Carrier-Grade NAT (CGN) Deployment with BGP/MPLS IP VPNs

Abstract

This document specifies a framework to integrate a Network Address Translation (NAT) layer into an operator’s network to function as a Carrier-Grade NAT (also known as CGN or Large-Scale NAT). The CGN infrastructure will often form a NAT444 environment as the subscriber home network will likely also maintain a subscriber-side NAT function. Exhaustion of the IPv4 address pool is a major driver compelling some operators to implement CGN. 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 that allows the CGN platform to be integrated into the network, meeting the connectivity needs of the subscriber while being mindful of not disrupting existing services and meeting the technical challenges that CGN brings. The model included in this document utilizes BGP/MPLS IP VPNs, which allow for virtual routing separation, helping ease the CGN’s impact on the network. This document does not intend to defend the merits of CGN.

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/rfc7289.
# Table of Contents

1. Introduction ................................................... 4  
   1.1. Acronyms and Terms ........................................ 4  
2. Existing Network Considerations ............................. 5  
3. CGN Network Deployment Requirements ....................... 5  
   3.1. Centralized versus Distributed Deployment ............... 6  
   3.2. CGN and Traditional IPv4 Service Coexistence ........... 7  
   3.3. CGN Bypass ............................................... 7  
   3.4. Routing Plane Separation ................................ 8  
   3.5. Flexible Deployment Options .............................. 8  
   3.6. IPv4 Overlap Space ....................................... 9  
   3.7. Transactional Logging for CGN Systems .................. 9  
   3.8. Base CGN Requirements ................................... 9  
4. BGP/MPLS IP VPN-Based CGN Framework ....................... 9  
   4.1. Service Separation ....................................... 11  
   4.2. Internal Service Delivery ................................ 12  
      4.2.1. Dual-Stack Operation .............................. 14  
   4.3. Deployment Flexibility ................................... 16  
   4.4. Comparison of BGP/MPLS IP VPN Option versus Other  
      CGN Attachment Options ................................. 16  
      4.4.1. Policy-Based Routing .............................. 16  
      4.4.2. Traffic Engineering ............................... 17  
      4.4.3. Multiple Routing Topologies ....................... 17  
   4.5. Multicast Considerations ................................ 17  
5. Experiences .................................................... 17  
   5.1. Basic Integration and Requirements Support ............. 17  
   5.2. Performance ............................................. 18  
6. Security Considerations ....................................... 18  
7. BGP/MPLS IP VPN CGN Framework Discussion .................... 18  
8. Acknowledgements ............................................. 19  
9. References .................................................... 19  
   9.1. Normative Reference .................................... 19  
   9.2. Informative References .................................. 19
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 subscriber 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 [RFC6888] to help extend the connectivity matrix once IPv4 address caches run out on the local operator. CGN deployments will most often be added into operator networks that already have active IPv4 and/or IPv6 services.

The addition of the CGN introduces a translation layer that is controlled and administered by an operator and that should be added in a manner that minimizes disruption to existing services. The CGN system addition may also include interworking in a dual-stack environment where the IPv4 path requires translation.

This document shows how BGP/MPLS IP VPNs as described in [RFC4364] can be used to integrate the CGN infrastructure solving key integration challenges 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.

1.1. Acronyms and Terms

Acronyms and terms used in this document are defined in the list below.

CGN - Carrier-Grade NAT

DOCSIS - Data Over Cable Service Interface Specification

CMTS - Cable Modem Termination System

DSL - Digital Subscriber Line

BRAS - Broadband Remote Access Server

GGSN - Gateway GPRS Support Node

GPRS - General Packet Radio Service

ASN-GW - Access Service Network Gateway
GRT - Global Routing Table

Internal Realm - Addressing and/or network zone between the Customer Premises Equipment (CPE) and CGN as specified in [RFC6888]

External Realm - Public-side network zone and addressing on the Internet-facing side of the CGN as specified in [RFC6888]

2. Existing Network Considerations

The selection of CGN may be made by an operator based on a number of factors. The overall driver to use CGN may be the depletion of IPv4 address pools, which leaves little to no addresses for a growing IPv4 service or connection demand growth. IPv6 is considered the strategic answer for IPv4 address depletion; however, the operator may independently decide that CGN is needed to supplement IPv6 and address their particular IPv4 service deployment needs.

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

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 architectural requirements that are of importance. Preliminary architectural requirements may require all or some of those captured in the list below. Each of the architectural requirement items listed is expanded upon in the following subsections. It should be noted that architectural CGN requirements are additive to base CGN functional requirements contained in [RFC6888]. The assessed architectural requirements for deployment are:

- Support distributed (sparse) and centralized (dense) deployment models. See Section 3.1

- Allow coexistence with traditional IPv4-based deployments, which provide global-scoped IPv4 addresses to CPEs. See Section 3.2.

- Provide a framework for CGN bypass supporting non-translated flows between endpoints within a provider’s network. See Section 3.3.
- Provide a routing framework that allows the segmentation of routing control and forwarding paths between CGN-mediated and non-CGN-mediated flows. See Section 3.4.

