

INTERNET-DRAFT  
Intended Status: Informational  
Expires: July 1, 2014

Cameron Byrne  
T-Mobile US  
December 28, 2013

IPv6 Transitional Technology IPv4 Prefix  
draft-byrne-v6ops-clatip-01

Abstract

DS-Lite [RFC6333] directs IANA to reserve 192.0.0.0/29 for the B4 element. This memo generalizes that reservation to include other cases where a non-routed IPv4 interface must be numbered in an IPv6 transition solution.

Status of this Memo

This Internet-Draft is submitted to IETF 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/lid-abstracts.html>

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

Copyright and License Notice

Copyright (c) 2013 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	The Case of 464XLAT . . . . .	3
3.	Choosing 192.0.0.0/29 . . . . .	3
4	Security Considerations . . . . .	3
5	IANA Considerations . . . . .	3
6	References . . . . .	3
	6.1 Normative References . . . . .	4
	Authors' Addresses . . . . .	4

## 1 Introduction

DS-Lite [RFC6333] directs IANA to reserve 192.0.0.0/29 for the B4 element. This memo generalizes that IANA reservation to include other cases where a non-routed IPv4 interface must be numbered in an IPv6 transition solutions. IANA shall list 192.0.0.0/29 to be reserved for IPv6 Transitional Technology IPv4 Prefix. The result is that 192.0.0.0/29 may be used in any system that requires IPv4 addresses for backward compatibility with IPv4 communications, but does not emit IPv4 packets "on the wire".

## 2 The Case of 464XLAT

464XLAT [RFC6877] describes an architecture for providing IPv4 communication over an IPv6-only access network. One of the methods described in [RFC6877] is for the client side translator (CLAT) to be embedded in the host, such as a smartphone. In this scenario, the host must have an IPv4 address configured to present to the network stack and for applications to bind sockets.

## 3. Choosing 192.0.0.0/29

To avoid conflicts with any other network that may communicate with the CLAT, a locally unique address must be assigned.

IANA has defined a well-known range, 192.0.0.0/29, in [RFC6333], which is dedicated for DS-lite. As defined in [RFC6333], this subnet is only present between the B4 and the AFTR and never emits packets from this prefix "on the wire". 464XLAT has the same need for a non-routed IPv4 prefix. It is most prudent and effective to generalize 192.0.0.0/29 for the use of supporting IPv4 interfaces in IPv6 transition technologies rather than reserving a prefix for every possible solution.

## 4 Security Considerations

No new security considerations beyond what is described [RFC6333] and [RFC6877].

## 5 IANA Considerations

IANA is directed to generalize the reservation of 192.0.0.0/29 from DS-lite to "IPv6 Transitional Technology IPv4 Prefix".

## 6 References

## 6.1 Normative References

- [RFC6333] Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", RFC6333, August 2011.
- [RFC6877] Mawatari, M., Kawashima, M., and C. Byrne, "464XLAT: Combination of Stateful and Stateless Translation", RFC6877, April 2013.

## Authors' Addresses

Cameron Byrne  
Bellevue, WA, USA  
Email: Cameron.Byrne@T-Mobile.com

Network Working Group  
Internet-Draft  
Intended status: Informational  
Expires: May 08, 2014

G. Chen  
H. Deng  
China Mobile  
D. Michaud  
Rogers  
J. Korhonen  
Renesas Mobile  
M. Boucadair  
France Telecom  
A. Vizdal  
Deutsche Telekom AG  
C. Byrne  
T-Mobile USA  
November 04, 2013

IPv6 Roaming Behavior Analysis  
draft-chen-v6ops-ipv6-roaming-analysis-02

Abstract

This document intends to enumerate failure cases when a IPv6 subscriber roams into visited network areas. The investigations on those failed cases reveal the causes in order to notice improper configurations, equipment's incomplete functions or inconsistent IPv6 strategy.

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 May 08, 2014.

Copyright Notice

Copyright (c) 2013 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 . . . . .	2
2. Roaming Architecture Descriptions . . . . .	3
3. Roaming Scenario Overview . . . . .	4
4. Failure Cases Descriptions . . . . .	5
4.1. Failure Case 1: Incompatible with Extended PDP/PDN Type .	5
4.2. Failure Case 2: Splitting Dual-stack Bearer . . . . .	6
4.3. Failure Case 3: Shortage of IPv6 support . . . . .	7
4.4. Failure Case 4: Fallback Incapability . . . . .	7
4.5. Failure Case 5: 464xlat Support . . . . .	7
5. Discussions . . . . .	8
6. IANA Considerations . . . . .	9
7. Security Considerations . . . . .	9
8. Acknowledgements . . . . .	9
9. References . . . . .	9
9.1. Normative References . . . . .	9
9.2. Informative References . . . . .	10
Authors' Addresses . . . . .	11

## 1. Introduction

IPv6 has been deployed globally to overcome the IPv4 depletion. Operators likely start or plan to upgrade the networks that allow IPv6 subscribers to access. As the dramatical uses of Internet services with a mobile access, IPv6 is an essential part to be considered in the mobile network evolution. 3rd Generation Partnership Project (3GPP) published the IPv6 migration guidance [TR23.975], which describes different technical evolution paths. In general, operators may deploy dual-stack or IPv6 single-stack depending on network's conditions. It has been observed that those deployments are rolled out in multiple provisioning domains. In the early IPv6 stage, a mobile subscriber roaming around the different areas may experience service degradations or interruptions due to the inconsistent configurations and incomplete functions in the networks nodes. This memo intends to document the observed failed cases and analyze the causes. It's expected that operators could notice the issues and prevent potential risks.

## 2. Roaming Architecture Descriptions

The roaming process could be triggered in the following scenarios:

- o International roaming: a mobile subscriber may entry a visited network, where different PLMN identity is used. The subscribers could either in an automatic mode or a manual mode to attach a PLMN cell.
- o Intra-PLMN mobility: a subscriber moves to a visited network as that of the Home Public Land Mobile Network (HPLMN). However, the subscriber profiles may not be stored in the area. Once the subscriber attaches to the network, the subscriber profile should be extracted from the home network for the network registration.

When a mobile device is turned on or is transferred via a handover to a visited network, the mobile device will scan all radio channels and find available Public Land Mobile Networks (PLMNs) to attach. Serving GPRS Support Node (SGSN) or Mobility Management Entity (MME) in the visited networks must contact the home Home Location Register(HLR) or Home Subscriber Server(HSS) and obtain the subscriber profile. Once the authentication and registration process is completed, the PDP activation and traffic flows may be operated differently according to the subscriber data configuration. Two modes have been shown at the figure to illustrate, that are "Home routed traffic" and "Local breakout".

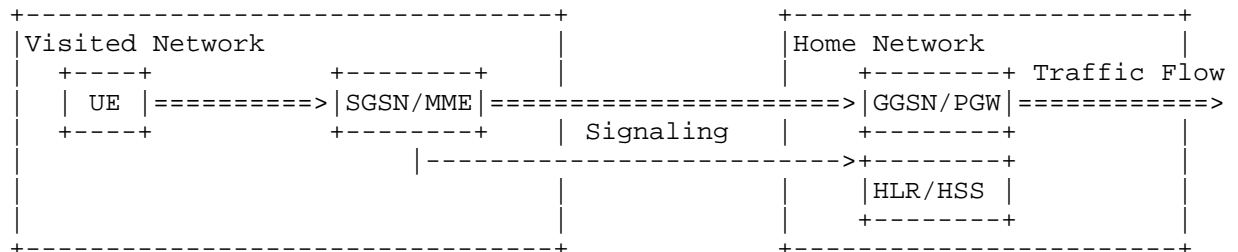
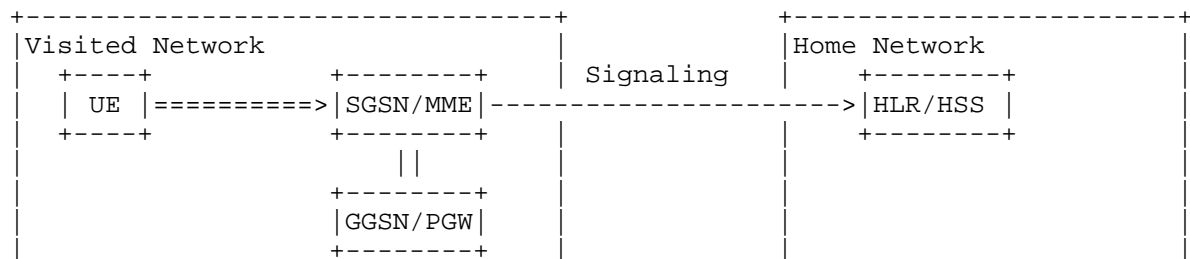


Figure 1: Home Routed Traffic



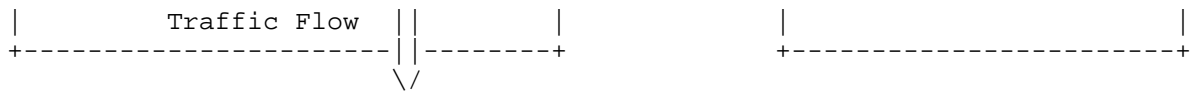


Figure2: Local Breakout

In the home routed mode, subscribers will activate the PDP/PDN context and get address from the home network. All traffic would be routed back to the home networks. That is the default case for an international roaming except for the IP Multimedia Subsystem (IMS) scenario.

