Routing Area Working Group
Internet-Draft
Intended status: Standards Track
Expires: September 5, 2015

S. Litkowski
Orange Business Service

March 4, 2015

Link State protocols SPF trigger and delay algorithm impact on IGP microloops

draft-litkowski-rtgwg-spf-uloop-pb-statement-02

Abstract

A micro-loop is a packet forwarding loop that may occur transiently among two or more routers in a hop-by-hop packet forwarding paradigm.

In this document, we are trying to analyze the impact of using different Link State IGP implementations in a single network in regards of microloops. The analysis is focused on the SPF triggers and SPF delay algorithm.

Requirements Language

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 [RFC2119].

Table of Contents

1. Introduction

Link State IGP protocols are based on a topology database on which a SPF (Shortest Path First) algorithm like Dijkstra is implemented to find the optimal routing paths.

Specifications like IS-IS ([RFC1195]) propose some optimization of the route computation (See Appendix C.1) but not all the implementations are following those not mandatory optimizations.

We will call SPF trigger, the events that would lead to a new SPF computation based on the topology.

Link State IGP protocols, like OSPF ([RFC2328]) and IS-IS ([RFC1195]), are using plenty of timers to control the router behavior in case of churn : SPF delay, PRC delay, LSP generation delay, LSP flooding delay, LSP retransmission interval ...

Some of those timers are standardized in protocol specification, some are not especially the SPF computation related timers.

For non standardized timers, implementations are free to implement it in any way. For some standardized timer, we can also see that rather than using static configurable values for such timer , implementations may offer dynamically adjusted timers to help controlling the churn.

We will call SPF delay, the delay timer that exists in most implementations that makes codes to wait before running SPF computation after a SPF trigger is received.

A micro-loop is a packet forwarding loop that may occur transiently among two or more routers in a hop-by-hop packet forwarding paradigm. We can observe that these micro-loops are formed when two routers do not update their Forwarding Information Base (FIB) for a certain prefix at the same time. The micro-loop phenomenon is described in [I-D.ietf-rtgwg-microloop-analysis].

Routers have more and more powerful controlplane and dataplane that reduce the Control plane to Forwarding plane overhead during the convergence process. Even if FIB update is still reasonably the highest contributor in the convergence time for large network, its duration is reducing more and more and may become comparable to protocol timers. This is particular true in small and medium networks.

In multi vendor networks, using different implementations of a link state protocol may favor micro-loops creation during convergence time due to deprecancies of timers. Service Providers are already aware to use similar timers for all the network as best practice, but sometimes it is not possible due to limitation of implementations.

This document will present why it sounds important for service provider to have consistent implementations of Link State protocols across vendors. We are particularly analyzing the impact of using different Link State IGP implementations in a single network in regards of microloops. The analysis is focused on the SPF triggers and SPF delay algorithm in a first step.

This document is only stating the problem, and defining some work items but its not intented to provide a solution.

2. Problem statement

```
                                    A ---- B
                                    |      |
                                 10 |      | 10
                                    |      |
                                    C ---- D
                                    |  2   |
                                    Px     Px
```

Figure 1

In the figure above, A uses primarily the AC link to reach C. When the AC link fails, IGP convergence occurs. If A converges before B, A will forward traffic to C through B, but as B as not converged yet, B will loop back traffic to A, leading to a microloop.

The micro-loop appears due to the asynchronous convergence of nodes in a network when a event occurs.

Multiple factors (and combination of these factors) may increase the probability for a micro-loop to appear :

- o delay of failure notification : the more B is advised of the failure later than A, the more a micro-loop may appear.

- o SPF delay : most of the implementations supports a delay for the SPF computation to try to catch as many events as possible. If A uses a SPF delay timer of x msec and B uses a SPF delay timer of y msec and x < y, B would start converging after A leading to a potential microloop.

- o SPF computation time : mostly a matter of CPU power and optimizations like incremental SPF. If A computes SPF faster than B, there is a chance for a microloop to appear. CPUs are today faster enough to consider SPF computation time as negligeable (order of msec in a large network).

- o RIB and FIB prefix insertion speed or ordering : highly implementation dependant.

This document will focus on analysis SPF delay (and associated triggers).

