

Network Working Group
Internet Draft
Intended status: Informational
Expires: April 24, 2011

Qiong Sun
Chongfeng Xie
China Telecom
Xing Li
CongXiao Bao
CERNET Center/Tsinghua University
Ming Feng
China Telecom
March 6, 2011

**Considerations for Stateless Translation (IVI/dIVI) in Large SP
Network
draft-sunq-v6ops-ivi-sp-02.txt**

Abstract

With the approaching exhaustion of IPv4 address space, large-scale SPs are now faced with the only real option to deploy IPv6 in a timely manner. In order to achieve smooth transition to IPv6, migration tools should be introduced for different deployment models. Among different IPv6 transition mechanisms, dIVI is a prefix-specific and stateless address mapping method which can directly translate IPv4 packet to IPv6 packet. This document describes the challenges and requirements for large SP to deploy IPv6 in operational network, the experimental results of dIVI in our laboratory and the considerations for dIVI deployment in large SP operational network.

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), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>

This Internet-Draft will expire on September 6,2011.

Copyright Notice

Copyright (c) 2011 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	3
2.	Terminologies	3
3.	Problem Statement	3
4.	Laboratory experiment.....	5
4.1.	Experiment environment.....	6
4.2.	Experiment configuration.....	7
4.3.	Experiment results.....	7
5.	dIVI Deployment Scenario.....	9
5.1.	Network Architecture for Large SP Network.....	9
5.2.	dIVI Deployment Scenario in Operational Network.....	11
6.	Considerations for dIVI deployment.....	12
6.1.	Addressing	12
6.2.	Routing	13
6.3.	DNS	13
6.4.	AAA and User Management.....	13
6.5.	Network management.....	14
6.6.	dIVI CPE Requirements and Configuration	14
6.7.	dIVI Xlate Placement in Large SP Network	14
6.8.	ALG consideration	15
7.	Security Considerations.....	15
8.	IANA Considerations	15
9.	References	15
9.1.	Normative References.....	15
10.	Acknowledgments	16

Sun

Expires September 6, 2011

[Page 2]

1. Introduction

The dramatic growth of the Internet is accelerating the exhaustion of available IPv4 addressing pool. It is widely accepted that IPv6 is the only real option on the table for the continued growth of the Internet. However, IPv6 deployment is a huge systematic project and a lot of challenges will arise especially in large SP operational network.

In order to achieve smooth transition to IPv6, migration tools should be introduced for different deployment models, among which dIVI is a stateless translation mechanism with good scalability. This document describes the challenges and requirements for large SPs in IPv6 transition period. Then, we introduce dIVI experimental results in our laboratory. And finally, the considerations for designing and deploying dIVI in operational network are discussed in terms of addressing and routing, DNS deployment requirement, AAA support and user management, network management, CPE requirement and Xlate placement.

2. Terminologies

This document uses the terminologies defined in [I-D.ietf-behave-v6v4-framework], [[I-D.ietf-behave-v6v4-xlate](#)], [I-D.ietf-behave-address-format], [[I-D.ietf-behave-v6v4-xlate-stateful](#)], and [I-D.xli-behave-ivi].

The key words MUST, MUST NOT, REQUIRED, SHALL, SHALL NOT, SHOULD, SHOULD NOT, RECOMMENDED, MAY, and OPTIONAL, when they appear in this document, are to be interpreted as described in [[RFC2119](#)].

3. Problem Statement

It is well known that the pool of public IPv4 addressing is nearing its exhaustion. The '/8' IANA blocks for Regional Internet Registries (RIRs) are projected to 'run-out' towards the end of 2011. Credible estimates based on past behavior suggest that the RIRs will exhaust their remaining address space by early 2012, apart from the development of a market in IPv4 address space. Thus, it will become much more difficult to get available public IPv4 addresses. In the same time, a lot of emerging applications, e.g. Apple's iPad, Motion's BlackBerry, etc. will quickly deplete the available addresses. This has led to a heightened awareness among the providers to consider introducing IPv6 to keep the Internet operational.

