Internet Area Working Group
Internet Draft
Intended status: Informational
Expires: January 2011

V.Kuarsingh
J.Cianfarani
Rogers Communications

July 4, 2010

# NAT44/LSN Deployment Options and Experiences
draft-kuarsingh-lsn-deployment-00.txt

Abstract

This document specifies NAT44 [RFC3022] with Large Scale NAT [draft-nishitani-cgn-04] integration options along with production model experience. The NAT44/LSN implementation is associated with the NAT444 [draft-shirasaki-nat444-01] model. Service Providers are preparing for IPv4 address depletion by enabling IPv6 and/or extending connectivity for IPv4 to support legacy Internet end points. This document provides practical integration options for Large Scale NAT systems which enable provider NAT44 and is applicable primarily to service providers. The document does not intend to argue the merits of NAT444 versus other IPv4 run out technology options

Table of Contents

## 1. Introduction

The majority of service providers are planning on or implementing technologies to mitigate the impact of IPv4 address depletion. Many service providers are planning on implementing a service provider NAT44 infrastructure independently or alongside IPv6.

NAT444 [draft-shirasaki-nat444-01] is a technology model which deals with IPv4 address depletion by providing a framework which implements a service provider controlled and administrated NAT44 translation infrastructure.

The NAT444 model can be implemented in an incremental fashion to complement the existing IPv4 legacy services. NAT44/LSN can also be implemented as part of an IPv6 (Dual Stack) offering allowing for connectivity to the IPv6 Internet.

This document shows how MPLS/VPNs [RFC4364] can be used to provide NAT44/LSN services by solving key problems faced by service providers. Although other integration models may exist, the framework described herein has shown to be successful when tested and characterized in real production models.

## 2. Motivation

Providers may choose to select NAT44/LSN as an initial stage to deal with IPv4 address depletion due to the limited IPv6 technology support in existing end user equipment. Specifically, many devices within the customer premise may not initially support IPv6 or may never support IPv6. Cost, business constraints and other technical reasons may also motivate providers to initially offer NAT44/LSN services.

The selection of NAT44/LSN to connect IPv4 endpoints to the IPv4 Internet does not preclude service providers from making IPv6 available to customers. Dual Stack implementations are feasible utilizing both native IPv4 and NAT44/LSN based IPv4 connectivity. IPv6 connectivity can co-exist with NAT44/LSN, including 6RD [RFC5569] deployments, which may utilize private space for IPv4.

Initially providing a NAT44/LSN deployment in a dual stack connection model versus other options such as DS-Lite [draft-ietf-softwires-dual-stack-lite-04] allows providers to ease the burden on networks as services, back office, billing, policy and security systems are migrated to support IPv6 or IPv4/IPv6 tunneling options. NAT444 based

connectivity utilizing Large Scale NAT systems can be evolved in the future once IPv6 has matured within the service provider network and customer environments.

## 3. NAT44/LSN Deployment Requirements

If a service provider is considering a NAT44 deployment with Large Scale NAT (NAT44/LSN), there are a number of basic requirements which are of importance. Preliminary requirements may include the NAT44/LSN system capability to:

- Support distributed (sparse) and centralized (dense) Deployment Modes;
- Allow co-existence with traditional IPv4 based deployments, which provide global scoped IPs to CPEs;
- Provide a framework for NAT 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 NAT44/LSN and non-NAT44/LSN mediated flows;
- Provide flexibility for providers 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 NAT44/LSN provider network (between various translation realms);
- 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 ISP Shared Space [draft-shirasaki-isp-shared-addr-04] deployment modes if applicable;

- o Allows for the enablement of NAT44/LSN functionality (if required) while still minimizing costs and customer impact to the best extend possible;

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

## 3.1. Centralized vs Distributed Modes

Centralized deployments of LSN (longer proximity to end user and/or higher densities of subscribers/connections to LSN instances) differ from distributed deployments of LSN (closer proximity to end user and/or lower densities of subscribers/connections to LSN instances). Service providers will likely deploy LSN 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 providers who already translate most IPv4 traffic and will migrate to a NAT44/LSN 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 NAT44/LSN system increases. The extent of the expansion of the NAT44/LSN 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 world.