3. SPF trigger strategies

Depending of the change advertised in LSP/LSA, the topology may be affected or not. An implementation can decide to not run SPF (and only run IP reachability) if the advertised change is not affecting topology.

Different strategies exists to trigger SPF :

1. Always run full SPF whatever the change to process.

2. Run only Full SPF when required : e.g. if a link fails, a local node will run an SPF for its local LSP update. If the LSP from the neighbor (describing the same failure) is received after SPF has started, the local node can decide that a new full SPF is not required as the topology has not change.

3. If topology does not change, only recompute reachability.

As pointed in Section 1, SPF optimization are not mandatory in specifications, leading to multiple strategies to be implemented.

4. SPF delay strategies

Implementations of link state routing protocols use different strategies to delay SPF :

1. Two steps.

2. Exponential backoff.

4.1. Two step SPF delay

The SPF delay is managed by four parameters :

- Rapid delay : amount of time to wait before running SPF.

- Rapid runs : amount of consecutive SPF runs that can run using rapid delay. When amount is exceeded router moves to slow delay.

- Slow delay : amount of time to wait before running SPF.

- Wait time : amount of time to wait without events before going back to rapid delay.

Example : Rapid delay = 50msec, Rapid runs = 3, Slow delay = 1sec, Wait time = 2sec

```
   SPF delay time
       ^
       |
       |
   SD- |             x xx x
       |
       |
       |
   RD- |   x  x   x                   x
       |
       +---------------------------------> Events
           |  |   |  | || |            |
                            < wait time >
```

4.2. Exponential backoff

The algorithm has two mode : fast mode and backoff mode. In backoff mode, the SPF delay is increasing exponentially at each run. The SPF delay is managed by four parameters :

- o First delay : amount of time to wait before running SPF. This delay is used on when SPF is in fast mode.

- o Incremental delay : amount of time to wait before running SPF. This delay is used on when SPF is in backoff mode and increments exponentially at each SPF run.

- o Maximum delay : maximum amount of time to wait before running SPF.

- o Wait time : amount of time to wait without events before going back to fast mode.

Example : First delay = 50msec, Incremental delay = 50msec, Maximum delay = 1sec, Wait time = 2sec

```
   SPF delay time
       ^
   MD- |               xx x
       |
       |
       |
       |
       |
       |             x
       |
       |
       |
       |          x
       |
   FD- |   x  x                        x
   ID  |
       +---------------------------------> Events
           |  |   |  |  || |            |
                           < wait time >
          FM->BM -------------------->FM
```

5. Mixing strategies

```
                              S ---- E
                              |      |
                           10 |      | 10
                              |      |
                              D ---- A
                              |  2
                              Px
```

Figure 2

In the diagram above, we consider a flow of packet from S to D. We consider that S is using optimized SPF triggering (Full SPF is triggered only when necessary), and two steps SPF delay (rapid=150ms,rapid-runs=3, slow=1s). As implementation of S is optimized, Partial Reachability Computation (PRC) is available. We consider the same timers as SPF for delaying PRC. We consider that E is using a SPF trigger strategy that always compute Full SPF and exponential backoff strategy for SPF delay (start=150ms, inc=150ms, max=1s)

We also consider the following sequence of events (note : the timescale does not intend to represent a real router timescale where jitters are introduced to all timers) :

