Deploying Routing Protocol Resiliency with Graceful Restart

Session RST-2004
Objective

• Graceful restart extensions introduced in routing protocols to improve network availability
• Design consideration for successful deployment
• Configuration commands

Agenda

• Introduction
• OSPF Graceful Restart (OSPF NSF)
• ISIS Graceful Restart (ISIS NSF)
• BGP Graceful Restart
• EIGRP NSF
• General Deployment Guidelines
Network Resiliency Tool Kit

**Application Level Resiliency**
- CoS/QoS, Diff-Serv TE, Global Server Load Balancing, Stateful NAT, Stateful IPSec

**Protocol Level Resiliency**
- Load Balancing, Route Reflector Redundancy, HSRP, Graceful Restart (GR) in BGP, OSPF NSF, ISIS NSF, LDP GR, MPLS, Route Dampening

**Link Level Resiliency**
- MLPPP, T1IMA, APS, Fast Re-route, Resilient Packet Ring, EtherChannel, LFI, POS Channel

**Device Level Resiliency**
- Redundant Processors (RP), Switch Fabric, Line Cards, Ports, Power, NSF/SSO, Control Traffic Prioritization

Focus of This Session...

**Application Level Resiliency**
- CoS/QoS, Diff-Serv TE, Global Server Load Balancing, Stateful NAT, Stateful IPSec

**Protocol Level Resiliency**
- Load Balancing, Route Reflector Redundancy, HSRP, Graceful Restart (GR) in BGP, OSPF NSF, ISIS NSF, LDP GR, MPLS, Route Dampening

**Link Level Resiliency**
- MLPPP, T1IMA, APS, Fast Re-route, Resilient Packet Ring, EtherChannel, LFI, POS Channel

**Device Level Resiliency**
- Redundant Processors (RP), Switch Fabric, Line Cards, Ports, Power, NSF/SSO, Control Traffic Prioritization
Problem Definition

- Requirements for network availability is high in networks where Aggregation routers can become a single point of failure

  Business critical traffic, like voice, is carried on an IP network and network availability is important

Aggregation Edge Is Vulnerable

- Single point of failure for 100’s to 1000’s of circuit terminations
- Redundant components used: fans, power, fabric, route processors
**Redundant Route Processor Details**

- **Active Route Processor**
  - Standby RP is in "hot standby".
  - Chassis and line card states sync.
  - Line protocols—ATM, FR etc. in sync.
  - Forwarding table is in sync.

- **Standby Route Processor**
  - Standby RP is in "hot boot".
  - Startup and running config. in sync.
  - During failover line protocols reset.
  - Forwarding table is NOT in sync.

- **RPR**
  - Standby RP is in "cold boot".
  - Start up configs are in sync.
  - Running configs are not in sync.
  - During fail over standby resets.
  - line cards and restarts system.

**Non-Stop Forwarding**

- **Active RP Stores Routes in Routing Information Base**
  - Forwarding Information Base
  - RP and LCs Forward Packets Based on FIB

- **RP Dies**
  - FIB is transferred to Line Cards and/or Standby RP
  - Standby RP Becomes Active
  - Standby RP or Line Card Forwards Packets while Converging

- **FIB is preserved for Non-Stop**
- **Forwarding (NSF)**
- **Graceful restart/NSF mechanism used to re-converge**
How Does the Redundant RP Handle Routing Protocols?

• Using the Graceful restart (or NSF) mechanisms

What Is Graceful Restart?

• Under certain failure conditions when a routing process restarts it seeks the help of peer routers to re-learn routes and resume neighbor relationship while:
  a) The data traffic continues to be routed between the restarting router and peers
  b) The peer does not pre-maturely declare the restarting router dead
Cisco Implementation of Graceful Restart

• The failure conditions are applicable in platforms with dual Route Processors (RP) and the conditions force a switch over from active to standby RP
• Restarting router is in SSO mode
• Data forwarding tables are preserved in peer router

Factors Causing Failure Conditions

• A fault condition that causes the active RP to crash or reboot—automatic switchover
• The user can force the switchover from active to standby—manual switchover
• The active RP is declared dead (not responding)—automatic switchover
What Do We Accomplish with this Feature?

- Increase network availability when the router is a single point of failure
- Increase network availability when business critical functions are likely victims of failure—ex: Net Commerce
- Reduce packet drops to almost zero!