In the local breakout mode, the subscriber address will be assigned from the visited network. The traffic flow would directly offloaded locally at a network node close to that device's point of attachment in the visited networks. Therefore, more efficient route would be achieved. The following will describe the cases where there is local breakout mode adopted.

- o Operators may add the APN-OI-Replacement flag defined in 3GPP [TS29.272] into user's subscription-data. The visited network indicates a local domain name to replace the user requested Access Point Name (APN). As the consequence, the traffic would be steered to the visited network. Those functions are normally deployed for the Intra-PLMN mobility cases.
- o Operators could also configure VPLMN-Dynamic-Address-Allowed flag[TS29.272] in the user profile to enable local breakout mode in Visited Public Land Mobile Networks (VPLMNs).
- o 3GPP specified Selected IP Traffic Offload (SIPTO) function[TS23.401] since Release 10 in order to get efficient route paths. It enables an operator to offload certain types of traffic at a network node close to that device's point of attachment to the access network.
- o GSMA has defined RAVEL[IR.65] as IMS international roaming architecture. Local breakout mode has been adopted for the roaming architecture.

### 3. Roaming Scenario Overview

3GPP specified three types of Packet Data Protocol (PDP)/Packet Data Networks (PDN) to describe each connection, i.e. PDP/PDN Type IPv4, PDP/PDN Type IPv6 and PDP/PDN Type IPv4v6. User devices can be set to request a particular PDP/PDN Type. Those PDP/PDN types should also be restored in Home Subscriber Server (HSS) as a part of subscriber profile, as defined in [TS29.272]. When a subscriber



roams to a visited network, the new visited network notices that it is not registered with its own system, and attempts to identify its home network. Afterwards, the visited network will contact the home network and request the subscriber profile from HSS. In this process, service may be provided in a home routed or local breakout mode. The IP address can be allocated from home network or visited network accordingly. There may be a mismatch between the subscriber request and network capability. The following table lists the potential failure cases.

UE Request	Visited Network Capability	Home routed	Local Breakout
Dual stack	IPv4-only	Failure case 1	Failure case 1
Dual stack	IPv4-only/IPv6-only	Failure case 1	Failure case 2
Dual stack	IPv6-only	Failure case 1	Failure case 3
IPv6-only	IPv4-only	OK	Failure case 4
IPv6-only with 464xlat	Dual stack	OK	Failure case 5
IPv6-only with 464xlat	IPv6-only	OK	OK
IPv4-only	Dual stack	OK	OK

Table 1: Roaming Scenario Descriptions

#### 4. Failure Cases Descriptions

##### 4.1. Failure Case 1: Incompatible with Extended PDP/PDN Type

A mobile device in a dual-stack network likely requests PDP/PDN type IPv4v6 to allocate address. Such PDP/PDN type should be understandable in the network nodes, including Serving GPRS Support Node(SGSN), Gateway GPRS Support Node(GGSN), Mobility Management Entity (MME), Serving Gateway(SGW), PDN Gateway(PGW), Home Location Registrar(HLR) and Home Subscriber Server(HSS). When a subscriber roams to the IPv4 network, the visited SGSN or MME has to communicate with HLR/HSS in the home land to retrieve the subscriber profile. The issue we observe is that multiple SGSN/MME will be unable to correctly process a subscriber profile received in the Insert Subscriber Data procedure if it contains an Ext-PDP-Type defined in

3GPP [TS29.002]. Therefore, it will likely refuse the subscriber registration.

Operators may have to remove the PDP/PDN type IPv4v6 from HLR/HSS in home networks, that will restrict UEs only initiates IPv4 PDP or IPv6 PDP activation. In order to avoid this situation, operators should make a comprehensive roaming agreement to support IPv6 and ensure that aligns with GSMA document, e.g [IR.33], [IR.88] and [IR.21]. Since the agreement requires visited operators to upgrade all SGSN nodes, some short- or medium-term solutions have been implemented to fix the issue. There are some specific configurations in HLR/HSS of home network. Multiple PDP/PDN subscription information will be added in the subscriber profiles, for example it may include both PDP/PDN type IPv4 and PDP/PDN type IPv4v6 for a user profile. Once the HLR/HSS receives an Update Location message from visited SGSN/MME, only the subscription data with PDP/PDN type IPv4 will be sent to SGSN/MME in the Insert Subscriber Data procedure. It guarantee the user profile could compatible with visited SGSN/MME capability.

#### 4.2. Failure Case 2: Splitting Dual-stack Bearer

Dual-stack capability can also be provided in a early mobile network(i.e. Pre-Release 8 network) using separate PDP/PDN activations. That means only a single IPv4 and IPv6 PDP/PDN can be initiated to allocate IPv4 and IPv6 address separately. Once a UE with PDP/PDN type IPv4v6 request roams to those networks, same issue described in failure case 1 will be occurred if the UE initiate a network attachment process.

If networks could allow UE to make a success attachment, a roaming subscriber with IPv4v6 PDP/PDN type should change the request to two separated PDP/PDN request with single IP version in order to achieve equivalent results. This restriction may be occurred in the below two cases.

- o The GGSN/PGW preferences dictate the use of IPv4 addressing only or IPv6 prefix only for a specific APN.
- o The SGSN/MME does not set the Dual Address Bearer Flag due to the operator using single addressing per bearer to support interworking with nodes of earlier releases

Above process would likely double PDP/PDN allocation costs. Some operators may only allow one PDP/PDN is alive for each subscriber. For example, IPv6 PDP/PDN would be rejected if the subscriber has an active IPv4 PDP/PDN. Therefore, the subscriber will lost IPv6 connection in the visited network. Even the two parallel PDP/PDN activations are allowed, it will require additional correlation of

those two sessions of single IP version on the charging system. If there are Policy and Charging Rules Function(PCRF)/Policy and Charging Enforcement Function (PCEF) deployed, the system would treat IPv4 and IPv6 session as independent and perform different Quality of Service(QoS) policies. The subscriber may have unstable experiences due to different behaviors on each IP version connection.

#### 4.3. Failure Case 3: Shortage of IPv6 support

Some operators may adopt IPv6-only configuration for the IMS service, e.g. Voice over LTE(VoLTE) or Rich Communication Suite (RCS). Since IMS roaming architecture will offload all traffic in the visited network, a dual-stack subscriber can only be assigned with IPv6 address. There is no IPv4 address returned. It requires all the IMS based applications should be IPv6 enable. A translation-based method, for example Bump-in-the-host (BIH)[RFC6535] and 464xlat[RFC6877] , may help to address the issue if there is IPv6 compatibility problems. Operators may could automatically enable the function in a IPv6-only network and disable in a dual-stack or IPv4 network.

#### 4.4. Failure Case 4: Fallback Incapability

3GPP specified the PDP/PDN type IPv6 as early as PDP/PDN type IPv4. Therefore, the IPv6 single PDP/PDN type has been well supported and interpretable in the 3GPP network nodes. When a subscriber requests PDP/PDN type IPv6, the network should only return the expected IPv6 address. Otherwise, the request should be dropped and the error code should be sent to the user. Roaming to IPv4-only networks with IPv6 PDP/PDN request would fail to get addresses. A proper fallback is desirable however the behavior is implementation specific. There are some UE have the ability to provide a different configuration for home network and visited network respectively. It guarantees UE will always initiate PDP/PDN type IPv4 in the roaming area. Android system solves the issue by setting the roaming Access Point Name(APN). The mobile terminal is allowed to ignore the original requested protocol and always adhere to IPv4 when roaming. Those fallback mechanisms are deserved to be implemented timely.

#### 4.5. Failure Case 5: 464xlat Support

464xlat[RFC6877] is proposed to address IPv4 compability issue in a IPv6 single-stack environment. The function on a mobile terminal likely gets along with PDP/PDN IPv6 type request to cooperate with a remote NAT64[RFC6146] gateway. 464xlat may use the mechanism defined in [I-D.ietf-behave-nat64-discovery-heuristic] to automatically discover NAT64 prefixes. Those behaviors depend on the network deployment. If the DNS64 or NAT64 is not deployed in the visited

networks, 464xlat may be failed to perform. Considering the various network's situations, operators may adopt 464xlat in the home networks and use IPv4-only in the roaming networks with different roaming profile configurations.

