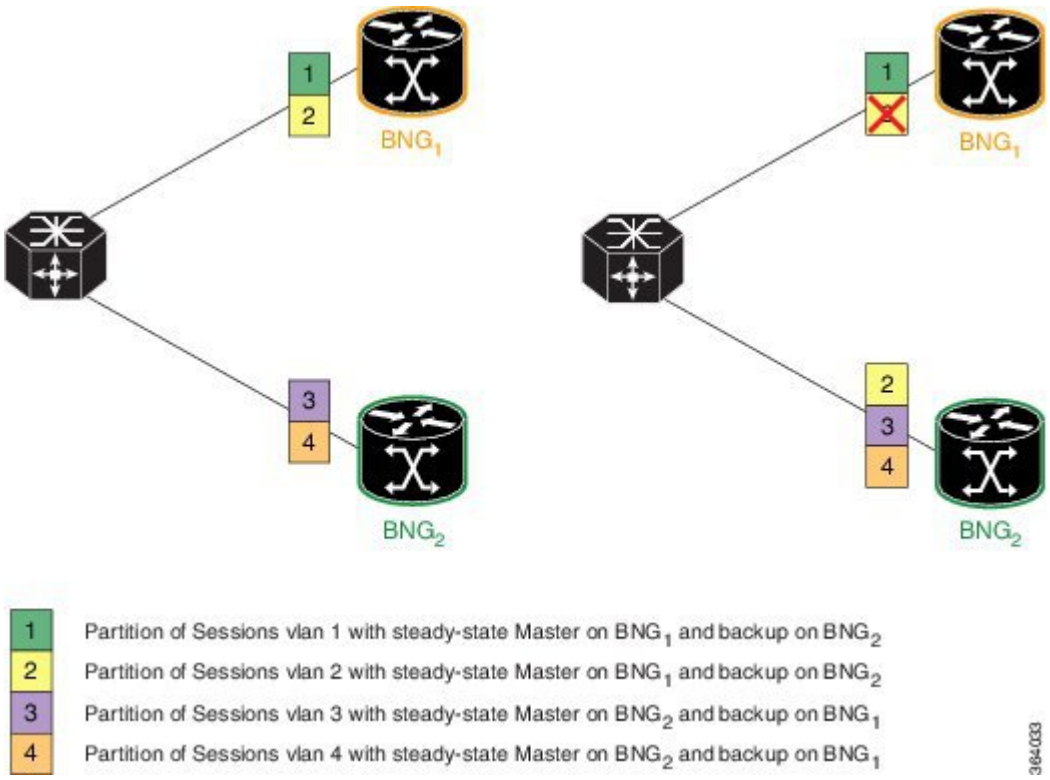
- An interface failure gets detected (using object-tracking of the access-interface) through MC-LAG.
- MC-LAG and SRG for each VLAN on BNG2 gets the master role.
- All 32K sessions are switched to BNG2.
- BNG2 sees a session termination count of 32K.

- Active-active mode:

  In this mode, a BNG can be master for one SRG and a slave for another SRG at the same time.

  This figure shows an active-active mode of session distribution across SRGs:

Figure 4: Active-active Mode of Session Distribution



In figure a:

- Sessions are associated with partitions (VLAN 1, 2) on BNG1, with each VLAN mapped to separate SRG configured as master role.
- Sessions are associated with partitions (VLAN 3, 4) on BNG2, with each VLAN mapped to separate SRG configured as master role.
- Each VLAN has 8K sessions terminated on it.
- Each BNG has 16K sessions terminated on it.

In figure b:

- The interface associated with VLAN 2 on BNG1 goes down.
- Sessions associated with partitions (VLAN 2) on BNG1 are switched to BNG2.
- BNG1 sees a session termination count of 8K and BNG2 sees a session termination count of 24K.

Benefits of BNG Geo Redundancy

Major benefits of BNG Geo Redundancy include:

- Supports various redundancy models such as 1:1 (active-active) and M:N, including M:1.
- Provides flexible redundancy pairing on access-link basis.