## 3.2. NAT44/LSN and Traditional IPv4 Service Co-existence

Newer NAT44/LSN 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) with direct forwarding towards Internet peering points while NAT444 mediated flows require access to a translator.

The new NAT44/LSN 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.

Traffic flow and forwarding efficiency is considered important since networks are under considerable demand to delivery more and more bandwidth without the luxury of needless inefficiencies which can be introduced with NAT44/LSN.

### 3.3. NAT444 By-Pass

The NAT44/LSN 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 provider offered DNS Caching, DHCP Services, NTP Services, Web Caching and other services local to the provider network.

The provider may want to leverage opportunities to offer third parties a platform to also provide end device services without translation. NAT44/LSN By-pass can be accomplished in many ways, but a simplistic, deterministic and scalable model is preferred.

### 3.4. Routing Plane Separation

Many providers will want to engineer traffic separately for NAT44/LSN flows versus flows which are part of the more traditional IPv4 environment. Many times the routing of these two major flow types differ, therefore routing separation may be required.

Routing plane separation also allows the provider 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 (i.e. ISP Shared Space, BOGON Space or others).

### 3.5. Flexible Deployment Options

Service providers operate complex routing environments and offer a variety of IPv4 based services. Many provider environments utilize distributed peering infrastructures for transit and peering and these may span large geographical regions. A NAT44/LSN solution should offer the provider the ability to place LSN translation points at various points within their network.

The NAT44/LSN 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 migration 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 grown over time to meet organic IPv4 demand growth. Providers will want to seek deployment models which are conducive to meeting these goals as well.

### 3.6. IPv4 Overlap Space

IP address overlap for NAT44/LSN translation realms may be required if insufficient IPv4 addresses are available within the service provider environment. The NAT44/LSN deployment should provide mechanisms to enable such an option if required.

### 3.7. Transactional Logging for LSN Systems

NAT44/LSN 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 LSN 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 LSN hardware and/or exported to an external system for processing. Providers may need to keep track of this information (securely) to meet regulatory and/or legal obligations.

## 4. MPLS/VPN based NAT44/LSN Framework

The MPLS/VPN [RFC4364] framework for NAT44/LSN segregates the 'pre-translated' realms within the service provider space into Layer-3 MPLS/VPNs. The provider can deploy a single realm for all NAT44/LSN 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 and routing plane.

The MPLS/VPN infrastructure provides control plane and forwarding separation for the traditional IPv4 service environment and NAT44/LSN environment(s). The separation allows for routing information (such as default routes) to be propagated separately for these two major service classes. Traffic can be efficiently routed to the Internet for normal flows, and routed directly to translators for NAT44/LSN mediated flows. Although many providers 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 realm.

The physical location of the VRF Termination point and LSN can vary and be located anywhere within the provider's MPLS enabled network. This model fully virtualizes the translation service forwarding from the base IPv4 forwarding environment which will likely carry Internet bound traffic. The base IPv4 environment can continue to service traditional IPv4 customer flows plus post translated NAT44/LSN flows.

Figure 1 provides a view of the basic model. The Access node provides CPE access to either the NAT44/LSN VRF or the Global Routing Table, depending on whether the customer receives a private or public IP. Translation mediated traffic follows an MPLS LSP which can be setup dynamically and can span one hop, or many hops (with not 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.

If more then one VRF (translation realm) is used within the provider space, each VPN instance can manage NAT44/LSN flows independently for the respective realm. Various redundancy models can be used within this architecture to support failover from one physical LSN 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.

Since the MPLS/VPN based traffic (LSPs) would share a common topology base with traditional IPv4 services, QoS techniques (if used) on the base IPv4 traffic flows can be applied to the MPLS based traffic for

consistency. IP specific policy can be applied inside the VPN or applied to the MPLS forwarded (label switched) packet depending on the provider's hardware and system wide capabilities.