As an alternative solution, an AAA Server could be deployed to connect with GGSN/PGW. Once the GGSN/PGW receive the session creation requests, it will initiate an Access-Request to an AAA server via Radius protocol. The Access-Request contains subscriber and visited network information, e.g. PDP/PDN Type, International Mobile Equipment Id (IMEI), Software Version(SV) and visited SGSN/MME location code, etc. The AAA server could take IMEI and SV components to verify if device has 464XLAT support. Combining with the visited network information, the AAA server will ultimately decide to enable 464xlat in an IPv6-only roaming or fallback to IPv4.

## 5. Discussions

The dual-stack deployment is recommended in most cases. However, it may take some times in a mobile environment. 3GPP didn't specify PDP/PDN type IPv4v6 in the early release. Such PDP/PDN type is supported in new-built Long Term Evolution(LTE)/System Architecture Evolution(SAE) network, but didn't support well in the third generation network. The situations may cause the roaming issues dropping the attachment from dual-stack subscribers in the case of LTE to 3G and IPv6-enabled 3G to IPv4 3G. Operators may have to adopt temporary solution unless all the interworking nodes(i.e. SSGN and SGW) in the visited network have been upgraded to support Ext-PDP-Type feature.

As an alternative solution for dual-stack, operators may change a unified PDP/PDN request into two separated single IP version requests. However, this approach is problematic in the Charging records and QoS policy enforcement. In addition, it doubles the PDP resource uses. It may be unappealing for the deployment.

Conversely, some operators may choose PDP/PDN Type IPv6 to start the communications in home networks and use different profile in the roaming area. Since PDP/PDN Type IPv6 has been introduced in 3GPP early release, it didn't require upgrading on the interworking nodes to make compatibility. The proper IPv4 fallback mechanism should be supported either on the mobile terminal or network equipment.

A roaming to IPv6-only network occurs when operators deploy roaming function for IMS service. A dual-stack capable device could implement translation-based function to support the IPv4 applications. Those inserted translation function can be turned off properly when the terminals roam back to dual-stack or IPv4 networks. Operators can also deploy AAA servers to make final decision

## 6. IANA Considerations

This document makes no request of IANA.

## 7. Security Considerations

The draft didn't introduce additional security concerns to the networks.

## 8. Acknowledgements

The authors would like to thank V6ops chairs(Fred Baker and John Brzozowski) to encourage us to continue the work. This document is the result of the IETF V6ops IPv6-Roaming design team effort.

## 9. References

### 9.1. Normative References

- [I-D.ietf-behave-nat64-discovery-heuristic]  
Savolainen, T., Korhonen, J., and D. Wing, "Discovery of the IPv6 Prefix Used for IPv6 Address Synthesis", draft-ietf-behave-nat64-discovery-heuristic-17 (work in progress), April 2013.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC6052] Bao, C., Huitema, C., Bagnulo, M., Boucadair, M., and X. Li, "IPv6 Addressing of IPv4/IPv6 Translators", RFC 6052, October 2010.
- [RFC6146] Bagnulo, M., Matthews, P., and I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", RFC 6146, April 2011.
- [RFC6147] Bagnulo, M., Sullivan, A., Matthews, P., and I. van Beijnum, "DNS64: DNS Extensions for Network Address Translation from IPv6 Clients to IPv4 Servers", RFC 6147, April 2011.

- [RFC6535] Huang, B., Deng, H., and T. Savolainen, "Dual-Stack Hosts Using "Bump-in-the-Host" (BIH)", RFC 6535, February 2012.
- [RFC6877] Mawatari, M., Kawashima, M., and C. Byrne, "464XLAT: Combination of Stateful and Stateless Translation", RFC 6877, April 2013.

## 9.2. Informative References

- [IR.21] Global System for Mobile Communications Association, GSMA., "Roaming Database, Structure and Updating Procedures", July 2012.
- [IR.33] Global System for Mobile Communications Association, GSMA., "GPRS Roaming Guidelines", July 2012.
- [IR.65] Global System for Mobile Communications Association, GSMA., "IMS Roaming & Interworking Guidelines", May 2012.
- [IR.88] Global System for Mobile Communications Association, GSMA., "LTE Roaming Guidelines", January 2012.
- [RFC6052] Bao, C., Huitema, C., Bagnulo, M., Boucadair, M., and X. Li, "IPv6 Addressing of IPv4/IPv6 Translators", RFC 6052, October 2010.
- [RFC6586] Arkko, J. and A. Keranen, "Experiences from an IPv6-Only Network", RFC 6586, April 2012.
- [TR23.975] 3rd Generation Partnership Project, 3GPP., "IPv6 migration guidelines", June 2011.
- [TS23.401] 3rd Generation Partnership Project, 3GPP., "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access v9.00", March 2009.
- [TS29.002] 3rd Generation Partnership Project, 3GPP., "Mobile Application Part (MAP) specification v9.00", December 2009.
- [TS29.272]

3rd Generation Partnership Project, 3GPP., "Mobility Management Entity (MME) and Serving GPRS Support Node (SGSN) related interfaces based on Diameter protocol v9.00", September 2009.

Authors' Addresses

Gang Chen  
China Mobile  
53A,Xibianmennei Ave.,  
Xuanwu District,  
Beijing 100053  
China

Email: phdgang@gmail.com

Hui Deng  
China Mobile  
53A,Xibianmennei Ave.,  
Xuanwu District,  
Beijing 100053  
China

Email: denghui@chinamobile.com

Dave Michaud  
Rogers

Email: Michaud@rci.rogers.com

Jouni Korhonen  
Renesas Mobile  
Porkkalankatu 24  
FIN-00180 Helsinki, Finland

Email: jouni.nospam@gmail.com

Mohamed Boucadair  
France Telecom  
No.32 Xuanwumen West Street  
Rennes,  
35000  
France

Email: mohamed.boucadair@orange.com

Vizdal Ales  
Deutsche Telekom AG  
Tomickova 2144/1  
Prague 4, 149 00  
Czech Republic

Email: ales.vizdal@t-mobile.cz

Cameron Byrne  
T-Mobile USA  
Bellevue  
Washington 98105  
USA

Email: cameron.byrne@t-mobile.com



IPv6 Operations  
Internet-Draft  
Intended status: Informational  
Expires: June 8, 2014

M. Gysi  
Swisscom  
G. Leclanche  
Viagenie  
E. Vyncke, Ed.  
Cisco Systems  
R. Anfinson  
Altibox  
December 05, 2013

Balanced Security for IPv6 Residential CPE  
draft-ietf-v6ops-balanced-ipv6-security-01

Abstract

This document describes how an IPv6 residential Customer Premise Equipment (CPE) can have a balanced security policy that allows for a mostly end-to-end connectivity while keeping the major threats outside of the home. It is documenting an existing IPv6 deployment by Swisscom and allows all packets inbound/outbound EXCEPT for some layer-4 ports where attacks and vulnerabilities (such as weak passwords) are well-known. The policy is a proposed set of rules that can be used as a default setting. The set of blocked inbound and outbound ports is expected to be updated as threats come and go.

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 June 8, 2014.

## Copyright Notice

Copyright (c) 2013 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 . . . . .	2
2. Threats . . . . .	3
3. Overview . . . . .	3
3.1. Rules for Balanced Security Policy . . . . .	4
3.2. Rules Example for Layer-4 Protection: Swisscom Implementation . . . . .	5
4. IANA Considerations . . . . .	6
5. Security Considerations . . . . .	6
6. Acknowledgements . . . . .	7
7. Informative References . . . . .	7
Authors' Addresses . . . . .	8

## 1. Introduction

Internet access in residential IPv4 deployments generally consists of a single IPv4 address provided by the service provider for each home. The residential CPE then translates the single address into multiple private IPv4 addresses allowing more than one device in the home, but at the cost of losing end-to-end reachability. IPv6 allows all devices to have a globally unique IP address, restoring end-to-end reachability directly between any device. Such reachability is very powerful for ubiquitous global connectivity, and is often heralded as one of the significant advantages to IPv6 over IPv4. Despite this, concern about exposure to inbound packets from the IPv6 Internet (which would otherwise be dropped by the address translation function if they had been sent from the IPv4 Internet) remain.

This difference in residential default internet protection between IPv4 and IPv6 is a major concern to a sizable number of ISPs and the security policy described in this document addresses this concern without damaging IPv6 end-to-end connectivity.

The security model provided in this document is meant to be used as a pre-registered setting and potentially default one for IPv6 security in CPEs. The model departs from the "simple security" model described in [RFC6092]. It allows most traffic, including incoming unsolicited packets and connections, to traverse the CPE unless the CPE identifies the traffic as potentially harmful based on a set of rules. This policy has been deployed as a default setting in Switzerland by Swisscom for residential CPEs.

This document can be applicable to off-the-shelves CPE as well as to managed Service Provider CPE or for mobile Service Providers (where it can be centrally implemented).

