Prefix Unreachable Announcement

draft-wang-lsr-prefix-unreachable-annoucement-06

A. Wang (China Telecom)
G. Mishra (Verizon)
Z. Hu (Huawei Technologies)
Y. Xiao (Huawei Technologies)

IETF-111, July 2021

PUA Mechanism

- Upon receiving the node/link failure information, which prefix is within the range of advertised summary address, the ABR or L1/L2 border router will:
  - Generate one new summary address, with the failure prefix associated, but set its originator information to NULL.
  - For ISIS, we use “IPv4/IPv6 Source Router ID” sub-TLV, which is defined in RFC 7794
  - For OSPF, we use “Prefix Originator Sub-TLV”, which is defined in draft-ietf-lsr-ospf-prefix-originator
  - Such summary message will be flooded across the boundary as normal OSPF/IS-IS procedures.

Updated Action based on PUA message

- For node failure scenario
  - When node within one area receives the PUA message from All of its ABRs, it will trigger the switchover of the control plane, which is run on top of it.

- For link failure/network partition scenario
  - When only some of the ABRs can't reach the failure prefix, the ABRs that can reach this prefix should advertise the specific route to this PUA prefix.
  - Same procedures as RIFT.

Conclusions and Further Actions

- Consensus on the scenarios
- Intense discussions on the list for the solution
  - Focus now on the control plane notification
  - Reuse the existing flooding message/mechanism
  - Increment deployment
  - Only the ABR and the receivers of PUA should aware such mechanism.

- Further Actions:
  - Enough interests on this topic.
  - Request the WG adoption call.
  - The implementation and deployment of PUA can give some guides for other solutions.
  - Detail introduction can refer the backup slides.

Passive Interface Attribute

draft-wang-lsr-passive-interface-attribute-08

Aijun Wang (China Telecom)

Zhibo Hu(Huawei)

Gyan Mishra(Verizon)

Jinsong Sun(ZTE)

IETF 111@Online, July 2021

What The Proposal Want To Solve?

- Passive Interfaces are used commonly in the network
  - Within data center, they are used for the VLAN interfaces that serving the layer 2 broadcast domain.
  - In the inter-AS boundary, they are used to protect each domain from IGP flapping that caused by other domain.
  - In the edge compute scenario that described by https://datatracker.ietf.org/doc/draft-dunbar-lsr-5g-edge-compute/00/,

  the passive interface is used to connect the server to the network.

- But currently, there is no suitable place to advertise the passive interfaces and their associated attributes. Refer to existing solutions slide

OSPFv2 Extension Proposal

- RFC7684 defines OSPFv2 Extended Link Opaque LSA to contain additional link attribute TLV
- Currently, only OSPFv2 Extended Link TLV is defined.
- We propose define OSPFv2 Extended Stub-Link TLV to contain the stub-link related sub TLV.

- In updated version 08, we add the “Link Type” and “Metric” fields to the newly defined TLV.
- The “Link Type” field can be used to indicate the subdivision of the Passive Interface, for further potential usages
- Existing sub-TLV that defined within "OSPFv2 Extended Link TLV Sub-TLV" can be included if necessary.

The OSPFv2 Extended Stub-Link TLV has the following format:

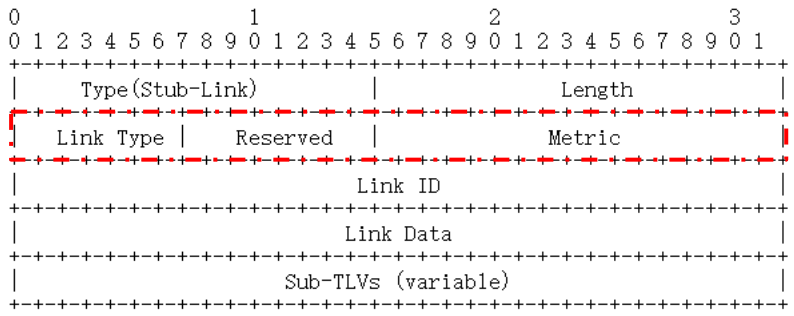


Figure 1: OSPFv2 Extended Stub-Link TLV

Type: The TLV type. The value is 2(TBD) for this stub-link type

Length: Variable, dependent on sub-TLVs

Link Type: Define the type of the stub-link. This document defines the followings type:

