                          Restart Signaling for
            Intermediate System to Intermediate System (IS-IS)

Status of this Memo

Copyright Notice

Abstract

   This document describes a mechanism for a restarting router to signal
   to its neighbors that it is restarting, allowing them to reestablish
   their adjacencies without cycling through the down state, while still
   correctly initiating database synchronization.

   This document additionally describes a mechanism for a restarting
   router to determine when it has achieved LSP database synchronization
   with its neighbors and a mechanism to optimize LSP database
   synchronization, while minimizing transient routing disruption when a
   router starts.

Table of Contents

1. Conventions used in this Document

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
   document are to be interpreted as described in BCP 14, RFC-2119 [3].

   If the control and forwarding functions in a router can be maintained
   independently, it is possible for the forwarding function state to be
   maintained across a resumption of control function operations. This
   functionality is assumed when the terms "restart/restarting" are used
   in this document.

   The terms "start/starting" are used to refer to a router in which the
   control function has either commenced operations for the first time
   or has resumed operations but the forwarding functions have not been
   maintained in a prior state.

   The terms "(re)start/(re)starting" are used when the text is
   applicable to both a "starting" and a "restarting" router.

2. Overview

   The Intermediate System to Intermediate System (IS-IS) routing
   protocol [RFC 1195, ISO/IEC 10589] is a link state intra-domain
   routing protocol. Normally, when an IS-IS router is restarted,
   temporary disruption of routing occurs due to events in both the
   restarting router and the neighbors of the restarting router.

   The router which has been restarted computes its own routes before
   achieving database synchronization with its neighbors. The results
   of this computation are likely to be non-convergent with the routes
   computed by other routers in the area/domain.

   Neighbors of the restarting router detect the restart event and cycle
   their adjacencies with the restarting router through the down state.
   The cycling of the adjacency state causes the neighbors to regenerate
   their LSPs describing the adjacency concerned. This in turn causes a
   temporary disruption of routes passing through the restarting router.

   In certain scenarios, the temporary disruption of the routes is
   highly undesirable. This document describes mechanisms to avoid or
   minimize the disruption due to both of these causes.

   When an adjacency is reinitialized as a result of a neighbor
   restarting, a router does three things:

   1. It causes its own LSP(s) to be regenerated, thus triggering SPF
      runs throughout the area (or in the case of Level 2, throughout
      the domain).

   2. It sets SRMflags on its own LSP database on the adjacency
      concerned.

   3. In the case of a Point-to-Point link, it transmits a (set of)
      CSNP(s) over the adjacency.

   In the case of a restarting router process, the first of these is
   highly undesirable, but the second is essential in order to ensure
   synchronization of the LSP database.

   The third action above minimizes the number of LSPs which must be
   exchanged and, if made reliable, provides a means of determining when
   the LSP databases of the neighboring routers have been synchronized.
   This is desirable whether the router is being restarted or not (so
   that the overload bit can be cleared in the router's own LSP, for
   example).

   This document describes a mechanism for a restarting router to signal
   that it is restarting to its neighbors, and allow them to reestablish
   their adjacencies without cycling through the down state, while still
   correctly initiating database synchronization.

   This document additionally describes a mechanism for a restarting
   router to determine when it has achieved LSP database synchronization
   with its neighbors and a mechanism to optimize LSP database
   synchronization and minimize transient routing disruption when a
   router starts.

   It is assumed that the three-way handshake [4] is being used on
   Point-to-Point circuits.

3. Approach

3.1. Timers

   Three additional timers, T1, T2, and T3 are required to support the
   functionality defined in this document.

   An instance of the timer T1 is maintained per interface, and
   indicates the time after which an unacknowledged (re)start attempt
   will be repeated. A typical value might be 3 seconds.

   An instance of the timer T2 is maintained for each LSP database
   present in the system, i.e., for a Level1/2 system, there will be an
   instance of the timer T2 for Level 1 and an instance for Level 2.
   This is the maximum time that the system will wait for LSPDB
   synchronization. A typical value might be 60 seconds.