## 2. Threats

For a typical residential network connected to the Internet over a broadband or mobile connection, the threats can be classified into:

- o denial of service by packet flooding: overwhelming either the access bandwidth or the bandwidth of a slower link in the residential network (like a slow home automation network) or the CPU power of a slow IPv6 host (like networked thermostat or any other sensor type nodes);
- o denial of service by Neighbor Discovery cache exhaustion [RFC6583]: the outside attacker floods the inside prefix(es) with packets with a random destination address forcing the CPE to exhaust its memory and its CPU in useless Neighbor Solicitations;
- o denial of service by service requests: like sending print jobs from the Internet to an ink jet printer until the ink cartridge is empty or like filing some file server with junk data;
- o unauthorized use of services: like accessing a webcam or a file server which are open to anonymous access within the residential network but should not be accessed from outside of the home network or accessing to remote desktop or SSH with weak password protection;
- o exploiting a vulnerability in the host in order to get access to data or to execute some arbitrary code in the attacked host;
- o trojanized host (belonging to a Botnet) can communicate via a covert channel to its master and launch attacks to Internet targets.

## 3. Overview

The basic goal is to provide a pre-defined security policy which aims to block known harmful traffic and allow the rest, restoring as much of end-to-end communication as possible. This pre-defined policy should be centrally updated, as threats are changing over time. It could also be a member of a list of pre-defined security policies available to an end-customer, for example together with "simple security" from [RFC6092] and a "strict security" policy denying access to all unexpected input packets.

### 3.1. Rules for Balanced Security Policy

These are an example set of generic rules to be applied. Each would normally be configurable, either by the user directly or on behalf of the user by a subscription service. This document does not address the statefulness of the filtering rules as its main objective is to present an approach where some protocols (identified by layer-4 ports) are assumed weak or malevolent and therefore are blocked while all other protocols are assumed benevolent and are permitted.

If we name all nodes on the residential side of the CPE as 'inside' and all nodes on the Internet as 'outside', and any packet sent from outside to inside as being 'inbound' and 'outbound' in the other direction, then the behavior of the CPE is described by a small set of rules:

1. Rule RejectBogon: apply ingress filtering in both directions per [RFC3704] and [RFC2827] for example with unicast reverse path forwarding (uRPF) checks (anti-spoofing) for all inbound and outbound traffic (implicitly blocking link-local and ULA in the same shot), as described in Section 2.1 Basic Sanitation and Section 3.1 Stateless Filters of [RFC6092];
2. Rule AllowManagement: if the CPE is managed by the SP, then allow the management protocols (SSH, SNMP, syslog, TR-069, IPfix, ...) from/to the SP Network Operation Center;
3. Rule ProtectWeakServices: drop all inbound and outbound packets whose layer-4 destination is part of a limited set (see Section 3.2), the intent is to protect against the most common unauthorized access and avoid propagation of worms; an advanced residential user should be able to modify this pre-defined list;

4. Rule Openess: allow all unsolicited inbound packets with rate limiting the initial packet of a new connection (such as TCP SYN, SCTP INIT or DCCP-request, not applicable to UDP) to provide very basic protection against SYN port and address scanning attacks. All transport protocols and all non-deprecated extension headers are accepted. This is a the major deviation from REC-11, REC-17 and REC-33 of [RFC6092].
5. All requirements of [RFC6092] except REC-11, REC-18 and REC-33 must be supported.

### 3.2. Rules Example for Layer-4 Protection: Swisscom Implementation

As of 2013, Swisscom has implemented the rule ProtectWeakService as described below. This is meant as an example and must not be followed blindly: each implementer has specific needs and requirements. Furthermore, the example below will not be updated as time passes, whereas threats will evolve.

Transport	Port	Description
tcp	22	Secure Shell (SSH)
tcp	23	Telnet
tcp	80	HTTP
tcp	3389	Microsoft Remote Desktop Protocol
tcp	5900	VNC remote desktop protocol

Table 1: Drop Inbound

Transport	Port	Description
tcp-udp	88	Kerberos
tcp	111	SUN Remote Procedure Call
tcp	135	MS Remote Procedure Call
tcp	139	NetBIOS Session Service
tcp	445	Microsoft SMB Domain Server
tcp	513	Remote Login
tcp	514	Remote Shell
tcp	548	Apple Filing Protocol over TCP
tcp	631	Internet Printing Protocol
udp	1900	Simple Service Discovery Protocol
tcp	2869	Simple Service Discovery Protocol
udp	3702	Web Services Dynamic Discovery
udp	5353	Multicast DNS
udp	5355	Link-Lcl Mcast Name Resolution

-----+

Table 2: Drop Inbound and Outbound

Choosing services to protect is not an easy task, and as of 2013 there is no public service proposing a list of ports to use in such a policy. The Swisscom approach was to think in terms of services, by defining a list of services that are LAN-Only (ex: Multicast DNS) whose communication is denied by the policy both inbound and outbound, and a list of services that are known to be weak or vulnerable like management protocols that could be activated unbeknownst to the user.

The process used to set-up and later update the filters is out of scope of this document. The update of the specific rules could be done together with a firmware upgrade or by a policy update (for example using Broadband Forum TR-069).

Among other sources, [DSHIELD] was used by Swisscom to set-up their filters. Another source of information could be the appendix A of [TR124]. The L4-filter as described does not block GRE tunnels ([RFC2473]) so this is a deviation from [RFC6092].

Note: the authors believe that with a dozen of rules only, a naive and unaware residential subscriber would be reasonably protected. Of course, technically-aware subscribers should be able to open other applications (identified by their layer-4 ports or IP protocol numbers) through their CPE using some kind of user interface or even to select a completely different security policy such as the open or 'closed' policies defined by [RFC6092]. This is the case in the Swisscom deployment.

It is worth mentioning that PCP ([RFC6887]), UPnP ([IGD]) and similar protocols can also be used to dynamically override the default rules.

## 4. IANA Considerations

There are no extra IANA consideration for this document.

## 5. Security Considerations

The security policy protects from the following type of attacks:

- o Unauthorized access because vulnerable ports are blocked

Depending on the extensivity of the filters, certain vulnerabilities could be protected or not. It does not preclude the need for end-devices to have proper host-protection as most of those devices

(smartphones, laptops, etc.) would anyway be exposed to completely unfiltered internet at some point of time. The policy addresses the major concerns related to the loss of stateful filtering imposed by IPV4 NAT when enabling public globally reachable IPv6 in the home.

To the authors' knowledge, there has not been any incident related to this deployment in Swisscom network, and no customer complaints have been registered.

This set of rules cannot help with the following attacks:

- o Flooding of the CPE access link;
- o Malware which is fetched by inside hosts on a hostile web site (which is in 2013 the majority of infection sources).

## 6. Acknowledgements

The authors would like to thank several people who initiated the discussion on the `ipv6-ops@lists.cluonet.de` mailing list and others who provided us valuable feedback and comments, notably: Tore Anderson, Rajiv Asati, Fred Baker, Lorenzo Colitti, Paul Hoffman, Merike Kaeo, Simon Leinen, Eduard Metz, Martin Millnert, Benedikt Stockebrand. Thanks as well to the following SP that discussed with the authors about this technique: Altibox, Swisscom and Telenor.

## 7. Informative References

- [DSHIELD] DShield, "Port report: DShield", <<https://secure.dshield.org/portreport.html?sort=records>>.
- [IGD] UPnP Forum, "WANIPConnection:2 Service", December 20110, <<http://upnp.org/specs/gw/UPnP-gw-WANIPConnection-v2-Service.pdf>>.
- [RFC2473] Conta, A. and S. Deering, "Generic Packet Tunneling in IPv6 Specification", RFC 2473, December 1998.
- [RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, May 2000.
- [RFC3704] Baker, F. and P. Savola, "Ingress Filtering for Multihomed Networks", BCP 84, RFC 3704, March 2004.

- [RFC6092] Woodyatt, J., "Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service", RFC 6092, January 2011.
- [RFC6583] Gashinsky, I., Jaeggli, J., and W. Kumari, "Operational Neighbor Discovery Problems", RFC 6583, March 2012.
- [RFC6887] Wing, D., Cheshire, S., Boucadair, M., Penno, R., and P. Selkirk, "Port Control Protocol (PCP)", RFC 6887, April 2013.
- [TR124] Broadband Forum, "Functional Requirements for Broadband Residential Gateway Devices", December 2006, <<http://www.broadband-forum.org/technical/download/TR-124.pdf>>.

