

INTERNET-DRAFT
Intended Status: Standard Track
Expires: September 4, 2018

Y. Zhang
M. Sun
Huawei Technologies
F. Gao
Baidu Inc
March 5, 2018

A method to assign IP-address automatically in data center construction
[draft-zhang-rift-network-ip-assignment-00](#)

Abstract

IP assignment is always a big challenge in the construction of an enterprise-scale data center. For two interfaces connected by a cable, if their assigned IP addresses do not belong to a same subnet, data then cannot be transmitted via this cable. An error in a cable connection could make the IP addresses of the cable's two ends not in a same subnet, resulting in a reduction in network throughput. Furthermore, it is difficult to find and repair this error connection among numerous cables. A method to address this error in switch connection would save a great amount of time and cost in modern data center construction. Here, this draft introduces an approach to assign the IP addresses automatically for two connected switches, which allows data to be transmitted via a cable that is connecting two arbitrary interfaces from two switches.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/lid-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>

Copyright and License Notice

Copyright (c) 2018 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1	Introduction	3
1.1	Background	3
1.1	A typical method to assign IPs to interfaces	3
2	IP assigned automatically during switch connection	4
2.1	One switch with one logic node	4
2.2	One switch with more than one logic node	6
3	Security Considerations	8
4	References	8
4.1	Normative References	8
4.2	Informative References	8
	Authors' Addresses	8

1 Introduction

1.1 Background

'Three-layer network' becomes more popular as the scale of a modern data center increases continuously. Data can be transmitted in a 'three-layer network' by using the Border Gateway Protocol (BGP) [RFC4271] as the routing protocol [RFC 7938]. For two interfaces connected by a cable, if their assigned IP addresses belongs to a same subnet, data then can be transmitted via this cable. Otherwise, the destination address cannot be resolved via address resolution protocol (ARP) [RFC 2119]. This leads to an interruption of BGP, which contains the message of allowed data type for switches. As there is a strong correlation between switch interfaces and IP addresses, a wrong wire connection between the interfaces will result in a reduction in the network throughput and even un-reachable for some parts of the network.

1.1 A typical method to assign IPs to interfaces

To avoid a wrong wire connection between two interfaces, a typical method is labelling the cable in its two ends during cable deployment. As shown in Figure 1, the cable's two ends are attached with two labels containing the identifications of local switch and interface, and the identifications of peer switch and interface, respectively. After the set-up of wire connection, IP addresses belong to a same subnet are then assigned to the two connected interfaces.

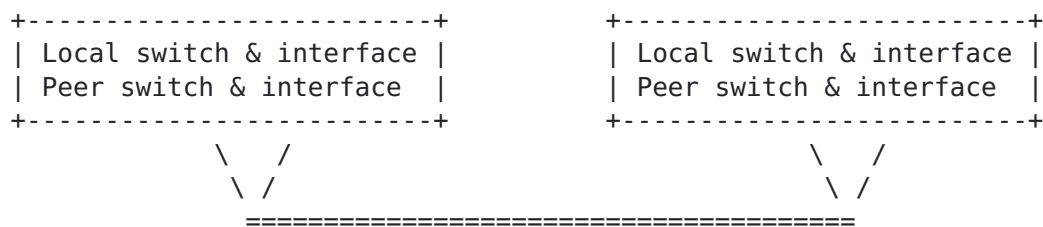


Figure 1: Attached labels in a cable's two ends.

This method could effectively reduce the occurrence rate of error connection. However, the steps of cable labelling manually and insertion into a certain interface, will cost a huge amount of time. Furthermore, if there comes an error connection, it is also very difficult to find the position of the error connection.

This draft presents an approach to assign IP addresses to switch interfaces automatically when the interfaces are connected. With this approach, message can be transmitted when a cable is connected to the

right switch, regardless of the interface it is connected.

2 IP assigned automatically during switch connection

An innovation of this approach is that IP address is saved on a logic node instead of being assigned directly to a certain interface. The logic node describes the information of logic IDs of local switch, and logic IDs of peer switches, and corresponding IP addresses for local interfaces. It is not correlated to a specific interface. Any interface of peer switch can be connected to an arbitrary local interface. Logic ID could be sent from peer interface to local interface by a link layer protocol such as the link-layer discover protocol (LLDP) [[LLDP](#)]. If the logic ID can be looked up in logic nodes, IP address would be assigned to the local interface.