- o 0: Reserved
- o 1: AS boundary link
- o 2: Loopback link
- o 3: Vlan interface link
- o 4-255: For future extension

OSPFv3 and ISIS Extension Proposal

- For OSPFv3 and ISIS, the extension proposal are similar.
- OSPFv3, defines the Router-Stub-Link TLV to describe a single router passive interface. This TLV should only be contained within the E-Router-LSA.
- ISIS, one new top TLV(Stub-Link TLV) within ISIS is proposed, as the followings:

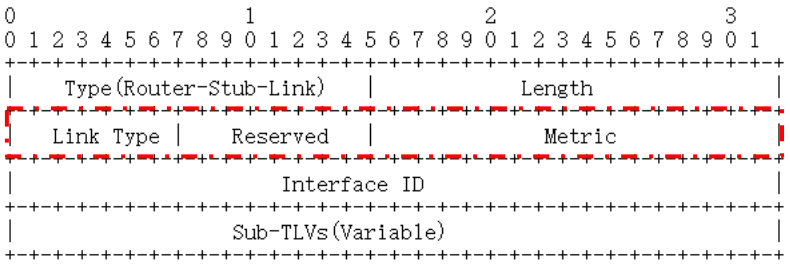


Figure 2: OSPFv3 Router-Stub-Link TLV

Type: OSPFv3 Extended-LSA TLV Type. Value is 10(TBD) for Router-Stub-Link TLV.

Length: Variable, dependent on sub-TLVs

Link Type: Define the type of the stub-link. This document defines the followings type:

- 0: Reserved
- 1: AS boundary link
- 2: Loopback link
- 3: Vlan interface link
- 4-255: For future extension

Metric: Link metric used for inter-AS traffic engineering.

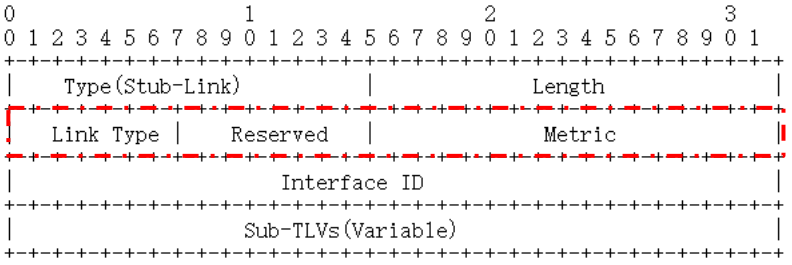


Figure 3: ISIS Stub-Link TLV

Type: ISIS TLV Codepoint. Value is 28(TBD) for stub-link TLV.

Length: Variable, dependent on sub-TLVs

Link Type: Define the type of the stub-link. This document defines the followings type:

- 0: Reserved
- 1: AS boundary link
- 2: Loopback link
- 3: Vlan interface link
- 4-255: For future extension

Metric: Link metric used for inter-AS traffic engineering.