- t0 : a prefix is declared down in the network.
- t0+200ms : the prefix is declared as up.
- t0+400ms : a prefix is declared down in the network.
- t0+1000ms : S-D link fails.

```
  S timescale            E timescale              Event timescale
     |                      |                          |
     |                      |                          | <- t0 Event
     | Schedule PRC (150ms) | Schedule SPF (150ms)     |
         |                      |                          |
         |                      |                          |
         |                      |                          |
         | PRC starts           | SPF starts               |
         | PRC ends             |                          |
         | RIB/FIB starts       | SPF ends                 |
     |                      | RIB/FIB starts           |
         | RIB/FIB ends         |                          |
         |                      | RIB/FIB ends             | t0+180ms
         |                      |                          |
         |                      |                          | < - t0+200ms Event
     | Schedule PRC (150ms) | Schedule SPF (150ms)     |
     |                      |                          |
         |                      |                          |
         |                      |                          |
         | PRC starts           | SPF starts               |
         | PRC ends             |                          |
         | RIB/FIB starts       | SPF ends                 |
     |                      | RIB/FIB starts           |
         | RIB/FIB ends         |                          |
         |                      | RIB/FIB ends             | t0+380ms
     |                      |                          | < - t0+400ms Event
     | Schedule PRC (300ms) | Schedule SPF (300ms)     |
     |                      |                          |
         |                      |                          |
         |                      |                          |
     |                      |                          |
         |                      |                          |
         |                      |                          |
         | PRC starts           | SPF starts               |
         | PRC ends             |                          |
         | RIB/FIB starts       | SPF ends                 |
     |                      | RIB/FIB starts           |
         | RIB/FIB ends         |                          |
         |                      | RIB/FIB ends             | t0+730ms
         |                      |                          |
```

```
        |                       |                         |
        |                       |                         |
    |                       |                         |
        |                       |                         | < - t0+1000ms Event
    | Schedule SPF (150ms)  | Schedule SPF (600ms)    |
    |                       |                         |
        |                       |                         |
        | SPF starts            |                         |
    |                       |                         |
        | SPF ends              |                         |
        | RIB/FIB starts        |                         |
        |                       |                         |      }
        | RIB/FIB ends          |                         |  }
        |                       |                         |  }
    |                       |                         |  }
        |                       |                         |  }
        |                       |                         |  }
        |                       |                         |  } Micro-loop creati
on
        |                       |                         |  }
        |                       | SPF starts              |  }
        |                       |                         |  }
        |                       | SPF ends                |  }
        |                       | RIB/FIB starts          |  }
        |                       |                         |  }
        |                       | RIB/FIB ends            |  }

                       Figure 3
```

In the figure above, we can see that due to deprecancies in SPF management, after multiple events (different types of event), SPF delays are completely misaligned between nodes leading to long microloop creation.

The same issue can also appear with only single type of events as displayed below :

```
  S timescale             E timescale               Event timescale
    |                       |                         |
    |                       |                         | < - t0 Event remote link
 down
    | Schedule SPF (150ms)  | Schedule SPF (150ms)    |
        |                       |                         |
        |                       |                         |
        |                       |                         |
        | PRC starts            | SPF starts              |
        | PRC ends              |                         |
        | RIB/FIB starts        | SPF ends                |
    |                       | RIB/FIB starts          |
```

```
        | RIB/FIB ends          |                          |
        |                       | RIB/FIB ends             | t0+180ms
        |                       |                          |
        |                       |                          | < - t0+200ms Event r
emote link down
    | Schedule SPF (150ms)  | Schedule SPF (150ms)     |
    |                       |                          |
        |                       |                          |
        |                       |                          |
        | SPF starts            | SPF starts               |
        | SPF ends              |                          |
        | RIB/FIB starts        | SPF ends                 |
    |                       | RIB/FIB starts           |
        | RIB/FIB ends          |                          |
        |                       | RIB/FIB ends             | t0+380ms
    |                       |                          | < - t0+400ms Event remot
e link change
    | Schedule SPF (150ms)  | Schedule SPF (300ms)     |
    |                       |                          |
        |                       |                          |
        | SPF starts            |                          |
    |                       |                          |
        | SPF ends              |                          |
        | RIB/FIB starts        |                          |
        |                       | SPF starts               |  }
        | RIB/FIB ends          |                          |  }
        |                       | SPF ends                 |  } micro-loop creati
on
    |                       | RIB/FIB starts           |  }
        |                       |                          |  }
        |                       | RIB/FIB ends             | t0+730ms
        |                       |                          |
        |                       |                          |
        |                       |                          |
    |                       |                          |
        |                       |                          | < - t0+1000ms Event
    | Schedule SPF (1s)     | Schedule SPF (600ms)     |
    |                       |                          |
        |                       |                          |
        |                       |                          |
    |                       |                          |
        |                       |                          |
        |                       |                          |
        |                       |                          |
        |                       |                          |
        |                       |                          |
    |                       |                          |
        |                       |                          |
        |                       |                          |
        |                       |                          |
        |                       |                          |
```

```
        |                       | SPF starts               |
        |                       |                          |
        |                       | SPF ends                 |
        |                       | RIB/FIB starts           |
        |                       |                          |  }
        |                       | RIB/FIB ends             |  }
        |                       |                          |  }
        |                       |                          |  }
        |                       |                          |  } microloop creatio
n
        |                       |                          |  }
        |                       |                          |  }
        |                       |                          |  }
        | SPF starts            |                          |  }
    |                       |                          |  }
        | SPF ends              |                          |  }
        | RIB/FIB starts        |                          |  }
        |                       |                          |  }
        | RIB/FIB ends          |                          | t0 + 2030ms
```