## Authors' Addresses

Martin Gysi  
Swisscom  
Binzring 17  
Zuerich 8045  
Switzerland

Phone: +41 58 223 57 24  
Email: [Martin.Gysi@swisscom.com](mailto:Martin.Gysi@swisscom.com)

Guillaume Leclanche  
Viagenie  
246 Aberdeen  
Quebec, QC G1R 2E1  
Canada

Phone: +1 418 656 9254  
Email: [guillaume.leclanche@viagenie.ca](mailto:guillaume.leclanche@viagenie.ca)

Eric Vyncke (editor)  
Cisco Systems  
De Kleetlaan 6a  
Diegem 1831  
Belgium

Phone: +32 2 778 4677  
Email: [evyncke@cisco.com](mailto:evyncke@cisco.com)



Ragnar Anfinssen  
Altibox  
Breiflaaaveien 18  
Stavanger 4069  
Norway

Phone: +47 93488235  
Email: Ragnar.Anfinssen@altibox.no

Network Working Group  
Internet-Draft  
Intended status: Informational  
Expires: June 19, 2016

D. Binet  
M. Boucadair  
Orange  
A. Vizdal  
Deutsche Telekom AG  
G. Chen  
China Mobile  
N. Heatley  
EE  
R. Chandler  
eircom | meteor  
D. Michaud  
Rogers Communications  
D. Lopez  
Telefonica I+D  
W. Haeffner  
Vodafone  
December 17, 2015

An Internet Protocol Version 6 (IPv6) Profile for 3GPP Mobile Devices  
draft-ietf-v6ops-mobile-device-profile-24

## Abstract

This document defines a profile that is a superset of that of the connection to IPv6 cellular networks defined in the IPv6 for Third Generation Partnership Project (3GPP) Cellular Hosts document. This document defines an IPv6 profile that a number of operators recommend in order to connect 3GPP mobile devices to an IPv6-only or dual-stack wireless network (including 3GPP cellular network) with a special focus on IPv4 service continuity features.

Both mobile hosts and mobile devices with capability to share their 3GPP mobile connectivity are in scope.

## 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 June 19, 2016.

#### Copyright Notice

Copyright (c) 2015 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 . . . . .	2
1.1. Terminology . . . . .	4
1.2. Scope . . . . .	4
2. Connectivity Recommendations . . . . .	6
3. Recommendations for Cellular Devices with LAN Capabilities .	10
4. Advanced Recommendations . . . . .	12
5. Security Considerations . . . . .	14
6. IANA Considerations . . . . .	15
7. Acknowledgements . . . . .	15
8. References . . . . .	15
8.1. Normative References . . . . .	15
8.2. Informative References . . . . .	17
Authors' Addresses . . . . .	20

#### 1. Introduction

IPv6 deployment in Third Generation Partnership Project (3GPP) mobile networks is the only viable solution to the exhaustion of IPv4 addresses in those networks. Several mobile operators have already deployed IPv6 [RFC2460] or are in the pre-deployment phase. One of the major hurdles as perceived by some mobile operators is the lack of availability of working IPv6 implementation in mobile devices (e.g., Section 3.3 of [OECD]).

[RFC7066] lists a set of features to be supported by cellular hosts to connect to 3GPP mobile networks. In the light of recent IPv6

production deployments, additional features to facilitate IPv6-only deployments while accessing IPv4-only services should be considered. This document fills this void. Concretely, this document lists means to ensure IPv4 service over an IPv6-only connectivity given the adoption rate of this model by mobile operators. Those operators require that no service degradation is experienced by customers serviced with an IPv6-only model compared to the level of service of customers with legacy IPv4-only devices.

This document defines an IPv6 profile for mobile devices listing specifications produced by various Standards Developing Organizations (including 3GPP, IETF, and GSMA). The objectives of this effort are:

1. List in one single document a comprehensive list of IPv6 features for a mobile device, including both IPv6-only and dual-stack mobile deployment contexts. These features cover various packet core architectures such as GPRS (General Packet Radio Service) or EPC (Evolved Packet Core).
2. Help Operators with the detailed device requirement list preparation (to be exchanged with device suppliers). This is also a contribution to harmonize Operators' requirements towards device vendors.
3. Vendors to be aware of a set of features to allow for IPv6 connectivity and IPv4 service continuity (over an IPv6-only transport).

The recommendations do not include 3GPP release details. For more information on the 3GPP releases detail, the reader may refer to Section 6.2 of [RFC6459]. More details can be found at [R3GPP].

Some of the features listed in this profile document could require to activate dedicated functions at the network side. It is out of scope of this document to list these network-side functions.

A detailed overview of IPv6 support in 3GPP architectures is provided in [RFC6459]. IPv6-only considerations in mobile networks are further discussed in [RFC6342].

This document is organized as follows:

- o Section 2 lists generic recommendations including functionalities to provide IPv4 service over an IPv6-only connectivity.
- o Section 3 enumerates a set of recommendations for cellular devices with Local Area Network (LAN) capabilities (e.g., CE routers

(Customer Edge routers) with cellular access link, dongles with tethering features).

- o Section 4 identifies a set of advanced recommendations to fulfill requirements of critical services such as VoLTE (Voice over Long Term Evolution (LTE)).

### 1.1. Terminology

This document makes use of the terms defined in [RFC6459]. In addition, the following terms are used:

- o 3GPP cellular host (or cellular host for short): denotes a 3GPP device which can be connected to 3GPP mobile networks.
- o 3GPP cellular device (or cellular device for short): refers to a cellular host which supports the capability to share its 3GPP mobile connectivity.
- o IPv4 service continuity: denotes the features used to provide access to IPv4-only services to customers serviced with an IPv6-only connectivity. A typical example of IPv4 service continuity technique is NAT64 (Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers, [RFC6146]).

PREFIX64 denotes an IPv6 prefix used to build IPv4-converted IPv6 addresses [RFC6052].

### 1.2. Scope

A 3GPP mobile network can be used to connect various user equipments such as a mobile telephone or a CE router. Because of this diversity of terminals, it is necessary to define a set of IPv6 functionalities valid for any node directly connecting to a 3GPP mobile network. This document describes these functionalities.

Machine-to-machine (M2M) devices profile is out of scope.

This document is structured to provide the generic IPv6 recommendations which are valid for all nodes, whatever their function (e.g., host or CE router) or service (e.g., Session Initiation Protocol (SIP, [RFC3261])) capability. The document also contains sections covering specific functionalities for devices providing some LAN functions (e.g., mobile CE router or broadband dongles).

The recommendations listed below are valid for both 3GPP GPRS and 3GPP EPS (Evolved Packet System). For EPS, PDN-Connection term is

used instead of PDP-Context. Other non-3GPP accesses [TS.23402] are out of scope of this document.

This profile is a superset of that of the IPv6 profile for 3GPP Cellular Hosts [RFC7066], which is in turn a superset of IPv6 Node Requirements [RFC6434]. It targets cellular nodes, including GPRS and EPC (Evolved Packet Core), that require features to ensure IPv4 service delivery over an IPv6-only transport in addition to the base IPv6 service. Moreover, this profile also covers cellular CE routers that are used in various mobile broadband deployments. Recommendations inspired from real deployment experiences (e.g., roaming) are included in this profile. Also, this profile sketches recommendations for the sake of deterministic behaviors of cellular devices when the same configuration information is received over several channels.

For conflicting recommendations in [RFC7066] and [RFC6434] (e.g., Neighbor Discovery Protocol), this profile adheres to [RFC7066]. Indeed, the support of Neighbor Discovery Protocol is mandatory in 3GPP cellular environment as it is the only way to convey IPv6 prefix towards the 3GPP cellular device. In particular, MTU (Maximum Transmission Unit) communication via Router Advertisement must be supported since many 3GPP networks do not have a standard MTU setting.

This profile uses a stronger language for the support of Prefix Delegation compared to [RFC7066]. The main motivation is that cellular networks are more and more perceived as an alternative to fixed networks for home IP-based services delivery; especially with the advent of smartphones and 3GPP data dongles. There is a need for an efficient mechanism to assign larger prefixes to cellular hosts so that each LAN segment can get its own /64 prefix and multi-link subnet issues to be avoided. The support of this functionality in both cellular and fixed networks is key for fixed-mobile convergence.

The use of address family dependent Application Programming Interfaces (APIs) or hard-coded IPv4 address literals may lead to broken applications when IPv6 connectivity is in use. As such, means to minimize broken applications when the cellular host is attached to an IPv6-only network should be encouraged. Particularly, (1) name resolution libraries (e.g., [RFC3596]) must support both IPv4 and IPv6; (2) applications must be independent of the underlying IP address family; (3) and applications relying upon Uniform Resource Identifiers (URIs) must follow [RFC3986] and its updates. Note, some IETF specifications (e.g., SIP [RFC3261]) contains broken IPv6 Augmented Backus-Naur Form (ABNF) and rules to compare URIs with embedded IPv6 addresses; fixes (e.g., [RFC5954]) must be used instead.

The recommendations included in each section are listed in a priority order.