   A single instance of the timer T3 is maintained for the entire
   system. It indicates the time after which the router will declare
   that it has failed to achieve database synchronization (by setting
   the overload bit in its own LSP). This is initialized to 65535
   seconds, but is set to the minimum of the remaining times of received
   IIHs containing a restart TLV with the RA set and an indication that
   the neighbor has an adjacency in the "UP" state to the restarting
   router.

   NOTE: The timer T3 is only used by a restarting router.

3.2. Restart TLV

   A new TLV is defined to be included in IIH PDUs. The presence of
   this TLV indicates that the sender supports the functionality defined
   in this document and it carries flags that are used to convey
   information during a (re)start. All IIHs transmitted by a router
   that supports this capability MUST include this TLV.

   Type   211
   Length # of octets in the value field (1 to (3 + ID Length))
   Value

                                       No. of octets
      +-----------------------+
      | Flags | 1
      +-----------------------+
      | Remaining Time | 2
      +-----------------------+
      | Restarting Neighbor ID| ID Length
      +-----------------------+

   Flags (1 octet)

       0 1 2 3 4 5 6 7
      +--+--+--+--+--+--+--+--+
      | Reserved |SA|RA|RR|
      +--+--+--+--+--+--+--+--+

   RR - Restart Request
   RA - Restart Acknowledgement
   SA - Suppress adjacency advertisement

   (Note: Remaining fields are required when the RA bit is set)

   Remaining Time (2 octets)

      Remaining holding time (in seconds)

   Restarting Neighbor System ID (ID Length octets)

   The system ID of the neighbor to which an RA refers. Note:
   Implementations based on earlier versions of this document may not
   include this field in the TLV when the RA is set. In this case, a
   router which is expecting an RA on a LAN circuit SHOULD assume that
   the acknowledgement is directed at the local system.

3.2.1. Use of RR and RA Bits

   The RR bit is used by a (re)starting router to signal to its
   neighbors that a (re)start is in progress, that an existing adjacency
   SHOULD be maintained even under circumstances when the normal
   operation of the adjacency state machine would require the adjacency
   to be reinitialized, to request a set of CSNPs, and to request
   setting of the SRMflags.

   The RA bit is sent by the neighbor of a (re)starting router to
   acknowledge the receipt of a restart TLV with the RR bit set.

   When the neighbor of a (re)starting router receives an IIH with the
   restart TLV having the RR bit set, if there exists on this interface
   an adjacency in state "UP" with the same System ID, and in the case
   of a LAN circuit, with the same source LAN address, then,
   irrespective of the other contents of the "Intermediate System
   Neighbors" option (LAN circuits) or the "Point-to-Point Three-Way
   Adjacency" option (Point-to-Point circuits):

   a) the state of the adjacency is not changed. If this is the first
      IIH with the RR bit set that this system has received associated
      with this adjacency, then the adjacency is marked as being in
      "Restart mode" and the adjacency holding time is refreshed -
      otherwise the holding time is not refreshed. The "remaining time"
      transmitted according to (b) below MUST reflect the actual time
      after which the adjacency will now expire. Receipt of a normal
      IIH with the RR bit reset will clear the "Restart mode" state.
      This procedure allows the restarting router to cause the neighbor
      to maintain the adjacency long enough for restart to successfully
      complete while also preventing repetitive restarts from
      maintaining an adjacency indefinitely. Whether an adjacency is
      marked as being in "Restart mode" or not has no effect on
      adjacency state transitions.

   b) immediately (i.e., without waiting for any currently running timer
      interval to expire, but with a small random delay of a few 10s of
      milliseconds on LANs to avoid "storms") transmit over the
      corresponding interface an IIH including the restart TLV with the
      RR bit clear and the RA bit set, in the case of Point-to-Point
      adjacencies having updated the "Point-to-Point Three-Way
      Adjacency" option to reflect any new values received from the
      (re)starting router. (This allows a restarting router to quickly
      acquire the correct information to place in its hellos.) The
      "Remaining Time" MUST be set to the current time (in seconds)
      before the holding timer on this adjacency is due to expire. If
      the corresponding interface is a LAN interface, then the
      Restarting Neighbor System ID SHOULD be set to the System ID of

the router from whom the IIH with the RR bit set was received.
This is required to correctly associate the acknowledgement and
holding time in the case where multiple systems on a LAN restart
at approximately the same time. This IIH SHOULD be transmitted
before any LSPs or SNPs are transmitted as a result of the receipt
of the original IIH.