Newly defined Sub-TLV

- One new sub-TLV to describe the IP address information that associated with the passive interface is defined:

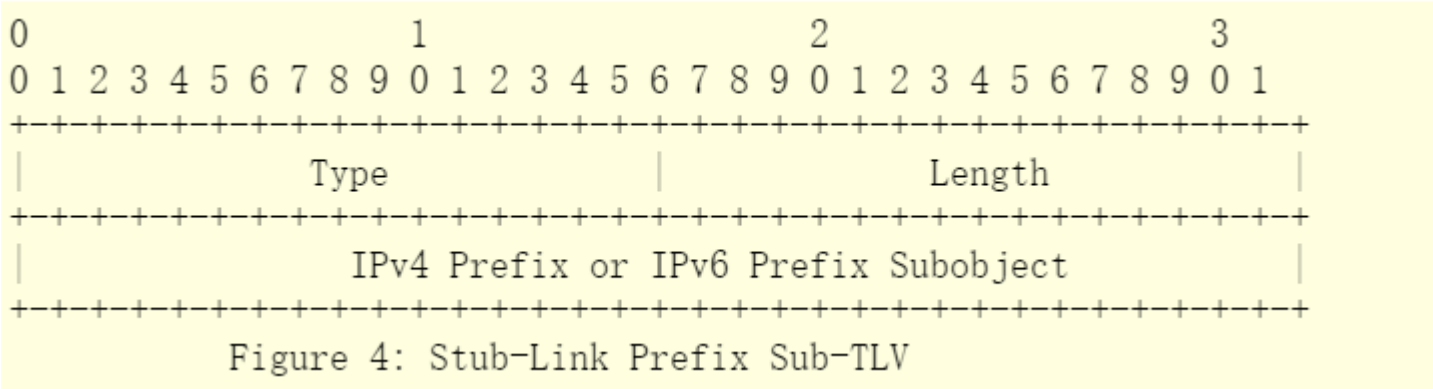


Figure 4: Stub-Link Prefix Sub-TLV

- Suboject is defined within RFC3209.
- Propose one independent Registry CodePoint “Stub-Link Attribute”, which can be referred by OSPFv2/v3 and ISIS

Further Plan

- Comments?
- Thanks all experts for past review(Acee, Peter, Tony, Les, Jeff etc).
- Co-Authors are welcome(Jinsong Sun from ZTE joined already).
- Adopt as WG Document?

Wangaj3@Chinatelecom.cn
Huzhibo@Huawei.com
Gyan.s.Mishra@Verizon.com
Sun.jinsong@ZTE.com.cn

IETF111@Online

Existing Possible Solutions

- ISIS(RFC5029) defines Link-Attribute Sub-TLV, but this sub-TLV can only be carried within the TLV 22, which is used to described the attached router.
- OSPFv2(RFC2328) defines link type within Router LSA, the type 3 can be used to describe the stub link(passive interface). But
- OSPFv3(RFC5340) has removed type 3 link type.

- **It is necessary to extend the OSPFv2/v3 and ISIS to transfer the passive interface and their related attributes**