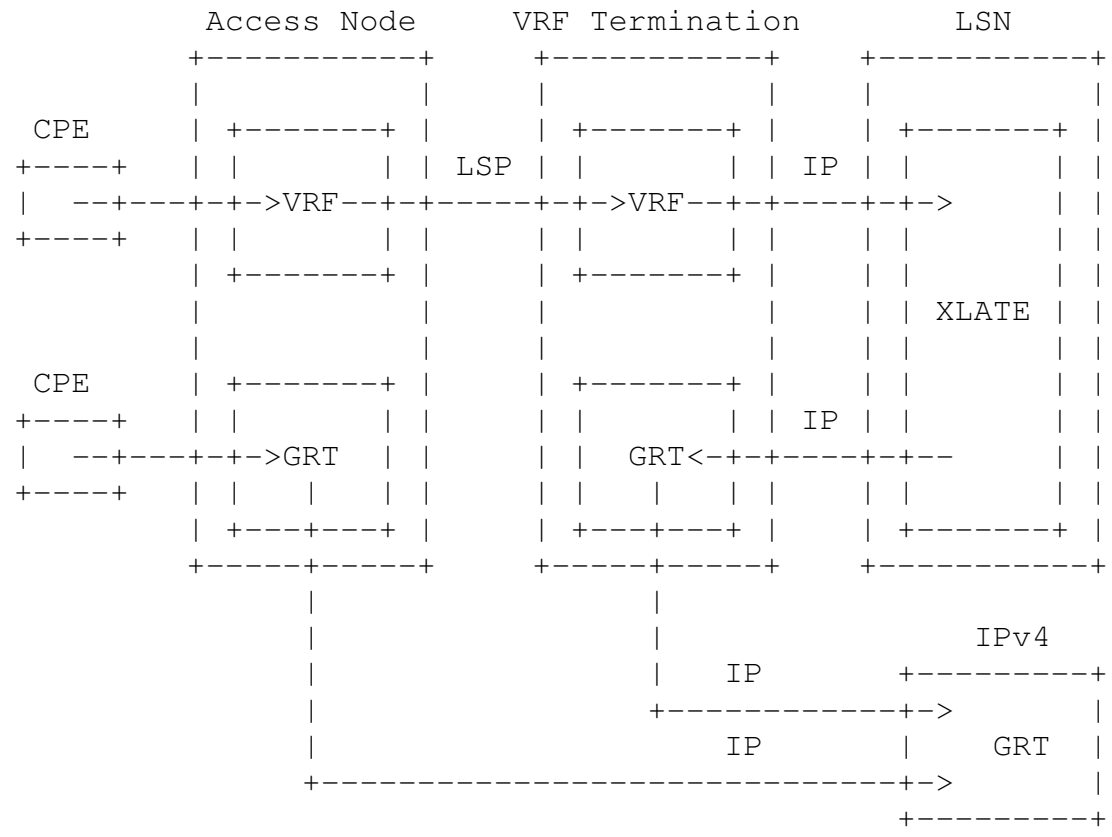


Figure 1 Basic MPLS/VPN NAT44/LSN Model

## 4.1. Service Separation

The MPLS/VPN NAT44/LSN 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 provider networks. Service separation can be accomplished on common access technology hardware

including those used for DOCSIS [CMTS], Ethernet Access, DSL [BRAS], and Mobile Access [GGSN/AGNGW] architectures.

### 4.2. Internal Service Delivery

Internal services can be delivered directly to the privately addressed endpoint within the NAT44/LSN domain without translation. This can be accomplished using direct route exchange (import/export) between the NAT44/LSN VRFs and the Services VRFs. The previous statement assumes the provider puts key services into a VRF for easy routed exchange. This model allows the provider to maintain separate forwarding rules for translated flows, which require a pass through the translator to reach an external network entity, 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 LSN 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 independent manner. Additionally, the provider can allow NAT44/LSN endpoints to be accessible via an untranslated path reducing the complexities of provider initiated management flows. This last point is of key interest since NAT44 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 LSN/translator and therefore is not subject to problematic behaviors related to NAT. The services VRF contains routing information which can be “imported” into the access node VRF and the LSN VRF routing information can be “imported” into the services VRF.