This document is not a standard, and conformance with it is not required in order to claim conformance with IETF standards for IPv6. Compliance with this profile does not require the support of all enclosed items. Obviously, the support of the full set of features may not be required in some deployment contexts. However, the authors believe that not supporting relevant features included in this profile (e.g., Customer Side Translator (CLAT, [RFC6877])) may lead to a degraded level of service.

## 2. Connectivity Recommendations

This section identifies the main connectivity recommendations to be followed by a cellular host to attach to a network using IPv6 in addition to what is defined in [RFC6434] and [RFC7066]. Both dual-stack and IPv6-only deployment models are considered. IPv4 service continuity features are listed in this section because these are critical for Operators with an IPv6-only deployment model. These recommendations apply also for cellular devices (see Section 3).

C\_REC#1: In order to allow each operator to select their own strategy regarding IPv6 introduction, the cellular host must support both IPv6 and IPv4v6 PDP-Contexts [TS.23060].

IPv4, IPv6 or IPv4v6 PDP-Context request acceptance depends on the cellular network configuration.

C\_REC#2: The cellular host must comply with the behavior defined in [TS.23060] [TS.23401] [TS.24008] for requesting a PDP-Context type.

In particular, the cellular host must request by default an IPv6 PDP-Context if the cellular host is IPv6-only and request an IPv4v6 PDP-Context if the cellular host is dual-stack or when the cellular host is not aware of connectivity types requested by devices connected to it (e.g., cellular host with LAN capabilities as discussed in Section 3):

- \* If the requested IPv4v6 PDP-Context is not supported by the network, but IPv4 and IPv6 PDP types are allowed, then the cellular host will be configured with an IPv4 address or an IPv6 prefix by the network. It must initiate another PDP-Context activation of the other address family in addition to the one already activated for a given APN (Access Point Name). The purpose of

initiating a second PDP-Context is to achieve dual-stack connectivity by means of two PDP-Contexts.

- \* If the subscription data or network configuration allows only one IP address family (IPv4 or IPv6), the cellular host must not request a second PDP-Context to the same APN for the other IP address family.

The network informs the cellular host about allowed PDP types by means of Session Management (SM) cause codes. In particular, the following cause codes can be returned:

- \* cause #50 "PDP type IPv4 only allowed". This cause code is used by the network to indicate that only PDP type IPv4 is allowed for the requested PDN connectivity.
- \* cause #51 "PDP type IPv6 only allowed". This cause code is used by the network to indicate that only PDP type IPv6 is allowed for the requested PDN connectivity.
- \* cause #52 "single address bearers only allowed". This cause code is used by the network to indicate that the requested PDN connectivity is accepted with the restriction that only single IP version bearers are allowed.

The text above focuses on the specification (excerpt from [TS.23060] [TS.23401] [TS.24008]) which explains the behavior for requesting IPv6-related PDP-Context(s).

C\_REC#3: The cellular host must support the PCO (Protocol Configuration Options) [TS.24008] to retrieve the IPv6 address(es) of the Recursive DNS server(s).

The 3GPP network communicates parameters by means of the protocol configuration options information element when activating, modifying or deactivating a PDP-Context. PCO is a convenient method to inform the cellular host about various services, including DNS server information. It does not require additional protocol to be supported by the cellular host and it is already deployed in IPv4 cellular networks to convey such DNS information.

C\_REC#4: The cellular host must support IPv6 aware Traffic Flow Templates (TFT) [TS.24008].



Traffic Flow Templates are employing a packet filter to couple an IP traffic with a PDP-Context. Thus a dedicated PDP-Context and radio resources can be provided by the cellular network for certain IP traffic.

C\_REC#5: If the cellular host receives the DNS information in several channels for the same interface, the following preference order must be followed:

1. PCO
2. RA
3. DHCPv6

The purpose of this recommendation is to guarantee for a deterministic behavior to be followed by all cellular hosts when the DNS information is received in various channels.

C\_REC#6: Because of potential operational deficiencies to be experienced in some roaming situations, the cellular host must be able to be configured with a home PDP-Context type(s) and a roaming PDP-Context type(s). The purpose of the roaming profile is to limit the PDP type(s) requested by the cellular host when out of the home network. Note that distinct PDP type(s) and APN(s) can be configured for home and roaming cases.

A detailed analysis of roaming failure cases is included in [RFC7445].

The configuration can be either local to the device or be managed dynamically using, for example, Open Mobile Alliance (OMA) management. The support of dynamic means is encouraged.

C\_REC#7: In order to ensure IPv4 service continuity in an IPv6-only deployment context, the cellular host should support a method to learn PREFIX64(s).

In the context of NAT64, IPv6-enabled applications relying on address referrals will fail because an IPv6-only client will not be able to make use of an IPv4 address received in a referral. This feature allows to solve the referral problem (because an IPv6-enabled application can construct IPv4-embedded IPv6 addresses [RFC6052]) and, also, to distinguish between IPv4-converted IPv6 addresses and native IPv6 addresses.

In other words, this feature contributes to offload both CLAT module and NAT64 devices. Refer to Section 3 of [RFC7051] for an inventory of the issues related to the discovery of PREFIX64(s).

In PCP-based environments, cellular hosts should follow [RFC7225] to learn the IPv6 Prefix used by an upstream PCP-controlled NAT64 device. If PCP is not enabled, the cellular host should implement the method specified in [RFC7050] to retrieve the PREFIX64.

C\_REC#8: In order to ensure IPv4 service continuity in an IPv6-only deployment context, the cellular host should implement the Customer Side Translator (CLAT, [RFC6877]) function in compliance with [RFC6052][RFC6145][RFC6146].

CLAT function in the cellular host allows for IPv4-only application and IPv4-referrals to work on an IPv6-only connectivity. The more applications are address family independent, the less CLAT function is solicited. CLAT function requires a NAT64 capability [RFC6146] in the network.

The cellular host should only invoke the CLAT in the absence of the IPv4 connectivity on the cellular side, i.e., when the network does not assign an IPv4 address on the cellular interface. Note, NAT64 assumes an IPv6-only mode [RFC6146].

The IPv4 Service Continuity Prefix used by CLAT is defined in [RFC7335].

CLAT and/or NAT64 do not interfere with native IPv6 communications.

CLAT may not be required in some contexts, e.g., if other solutions such as Bump-in-the-Host (BIH, [RFC6535]) are supported.

The cellular device can act as a CE router connecting various IP hosts on a LAN segment; it is also the case with the use of WLAN (Wireless LAN) tethering or WLAN hotspot from the cellular device. Some of these IP hosts can be dual-stack, others are IPv6-only or IPv4-only. IPv6-only connectivity on the cellular device does not allow IPv4-only sessions to be established for hosts connected on the LAN segment of the cellular device. IPv4 session establishment

initiated from hosts located on LAN segment side and destined for IPv4 nodes must be maintained. A solution is to integrate the CLAT function to the LAN segment in the cellular device.

- C\_REC#9: The cellular host may be able to be configured to limit PDP type(s) for a given APN. The default mode is to allow all supported PDP types. Note, C\_REC#2 discusses the default behavior for requesting PDP-Context type(s).

This feature is useful to drive the behavior of the UE to be aligned with: (1) service-specific constraints such as the use of IPv6-only for VoLTE (Voice over LTE), (2) network conditions with regards to the support of specific PDP types (e.g., IPv4v6 PDP-Context is not supported), (3) IPv4 sunset objectives, (4) subscription data, etc.

Note, a cellular host changing its connection between an IPv6-specific APN and an IPv4-specific APN will interrupt related network connections. This may be considered as a brokenness situation by some applications.

The configuration can be either local to the device or be managed dynamically using, for example, Open Mobile Alliance (OMA) management. The support of dynamic means is encouraged.

### 3. Recommendations for Cellular Devices with LAN Capabilities

This section focuses on cellular devices (e.g., CE router, smartphones or dongles with tethering features) which provide IP connectivity to other devices connected to them. In such case, all connected devices are sharing the same 2G, 3G or LTE connection. In addition to the generic recommendations listed in Section 2, these cellular devices have to meet the recommendations listed below.

- L\_REC#1: For deployments requiring to share the same /64 prefix, the cellular device should support [RFC7278] to enable sharing a /64 prefix between the 3GPP interface towards the GGSN/PGW (WAN interface) and the LAN interfaces.

Prefix Delegation (refer to L\_REC#2) is the target solution for distributing prefixes in the LAN side but, because the device may attach to earlier 3GPP release networks, a mean to share a /64 prefix is also recommended [RFC7278].

[RFC7278] must be invoked only if Prefix Delegation is not in use.

- L\_REC#2: The cellular device must support Prefix Delegation capabilities [RFC3633] and must support Prefix Exclude Option for DHCPv6-based Prefix Delegation as defined in [RFC6603]. Particularly, it must behave as a Requesting Router.

