softwire R. Maglione, Ed.
Internet-Draft W. Dec
Intended status: Informational Cisco Systems
Expires: April 16, 2016 I. Leung
Rogers Communications
E. Mallette
Bright House Networks
October 14, 2015

Use cases for MAP-T
draft-maglione-softwire-map-t-scenarios-06

Abstract

The Softwire working group standardized both encapsulation and translation based stateless IPv4/IPv6 solutions in order to be able to provide IPv4 connectivity to customers in an IPv6-Only environment.

The purpose of this document is to describe some operational use cases that would benefit from a translation based approach and highlights the operational benefits that a translation based solution would allow.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

The Softwire working group standardized both encapsulation [RFC7597] and translation [RFC7599] based stateless IPv4/IPv6 solutions developed for the purposes of offering IPv4 connectivity to the customers in an IPv6-Only environment.

There are deployment scenarios that may benefit equally from an encapsulated or translated form of an IPv4/IPv6 stateless addressing solution. There are, however, use cases where using a translation approach could lead to significant operational benefits and potential savings for the operators.

This document describes some use cases that would take advantage of a translation based solution, by highlighting the operational benefits that a translation based approach would allow.

2. Operational Service Policy Use Cases

In Broadband Networks it is common practice for Operators to apply per-subscriber policies on subscriber traffic at the network edge such as a BNG (Broadband Network Gateway), CMTS (Cable Modem Termination System), PGW (PDN Gateway) or like device. Various services may require the application of different policies at these services edges.

Typically a policy would include a classification function and an action function.

- Service flow classification may occur based on any combination of the following:

  - Layer-3 identifiers such as source, destination address, protocol or next header, DSCP or Traffic Class;

  - Layer-4 identifiers such as source or destination port;

  - service type/destination (i.e. Internet, network service, or other service)

- Actions may be provisioned against the classified traffic; the following are some examples of actions:

  - application of different QoS treatment (could be rate-limit, drop, redirect,.. etc) based on Layer 3 or higher layer (Layer 4-7) classification from devices like deep packet inspection appliances;

  - Service flow redirection on selected types of traffic (i.e. Web portal);

  - Service flow caching on selected types of traffic.

The rationale for applying such policy at the network edge is based on how tightly coupled this layer of the network is with many key systems within the operators network such as RADIUS, DHCP, access technology awareness and ability to implement subscriber awareness.

In many common deployments today, the customer's policies are maintained in RADIUS server or enforced through other provisioned data in co-operation with service activation such as DHCP and bootstrap configuration. In a cable operator network, while much of the heavily lifting of subscriber management is embedded on the CMTS or OLT, the reality is that classification is shared across CMTS and cable-modem (CM) or across OLT and optical network unit (ONU.) The

CM and ONU classification capabilities are not as robust and flexible as the upstream CMTS, OLT and/or assisting edge router. The implications of that are that the CPE may need to be replaced with a device that has the capability to classify on a larger packet header.

An additional point to consider is that the edge network nodes are also often fitted with, or co-located with higher functioning appliances that employ Deep Packet Inspection and distributed caches used to enhance service performance.

2.1. Network/Transport Layer Classification classifiers

Most of the policies described in Section 2 require the use of network and transport layer classification and filtering mechanisms such as classifiers at the network edge. The application of classifiers and other network layer classification functions on selected subscriber flows are often applied by a AAA server, gleaned from configuration information, provisioned from per-CM DOCSIS configuration files generated from the operator OSS, or sent by a policy control function (PCRF, PCMM, etc).

This section will explain why the application of some types of classifiers (like Layer 3 destination based classifiers and - Layer 3 plus Layer 4 - classifiers,) can be deployed in a more simplistic fashion when using a translated form of a stateless IPv4/IPv6 transition technology such as MAP-T [RFC7599].

A key characteristic of MAP-T is the mapping of the IPv4 address of any destination into the IPv6 destination address, by means of IPv4 to IPv6 mapping rules. This mapping means that the subscriber flows are native IPv6 flows within the operators network. The ability to use a standard IPv6 classifier to identify interesting traffic for classification is well aligned with traditional traffic identification capabilities using IPv4 based classifiers. Such classifiers can be easily applied at the access edge as a standard function commonly available on most platforms deployed.