Which Platforms Support these Features

- Dual RP routers currently supporting these features include 7500, 10000 and 12000
  More platforms along the way ex: 7300, 7600
- Peer routers that support 12.0(22)S, 12.2(T) and 12.2S (release dependent code)
Terms Frequently Used in this Session...

- NSF Capable Router (restarting router)
  A router that preserves its forwarding table and rebuilds its routing topology after an RP switch over; currently a dual RP router
  ex: Cisco 7500, 10000, 12000

- NSF Aware Router (peer)
  A router that assists an NSF capable during restart and can preserve routes reachable via the restarting router
  ex: Cisco 7200, 3600, 2600, 1700

- NSF Unaware Router
  A router that is not capable of assisting an NSF Capable router during an RP switchover

- NSF Capable Router is NSF Aware too!!!!

Relationship Building Exercise—1

GR (NSF/SSO) Capable Router

“\( I \) Can Preserve My Forwarding Table during Restart”

Agreement

GR (NSF) Aware Peer

During Restart
  – I Will Preserve My Forwarding Table
  – I Will Not Declare You Dead
  – I Will Not Inform My Neighbors
Relationship Building Exercise—2

GR (NSF/SSO) Capable Router
- I Have Restarted
- I Will Use Your Knowledge to Build My Database

GR (NSF) Aware Peer
- Restart Notification and Acknowledgement
- OK. I Acknowledge. I Will Stick to My Agreement
- Knowledge Transfer
- This Is My Knowledge of the Network
- Updates

Networks without NSF/SSO and Graceful Restart

1. Router Restarts
2. Adjacency/Peer Relation Fails
3. Peer Removes All Associated Routes from the Routing Table
4. Peer Informs the Neighbors about the Change
5. Restarting Router Re-Establishes Adjacency
6. Peer Adds Associated Routes
7. Peer Informs Neighbors about the Change

4,7 Cause Route Flaps and Network Instability
Traffic Is Interrupted

A) During Failover
1. rtrB
2. rtrC
3. DEL
4. rtrC

B) During Recovery
5. rtrB
6. ADD
7. rtrC
Routing Protocol Operation with NSF/SSO and Graceful Restart

1. Routers Establish Peer Relation and Exchange Graceful Restart Capability
2. Router Restarts
3. Peer Relation Is Lost. Peer DOES NOT Remove Routes from Table
4. Peer DOES NOT Inform Neighbors
5. Restarting Router Re-Establishes Adjacency
6. Peer Updates Restarting Router with It’s Routing Information
7. Restarting Router Sends Routing Updates to the Peer

A) During Failover
- 2. Router Restarts
- 1, 3. Peer Relation Is Lost. Peer DOES NOT Remove Routes from Table
- 4. Peer DOES NOT Inform Neighbors
- 3. Peer Relation Is Lost. Peer DOES NOT Remove Routes from Table
- 5. Restarting Router Re-Establishes Adjacency
- 6. Peer Updates Restarting Router with It’s Routing Information
- 7. Restarting Router Sends Routing Updates to the Peer

No Route Flaps during Recovery
Traffic Flow Is not Interrupted

B) During Recovery
- 2. Router Restarts
- 1, 3. Peer Relation Is Lost. Peer DOES NOT Remove Routes from Table
- 4. Peer DOES NOT Inform Neighbors
- 3. Peer Relation Is Lost. Peer DOES NOT Remove Routes from Table
- 5. Restarting Router Re-Establishes Adjacency
- 6. Peer Updates Restarting Router with It’s Routing Information
- 7. Restarting Router Sends Routing Updates to the Peer

RP Redundancy Configuration

RouterA(config)#redundancy
RouterA(red-config)#mode?
rpr Route Processor Redundancy
rpr-plus Route Processor Redundancy Plus
sso Stateful Switchover (Hot Standby)
RP Redundancy Show Commands

RouterA# sh redundancy state

my state = 13 -ACTIVE
peer state = 8 -STANDBY HOT
Mode = Duplex
Unit ID = 6

Redundancy Mode = SSO
Maintenance Mode = Disabled
Manual Swact = Enabled
Communications = Up
client count = 18
client_notification_TMR = 30000 milliseconds
RF debug mask = 0x0

Agenda

• Introduction
• OSPF Graceful Restart (OSPF NSF)
• ISIS Graceful Restart (ISIS NSF)
• BGP Graceful Restart
• EIGRP NSF
• General Deployment Guidelines
OSPF NSF

