OPSAWG
Internet-Draft
Intended status: Informational
Expires: August 23, 2012

V. Kuarsingh, Ed.
J. Cianfarani
Rogers Communications

February 20, 2012

CGN Deployment with MPLS/VPNs
draft-kuarsingh-lsn-deployment-06

Abstract

This document specifies a framework to integrate a Network Address Translation layer into an operator's network to function as a Carrier Grade NAT (also known as CGN or Large Scale NAT). CGN is a concept also described in [I-D.ietf-behave-lsn-requirements] and describes the model as a dual layer translation model. Although operators may wish to deploy IPv6 to strategically overcome IPv4 exhaustion, near term needs may not be satisfied with an IPv6 deployment alone. This document provides a practical integration model which allows CGN to be integrated into the network meeting the connectivity needs of the customer while being mindful of not disrupting existing services and meeting the technical challenges that CGN brings. The model includes the use of MPLS/VPNs defined in [RFC4364] as a tool to achieve this goal. This document does not intend to defend the merits of CGN.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on August 23, 2012.

Copyright Notice

Copyright (c) 2012 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

Operators are faced with near term IPv4 address exhaustion challenges. Many operators may not have a sufficient amount of IPv4 addresses in the future to satisfy the needs of their growing customer base. This challenge may also be present before or during an active transition to IPv6 somewhat complicating the overall problem space.

To face this challenge, operators may need to deploy CGN (Carrier Grade NAT) as described in [I-D.ietf-behave-lsn-requirements] to help extend the connectivity matrix once IPv4 addresses run out in the network. CGN's addition to the network requires integration in an often running state environment with working IPv4 and/or IPv6 services.

The addition of the CGN introduces an operator controlled and administered translation layer which needs to be added in a manner which does not overly disrupt existing services. This addition may also include interworking in a dual stack environment where the IPv4 path requires translation.

This document shows how MPLS/VPNs as described in [RFC4364] can be used to integrate the CGN infrastructure solving key problems faced by the operator. This model has also been tested and validated in real production network models and allows fluid operation with existing IPv4 and IPv6 services.

2. Motivation

The selection of CGN may be made by an operator based on a number of factors. The overall driver may be the depletion of IPv4 address pools which leaves little to no addresses for IPv4 service growth. IPv6 is considered the strategic answer, but it's applicability and usefulness in many networks is limited by the current access network and consumer home network. These environments often are filled with IPv4-Only equipment which may not be upgradable to IPv6.

The ability to replace IPv4-Only equipment may be out of the control of the operator, and even when it's in the administrative control; it poses both cost and technical challenges as operators build out massive programs for equipment retirement or upgrade. Theses issues leave an operator in a precarious position which may lead to the decision to deploy CGN. Other address IPv4 sharing options do exist which are more architecturally desirable, but the practical and workable approach in many cases is a CGN deployment using NAT444.

If the operator as has chosen to deploy CGN, they should this in a manner as not to negatively impact the existing IPv4 or IPv6 customer base. This will include solving a number of challenges since customers who's connections require translation will have network routing and flow needs which are different from legacy IPv4 connections.

The solution will also need to work in a dual stack environment where other options such as DS-Lite [RFC6333] are not yet viable. Even technologies like 6RD [RFC5969] still require an IPv4 connectivity path to service the customer endpoint. The solution will need to address basic Internet connectivity, on-net service offerings, back office management, billing, policy and security models already in place within the operator's network. CGN will often integrate quite readily with the aforementioned requirements where as other transition mechanism may not due to the requirements to support IPv6 as the base protocol for IPv4 connectivity.

3. CGN Network Deployment Requirements

If a service provider is considering a CGN deployment with a provider NAT44 function, there are a number of basic requirements which are of importance. Preliminary requirements may require the following from the incoming CGN system architecture:

- Support distributed (sparse) and centralized (dense) deployment models;

- Allow co-existence with traditional IPv4 based deployments, which provide global scoped IPs to CPEs;

- Provide a framework for CGN by-pass supporting non-translated flows between endpoints within a provider's network;

- Provide routing framework which allows the segmentation of routing control and forwarding paths between CGN and non-CGN mediated flows;

- Provide flexibility for operators to modify their deployments over time as translation demands change (connections, bandwidth, translation realms/zones and other vectors);