2.1 One switch with one logic node

Figure 2 illustrates an example of setting up two switches in a three-layer network with presented approach. The set-up procedure can be divided into four steps as described bellow.

- 1) The two switches are named as LogicID11 and LogicID21, respectively. It is worth noting that one switch could be named with more than one logic ID (In this case, each switch is named with one logic ID). According to the plan of switch connection and IP assignment, IP addresses are assigned to some logic nodes in switches. The logic node contains the information of logic ID of local switch, logic ID of peer switch, and IP address for local interface.
- 2) Switches are connected by one cable with little consideration of which interface it is used.
- 3) Using a link-layer protocol, the logic ID of local switch is sent from the local interface to the peer interface when the two switches are wire-connected.
- 4) After receiving the logic ID from peer interface, it will be looked up in the logic nodes. The matched logic node will assign the corresponding IP address to the local interface.

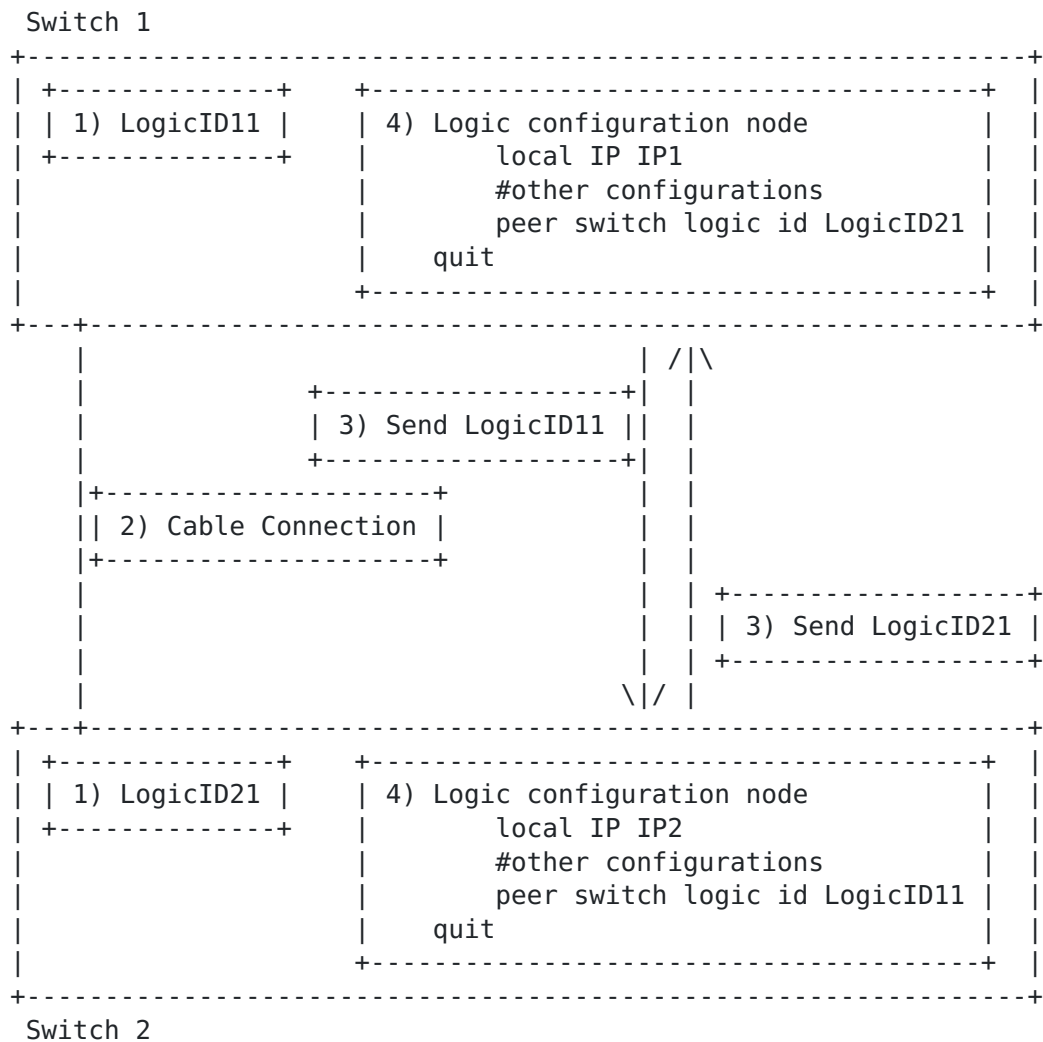


Figure 2: Schematic diagram of the approach for auto-assignment of IP addresses for wired switches.

