Internet Engineering Task Force (IETF)
Request for Comments: 6908
Category: Informational
ISSN: 2070-1721

Y. Lee
Comcast
R. Maglione
Cisco Systems
C. Williams
MCSR Labs
C. Jacquenet
M. Boucadair
France Telecom

March 2013

Deployment Considerations for Dual-Stack Lite

Abstract

This document discusses the deployment issues of and the requirements for the deployment and operation of Dual-Stack Lite (DS-Lite). This document describes the various deployment considerations and applicability of the DS-Lite architecture.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6908.

Table of Contents

1. Overview

DS-Lite [RFC6333] is a transition technique that enables operators to multiplex public IPv4 addresses while provisioning only IPv6 to users. DS-Lite is designed to continue offering IPv4 services while operators upgrade their networks incrementally to IPv6. DS-Lite combines IPv4-in-IPv6 softwire [RFC2473] and Network Address Translator IPv4/IPv4 (NAT44) [RFC3022] to enable more than one user to share a public IPv4 address.

While Appendix A of [RFC6333] explains how to deploy DS-Lite within specific scenarios, the purpose of this document is to describe problems that arise when deploying DS-Lite and what guidance should be taken to mitigate those issues. The information is based on real deployment experience and is compiled in one comprehensive document so that operators aren't required to search through various RFCs deciding which sections are applicable and impact their DS-Lite deployment.

2. AFTR Deployment Considerations

2.1. Interface Consideration

Address Family Transition Router (AFTR) is a network element that is deployed inside the operator's network. An AFTR can be a stand-alone device or be embedded into a router. The AFTR is the IPv4-in-IPv6 tunnel termination point and the NAT44 device. It is deployed at the IPv4-IPv6 network border where the tunnel interface is IPv6 and the external NAT44 interface is IPv4. The Basic Bridging BroadBand (B4) element [RFC6333] is a function implemented on a dual-stack-capable node (either a host device or a home gateway) that creates a tunnel to an AFTR. Although an operator can configure both softwire tunnel termination and interface for NAT44 functions on a single physical interface (yet, keep them logically separated), there are scenarios we recommend to configure two individual interfaces (i.e., one dedicated for IPv4 and one dedicated for IPv6) to segregate the functions.

- The access network between the B4 and AFTR is an IPv6-only network, and the network between the AFTR and IPv4 network is an IPv4-only network. In this deployment scenario, the AFTR interface to the IPv6-only network and the interface to the IPv4 network should use two physical interfaces on the AFTR.

- Operators may use Operations Support System (OSS) tools (e.g., Multi Router Traffic Grapher) to collect interface data packet count information. If an operator wants to separate the softwire function and NAT44 function on different physical interfaces for

collecting a data packet count, and the AFTR does not support packet count for logical interfaces, they should use two physical interfaces on the AFTR.

2.2. MTU and Fragmentation Considerations

DS-Lite is part tunneling protocol. Tunneling introduces overhead to the packet and decreases the effective MTU size after encapsulation. DS-Lite users may experience problems with applications such as not being able to download Internet pages or transfer large files.

Since fragmentation and reassembly is not optimal, the operator should do everything possible to eliminate the need for it. If the operator uses simple IPv4-in-IPv6 softwire [RFC2473], it is recommended that the MTU size of the IPv6 network between the B4 and the AFTR accounts for the additional overhead (40 bytes). If the access network MTU size is fixed and cannot be changed, the operator should be aware that the B4 and the AFTR must support fragmentation as defined in [RFC6333]. The operator should also be aware that reassembly at the Tunnel Exit-Point is resource intensive as a large number of B4 may terminate on the same AFTR. Scalability of the AFTR is advised in this scenario.