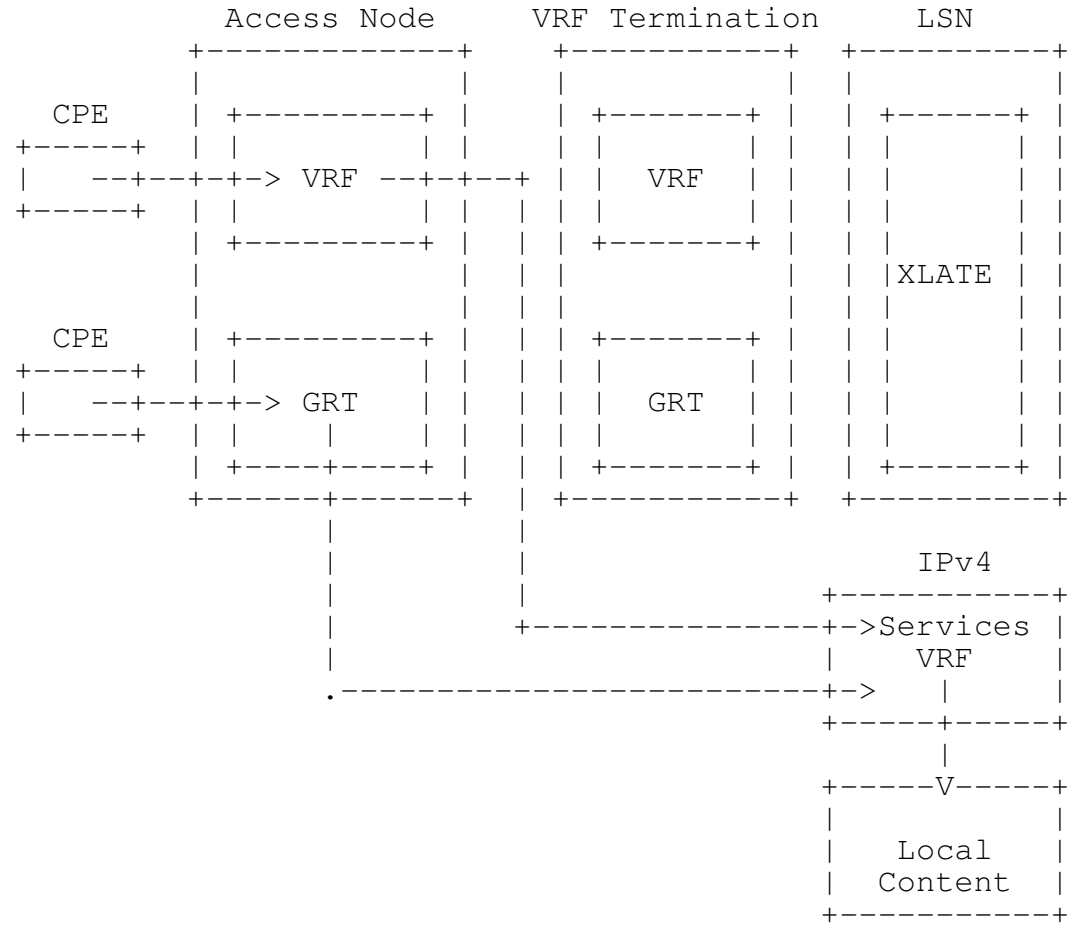


Figure 2 Internal Services and NAT44/LSN By-Pass

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

## 4.3. Dual Stack Operation

The MPLS/VPN NAT44/LSN model can also be used in conjunction with IPv4/IPv6 dual stack service modes. Since many providers will use LSNs on an interim basis while IPv6 matures within the global Internet, a dual stack option is of strategic importance. Providers can offer this dual stack service for both traditional IPv4 (global IP) endpoints and NAT44/LSN mediated endpoints.

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

while allowing IPv4 flows to remain within the IPv4 LSN VRF for translator services.

The figure below shows how IPv4 translation services can be provided alongside IPv6 based services. The model shown allows the provider to enable NAT44/LSN 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.