c) if the corresponding interface is a Point-to-Point interface, or
if the receiving router has the highest LnRouterPriority (with
highest source MAC address breaking ties) among those routers to
which the receiving router has an adjacency in state "UP" on this
interface whose IIHs contain the restart TLV, excluding
adjacencies to all routers which are considered in "Restart mode"
(note the actual DIS is NOT changed by this process), initiate the
transmission over the corresponding interface of a complete set of
CSNPs, and set SRMflags on the corresponding interface for all
LSPs in the local LSP database.

Otherwise (i.e., if there was no adjacency in the "UP" state to the
system ID in question), process the IIH as normal by reinitializing
the adjacency and setting the RA bit in the returned IIH.

3.2.2. Use of the SA Bit

The SA bit is used by a starting router to request that its neighbor
suppress advertisement of the adjacency to the starting router in the
neighbor's LSPs.

A router which is starting has no maintained forwarding function
state. This may or may not be the first time the router has started.
If this is not the first time the router has started, copies of LSPs
generated by this router in its previous incarnation may exist in the
LSP databases of other routers in the network. These copies are
likely to appear "newer" than LSPs initially generated by the
starting router due to the reinitialization of LSP fragment sequence
numbers by the starting router. This may cause temporary blackholes
to occur until the normal operation of the update process causes the
starting router to regenerate and flood copies of its own LSPs with
higher sequence numbers. The temporary blackholes can be avoided if
the starting router's neighbors suppress advertising an adjacency to
the starting router until the starting router has been able to
propagate newer versions of LSPs generated by previous incarnations.

When a router receives an IIH with the restart TLV having the SA bit
set, if there exists on this interface an adjacency in state "UP"
with the same System ID, and in the case of a LAN circuit, with the
same source LAN address, then the router MUST suppress advertisement
of the adjacency to the neighbor in its own LSPs. Until an IIH with

the SA bit clear has been received, the neighbor advertisement MUST
continue to be suppressed. If the adjacency transitions to the "UP"
state, the new adjacency MUST NOT be advertised until an IIH with the
SA bit clear has been received.

Note that a router which suppresses advertisement of an adjacency
MUST NOT use this adjacency when performing its SPF calculation. In
particular, if an implementation follows the example guidelines
presented in [2] Annex C.2.5 Step 0:b) "pre-load TENT with the local
adjacency database", the suppressed adjacency MUST NOT be loaded into
TENT.

3.3. Adjacency (Re)Acquisition

   Adjacency (re)acquisition is the first step in (re)initialization.
   Restarting and starting routers will make use of the RR bit in the
   restart TLV, though each will use it at different stages of the
   (re)start procedure.

3.3.1. Adjacency Reacquisition During Restart

   The restarting router explicitly notifies its neighbor that the
   adjacency is being reacquired, and hence that it SHOULD NOT
   reinitialize the adjacency. This is achieved by setting the RR bit
   in the restart TLV. When the neighbor of a restarting router
   receives an IIH with the restart TLV having the RR bit set, if there
   exists on this interface an adjacency in state "UP" with the same
   System ID, and in the case of a LAN circuit, with the same source LAN
   address, then the procedures described in 3.2.1 are followed.

   A router that does not support the restart capability will ignore the
   restart TLV and reinitialize the adjacency as normal, returning an
   IIH without the restart TLV.

   On restarting, a router initializes the timer T3, starts the timer T2
   for each LSPDB, and for each interface (and in the case of a LAN
   circuit, for each level) starts the timer T1 and transmits an IIH
   containing the restart TLV with the RR bit set.

   On a Point-to-Point circuit the restarting router SHOULD set the
   "Adjacency Three-Way State" to "Init", because the receipt of the
   acknowledging IIH (with RA set) MUST cause the adjacency to enter the
   "UP" state immediately.

