Internet Engineering Task Force (IETF)
Request for Comments: 6674
Category: Standards Track
ISSN: 2070-1721

F. Brockners
S. Gundavelli
Cisco
S. Speicher
Deutsche Telekom AG
D. Ward
Cisco

July 2012

Gateway-Initiated Dual-Stack Lite Deployment

Abstract

Gateway-Initiated Dual-Stack Lite (GI-DS-Lite) is a variant of Dual-Stack Lite (DS-Lite) applicable to certain tunnel-based access architectures. GI-DS-Lite extends existing access tunnels beyond the access gateway to an IPv4-IPv4 NAT using softwires with an embedded Context Identifier that uniquely identifies the end-system to which the tunneled packets belong. The access gateway determines which portion of the traffic requires NAT using local policies and sends/receives this portion to/from this softwire.

Status of This Memo

This is an Internet Standards Track document.

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). Further information on Internet Standards is available in 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/rfc6674.

Table of Contents

1. Overview

Gateway-Initiated Dual-Stack Lite (GI-DS-Lite) is a variant of Dual-Stack Lite (DS-Lite) [RFC6333], applicable to network architectures that use point-to-point tunnels between the access device and the access gateway. The access gateway in these models is designed to serve large numbers of access devices. Mobile architectures based on Mobile IPv6 [RFC6275], Proxy Mobile IPv6 [RFC5213], or GPRS Tunnelling Protocol (GTP) [TS29060], as well as broadband architectures based on PPP or point-to-point VLANs as defined by the Broadband Forum [TR59][TR101], are examples of this type of architecture.

The DS-Lite approach leverages IPv4-in-IPv6 tunnels (or other tunneling modes) for carrying the IPv4 traffic from the customer network to the Address Family Transition Router (AFTR). An established softwire between the AFTR and the access device is used for traffic-forwarding purposes. This makes the inner IPv4 address irrelevant for traffic routing and allows sharing private IPv4 addresses [RFC1918] between customer sites within the service provider network.

Similarly to DS-Lite, GI-DS-Lite enables the service provider to share public IPv4 addresses among different customers by combining tunneling and NAT. It allows multiple access devices behind the access gateway to share the same private IPv4 address [RFC1918]. Rather than initiating the tunnel right on the access device, GI-DS-Lite logically extends the already existing access tunnels beyond the access gateway towards the AFTR using a tunneling mechanism with semantics for carrying context state related to the encapsulated traffic. This approach results in supporting overlapping IPv4 addresses in the access network, requiring no changes to either the access device or the access architecture. Additional tunneling overhead in the access network is also omitted. If, for example, an encapsulation mechanism based on Generic Routing Encapsulation (GRE) is chosen, it allows the network between the access gateway and the AFTR to be either IPv4 or IPv6 and allows the operator to migrate to IPv6 in incremental steps.

2. Conventions

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 [RFC2119].

The following abbreviations are used within this document:

AFTR: Address Family Transition Router. An AFTR combines IP-in-IP tunnel termination and IPv4-IPv4 NAT.

AD: Access Device. It is the end host, also known as the mobile node in mobile architectures.

AG: Access Gateway. A gateway in the access network, such as LMA, Home Agent (HA), GGSN, or PDN-GW in mobile architectures.

CID: Context Identifier

DS-Lite: Dual-Stack Lite

GGSN: Gateway GPRS Support Node

GI-DS-Lite: Gateway-Initiated DS-Lite

NAT: Network Address Translator

PDN-GW: Packet Data Network Gateway

SW: Softwire [RFC4925]

SWID: Softwire Identifier

3. Gateway-Initiated DS-Lite

This section provides an overview of Gateway-Initiated DS-Lite (GI-DS-Lite). Figure 1 outlines the generic deployment scenario for GI-DS-Lite. This generic scenario can be mapped to multiple different access architectures, some of which are described in Appendix A.