Back to anchor

Reference Slides

Prefix Unreachable Announcement

draft-wang-lsr-prefix-unreachable-annoucement-06

A. Wang (China Telecom)
G. Mishra (Verizon)
Z. Hu (Huawei Technologies)
Y. Xiao (Huawei Technologies)

IETF-111, July 2021

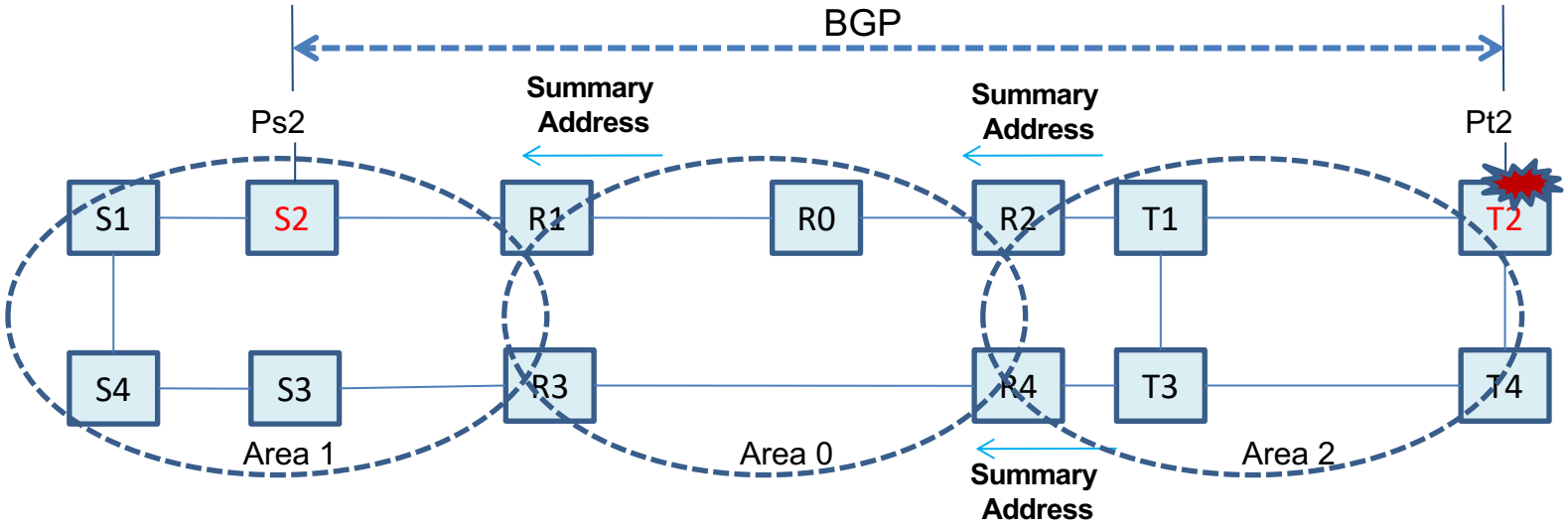
Motivation & Problem Statement

- Summarization of Inter-Area types routes propagated into the backbone area for flood reduction are made up of component prefixes.
- It is these component prefixes that the “Prefix Unreachability Announcement” tracks to ensure traffic is not “black hole” sink routed due to a PE failure.
- The PUA mechanism ensures immediate control plane convergence with PE node switchover when area is partitioned to avoid black hole of traffic.
- This draft provides a control plane signaling mechanism to detect the component prefix failures that are part of a summary prefix to force immediate control plane convergence to an alternate path.

Updated Contents

- Updated Scenarios
- Updated Action based PUA message
- Further Action

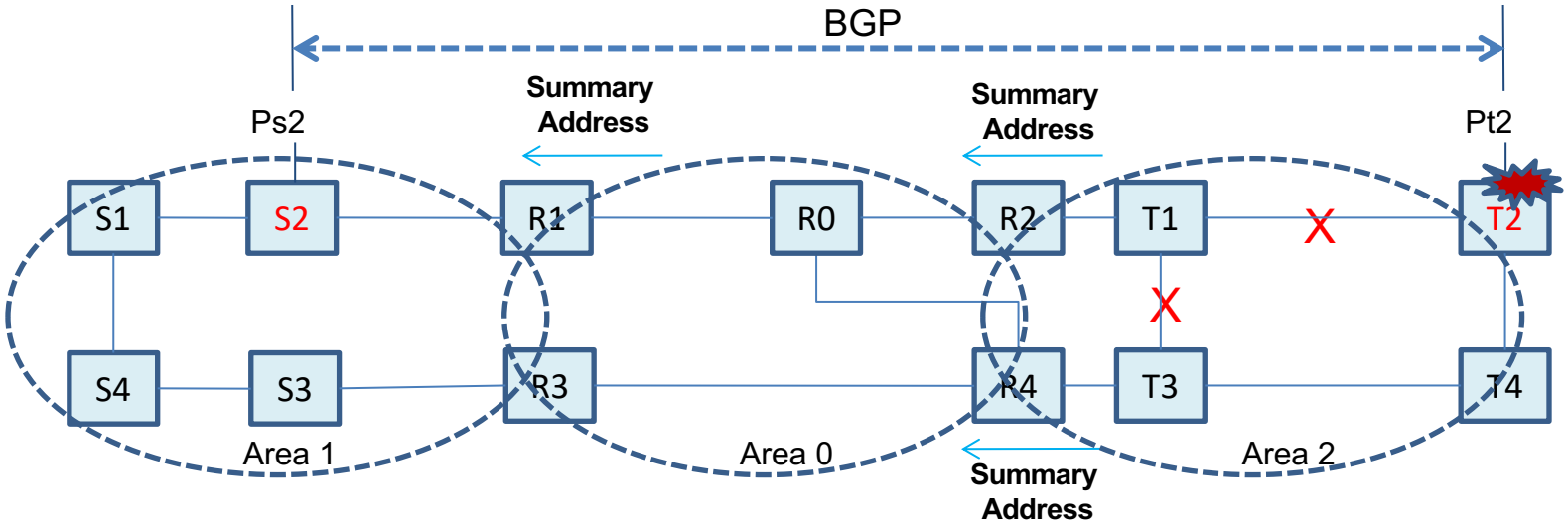
Updated Scenarios(1/2)