2.3. Logging at the AFTR

A source-specific log is essential for backtracking specific hosts when a problem is identified with one of the AFTR's NAT-ed addresses. The source-specific log contains the B4 IPv6 source address, transport protocol, source port, and source IPv4 address after it has been NAT-ed. Using the source-specific log, operators can uniquely identify a specific host when a DS-Lite host experiences problems accessing the IPv4 network. To maximize IPv4 shared ratio, an operator may configure a short timeout value for NAT44 entries. This will result in a large number of logs created by the AFTR. For operators who desire to aggregate the logs, they can configure the AFTR to preallocate a range of ports to each B4. This range of ports will be used in the NAT44 function, and the AFTR will create one log entry for the whole port range. This aggregation can significantly reduce the log size for source-specific logging.

Some operators may require logging both source and destination information for a host's connections. This is called a destination-specific log. A destination-specific log contains the B4's IPv6 address, transport protocol, source port, source IPv4 address after it has been NAT-ed, destination port, and destination IPv4 address. A destination-specific log is session-based; the operators should be aware that they will not be able to aggregate log entries. When using a destination-specific log, the operator must be careful of the

large number of log entries created by the AFTR. Some AFTR implementations may keep the logs in their main memory. This may be CPU and memory resource intensive. The operators should configure the AFTR to periodically send logs to storage facility and then purge them from the AFTR.

2.4. Blacklisting a Shared IPv4 Address

The AFTR is a NAT device. It enables multiple B4s to share a single public IPv4 address. [RFC6269] discusses some considerations when sharing an IPv4 address. When a public IPv4 address is blacklisted by a remote peer, this may affect multiple users or hosts. Operators deploying DS-Lite should be aware that Internet hosts may not be aware that a given single IPv4 address is actually shared by multiple B4s. A content provider might block services for a shared IPv4 address and this would then impact all B4s sharing this particular IPv4 address. The operator would be likely to receive calls related to service outage and would then need to take appropriate corrective actions. [RFC6302] describes necessary information required to identify a user or host in shared address environment. It is also worth mention that [NAT-REVEAL] analyses different approaches to identify a user or host in a shared address environment.

2.5. AFTR's Policies

There are two types of AFTR policies:

- Outgoing Policies apply to packets originating from B4 to the AFTR. These policies should be provisioned on the AFTR's IPv6 interface that is connected to the B4s.

- Incoming Policies apply to packets originating from IPv4 networks to B4s. These policies should be provisioned on the IPv4 interface connected to the IPv4 network.

2.5.1. Outgoing Policy

Outgoing Policies may include Access Control List (ACL) and Quality of Service (QoS) settings. These policies control the packets from B4s to the AFTR. For example, the operator may configure the AFTR only to accept B4 connections that originated from specific IPv6 prefixes configured in the AFTR. More discussion of this use case can be found in Section 2.12. An operator may configure the AFTR to give priority to the packets marked by certain Differentiated Services Code Point (DSCP) values [RFC2475]. Furthermore, an AFTR may also apply an Outgoing Policy to limit the rate of port allocation for a single B4's IPv6 address.

Some operators offer different service level agreements (SLAs) to users to meet their requirements. Some users may require more ports and some may require different service priority. In this deployment scenario, the operator can implement Outgoing Policies specified to a user's B4 or a group of B4s sharing the same policies.

2.5.2. Incoming Policy

Similar to the Outgoing Policy, an Incoming Policy may also include ACL and QoS settings. The Outgoing Policy controls packets coming from the IPv4 network to the B4s. Incoming packets are normally treated equally, so these policies are globally applied. For example, an operator wants to use a predefined DSCP value to signal the IPv6 access network to apply certain traffic policies. In this deployment scenario, the operator can configure the AFTR to mark the incoming packets with the predefined DSCP value. This policy will apply to all incoming packets from the IPv4 network.

2.6. AFTR Impacts on Accounting Process

This section discusses IPv4 and IPv6 traffic accounting in the DS-Lite environment. In a typical broadband access scenario (e.g., DSL or Cable), the B4 is embedded in a Residential Gateway. The edge router for the B4s in the provider's network is an IPv6 edge router. The edge router is usually responsible for IPv6 accounting and the user management functions such as authentication, authorization, and accounting (AAA). However, given the fact that IPv4 traffic is encapsulated in an IPv6 packet at the B4 and only decapsulated at the AFTR, the edge router will require additional functionality to associate IPv4 accounting information to the B4 IPv6 address. If DS-Lite is the only application using the IPv4-in-IPv6 protocol in the IPv6 access network, the operator can configure the edge router to check the IPv6 Next Header field in the IPv6 header, identify the protocol type (i.e., 0x04), and collect IPv4 accounting information.

