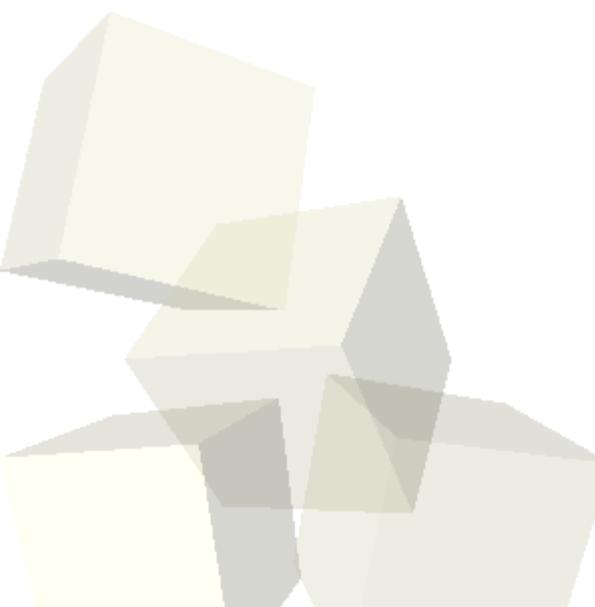




# PIM adjacencies and multicast blackhole mitigation issues

[draft-morin-mboned-mcast-blackhole-mitigation-01](#)

Thomas Morin  
Gregory Cauchie  
(Orange Labs)





# Context / Problem statement

## ■ Context

- ◆ Multicast more and more deployed
- ◆ Focus on multicast QoS / convergence

## ■ It can occur that the unicast routing advertises a link while the PIM-SM adjacency on a link is not ready yet, e.g.:

- ◆ if PIM Hellos were not exchanged yet
- ◆ or if PIM is not configured on both sides (not yet, misconfig)
- ◆ some bug or issue in the setup of the adjacency

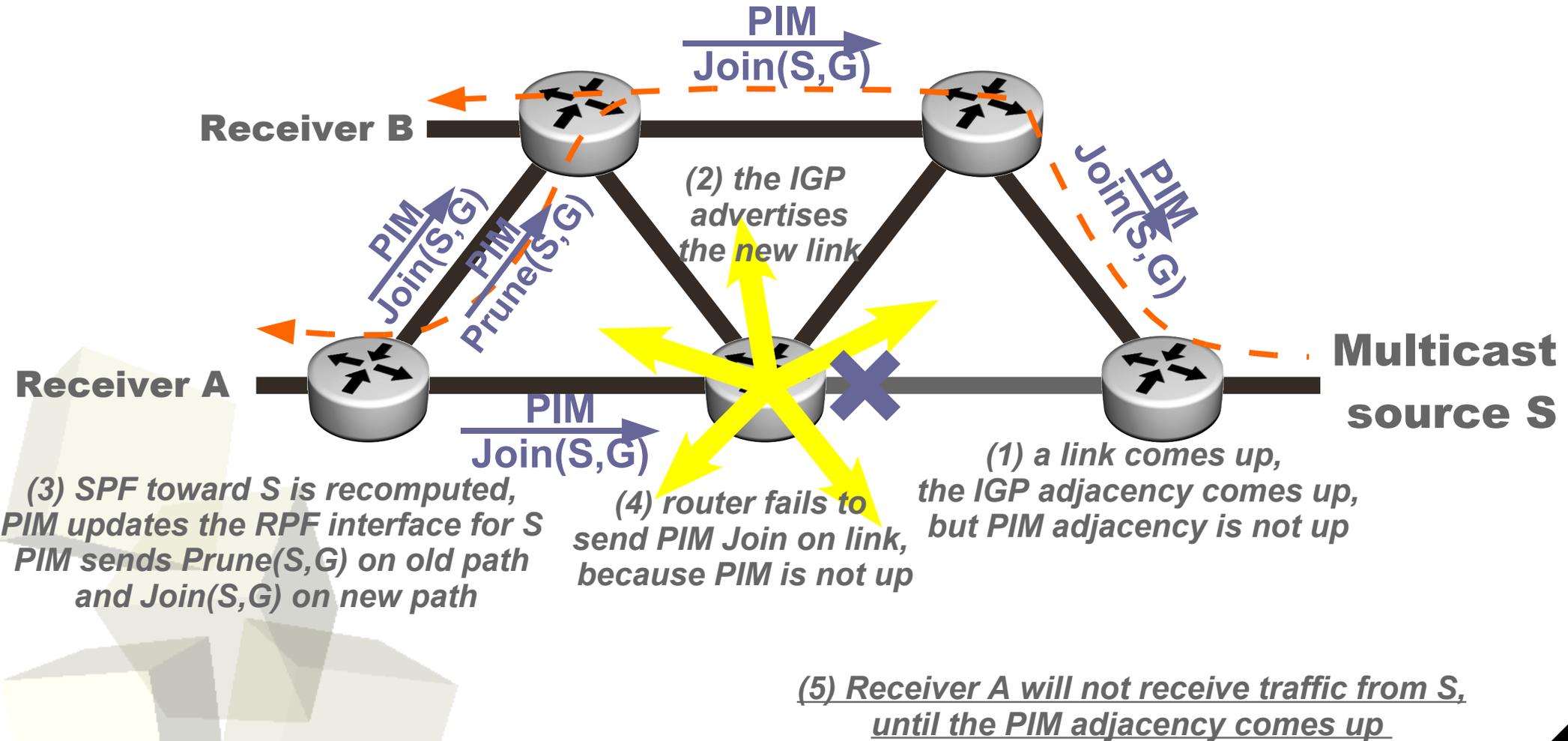
## ■ What happens:

- ◆ the SPF computed by unicast routing uses a link on which PIM is not ready
- ◆ PIM Joins propagate along this path...
- ◆ ...but fail at the router before that link...
- ◆ **...resulting in a traffic blackhole**



# Illustration

(0) Initially, A and B receive multicast sent by multicast source S toward group address G.