Figure 4

6. Proposed work items

In order to enhance the current LinkState IGP behavior, authors would encourage working on standardization of some behaviors.

Authors are proposing the following work items :

- o Standardize SPF trigger strategy.
- o Standardize computation timer scope : single timer for all computation operations, separated timers ...
- o Standardize "slowdown" timer algorithm including its association to a particular timer : authors of this document does not presume that the same algorithm must be used for all timers.

Using the same event sequence as in figure 2, we may expect fewer and/or shorter microloops using standardized implementations.

```
          S timescale             E timescale               Event timescale
    |                       |                          |
    |                       |                          | < - t0 Event
    | Schedule PRC (150ms)  | Schedule PRC (150ms)     |
        |                       |                          |
        |                       |                          |
```

```
        |                       |                         |
        | PRC starts            | PRC starts              |
        | PRC ends              |                         |
        | RIB/FIB starts        | PRC ends                |
    |                       | RIB/FIB starts          |
        | RIB/FIB ends          |                         |
        |                       | RIB/FIB ends            | t0+180ms
        |                       |                         |
        |                       |                         | < - t0+200ms Event
    | Schedule PRC (150ms)  | Schedule PRC (150ms)    |
    |                       |                         |
        |                       |                         |
        |                       |                         |
        | PRC starts            | PRC starts              |
        | PRC ends              |                         |
        | RIB/FIB starts        | PRC ends                |
    |                       | RIB/FIB starts          |
        | RIB/FIB ends          |                         |
        |                       | RIB/FIB ends            | t0+380ms
    |                       |                         | < - t0+400ms Event
    | Schedule PRC (300ms)  | Schedule PRC (300ms)    |
    |                       |                         |
        |                       |                         |
        |                       |                         |
    |                       |                         |
        |                       |                         |
        |                       |                         |
        | PRC starts            | PRC starts              |
        | PRC ends              |                         |
        | RIB/FIB starts        | PRC ends                |
    |                       | RIB/FIB starts          |
        | RIB/FIB ends          |                         |
        |                       | RIB/FIB ends            | t0+730ms
        |                       |                         |
        |                       |                         |
        |                       |                         |
    |                       |                         |
        |                       |                         | < - t0+1000ms Event
    | Schedule SPF (150ms)  | Schedule SPF (150ms)    |
    |                       |                         |
        |                       |                         |
        | SPF starts            |  SPF starts             |
    |                       |                         |
        | SPF ends              |                         |
        | RIB/FIB starts        |  SPF ends               |
        |                       |  RIB/FIB starts         |      } microloop cre
ation
        | RIB/FIB ends          |                         |  }
        |                       |  RIB/FIB ends           |
```

```
    |                        |                          |
        |                        |                          |
```

Figure 5

As displayed above, there could be some other parameters like router computation power, flooding timers that may also influence microloops. In the figure 5, we consider E to be a bit slower than S, leading to microloop creation. Despite of this, we expect that by aligning implementations at least on SPF trigger and SPF delay, service provider may reduce number or duration of microloops.

7. Security Considerations

This document does not introduce any security consideration.

8. Acknowledgements

9. IANA Considerations

This document has no action for IANA.

10. Normative References

[I-D.ietf-rtgwg-microloop-analysis]
Zinin, A., "Analysis and Minimization of Microloops in Link-state Routing Protocols", draft-ietf-rtgwg-microloop-analysis-01 (work in progress), October 2005.

[RFC1195] Callon, R., "Use of OSI IS-IS for routing in TCP/IP and dual environments", RFC 1195, December 1990.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, April 1998.

Author’s Address

Stephane Litkowski
Orange Business Service

Email: stephane.litkowski@orange.com