Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 27, 2015

N. Wu
Z. Li
Huawei
S. Hares
Hickory Hill Consulting
October 24, 2014

Use Cases for an Interface to IGP Protocol
draft-wu-i2rs-igp-usecases-01

Abstract

A link-state routing protocol such as OSPF or IS-IS is an essential component for a routing system. With substantial effort on the IGP protocols, the infrastructure of the network has achieved high reliability. During past years they have been operated and maintained through typical CLI, SNMP and NETCONF. As modern networks become larger and more complex, the IGP protocol may require a programmatic interface which is able to facilitate additional control and observation in such networks.

Interface to the Routing System's (I2RS) is a standards-based interface which provides a programmatic way to control and observe the IGP protocol. I2RS can be used to operate, maintain and monitor the routing-related state. This document describes set of use cases for which I2RS can be used for IGP protocol. It is intended to provide a base for the solution draft describing information models and a set of interfaces to the IGP protocol.

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 RFC 2119 [RFC2119].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any

time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 23, 2015.

Table of Contents

1. Introduction

A link-state routing protocol such as OSPF[RFC2328] or IS-IS[ISO.10589.1992] is an essential component for a routing system. With substantial effort of IGP protocol, the infrastructure of network has achieved high levels of reliability. During past years they have been operated and maintained through typical CLI, SNMP and NETCONF. As modern networks become larger and more complex, the IGP protocol may require a programmatic interface which is capable of facilitating additional control and observation in such networks.

Interface to the Routing System's (I2RS) [I-D.ietf-i2rs-architecture] architecture specifies common, standards-based programmatic interfaces which is an elegant way to control and observe the IGP protocol. The I2RS interface can be used to operate, maintain and monitor the routing-related state. The I2RS decribed here is aimed to co-exist with current control and diagnose mechanism such as CLI, SNMP and NETCONF instead of obseleting them. Acutally the I2RS can enhance these existing mechanism by defining a standardized set of programmatic interfaces to enable flexible manipulation, inquiry and analysis of the IGP protocol. The use cases described in this document cover the following aspects of IGP: network planning, path engineering and tracking of protocol events. The purpose here is to gain the rough consensus from the community that the I2RS IGP extensions fit within the overall I2RS architecture. It is intended to provide a base for the solution draft describing information models and a set of interfaces to the IGP protocol.

2. IGP Network Planning

With the growing size of modern network, more and more nodes and links in network are deployed with IGP protocol. A network containing 1000 IGP-enable nodes is not rare nowadays. As the consequence of this network inflation, some drawbacks can be easily introduced into the network. For example, link-state protocols depend on flooding mechanism to advertise link-state related information and keep the database updated. Too many nodes can periodically produce large amounts of link-state information which can burden the forwarding plane and jeopardize the reliability of IGP adjacencies. The number of adjacencies, links and routes involved into IGP network consumes forwarding and storage resources of the routing elements in the network. The I2RS Clients may be connected to by applications wishing to use the I2RS Client-Agent protocol to deploy IGP protocol in an efficient, scalable and interoperable manner.

2.1. Identification Allocation

IGP routers are identified by one identification (router-id or system-id) which MUST be unique for each router in the AS. It is increasingly common to observe that many subtle issues are introduced because of this identification conflict. Since this identification is inherited from interface ip address or configured manually, it is prone to conflict with another router located in remote network segment.

In the routing domain of "Bit Index Explicit Replication" (BIER)[I-D.wijnands-bier-architecture], it is essential for each Bit-Forwarding-Router(BFR) to have an unique BFR-id that MUST be in one specific numeric range. It is very likely that confliction can be observed quite often when those IDs are allocated in a distributed manner.

The I2RS MAY help to alleviate this situation by introducing certain application which is responsible for allocating identification. Though the mechanism used to allocate unique identification is out of the scope of this document.

2.2. Domain Partition

As stated above, huge network is harder to operate and maintain, what is more, is susceptible to topology turbulence which can degrage the quality of service provided by IGP protocol. Link-state protocols(OSPF or IS-IS) introduce routing hierarchy to solve this kind of problems. Some devices have limited CPU or storage resources and cannot hold all link-state information. These devices may need to be transferred to a limited IGP domain which holds part of the link-state information.

The I2RS may guide this partition process after considering different conditions including the number of routers, adjacency, links and routes, CPU and storage resouce of corresponding routers and also their geography location.

2.3. Route Manipulation

Searching entries in the Routing Information Base(RIB) is a fundamental operation in routing system. In order to speed up the searching process and saving storage resources, the RIB may contain only part of the routing table entries provided the network reachability is not compromised. The reduction of the routing table is achieved via route manipulation. The interface addresses of a router can be suppressed for sake of less entries or secure entries. The policy SHOULD be deployed carefully to summarize and filter those

routing information crossing the domain border through the way of generation or redistribution.

