Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: July 5, 2009

R. Gagliano
LACNIC

January 2009

IPv6 Deployment in Internet Exchange Points (IXPs)
draft-rgaglian-v6ops-v6inixp-01.txt

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/ietf/1id-abstracts.txt.

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

This Internet-Draft will expire on July 5, 2009.

Copyright Notice

Abstract

This document provides a guide for IPv6 deployment in Internet Exchange Points (IXP). It includes information about the switching

fabric configuration, the addressing plan options and general organizational tasks to be performed. IXP are mainly a layer 2 device (the switching fabric) and in many case the best recommendations state that IPv6 traffic and management should not be handled differently than in IPv4.

Table of Contents

1. Introduction 3
1.1. Requirements Language 3
2. Switch Fabric Configuration 3
3. Addressing Plan 4
4. Reverse DNS 6
5. Route Server Configuration 6
6. Internal and External Services support 6
7. IXP Policies and IPv6 7
8. Multicast IPv6 7
9. IANA Considerations 7
10. Security Considerations 7
11. Acknowledgements 7
12. References 7
12.1. Normative References 7
12.2. Informative References 8
Author's Address 8

1. Introduction

Most Internet Exchange Points (IXP) work on the Layer 2 level, making the adoption of IPv6 an easy task. However, IXPs normally implement additional services such as statistics, route servers, looking glasses, broadcast control and others that may be impacted by the implementation of IPv6. This document gives guidance on the impact of IPv6 on a new or an existing IXP that may or may not fit any particular deployment. The document assumes an Ethernet switch fabric, although other layer 2 canfigurations can be deployed.

1.1. 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].

2. Switch Fabric Configuration

An Ethernet based IXP switching fabric implements IPv6 over Ethernet as described in RFC 2464 [RFC2464], therefore the switching of IPv6 traffic happens in the same way as in IPv4. However, some management functions require explicit IPv6 support. Such functionalities may include: switch management, SNMP support and flow analysis tools.

There are two common configurations of IXP switch ports to support IPv6:

1. dual stack VLAN: both IPv4 and IPv6 traffic share a common VLAN. No extra configuration is required in the switch. Typically participants will configure dual stack interfaces in this scenario but independent port can be an option.

2. independent VLAN: an IPv6 VLAN is created for IPv6 traffic. If IXP participants are already using VLAN tagging on the interfaces of their routers which face the IXP switch, this only requires configuring one additional VLAN tag across the interconnection. If participants are using untagged interconnections with the IXP switch and wish to continue doing so, they will need to use a separate physical port to access the IPv6-specific VLAN.

The "independent VLAN" configuration provides a physical separation for IPv4 and IPv6 traffic simplifying the separated analysis for IPv4 and IPv6 traffic.However, it can be more costly in both capital expenses (if new ports are needed) and operational expends.

The "dual stack" configuration allows a quick and cost-free start-up for IPv6 support in the IXP. It also avoids transforming untagged ports into tagged ports. Traffic split for statistical analysis may be done using flows techniques such as in IPFIX [RFC5101] considering the different ether-types (0x0800 for IPv4 and 0x86DD for IPv6).

The support for jumbo frames MTU should be evaluated. The only technical requirement for IPv6 referring link MTUs is that it needs to be greater than or equal to 1280 octets [RFC2460]. Most IXPs support MTUs of 1500, 4470, or 9216 bytes, so this typically requires no change of configuration.

3. Addressing Plan

Regional Internet Registries (RIRs) have specific address policies to allocate Provider Independent (PI) IPv6 address to IXPs. Those allocations are usually /48 prefixes [RIR_IXP_POLICIES]. Depending on the country and region of operation, address allocations may be provided by NIRs (National Internet Registries).

From the allocated prefix, following the recommendations of RFC 4291 [RFC4291], a /64 prefix should be allocated for each of the exchange point Local Area Networks (LANs). A /48 prefix allows the addressing of 65536 LANs. Longer prefixes (/65-/127), are technically feasible using static address configuration, but should be avoided, in order to keep EUI-64 compatibility.

The common practice for Interface Identifiers (IID) configuration is to use static configuration, disallowing auto-configuration on every interface. Also, on a LAN where all its participants are typically routers, it is important that every node has it's router advertisement messages RFC 4861 [RFC4861] turned off. The goal is that none of the remaining routers configure it-selves a default ICMPv6 route by accident. A scanning device can be set up at the IXP LANs to monitor link-local multicast traffic (addresses ff02::/16), allowing only ICMPv6 Neighbor Solicitation, Neighbor Advertisement messages and MLD (Multicast Listener Discovery) if multicast peering is permited in any particular VLAN.

When selecting the use of static IIDs, there are different options on how to "intelligently" fill its 64 bits (or 16 hexadecimal characters) in order to help both IXPs and paricipants network operations. A non exhausted list of possible IID selection mechanisms follows:

1. Some IXPs like to include the participants' ASN number decimal encoding inside each IPv6 address. The ASN decimal number number is used as the BCD (binary code decimal) encoding of the upper part of the IID such as shown in this example:

   * IXP LAN prefix: 2001:DB8::/64

   * ASN: 64496

   * IPv6 Address: 2001:DB8::6449:6000:0000:0001/64 or its equivalent representation 2001:DB8::6449:6000:0:1/64

   In this representation each ASN may require a maximum of 10 characters, as 16 characters are available, up to 2^24 IPv6 addresses can be configured per ASN.