It has been widely accepted that the end goal of IPv6 transition is to achieve an end-to-end IPv6-only network, and IPv4 can eventually

Sun

Expires September 6, 2011

[Page 3]

be turned off. However, since it will have impact on almost the entire world, it will take a considerable period of time to reach the ultimate goal. As a result, IPv4 and IPv6 need to coexist during the whole transition period. In this document, we mainly focus on IPv6 migration issues from large ISP's point of view. In order to facilitate smooth IPv6 migration, some factors need to be taken into consideration especially for large SPs. There are ten major requirements:

1. It should deploy in an incremental fashion and the overall transition process should be stable and operational.
2. It should largely reduce IPv4 public address consumption.
3. It should accelerate the deployment of IPv6, rather than just prolonging the lifecycle of IPv4 by introducing multiple layers of NAT.
4. There should be no perceived degradation of customer experience. As a result, IPv6 transition mechanisms should provide IPv4 service continuity.
5. It should achieve scalability, simplicity and high availability, especially for large-scale SPs.
6. It should have user management and network management ability.
7. It should support user authentication, authorization and accounting as an essential part of operational network.
8. Most ISPs need some kind of mechanisms to trace the origin of traffic in their networks. This should also be available for IPv6 traffic.
9. It should have good throughput performance and massive concurrent session support.
10. It should maintain the deployment concepts and business models which have been proven and used with existing revenue generating IPv4 services.

All existing IPv6 transition mechanisms can be widely divided into three categories: dual-stack solution, translation-based solution and tunnel-based solution. In this document, we mainly concentrate on stateless translation mechanism: dIVI. The original stateless IPv4/IPv6 translation (stateless 1:1 IVI) is scalable, [I-D.ietf-behave-v6v4-framework], [[I-D.ietf-behave-v6v4-xlate](#)], [I-D.ietf-

Sun

Expires September 6, 2011

[Page 4]

behave-address-format],[I-D.xli-behave-ivi]. But it cannot use the IPv4 addresses effectively. The stateless dIVI[I-D.xli-behave-divi] is a double translation mechanism which includes a 1:N stateless translator and a 1:1 Hgw translator. The 1:N stateless translator is implemented in the border between the IPv6 network and the IPv4 Internet. It translates the packets between IPv4 and IPv6 with the 1:N stateless address mapping. The 1:1 Hgw translator is implemented between an IPv6 network and user's end system. It translates the packets between IPv4 and IPv6 with 1:1 stateless address mapping. In addition, the home gateway translator maps random source port numbers to restricted port number based on the extended IPv4-translatable address format and keeps the mapping table in database for the port number mapping of the retuning packets and all the packets in the same session.

dIVI support bidirectional communication initiated from IPv4 and IPv6. It can be deployed in an IPV6-only access network, in which operational and maintenance cost can be reduced. It has very good scalability and can largely reduce IPv4 address consumption.

In this document, we firstly demonstrate the laboratory experimental results of dIVI in [section 4](#). We can see that dIVI can support most of the current IPv4 applications in IPv6-only access network, while largely reducing IPv4 address consumption. And then dIVI deployment model and considerations in large operational network are proposed in [section 5](#) and [section 6](#) respectively. Some important factors need to be taken into account when introducing dIVI. Since most challenges in dIVI have no big difference compared to an IPV6-only environment, we strongly recommend that related network elements should take the corresponding modifications in order to guarantee the IPv6 transition process to be operational and manageable.

4. Laboratory experiment

We have tested dIVI using the prototype in our laboratory. The major objective listed in the following is to verify the functionality and performance of dIVI:

- 0 Verify how to deploy dIVI in practical network.
- 0 Verify what applications can be used in dIVI.
- 0 Verify what Operating Systems can be supported in dIVI.
- 0 Verify the effect of user experience with limited ports.
- 0 Verify the performance of dIVI.

Sun

Expires September 6, 2011

[Page 5]

4.1. Experiment environment

We have tested dIVI in our laboratory. The network topology is depicted in Figure 1.