- Provide flexibility for operators to modify their deployments over time as translation demands change (connections, bandwidth, translation realms/zones, and other vectors). See Section 3.5.

- Flexibility should include integration options for common access technologies such as DSL (BRAS), DOCSIS (CMTS), Mobile (GGSN/PGW/ASN-GW), and direct Ethernet. See Section 3.5.

- Support deployment modes that allow for IPv4 address overlap within the operator’s network (between various translation realms or zones). See Section 3.6.

- Allow for evolution to future dual-stack and IPv4/IPv6 transition deployment modes. See Section 3.5.

- Transactional logging and export capabilities to support auxiliary functions, including abuse mitigation. See Section 3.7.

- Support for stateful connection synchronization between translation instances/elements (redundancy). See Section 3.8.

- Support for CGN Shared Address Space [RFC6598] deployment modes if applicable. See Section 3.6.

- Allow for the enablement of CGN functionality (if required) while still minimizing costs and subscriber impact to the best extend possible. See Section 3.8.

Other requirements may be assessed on an operator-by-operator basis, but those listed above may 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 may likely deploy CGN translation points more centrally during initial phases if the early system demand is low. Early deployments may 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 that may drive heavy usage in initial stages may include operators that
already translate a significant portion of their IPv4 traffic, operators that may transition to a CGN implementation from legacy translation mechanisms (i.e., traditional firewalls), or operators that build a greenfield deployment that may see quick growth in the number of new IPv4 endpoints that require Internet connectivity.

Over time, some providers may need to expand and possibly distribute the translation points if 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). The overall demand for CGN resources will probably follow a bell-like curve with a growth, peak, and decline period.

3.2. CGN and Traditional IPv4 Service Coexistence

Newer CGN-serviced endpoints will exist alongside endpoints served by traditional IPv4 globally routed addresses. Operators will need to rationalize these environments since both have distinct forwarding needs. Traditional IPv4 services will likely require (or be best served by) direct forwarding toward Internet peering points while CGN-mediated flows require access to a translator. CGN-mediated and non-CGN-mediated flows pose 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.

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

3.3. CGN Bypass

The CGN environment is only needed for flows with translation requirements. Many flows that remain within the operator’s network do not require translation. Such services include operator-offered DNS services, DHCP services, NTP services, web caching, email, news, and other services that 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 bypass 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 that 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], Shared Address Space [RFC6598], or other IPv4 space that 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 external routing infrastructures for transit and peering, and these may span large geographical areas. A CGN solution should offer operators the 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 or change as noted in [RFC6264]. In turn, the deployment will need to then adapt as translation demand decreases due to 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 subscriber base.

Depending on hardware capabilities, security practices, and IPv4 address availability, the translation environments may need to be segmented and/or scaled over time to meet organic IPv4 demand growth. Operators may also want to choose models that support transition to other translation environments such as Dual-Stack Lite (DS-Lite) [RFC6333] and/or Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers (NAT64) [RFC6146]. Operators will want to seek deployment models that are conducive to meeting these goals as well.
3.6. IPv4 Overlap Space

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

3.7. Transactional Logging for CGN Systems

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

If needed, CGN systems should be able to generate logs that identify internal-realm host parameters (i.e., IP/Port) and associate them to external-realm parameters imposed by the translator. The logged information should be stored on the CGN hardware and/or exported to another system for processing. The operator may choose to also enable mechanisms to help reduce logging, such as block allocation of UDP and TCP ports or deterministic translation options, e.g., [CGN-DEPLOYMENTS].

Operators may be legally obligated to keep track of translation information. The operator may need to utilize their standard practices in handling sensitive customer data when storing and/or transporting such data. Further information regarding CGN logging requirements can be found in Section 4 of [RFC6888].

3.8. Base CGN Requirements

Whereas the requirements above represent assessed architectural requirements, the CGN platform will also need to meet the base CGN requirements of a CGN function. Base requirements include functions such as Bulk Port Allocation and other CGN device-specific functions. These base CGN platform requirements are captured in [RFC6888].

4. BGP/MPLS IP VPN-Based CGN Framework

The BGP/MPLS IP VPN [RFC4364] framework for CGN segregates the internal realms within the service-provider space into Layer 3 MPLS-based 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 Route Distinguisher / Route Target (RD/RT) combination, routing plane, and forwarding behaviors.
The BGP/MPLS IP 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-based and non-CGN-based subscriber 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 that require translation must obviously be routed first to a translator, so a default route is acceptable for the internal realms.