- Competing drafts proposed in IETF
- Cisco calls its implementation OSPF NSF, others call their implementation OSPF hitless restart
- Cisco implementation is Cisco IOS:12.0(22)S, 12.2T, 12.2S (release and device dependent)

OSPF NSF Operation Summary

OSPF NSF-Capable Router

- OSPF NSF Capability Exchange
- Router Restarts
- Send Restart Notification
- Request Database Information
- Run SPF Calculation
- Update Forwarding Table
- Send Updates Only

OSPF NSF-Aware Peer

- OSPF NSF Capability Exchange
- OSPF Hello with LR Bit Set
- Hello with Restart (RS) Bit
- Hello w/o Restart (RS) Bit
- Link State Database Request
- Send Link State Database
- Send Updates

CONVERGED!
OSPF NSF Deployment Consideration

- OSPF NSF will terminate on a segment if a non-NSF aware router is detected on the segment; NSF will continue on other segments.
- OSPF NSF can be terminated on an entire router if one Non-NSF router is found on any segment, using an optional command:
  
  ```
  router ospf 100
  nsf enforce global
  ```

OSPF NSF Deployment Consideration

- R2 load balances traffic from subnet A to subnet B.
- R2 undergoes a restart.
- R2 is NOT configured with `nsf enforce global`.
- R2 does not perform OSPF NSF with R3 since R4 is non-NSF.
- R2 retains the forwarding information for routes to subnet B and load balances traffic during restart.

ALL Routers Are NSF Aware
Unless Indicated
OSPF NSF Deployment Consideration (Cont.)

- Traffic from subnet B to subnet A is interrupted during R2 switchover
- R2 does not perform OSPF NSF with R3 as mentioned
- R3 and R4 reset adjacency with R2
- Traffic during this period uses R8
- This will create momentary asymmetric routing

ALL Routers are NSF Aware Unless Indicated

OSPF NSF Configuration

```
Router A (Dual RP)
2.2.2.1            3.3.3.1
192.10.0.x/16     192.10.0.x/16

Router A
interface GigabitEthernet1/0/0
ip address 192.10.0.2 255.255.0.0
router ospf 1
router-id 2.2.2.1
nsf
network 192.10.0.0 0.0.255.255 area 0

Router B
interface GigabitEthernet5/0/0
ip address 192.10.0.3 255.255.0.0
router ospf 1
router-id 3.3.3.1
nsf
network 192.10.0.0 0.0.255.255 area 0
```
OSPF NSF Show Commands

• The following command can be used to verify OSPF NSF capability on a router

RouterA# sh ip ospf
Routing Process "ospf 1" with ID 2.2.2.1 and Domain ID 0.0.0.1
Supports only single TOS(TOS0) routes
Supports opaque LSA
SPF schedule delay 5 secs, Hold time between two SPF's 10 secs
Minimum LSA interval 5 secs. Minimum LSA arrival 1 secs
*
*
Number of areas in this router is 1. 1 normal 0 stub 0 nssa
External flood list length 0
Non-Stop Forwarding enabled, last NSF restart 00:02:51 ago (took 37 secs)
>GLOBAL: ospf_nsf_active: 0x0 router_rp_is_standby 0x0
>GLOBAL: router_in_sso_mode: 0x1

OSPF NSF Show Commands

• Following command can be used to verify if peer is NSF aware

RouterA# sh ip ospf neighbor detail
Neighbor 3.3.3.1, interface address 192.10.0.3
In the area 0 via interface GigabitEthernet1/0/0
Neighbor priority is 1, State is FULL, 7 state changes
DR is 192.10.0.3 BDR is 192.10.0.2
Options is 0x52
LLS Options is 0x1 (LR), last OOB-Resync 00:03:06 ago
Dead timer due in 00:00:37
Neighbor is up for 00:00:32
Index 1/1, retransmission queue length 0, number of retransmission 0
First 0x0(0)/0x0(0) Next 0x0(0)/0x0(0)
Last retransmission scan length is 0, maximum is 0
Last retransmission scan time is 0 msec, maximum is 0 msec
**OSPF NSF Show Commands and Output**

- During database sync. any self originated LSA's are marked stale and kept in database

```
RouterA#show ip ospf database
OSPF Router with ID (2.2.2.1) (Process ID 1)
Router Link States (Area 0)
Link ID ADV Router Age Seq# Checksum Link count
1.1.1.1 1.1.1.1 350 0x80000004 0xCADC 2
2.2.2.1 s2.2.2.1 282 0x80000004 0xED99 1
3.3.3.1 3.3.3.1 280 0x80000004 0x3B4A 2
```