- Flexibility should include integration options for common access technologies such as DSL (BRAS), DOCSIS (CMTS), Mobile (GGSN/PGW/ ASN-GW), and Ethernet access;

- Support deployment modes that allow for IPv4 address overlap within the operator's network (between various translation realms or zones);

- Allow for evolution to future dual-stack and IPv4/IPv6 transition deployment modes;

- Transactional logging and export capabilities to support auxiliary functions including abuse mitigation;

- Support for stateful connection synchronization between translation instances/elements (redundancy);

- Support for CGN Shared Space [I-D.weil-shared-transition-space-request] deployment modes if applicable;

- Allows for the enablement of CGN functionality (if required) while still minimizing costs and customer impact to the best extend possible;

Other requirements may be assessed on a operator-by-operator basis, but those listed above should be considered for any given deployment architecture.

3.1. Centralized versus Distributed Deployment

Centralized deployments of CGN (longer proximity to end user and/or higher densities of subscribers/connections to CGN instances) differ from distributed deployments of CGN (closer proximity to end user and/or lower densities of subscribers/connections to CGN instances). Service providers will likely deploy CGN translation points more centrally during initial phases. Early deployments will likely see light loading on these new systems since legacy IPv4 services will continue to operate with most endpoints using globally unique IPv4 addresses. Exceptional cases which may drive heavy usage in initial stages may include operators who already translate most IPv4 traffic and will migrate to a CGN implementation from legacy firewalls; or a green field deployment which may see quick growth in the number of new IPv4 endpoints which require Internet connectivity.

Over time, most providers will likely need to expand and possibly distribute the translation points as demand for the CGN system increases. The extent of the expansion of the CGN infrastructure will depend on factors such as growth in the number of IPv4 endpoints, status of IPv6 content on the Internet and the overall progress globally to an IPv6-dominate Internet (reducing the demand for IPv4 connectivity).

3.2. CGN and Traditional IPv4 Service Co-existence

Newer CGN serviced endpoints will exist alongside endpoints served by traditional IPv4 global IPs. Providers will need to rationalize these environments since both have distinct forwarding needs. Traditional IPv4 services will likely require (or be best served) direct forwarding towards Internet peering points while CGN mediated flows require access to a translator. CGN and non-CGN mediated flows post two fundamentally different forwarding needs.

The new CGN environments should not negatively impact the existing IPv4 service base by forcing all traffic to translation enabled network points since many flows do not require translation and this would reduce performance of the existing flows. This would also require massive scaling of the CGN which is a cost and efficiency concern as well.

Traffic flow and forwarding efficiency is considered important since networks are under considerable demand to deliver more and more bandwidth without the luxury of needless inefficiencies which can be introduced with CGN.

3.3. CGN By-Pass

The CGN environment is only needed for flows with translation requirements. Many flows which remain in a service provider environment, do not require translation. Such services include operator offered DNS Services, DHCP Services, NTP Services, Web Caching, Mail, News and other services which are local to the operator’s network.

The operator may want to leverage opportunities to offer third parties a platform to also provide services without translation. CGN By-pass can be accomplished in many ways, but a simplistic, deterministic and scalable model is preferred.

3.4. Routing Plane Separation

Many operators will want to engineer traffic separately for CGN flows versus flows which are part of the more traditional IPv4 environment. Many times the routing of these two major flow types differ, therefore route separation may be required.

Routing plane separation also allows the operator to utilize other addressing techniques, which may not be feasible on a single routing plane. Such examples include the use of overlapping private address space [RFC1918] or use of other IPv4 space which may overlap globally within the operator’s network.

3.5. Flexible Deployment Options

Service providers operate complex routing environments and offer a variety of IPv4 based services. Many operator environments utilize distributed peering infrastructures for transit and peering and these may span large geographical areas and regions. A CGN solution should offer the operator an ability to place CGN translation points at various points within their network.

The CGN deployment should also be flexible enough to change over time as demand for translation services increase. In turn, the deployment will need to then adapt as translation demand decreases caused by the transition of flows to IPv6. Translation points should be able to be placed and moved with as little re-engineering effort as possible minimizing the risks to the customer base.