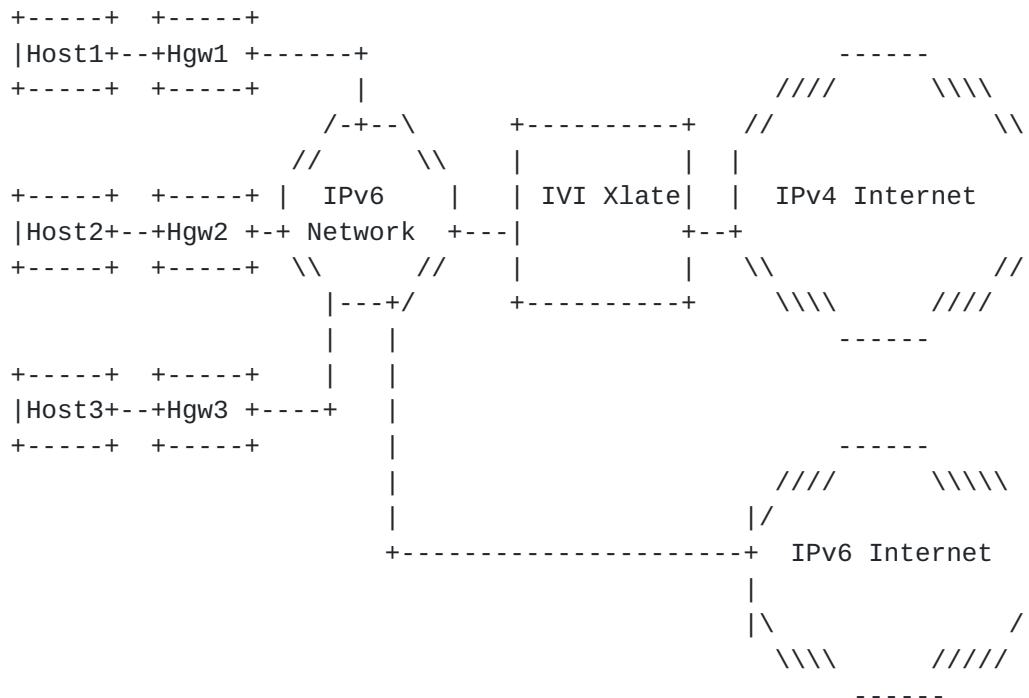


Figure 1 dIVI topology in the test

There are three key components in the test:

- o The Hosts are dual-stack customers, who could run IPv4 application, IPv6 application or dual stack application.
- o The Home Gateways (Hgw) are dIVI translator in user side. It translates the packets between IPv4 and IPv6 with 1:1 stateless address mapping, and maps random source port numbers to restricted port number.
- o The Xlate translates the packets between IPv4 and IPv6 with the 1:N stateless address mapping.

The network between Hgw and Xlate is IPv6-only, and the network behind Hgw is dual-stack. Thus, the hosts behind Hgw can communicate with both IPv4 Internet and IPv6 Internet.

Sun

Expires September 6, 2011

[Page 6]

4.2. Experiment configuration

For address configuration, each host will use two IPv6 addresses: one is IIVI6 address which is synthesized in Hgw with the IIVI4 address and port-related information, and the other is non-IIVI IPv6 address which is used for native IPv6 communication. We should notice that only non-IIVI IPv6 address needs be allocated to end users. Besides, each user will get an IIVI4 address from Hgw.

For routing configuration, both IIVI address and non-IIVI address need to be imported into the IPv6-only network.

For port configuration, since there are 65536 TCP/UDP ports for each IP address, and in fact one can use hundreds only for normal applications, so one IPv4 address can be shared by multiple customers. In our experiment, we selected ratio to be 128. That is to say, one IPv4 address is shared by 128 users, and there are 512 available ports per user.

For DNS configuration, there is no need to have additional DNS64 for dIIVI. Only an IPv6 DNS server with A/AAAA records is needed and the DNS address is manually configured in Hgw. Besides, Hgw has implemented DNS Proxy and it will convert an IPv4 DNS request/response to IPv6 DNS request/response.

For ALG configuration, there is no need to deploy specific ALG for IPv4 applications in dIIVI.

4.3. Experiment results

In our laboratory, we mainly have taken the following tests:

- o Application test: The applications we have tested include web, email, Instant Message, ftp, telnet, SSH, video, Video Camera, P2P, online game, Voip, VPN and so on.
- o Operating System test: The OS we have tested includes Win7, VISTA, windows XP.
- o Performance test: We have measured the parameters of concurrent session number, throughput performance.

Sun

Expires September 6, 2011

[Page 7]

The experimental results are listed as follows:

Type	Experiment Result
Application test	dIVI hosts can support web, email, im, ftp, telnet, SSH, video, Video Camera, P2P, online game, voip, and so on.
Operating System test	dIVI can widely support Win7, VISTA, windows XP.
Performance test	The performance test for dIVI Xlate is carried out on a normal PC. Due to limitation of the PC hardware, the overall throughput is not quite good. However, it can still support more than one hundred million concurrent sessions.