Alternatively, the AFTR may perform accounting for IPv4 traffic. However, operators must be aware that this will introduce some challenges, especially in DSL deployment. In DSL deployment, the AAA transaction normally happens between the edge router (i.e., Broadband Network Gateway) and AAA server. [RFC6333] does not require the AFTR to interact with the AAA server or edge router. Thus, the AFTR may not have the AAA parameters (e.g., Account Session ID) associated with B4s to generate an IPv4 accounting record. IPv4 traffic accounting at the AFTR is not recommended when the AAA parameters necessary to generate complete IPv4 accounting records are not available. The accounting process at the AFTR is only necessary if the operator requires separating per-B4 accounting records for IPv4 and IPv6 traffic. If the per-B4 IPv6 accounting records, collected

by the edge router, are sufficient, then the additional complexity of enabling IPv4 accounting at the AFTR is not required. It is important to notice that, since the IPv4 traffic is encapsulated in IPv6 packets, the data collected by the edge router for IPv6 traffic already contains the total amount of traffic (i.e., IPv4 and IPv6).

Even if detailed accounting records collection for IPv4 traffic may not be required, it would be useful for an operator, in some scenarios, to have information that the edge router generates for the IPv6 traffic. This information can be used to identify the AFTR who is handling the IPv4 traffic for that B4. This can be achieved by adding additional information to the IPv6 accounting records. For example, operators can use RADIUS attribute information specified in [RFC6519] or a new attribute to be specified in Internet Protocol Detailed Record (IPDR).

2.7. Reliability Considerations of AFTR

For robustness, reliability, and load distribution purposes, operators may deploy multiple AFTRs. In such cases, the IPv6 prefixes and algorithm to build the tunneling mechanisms configured on each of these AFTRs will be the same. In [RFC6333], Appendix A.3 mentions that High Availability (HA) is the operator's responsibility. Since DS-Lite is a stateful mechanism, all requirements for load-balancing and failover mechanisms apply. There are many ways to implement HA in a stateful mechanism; the most common are Cold Standby mode and Hot Standby mode. More discussion on deploying these two modes for NAT can be found in [NAT-STANDBY]. In Cold Standby mode, the AFTR states are not replicated from the Primary AFTR to the Backup AFTR. When the Primary AFTR fails, all the existing established sessions will be flushed out. The internal hosts are required to reestablish sessions with the external hosts. In Hot Standby mode, the session's states are replicated on-the-fly from the Primary AFTR to the Backup AFTR. When the Primary AFTR fails, the Backup AFTR will take over all the existing established sessions. In this mode, the internal hosts are not required to reestablish sessions with the external hosts.

For operators, the decision to use Cold Standby mode or Hot Standby mode depends on the trade-off between capital cost and operational cost. Cold Standby mode does not require a Backup Standby AFTR to synchronize session states. This simplifies the operational model. When the Primary AFTR goes down, any AFTR with extra capacity can take over. Hot Standby mode provides a smoother failover experience to users; the cost for the operators is more careful failover planning. For most deployment scenarios, we believe that Cold Standby mode should be sufficient enough and is thus recommended.

2.8. Strategic Placement of AFTR

In the DS-Lite environment, the AFTR is the logical next-hop router of the B4s to access the IPv4 network, so the placement of the AFTR will affect the traffic flows in the access network and overall network design. In general, there are two placement models to deploy an AFTR. Model One deploys the AFTR at the edge of the network to cover a small region. Model Two deploys the AFTR at the core of the network to cover a large region.

When an operator considers where to deploy the AFTR, the operator must make trade-offs. The AFTR in Model One serves fewer B4s; thus, it requires a less powerful AFTR. Moreover, the traffic flows are more evenly distributed to the AFTRs. However, it requires deploying more AFTRs to cover the entire network. Often, the operation cost increases proportionally with the amount of network equipment.