In contrast, a solution utilizing an IP tunnel based transport (MAP-E [RFC7597] or DS-Lite [RFC6333]), effectively hides the payload's IP layer information, making it difficult to identify by means of an IPv6 classifier . The operator in the latter case (tunneled option) would need additional functionality to classify the same subscriber flows which may not be available on the deployed platforms.

The classifier use case is further extended when considering that many traffic classifications are made using transport layer (Layer 4) information. This is common in operator networks that often apply differential traffic treatment to different services that typically

operate using well defined TCP/UDP ports. In the MAP-T deployment case, these ports are available for classification matching using the same standard access edge node capabilities using IPv6 classifiers. In the case where tunneled forms of a solution are used, these higher layer ports are hidden from the network (base IP layer) and special functionality to correctly classify these service flows is required.

The ability to apply classifiers at the access edge node allows the operator to not only use standard IPv6 classifier functionality, but also use same mechanisms (RADIUS interface parameters/system, or DOCSIS configuration classifier parameters) for applying such classifiers. I.e. custom RADIUS interface extensions or custom DOCSIS classifier extensions to deal with the classifier semantics of an IP tunnel based transport are not required.

2.2. Device Configuration (DOCSIS)

Some access technologies, like DOCSIS, require a modem configuration file for network operation. These configuration files often contain access control and classification information that uses IPv4 and/or IPv6 network and transport layer information.

MAP-T allows use of standard IPv6 classifiers within these configuration files permitting the continued use of the well-known service architecture. Translation technologies which use tunneling may require the operator to update how services are managed as information needed to enforce policy is not longer viewable by the Cable Modem or upstream CMTS. The operator in this case may need to build new service capabilities higher up in the network after the network translator to apply the full range of polices for the subscriber base.

2.3. Service Flow management using Deep Packet Inspection

Several Service Providers today use Deep Packet Inspection devices located at the network edge (such as a BNG) in order to inspect the subscriber's traffic for different purposes: profiling the user's behavior, and classifying the traffic based on higher layer information and/or traffic signatures.

Deep packet inspection devices available today in the market and, more importantly, those already deployed in operator's network may not be able to analyze encapsulated traffic, like IPinIP, and to correlate the inner packet's contents to the outer packet's "subscriber" context - this limitation is consistent across multiple vendors. In order to overcome this limitation when using IP tunnel based transports, without resorting to costly network upgrades, dedicated DPI devices need to be applied at a point in the network

where the IP tunnel transport has been stripped and the payload is directly available for native processing. This not only changes the network architecture, but it increases the number of DPI's devices required: one for IPv6 traffic at the access edge, the other at a location where the IPv4 traffic is exposed (typically a separate Location). In addition the operator would need to enforce policies at two architecturally separate places in the network. Furthermore, even with these changes enacted, there remains a critical problem of correlating traffic to a given subscriber: in encapsulation based solutions, the IPv4 address information in the payload is not sufficient to uniquely identify a subscriber given that an IPv4 address will not be unique. As such, additional mechanisms and changes to the accounting infrastructure need to be introduced which when combined with all the previous aspects makes this solution operationally complex.

With MAP-T operators can continue using the current architectural model with DPI devices installed at the access edge; the only requirement would be to have the same device able to recognize specific applications on the native IPv6 transport, which DPI devices based on application signatures are capable of doing. Thus with MAP-T it doesn't matter that an operator might provision the same IPv4 address across multiple subscribers. In addition with MAP-T the access edge would remain the single enforcement point for all user's policies for all traffic. This would allow the operators to continue using a consistent architecture and set of accounting tools for their network.

2.4. Service Flow Redirection Policies (Web-redirection)

Redirecting the user's traffic to web portal is a common practice in Service Provider networks. For example, it is common for operators to inform users about new services, service advisories and/or access to account changes using web-reduction techniques activated on http traffic. In current deployments web-redirection occurs at the Edge node level, where the subscriber's traffic first hits the IP network. The activation/de- activation of redirection policy on selected subscribers may be driven by the AAA/RADIUS through specific RADIUS attributes. In current deployments web-redirection occurs at the Edge node level, where the subscriber's traffic first hits the IP network. The activation/de- activation of redirection policy on selected subscribers may be driven by the AAA/RADIUS through specific RADIUS attributes.