In Figure 1, access devices (AD-1 and AD-2) are connected to the access gateway using some form of tunnel technology and the same is used for carrying IPv4 (and optionally IPv6) traffic of the access device. These access devices may also be connected to the access gateway over point-to-point links. The details on how the network delivers the IPv4 address configuration to the access devices are specific to the access architecture and are outside the scope of this document. With GI-DS-Lite, the access gateway and AFTR are connected by a softwire [RFC4925]. The softwire is identified by a softwire identifier (SWID). The SWID does not need to be globally unique, i.e., different SWIDs could be used to identify a softwire at the different ends of a softwire. The form of the SWID depends on the tunneling technology used for the softwire. The SWID could, for

example, be the endpoints of a GRE tunnel or a VPN-ID. See Section 6 for details. A Context Identifier (CID) is used to multiplex flows associated with the individual access devices onto the softwire. It is deployment dependent whether the flows from a particular AD are identified using the source IP address or an access tunnel identifier. Local policies at the access gateway determine which part of the traffic received from an access device is tunneled over the softwire to the AFTR. The combination of CID and SWID must be unique between the access gateway and AFTR to identify the flows associated with an AD. The CID is typically a 32-bit-wide identifier and is assigned by the access gateway. It is retrieved either from a local or remote (e.g., Authentication, Authorization, and Accounting (AAA)) repository. Like the SWID, the embodiment of the CID depends on the tunnel mode used and the type of the network connecting the access gateway and AFTR. If, for example, GRE [RFC2784] with GRE Key and Sequence Number extensions [RFC2890] is used as the softwire technology, the network connecting the access gateway and AFTR could be a IPv4-only, IPv6-only, or dual-stack IP network. The CID would be carried within the GRE Key field. Section 6 details the different softwire types supported with GI-DS-Lite.

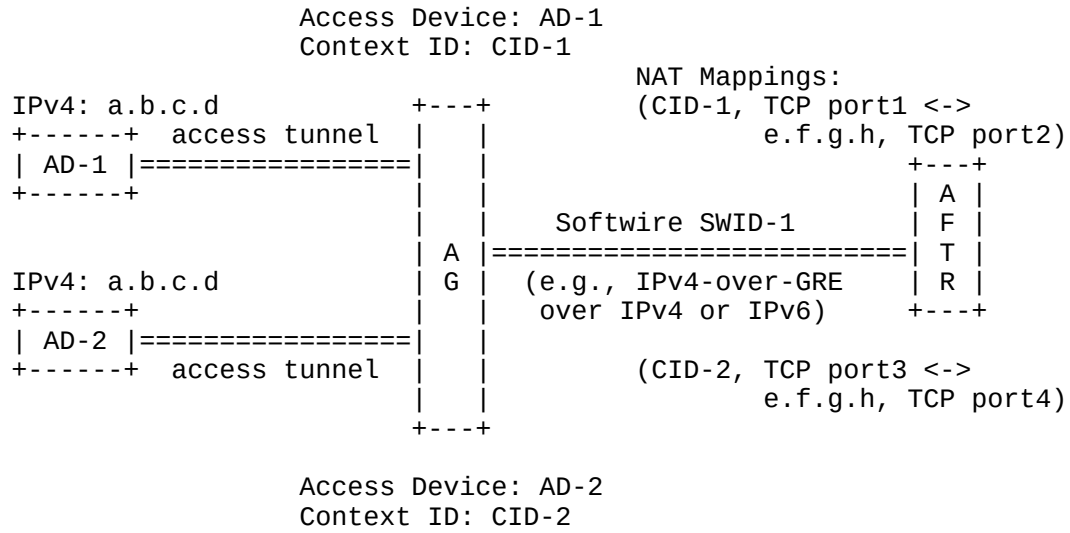

Figure 1: Gateway-Initiated Dual-Stack Lite Reference Architecture

The AFTR combines softwire termination and IPv4-IPv4 NAT. The NAT binding of the AD's address could be assigned autonomously by the AFTR from a local address pool, configured on a per-binding basis (either by a remote control entity through a NAT control protocol or through manual configuration), or derived from the CID (e.g., the CID, if 32 bits wide, could be mapped 1:1 to an external IPv4

address). A simple example of a translation table at the AFTR is shown in Figure 2. The choice of the appropriate translation scheme for a traffic flow can take into account parameters such as destination IP address, incoming interface, etc. The IP address of the AFTR, which will either be an IPv6 or an IPv4 address depending on the transport network between the access gateway and the AFTR, is configured on the access gateway. A variety of methods, such as out-of-band mechanisms or manual configuration, apply.