Figure 2 dIVI test result

From the experiment, we can have the following conclusions:

1. dIVI can have good scalability. Xlate does not need to maintain any session state, and only limited session states have to be maintained in Hgw for its customer.
2. dIVI can be deployed in an incremental way. The most complicated part of dIVI is addressing and routing. The configuration for DNS and ALG is very simple.
3. dIVI can support a majority of current IPv4 applications.
4. dIVI can support a variety of Operating Systems.
5. With the ratio of 128 (512 maximum concurrent sessions), there is no perceived degradation of customer experience.

However, in the current status of equipment, e.g. BRAS, end system, etc., the support for IPv6-only function has not been fully accomplished. Therefore, there are still some limitations which would be improved in the next version of dIVI development when deploying dIVI prototype in practical operational network:

Sun

Expires September 6, 2011

[Page 8]

1. Address assignment process have not been integrated with existing address allocation system.
2. Currently, IVI routing entries are configured manually.
3. Hgw has not integrated PPPoE functionality with dIVI functionality.
4. AAA system has not supported IVI-related (or IPv6-only) functions.

With regard to the listed IPv6 transition requirements in [section 3](#), most of them can be satisfied by dIVI, except for the requirement of network management and user management. These two points should be paid special attention for large SPs, which will be further discussed in [section 6](#).

5. dIVI Deployment Scenario

In order to investigate the ways to deploy dIVI in operational network, we firstly briefly discuss network architecture for large SP network. Then dIVI deployment model is introduced.

5.1. Network Architecture for Large SP Network

In large SPs, the generic network topology can be divided into four main parts (as depicted in Figure3): the Customer Premises Network (CPN), the Access Network (AN), the Metro Area Network (MAN), and the Backbone Network.

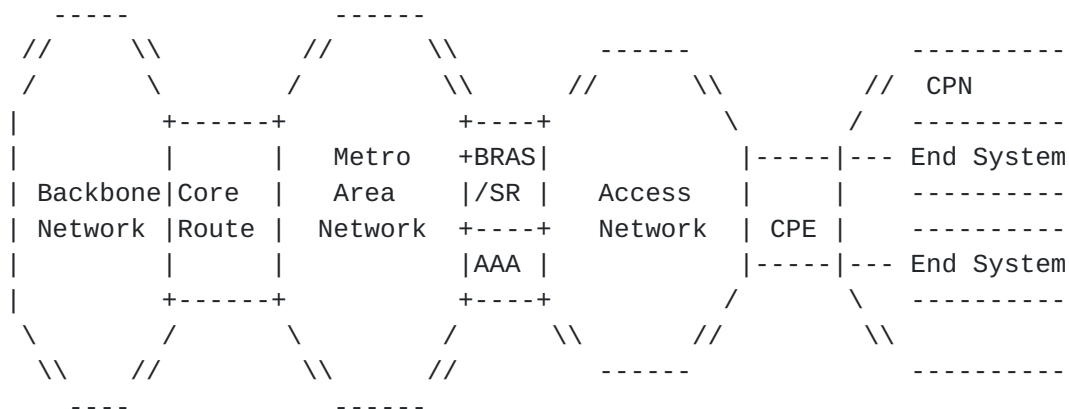


Figure 3 Network Architecture for Large SP Network

Sun

Expires September 6, 2011

[Page 9]

1. CPN is the part of the network by a customer when connecting to an ISP's network which includes the CPE and the last hop link.
2. In Access Network, a very wide variety of access technologies can be used, including ADSL, Ethernet, PON, ATM, WIFI, etc.
3. MAN is the aggregated network for customers in one single metro, with the vast range of size. In most metro networks, BRAS is connected to Core Router directly, while for a small portion of large metro networks, BRAS is connected to Core Routers via aggregated routers.
4. Backbone network is to offer transit service between MANs and other ISPs.

There are typically two network models for fixed broadband access service: one is to access using PPPoE/PPPoA authentication method while the other is to use IPoE. The first one is usually applied to Residential network and SOHO networks. Subscribers in CPNs can access broadband network by PPP dial-up authentication. BRAS is the key network element which takes full responsibility of IP address assignment, user authentication, traffic aggregation, PPP session termination, etc. Then IP traffic is forwarded to Core Routers through Metro Area Network, and finally transited to external Internet via Backbone network.