- Works with multiple access networks such as MC-LAG, dual-home and OLT rings.
- Supports various types of subscribers such as IPv4, IPv6 and dual-stack IPoE sessions.
- Works for RP (bundle and virtual access-links) based subscribers.
- Provides failure protection to access link failures, LC failures, RP failures and chassis failures.
- Performs automatic switchovers during dynamic failures or planned events such as maintenance, upgrades and transitions.
- Co-exists with other high availability (HA) or redundancy mechanisms.
- Does switchover of the impacted session group only; other session groups remain on the same BNG.
- Provides fast convergence and rapid setup of sessions, with minimal subscriber impact during switchover.
- Provides automatic routing convergence towards core and efficient address pool management.
- Provides seamless switchover for subscriber CPE without the need for any signaling.
- Integrates with RADIUS or policy and charging rule function (PCRF) systems.
- Provides minimal to zero incremental load on back end servers and PCRFs during normal operations and switchover.
- Does not impact session scale and call-per-second (CPS) during normal operation.

Supported Features in BNG Geo Redundancy

Supported Features in BNG Geo Redundancy

These access topologies are supported:

- MC-LAG topology (recommended only for IPv4 BNG sessions).
- Dual-home bundle interfaces with SRG vMAC using CFM or EFD fault detection and MST-AG for blocking.
- Ring bundle interfaces with SRG vMAC using CFM or EFD fault detection and MST-AG for blocking.
- Other access topologies and design variations may also be used for this feature.

These base geo redundancy features are supported:

- RP subscribers.
- Multiple SRG groups to different peer routers.
- Setting up peering statically through IPv4 or IPv6 TCP sessions.
- Hot-standby mode for slave (that is, subscribers provisioned in hardware on the slave as they are synchronized).
- Dynamic role negotiation between peers.
- Manual SRG switchover through command line interface (CLI).

- Dynamic failure detection using object tracking (link up-down, route and IPSLA tracking).
- Hold timer for dynamic switchover or switchback.
- Protocol bindings alone synchronized to slave; whereas AAA authorization for subscriber profile download performed by slave.
- Full BNG scale support (that is, half the scale number with redundancy).
- G.8032 (dual-home and ring) access technologies.

These DHCP features are supported:

- DHCPv6 IA-NA and IA-PD support for L2 connected sessions.
- DHCPv4 support for L2 connected sessions.
- DHCPv4 or DHCPv6 dual-stack support.
- DHCP proxy mode.
- Session initiation through DHCPv4 or DHCPv6 protocol.

Unsupported Features and Restrictions for BNG Geo Redundancy

This section lists the unsupported features and restrictions for BNG geo redundancy.

These are not supported in BNG geo redundancy:

- IPoE packet-triggered sessions.
- Routed (L3 connected) sessions
- Multicast
- Use of Neighbor Discovery (ND - SLAAC) feature for subscribers.
- SRG for ambiguous VLAN BNG sessions.
- SRG between Cisco IOS XR 64-bit BNG node and 32-bit BNG node.
- Both RP and LC subscribers do not support enabling fast switchover for subscriber framed-routes.

These are planned to be fully qualified only in future releases of Cisco IOS XR Software:

- Warm-standby slave mode.
- Line card (LC) based subscribers (that is, using physical port sub-interfaces).
- DHCP server mode.
- Pseudowire Headend (PWHE)

BNG Geo Redundancy Configuration Guidelines

While configuring BNG geo redundancy, certain guidelines must be followed in these areas:

- BNG Configuration Consistency

- Access-link Integration
- Core Routing Integration
- RADIUS-PCRF Integration

BNG Configuration Consistency

- Geo redundancy feature infrastructure synchronizes individual subscriber session state from master to slave. But, it does not synchronize the BNG related configurations (namely dynamic-template, DHCP profiles, policy-maps, access-interface configurations, external RADIUS or DHCP server and so on).
- For successful synchronization and setup of subscriber sessions between the two BNGs, it is mandatory that the relevant BNG configurations must be identical on the two routers and on the access-interfaces pairs in the SRG.
- While the access-interfaces or their types (or both) may vary between the paired BNGs, their outer-VLAN tag (that is, S-VLAN imposed by the access or aggregation devices) must be identical.
- Inconsistencies in base BNG or SRG configurations may result in synchronization failure and improper setup of sessions on the slave.