**Net Link States (Area 0)**

*Stale*

- Once database sync. is complete router A builds new self-originated LSA and clears stale flag

```
RouterA#show ip ospf database
OSPF Router with ID (2.2.2.1) (Process ID 1)
Router Link States (Area 0)
Link ID ADV Router Age Seq# Checksum Link count
1.1.1.1 1.1.1.1 353 0x80000004 0xCADC 2
2.2.2.1 2.2.2.1 285 0x80000004 0xED99 1
3.3.3.1 3.3.3.1 283 0x80000004 0x3B4A 2
```

**Net Link States (Area 0)**

- After flush timer expires OSPF flushes all LSAs still present in the database with "stale" flag set; the NSF flush timer has a period of 60 seconds

*Jan 4 00:13:22.643: OSPF: NSF flushing uninstalled self-originated LSAs in area 0*
Agenda

- Introduction
- OSPF Graceful Restart (OSPF NSF)
- ISIS Graceful Restart (ISIS NSF)
- BGP Graceful Restart
- General Deployment Guidelines

ISIS NSF

- Cisco’s ISIS NSF implementation comes in two flavors
  - IETF version: based on IETF draft: draft-ietf-isis-restart-0X
  - Cisco version
- The difference between them
  - IETF version depends on neighbors to rebuild the routing table
  - With Cisco version peer can be non-NSF aware
**ISIS NSF Operation Summary (IETF Version)**

**ISIS NSF-Capable Router**
- ISIS NSF Capability Exchange
- **Router Restarts**
  - Send Restart Notification
  - Run SPF Calculation
  - Update Forwarding Table
  - Send Updates Only

**ISIS NSF-Aware Peer**
- ISIS NSF Capability Exchange
- Hello with Restart Options Fields
- Hello with Restart Request Bit Set
- Hello with Restart Acknowledge Bit Set
- Send CSNP
- Send Link State Packets
- Send Updates

**CONVERGED!**

**ISIS NSF Operation Summary (Cisco Version): Point-to-Point Link**

**ISIS NSF-Capable Router**
- Periodic Hello. No Restart Capability Exchange
- **Router Restarts**
  - After Restart Send a CSNP with a Unique LSP
  - Run SPF Calculation
  - Update Forwarding Table
  - Send Updates Only

**Non NSF-Aware Peer**
- Periodic Hello Exchange
- Request Details of this Unique LSP
- Send LSP for All Data Missing in the CSNP
- Update Forwarding Table

**CONVERGED!**

Copyright © 2003, Cisco Systems, Inc. All rights reserved. Printed in USA.
Presentation_ID.scr
**ISIS NSF: Configuration Command**

```
router(config)# router isis
router(config-router)# nsf [cisco/ietf]
```

- **IETF Draft-Based**
- **Cisco Internal Implementation**

**Show and Debug Commands: IETF Version**

```
show clns neighbor detail
```

```
Router#show clns nei detail
System Id Interface SNPA State Holdtime Type Protocol
esr2 PO1/0/0 *HDLC* Up 24 L2 IS-IS
Area Address(es): 49.0002
IP Address(es): 180.10.10.1*
Uptime: 00:02:27
NSF capable
```
**Show and Debug Commands:**
**IETF Version**

```
show isis NSF

Router#show isis NSF
NSF is ENABLED, mode 'ietf'

NSF pdb state:
NSF L1 active interfaces: 0
NSF L1 active LSPs: 0
8.2.2.1.1.1 NSF interfaces awaiting L1 CSNP: 0
Awaiting L1 LSPs:
NSF L2 active interfaces: 0
NSF L2 active LSPs: 0
NSF interfaces awaiting L2 CSNP: 0
Awaiting L2 LSPs:
```

**Agenda**

- Introduction
- OSPF Graceful Restart (OSPF NSF)
- ISIS Graceful Restart (ISIS NSF)
- BGP Graceful Restart
- EIGRP NSF
- General Deployment Guidelines
BGP Graceful Restart

- IETF draft:draft-ietf-idr-restart-06.txt
- Provides a graceful recovery mechanism for a restarting BGP process
- Cisco implementation is Cisco IOS:12.0(22)S, 12.2T, 12.2S (release and device dependent)
- Requires a graceful restart aware neighbor
- Graceful restart capable routers are 7500, 12000, 10000