The second network scenario is usually applied to large enterprise networks. Subscribers in CPNs can access broadband network by IPoE authentication. IP address is normally assigned by DHCP server, or static configuration.

5.2. dIVI Deployment Scenario in Operational Network

The deployment model of dIVI in operational network is depicted in Figure4.

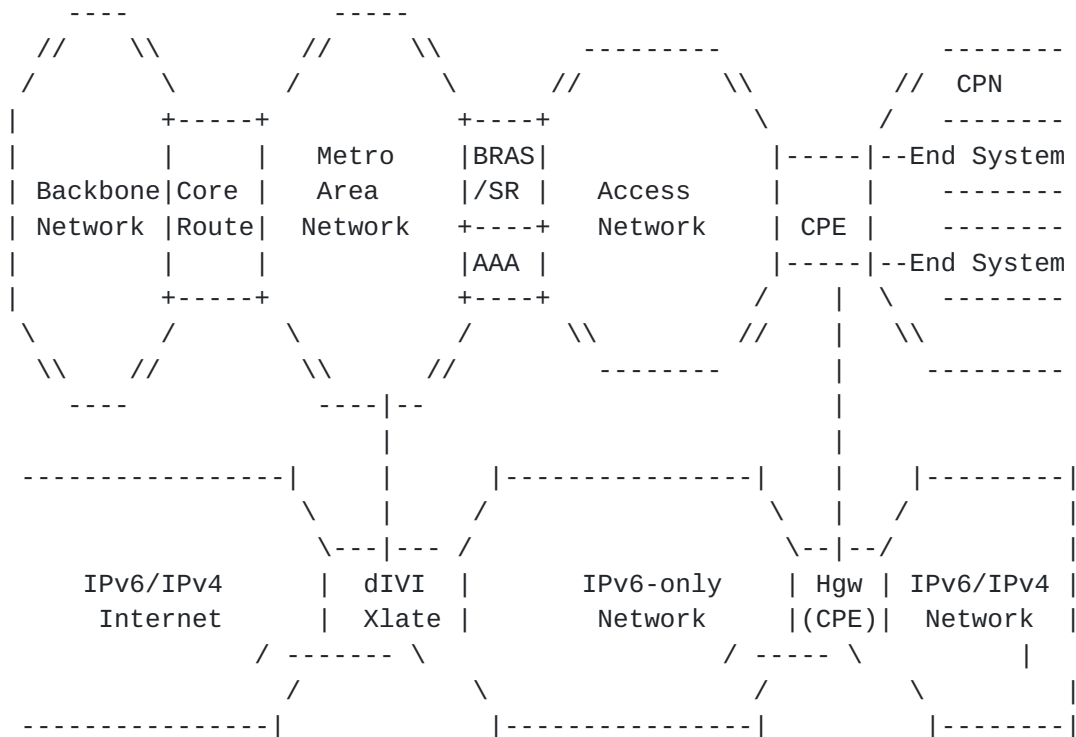


Figure 4 dIVI Deployment in Operational Network

In stateless dIVI, the network between Hgw and Xlate is an IPv6-only network, in which the operational and maintenance cost can be greatly reduced. The access network behind Hgw can either be IPv4-only or dual-stack. Thus, IPv4-only system and dual-stack system can communication with IPv4 Internet using shared IPv4 address by dIVI and the dual-stack system can also communicate with IPv6 Internet directly.

In operational network, Hgw can usually be integrated with CPE, while Xlate can be in someplace of MAN or Backbone Network. Subscribers can get IPv6 address from BRAS/SR after user authentication stage. Then, IVI-related route should be imported into the IPv6-only network between Xlate and Hgw. The detailed considerations for dIVI deployment will be discussed in [section 6](#).

Sun

Expires September 6, 2011

[Page 11]

6. Considerations for dIVI deployment

This section describes the considerations for dIVI deployment in large operational network.

The major differences between dIVI deployment in laboratory and operational network lie in:

1. Operational network is a commercial network with strict user management requirement, while in laboratory it is simple and straightforward.
2. Operational network has different kinds of network equipments, e.g. BRAS/SR, CPE, Radius, etc. It would be more difficult to take modifications on all of these equipments.
3. Operational network has a large number of customers. Thus, it would be impossible to take manual configuration for all customers.