Softwire-ID/Context-ID/IPv4/Port	Public IPv4/Port
SWID-1/CID-1/a.b.c.d/TCP-port1	e.f.g.h/TCP-port2
SWID-1/CID-2/a.b.c.d/TCP-port3	e.f.g.h/TCP-port4

Figure 2: Example Translation Table at the AFTR

GI-DS-Lite does not require a 1:1 relationship between an access gateway and AFTR, but more generally applies to (M:N) scenarios, where M access gateways are connected to N AFTRs. Multiple access gateways could be served by a single AFTR. AFTRs could be dedicated to specific groups of access-devices, groups of access gateways, or geographic regions. An AFTR could be, but does not have to be, co-located with an access gateway.

4. Protocol and Related Considerations

- Depending on the embodiment of the CID (e.g., for GRE encapsulation with GRE Key), the NAT binding entry maintained at the AFTR, which reflects an active flow between an access device inside the network and a node in the Internet, SHOULD be extended to include the CID and the identifier of the softwire (SWID).

- When creating an IPv4-to-IPv4 NAT binding for an IPv4 packet flow received from the access gateway over the softwire, the AFTR SHOULD associate the CID with that NAT binding. It SHOULD use the combination of CID and SWID as the unique identifier and use those parameters in the NAT binding entry.

- When forwarding a packet to the access device, the AFTR SHOULD obtain the CID from the NAT binding associated with that flow. For example, in case of GRE encapsulation, it SHOULD add the CID to the GRE Key and Sequence Number extension of the GRE header and tunnel it to the access gateway.

- On receiving any packet from the softwire, the AFTR SHOULD obtain the CID from the incoming packet and use it for performing NAT binding lookup and packet translation before forwarding the packet.

- The access gateway, on receiving any IPv4 packet from the access device, SHOULD lookup the CID for that access device. In case of GRE encapsulation, it can, for example, add the CID to the GRE Key and Sequence Number extension of the GRE header and tunnel it to the AFTR.

- On receiving any packet from the softwire, the access gateway SHOULD obtain the CID from the packet and use it for making the forwarding decision.

- When encapsulating an IPv4 packet, the access gateway and AFTR SHOULD use its Diffserv Codepoint (DSCP) to derive the DSCP (or MPLS Traffic-Class Field in the case of MPLS) of the softwire.

5. Softwire Management and Related Considerations

The following are the considerations related to the operational management of the softwire between the AFTR and access gateway.

- The softwire between the access gateway and the AFTR MAY be created at system startup time OR dynamically established on-demand. It is deployment dependent which Operations, Administration, and Maintenance (OAM) mechanisms (such as ICMP, Bidirectional Forwarding Detection (BFD) [RFC5880], or Label Switched Path (LSP) ping [RFC4379]) are employed by the access gateway and AFTR for softwire health management and corresponding protection strategies.

- The softwire peers MAY be provisioned to perform policy enforcement, such as for determining the protocol type or overall portion of traffic that gets tunneled or for any other settings related to quality of service. The specific details on how this is achieved or the types of policies that can be applied are outside the scope for this document.

- The softwire peers SHOULD use the correct path MTU value for the tunnel path between the access gateway and the AFTR. This value MAY be statically configured at softwire creation time or dynamically discovered using the standard path MTU discovery techniques.

- An access gateway and an AFTR can have multiple softwires established between them (e.g., to separate address domains, provide for load-sharing, etc.).