Graceful Restart BGP Operation Summary

<table>
<thead>
<tr>
<th>BGP GR-Capable Router</th>
<th>BGP GR-Aware Peer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Router Restarts</td>
<td></td>
</tr>
<tr>
<td>Send Restart Notification</td>
<td></td>
</tr>
<tr>
<td>Session Established</td>
<td></td>
</tr>
<tr>
<td>Performs Best Path Selection when EoR Is Received</td>
<td></td>
</tr>
</tbody>
</table>

| OPEN w/ Graceful Restart Capability 64 | Acknowledge Restart, Mark Routes Stale, Start Restart Timer |
| OPEN w/ Restart Bit Set | Stop Restart Timer, Start Stale-path Timer |
| OPEN w/ Capability 64 | Stop Stale-path Timer, Delete Stale Prefixes and Refresh with New Ones |
| Send BGP Hello | Send Initial Updates, End of RIB (EoR) |
| Send Updates+ EoR | CONVERGED! |
BGP Graceful Restart Timers

- Restart timers are used by peers to set the amount of time it waits for a restarting router to establish a BGP session after it has indicated a restart
- Stalepath timers are used by peers to set the amount of time it waits to receive an End of RIB marker (end of RIB indicates the neighbor has converged) from the restarting router

BGP Graceful Restart Timers

- Important to keep restart timer below hold time
- Default values
  - BGP hold time 180 seconds (3 x 60 sec keepalive)
  - Restart timer default 120 seconds
  - Stale path timer default 360 seconds
- Restart timer is advertised to the peer
- Stale path timer is used internally by the router
BGP GR: Deployment Consideration—1

- Consider routes between R1 and R2 when R1 undergoes graceful restart
  - R1 preserves all routes to AS200 and continues forwarding traffic
  - R2 reaches AS100 via AS300
  - All traffic from R2 to R1 goes via R3
  - All traffic from R1 to R2 goes directly
  - This can lead to temporary asymmetric routing
  - No packet loss will be experienced from R1 to R2
  - Some packet loss from R2 to R1 during the re-convergence

BGP GR: Deployment Consideration—2

- Consider routes between R1 and R4 when R1 undergoes graceful restart
  - R1 preserves all routes to AS400 and continues forwarding traffic
  - R4 removes all routes to A
  - R1 continues to forward traffic to R4
  - R4 does not forward traffic to R1 till R1 re-converges
BGP GR: Deployment Consideration—3

- Consider routes between R1 and R3 when R1 undergoes graceful restart
  - R1 preserves all routes to AS300 and continues forwarding traffic
  - R3 preserves all routes to AS100 and continues forwarding traffic

BGP GR: Deployment Consideration—4

- Consider the RR being NSF/SSO capable when the RR reloads
  - RR preserves its forwarding table
  - R1, RC1 and RC2 preserve their forwarding table
  - R1, RC1 and RC2 continue forwarding traffic

- Other considerations for RR being an NSF/SSO router discussed next
• If the RR reflects large number of prefixes it may take longer to converge

Remember

(1) Peer waits to receive the initial routing table with EoR marker

(2) Peer uses the stalepath timer to decide how long to wait

All peers to the RR need to adjust their stalepath timer (default is 360 seconds)

• In an iBGP network the IGP used within the AS should also be NSF capable and NSF aware as applicable

The IGP protocol determines the route to the BGP next hop address. Hence non stop forwarding of BGP traffic requires the non-stop forwarding capability for IGP
BGP Graceful Restart Commands

R18C12KRP(config)#router bgp 100
R18C12KRP(config-router)# bgp graceful-restart
R18C12KRP(config-router)# bgp graceful-restart restart-time 120
R18C12KRP(config-router)# bgp graceful-restart stalepath-time 360

R18C12KRP#sh ip bgp nei
BGP neighbor is 10.10.104.1, remote AS 100, internal link
BGP version 4, remote router ID 10.10.104.1
BGP state = Established, up for 00:00:10
Last read 00:00:09, hold time is 180, keepalive interval is 60 seconds
Neighbor capabilities:
Route refresh: advertised and received(new)
Address family IPv4 Unicast: advertised and received
Graceful Restart Capability: advertised and received
Remote Restart timer is 140 seconds
Address families preserved by peer:
IPv4 Unicast
Agenda

• Introduction
• OSPF Graceful Restart (OSPF NSF)
• ISIS Graceful Restart (ISIS NSF)
• BGP Graceful Restart
• EIGRP NSF
• General Deployment Guidelines