In this section, we try to outline considerations for dIVI deployment on large SP network. Some of the features are not specific to dIVI, but rather to a general requirement on all IPv6 transition techniques.

6.1. Addressing

In dIVI, there is no need to allocate IPv6 address explicitly to end users. Thus, the process of IPv6 address assignment can be integrated with existing IPv6 address allocation process. Only CPE will need to get IPv4 address, reallocate it to end user, do port-mapping and traffic translation with port-related information. Here are some basic considerations in dIVI addressing:

- o Determine IPv6 prefix for each Xlate. Operators should use its own prefix as an IPv6 prefix, i.e. pref=2001:db8:a4a6::/48, in order to perform stateless translation. Address allocation process in BRAS/SR should be consistent with Xlate.
- o Determine the embedded IPv4 address and port multx ratio. Operators should estimate the scale of subscribers in a certain region, evaluate the number of remaining IPv4 address, and analyze the number of concurrent ports. It is a tradeoff between multx ratio and concurrent port numbers. The bigger the multx ratio is, the more an IPv4 address can be shared by multiple subscribers and the less concurrent port number can be used per subscriber. From the above test in our laboratory, we choose multx ratio to be 128 and it is enough for current usage.

Sun

Expires September 6, 2011

[Page 12]

- o Determine the ways to distribute the configuration profile including IVI4 address and port multx ratio to Hgw automatically, either by extended DHCP option, or other protocols.

6.2. Routing

In dIVI, IVI4(i) and IVI6(i) will be aggregated to ISP's IPv4 address and ISP's IPv6 address. They will not affect the global IPv4 and IPv6 routing tables

In dIVI deployment model, Hgws are normally configured with a default route that points to the BRAS/SR. The routers between BRAS/SR and Xlate run IPv6 dynamic routing protocols (IGP or BGP), and routers in the upper level of Xlate run IPv4 dynamic routing protocols. Therefore, the aggregated IVI6 routing directing to the upper routers will be learned/inserted by in IPv6-only domain. And the IVI6 route directing to Hgws should also be configured in BRAS/SR.

6.3. DNS

In dIVI, there is no DNS64/DNS46 needed anymore. An IPv6 DNS server is needed to process IPv6 DNS request/response, and the address of IPv6 DNS server should be passed to Hgw.

Since there is no IPv4 DNS server in IPv6-only network, it is recommended that Hgw should implement IPv4-to-IPv6 DNS Proxy to convert an IPv4 DNS request/response to IPv6 DNS request/response accordingly.

6.4. AAA and User Management

User authentication can be used to control who can have the dIVI connectivity service. This is not always required when a customer of IPv4 service automatically can have access to the dIVI service. However, it is highly recommended that IPv6-only customers should be authenticated separately. It is good for security, trouble shooting, user accounting, etc. There are some major points that AAA systems need to be taken into consideration:

- o User authentication function needs to be extended to support the identification of IPv6-only subscriber, with additional dIVI-related profile for subscribers, e.g. IVI6 address, IVI4 address, non-IVI address, etc.
- o User accounting and management function needs to be extended to identify dIVI user (or IPv6-only user) separately.

Sun

Expires September 6, 2011

[Page 13]

In summary, the major challenge of dIVI for the AAA and User Management is no big difference compared to an IPv6-only environment.

6.5. Network management

There are two issues to manage dIVI in operational network:

- o Manage IPv6-only network. Operators should be able to manage IPv6-only network, including IPv6 MIB modules, Netflow Records, log information, etc.
- o Manage the translation process. There are some exceptions that the MIB modules need to add dIVI related features, e.g. dIVI device management, dIVI traffic monitoring, etc.

6.6. dIVI CPE Requirements and Configuration

In dIVI, CPE is an important network element. It should perform DHCP server, integrated user authentication function, traffic translation and port mapping, DNS proxy, etc. The major operations in dIVI CPE include:

- o Address assignment: dIVI CPE should support IPv4 address assignment by DHCP process to end users. It should also support IPv6 address assignment, either by stateful DHCP or stateless auto-configuration.
- o Integrated user authentication function: dIVI CPE should integrate with existing user authentication function, e.g. PPPoE/PPPoA, etc.
- o DNS: CPE should enable [RFC 5006](#), well-known addresses, and DHCPv6 in order to maximize the likelihood of dIVI Hgw being able to use DNS without manual configuration. Besides, dIVI CPE should also support IPv4-to-IPv6 DNS proxy.