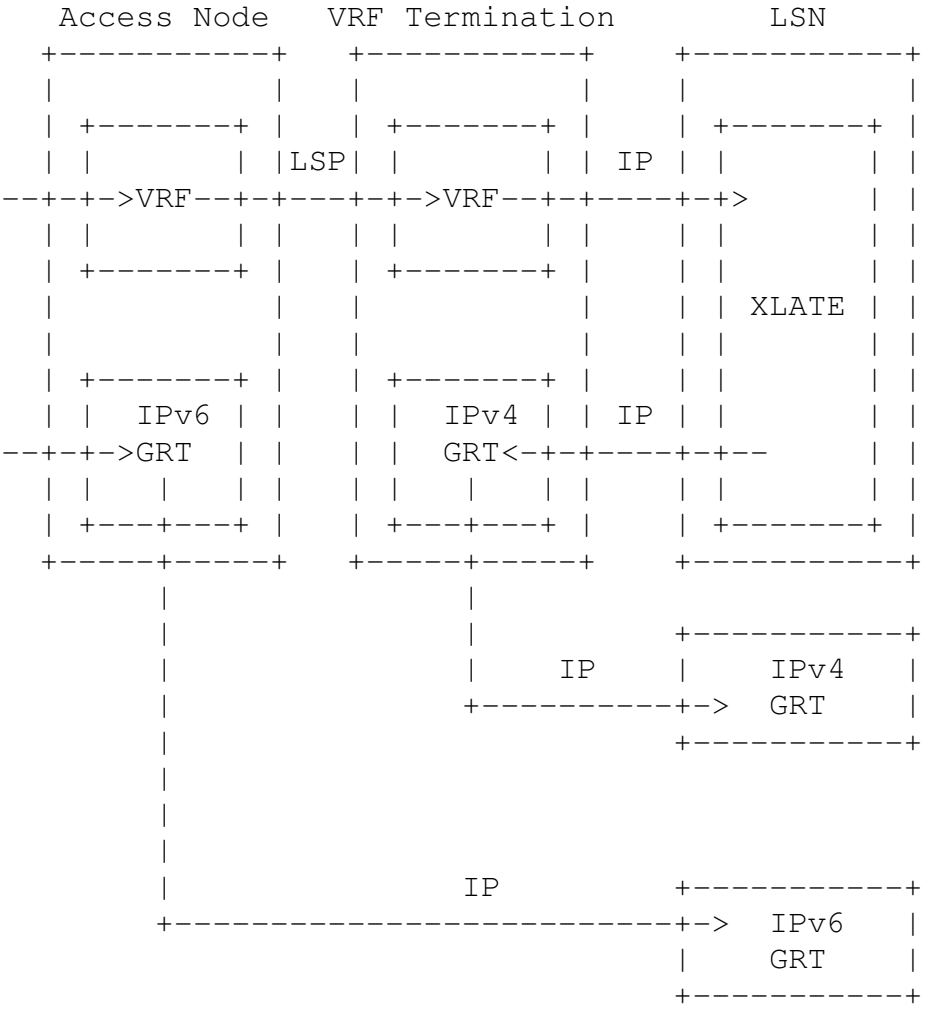


Figure 3 NAT44/LSN with IPv6 Dual Stack Operation

## 4.4. Deployment Flexibility

The LSN 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, physical locations 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 a service provider 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 not specific need to diversify the existing infrastructure in most cases.

### 4.5. Comparison of MPLS/VPN Option versus other LSN Attachment Options

Other integration architecture options exist which can attach NAT44/LSN 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

#### 4.5.1. IEEE 802.1Q

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

This option is most effective if LSN instances are placed directly upstream of the access node. Distributed LSN instance placement is not likely an initial stage of the NAT44/LSN deployment due to cost and demand factors.

#### 4.5.2. Policy Based Routing

Policy Based Routing (Static Routing) provides another option to direct NAT44/LSN 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.

#### 4.5.3. Traffic Engineering

Traffic Engineering can also be used to direct traffic from an access node towards a translator. Traffic Engineering, like PBR, 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 LSN translation points are augmented or moved.

#### 4.5.4. Multiple Routing Topologies

Multiple routing topologies can be used to direct NAT44/LSN based flows to translators. This option would achieve the same basic goal as the MPLS/VPN option but with additional implementation overhead. Since provider based translation is expected to have an unknown lifecycle, and may see various degrees of demand (dependant on providers 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 NAT44/LSN.

# 5. Experiences

## 5.1. Basic Integration and Requirements Support

The MPLS/VPN NAT44/LSN environment has been successfully integrated into real network environments utilizing existing network service delivery mechanisms. It appears to solve many issues related to provider based translation environments, while still subject to problematic behaviors inherent within NAT44 (and by extension NAT444).