EIGRP NSF

• Will be supported soon in Cisco IOS®
EIGRP NSF Operation Summary

- **EIGRP NSF-Capable Router**
  - EIGRP NSF Capability Exchange
  - Set Signal Timer; Send Restart Notification
  - Stop Signal Timer; Start Convergence Timer
  - Stop Convergence Timer
  - Calculate Best Path
  - Send Startup Updates + EOT

- **EIGRP NSF-Aware Peer**
  - EIGRP NSF Capability Exchange
  - Hello with Restart Bit Set
  - Hello Response w/o Restart Bit Set
  - Null Restart + INIT
  - Restart + Update Packets
  - Startup updates + End of Table (EOT) Packet
  - Startup Update + EOT

**CONVERGED!**

EIGRP NSF Timers

- **On restart router**
  - Signal timer: Used to send Hello with Restart bit set; when this timer expires Hellos’ are sent without Restart bit set
  - Convergence timer: Used to set the amount of time the restarting router waits to receive EOT marker from peers

- **On the peer**
  - Route hold timer: Used by peer to indicate the amount of time the peer waits to receive routing updates and EOT marker from restarting router
EIGRP NSF: Configuration Commands

- On restarting router
  - `router eigrp 100`
  - `nsf`
  - `timers nsf signal`
  - `timers nsf converge`
- On peer
  - `router eigrp 100`
  - `nsf`
  - `timers nsf route-hold`

Agenda

- Introduction
- OSPF Graceful Restart (OSPF NSF)
- ISIS Graceful Restart (ISIS NSF)
- BGP Graceful Restart
- EIGRP NSF
- General Deployment Guidelines
General Deployment Guidelines—1
Route Processor Combination

• When mixing RP types:
  
  Cisco 12000 Series Internet Router: GRP and GRP-B RPs can be used together; PRP can be used only with another PRP

  Cisco 10000 Series Internet Router: two PRE-1s must be used; original PRE does not support NSF/SSO

  Cisco 7500: RSP-2 and RSP-4 can be used in combination; RSP-8 and RSP-16 can also be used in combination; RSP-8 or RSP-16 cannot be mixed with an RSP-2 or an RSP-4

General Deployment Guidelines—2
Fast Convergence and Graceful Restart

• Fast convergence is used to reduce the time taken to detect link failure and provide faster convergence on alternate routes

  Configured in IGP or BGP protocol

  Implemented by reducing hello and hold time interval

  When the hold time expires peer is declared as unreachable
• For graceful restart to succeed the restarting router has to re-establish relation with neighbor before hold time expires

• NSF/SSO deployed on routers with fast convergence enabled will benefit from both
  Faster convergence during link failures
  NSF/SSO and graceful convergence during RP failover

• Factors to consider for fast convergence + NSF/SSO
  On point to point links L2 failure detections are faster and fast convergence may not be needed
  In some L2 protocols like POS, the underlying mechanisms provide failure detection in milliseconds, hence fast convergence is not needed
General Deployment Guidelines—2
Fast Convergence and Graceful Restart

• Factors to consider for fast convergence
+ NSF/SSO (Cont.)

In OSPF NSF implementation

Deploy with 12.0(25)S or 12.2(15)T

On 7500 routers, RSP 8 and lower, dead interval should be no less than 8 seconds for NSF behavior

On other platforms, dead interval can be lower

In ISIS NSF implementation

On 7500 routers, RSP 8 and lower, dead interval should be no less than 9 seconds for NSF behavior

On other platforms, dead interval can be lower
Summary

• Graceful restart provides a mechanism for a restarting RP to continue forwarding packets while undergoing re-convergence

• Graceful restart will help reduce packet loss and eliminate routing instability under certain failure conditions