Depending on hardware capabilities, security practices and IPv4 address availability, the translation environments my need to be segmented and/or scaled over time to meet organic IPv4 demand growth. Operators will want to seek deployment models which are conducive to meeting these goals as well.

3.6. IPv4 Overlap Space

IP address overlap for CGN translation realms may be required if insufficient IPv4 addresses are available within the service provider environment to assign internally unique IPs to the CGN customer base . The CGN deployment should provide mechanisms to manage IPv4 overlap if required.

3.7. Transactional Logging for LSN Systems

CGNs may require transactional logging since the source IP and related transport protocol information is not easily visible to external hosts and system.

If needed, the CGN systems should be able to generate logs which identify 'internal' host parameters (i.e. IP/Port) and associated them to external translated parameters imposed by the translator. The logged information should be stored on the CGN hardware and/or exported to an external system for processing. Operators may need to keep track of this information (securely) to meet regulatory and/or legal obligations. Further information can be found in [I-D.ietf-behave-lsn-requirements] with respect to CGN logging requirements (Logging Section).

3.8. Additional CGN Requirements

The CGN platform will also need to meet the needs of additional requirements such as Bulk Port Allocation and other CGN device specific functions. These additional requirements are captured within [I-D.ietf-behave-lsn-requirements].

4. MPLS/VPN based CGN Framework

The MPLS/VPN [RFC4364] framework for CGN segregates the 'pre-translated' realms within the service provider space into layer-3 MPLS/VPNs. The operator can deploy a single realm for all CGN based flows, or can deploy multiple realms based on translation demand and other factors such as geographical proximity. A realm in this model refers to a 'VPN' which shares a unique RD/RT combination, routing plane and forwarding behaviours.

The MPLS/VPN infrastructure provides control plane and forwarding separation for the traditional IPv4 service environment and CGN environment(s). The separation allows for routing information (such as default routes) to be propagated separately for CGN and non-CGN based customer flows. Traffic can be efficiently routed to the Internet for normal flows, and routed directly to translators for CGN mediated flows. Although many operators may run a "default-route-free" core, IPv4 flows which require translation must obviously be routed first to a translator, so a default route is acceptable for the pre-translated realms.

The physical location of the VRF Termination point for a MPLS/VPN enabled CGN can vary and be located anywhere within the operator's network. This model fully virtualizes the translation service from the base IPv4 forwarding environment which will likely carrying Internet bound traffic. The base IPv4 environment can continue to service traditional IPv4 customer flows plus post translated CGN flows.

Figure 1 provides a view of the basic model. The Access node provides CPE access to either the CGN VRF or the Global Routing Table, depending on whether the customer receives a private or public IP. Translator mediated traffic follows an MPLS LSP which can be setup dynamically and can span one hop, or many hops (with no need for complex routing policies). Traffic is then forwarded to the translator (shown below) which can be an external appliance or integrated into the VRF Termination (Provider Edge) router. Once traffic is translated, it is forwarded to the global routing table for general Internet forwarding. The Global Routing table can also be a separate VRF (Internet Access VPN/VRF) should the provider

choose to implement their Internet based services in that fashion. The translation services are effectively overlaid onto the network, but are maintained within a separate forwarding and control plane.

