Graceful Restart, Non Stop Routing and IGP routing protocol timer Manipulation

Last updated: June 2008

Graceful Restart (GR) (also known as Non Stop Forwarding (NSF)) and Non Stop Routing (NSR) are two different mechanisms to prevent routing protocol reconvergence during a processor switchover.

Traditionally, when a networking device restarts, all routing peers associated with that device detect the device has gone down and routes from that peer are removed. The session is re-established when the device completes the restart. This transition results in removal and re-insertion of routes, which could spread across multiple routing domains. This was required because of the inability of the restarting device to forward traffic during the reload period. Today, dual processor systems which support Stateful Switch Over (SSO) or In-Service Software Upgrades (ISSU) can continue to forward traffic while restarting the control plane on the second processor. In this case, route removal and insertion caused by routing protocol restarts is no longer necessary, creating unnecessary routing instabilities, which are detrimental to the overall network performance. Graceful Restart and Non Stop Routing suppress routing changes on peers to SSO-enabled devices during processor switchover events (SSO or ISSU), reducing network instability and downtime.

GR and NSR, when coupled with SSO provide the foundation for fast recovery from a processor failure and allow the use of ISSU to perform software upgrades with little downtime. SSO is necessary to handle other non routing protocol related items needed for the router to operate following a switchover. These include syncing the complete router configuration, Cisco Express Forwarding (CEF) forwarding entries and other needed information to the standby processor.

Graceful Restart and Non Stop Routing both allow for the forwarding of data packets to continue along known routes while the routing protocol information is being restored (in the case of Graceful Restart) or refreshed (in the case of Non Stop Routing) following a processor switchover.

When Graceful Restart is used, peer networking devices are informed, via protocol extensions prior to the event, of the SSO capable routers ability to perform graceful restart. The peer device must have the ability to understand this messaging. When a switchover occurs, the peer will continue to forward to the switching over router as instructed by the GR process for each particular protocol, even though in most cases the peering relationship needs to be rebuilt. Essentially, the peer router will give the switching over router a "grace" period to re-establish the neighbor relationship, while continuing to forward to the routes from that peer. Graceful Restart is available today for OSPF, ISIS, EIGRP, LDP and BGP. Standards are defined for OSPF, ISIS, BGP and LDP to ensure vendor interoperability.

When Non Stop Routing is used, peer networking devices have no knowledge of any event on the switching over router. All information needed to continue the routing protocol peering state is transferred to the standby processor so it can "pick up" immediately upon a switchover.

NSR is available today in Cisco IOS for ISIS and BGP. Unlike Graceful Restart, Non Stop Routing uses more system resources due to the information transfer to the standby processor. Standards are not necessary in the case of Non Stop Routing, as the solution does not require any additional communication with protocol peers. NSR is desirable in cases where the routing protocol peer doesn't support the RFCs necessary to support Graceful Restart, however it comes at a cost of using more system resources than would be if the same session used Graceful Restart. Often a simple software upgrade on the peer will allow the use of Graceful Restart over Non Stop Routing.

Some protocols (BGP for instance) operate in a "hybrid" mode in Cisco IOS. In a typical PE type role, the majority of BGP routes will be learned from the network Route Reflectors over iBGP sessions, and those route reflectors are in the control of the operator. It is recommended to operate these sessions in GR mode and only run NSR to non managed CE type connections, since their software may not be easily changed to support GR. In addition, single CE connections generally have a relatively fewer number of routes, reducing the load on the router should NSR be necessary.

The following table shows various GR and NSR modes and the versions which support was added:

Table 1.

Protocol and Mode	Support starts with:
OSPF Cisco GR	Cisco IOS Software Release 12.0(22)S
OSPF IETF GR	Release 12.2(31)SB
ISIS IEFT GR	Release 12.0(22)S
ISIS NSR	Release 12.0(22)S
EIGRP GR	Release 12.2(25)S
LDP GR	Release 12.2(25)S
BGP GR	Release 12.0(22)S
BGP NSR	Release 12.2(25)SB

Specific CLI configurations may refer to the Graceful Restart process as "NSF", or Non-Stop Forwarding. The two terms should be considered equal and interchangeable.