2. Although BCD encoding is more "human-readable", some IXPs prefer to use the hexadecimal encoding of the ASNs number as the upper part of the IID as follow:

   * IXP LAN prefix: 2001:DB8::/64

   * ASN: 64496 (DEC) or FBF0 (HEX)

   * IPv6 Address: 2001:DB8::0000:FBF0:0000:0001/64 or its equivalent representation 2001:DB8::FBF0:0:1/64

   In this representation each ASN may require a maximum of 8 characters, as 16 characters are available, up to 2^32 IPv6 addresses can be configured per ASN.

3. A third scheme for statically assigning IPv6 addresses on an IXP LAN could be to relate some portion of a participant's IPv6 address to its correspondant IPv4 address. In the following example, the last three decimals of the IPv4 address are copied to the last hexadecimals of the IPv6 address, using the decimal number as the BCD encoding for the last three characters of the IID such as in the following example:

   * IXP LAN prefix: 2001:DB8::/64

   * IPv4 Address: 240.0.20.132/23

   * IPv6 Address: 2001:DB8::132/64

4. A fourth configutation might be based on the IXPs ID for that participant.

The current practice that applies to IPv4 about publishing IXP allocations to the DFZ (Default Free Zone) should also apply to the IPv6 allocation (normally a /48 prefix). IXP external services (such as dns, web pages, ftp servers) could be part of this prefix.

4. Reverse DNS

PTR records for all addresses assigned to participants should be included in the IXP reverse zone under "ip6.arpa".

5. Route Server Configuration

Some IXPs may offer a Route Server service, either for Multi-Lateral (ML) Peering Agreements or for a looking glass service. IPv6 support needs to be added to the router used as BGP end point. The equipment should be able to transport IPv6 traffic and to support Multi-protocol BGP (MP-BGP) extensions for IPv6 address family (RFC 2545 [RFC2545] and RFC 4760 [RFC4760]).

A good practice is to have IPv6 SAFI (Subsequent Address Family Identifiers) information carried over sessions established also on top of the IPv6 IP/TCP stack and independently of the IPv4 sessions. This configuration allows that in the event of IPv6 reachability issues to any IPv6 peer, the specific session will be turned down and the IPv4 session to the same peer will not be affected. Please consider the use of MD5 (even better IPSEC) to authenticate the BGP sessions.

The Router-Server or Looking Glass external service should be available for external IPv6 access, either by an IPv6 enabled web page or an IPv6 enabled console server.

6. Internal and External Services support

Some external services that need to have IPv6 support are Traffic Graphics, DNS, FTP, Web and Looking Glass. Other external services such as NTP servers, or SIP Gateways need to be evaluated as well. In general, each service that is accessed through IPv4 or that handle IPv4 addresses should be compatible with IPv6.

Internal services are also important when considering IPv6 adoption at an IXP. Such services may not deal with IPv6 traffic but may handle IPv6 addresses; that is the case of provisioning systems, logging tools and statistics analysis tools. Databases and tools needs to be evaluated to determinate its IPv6 support level.

7. IXP Policies and IPv6

IXP Policies may need to be revised as any mention of IP should be clarified if it refers to IPv4, IPv6 or both. The current interpretation is that IP refers to the Internet Protocol, independently of the its version (i.e. both IPv4 and IPv6). In any case contracts and policies should be reviewed for any occurrence of IP and/or IPv4 and replace it with the appropriate IP, IPv4 and/or IPv6 language.

8. Multicast IPv6

Multicast IPv6 is not different from an IXP perspective than Multicast IPv4. The IXP may decide to use a reserved VLAN for Multicast traffic or to exchange that traffic in the same VLAN as the unicast traffic. Link-local multicast traffic should be monitored as this traffic should be reduced to ICMPv6 Neighbor Discovery RFC 4861 [RFC4861] and MLD (Multicast Listener Discovery) Protocol (MLDv2) RFC 3810 [RFC3810].

9. IANA Considerations

This memo includes no request to IANA.

10. Security Considerations

This memo includes no Security Considerations.

11. Acknowledgements

The author would like to thank the contributions from Bill Woodcock (PCH), Martin Levy (Hurricane Electric), Carlos FriaAas of FCCN (GIGAPIX), Arien Vijn (AMS-IX) and Louis Lee (Equinix).

12. References

12.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC2464] Crawford, M., "Transmission of IPv6 Packets over Ethernet Networks", RFC 2464, December 1998.

[RFC2545] Marques, P. and F. Dupont, "Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing", RFC 2545, March 1999.

[RFC3810] Vida, R. and L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, January 2007.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, September 2007.

[RFC5101] Claise, B., "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information", RFC 5101, January 2008.

12.2. Informative References

[RIR_IXP_POLICIES]
Numbers Support Organization (NRO)., "RIRs Allocations Policies for IXP. NRO Comparison matrix", 2008, <http://www.nro.net/documents/comp-pol.html>.

Author's Address

Roque Gagliano
LACNIC
Rambla Rep Mexico 6125
Montevideo, 11400
UY

Phone: +598 2 4005633
Email: roque@lacnic.net