6. Softwire Embodiments

Which tunnel technologies can be applied for the softwire connecting an access gateway and AFTR are dependent on the deployment and the requirements. GRE encapsulation with GRE Key extension, MPLS VPNs [RFC4364], or plain IP-in-IP encapsulation can be used. Softwire identification and CID depend on the tunneling technology employed:

- GRE with GRE Key: SWID is the tunnel identifier of the GRE tunnel between the access gateway and the AFTR. The CID is the GRE Key associated with the AD.

- MPLS VPN: The SWID is a generic identifier that uniquely identifies the VPN at either the access gateway or AFTR. Depending on whether the access gateway or AFTR is acting as customer edge (CE) or provider edge (PE), the SWID could, for example, be an attachment circuit identifier, an identifier representing the set of VPN route labels pointing to the routes within the VPN, etc. The AD's IPv4 address is the CID. For a given VPN, the AD's IPv4 address must be unique.

- IPv4/IPv6-in-MPLS: The SWID is the top MPLS label. CID might be the next MPLS label in the stack, if present, or the IP address of the AD.

- IPv4-in-IPv4: SWID is the outer IPv4 source address. The AD's IPv4 address is the CID. For a given outer IPv4 source address, the AD's IPv4 address must be unique.

- IPv4-in-IPv6: SWID is the outer IPv6 source address. If the AD's IPv4 address is used as CID, the AD's IPv4 address must be unique. If the IPv6 Flow Label [RFC6437] is used as CID, the IPv4 addresses of the ADs may overlap. Given that the IPv6 Flow Label is 20 bits wide, which is shorter than the recommended 32-bit CID, large-scale deployments may require additional scaling considerations. In addition, one should ensure sufficient randomization of the IP Flow Label to avoid possible interference with other uses of the IP Flow Label, such as Equal Cost Multipath (ECMP) support.

Figure 3 gives an overview of the different tunnel modes as they apply to different deployment scenarios. "x" indicates that a certain deployment scenario is supported. The following abbreviations are used:

- IPv4 address
  - "up": Deployments with "unique private IPv4 addresses" assigned to the access devices are supported.
  - "op": Deployments with "overlapping private IPv4 addresses" assigned to the access devices are supported.
  - "s": Deployments where all access devices are assigned the same IPv4 address are supported.
- Network-type
  - "v4": The access gateway and AFTR are connected by an IPv4-only network.
  - "v6": The access gateway and AFTR are connected by an IPv6-only network.
  - "v4v6": The access gateway and AFTR are connected by a dual-stack network, supporting IPv4 and IPv6.
  - "MPLS": The access gateway and AFTR are connected by an MPLS network

Softwire	IPv4 address			Network-type			
	up	op	s	v4	v6	v4v6	MPLS
GRE with GRE Key	x	x	x	x	x	x	
MPLS VPN	x	x					x
IPv4/IPv6-in-MPLS	x	x	x				x
IPv4-in-IPv4	x			x			
IPv4-in-IPv6	x				x		
IPv4-in-IPv6 w/ FL	x	x	x		x		

Figure 3: Tunnel Modes and Their Applicability

7. Security Considerations

The approach specified in this document allows the use of Dual-Stack Lite for tunnel-based access architectures. Rather than initiating the tunnel from the access device, GI-DS-Lite logically extends the already existing access tunnel beyond the access gateway towards the AFTR and builds a virtual softwire between the AFTR and the access device. This approach requires the use of an additional Context Identifier in the AFTR and at the access gateway, which is required for making IP packet-forwarding decisions.

If a packet is received with an incorrect Context Identifier at the access gateway/AFTR, it will be associated with an incorrect access device. Therefore, care must be taken to ensure an IP packet tunneled between the access gateway and the AFTR is carried with the Context Identifier of the access device associated with that IP packet. The Context Identifier is not carried from the access device, and it is not possible for one access device to claim the Context Identifier of some other access device. However, it is possible that an on-path attacker between the access gateway and the AFTR can potentially modify the Context Identifier in the packet, resulting in association of the packet to an incorrect access device. This threat is no different from an on-path attacker modifying the source/destination address of an IP packet. However, this threat can be prevented by enabling IPsec security with integrity protection turned on, between the access gateway and the AFTR, that will ensure the correct binding of the Context Identifier and the inner packet. This specification does not require any new security considerations other than those specified in the Dual-Stack Lite specification [RFC6333] and in the security considerations specified for the given access architecture, such as Proxy Mobile IPv6, leveraging this transitioning scheme.

