Network Working Group
Internet-Draft
Intended status: Informational
Expires: April 24, 2014

Z. Li
H. Chen
G. Yan
Huawei Technologies

October 21, 2013

An Architecture of Central Controlled Interior Gateway Protocol (IGP)
draft-li-rtgwg-cc-igp-arch-00

Abstract

As the Software Defined Networks (SDN) solution develops, IGP will be extended to support central control. This document introduces an architecture of using IGP for central controlling. Some use cases under this new framework are also discussed. For specific use cases, making necessary extensions in IGP are required.

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

Copyright Notice

Table of Contents

1. Introduction

Interior Gateway Protocol (IGP) is a protocol for exchanging routing information between gateways (hosts with routers) within an autonomous network (for example, a system of corporate local area networks). The routing information can then be used by the Internet Protocol (IP) or other network protocols to specify how to route transmissions.

The internet is the most popular network, it is a distributed system. Depending on its configuration, each network device communicates with its neighbor, generates the FIB, and forwards the packet hop by hop. As the rise of SDN, central controlled IGP is becoming more important and new requirements for IGP are proposed as follows:

1. Build the central control architecture between the controller and the client. It includes building connectivity, collecting the topology, and dividing multiple areas automatically, etc.

2. Many new applications are emerging under the central controlled framework, such as network virtualization, centralized MPLS TE calculation, segment routing, etc. These new applications bring extension requirement to IGP.

This document defines an IGP-Based Central Control architecture and then use cases and corresponding IGP extensions under this architecture are described.

2. Terminology

BGP: Border Gateway Protocol

IGP: Interior Gateway Protocol

IS-IS: Intermediate System-Intermediate System

OSPF: Open Shortest Path First

SDN: Software Defined Network

3. Architecture

3.1. Reference Model

The following figure depicts a typical architecture of central controlled IGP. It consists of two essential network elements: IGP Controller and IGP Client. IGP Controller controls all the IGP Clients within its administrative domain by communicating with them. And the controller will also exchange the information each other through some protocol extensions which is out of scope of this document.

```
   +------------------------------+    +------------------------------+
   |   Central Control Domain 1   |    |   Central Control Domain 2   |
   |         +----------+         |    |         +----------+         |
   |         |          |         |    |         |          |         |
   |         |   IGP    |------------------------|   IGP    |         |
   |         |Controller|         |    |         |Controller|         |
   |         |          |         |    |         |          |         |
   |         +----------+         |    |         +----------+         |
   |         /          \         |    |         /          \         |
   |        /            \        |    |        /            \        |
   |       /              \       |    |       /              \       |
   | +--------+        +--------+ |    | +--------+        +--------+ |
   | | NODE 1 |        | NODE n | |    | | NODE 1 |        | NODE n | |
   | |        | ...... |        | |    | |        | ...... |        | |
   | |  IGP   |        |  IGP   | |    | |  IGP   |        |  IGP   | |
   | | CLIENT |        | CLIENT | |    | | CLIENT |        | CLIENT | |
   | +--------+        +--------+ |    | +--------+        +--------+ |
   |                              |    |                              |
   +------------------------------+    +------------------------------+
```

Figure 1: An Architecture of Central Controlled IGP

3.2. Deployment Mode

IGP Controller can run on a general-purpose server or a network device. If IGP Controller runs on a network device, it supports both central-controlled functionality and forwarding functionality. In this scenario, besides the control central point, the IGP controller can also work as a forwarding central point to receive traffic from one node and forward to other nodes. The forwarding model in this scenario is just like hub-spoke forwarding model. If IGP Controller runs on a server, it will not involve in the actual forwarding. It only works as the control central point to control the forwarding behaviors of the nodes. In this scenario the traffic will be distributed in the controlled nodes.

More than one controller can be deployed in a central control domain. These controllers can work on master-slave mode or load-sharing mode.

3.3. Requirement of IGP Extensions

Building a IGP-based Central Controlled Framework needs extensions to IGP, I2RS etc.

3.3.1. Building Connectivity

IGP protocol is very important to establish connective in the central control domain. When a new device connects to the this domain, the connectivity with the other node and the controller should be built at first. The procedures should be automated since the number of devices in this domain can be huge. Base on this initialization process, the controller can download the necessary configuration to this new node to drive it to set up adjacency with its neighbors and the controller. Then the topology information can be synchronized in the central control domain and the connectivity can be built.

3.3.2. Roles Auto-Discovery

In the central control domain, there are two basic roles: IGP controller and IGP client. The controller can centrally configure the client role through I2RS interface. The role information should be flooded through IGP extensions to support the auto discovery functionality.