   On a LAN circuit the LAN-ID assigned to the circuit SHOULD be the
   same as that used prior to the restart. In particular, for any
   circuits for which the restarting router was previously DIS, the use
   of a different LAN-ID would necessitate the generation of a new set

of pseudonode LSPs, and corresponding changes in all the LSPs
referencing them from other routers on the LAN. By preserving the
LAN-ID across the restart, this churn can be prevented. To enable a
restarting router to learn the LAN-ID used prior to restart, the
LAN-ID specified in an IIH with RR set MUST be ignored.

Transmission of "normal" IIHs is inhibited until the conditions
described below are met (in order to avoid causing an unnecessary
adjacency initialization). Upon expiry of the timer T1, it is
restarted and the IIH is retransmitted as above.

When a restarting router receives an IIH a local adjacency is
established as usual, and if the IIH contains a restart TLV with the
RA bit set (and on LAN circuits with a Restart Neighbor System ID
which matches that of the local system), the receipt of the
acknowledgement over that interface is noted. When the RA bit is set
and the state of the remote adjacency is "UP", then the timer T3 is
set to the minimum of its current value and the value of the
"Remaining Time" field in the received IIH.

On a Point-to-Point link, receipt of an IIH not containing the
restart TLV is also treated as an acknowledgement, since it indicates
that the neighbor is not restart capable. However, since no CSNP is
guaranteed to be received over this interface, the timer T1 is
cancelled immediately without waiting for a complete set of CSNP(s).
Synchronization may therefore be deemed complete even though there
are some LSPs which are held (only) by this neighbor (see section
3.4). In this case we also want to be certain that the neighbor will
reinitialize the adjacency in order to guarantee that the SRMflags
have been set on its database, thus ensuring eventual LSPDB
synchronization. This is guaranteed to happen except in the case
where the Adjacency Three-Way State in the received IIH is "UP" and
the Neighbor Extended Local Circuit ID matches the extended local
circuit ID assigned by the restarting router. In this case the
restarting router MUST force the adjacency to reinitialize by setting
the local Adjacency Three-Way State to "DOWN" and sending a normal
IIH.

In the case of a LAN interface, receipt of an IIH not containing the
restart TLV is unremarkable since synchronization can still occur so
long as at least one of the non-restarting neighboring routers on the
LAN supports restart. Therefore T1 continues to run in this case.
If none of the neighbors on the LAN are restart capable, T1 will
eventually expire after the locally defined number of retries.

In the case of a Point-to-Point circuit, the "LocalCircuitID" and
"Extended Local Circuit ID" information contained in the IIH can be
used immediately to generate an IIH containing the correct 3-way

handshake information. The presence of "Neighbor Extended Local
Circuit ID" information which does not match the value currently in
use by the local system is ignored (since the IIH may have been
transmitted before the neighbor had received the new value from the
restarting router), but the adjacency remains in the initializing
state until the correct information is received.

In the case of a LAN circuit, the source neighbor information (e.g.,
SNPAAddress) is recorded and used for adjacency establishment and
maintenance as normal.

When BOTH a complete set of CSNP(s) (for each active level, in the
case of a point-to-point circuit) and an acknowledgement have been
received over the interface, the timer T1 is cancelled.

Once the timer T1 has been cancelled, subsequent IIHs are transmitted
according to the normal algorithms, but including the restart TLV
with both RR and RA clear.

If a LAN contains a mixture of systems, only some of which support
the new algorithm, database synchronization is still guaranteed, but
the "old" systems will have reinitialized their adjacencies.

If an interface is active, but does not have any neighboring router
reachable over that interface, the timer T1 would never be cancelled,
and according to clause 3.4.1.1, the SPF would never be run.
Therefore timer T1 is cancelled after some pre-determined number of
expirations (which MAY be 1).

3.3.2. Adjacency Acquisition During Start

The starting router wants to ensure that in the event that a
neighboring router has an adjacency to the starting router in the
"UP" state (from a previous incarnation of the starting router), this
adjacency is reinitialized. The starting router also wants
neighboring routers to suppress advertisement of an adjacency to the
starting router until LSP database synchronization is achieved. This
is achieved by sending IIHs with the RR bit clear and the SA bit set
in the restart TLV. The RR bit remains clear and the SA bit remains
set in subsequent transmissions of IIHs until the adjacency has
reached the "UP" state and the initial T1 timer interval (see below)
has expired.