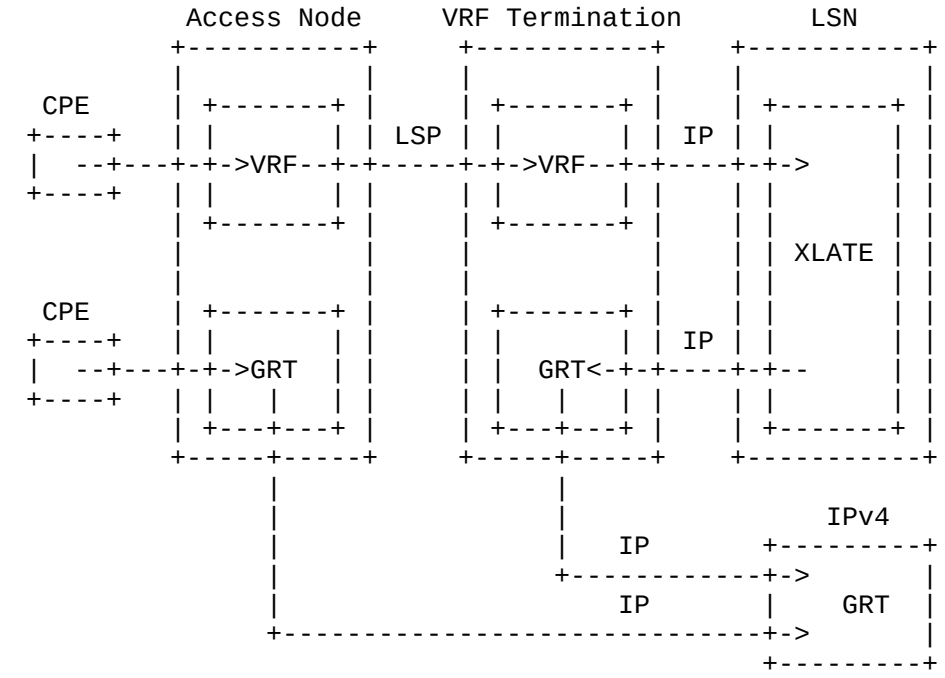

Figure 1: Basic MPLS/VPN CGN Model

If more then one VRF (translation realm) is used within the operator's network, each VPN instance can manage CGN flows independently for the respective realm. Various redundancy models can be used within this architecture to support failover from one physical CGN hardware instance to another. If state information needs to be passed or maintained between hardware instances, the vendor would need to enable this feature in a suitable manner.

4.1. Service Separation

The MPLS/VPN CGN framework supports route separation. The traditional IPv4 flows can be separated at the access node (Initial Layer 3 service point) from those which require translation. This type of service separation is possible on common technologies used for Internet access within many operator networks. Service separation can be accomplished on common access technology including

those used for DOCSIS (CMTS), Ethernet Access, DSL (BRAS), and Mobile Access (GGSN/ASN-GW) architectures.

4.2. Internal Service Delivery

Internal services can be delivered directly to the privately addressed endpoint within the CGN domain without translation. This can be accomplished using direct route exchange (import/export) between the CGN VRFs and the Services VRFs. The previous statement assumes the provider puts key services into a VRF for simple route exchange. This model allows the provider to maintain separate forwarding rules for translated flows, which require a pass through the translator to reach external network entities, versus those flows which need to access internal services. This operational detail can be advantageous for a number of reasons.

First, the provider can reduce the load on the translator since internal services do not need to be factored into the scaling of the CGN hardware. Secondly, more direct forwarding paths can be maintained providing better network efficiency. Thirdly, geographic locations of the translators and the services infrastructure can be deployed in a location in an independent manner. Additionally, the operator can allow CGN subject endpoints to be accessible via an untranslated path reducing the complexities of provider initiated management flows. This last point is of key interest since NAT removes transparency to the end device in normal cases.

Figure 2 below shows how internal services are provided untranslated since flows are sent directly from the access node to the services node/VRF via an MPLS LSP. This traffic is not forwarded to the CGN translator and therefore is not subject to problematic behaviours related to NAT. The services VRF contains routing information which can be "imported" into the access node VRF and the CGN VRF routing information can be "imported" into the Services VRF.