OSPF Prefix Unreachable Scenario (Node Failure)

- ✓ ABR R2/R4 do the summary action, send only the summary address to Area 0, Area 1.
- ✓ S2 has BGP session with T2, which provides the control connection for VPN services between them.
- ✓ When node T2 is failure, the summary address is still advertised. S2 doesn't know this until the BGP keep alive timer is timeout.
- ✓ Service Traffic will be breakout during this duration.

Updated Scenarios(2/2)



- ✓ ABR R2/R4 do the summary action, send only the summary address to Area 0, Area 1.
- ✓ S2 has BGP session with T2, which provides the control connection for VPN services between them.
- ✓ When link between T1/T2 and T1/T3 are broken, R2 can't reach T2, but it still announces the summary address. R0 still takes R2 as the next hop to T2.
- ✓ Traffic to T2 will be broken at ABR R2.

PUA Mechanism

- Upon receiving the node/link failure information, which prefix is within the range of advertised summary address, the ABR or L1/L2 border router will:
  - Generate one new summary address, with the failure prefix associated, but set its originator information to NULL.
  - For ISIS, we use “IPv4/IPv6 Source Router ID” sub-TLV, which is defined in RFC 7794
  - For OSPF, we use “Prefix Originator Sub-TLV”, which is defined in draft-ietf-lsr-ospf-prefix-originator
  - Such summary message will be flooded across the boundary as normal OSPF/IS-IS procedures.

Updated Action based on PUA message

- For scenario 1(node failure)
  - When node within one area receives the PUA message from All of its ABRs, it will trigger the switchover of the control plane, which is run on top of it.
  - For scenario 1, the BGP session between S2/T2 will be notified, S2 can then begin the BGP session switchover immediately.

- For scenario 2(link failure/network partition)
  - When only some of the ABRs can't reach the failure prefix, the ABRs that can reach this prefix should advertise the specific route to this PUA prefix.
  - Same procedures as RIFT.

Implementation Consideration

- Considering the balance of reachable information and unreachable information announcement capabilities, the implementation of this mechanism should set one MAX_Address_Announcement (MAA) threshold to control the advertisement of PUA and summary address.
  - If the number of unreachable prefixes is less than MAA, the ABR should advertise the summary address and the PUA.
  - If the number of reachable address is less than MAA, the ABR should advertise the detail reachable address only.
  - If the number of reachable prefixes and unreachable prefixes exceeds MAA, then advertises the summary address with MAX metric.

Further Action

- Comments?
- Adopt as WG document?

wangaj3@chinatelecom.cn
gyan.s.mishra@verizon.com
huzhibo@huawei.com
xiaoyaqun@Huawei.com

IETF111@Online(Virtual)