The AFTR in Model Two covers a larger area; thus, it serves more B4s. The operator could deploy only a few AFTRs to support the entire user base. However, this model requires a more powerful AFTR to sustain the load at peak hours. Since the AFTR would support B4s from different regions, the AFTR would be deployed closer to the core network.

DS-Lite framework can be incrementally deployed. An operator may consider starting with Model Two. When the demand increases, the operator can push the AFTR closer to the edge, which would effectively become Model One.

2.9. AFTR Considerations for Geographically Aware Services

By centralizing public IPv4 addresses in the AFTR, remote services can no longer rely on an IPv4 address and IPv4 routing information to derive a host's geographical information. For example, the IPv6 access network and the AFTR may be in two different cities. If the remote services rely on the IPv4 address to locate a host, they may have thought the host was in a different city. [RFC6269] Section 7 describes the problem in more detail. Applications could explicitly ask users to enter location information, such as postal code or telephone number, before offering geographical service. In contrast, applications could use HTTP-Enabled Location Delivery (HELD) [RFC5985] to get the location information from the Location Information Server and give this information to the remote peer. [RFC6280] describes an architecture to enable location-based services. However, to mitigate the impact, we recommend that operators deploy the AFTR as close to B4s as possible.

2.10. Impacts on QoS Policy

This section describes the application of [RFC2983] to the DS-Lite deployment model. Operators must ensure that the QoS policy that is in place operates properly within the DS-Lite deployment. In this regard, operators commonly use DSCP [RFC2475] to classify and prioritize different types of traffic in their networks. DS-Lite tunnel can be seen as a particular case of uniform conceptual tunnel model, as described in Section 3.1 of [RFC2983]. The uniform model views an IP tunnel only as a necessary mechanism to forward traffic to its destination: the tunnel has no significant impact on traffic conditioning. In this model, any packet has exactly one DSCP field that is used for traffic conditioning at any point, and it is the field in the outermost IP header. In the DS-Lite model, this is the Traffic Class field in the IPv6 header. According to [RFC2983], implementations of this model copy the DSCP value to the outer IP header at encapsulation and copy the outer header's DSCP value to the inner IP header at decapsulation.

Operators should use this model by provisioning the network such that the AFTR copies the DSCP value in the IPv4 header to the Traffic Class field in the IPv6 header, after the encapsulation for the downstream traffic. Similarly, the B4 copies the DSCP value in the IPv4 header to the Traffic Class field to the IPv6 header, after the encapsulation for the upstream traffic. Traffic identification and classification can be done by examining the outer IPv6 header in the IPv6 access network.

2.11. Port Forwarding Considerations

Some applications behind the B4 require the B4 to accept incoming requests. If the remote application wants to communicate to the application behind the B4, the remote application must know both the NAT-ed IPv4 address used by the B4 and the IPv4 destination port. Some applications use Universal Plug and Play (UPnP) (e.g., popular gaming consoles) or Interactive Community Establishment (ICE) [RFC5245] to request incoming ports. Some applications rely on Application Level Gateway (ALG) or manual port configuration to reserve a port in the NAT. For the DS-Lite deployment scenario whereby the B4 does not own a full IPv4 address, the operator will manage port-forwarding in the serving AFTR. Operators may use Port Control Protocol (PCP) [PCP-BASE] as guidance to provide port forwarding service. Operators will deploy PCP client in the B4s. PCP permits the PCP server to be deployed in a stand-alone server. However, we recommend that operators consider deploying the PCP server in the AFTR. This will ease the overhead to design a global configuration for the PCP server for many AFTRs because each PCP server will be dedicated to the collocated AFTR.