3.3.3. Choosing Controller

After the roles of the elements are discovered, if there are multiple controllers in the domain, the client can determine which controller to join by its own, or the controllers can determine which controller the clients should join and set the configuration on the nodes through I2RS interface. When determine the controller to be joined, the work mode (master-slave, load-sharing, etc.) of multiple controllers, service type and some other constraints needs to be taken into account.

3.3.4. High Availability

In the IGP-based Central Controlled framework, IGP Controller plays a key role. To avoid one-point-failure of IGP Controller, it is possible to run redundant IGP Controllers for high availability.

Information should be synchronized between the controllers through necessary mechanisms or protocol extensions other than IGP. When the Primary IGP Controller failed, the Backup IGP Controllers will take over the work of the Primary IGP Controller.

To ensure IGP route persistence in case of occurrence of IGP Controller failure, the new Primary IGP Controller SHOULD perform resynchronization with IGP Clients.

When IGP Client loses connection with Primary IGP Controller, it SHOULD following IGP Graceful Restart routine.

3.3.5. Security

In IGP-based Central Controlled framework, it SHOULD be ensured that communications between IGP Controllers and IGP Clients conform to network security policy. The communication key used on IGP Client can be configured through I2RS or other way.

4. Usecases

In IGP-based Central Controlled framework, new use cases which are difficult to be supported in traditional networks are emerging. In some specific use cases, extension and enhancement of IGP protocol are necessary.

4.1. Network Topology Acquirement

In traditional network, it is very difficult for the application to get and use the topology. The application has to depend multiple protocols such as OSPF, ISIS, LLDP, etc. In some scenarios, the application has to communicate with these protocols directly. In the IGP-based central controlled framework, the topology acquirement procedures SHOULD be simplified. All topology related information SHOULD be able to be collected by IGP. Thus the complexity of network operation and management can be reduced. In the IGP-base central controlled framework, the controller can get the whole topology information of the central control domain which can be easily provided for applications through public interface.

4.2. Automated Dividing Multiple Domains

When there are mass devices in the network, not only LSDB synchronization, but also route convergence, will be big pressure for any device, so the network has to be divided into multiple domains. In the IGP-based central controlled the framework, the division can be done automatically by the controller which can calculate reasonable scale for IGP domains based on the whole network information and the possible constraints. IGP adjacency is only set up between nodes in the same IGP domain. The adjacency SHOULD not set up between nodes in different IGP domains. Thus the pressure on the nodes for LSDB synchronization can be reduced and route convergence performance can be improved. The configuration about domain division can be set through I2RS interfaces from the controller to the clients. The architecute for dividing multiple domains with the central controller is shown in the figure 2.

```
       +----------------------------------------------------------+
       |                                                          |
       |                      IGP Controller                      |
       |                                                          |
       +----------------------------------------------------------+
         |                 |                 |                 |
         |                 |                 |                 |
   +-----|-----------------|------+    +-----|-----------------|------+
   |     |   IGP Domain 1  |      |    |     |   IGP Domain 2  |      |
   |     |                 |      |    |     |                 |      |
   |     |                 |      |    |     |                 |      |
   | +--------+        +--------+ |    | +--------+        +--------+ |
   | | NODE 1 |        | NODE n | |    | | NODE 1 |        | NODE n | |
   | |        | ...... |        | |    | |        | ...... |        | |
   | |  IGP   |        |  IGP   | |    | |  IGP   |        |  IGP   | |
   | | CLIENT |        | CLIENT | |    | | CLIENT |        | CLIENT | |
   | +--------+        +--------+ |    | +--------+        +--------+ |
   |                              |    |                              |
   +------------------------------+    +------------------------------+
```

Figure 2: Automatic Division of Multiple IGP Domains

4.3. Centralized MPLS TE

In the IGP-based Central Controlled framework, the controller can implement better traffic engineering functionality because it can calculate more reasonable path based on complete topology information and state information of the whole network. Centralized MPLS TE calculation can avoid the flaw of non-best path proposed by the existing distributed MPLS TE calculation.

In order to support centralized MPLS TE path calculation, IGP SHOULD be able to collect more information from the network. There are two types of information for IGP to collect:

1. Static configuration: In traditional network, MPLS TE attributes should be configured on the link such as maximum reservable bandwidth, color, TE metric, etc. These information will be flooded in the work for MPLS TE path calculation. In the IGP-based Central Controlled framework, these configuration can be set by the controller. This means it is not necessary for the controller to get the TE link information through IGP flooding process. For the reason of compatibility, the IGP flooding process of MPLS TE link information can be kept in the central controlled framework. On the other hand, it provides a possible way for the inconsistency check on the configuration.