Receipt of an IIH with the RR bit clear will result in the
neighboring router utilizing normal operation of the adjacency state
machine. This will ensure that any old adjacency on the neighboring
router will be reinitialized.

Upon receipt of an IIH with the SA bit set, the behavior described in
3.2.2 is followed.

Upon starting, a router starts timer T2 for each LSPDB.

For each interface (and in the case of a LAN circuit, for each
level), when an adjacency reaches the "UP" state, the starting router
starts a timer T1 and transmits an IIH containing the restart TLV
with the RR bit clear and SA bit set. Upon expiry of the timer T1,
it is restarted and the IIH is retransmitted with both RR and SA bits
set (only the RR bit has changed state from earlier IIHs).

Upon receipt of an IIH with the RR bit set (regardless of whether the
SA is set or not), the behavior described in 3.2.1 is followed.

When an IIH is received by the starting router and the IIH contains a
restart TLV with the RA bit set (and on LAN circuits with a Restart
Neighbor System ID which matches that of the local system), the
receipt of the acknowledgement over that interface is noted.

On a Point-to-Point link, receipt of an IIH not containing the
restart TLV is also treated as an acknowledgement, since it indicates
that the neighbor is not restart capable. Since the neighbor will
have reinitialized the adjacency, this guarantees that SRMflags have
been set on its database, thus ensuring eventual LSPDB
synchronization. However, since no CSNP is guaranteed to be received
over this interface, the timer T1 is cancelled immediately without
waiting for a complete set of CSNP(s). Synchronization may therefore
be deemed complete even though there are some LSPs which are held
(only) by this neighbor (see section 3.4).

In the case of a LAN interface, receipt of an IIH not containing the
restart TLV is unremarkable since synchronization can still occur so
long as at least one of the non-restarting neighboring routers on the
LAN supports restart. Therefore T1 continues to run in this case.
If none of the neighbors on the LAN are restart capable, T1 will
eventually expire after the locally defined number of retries. The
usual operation of the update process will ensure that
synchronization is eventually achieved.

When BOTH a complete set of CSNP(s) (for each active level, in the
case of a point-to-point circuit) and an acknowledgement have been
received over the interface, the timer T1 is cancelled. Subsequent
IIHs sent by the starting router have the RR and RA bits clear and
the SA bit set in the restart TLV.

Timer T1 is cancelled after some pre-determined number of expirations
(which MAY be 1).

When the T2 timer(s) are cancelled or expire, transmission of
"normal" IIHs (with RR, RA, and SA bits clear) will begin.

3.3.3. Multiple Levels

A router which is operating as both a Level 1 and a Level 2 router on
a particular interface MUST perform the above operations for each
level.

On a LAN interface, it MUST send and receive both Level 1 and Level 2
IIHs and perform the CSNP synchronizations independently for each
level.

On a point-to-point interface, only a single IIH (indicating support
for both levels) is required, but it MUST perform the CSNP
synchronizations independently for each level.

3.4. Database Synchronization

When a router is started or restarted it can expect to receive a (set
of) CSNP(s) over each interface. The arrival of the CSNP(s) is now
guaranteed, since an IIH with the RR bit set will be retransmitted
until the CSNP(s) are correctly received.

The CSNPs describe the set of LSPs that are currently held by each
neighbor. Synchronization will be complete when all these LSPs have
been received.

When (re)starting, a router starts an instance of timer T2 for each
LSPDB as described in 3.3.1 or 3.3.2. In addition to normal
processing of the CSNPs, the set of LSPIDs contained in the first
complete set of CSNP(s) received over each interface is recorded,
together with their remaining lifetime. In the case of a LAN
interface, a complete set of CSNPs MUST consist of CSNPs received
from neighbor(s) which are not restarting. If there are multiple
interfaces on the (re)starting router, the recorded set of LSPIDs is
the union of those received over each interface. LSPs with a
remaining lifetime of zero are NOT so recorded.