8. Acknowledgements

The authors would like to acknowledge the discussions on this topic with Mark Grayson, Jay Iyer, Kent Leung, Vojislav Vucetic, Flemming Andreasen, Dan Wing, Jouni Korhonen, Teemu Savolainen, Parviz Yegani, Farooq Bari, Mohamed Boucadair, Vinod Pandey, Jari Arkko, Eric Voit, Yiu L. Lee, Tina Tsou, Guo-Liang Yang, Cathy Zhou, Olaf Bonness, Paco Cortes, Jim Guichard, Stephen Farrell, Pete Resnik, and Ralph Droms.

9. References

9.1. Normative References

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

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

[RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 2784, March 2000.

[RFC2890] Dommety, G., "Key and Sequence Number Extensions to GRE", RFC 2890, September 2000.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, February 2006.

[RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures", RFC 4379, February 2006.

[RFC5213] Gundavelli, S., Leung, K., Devarapalli, V., Chowdhury, K., and B. Patil, "Proxy Mobile IPv6", RFC 5213, August 2008.

[RFC5555] Soliman, H., "Mobile IPv6 Support for Dual Stack Hosts and Routers", RFC 5555, June 2009.

[RFC5844] Wakikawa, R. and S. Gundavelli, "IPv4 Support for Proxy Mobile IPv6", RFC 5844, May 2010.

[RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, June 2010.

[RFC6275] Perkins, C., Johnson, D., and J. Arkko, "Mobility Support in IPv6", RFC 6275, July 2011.

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

[RFC6437] Amante, S., Carpenter, B., Jiang, S., and J. Rajahalme, "IPv6 Flow Label Specification", RFC 6437, November 2011.

9.2. Informative References

[NAT-CONTROL] Brockners, F., Bhandari, S., Singh, V., and V. Fajardo, "Diameter Network Address and Port Translation Control Application", Work in Progress, April 2012.

[RFC4925] Li, X., Dawkins, S., Ward, D., and A. Durand, "Softwire Problem Statement", RFC 4925, July 2007.

[TR101] Broadband Forum, "TR-101: Migration to Ethernet-Based DSL Aggregation", April 2006.

[TR59] Broadband Forum, "TR-059: DSL Evolution - Architecture Requirements for the Support of QoS-Enabled IP Services", September 2003.

[TS23060] 3GPP, "Technical Specification Group Services and System Aspects; General Packet Radio Service (GPRS); Service description; Stage 2, V11.2.0", TS 23.060, 2012.

[TS23401] 3GPP, "Technical Specification Group Services and System Aspects; General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access, V11.1.0", TS 23.401, 2012.

[TS29060] 3GPP, "Technical Specification Group Core Network and Terminals; General Packet Radio Service (GPRS); GPRS Tunnelling Protocol (GTP), V11.3.0", TS 29.060, 2012.

Appendix A. GI-DS-Lite Deployment

A.1. Connectivity Establishment: Example Call Flow

Figure 4 shows an example call flow - linking access tunnel establishment on the access gateway with the softwire to the AFTR. This simple example assumes that traffic from the AD uses a single access tunnel and that the access gateway will use local policies to decide which portion of the traffic received over this access tunnel needs to be forwarded to the AFTR.