Cellular networks are more and more perceived as an alternative to fixed broadband networks for home IP-based services delivery; especially with the advent of smartphones and 3GPP data dongles. There is a need for an efficient mechanism to assign larger prefixes (other than /64s) to cellular hosts so that each LAN segment can get its own /64 prefix and multi-link subnet issues to be avoided.

In case a prefix is delegated to a cellular host using DHCPv6, the cellular device will be configured with two prefixes:

- (1) one for 3GPP link allocated using stateless address autoconfiguration (SLAAC) mechanism and
- (2) another one delegated for LANs acquired during Prefix Delegation operation.

Note that the 3GPP network architecture requires both the WAN (Wide Area Network) and the delegated prefix to be aggregatable, so the subscriber can be identified using a single prefix.

Without the Prefix Exclude Option, the delegating router (GGSN/PGW) will have to ensure [RFC3633] compliancy (e.g., halving the delegated prefix and assigning the WAN prefix out of the 1st half and the prefix to be delegated to the terminal from the 2nd half).

Because Prefix Delegation capabilities may not be available in some attached networks, L\_REC#1 is strongly recommended to accommodate early deployments.

- L\_REC#3: The cellular CE router must be compliant with the requirements specified in [RFC7084].

There are several deployments, particularly in emerging countries, that relies on mobile networks to provide

broadband services (e.g., customers are provided with mobile CE routers).

Note, this profile does not require IPv4 service continuity techniques listed in Section 4.4 of [RFC7084] because those are specific to fixed networks. IPv4 service continuity techniques specific to the mobile networks are included in this profile.

This recommendation does not apply to handsets with tethering capabilities; it is specific to cellular CE routers in order to ensure the same IPv6 functional parity for both fixed and cellular CE routers. Note, modern CE routers are designed with advanced functions such as link aggregation that consists in optimizing the network usage by aggregating the connectivity resources offered via various interfaces (e.g., Digital Subscriber Line (DSL), LTE, WLAN, etc.) or offloading the traffic via a subset of interfaces. Ensuring IPv6 features parity among these interface types is important for the sake of specification efficiency, service design simplification and validation effort optimization.

L\_REC#4: If a RA MTU is advertised from the 3GPP network, the cellular device should send RAs to the downstream attached LAN devices with the same MTU as seen on the mobile interface.

Receiving and relaying RA MTU values facilitates a more harmonious functioning of the mobile core network where end nodes transmit packets that do not exceed the MTU size of the mobile network's GTP (GPRS Tunnelling Protocol) tunnels.

[TS.23060] indicates providing a link MTU value of 1358 octets to the 3GPP cellular device will prevent the IP layer fragmentation within the transport network between the cellular device and the GGSN/PGW. More details about link MTU considerations can be found in Annex C of [TS.23060].

#### 4. Advanced Recommendations

This section identifies a set of advanced recommendations to fulfill requirements of critical services such as VoLTE. These recommendations apply for mobile hosts, including mobile devices.

A\_REC#1: The cellular host must support ROHC RTP Profile (0x0001) and ROHC UDP Profile (0x0002) for IPv6 ([RFC5795]). Other ROHC profiles may be supported.

Bandwidth in cellular networks must be optimized as much as possible. ROHC provides a solution to reduce bandwidth consumption and to reduce the impact of having bigger packet headers in IPv6 compared to IPv4.

"RTP/UDP/IP" ROHC profile (0x0001) to compress RTP packets and "UDP/IP" ROHC profile (0x0002) to compress RTCP packets are required for Voice over LTE (VoLTE) by IR.92.4.0 section 4.1 [IR92]. Note, [IR92] indicates that the host must be able to apply the compression to packets that are carried over the voice media dedicated radio bearer.

A\_REC#2: The cellular host should support PCP [RFC6887].

The support of PCP is seen as a driver to save battery consumption exacerbated by keepalive messages. PCP also gives the possibility of enabling incoming connections to the cellular device. Indeed, because several stateful devices may be deployed in wireless networks (e.g., NAT64 and/or IPv6 Firewalls), PCP can be used by the cellular host to control network-based NAT64 and IPv6 Firewall functions which will reduce per-application signaling and save battery consumption.

According to [Power], the consumption of a cellular device with a keep-alive interval equal to 20 seconds (that is the default value in [RFC3948] for example) is 29 mA (2G)/34 mA (3G). This consumption is reduced to 16 mA (2G)/24 mA (3G) when the interval is increased to 40 seconds, to 9.1 mA (2G)/16 mA (3G) if the interval is equal to 150 seconds, and to 7.3 mA (2G)/14 mA (3G) if the interval is equal to 180 seconds. When no keep-alive is issued, the consumption would be 5.2 mA (2G)/6.1 mA (3G). The impact of keepalive messages would be more severe if multiple applications are issuing those messages (e.g., SIP, IPsec, etc.).

PCP allows to avoid embedding ALGs (Application Level Gateways) at the network side (e.g., NAT64) to manage protocols which convey IP addresses and/or port numbers (see Section 2.2 of [RFC6889]). Avoiding soliciting ALGs allows for more easiness to make evolve a service independently of the underlying transport network.

A\_REC#3: In order for host-based validation of DNS Security Extensions (DNSSEC) to continue to function in an IPv6-only connectivity with NAT64 deployment context, the cellular host should embed a DNS64 function ([RFC6147]).

This is called "DNS64 in stub-resolver mode" in [RFC6147].

As discussed in Section 5.5 of [RFC6147], a security-aware and validating host has to perform the DNS64 function locally.

Because synthetic AAAA records cannot be successfully validated in a host, learning the PREFIX64 used to construct IPv4-converted IPv6 addresses allows the use of DNSSEC [RFC4033] [RFC4034], [RFC4035]. Means to configure or discover a PREFIX64 are required on the cellular device as discussed in C\_REC#7.

[RFC7051] discusses why a security-aware and validating host has to perform the DNS64 function locally and why it has to be able to learn the proper PREFIX64(s).

A\_REC#4: When the cellular host is dual-stack connected (i.e., configured with an IPv4 address and IPv6 prefix), it should support means to prefer native IPv6 connection over connection established through translation devices (e.g., NAT44 and NAT64).

When both IPv4 and IPv6 DNS servers are configured, a dual-stack host must contact first its IPv6 DNS server. This preference allows to offload IPv4-only DNS servers.

Cellular hosts should follow the procedure specified in [RFC6724] for source address selection.

## 5. Security Considerations

The security considerations identified in [RFC7066] and [RFC6459] are to be taken into account.

In the case of cellular CE routers, compliance with L\_REC#3 entails compliance with [RFC7084], which in turn recommends compliance with Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service [RFC6092]. Therefore, the security considerations in Section 6 of [RFC6092] are relevant. In particular, it bears repeating here that the true impact of stateful filtering may be a reduction in security,

and that IETF make no statement, expressed or implied, as to whether using the capabilities described in any of these documents ultimately improves security for any individual users or for the Internet community as a whole.

The cellular host must be able to generate IPv6 addresses which preserve privacy. The activation of privacy extension (e.g., using [RFC7217]) makes it more difficult to track a host over time when compared to using a permanent Interface Identifier. Tracking a host is still possible based on the first 64 bits of the IPv6 address. Means to prevent against such tracking issues may be enabled in the network side. Note, privacy extensions are required by regulatory bodies in some countries.

Host-based validation of DNSSEC is discussed in A\_REC#3 (see Section 4).

## 6. IANA Considerations

This document does not require any action from IANA.

## 7. Acknowledgements

Many thanks to C. Byrne, H. Soliman, H. Singh, L. Colliti, T. Lemon, B. Sarikaya, M. Mawatari, M. Abrahamsson, P. Vickers, V. Kuarsingh, E. Kline, S. Josefsson, A. Baryun, J. Woodyatt, T. Kossut, B. Stark, and A. Petrescu for the discussion in the v6ops mailing list and for the comments.

Thanks to A. Farrel, B. Haberman, and K. Moriarty for the comments during the IESG review.

Special thanks to T. Savolainen, J. Korhonen, J. Jaeggli, F. Baker, L.M. Contreras Murillo, and M. Abrahamsson for their detailed reviews and comments.

## 8. References

### 8.1. Normative References

- [IR92] GSMA, "IR.92.V4.0 - IMS Profile for Voice and SMS", March 2011, <<http://www.gsma.com/newsroom/ir-92-v4-0-ims-profile-for-voice-and-sms>>.
- [RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <<http://www.rfc-editor.org/info/rfc2460>>.