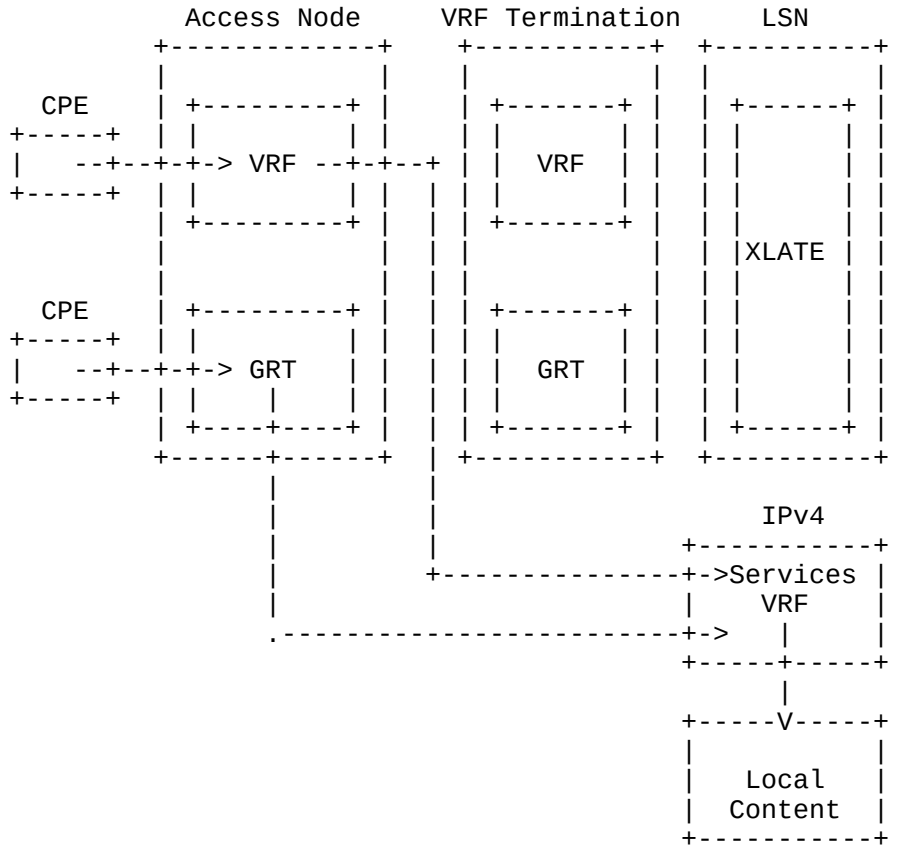

Figure 2: Internal Services and CGN By-Pass

This demonstrates the ability to offer CGN By-Pass in a simple and deterministic manner without the need of policy based routing or traffic engineering.

4.2.1. Dual Stack Operation

The MPLS/VPN CGN model can also be used in conjunction with IPv4/IPv6 dual stack service modes. Since many providers will use CGNs on an interim basis while IPv6 matures within the global Internet or due to technical constraints, a dual stack option is of strategic importance. Operators can offer this dual stack service for both traditional IPv4 (global IP) endpoints and CGN mediated endpoints.

Operators can separate the IP flows for IPv4 and IPv6 traffic, or use other routing techniques to move IPv6 based flows towards the GRT

(Global Routing Table or Instance) while allowing IPv4 flows to remain within the IPv4 CGN VRF for translator services.

The Figure 3 below shows how IPv4 translation services can be provided alongside IPv6 based services. The model shown allows the provider to enable CGN to manage IPv4 flows (translated) and IPv6 flows are routed without translation efficiently towards the Internet. Once again, forwarding of flows to the translator does not impact IPv6 flows which do not require this service.