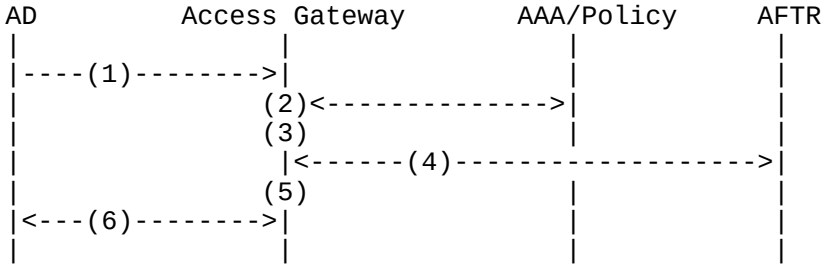

Figure 4: Example Call Flow for Session Establishment

1. The access gateway receives a request to create an access tunnel endpoint.

2. The access gateway authenticates and authorizes the access tunnel. Based on local policy or through interaction with the AAA/Policy system, the access gateway recognizes that IPv4 service should be provided using GI-DS-Lite.

3. The access gateway creates an access tunnel endpoint. The access tunnel links AD and access gateway.

4. (Optional): The access gateway and the AFTR establish a control session between themselves. This session can, for example, be used to exchange accounting or NAT-configuration information. Accounting information could be supplied to the access gateway, AAA/Policy, or other network entities that require information about the externally visible address/port pairs of a particular access device. The Diameter NAT Control Application [NAT-CONTROL] could, for example, be used for this purpose.

5. The access gateway allocates a unique CID and associates those flows received from the access tunnel that need to be tunneled towards the AFTR with the softwire linking the access gateway and AFTR. Local forwarding policy on the access gateway determines which traffic will need to be tunneled towards the AFTR.

6. The access gateway and AD complete the access tunnel establishment. Depending on the procedures and mechanisms of the corresponding access network architecture, this step can include the assignment of an IPv4 address to the AD.

A.2. GI-DS-Lite Applicability: Examples

The section outlines deployment examples of the generic GI-DS-Lite architecture described in Section 3.

- Access architectures based on Mobile-IP: In a network scenario based on Dual-Stack Mobile IPv6 (DSMIPv6) [RFC5555], the Mobile IPv6 home agent will implement the GI-DS-Lite access gateway function along with the dual-stack Mobile IPv6 functionality.

- Access architectures based on Proxy Mobile IPv6 (PMIPv6): In a PMIPv6 [RFC5213] scenario, the local mobility anchor (LMA) will implement the GI-DS-Lite access gateway function along with the PMIPv6 IPv4 support [RFC5844] functionality.

- GTP-based access architectures: Third Generation Partnership Project (3GPP) TS 23.401 [TS23401] and 3GPP TS 23.060 [TS23060] define mobile access architectures using GTP. For GI-DS-Lite, the PDN-GW or GGSN will also assume the access gateway function.

- Fixed Worldwide Interoperability for Microwave Access (WiMAX) architecture: If GI-DS-Lite is applied to fixed WiMAX, the Access Service Network Gateway (ASN-GW) will implement the GI-DS-Lite access gateway function.

- Mobile WiMAX: If GI-DS-Lite is applied to mobile WiMAX, the home agent will implement the access gateway function.

- PPP-based broadband access architectures: If GI-DS-Lite is applied to PPP-based access architectures, the Broadband Remote Access Server (BRAS) or Broadband Network Gateway (BNG) will implement the GI-DS-Lite access gateway function.

- In broadband access architectures using per-subscriber VLANs, the BNG will implement the GI-DS-Lite access gateway function.

Authors' Addresses

Frank Brockners
Cisco
Hansaallee 249, 3rd Floor
Duesseldorf, Nordrhein-Westfalen 40549
Germany

EMail: fbrockne@cisco.com

Sri Gundavelli
Cisco
170 West Tasman Drive
San Jose, CA 95134
USA

EMail: sgundave@cisco.com

Sebastian Speicher
Deutsche Telekom AG
Landgrabenweg 151
Bonn, Nordrhein-Westfalen 53277
Germany

EMail: sebastian.speicher@telekom.de

David Ward
Cisco
170 West Tasman Drive
San Jose, CA 95134
USA

EMail: wardd@cisco.com