• Coming soon...Graceful restart mechanisms in LDP and BGP MPLS

Appendix-A

• APS—Automatic Protection Switching
• CLI—Cisco Command Line Interface
• DBD—database Descriptor packet
• GRP—Gigabit Route Processor (Cisco 12000)
• GR—Graceful Restart
• CSNP—Complete Sequence Number Packets (used in ISIS)
• LFI—Link Fragmentation and Interleaving
• LLS—Link local signal in Hello and DBD data blocks in OSPF
• MLPPP—Multi-link PPP
• NSF—(Cisco’s) Non Stop Forwarding
• OOB— resync—Out of Band Resynchronization in OSPF NSF
• PPP—Point to Point Protocol
• PRP—Performance Route Processor (Cisco 12000)
• PRE—Performance Routing Engine (Cisco 10000)
• PSNP—Partial Sequence Number Packets (in ISIS)
• RIB—Routing Information Base (Routing Table)
• RP—Route Processor
• RPR—Route Processor Redundancy
• RSP—Route Switch Processor (Cisco 7500)
• SSO—(Cisco’s) Stateful Switchover
• SNP—Sequence Number Packets (SNP)
• SPF—Shortest Path First algorithm used by OSPF and ISIS
• TE—Traffic Engineering
• T1 IMA—T1 Inverse Multiplexing over ATM
• TLV—Type Length Value used in protocol headers
Recommended Sessions

• RST 2002: Deployment and Analysis of Link State Protocols
• RST 2003: Deployment and Analysis of BGP
• RST 2111/2112: Router platform Architecture and Analysis
• RST 3001: Deployment and Analysis of EIGRP Network

Reference Material

• Non-Stop Forwarding:  
  http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120limit/120s/120s22/nsf120s.pdf

• OSPF NSF:  

• BGP NSF:  

• Globally Resilient IP (GRIP) overview:  

• EIGRP NSF:  

• IETF: http://www.ietf.org
Please Complete Your Evaluation Form

Session RST-2004
Back-Up Slides

OSPF Restart Timers: An Illustration

Router B

• OSPF Hello from A
• A ReStarts
• A Starts NSF Wait Timer (= 20 Seconds)
• A Sends Fast Hello

Router B Time Line

• B receives Hello with w/ Restart Bit Set
• B Starts OOB Resync. Timer
• B Responds to Fast Hellos
• B Maintains FSM as FULL

Dead Interval 40 Sec.

OOB Resync. Timer 40 Sec.

Starting Router A

• B Receives DBD w/ Resync.
• B Resync. Bit Set.
• B Sends LSDB Resync. w/ A
• B Stops the OOB Resync. Timer

• DR/ BDR Election
• A sends DBD Packets with resync. Bit Set
• A Stops NSF Wait Timer
OSPF NSF Debug Commands

- Following debug commands can be used to track the events on a restarting router
  - debug ip ospf nsf detail
  - debug ip ospf adj
  - debug ip ospf hello

OSPF NSF Debug Output

- Router A restarts, sets NSF timer and sends fast hello’s (hello’s every 2 seconds)

```
*Jan 4 00:11:55.131: Routing Process "ospf 1", RIBID 0 mask 0x4
*Jan 4 00:11:55.131: NSF timer started for ospf process 1,
*Jan 4 00:11:57.131: OSPF: NSF interface count+: 1 (area 0),
  GigabitEthernet1/0/0
*Jan 4 00:11:57.131: OSPF: Send hello to 224.0.0.5 area 0 on
gigabitethernet1/0/0 from 192.10.0.2

....
*Jan 4 00:11:59.131: OSPF: Send hello to 224.0.0.5 area 0 on
gigabitethernet1/0/0 from 192.10.0
```
OSPF NSF Debug Output

• The neighbor replies back to “Fast Hellos” to establish neighbor relation

*Jan 4 00:11:57.919: OSPF: Rcv hello from 3.3.3.1 area 0 from GigabitEthernet1/0/0 192.10.0.3
*Jan 4 00:11:57.919: OSPF: 2 Way Communication to 3.3.3.1 on GigabitEthernet1/0/0, state 2WAY
*Jan 4 00:11:57.919: OSPF: NSF 2 Way Communication to 3.3.3.1 on GigabitEthernet1/0/0, state FULL

*Jan 4 00:11:59.143: OSPF: Rcv hello from 3.3.3.1 area 0 from GigabitEthernet1/0/0 192.10.0.3

OSPF NSF Debug Output

• NSF Timer stops after 20 seconds; OOB-Resync. Commences; DR/BDR lookup is done after NSF Timer expires