When sharing an IPv4 address, not all of the ports are available to a B4. Some restricted ports (i.e., 0-1023) are well known such as TCP port 25 and 80. Many users may want to be provisioned with the restricted ports. For fairness, we recommend that operators configure the AFTR and not allocate the restricted ports to regular DS-Lite B4s. This operation model ensures that DS-Lite B4s will have uniform configuration, which can simplify provisioning and operation. For users who want to use the restricted ports, operators can consider provisioning a full IPv4 address to those users' B4s. If an operator still wants to provision restricted ports to specific B4s, it may require implementing a static B4's configuration in the AFTR to match the B4's IPv6 address to the NAT rules. Alternatively, the B4 may dynamically allocate the ports, and the AFTR authenticates the session's request using PCP [PCP-BASE].

2.12. DS-Lite Tunnel Security

[RFC6333], Section 11 describes security issues associated with the DS-Lite mechanism. To restrict the service offered by the AFTR only to registered B4s, an operator can implement the Outgoing Policy on the AFTR's tunnel interface to accept only the IPv6 prefixes defined in the policy. For static provisioning, the operator will need to know in advance the IPv6 prefixes provisioned to the B4s for the softwire in order to configure the policy. To simplify operation, operators should configure the AFTRs in the same region with the same IPv6 prefixes' Outgoing Policy. The AFTRs will accept both regular connections and failover connections from the B4s in the same service region.

2.13. IPv6-Only Network Considerations

In environments where the operator wants to deploy the AFTR in an IPv6-only network, the AFTR nodes may not have direct IPv4 connectivity. In this scenario, the operator extends the IPv6-only boundary to the border of the network and only the border routers have IPv4 connectivity. For both scalability and performance purposes, the AFTR is located in the IPv6-only network closer to B4s. In this scenario, the AFTR has only IPv6 connectivity and must be able to send and receive IPv4 packets. Enhancements to the DS-Lite AFTR are required to achieve this. [DS-LITE] describes such issues and enhancements to DS-Lite in IPv6-only deployments.

3. B4 Deployment Considerations

In order to configure the IPv4-in-IPv6 tunnel, the B4 needs the IPv6 address of the AFTR. This IPv6 address can be configured using a variety of methods ranging from an out-of-band mechanism, manual configuration, and DHCPv6 option to RADIUS. If an operator uses

DHCPv6 to provision the B4, the B4 must implement the DHCPv6 option defined in [RFC6334]. If an operator uses RADIUS to provision the B4, the B4 must implement [RFC6519].

3.1. DNS Deployment Considerations

[RFC6333] recommends that the B4 send DNS queries to an external recursive resolver over IPv6. The B4 should implement a proxy resolver that will proxy a DNS query from IPv4 transport to the DNS server in the IPv6 network. [RFC6333] does not describe the DNS proxy behavior. In deployment, the operator must ensure that the DNS proxy implementation must follow [RFC5625]. This is important especially for operators who have deployed, or will consider deploying, DNSSEC [RFC4035].

Some operators may want to give hosts behind the B4 an IPv4 address of an external DNS recursive resolver. The B4 will treat the DNS packets as normal IP packets and forward them over the softwire. Note that there is no effective way to provision an IPv4 DNS address to the B4 over IPv6; operators who use this DNS deployment model must be aware that how to provision an IPv4 DNS address over an IPv6 network is undefined, so it will introduce additional complexity in B4 provisioning. Moreover, this will increase the load to the AFTR by creating entries in the NAT table for DNS sessions. Operators may deploy a local DNS caching resolver in the AFTR to reduce the load in the NAT table. Nonetheless, this DNS model is not covered in [RFC6333] and is not recommended.

3.2. IPv4 Service Monitoring

3.2.1. B4 Remote Management

B4 is connected to the IPv6 access network to offer IPv4 services. When users experience IPv4 connectivity issues, operators must be able to remotely access (e.g., TR-069) the B4 to verify its configuration and status. Operators should access B4s using native IPv6. Operators should not access B4 over the softwire.