Access-link Integration

- You must use only those dual-homing techniques where one side is up or active, and the other side is down or standby. Both sides must not be up and forwarding traffic at the same time.
- You must use access-tracking mechanism under the SRG to ensure that its BNG role is always in synchronization with its access-link. Without this, the data or control traffic may get black-holed.
- The access-tracking object used by the SRG must be same as the one used in the routing configuration for conditional advertisement of the subscriber summary route(s) corresponding to that SRG's subscriber address or subnet pool(s).
- Including multiple access-links (which do not fail or switchover their roles) together into a single SRG may be challenging, unless mechanisms are implemented to ensure that all these links change state even when one of them fails.

Core Routing Integration

- Redistribution of individual subscriber routes into the routing protocol is not recommended because it slows convergence in failure or switchover events.
- Recommended design option is to conditionally advertise the summary static route for the subscriber address/subnet pool(s) of the SRG into the core routing protocol, through access-tracking.
- You can also advertise from both routers with different preferences and use various fast-reroute techniques.
- To avoid core routing changes in certain failure conditions, there are options to re-route the traffic from the slave to the master (for example, a tunnel or inter-chassis link) for transient or prolonged intervals.
- Routing convergence and its correlation with access failures or convergence is a key to overall end-to-end service impact for subscribers. Multiple options exist to achieve sub-second intervals.

RADIUS-PCRF Integration

The backend policy and charging rule function (PCRF) system must send the CoA message to both master and slave nodes. The message can be sent to the slave either at the same time as it is sent to master, or it can be sent after the slave takes over the master role and sends the Accounting START message.

From Cisco IOS XR Software Release R5.3.1 and later, the backend PCRF sytem need to send the CoA message only to the master node.

Session Sync

Once the session is up on the master node, the entire session information gets synced to the slave node. This includes dynamic synchronization of updates such as CoA or service logon. This is applicable from Cisco IOS XR Software Release R5.3.1 and later.

Setting up BNG Subscriber Redundancy Group

Guidelines in setting up SRG

Setting up SRG is subjected to these guidelines:

- The configurations and subscriber policies applied on the two routers (where the SRG access-interfaces are dual homing) must be identical to ensure seamless session mirroring and switchover.
    - SRG IDs (group IDs) must be same across BNGs.
    - Access-interface names or types need not be the same across routers.
    - Interface mapping-IDs must be same for the acces-interfaces across BNGs.
    - Server configurations (namely, RADIUS and DHCP configurations), IP pools, subscriber policies and templates must be identical across routers.
- The database of SRGs are scoped to a particular control plane instance (that is, at RP or LC node level). Therefore, you cannot form a single SRG with member links across LCs or with a mix of virtual interfaces (for example, bundles) and physical ports.

Setting up a BNG subscriber redundancy group (SRG) involves these steps:

- Enable BNG Geo-Redundancy:

```
subscriber redundancy
 source-interface loopback1
```

- Setup SRG and specify peer IPv4 or IPv6 address:

```
subscriber redundancy
 group 1
  peer 1.1.1.2
```

- Specify access-interfaces or VLANs, and mapping IDs:

```
subscriber redundancy
```

```
 group 1
  interface-list
    interface Bundle-Ether1.10 id 210
```

- Setup access object tracking for SRG and summary subscriber route:

```
track mclag-be1
 type line-protocol state
  interface bundle-ether1

subscriber redundancy
 group 1
  access-tracking mc-lag-be1

router static
 address-family ipv4 unicast
  200.0.0.0/16 Null0 track mc-lag-be1
```

Some optional configurations such as **preferred-role**, **slave-mode** and **hold-timer** also exist for SRG.

Geo Redundancy for PPPoE Sessions

BNG supports geo redundancy for PPPoE-PPP Termination and Aggregation (PPPoE-PTA) and PPPoE-L2TP Access Concentrator (PPPoE-LAC) sessions.