- What the PIM-SM specs say (RFC4601):
  - ♦ when a link comes up, wait [0-5s] before sending a Hello
    - delay is overridden if the router need to send a Join on the link
  - ♦ **a router neighbor waits [0-5s] before sending a Hello in reply to a new neighbor**
  - ♦ if need to send a Join to a neighbor and no Hello was sent yet on the interface, send a Hello now before sending the Joins
  - ♦ not discussed:
    - *nothing said about whether or not a router needs to have received a Hello from a neighbor before sending a Join*
- Ignoring Hellos ?
  - ♦ a router might send/process a join without requiring having exchanged Hellos with the neighbor
  - ♦ **but** Hellos carry options that are meant to be extended, and new options may impact how Joins are sent or processed
  - ♦ ignoring Hello seem only viable as a temporary workaround to the issue

# Improving PIM adjacency setup time [1/2]

## ■ Improving adjacency setup time

- ♦ reason for the Triggered\_Hello\_Delay is avoiding “Hello storms”
  - is it needed on point to point links ?
  - LAN case: can we find a way to improve adjacency setup time AND avoid Hello storms ?
- ♦ proposal here reflects discussion with Bill Fenner and Mark Handley and Dino Farinacci in Vancouver

## ■ Point to point case

- ♦ propose a lower default Triggered\_Hello\_Delay for links that are point to point : 50ms ?

## ■ LAN case

- ♦ when a downstream router needs to send a Join to an upstream router with which Hellos haven't been exchanged yet, it could unicast a Hello to this upstream router, and the neighbor would unicast back a Hello, after what the downstream router can send his Joins

## ■ Side note on the LAN case

- ♦ some router do implement the workaround consisting in ignoring Hellos before sending or processing a Join
- ♦ the following case would be possible
  - downstream router implements the improved procedures with unicast Hellos
  - upstream router implements the temporary workaround, but won't send back an instant unicast Hello
  - what would happen is that the downstream router will wait for the unicast Hello in reply to his, for up to 5s, and only then will send his Joins, though the upstream would have successfully processed them earlier !
- ♦ fix
  - wait for the unicast Hello “reply”, **but not too long**
  - after sending the unicast Hello, initiate a short timer (e.g 100ms), and send the Joins if the timer expires before the unicast Hello has been received from the neighbor



# Back to the blackholing issues

- Improving PIM Adjacency setup time solves one of the blackholing cases
- But not all:
  - ◆ PIM configuration mismatch
    - PIM not configured on the link
    - not configured on both sides yet
  - ◆ Any kind of issue where...
    - PIM Hello are not properly exchanged (bug)
    - issue/bug with a PIM process
- What can be done to cover such cases ?





# Back to the blackholing issues

- We would like a minimal impact on unicast
  - ♦ raising an IGP link cost before PIM is ready on a link would have an impact on unicast routing
- A possible solution is to...
  - ♦ use a multi-topology IGP (or multi-instance)
  - ♦ make PIM follow the multicast-dedicated IGP topology
  - ♦ make the IGP use some “PIM adjacency ready” condition to advertise/not-advertise a link in the multicast topology
- Advantages
  - ♦ low impact on unicast routing
    - only increase the IGP LSP size
  - ♦ purely local behavior
  - ♦ no need to extend IGP specifications



# When to advertise a link?

- Multiple possible criteria to decide to advertise a link in the multicast-dedicated topology
- By increasing order of expected implementation “complexity” :
  - (1) **have PIM be configured on the link**
    - **enough to cover the PIM configuration mismatch cases**
  - (2) having sent and received PIM Hellos on the link
    - could cover possible bugs
  - (3) neighbor not currently being in graceful restart operations
    - ??
  - More...?
- A possible conclusion
  - (1) is easy/cheap to do, do it all the time
  - implement (2) if you don't have a solution to have quick PIM adjacency setup
  - Implement (3) or more, if you have needs/drivers for it



- Not specific to an IGP, the same problem happens with BGP unicast routing:
  - ◆ a BGP neighbor advertise a route to a unicast source on a link where PIM is not ready yet
- The proposed approach can be generalized:
  - ◆ use non-congruent unicast routing
    - in an IGP : use multi-topology IGP (or multi-instance)
    - in the i/eBGP case : use SAFI 2 BGP routes
    - applicable to the context of multicast in a VPN (SAFI 129)
  - ◆ take into account the PIM status on a link to...
    - ...advertise the link in the IGP (IGP case)
    - ...accept/advertise BGP routes on this link (BGP case)





- Proposed approach would be to
  - ◆ Update PIM specs to provide an improved adjacency setup time
  - ◆ Propose a BCP document recommending to
    - apply criteria (1) when an MT IGP is used for PIM routing
    - apply criteria (2) when an MT IGP is used for PIM routing and improved setup time procedures aren't provided
- Question to the working group:
  - ◆ Interest in adopting a draft PIM improving adjacency setup time ?
  - ◆ Feeling on which is the right working group for a BCP doc : pim ? mboned ?
- Feedback is welcome on the proposed approach

*Questions ? Comments ?*