If MAP-T is used the redirection of both IPv6 and IPv4 traffic can be kept at the Edge of the network with the same configuration currently used and by simply translating the Server's address in IPv6 with known mapping rules. In case of tunnel based solution the

redirection of IPv6 and IPv4 cannot occur in a single place, because the redirection of IPv4 traffic must be implemented at or after the v4/v6 gateway responsible for de-encapsulating the traffic. This approach not only would require deploying two separate infrastructures located in different places in order to achieve the redirection for both IPv6 and IPv4 traffic, but also it would not allow continuing using the AAA/RADIUS Server infrastructure in order to enforce the redirect policy at the subscriber's session.

2.5. Service Flow Caching

With the continuing growing of video traffic, especially considering the increase of http video traffic (YouTube like,) it is useful for the Service Providers to be able to cache the video stream at the Edge of the network in order to save bandwidth on upstream links. Using cache devices together with tunnel solutions would introduce similar challenges/issues as the ones described for DPI scenarios, in particular it would require applying caching functionality after the decapsulation point. Obviously this would not eliminate the benefits of the cache. Instead a MAP-T approach would allow caching the subscriber traffic at the edge of the network and gaining the bandwidth savings introduced by the caching. Crucially, any native IPv6 web-caches would be capable of processing IPv6 MAP-T traffic as fully native traffic.

In addition in some deployments today, Web Cache Control Protocol (WCCP) feature is used in order to redirect subscriber's traffic to the cache devices. When a subscriber requests a page from a web server (located in the Internet, in this case), the network node where the WCCP is active, sends the request to a Cache Engine. If the cache engine has a copy of the requested page in storage, the engine sends the user that page. Otherwise, the engine gets the requested page and the objects on that page from the web server, stores a copy of the page and its objects (caches them), and forwards the page and objects to the user. WCCP is another example of web redirect thus, the same considerations described in section Section 2.4 and the benefits introduced by MAP-T also apply here.

3. Technological Considerations

There are additional technological considerations which need to be analyzed by the operator when choosing which transition technology option they would like to deploy. This section describes some of those considerations.

3.1. Encapsulation and Translation Overhead

MAP-E adds an encapsulation tax of 40 bytes, while MAP-T adds a translation tax of 20 bytes (translating from a 20-byte IPv4 header to a 40-byte IPv6 header.) In the downstream direction (from network toward the CPE), with an average packet size of 1000-1100 bytes, the added encapsulation is under 4% in the case of MAP-E. In the case of MAP-T that encapsulation tax drops to about 2%.

In the upstream direction, with an average packet size of ~400 bytes, the effects of the encapsulation tax is more pronounced with an added 10% overhead for MAP-E and 5% additional overhead for MAP-T. As the upstream direction tends to be both (a) more heavily oversubscribed than is the downstream and (b) of lower performance, the greater the header tax the more it upsets the precariously balanced upstream/ downstream network loading models.

3.2. Efficient Utilization of the Access Network

Point-to-Multipoint access networks are common across network operators - DOCSIS (1.0, 1.1, 2.0, 3.0), EPON, 10G-EPON, GPON, XGPON,XGPON2, etc. This network type has been incredibly successful, as attested to by all the variants of point-to-multipoint networks deployed, primarily because of their cost effectiveness.

There are a couple challenges that are introduced by adding a significant amount of encapsulation overhead. These challenges affect MAP-T and MAP-E similarly; the effects from MAP-E are simply more pronounced.

The first challenge is that, commonly, point-to-multipoint networks have limited support for jumbo frames. The second challenge is one that results in reduction in effective capacity on the wire, which yields higher cost.

3.2.1. Jumbo Frame Support in the Access

Some access technologies natively support fragmentation, and as a result, can support "jumbo frames" up to a point. A max size IPv4 packet that fits into the payload of a standard-compliant Ethernet frame is 1500 bytes. In the context of this discussion a "jumbo frame" is any Ethernet frame that has more than 1500 bytes in the Ethernet payload. IEEE Std. 802.3 now specifies a larger frame size of up to 2000 bytes, referred to as an envelope frame, where the envelope frame, quoting from IEEE Std.802.3-2012 "is intended to allow inclusion of additional prefixes and suffixes required by higher layer encapsulation protocols. The encapsulation protocols may use up to 482 octets."