As LSPs are received (by the normal operation of the update process)
over any interface, the corresponding LSPID entry is removed (it is
also removed if an LSP arrives before the CSNP containing the
reference). When an LSPID has been held in the list for its
indicated remaining lifetime, it is removed from the list. When the
list of LSPIDs is empty and the timer T1 has been cancelled for all
the interfaces that have an adjacency at this level, the timer T2 is
cancelled.

At this point, the local database is guaranteed to contain all the
LSP(s) (either the same sequence number, or a more recent sequence
number) that were present in the neighbors' databases at the time of
(re)starting. LSPs that arrived in a neighbor's database after the
time of (re)starting may or may not be present, but the normal
operation of the update process will guarantee that they will
eventually be received. At this point, the local database is deemed
to be "synchronized".

Since LSPs mentioned in the CSNP(s) with a zero remaining lifetime
are not recorded, and those with a short remaining lifetime are
deleted from the list when the lifetime expires, cancellation of the
timer T2 will not be prevented by waiting for an LSP that will never
arrive.

3.4.1. LSP Generation and Flooding and SPF Computation

The operation of a router starting, as opposed to restarting, is
somewhat different. These two cases are dealt with separately below.

3.4.1.1. Restarting

In order to avoid causing unnecessary routing churn in other routers,
it is highly desirable that the router's own LSPs generated by the
restarting system are the same as those previously present in the
network (assuming no other changes have taken place). It is
important therefore not to regenerate and flood the LSPs until all
the adjacencies have been re-established and any information required
for propagation into the local LSPs is fully available. Ideally, the
information is loaded into the LSPs in a deterministic way, such that
the same information occurs in the same place in the same LSP (and
hence the LSPs are identical to their previous versions). If this
can be achieved, the new versions may not even cause SPF to be run in
other systems. However, provided the same information is included in
the set of LSPs (albeit in a different order, and possibly different
LSPs), the result of running the SPF will be the same and will not
cause churn to the forwarding tables.

In the case of a restarting router, none of the router's own LSPs are
transmitted, nor are the router's own forwarding tables updated while
the timer T3 is running.

Redistribution of inter-level information MUST be regenerated before
this router's LSP is flooded to other nodes. Therefore, the Level-n
non-pseudonode LSP(s) MUST NOT be flooded until the other level's T2
timer has expired and its SPF has been run. This ensures that any
inter-level information which is to be propagated can be included in
the Level-n LSP(s).

During this period, if one of the router's own (including
pseudonodes) LSPs is received, which the local router does not
currently have in its own database, it is NOT purged. Under normal
operation, such an LSP would be purged, since the LSP clearly should
not be present in the global LSP database. However, in the present
circumstances, this would be highly undesirable, because it could
cause premature removal of a router's own LSP - and hence churn in
remote routers. Even if the local system has one or more of the
router's own LSPs (which it has generated, but not yet transmitted),
it is still not valid to compare the received LSP against this set,
since it may be that as a result of propagation between Level 1 and
Level 2 (or vice versa), a further router's own LSP will need to be
generated when the LSP databases have synchronized.

During this period a restarting router SHOULD send CSNPs as it
normally would. Information about the router's own LSPs MAY be
included, but if it is included it MUST be based on LSPs which have
been received, not on versions which have been generated (but not yet
transmitted). This restriction is necessary to prevent premature
removal of an LSP from the global LSP database.

When the timer T2 expires or is cancelled indicating that
synchronization for that level is complete, the SPF for that level is
run in order to derive any information which is required to be
propagated to another level, but the forwarding tables are not yet
updated.

Once the other level's SPF has run and any inter-level propagation
has been resolved, the router's own LSPs can be generated and
flooded. Any own LSPs which were previously ignored, but which are
not part of the current set of own LSPs (including pseudonodes) MUST
then be purged. Note that it is possible that a Designated Router
change may have taken place, and consequently the router SHOULD purge
those pseudonode LSPs which it previously owned, but which are now no
longer part of its set of pseudonode LSPs.

When all the T2 timers have expired or been cancelled, the timer T3
is cancelled and the local forwarding tables are updated.