2. Running information: Some dynamic state information such as real traffic bandwidth, packet loss rate, delay, power consumption, etc. can be flooded through IGP extensions from the nodes to the controller. The running information can help the controller to calculate more reasonable path and calculate path for more constraints defined by applications.

4.4. MPLS Global Label Allocation

MPLS Global Label should be allocated centrally to guarantee all distributed network nodes can understand meaning of a specific global label in same. The IGP-based Central Controlled framework is particularly suitable to allocate MPLS Global Label through some necessary IGP extensions rather than traditional MPLS protocols(e.g. LDP, RSVP-TE etc.).

MPLS Global Label is defined in [I-D.li-mpls-global-label-framework] and related use cases are defined in [I-D.li-mpls-global-label-usecases].

The MPLS global label should be assigned centrally; each node in network should have same understanding about these labels. In the central control network, the global label will be handled by controller, and IGP protocol will flood these labels.

The extensions of IGP for MPLS global label include:

1. Collect the label capability of each node. The label capability is the global label space.

2. IGP Controller determines the COMMON label space for all its IGP Clients.

3. The controller will assign the global label for different services, and these label bindings will be flooded through IGP protocol to IGP clients.

4. IGP Client receives the MPLS Global Labels, and generates corresponding MPLS forwarding entries.

IGP is suitable for the use cases of MPLS global label in the intra-domain scenario. These use cases include MPLS virtual network and segment routing as defined in [I-D.li-mpls-global-label-usecases].

4.5. Virtual Link

When the IGP-based Central Controlled framework is applied, one possible scenario is partial deployment. That is, part of the existing network will be converted to be controlled in the central control mode. The application scenario is shown in the following figure:

```
                  +-------------------------------+
                  |    Central Control Domain     |
                  |         +----------+          |
                  |         |          |          |
                  |         |   IGP    |          |
                  |         |Controller|          |
                  |         |          |          |
                  |         +----------+          |
                  |         /          \          |
                  |        /            \         |
                  |       /              \        |
   +-----------+  | +--------+        +--------+  | +-----------+
   |Traditional|  | | NODE 1 |        | NODE n |  | |Traditional|
   |           |  | |        | ...... |        |  | |           |
   |   NODE    |----|  IGP   |        |  IGP   |----|   NODE    |
   |           |  | | CLIENT |        | CLIENT |  | |           |
   +-----------+  | +--------+        +--------+  | +-----------+
                  |                               |
                  +-------------------------------+
```

Figure 3: Partial Deployment of Central Controlled IGP

In this scenario, it is not necessary for the traditional nodes to learn the detailed topology information of the central control domain. The information flooded between the central control domain and the traditional nodes can be reduced. The central control domain can only advertise virtual links which connect the edge nodes in the domain that the traditional node can be aware of. The process can reduce the pressure of the traditional node for flooding and improve convergence performance.

In the central control domain, the controller can apply the policy defined by the applications to control whether the virtual link will be advertised to the outside and what metric is advertised to affect the route calculation of the outside network.

5. IANA Considerations

TBD.

6. Security Considerations

TBD.

7. References

7.1. Normative References

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

7.2. Informative References

[I-D.chen-ospf-ttz]
Chen, H., Li, R., Cauchie, G., Retana, A., Ning, S., Toy, M., and L. Liu, "OSPF Topology-Transparent Zone", draft-chen-ospf-ttz-06 (work in progress), July 2013.

[I-D.ietf-ospf-te-metric-extensions]
Giacalone, S., Ward, D., Drake, J., Atlas, A., and S. Previdi, "OSPF Traffic Engineering (TE) Metric Extensions", draft-ietf-ospf-te-metric-extensions-04 (work in progress), June 2013.

[I-D.li-mpls-global-label-framework]
Li, Z., Zhao, Q., and T. Yang, "A Framework of MPLS Global Label", draft-li-mpls-global-label-framework-00 (work in progress), July 2013.

[I-D.li-mpls-global-label-usecases]
Li, Z., Zhao, Q., and T. Yang, "Usecases of MPLS Global Label", draft-li-mpls-global-label-usecases-00 (work in progress), July 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.ylz-ospf-lsdb-sync-group]
Yan, G., Liu, Y., and X. Zhang, "OSPF Extensions for Link State Database Synchronization Group", draft-ylz-ospf-lsdb-sync-group-01 (work in progress), October 2013.

Authors' Addresses

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

Email: lizhenbin@huawei.com

Huaimo Chen
Huawei Technologies
Boston, MA
USA

Email: huaimo.chen@huawei.com

Gang Yan
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China

Email: yangang@huawei.com