6.7. dIVI Xlate Placement in Large SP Network

Normally, dIVI Xlate can be deployed in "centralized model" and "distributed model".

In "centralized model", dIVI Xlate could be deployed in the place of Core Router, or even in the entrance of ICP. Since dIVI is a stateless method with better scalability than stateful ones, it can handle numerous concurrent sessions.

In "distributed model", dIVI Xlate is usually be integrated with BRAS/SR. So each Xlate should be configured with its own IPv6 prefix

Sun

Expires September 6, 2011

[Page 14]

and is responsible for translating the traffic of its own region. The number of subscribers is normally limited, so does the number of IVI routing entries. However, the network infrastructure should still be upgraded to dual-stack support in MAN and backbone network, and so the decreased operational cost is rather limited. Besides, since the newly emerging customers might be distributed in the whole Metro area, we have to deploy Xlate on all BRAS/SRs. This will cost a lot in the initial phase of IPv6 transition period.

In summary, we strongly recommend adopting "centralized model" for dIVI. It is a cost-effective way and easy to manage.

6.8. ALG consideration

dIVI does not require ALG, this is a very important feature in the initial development phase of IPv6.

7. Security Considerations

There are no security considerations in this document.

8. IANA Considerations

This memo adds no new IANA considerations.

Note to RFC Editor: This section will have served its purpose if it correctly tells IANA that no new assignments or registries are required, or if those assignments or registries are created during the RFC publication process. From the author's perspective, it may therefore be removed upon publication as an RFC at the RFC Editor's discretion.

9. References

9.1. Normative References

[I-D.ietf-behave-address-format] C., Bao, Huitema, C., Bagnulo, M., Boucadair, M., and X.Li, "IPv6 Addressing of IPv4/IPv6 Translators", [draft-ietf-behave-address-format-10](#) (work in progress), August 2009.

[I-D.ietf-behave-dns64] Bagnulo, M., Sullivan, A., Matthews, P., and I. Beijnum, "DNS64: DNS extensions for Network Address Translation from IPv6 Clients to IPv4 Servers", [draft-ietf-behave-dns64-11](#) (work in progress), October 2009.

Sun

Expires September 6, 2011

[Page 15]

[I-D.ietf-behave-v6v4-framework] Baker, F., Li, X., Bao, C., and K. Yin, "Framework for IPv4/IPv6 Translation", [draft-ietf-behave-v6v4-framework-10](#) (work in progress), October 2009.

[I-D.ietf-behave-v6v4-xlate] Li, X., Bao, C., and F. Baker, "IP/ICMP Translation Algorithm", [draft-ietf-behave-v6v4-xlate-23](#) (work in progress), October 2009.

[I-D.ietf-behave-v6v4-xlate-stateful] Bagnulo, M., Matthews, P., I. Beijnum, "IP/ICMP Translation Algorithm", [draft-ietf-behave-v6v4-xlate-12](#) (work in progress), October 2009.

[I-D.xli-behave-divi] Li, X., Bao, C., and Zhang, H., "Address-sharing stateless double IVI", [draft-xli-behave-divi-01](#), April 29, 2010.

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

[10](#). Acknowledgments

The authors would like to thank to Fred Baker for his continuous suggestion around this topic over the years. Thanks to Qian Wang, Jie Hu and Fan Shi for useful feedback.

Authors' Addresses

Qiong SUN
China Telecom Beijing Research Institute
Room 708 No.118, Xizhimenneidajie, xicheng District Beijing 100035
China

Phone: <86 10 58552636>
Email: sunqiong@ctbri.com.cn

Chongfeng Xie
China Telecom Beijing Research Institute
Room 708 No.118, Xizhimenneidajie, xicheng District Beijing 100035
China

Phone: <86 10 58552116>
Email: xiechf@ctbri.com.cn

Xing Li
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University

Phone: <86 10 62785983>
Email: xing@cernet.edu.cn

Congxiao Bao
CERNET Center/Tsinghua University
Room 225, Main Building, Tsinghua University

Phone: <86 10 62785983>
Email: congxiao@cernet.edu.cn >

Ming Feng
China Telecom
No.31, Jinrong Ave,Xicheng District,100032

Phone: <86 10 58501428>
Email: fengm@chinatelecom.com.cn