The I2RS SHOULD facilitate reduction by allowing offline calculation to determine how to partition IGPs and where to place ABR and ASBRs. The I2RS cycle of the query of IGP information (see above) followed by downloading of a new temporary topologies.

3. IGP Path Engineering

Link-state protocol like IGP depend on Shortest Path First(SPF) algorithm to calculate its path to destinations. These SPF paths can dynamically adapt to the topology change from time to time without external involvement. Though this traditional mechanism works just fine, there are scenarioes in which external engagement needs to be involved into the decision process to fulfill special purpose.

3.1. LFA Precomputation and Adjustment

Loop-Free Alternates(LFA)[RFC5286] is deployed in pure IP and MPLS/ LDP networks to provide single-point-failure protection for unicast traffic. The goal of this technology is to reduce the packet loss that happens while routers converge after a topology change due to a failure. [I-D.ietf-rtgwg-lfa-manageability] provides operational feedback on LFA, highlights some limitations, and proposes a set of refinements to address those limitations. It also proposes required management specifications. In most of circumstances, operators will not be satisfied to know only the protection for links and prefixes. What they really hope is the overall protection for the whole network, especially for those high-value-added business. If lack of protection or protection coverage is not good enough, the operator may hope there are some ways to identify those weak points and the method to fix them up.

The I2RS MAY help to achieve the operator's hope by resorting to certain allowing applications to pre-computes the LFA backup of all links and prefixes in the network and calculating the protection coverage and recognizing optimization. Then an I2RS Client can deploy these new topology adjustments by sending the appropriate changes to the I2RS Agent that it will install in the routing place. The I2RS Agent can notify the I2RS Client (and the application) of the results of operation to provide a real-time feedback.

As showed below, traffic from Node-S to Node-D needs to pass Node-E. Under the circumstance of Link-SE's failure, the traffic can not be protected by Node-N since the metrics do not meet the demand of Inequality 1 from [RFC5286]. With the help from I2RS, the operator

can identify this weakness and may change the metric of Link-ND to gain LFA backup.

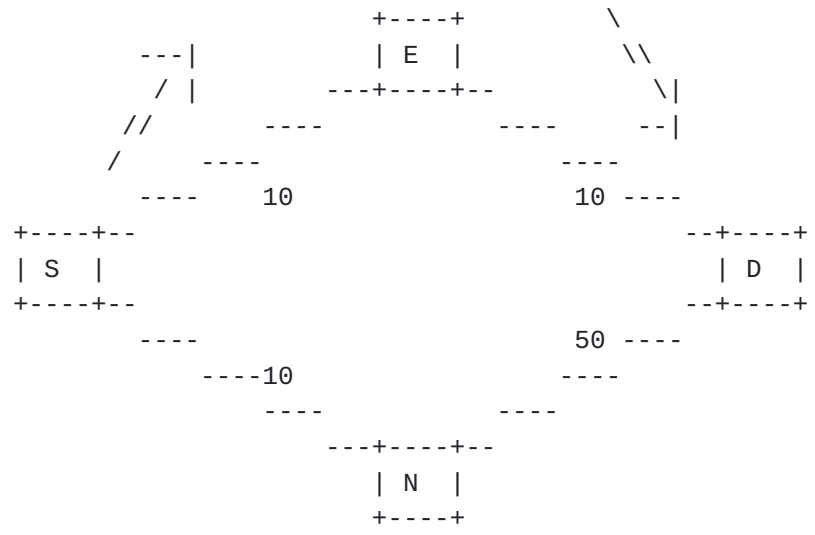

Figure 1: LFA precomputation and no backup available

3.2. Transient loop avoidance

Link-state protocols may need to reconverge when the network topology changes. During this phase packet loss and transient loops are frequently observed since inconsistent RIBs exist, even the reachability of the destinations is not compromised after the topology change. Drafts have been proposed to delay the reconvergence and update RIBs in a ordered manner among distributed routers. These methods may help to alleviate packet loss caused by transient loops while they are unable to help the packet loss during the delay.

Since the transient loops are introduced by inconsistent RIBs, the I2RS may help to avoid those loops by direct access to next hop information of route entries through pragmatic interfaces. By changing the next hops to be what they should be after reconvergence, transient loops can be avoided and no disturbance introduced when reconvergence happens.