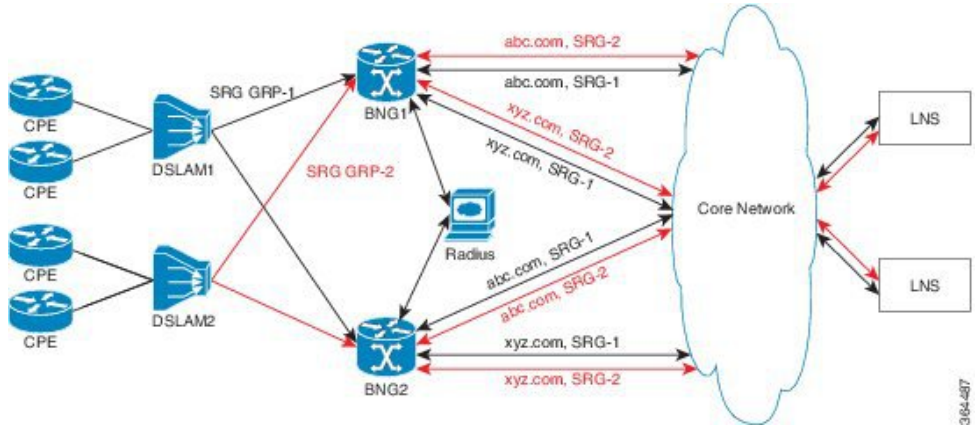
PPPoE-PTA Geo Redundancy

Geo redundancy behavior for the PPPoE-PTA sessions remains the same as for basic geo redundancy set up, except that the keepalives are disabled on the slave BNG node. The keepalives are sent only after the slave switches its role to master.

PPPoE-LAC Geo Redundancy

This figure shows a PPPoE-LAC Geo Redundancy set up with BNG

Figure 5: PPPoE-LAC Geo Redundancy Topology



For a PPPoE-LAC geo redundancy setup, the SRG is formed by grouping together the access-links on which LAC sessions are to arrive (co-exists with PTA). To enable SRG level redundancy switchover, tunnels for each SRG for each L2TP network server (LNS) must be setup. L2TP ensures that sessions belonging to different SRGs do not share the same tunnel even if they are going to the same LNS. The tunnel is set up on both master and slave nodes. By default, the tunnel is down on slave and it gets activated upon switchover. The BNG sync takes care of both tunnel and session-state sync from the master to the slave. The L2TP tunnel attributes and negotiated parameters are also synchronized through the BNG sync.

You must use this command in subscriber redundancy group configuration mode, to configure the source IP used for L2TP tunnel for subscribers coming under an SRG group:

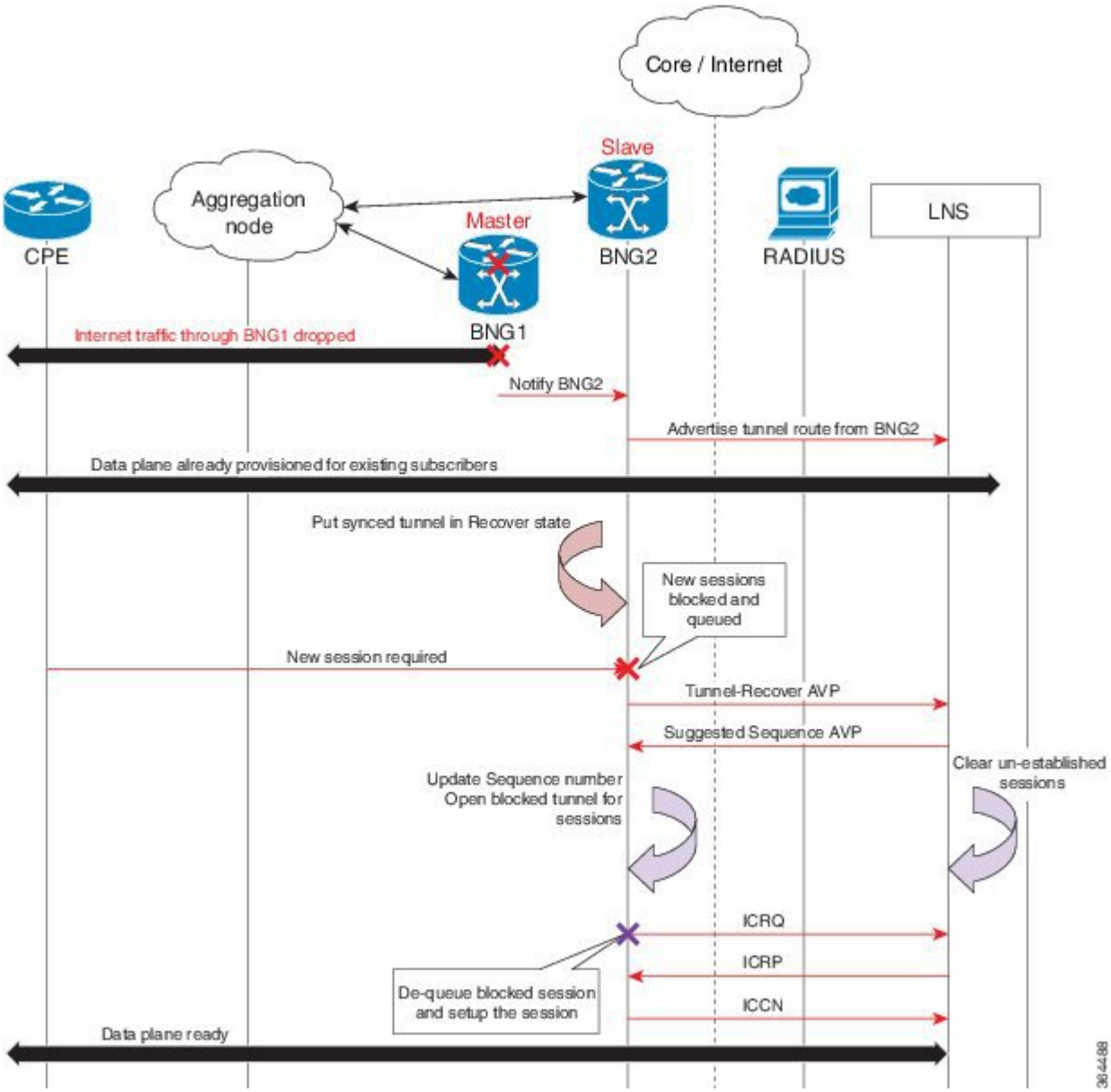
l2tp-source-ip *ipv4-address*

This ensures that there is a separate tunnel from each SRG group, in spite of having the same LNS.

PPPoE-LAC Session Switchover

This figure shows the call flow of PPPoE-LAC session switchover.

Figure 6: PPPoE-LAC Session Switchover



During switchover, the tunnel endpoint switches from the master (BNG1) to slave (BNG2) node as soon as the routing converges, and advertises the loopback address of slave (BNG2) to the LNS. The sessions and tunnels that are already provisioned on the data path on slave (BNG2) then seamlessly take over. The L2TP control plane on slave (BNG2) places the tunnel in re-sync state to recover the tunnel sequence number (Ns and Nr) during which only control messages are queued up for further processing. After the tunnel recovery, the LAC gets the sequence number from the LNS. The existing tunnels or sessions are not lost as the slave (BNG2) takes over. The signaling for the new session resumes and the queued requests also get processed. The unestablished sessions are then cleared off. For LNS, this switchover appears to be a convergence event where the tunnel has flapped.

Verification of Geo Redundancy for PPPoE Sessions

Listed below are some of the show commands that can be used to verify the Geo Redundancy configuration in BNG. For complete command reference, see the *Subscriber Commands*, *PPPoE Commands* and *PPPoE LAC-Specific Commands*, chapters in the *Cisco ASR 9000 Series Aggregation Services Router Broadband Network Gateway Command Reference*.

- **show subscriber redundancy group** *210*

```
Subscriber Redundancy Group ID: 210
  Description : <<not-configured>>

  Status                : Enabled
  Init-Role             : Master
  Negotiated-Role       : Master                Current-Role : Master

  Slave-mode            : Hot                      Hold Time : 15
  - - -
  - - -
Peer:
    11::2                                            Status : Established
    Role(Init/Neg/Cur): Slave/Slave/Slave
    Tracking Status   : Down
- - -
- - -
Switchover:
    Last Switchover   : 2014 Sep 12 07:12:11       Reason : Object Tracking Status
Change
- - -
- - -
Subscriber Session Statistics:
    Count             : 8000               Slave-Upd-Fail : 0
    Pending Update    : 0                  Pending Delete : 0
    Tunnel Count      : 0

Interface Count     : 1
    Bundle-Ether1.10                       Map-ID         : 210
```