This approach utilizes a link layer protocol to send the logic ID of local switch to peer switch. For an example of using LLDP, the LLDP frame not only contains the mandatory type-length-value (TLV) structures of Chassis ID, Port ID, Time-to-live and End of LLDPDU, but also contains a TLV structure of Logic ID as follows [[TLV-type](#)]:



2.2 One switch with more than one logic node

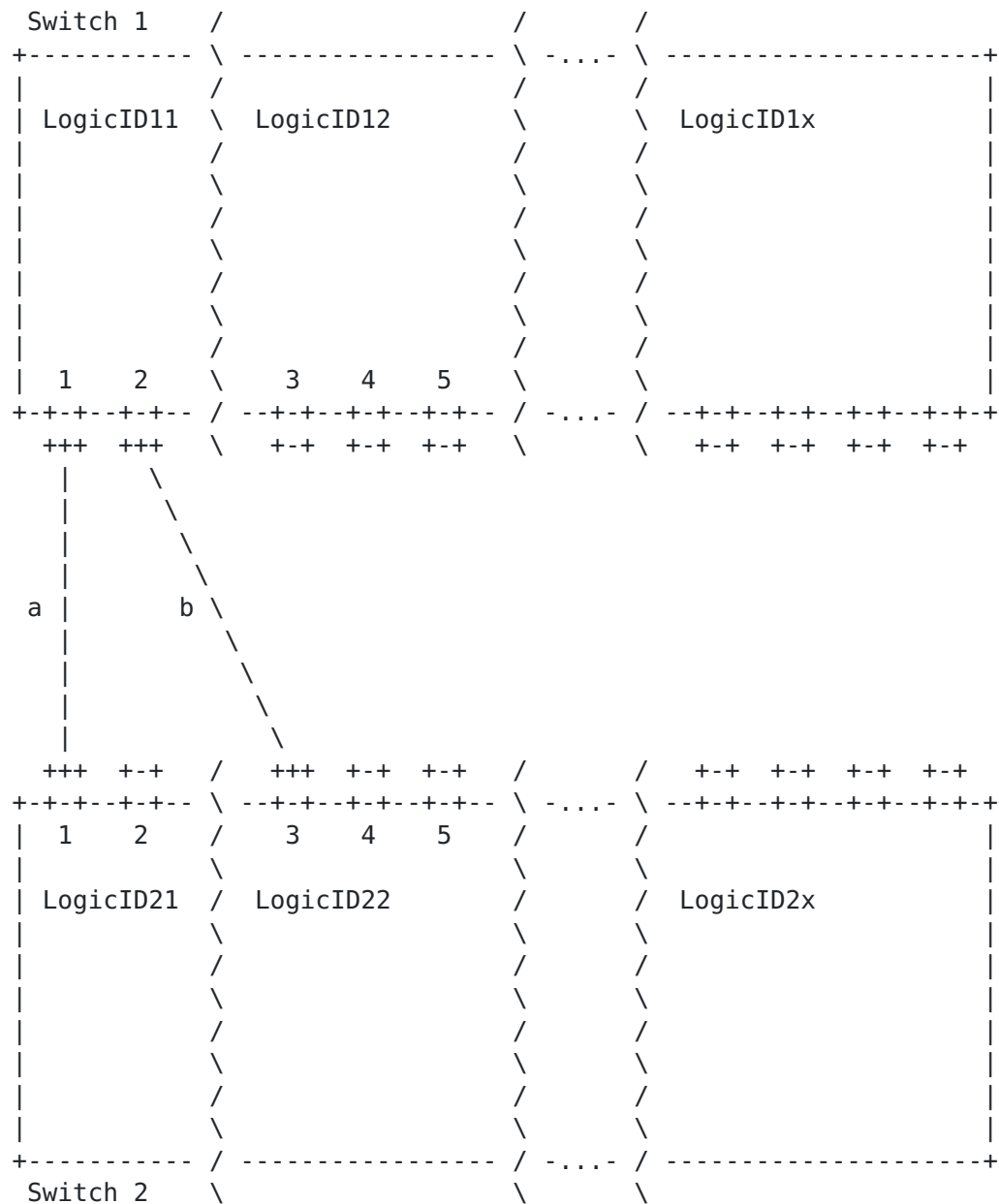


Figure 3: In an case of one switch is named with more than one logic ID.