*Jan 4 00:12:15.131: OSPF process 1: NSF RESTART TIMER period 1 expired
*Jan 4 00:12:17.131: OSPF: end of Wait on interface GigabitEthernet1/0/0
*Jan 4 00:12:17.131: OSPF: DR/BDR election on GigabitEthernet1/0/0
*Jan 4 00:12:17.131: OSPF: Elect BDR 2.2.2.1
*Jan 4 00:12:17.131: OSPF: Elect DR 3.3.3.1
*Jan 4 00:12:17.131: OSPF: Elect BDR 2.2.2.1
*Jan 4 00:12:17.131: OSPF: Elect DR 3.3.3.1
*Jan 4 00:12:17.131: DR: 3.3.3.1 (Id) BDR: 2.2.2.1 (Id)
*Jan 4 00:12:17.131: OSPF: OOB Resync scheduled for 3.3.3.1 on GigabitEthernet1/0/0
**OSPF NSF Debug Output**

- OOB-LSDB is progress; the "flag 0xF" in the output below indicates that the R-bit, along with the I-bit, M-bit, and MS-bit are set in DBD packet

*Jan 4 00:12:17.631: OSPF: Starting OOB-Resync with 3.3.3.1 address 192.10.0.3 on GigabitEthernet1/0/0 (requester)*
*Jan 4 00:12:22.631: OSPF: Send DBD (oob-resync) to 3.3.3.1 on GigabitEthernet1/0/0 seq 0x132F opt 0x52 flag 0xF len 32*
*Jan 4 00:12:22.635: OSPF: Rcv DBD (oob-resync) from 3.3.3.1 on GigabitEthernet1/0/0 seq 0x196F opt 0x52 flag 0xF len 32 mtu 1500 state EXSTART*
*Jan 4 00:12:22.643: OSPF: OOB-Resync completed with 3.3.3.1 address 192.10.0.3 on GigabitEthernet1/0/0*
*Jan 4 00:12:22.643: OSPF process 1: oob-resync completed for all neighbors*

**ISIS Debug Commands: IETF Version**

default isis nsf

The following debug messages report the steps of the IETF NSF IS-IS mechanism:

00:48:59: ISIS-NSF: Begin NSF proto processing
00:48:59: ISIS-NSF: Begin main RTR loop ...
00:48:59: ISIS-NSF: pdb state progression: state=Inactive/event=Restart/newstate=Init
00:48:59: ISIS-NSF: IETF NSF [UPD Init] - clock 0
00:49:00: ISIS-NSF: IETF NSF [UPD Init] - clock 1
00:49:01: ISIS-NSF: POS1/1/0 level-2 state progression: state=Running/event=Restart/
\[newstate=Restarting\]
00:49:01: ISIS-NSF: Inserting p2p NSF REQ on POS1/1/0 IIH
00:49:01: ISIS-NSF: IETF NSF [UPD Init] - clock 2
00:49:01: ISIS-NSF: POS1/1/0 level-2 state progression: state=Restarting/event=R A Rcvd/
\[newstate=RA Seen\]
00:49:01: ISIS-NSF: NSF restart ACK option received on p2p if from 0000.0000.0008
(POS1/1/0)
00:49:01: ISIS-NSF: pdb state progression: state=Init/event=LSP Complete/newstate=LSP
\[Completed\]
**ISIS Debug Commands: IETF Version**

**debug isis nsf (Cont.)**

[ snippets ]

00:49:01: ISIS-NSF: pdb state progression: state=LSP Completed/event=IF Complete/

newstate=SPF1

00:49:01: ISIS-NSF: POS1/1/0 level2 state progression: state=RA Seen/event=CSNP Rcvd/

newstate=Running

00:49:02: ISIS-NSF: IETF NSF [UPD SPF1] - clock 3

00:49:02: ISIS-NSF: pdb state progression: state=SPF1/event=Increment/newstate=SPF2 progression/state=SPF2/event=Increment/newstate=Complete

00:49:02: %CLNS-5-NSF_RESTART: ISIS NSF completion

00:49:02: ISIS-NSF: Begin main UPD loop ...

---

**BGP Graceful Restart Timers: An Illustration**

- **BGP Hold Time**: 180 Sec
- **Restart Timer**: 120 Sec
- **Stale Path Timer**: 180 Sec

**Router A**
- Receives OPEN from B w/o Restart Bit Set
- Sends Initial Routing Update to B with an EOR Marker
- Sends EOR Marker

**Router B**
- Gets Keepalive from A
- Gets OPEN from A w/ Restart Bit Set
- Marks Routes from A as Stale
- Starts the Restart Timer
- Sends an OPEN to A
- Receives OPEN from B w/o Restart Bit Set
- Sends Keepalive
- BGP Session Established

---

Copyright © 2003, Cisco Systems, Inc. All rights reserved. Printed in USA.

Presentation_ID.scr