- **show ppp interfaces**

```
Bundle-Ether2.1.pppoe16534 is up, line protocol is up
 SRG Role: Slave
  LCP: Open
     Keepalives enabled (60 sec, retry count 5)
     Local MRU: 1492 bytes
     Peer  MRU: 65531 bytes
  Authentication
     Of Peer: PAP (Completed as user1@domain.com)
     Of Us:   <None>
  IPCP: Open
     Local IPv4 address: 12.16.0.1
     Peer IPv4 address:  12.0.250.23
  IPv6CP: Initial
     Local IPv6 address: fe80::
     Peer IPv6 address:  fe80::
```

- **show pppoe interfaces**

```
Bundle-Ether2.1.pppoe16534 is Complete
```

```
 Session id: 16534
 Parent interface: Bundle-Ether2.1
 BBA-Group: BBA1
 Local MAC address: 0002.0003.0004
 Remote MAC address: 0000.6201.0103
 Outer VLAN ID: 10
 Tags:
   Service name: AGILENT
   Host-Uniq: 4 bytes, (000e0000)
 SRG-state: SRG-Standby
```

- **show vpdn**

```
RP/0/RSP0/CPU0:router# show vpdn session

SRG Role: Master
Subscriber label: 0x42, interface name: Bundle-Ether1.10.pppoe3
 user name: user1@lns2.com
 parent interface: Bundle-Ether1.10
 state: est last change: 00:01:01
 time to setup session:   0:2  (s:msec)
 conditional debug flags: 0
 L2TP data
  local end point: 11.1.1.1 remote end point: 19.9.9.2
  call serial number: 1970100002
  local tunnel id: 46813 remote tunnel id: 40849
  local session id: 36198 remote session id: 33437 remote port: 1701
  tunnel assigned id:
  tunnel client authentication id: LAC
  tunnel server authentication id: LNS
  tunnel authentication: disabled
  class attribute mask:
 Subscriber data
  NAS port id: 0/0/1/10
  NAS port type: Virtual PPPoE over VLAN
  physical channel id: 0
  Rx speed: 1000000000, Tx speed: 1000000000
 Configuration data
  table id: 0xe0000000, VRF id: 0x60000000, VPN id: 0:0
  VRF name: default
  dsl line info forwarding: disabled, l2tp busy timeout: 60
  TOS mode: default
```

Geo Redundancy Features

Peer Route Disable

Peer route disable is an enhancement in BNG geo redundancy whereby the user can disable the route on geo redundancy hot-standby peer. This disabling is so that the subscriber routes are not installed in the RIB even when the subscriber sessions are brought up on the standby peer. The subscriber routes are inserted into the RIB only when the BNG Geo-Redundancy state of peer changes to active. This ensures that only the master BNG, and not the slave BNG, routes the subscriber traffic in a scenario where access-interface is up on the standby peer. By disabling the routes, the hot-standby mode in BNG geo redundancy does not mandate the access-interface to be down on the standby peer any more.

To enable this feature, use the **peer route disable** command in subscriber redundancy group configuration mode.

Configuration Example

```
RP/0/RSP0/CPU0:router(config)# subscriber redundancy group 110
RP/0/RSP0/CPU0:router(config-subscr-red-group)# peer route-disable
```

Deployment Models for BNG Geo Redundancy

Multiple access networks are considered for BNG geo redundancy deployment scenarios. Some of the sample use cases are:

- Multi-chassis Link Aggregation (MC-LAG) - Two BNG boxes that are point-of-attachment (POA) devices, connected through MC-LAG either to a single Dual Homed Device (DHD) or to a DHD-pair using MC-LAG.
- Multiple Spanning Tree - Access Gateway (MST-AG):
  - Dual Homed Device using Bundle Interfaces - A single DHD with one bundle interface each to the two BNGs in active-active mode.
  - Ethernet Access Network-Ring - A physical ring (open or closed) that connects multiple OLTs (or L2 devices in general) to the two BNGs in active-active mode.