The physical location of the Virtual Routing and Forwarding (VRF) termination point for a BGP/MPLS IP 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 that will likely be carrying Internet-bound traffic. The base IPv4 environment can continue to service traditional IPv4 subscriber 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 (GRT), depending on whether the subscriber receives a private or public IP. Translator-mediated traffic follows an MPLS Label Switched Path (LSP) that can be set up dynamically and can span one hop or many hops (with no need for complex routing policies). Traffic is then forwarded to the translator, which can be an external appliance or integrated into the VRF Termination (Provider Edge) router. Once traffic is translated, it is forwarded to the GRT for general Internet forwarding. The GRT 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.
If more than one VRF (translation realm) is used within the operator's network, each VPN instance can manage CGN flows independently for the respective realm. The described architecture does not prescribe a single redundancy model that ensures network availability as a result of CGN failure. Deployments are able to select a redundancy model that fits best with their network design. 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 that 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 in one of two methods. The first method is the use of "route leaking", a method of exchanging routes between the CGN VRF and GRT; this method may also include reducing the overall number of VRFs in the system and exposing services in the GRT. The second method, which is described in detail within this section, is the use of a Services VRF. The second model is a more traditional extranet services model but requires more system resources to implement.

Using direct route exchange (import/export) between the CGN VRFs and the Services VRFs creates reachability using the aforementioned extranet model available in the BGP/MPLS IP VPN structure. 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 that need to access internal services. This operational detail can be advantageous for a number of reasons, such as service-access policies and endpoint identification. 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 (which may be quite large). Second, more direct forwarding paths can be maintained to provide better network efficiency. Third, geographic locations of the translators and the services infrastructure can be deployed in locations 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 behaviors related to NAT. The Services VRF contains routing information that can be "imported" into the access node VRF, and the CGN VRF routing information can be "imported" into the Services VRF.
An extension to the services delivery LSP is the ability to also provide direct subscriber-to-subscriber traffic flows between CGN zones. Each zone or realm may be fitted with separate CGN resources, but the subtending subscribers don’t necessarily need to be mediated (translated) by the CGN translators. This option, as shown in Figure 3, is easy to implement and can only be enabled if no IPv4 address overlap is used between communicating CGN zones.
The inherent capabilities of the BGP/MPLS IP VPN model demonstrates the ability to offer CGN bypass in a standard and deterministic manner without the need of policy-based routing or traffic engineering.

4.2.1. Dual-Stack Operation

The BGP/MPLS IP 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 they can use other routing techniques to move IPv6-based flows toward the GRT (Global Routing Table) while allowing IPv4 flows to remain within the IPv4 CGN VRF for translator services.

Figure 4 shows how IPv4 translation services can be provided alongside IPv6-based services. This model allows the provider to enable CGN to manage IPv4 flows (translated), and IPv6 flows are routed without translation efficiently toward the Internet. Once again, forwarding of flows to the translator does not impact IPv6 flows, which do not require this service.

Figure 4: 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 BGP/MPLS IP VPN Option versus Other CGN Attachment Options

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

- policy-based routing (static) to direct translation-bound traffic to a network-based translator;
- traffic engineering; and
- multiple routing topologies.

4.4.1. 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 (due to 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.
4.4.2. Traffic Engineering

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

4.4.3. 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 (dependent 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.

4.5. Multicast Considerations

When deploying BGP/MPLS IP VPNs as a service method for user-plane traffic to access CGN, one needs to be cognizant of current or future IP multicast requirements. User-plane IP Multicast that may originate outside of the VRF requires specific consideration. Adding the requirement for user plane IP multicast can potentially cause additional complexity related to importing and exporting the IP multicast routes in addition to suboptimal scaling and bandwidth utilization.

It is recommended to reference best practice and designs from [RFC6037], [RFC6513], and [RFC5332].

5. Experiences

5.1. 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 behaviors inherent within NAT.
The key issues that are solved or managed with the MPLS/VPN option include:

- Support for the centralized and distributed deployment model
- 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 than management of MPLS/VPNs)
- CGN bypass options (for internal and third-party services that 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

5.2. Performance

The MPLS/VPN CGN model was observed to support basic functions that are typically used by subscribers within an operator environment. A full review of the observed impacts related to CGN (NAT444) are covered in [RFC7021].

6. Security Considerations

An operator implementing CGN using BGP/MPLS IP VPNs should refer to Section 7 of [RFC6888] for security considerations related to CGN deployments. The operator should continue to employ the standard security methods in place for their standard MPLS deployment and can also refer to the Security Considerations section in [RFC4364], which discusses both control-plane and data-plane security.

7. BGP/MPLS IP VPN CGN Framework Discussion

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 NAT444/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 the issues related to deploying this technology in existing operator networks.

8. Acknowledgements

Thanks to the following people for their comments and feedback: Dan Wing, Chris Metz, Chris Donley, Tina TSOU, Christopher Liljenstolpe, and Tom Taylor.

Thanks to the following people for their participation in integrating and testing the CGN environment and for their guidance in regard to IPv6 transition: Syd Alam, Richard Lawson, John E. Spence, John Jason Brzozowski, Chris Donley, Jason Weil, Lee Howard, and Jean-Francois Tremblay.

9. References

9.1. Normative Reference


9.2. Informative References


Authors' Addresses

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

EMail: victor@jvknet.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