3.3. Capacity Planning

It is increasingly common to see Equal-Cost-Multipath(ECMP) is used the networks of SP, Enterprise and DataCenter to make efficient use the network bandwidth. The traffic is spread across as many ECMP paths as possible allowing growth (or shrinkage) without a physical capacity adjustment

The I2RS programmatic interface SHOULD allow the balancing of both ECMP traffic flows and end-to-end traffic flows in the IGP. The I2RS SHOULD support monitoring of the dynamic traffic flow in the network, and the query of the maximum capacity of the network. After some offline optimization occurs, the I2RS can be used to spread ECMP paths through the topology or aggregate traffic onto a single path so the rest of the devices may power off saving power (and money. One important thing to note here, topology changes triggered by capacity adjustment MAY cause transient forwarding loops of which MUST be taken care. And the specific solution for this issue is out of the scope of this document.

As pictured below, traffic from Node-A to Node-B is widely spread among all links and nodes between them. This can increase the whole capacity of this network. When the traffic decreased, the operator can use I2RS to adjust the metric of Link-AB to less than the current one then the traffic will be summarized on the Link-AB. As a result of this change, Node-C, Node-D and their links can be power off or used for other purpose.

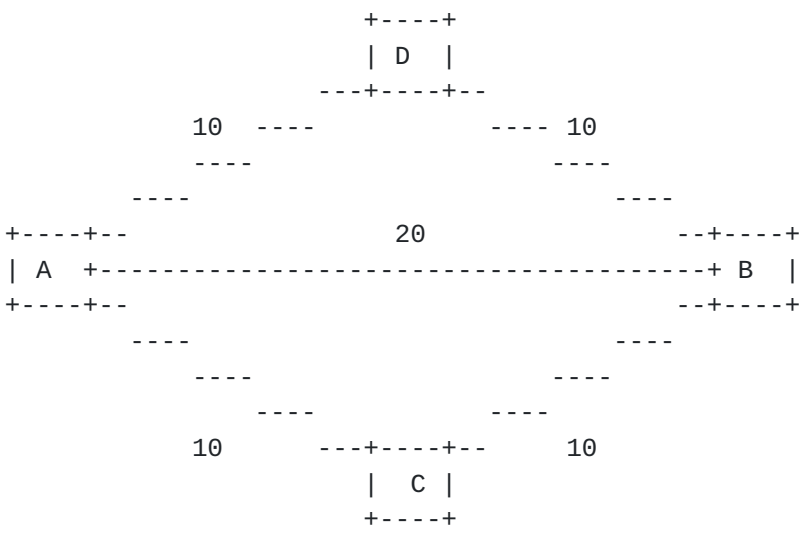

Figure 2: Capacity planning through topology adjustment

3.4. Traffic blackhole prevention

IGP hello packet is used to discover and maintain adjacencies among different devices. Without the deployment of fast detection techniques, one device has to wait for dozens of seconds before it realized the adjacency had broken. This kind of issue can cause one device is cut off from its network and lose connectivity completely. No matter planned or accidentally it may cause traffic blackhole before damage can be controlled.

Under the scenario of I2RS deployed, it is RECOMMENDED that the adjacency data of the other end side can be removed simultaneously or LSP can be updated directly by I2RS Agent when IGP is disabled or detached on one side. As a result traffic blackhole caused by silent broken adjacencies can be prevented.

4. IGP Events

As stated in [I-D.ietf-i2rs-architecture], it is practical for I2RS Clients to register for a range of notifications, and for the I2RS Agents to send notifications to a number of Clients. The I2RS Clients SHOULD use publish/subscribe mechanism to filter those events it is interested in. As regard for IGP protocol, these events MAY include topology changes, performance status and protocol statistics which are critical to operate and maintain IGP network with efficiency and scalability.

4.1. Topology Change Monitoring

Network topology information is the basis for futher operating and maintaining. It is very important and can be used in many scenarioes. Link-state protocol such as IGP is the recommended way to collect topology information.

Since many factors such as the status of interface, adjacency, node and etc can trigger the change of topology, the topology notification is reported to I2RS Clients at times. Considering lots of nodes and links in the network, these topology events can be massive. The I2RS SHOULD use the subscription mechanism to filter its interested events and use the publish mechanism to control the pace these events are notified. This precaution can protect the I2RS Client or even applications who depend on topology data from being drowned by massive duplicate events.

4.1.1. Router-ID conflict recovery

It is not rare to observe router-ID conflict in networks both intra and inter area, especially when different area merged. It is time-consuming and troublesome to detect and locate the place where this trouble happened. The frequently used solution is to rename one of the conflicted router-ID to a new one then reboot the involved IGP instance to force all adjacencies to rebuild and re-synchronize the Link-State-Database.