Interior Gateway Protocol (IGP) timer manipulation refers to the practice of reducing HELLO and HOLD-TIME timers in Open Shortest Path First (OSPF) or Intermediate-System to Intermediate-System (ISIS), which reduces the time to detect a failed routing neighbor, which in turn promotes quicker convergence when a link or router fails. In both OSPF and ISIS, neighbor adjacency is maintained by the periodic transmission of HELLO packets. Both protocols support the concept that a neighbor should be declared unavailable if it does not transmit a HELLO within a certain time interval. For configuration purposes, OSPF refers to this as the "dead-interval" and ISIS uses the term "hold timer". Either are also commonly referred to as "dead-interval" or "holddown-time"

It is important to understand potential feature interaction considerations when using SSO or ISSU and IGP timer manipulation, as the two methodologies have seemingly conflicting goals.

Stateful Switchover and In Service Software Upgrade aim to preserve the flow of traffic through a router during failure or upgrade, whereas reducing routing protocol timers aims to quickly redirect traffic away from a failure. It is important to carefully consider the individual network design goals and establish precedence for redundancy. Reducing timers on point to point links is less common than multipoint, as link flaps will immediately signal the routing protocol to begin reconvergence.

Multipoint environments are typically the areas of the network where interior routing protocol timers are reduced.

Following a processor transition, whether from a Stateful Switchover or In Service Software Upgrade, the new active processor will require a period of housekeeping prior to sending out the first IGP protocol Hello packet. In order to effectively decide on the appropriate IGP routing protocol timers, this period of time should be understood. The timeframe to complete the housekeeping depends on feature use, scale, platform and processor type. The following example illustrates a simple methodology that can be used to measure these times and therefore be confident that either Graceful Restart or Non-Stop Routing will maintain the network topology as intended.

First, ensure the best log and debug timestamping is enabled:

```
service timestamps debug datetime msec
service timestamps log datetime msec
```

Next, enable the appropriate debug to show when the first HELLO is sent. This debug should be enabled on the standby routing processor prior to the switchover. For OSPF, enable "debug ip ospf hello". For ISIS, enable "debug isis adj-packets".

The following log example is from a basic configured Cisco 10000 running Release 12.2(31)SB10.

```
19:19:15.422: %RED-5-REDCHANGE: PRE A now Non-participant(0x1C03 =>
0x1C23)
19:19:15.422: %REDUNDANCY-3-SWITCHOVER: RP switchover
(PEER_REDUNDANCY_STATE_CHANGE)
19:19:15.562: %IPCOIR-5-CARD_DETECTED: Card type 1gigethernet-hh-1
(0x390) in slot 1/0
19:19:15.930: %IPCOIR-2-CARD_UP_DOWN: Card in slot 1/0 is up. Notifying
1gigethernet-hh-1 driver.
19:19:16.442: %LINEPROTO-5-UPDOWN: Line protocol on Interface
Ethernet0/0/0, changed state to up
19:19:16.754: %C10KGE1H-6-SFP_OK: Interface GigabitEthernet1/0/0,
1000BASE-SX Gigabit ethernet module (SFP) inserted
19:19:17.058: %RED-5-REDCHANGE: PRE B now Active(0x1C23 => 0x1421)
19:19:18.010: %LINK-3-UPDOWN: Interface Null0, changed state to up
19:19:18.010: %LINK-3-UPDOWN: Interface FastEthernet0/0/0, changed state
to up
19:19:18.010: %LINK-3-UPDOWN: Interface GigabitEthernet1/0/0, changed
state to up
19:19:18.010: OSPF: Send hello to 224.0.0.5 area 0 on
GigabitEthernet1/0/0 from 192.168.1.1
19:19:18.010: OSPF: Rcv hello from 192.168.1.2 area 0 from
GigabitEthernet1/0/0 192.168.1.2
19:19:18.010: OSPF: Immediate hellos sent on GigabitEthernet1/0/0
limited by number non-full neighbors 0
```

```
19:19:18.010: OSPF: End of hello processing
```

From the data in this example, the router sent its first hello packet 2.678 seconds after switchover. Based on this information, if a hello period of 1 second is desired, a dead-interval of at least 7 seconds would be appropriate to allow the routing topology to maintain its stability during and after a SSO or ISSU switchover. While 7 seconds may seem longer than needed, a worst case scenario should be considered (ie: the switchover occurred just before the next hello packet was schedule to be sent).

Routing protocol timer manipulation and Stateful Switchover can co-exist, as long as care is taken through using appropriate timers, so that the switchover doesn't interfere with the normal IGP session timeout values.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA

C22-487228-00   07/08