If the timer T3 expires before all the T2 timers have expired or been
cancelled, this indicates that the synchronization process is taking
longer than the minimum holding time of the neighbors. The router's
own LSP(s) for levels which have not yet completed their first SPF
computation are then flooded with the overload bit set to indicate
that the router's LSPDB is not yet synchronized (and therefore other
routers MUST NOT compute routes through this router). Normal
operation of the update process resumes and the local forwarding
tables are updated. In order to prevent the neighbor's adjacencies

from expiring, IIHs with the normal interface value for the holding
time are transmitted over all interfaces with neither RR nor RA set
in the restart TLV. This will cause the neighbors to refresh their
adjacencies. The router's own LSP(s) will continue to have the
overload bit set until timer T2 has expired or been cancelled.

3.4.1.2. Starting

In the case of a starting router, as soon as each adjacency is
established, and before any CSNP exchanges, the router's own zeroth
LSP is transmitted with the overload bit set. This prevents other
routers from computing routes through the router until it has
reliably acquired the complete set of LSPs. The overload bit remains
set in subsequent transmissions of the zeroth LSP (such as will occur
if a previous copy of the router's own zeroth LSP is still present in
the network) while any timer T2 is running.

When all the T2 timers have been cancelled, the router's own LSP(s)
MAY be regenerated with the overload bit clear (assuming the router
is not in fact overloaded, and there is no other reason, such as
incomplete BGP convergence, to keep the overload bit set) and flooded
as normal.

Other LSPs owned by this router (including pseudonodes) are generated
and flooded as normal, irrespective of the timer T2. The SPF is also
run as normal and the RIB and FIB updated as routes become available.

To avoid the possible formation of temporary blackholes, the starting
router sets the SA bit in the restart TLV (as described in 3.3.2) in
all IIHs that it sends.

When all T2 timers have been cancelled, the starting router MUST
transmit IIHs with the SA bit clear.

4. State Tables

This section presents state tables which summarize the behaviors
described in this document. Other behaviors, in particular adjacency
state transitions and LSP database update operation, are NOT included
in the state tables except where this document modifies the behaviors
described in [2] and [4].

The states named in the columns of the tables below are a mixture of
states that are specific to a single adjacency (ADJ suppressed, ADJ
Seen RA, ADJ Seen CSNP) and states which are indicative of the state
of the protocol instance (Running, Restarting, Starting, SPF Wait).

   Three state tables are presented from the point of view of a running
   router, a restarting router, and a starting router.

4.1. Running Router

```
   Event        | Running             | ADJ suppressed
   ==============================================================
    RX RR       | Maintain ADJ State  |
                | Send RA             |
                | Set SRM,send CSNP   |
                |  (Note 1)           |
                | Update Hold Time,   |
                |  set Restart Mode   |
                |  (Note 2)           |
   -------------+---------------------+-----------------------
    RX RR clr   | Clr Restart mode    |
   -------------+---------------------+-----------------------
    RX SA       | Suppress IS neighbor|
                |   TLV in LSP(s)     |
                | Goto ADJ Suppressed |
   -------------+---------------------+-----------------------
    RX SA clr   |                     |Unsuppress IS neighbor
                |                     |   TLV in LSP(s)
                |                     |Goto Running
   ==============================================================
```

   Note 1: CSNPs are sent by routers in accordance with Section 3.2.1c
   Note 2: If Restart Mode clear

4.2. Restarting Router

Event	Restarting	ADJ Seen RA	ADJ Seen CSNP	SPF Wait
Router restarts	Send IIH/RR ADJ Init Start T1,T2,T3			
RX RR	Send RA			
RX RA	Adjust T3 Goto ADJ Seen RA		Cancel T1 Adjust T3	
RX CSNP set	Goto ADJ Seen CSNP	Cancel T1		
RX IIH w/o Restart TLV	Cancel T1 (Point- to-point only)			
T1 Expires	Send IIH/RR Restart T1	Send IIH/RR Restart T1	Send IIH/RR Restart T1	
T1 Expires nth time	Send IIH/ normal	Send IIH/ normal	Send IIH/ normal	
T2 expires	Trigger SPF Goto SPF Wait			
T3 expires	Set OL Flood local LSPs Update fwd plane			
LSP DB Sync	Cancel T2, and T3 Trigger SPF Goto SPF wait			
All SPF done				Clear OL Update fwd plane Flood local LSPs Goto Running