3.2.2. IPv4 Connectivity Check

The DS-Lite framework provides IPv4 services over the IPv6 access network. Operators and users must be able to check the IPv4 connectivity from the B4 to its AFTR using ping and IPv4 traceroute. The AFTR should be configured with an IPv4 address to enable a PING test and a Traceroute test. Operators should assign the same IPv4 address (e.g., 192.0.0.2/32 [RFC6333]) to all AFTRs. IANA has allocated the 192.0.0.0/29 network prefix to provide IPv4 addresses for this purpose [RFC6333].

4. Security Considerations

This document does not present any new security issues. [RFC6333] discusses DS-Lite related security issues.

5. Acknowledgements

Thanks to Mr. Nejc Skoberne and Dr. Maoke Chen for their thorough review and helpful comments. We also want to thank Mr. Hu Jie for sharing his DS-Lite deployment experience with us. He gave us recommendations of what his company learned while testing DS-Lite in the production network.

6. References

6.1. Normative References

[RFC6333] Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", RFC 6333, August 2011.

[RFC6334] Hankins, D. and T. Mrugalski, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6) Option for Dual-Stack Lite", RFC 6334, August 2011.

[RFC6519] Maglione, R. and A. Durand, "RADIUS Extensions for Dual-Stack Lite", RFC 6519, February 2012.

6.2. Informative References

[DS-LITE] Boucadair, M., Jacquenet, C., Grimault, J., Kassi-Lahlou, M., Levis, P., Cheng, D., and Y. Lee, "Deploying Dual-Stack Lite in IPv6 Network", Work in Progress, April 2011.

[NAT-REVEAL] Boucadair, M., Touch, J., Levis, P., and R. Penno, "Analysis of Solution Candidates to Reveal a Host Identifier (HOST_ID) in Shared Address Deployments", Work in Progress, March 2013.

[NAT-STANDBY] Xu, X., Boucadair, M., Lee, Y., and G. Chen, "Redundancy Requirements and Framework for Stateful Network Address Translators (NAT)", Work in Progress, October 2010.

[PCP-BASE] Wing, D., Cheshire, S., Boucadair, M., Penno, R., and P. Selkirk, "Port Control Protocol (PCP)", Work in Progress, November 2012.

[RFC2473] Conta, A. and S. Deering, "Generic Packet Tunneling in IPv6 Specification", RFC 2473, December 1998.

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, December 1998.

[RFC2983] Black, D., "Differentiated Services and Tunnels", RFC 2983, October 2000.

[RFC3022] Srisuresh, P. and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", RFC 3022, January 2001.

[RFC4035] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "Protocol Modifications for the DNS Security Extensions", RFC 4035, March 2005.

[RFC5245] Rosenberg, J., "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols", RFC 5245, April 2010.

[RFC5625] Bellis, R., "DNS Proxy Implementation Guidelines", BCP 152, RFC 5625, August 2009.

[RFC5985] Barnes, M., "HTTP-Enabled Location Delivery (HELD)", RFC 5985, September 2010.

[RFC6269] Ford, M., Boucadair, M., Durand, A., Levis, P., and P. Roberts, "Issues with IP Address Sharing", RFC 6269, June 2011.

[RFC6280] Barnes, R., Lepinski, M., Cooper, A., Morris, J., Tschofenig, H., and H. Schulzrinne, "An Architecture for Location and Location Privacy in Internet Applications", BCP 160, RFC 6280, July 2011.

[RFC6302] Durand, A., Gashinsky, I., Lee, D., and S. Sheppard, "Logging Recommendations for Internet-Facing Servers", BCP 162, RFC 6302, June 2011.

Authors' Addresses

Yiu L. Lee
Comcast
One Comcast Center
Philadelphia, PA 19103
U.S.A.

EMail: yiu_lee@cable.comcast.com
URI: http://www.comcast.com

Roberta Maglione
Cisco Systems
181 Bay Street
Toronto, ON M5J 2T3
Canada

EMail: robmgl@cisco.com

Carl Williams
MCSR Labs
U.S.A.

EMail: carlw@mcsr-labs.org

Christian Jacquenet
France Telecom
Rennes
France

EMail: christian.jacquenet@orange.com

Mohamed Boucadair
France Telecom
Rennes
France

EMail: mohamed.boucadair@orange.com