Figure 3 presents the case in which one switch are named with more than one logic ID. For example, in Switch 1, Interface 1 and 2 belong to LogicID11, while Interface 3, 4 and 5 belong to LogicID12. Two cables marked as 'a' and 'b' connect the Interface 1 of Switch 1 to Interface 1 of Switch 2, and Interface 2 of Switch 1 to Interface 3 of Switch 2, respectively. After connection, logic ID of local switch

will be sent to peer switch. Then, the logic ID of peer switch, together with the logic ID of local switch, will be looked up in logic nodes. As shown in Table 1, IP address of 182.111.211.111 255.255.255.0 will be assigned to Interface 1 of Switch 1, while IP address of 182.111.222.111. 255.255.255.0 will be assigned to Interface 2 of Switch 1. An advantage of this approach is that correct IP address still can be assigned to local interface when an error connection between interfaces occurs. For example, when Interface 2, instead of Interface 1 of Switch 2 is connected to Interface 1 of Switch 1, IP address of 182.111.211.111 255.255.255.0 will also be assigned to Interface 1 of Switch 1. In that case, data still can be transmitted by cable 'a' without any problem in ARP. Furthermore, when Switch 2 is replaced by another one Switch 3 in the future, IP address will also be assigned to local and peer interfaces automatically if there are 'Switch 3'-related logic IDs in nodes.

Table 1: List of Logic nodes For Switch 1 in Figure 3.

Logic id of local switch	Logic id of peer switch	Interface configuration
11	21	undo portswitch ip address 182.111.221.111 255.255.255.0
11	22	undo portswitch ip address 182.111.222.111 255.255.255.0
12	21	undo portswitch ip address 182.111.221.112 255.255.255.0
12	22	undo portswitch ip address 182.111.222.112 255.255.255.0
.....		

Compared with conventional IP assignment method, this method assigns the IP addresses automatically for wired switches, which can significantly decrease the workload of cable deployment. Cable will not be required to connect to a specific interface, leading to a reduction in both working time and rate of error connection. When cable is changed from one interface to another one belonging to a same switch logic ID, network will work as normal without any necessary in changing setting.

3 Security Considerations

The design does not introduce any additional security concerns. General BGP security considerations are discussed in [[RFC4271](#)] and [[RFC4272](#)].

4 References

4.1 Normative References

- [RFC4271] Y. Rekhter, T. Li, and S. Hares, "A Border Gateway Protocol 4 (BGP-4)", [RFC 4271](#), January 2006.
- [LLDP] IEEE, "IEEE Standard for Local and metropolitan area networks station and Media Access Control Connectivity Discovery Corrigendum 2: Technical and Editorial Corrections", IEEE 802.1AB-2009/Cor 2-2015, DOI 10.1109/ieeestd.2015.7056401, March 2015.

4.2 Informative References

- [RFC7938] P. Lapukhov, A. Premji, and J. Mitchell, "BGP Routing in Data Centers", [RFC 7938](#), August 2016.
- [RFC2119] S. Bradner, "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [TLV-type] IEEE Std 802.1AB-2016, DOI: 10.1109/IEEESTD.2016.7433915, March 2016, <<http://ieeexplore.ieee.org/servlet/opac?punumber=7433913>>
- [RFC4272] S. Murphy, "BGP Security Vulnerabilities Analysis", [RFC 4272](#), January 2006.

Authors' Addresses

Yun Zhang
Huawei
101 Software Avenue,
Yuhuatai District, Nanjing, 210012
China

EMail: zhangyun45@huawei.com

Marcus Sun
Huawei
101 Software Avenue,
Yuhuatai District, Nanjing, 210012
China

EMail: marcus.sun@huawei.com

Feng Gao
BAIDU Inc.
10 shangdi shijie Haidian, Beijing
China

Email:gaofeng04@baidu.com