4.3. Starting Router

Event	Starting	ADJ Seen RA	ADJ Seen CSNP
Router starts	Send IIH/SA Start T1,T2		
RX RR	Send RA		
RX RA	Goto ADJ Seen RA		Cancel T1
RX CSNP Set	Goto ADJ Seen CSNP	Cancel T1	
RX IIH w no Restart TLV	Cancel T1 (Point-to-Point only)		
ADJ UP	Start T1 Send local LSPs w OL		
T1 Expires	Send IIH/RR and SA Restart T1	Send IIH/RR and SA Restart T1	Send IIH/RR and SA Restart T1
T1 Expires nth time	Send IIH/SA	Send IIH/SA	Send IIH/SA
T2 expires	Clear OL Send IIH normal Goto Running		
LSP DB Sync	Cancel T2 Clear OL Send IIH normal		

5. Security Considerations

   Any new security issues raised by the procedures in this document
   depend upon the ability of an attacker to inject a false but
   apparently valid IIH, the ease/difficulty of which has not been
   altered.

If the RR bit is set in a false IIH, neighbors who receive such an
IIH will continue to maintain an existing adjacency in the "UP" state
and may (re)send a complete set of CSNPs. While the latter action is
wasteful, neither action causes any disruption in correct protocol
operation.

If the RA bit is set in a false IIH, a (re)starting router which
receives such an IIH may falsely believe that there is a neighbor on
the corresponding interface which supports the procedures described
in this document. In the absence of receipt of a complete set of
CSNPs on that interface, this could delay the completion of (re)start
procedures by requiring the timer T1 to time out the locally defined
maximum number of retries. This behavior is the same as would occur
on a LAN where none of the (re)starting router's neighbors support
the procedures in this document and is covered in Sections 3.3.1 and
3.3.2.

If an SA bit is set in a false IIH, this could cause suppression of
the advertisement of an IS neighbor which could either continue for
an indefinite period, or occur intermittently with the result being a
possible loss of reachability to some destinations in the network
and/or increased frequency of LSP flooding and SPF calculation.

The possibility of IS-IS PDU spoofing can be reduced by the use of
authentication as described in [1] and [2], and especially the use of
cryptographic authentication as described in [5].

6. IANA Considerations

   This document defines the following IS-IS TLV that is listed in the
   IS-IS TLV code-point registry:

Type	Description	IIH	LSP	SNP
211	Restart TLV	y	n	n

7. Normative References

   [1] Callon, R., "OSI IS-IS for IP and Dual Environment", RFC 1195,
        December 1990.

   [2] ISO, "Intermediate system to Intermediate system routeing
        information exchange protocol for use in conjunction with the
        Protocol for providing the Connectionless-mode Network Service
        (ISO 8473)," ISO/IEC 10589:2002, Second Edition.

   [3] Bradner, S., "Key words for use in RFCs to Indicate Requirement
        Levels", BCP 14, RFC 2119, March 1997.

   [4] Katz, D. and R. Saluja, "Three-Way Handshake for IS-IS Point-
        to-Point Adjacencies", RFC 3373, September 2002.

   [5] Li, T. and R. Atkinson, "Intermediate System to Intermediate
        System (IS-IS) Cryptographic Authentication", RFC 3567, July
        2003.

8. Acknowledgements

   The authors would like to acknowledge contributions made by Jeff
   Parker, Radia Perlman, Mark Schaefer, Naiming Shen, Nischal Sheth,
   Russ White, and Rena Yang.

9. Authors' Addresses

   Mike Shand
   Cisco Systems
   250 Longwater Avenue,
   Reading,
   Berkshire,
   RG2 6GB
   UK
   Phone: +44 208 824 8690

   EMail: mshand@cisco.com

   Les Ginsberg
   Cisco Systems
   510 McCarthy Blvd.
   Milpitas, Ca. 95035 USA

   EMail: ginsberg@cisco.com

10. Full Copyright Statement

Intellectual Property

Acknowledgement