It MAY be possible to alleviate this issue with the help of programmatic I2RS interfaces. With the help of router information statistics, this conflict can be detected automatically. When one substantial conflict is on the horizon, no need to wait for mutual

re-origination happened, ID conflict can be found in collection of router information , no matter the conflict routers come from the same area or not. What is more, through I2RS interfaces and Agent, it is possible to rewrite one of the conflicted router-ID into a new one then reboot the routing- protocol instance.

4.2. Performance Monitoring

Since IGP protocol is essential to the whole network, the I2RS Clients SHOULD monitor about the protocol's running status before forwarding is impacted. Performance data can be collected through collecting static configuration and observing dynamic status. Dynamic data includes adjacency status, the number of entries in link-state database and in the routing table, the calculation status, the overload status, the graceful switch status and etc.

The I2RS Clients SHOULD subscribe to the I2RS Agent's notification of critical node events. For example, link-state database or routing table is under the status of overflow or the overflow status is released, the calculation continues for a long time, the system is under graceful reboot and etc.

4.2.1. Router number monitoring

Complaint can be heard frequently from clients about how many routers should be deployed in one area. The answer for this question is not very clear in vendor's guide since the product specification is only for reference and what's worse, those words like "usually", "roughly" or "most of the time" are often used from field engineers. As the consequence, it is always convenient for clients to deploy all the routers in one area, which may introduce scaling issue in future.

With the help of I2RS, it is possible to give out deployment suggestion or warning dynamically in the real-time manner. Based on the statistics of router number and system resource consuming, plus the ratio relationship between them, one notification or warning can be sent to I2RS Client. From there decision can be made to expand safely or have to shrink for precaution.

4.3. Protocol Statistics Monitoring

IGP protocol contains many useful statistics which can help to do trouble-shooting and maintain it. These statistics can be used by I2RS Clients to support diagnosing or analyzing tasks. For example, through subscribing packet dropped statistics, the I2RS Clients can figure it out the reason why some adjacencies do not succeed in connecting. Through subscribing the error statistics, the I2RS Clients can find out some link-state updating because of

authentication or checksum failure, which can further help to diagnose a configuration mistake or a subtle security attack happened.

5. IANA Considerations

This document includes no request to IANA.

6. Security Considerations

This document does not introduce any further security issues other than those discussed in [I-D.ietf-i2rs-architecture].

7. References

7.1. Normative References

[I-D.ietf-i2rs-architecture]
Atlas, A., Halpern, J., Hares, S., Ward, D., and T. Nadeau, "An Architecture for the Interface to the Routing System", draft-ietf-i2rs-architecture-05 (work in progress), July 2014.

[I-D.ietf-rtgwg-lfa-manageability]
Litkowski, S., Decraene, B., Filsfils, C., Raza, K., Horneffer, M., and p. psarkar@juniper.net, "Operational management of Loop Free Alternates", draft-ietf-rtgwg-lfa-manageability-04 (work in progress), August 2014.

[ISO.10589.1992]
International Organization for Standardization, "Intermediate system to intermediate system intra-domain-routing routine information exchange protocol for use in conjunction with the protocol for providing the connectionless-mode Network Service (ISO 8473)", ISO Standard 10589, 1992.

[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.

[RFC5286] Atlas, A. and A. Zinin, "Basic Specification for IP Fast Reroute: Loop-Free Alternates", RFC 5286, September 2008.

7.2. Informative References

[I-D.filsfils-spring-segment-routing-use-cases]
Filsfils, C., Francois, P., Previdi, S., Decraene, B., Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W., Tantsura, J., Kini, S., and E. Crabbe, "Segment Routing Use Cases", draft-filsfils-spring-segment-routing-use-cases-01 (work in progress), October 2014.

[I-D.ietf-isis-oper-enhance]
Shen, N., Li, T., Amante, S., and M. Abrahamsson, "IS-IS Operational Enhancements for Network Maintenance Events", draft-ietf-isis-oper-enhance-03 (work in progress), February 2013.

[I-D.li-ospf-ext-green-te]
Yan, G., Yang, J., and Z. Li, "OSPF Extensions for MPLS Green Traffic Engineering", draft-li-ospf-ext-green-te-01 (work in progress), October 2013.

[I-D.wijnands-bier-architecture]
Wijnands, I., Rosen, E., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast using Bit Index Explicit Replication", draft-wijnands-bier-architecture-01 (work in progress), October 2014.

Authors' Addresses

Nan Wu
Huawei
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: eric.wu@huawei.com

Zhenbin Li
Huawei
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: lizhenbin@huawei.com

Susan Hares
Hickory Hill Consulting
7453 Hickory Hill
Saline, CA 48176
USA

Email: shares@ndzh.com