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

- Centralized and Distributed Deployment model support
- Routing Plane Separation for NAT44/LSN 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)
- NAT44 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

## 5.2. Performance

The MPLS/VPN NAT44/LSN model was observed to support basic functions which are typically used by customers within a service provider 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 (mention Bit Torrent?)
- DNS (based on services extranet option, but was problematic when passed through a translator)

NAT44/LSNs 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 (subject to further validation)

## 6. Security Considerations

The same security considerations would typically exist for NAT44/LSN deployments when compared with traditional IPv4 based services.

If a provider plans to operation 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.

## 7. 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.

## 8. Conclusions

The MPLS/VPN delivery method for NAT44/LSN is an effective and scalable way to delivery mass translation services. The architecture avoids the complex requirements of traffic engineering and policy based routing. This is advantageous since the NAT44/LSN environment is expected to change over time and will potentially migrate to more progressive dual stack environments like DS-Lite over time.

The architecture solves many of this issues related to NAT444 as an overall translation model which are of concern to large Service Providers.

## 9. References

### 9.1. Normative References

[I-D.shirasaki-nat444-01] Yamagata, I., Shirasaki, Y., Nakagawa, A., Yamaguchi, J., Ashida, H., "NAT444", draft-shirasaki-nat444-01 (work in progress), March 2010

[I-D.nishitani-cgn-04] Yamagata, I., Nishitani, T., Miyakawa, S., Nakagawa, A., Ashida, H., "Common requirements for IP address sharing schemes", draft-nishitani-cgn-04 (work in progress), March 2010.

[I-D.draft-ietf-softwire-dual-stack-lite-04] Durand, A., "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", draft-ietf-softwire-dual-stack-lite-04(work in progress), March 8, 2010

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

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

## 9.2. Informative References

[I-D.azinger-additional-private-ipv4-space-issues-04] Azinger, M., Vegoda, L., "Additional Private IPv4 Space Issues", draft-azinger-additional-private-ipv4-space-issues-04 (work in progress), April 2010.

[I-D.ford-shared-addressing-issues-02] Ford, M,. Boucadair, M., Durand, A., Levis, P., Roberts, P., "Issues with IP Address Sharing", draft-ford-shared-addressing-issues-02 (work in progress), March 2010.

[I-D.shirasaki-isp-shared-addr] Yamagata, I., Miyakawa, S., Nakagawa, A., Yamaguchi, J., Ashida, H., "ISP Shared Address", draft-shirasaki-isp-shared-addr-04 (work in progress), March 2010.

[I-D.draft-wing-softwire-port-control-protocol-01] Wing, D., Penno, R., Boucadair, M., "Port Control Protocol (PCP)", draft-wing-softwires-port-control-protocol-01, March 2010.

[RFC5569] Despres, R "IPv6 Rapid Deployment on IPv4 Infrastructures (6rd), RFC5596, January 2010

[RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, January 2003.

[RFC5305] Smit, H. and T. Li, "IS-IS Extensions for Traffic Engineering", RFC 5305, October 2008.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC3272] Awduche, D., Chiu, A., Elwalid, A., Widjaja, I., Xiao, X., "Overview and Principles of Internet Traffic Engineering", RFC 3272, May 2002.

[RFC2702] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M., McManus, J., "Requirements for Traffic Engineering over MPLS", RFC 2702, September 1999.

[RFC4787] Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", BCP 127, RFC 4787, January 2007.

[RFC5382] Guha, S., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, "NAT Behavioral Requirements for TCP", BCP 142, RFC 5382, October 2008.

[RFC5508] Srisuresh, P., Ford, B., Sivakumar, S., and S. Guha, "NAT Behavioral Requirements for ICMP", BCP 148, RFC 5508, April 2009.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., Swallow, G., "RSVP-TE: Extensions to RSVP for LSP Tunnels" RFC 3209, December 2001

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

[RFC1918] Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot., G. J., Lear, E., "Address Allocation for Private Internets", RFC 1918, February 1994.

[RFC2881] Mitton, D., Beadles, M., “Network Access Server Requirements Next Generation”, RFC2881, July 2000.

## 10. Acknowledgments

Thanks to the following people for their participating in integrating and testing the NAT44/LSN 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

Authors' Addresses

Victor Kuarsingh
Rogers Communications Inc
8200 Dixie Road
Brampton, Ontario L6T 0C1
Canada

Email: victor.kuarsingh@rci.rogers.com

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

Email: john.cianfarani@rci.rogers.com