In the network access space, particularly one filled with legacy access products which may be 10 years old (or perhaps older), it is not uncommon to find products that just only support a max 1500 byte Ethernet payload. Some may support up to 1532 byte payload (1550 byte Ethernet frame), some 1582 byte payload (1600 byte Ethernet frame), though there's certainly not a uniform supported frame size past the 1500 byte payload.

Since MTU discovery isn't typically used for IPv4 in operator networks and since MTU discovery for IPv6 is not implemented on the IPv4 host stack requesting the communication, there's no effective way to tell the host stack to reduce the size of its IPv4 frame to accommodate the MAP-T or MAP-E overhead with the MTU frame size limitation of the specific access products. There are tools like Maximum Segment Size rewrite that can be implemented to help address the issue for a TCP payload but UDP payload will continue to be impaired.

Thus MAP-T is preferred as there are more deployed access products that could support a 1534-byte or 1538-byte Ethernet frame than can support a 1554-byte or 1558-byte Ethernet frame, which mandates fewer access product replacements.

3.2.2. Operator Added Packet Overhead and Service Level Agreements

One of the traditional challenges with adding additional packet overhead to a customer frame is that it becomes more challenging to provide customer the last-mile bandwidth in their SLA. This is a very simple overprovisioning problem when the maximum size frame is used, as the overhead in that case is a fixed ~1.5% or ~3% for MAP-T and MAP-E respectively.

However in the case of variable packet sizes, the added overhead from either MAP-T or MAP-E can become very significant - from a worse case of ~31% (MAP-T) and ~63% (MAP-E) to the ~1.5% or ~3%. This means that to provide the customer what they purchased operators will either provision more than the customer SLA to account for the added overhead or abide by the "not guaranteed" bandwidth response.

With the average upstream packet sizes being smaller, the 5% (MAP-T) or 10% (MAP-E) added overhead for the average upstream packet size could find itself in an overprovisioned QoS profile.

Many customers, particularly business customers, are very savvy and have a strong belief that when a network operator offers them an SLA, it's not an SLA at a specific packet size. This can be a significant operational difficulty for network operators, one with a real operational cost.

4. Conclusions

The use cases described in this document have highlighted a clear need for a MAP-T solution based on Service Providers' operational requirements.

This document showed that a MAP-T approach is not a duplication of any other existing IPv4/IPv6 migration mechanisms based on IP tunneling, but actually has capabilities to solve Service Provider's problems.

5. Acknowledgements

The authors would like to thank Victor Kuarsingh for his valuable comments and inputs to this document.

6. IANA Considerations

This document does not require any action from IANA.

7. Security Considerations

This document has no additional security considerations beyond those already identified in section 11 of [RFC7599]

8. Informative References

[RFC6333] Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", RFC 6333, DOI 10.17487/RFC6333, August 2011, <http://www.rfc-editor.org/info/rfc6333>.

[RFC7597] Troan, O., Ed., Dec, W., Li, X., Bao, C., Matsushima, S., Murakami, T., and T. Taylor, Ed., "Mapping of Address and Port with Encapsulation (MAP-E)", RFC 7597, DOI 10.17487/RFC7597, July 2015, <http://www.rfc-editor.org/info/rfc7597>.

[RFC7599] Li, X., Bao, C., Dec, W., Ed., Troan, O., Matsushima, S., and T. Murakami, "Mapping of Address and Port using Translation (MAP-T)", RFC 7599, DOI 10.17487/RFC7599, July 2015, <http://www.rfc-editor.org/info/rfc7599>.

Authors' Addresses

Roberta Maglione (editor)
Cisco Systems
Via Torri Bianche 8
Vimercate 20871
Italy

Email: robmgl@cisco.com

Wojciech Dec
Cisco Systems
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands

Email: wdec@cisco.com

Ida Leung
Rogers Communications
8200 Dixie Road
Brampton, ON L6T 0C1
CANADA

Email: Ida.Leung@rci.rogers.com

Edwin Mallette
Bright House Networks
4145 S. Faulkenburg Road
Riverview, Florida 33578
USA

Email: edwin.mallette@gmail.com