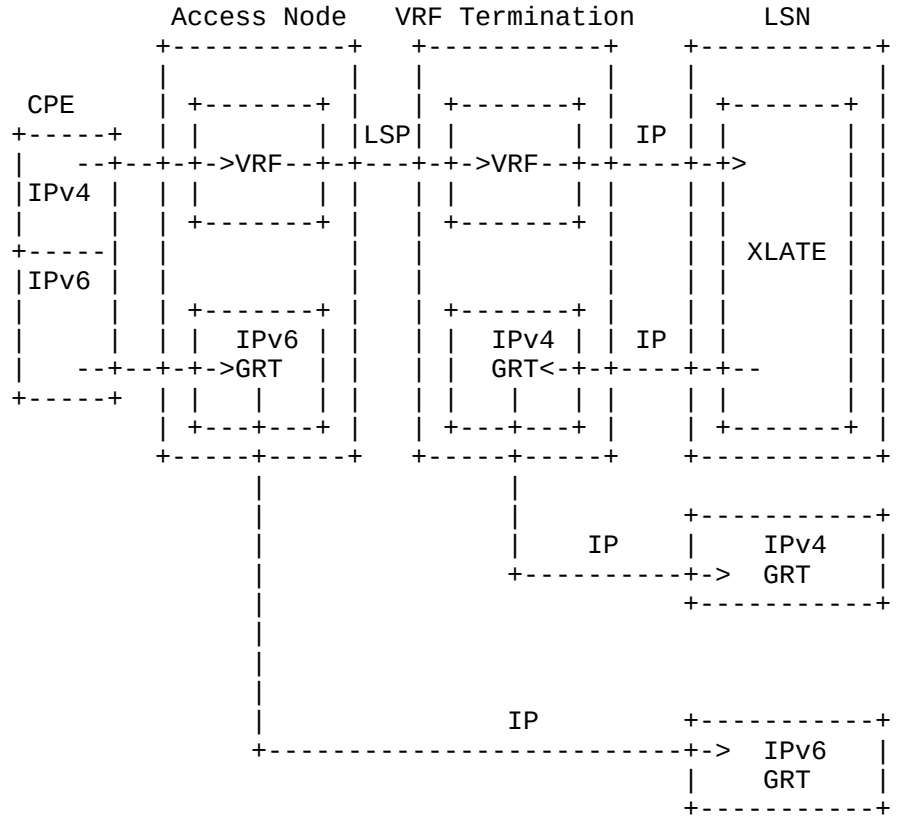

Figure 3: CGN with IPv6 Dual Stack Operation

4.3. Deployment Flexibility

The CGN translator services can be moved, separated or segmented (new translation realms) without the need to change the overall translation design. Since dynamic LSPs are used to forward traffic

from the access nodes to the translation points, the physical location of the VRF termination points can vary and be changed easily.

This type of flexibility allows the service provider to initially deploy more centralized translation services based on relatively low loading factors, and distribute the translation points over time to improve network traffic efficiencies and support higher translation load.

Although traffic engineered paths are not required within the MPLS/VPN deployment model, nothing precludes an operator from using technologies like MPLS with Traffic Engineering [RFC3031]. Additional routing mechanisms can be used as desired by the provider and can be seen as independent. There is no specific need to diversify the existing infrastructure in most cases.

4.4. Comparison of MPLS/VPN Option versus other CGN Attachment Options

Other integration architecture options exist which can attach CGN based service flows to a translator instance. Alternate options which can be used to attach such services include:

- IEEE 802.1Q for direct attachment to a next hop translator;

- Policy Based Routing (Static) to direct translation bound traffic to a network based translator;

- Traffic Engineering or;

- Multiple Routing Topologies

4.4.1. IEEE 802.1Q

IEEE 802.1Q can be used to associate separated traffic from the access node to the next hop router’s CGN instance. This technology option may limit the CGN placement to the next hop router unless a second technology option is paired with it to extend connectivity deeper in the network.

This option is most effective if CGN instances are placed directly upstream of the access node. Distributed CGN instance placement is not likely an initial stage of the CGN deployment due to cost and demand factors.

4.4.2. Policy Based Routing

Policy Based Routing (PBR) provides another option to direct CGN mediated flows to a translator. PBR options, although possible, are difficult to maintain (static policy) and must be configured throughout the network with considerable maintenance overhead.

More centralized deployments may be difficult or too onerous to deploy using Policy Based Routing methods. Policy Based Routing would not achieve route separation (unless used with others options), and may add complexities to the providers' routing environment.

4.4.3. Traffic Engineering

Traffic Engineering can also be used to direct traffic from an access node towards a translator. Traffic Engineering, like MPLS-TE, may be difficult to setup and maintain. Traffic Engineering provides additional benefits if used with MPLS by adding potentials for faster path re-convergence. Traffic Engineering paths would need to be updated and redefined overtime as CGN translation points are augmented or moved.

4.4.4. Multiple Routing Topologies

Multiple routing topologies can be used to direct CGN based flows to translators. This option would achieve the same basic goal as the MPLS/VPN option but with additional implementation overhead and platform configuration complexity. Since operator based translation is expected to have an unknown lifecycle, and may see various degrees of demand (dependant on operator IPv4 Global space availability and shift of traffic to IPv6), it may be too large of an undertaking for the provider to enabled this as their primary option for CGN.

5. Experiences

6. Basic Integration and Requirements Support

The MPLS/VPN CGN environment has been successfully integrated into real network environments utilizing existing network service delivery mechanisms. It solves many issues related to provider based translation environments, while still subject to problematic behaviours inherent within NAT.

Key issues which are solved or managed with the MPLS/VPN option include:

- Centralized and Distributed Deployment model support

- Routing Plane Separation for CGN flows versus traditional IPv4 flows

- Flexible Translation Point Design (can relocate translators and split translation zones easily)

- Low maintenance overhead (dynamic routing environment with little maintenance of separate routing infrastructure other then management of MPLS/VPNs)

- CGN By-pass options (for internal and third party services which exist within the provider domain)

- IPv4 Translation Realm overlap support (can reuse IP addresses between zones with some impact to extranet service model)

- Simple failover techniques can be implemented with redundant translators, such as using a second default route

7. Performance

The MPLS/VPN CGN model was observed to support basic functions which are typically used by customers within an operator environment. Examples of successful operation include:

- Traditional Web (HTTP) Surfing (client initiated)

- Internet Video Streaming

- HTTP Based Client Connections

- High Connection Count sites (i.e. Google Maps)

- Email Transaction Support (POP, IMAP, SMTP)

- Instant Messaging Support (Online Status, File transfers, text chat)

- ICMP Operation (client initiated Echo, Traceroute)

- Peer to Peer application support (download)

- DNS (based on services extranet option, but was problematic when passed through a translator)

CGNs are still subject to problematic connectivity even within the MPLS/VPN technology approach. Problems which arise, or are not inherently addressed in this model include:

- Inward services from the Internet to the CPE
- Web session tracking
- Restricting usage and/or access based on source IP
- Abuse mitigation (masquerade of potential offenders)
- Increased network or server IDS false positives
- Increased customer risk for session hijacking
- Exceeding firewall TCP/UDP limits
- Customer identification (external site)
- Poor source based load balancing
- Customer usage tracking / Ad insertion
- Other applications or operations may be negatively impacted

8. IANA Considerations

There are not specific IANA considerations known at this time with the architecture described herein. Should a provide choose to use non-assigned IP address space within their translation realms, then considerations may apply.

9. Security Considerations

The same security considerations would typically exist for CGN deployments when compared with traditional IPv4 based services. With the MPLS/VPN model, the operator would want to consider security issues related to offering IP services over MPLS.

If a provider plans to operate the pre-translation realm (CPE towards translator IPv4 zone) as a non-public like network, then additional security measures may be needed to secure this environment. It is however the position in this document that CGN realms are public domains which utilize non-Internet routable IP addresses for endpoint addressing.

10. Conclusions

The MPLS/VPN delivery method for a CGN deployment is an effective and scalable way to deliver mass translation services. The architecture avoids the complex requirements of traffic engineering and policy based routing when combining these new service flows to existing IPv4 operation. This is advantageous since the NAT44/CGN environments should be introduced with as little impact as possible and these environments are expected to change over time.

The MPLS/VPN based CGN architecture solves many of this issues related to deploying this technology in existing operator networks.

11. Acknowledgements

Thanks to the following people for their participating in integrating and testing the CGN environment: Chris Metz, Syd Alam, Richard Lawson, John E Spence.

Additional thanks for the following people for the guidance on IPv6 transition considerations: John Jason Brzozowski, Chris Donley, Jason Weil, Lee Howard, Jean-Francois Tremblay

12. References

12.1. Normative References

[I-D.ietf-behave-lsn-requirements]
Perreault, S., Yamagata, I., Miyakawa, S., Nakagawa, A., and H. Ashida, "Common requirements for Carrier Grade NATs (CGNs)", draft-ietf-behave-lsn-requirements-05 (work in progress), November 2011.

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

12.2. Informative References

[I-D.weil-shared-transition-space-request]
Weil, J., Kuarsingh, V., Donley, C., Liljenstolpe, C., and M. Azinger, "IANA Reserved IPv4 Prefix for Shared Address Space", draft-weil-shared-transition-space-request-15 (work in progress), February 2012.

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets",

BCP 5, RFC 1918, February 1996.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[RFC5969] Townsley, W. and O. Troan, "IPv6 Rapid Deployment on IPv4 Infrastructures (6rd) -- Protocol Specification", RFC 5969, August 2010.

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

Authors' Addresses

Victor Kuarsingh (editor)
Rogers Communications
8200 Dixie Road
Brampton, Ontario L6T 0C1
Canada

Email: victor.kuarsingh@gmail.com
URI: http://www.rogers.com

John Cianfarani
Rogers Communications
8200 Dixie Road
Brampton, Ontario L6T 0C1
Canada

Email: john.cianfarani@rci.rogers.com
URI: http://www.rogers.com