- [RFC3596] Thomson, S., Huitema, C., Ksinant, V., and M. Souissi, "DNS Extensions to Support IP Version 6", RFC 3596, DOI 10.17487/RFC3596, October 2003, <<http://www.rfc-editor.org/info/rfc3596>>.
- [RFC3633] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633, DOI 10.17487/RFC3633, December 2003, <<http://www.rfc-editor.org/info/rfc3633>>.
- [RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, DOI 10.17487/RFC3986, January 2005, <<http://www.rfc-editor.org/info/rfc3986>>.
- [RFC5795] Sandlund, K., Pelletier, G., and L-E. Jonsson, "The RObusT Header Compression (ROHC) Framework", RFC 5795, DOI 10.17487/RFC5795, March 2010, <<http://www.rfc-editor.org/info/rfc5795>>.
- [RFC5954] Gurbani, V., Ed., Carpenter, B., Ed., and B. Tate, Ed., "Essential Correction for IPv6 ABNF and URI Comparison in RFC 3261", RFC 5954, DOI 10.17487/RFC5954, August 2010, <<http://www.rfc-editor.org/info/rfc5954>>.
- [RFC6052] Bao, C., Huitema, C., Bagnulo, M., Boucadair, M., and X. Li, "IPv6 Addressing of IPv4/IPv6 Translators", RFC 6052, DOI 10.17487/RFC6052, October 2010, <<http://www.rfc-editor.org/info/rfc6052>>.
- [RFC6603] Korhonen, J., Ed., Savolainen, T., Krishnan, S., and O. Troan, "Prefix Exclude Option for DHCPv6-based Prefix Delegation", RFC 6603, DOI 10.17487/RFC6603, May 2012, <<http://www.rfc-editor.org/info/rfc6603>>.
- [RFC7066] Korhonen, J., Ed., Arkko, J., Ed., Savolainen, T., and S. Krishnan, "IPv6 for Third Generation Partnership Project (3GPP) Cellular Hosts", RFC 7066, DOI 10.17487/RFC7066, November 2013, <<http://www.rfc-editor.org/info/rfc7066>>.
- [TS.23060] 3GPP, "General Packet Radio Service (GPRS); Service description; Stage 2", September 2011, <<http://www.3gpp.org/DynaReport/23060.htm>>.

- [TS.23401] 3GPP, "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access", September 2011, <<http://www.3gpp.org/DynaReport/23401.htm>>.
- [TS.24008] 3GPP, "Mobile radio interface Layer 3 specification; Core network protocols; Stage 3", June 2011, <<http://www.3gpp.org/DynaReport/24008.htm>>.

## 8.2. Informative References

- [OECD] Organisation for Economic Cooperation and Development (OECD), "The Economics of the Transition to Internet Protocol version 6 (IPv6)", November 2014, <<http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/ICCP/CISP%282014%293/FINAL&docLanguage=En>>.
- [Power] Haverinen, H., Siren, J., and P. Eronen, "Energy Consumption of Always-On Applications in WCDMA Networks", April 2007, <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4212635>>.
- [R3GPP] 3GPP, "The Mobile Broadband Standard, Releases", 2015, <<http://www.3gpp.org/specifications/67-releases>>.
- [RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, DOI 10.17487/RFC3261, June 2002, <<http://www.rfc-editor.org/info/rfc3261>>.
- [RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <<http://www.rfc-editor.org/info/rfc3948>>.
- [RFC4033] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "DNS Security Introduction and Requirements", RFC 4033, DOI 10.17487/RFC4033, March 2005, <<http://www.rfc-editor.org/info/rfc4033>>.
- [RFC4034] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "Resource Records for the DNS Security Extensions", RFC 4034, DOI 10.17487/RFC4034, March 2005, <<http://www.rfc-editor.org/info/rfc4034>>.

- [RFC4035] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "Protocol Modifications for the DNS Security Extensions", RFC 4035, DOI 10.17487/RFC4035, March 2005, <<http://www.rfc-editor.org/info/rfc4035>>.
- [RFC6092] Woodyatt, J., Ed., "Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service", RFC 6092, DOI 10.17487/RFC6092, January 2011, <<http://www.rfc-editor.org/info/rfc6092>>.
- [RFC6145] Li, X., Bao, C., and F. Baker, "IP/ICMP Translation Algorithm", RFC 6145, DOI 10.17487/RFC6145, April 2011, <<http://www.rfc-editor.org/info/rfc6145>>.
- [RFC6146] Bagnulo, M., Matthews, P., and I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", RFC 6146, DOI 10.17487/RFC6146, April 2011, <<http://www.rfc-editor.org/info/rfc6146>>.
- [RFC6147] Bagnulo, M., Sullivan, A., Matthews, P., and I. van Beijnum, "DNS64: DNS Extensions for Network Address Translation from IPv6 Clients to IPv4 Servers", RFC 6147, DOI 10.17487/RFC6147, April 2011, <<http://www.rfc-editor.org/info/rfc6147>>.
- [RFC6342] Koodli, R., "Mobile Networks Considerations for IPv6 Deployment", RFC 6342, DOI 10.17487/RFC6342, August 2011, <<http://www.rfc-editor.org/info/rfc6342>>.
- [RFC6434] Jankiewicz, E., Loughney, J., and T. Narten, "IPv6 Node Requirements", RFC 6434, DOI 10.17487/RFC6434, December 2011, <<http://www.rfc-editor.org/info/rfc6434>>.
- [RFC6459] Korhonen, J., Ed., Soininen, J., Patil, B., Savolainen, T., Bajko, G., and K. Iisakkila, "IPv6 in 3rd Generation Partnership Project (3GPP) Evolved Packet System (EPS)", RFC 6459, DOI 10.17487/RFC6459, January 2012, <<http://www.rfc-editor.org/info/rfc6459>>.
- [RFC6535] Huang, B., Deng, H., and T. Savolainen, "Dual-Stack Hosts Using "Bump-in-the-Host" (BIH)", RFC 6535, DOI 10.17487/RFC6535, February 2012, <<http://www.rfc-editor.org/info/rfc6535>>.

- [RFC6724] Thaler, D., Ed., Draves, R., Matsumoto, A., and T. Chown, "Default Address Selection for Internet Protocol Version 6 (IPv6)", RFC 6724, DOI 10.17487/RFC6724, September 2012, <<http://www.rfc-editor.org/info/rfc6724>>.
- [RFC6877] Mawatari, M., Kawashima, M., and C. Byrne, "464XLAT: Combination of Stateful and Stateless Translation", RFC 6877, DOI 10.17487/RFC6877, April 2013, <<http://www.rfc-editor.org/info/rfc6877>>.
- [RFC6887] Wing, D., Ed., Cheshire, S., Boucadair, M., Penno, R., and P. Selkirk, "Port Control Protocol (PCP)", RFC 6887, DOI 10.17487/RFC6887, April 2013, <<http://www.rfc-editor.org/info/rfc6887>>.
- [RFC6889] Penno, R., Saxena, T., Boucadair, M., and S. Sivakumar, "Analysis of Stateful 64 Translation", RFC 6889, DOI 10.17487/RFC6889, April 2013, <<http://www.rfc-editor.org/info/rfc6889>>.
- [RFC7050] Savolainen, T., Korhonen, J., and D. Wing, "Discovery of the IPv6 Prefix Used for IPv6 Address Synthesis", RFC 7050, DOI 10.17487/RFC7050, November 2013, <<http://www.rfc-editor.org/info/rfc7050>>.
- [RFC7051] Korhonen, J., Ed. and T. Savolainen, Ed., "Analysis of Solution Proposals for Hosts to Learn NAT64 Prefix", RFC 7051, DOI 10.17487/RFC7051, November 2013, <<http://www.rfc-editor.org/info/rfc7051>>.
- [RFC7084] Singh, H., Beebe, W., Donley, C., and B. Stark, "Basic Requirements for IPv6 Customer Edge Routers", RFC 7084, DOI 10.17487/RFC7084, November 2013, <<http://www.rfc-editor.org/info/rfc7084>>.
- [RFC7217] Gont, F., "A Method for Generating Semantically Opaque Interface Identifiers with IPv6 Stateless Address Autoconfiguration (SLAAC)", RFC 7217, DOI 10.17487/RFC7217, April 2014, <<http://www.rfc-editor.org/info/rfc7217>>.
- [RFC7225] Boucadair, M., "Discovering NAT64 IPv6 Prefixes Using the Port Control Protocol (PCP)", RFC 7225, DOI 10.17487/RFC7225, May 2014, <<http://www.rfc-editor.org/info/rfc7225>>.

- [RFC7278] Byrne, C., Drown, D., and A. Vizdal, "Extending an IPv6 /64 Prefix from a Third Generation Partnership Project (3GPP) Mobile Interface to a LAN Link", RFC 7278, DOI 10.17487/RFC7278, June 2014, <<http://www.rfc-editor.org/info/rfc7278>>.
- [RFC7335] Byrne, C., "IPv4 Service Continuity Prefix", RFC 7335, DOI 10.17487/RFC7335, August 2014, <<http://www.rfc-editor.org/info/rfc7335>>.
- [RFC7445] Chen, G., Deng, H., Michaud, D., Korhonen, J., and M. Boucadair, "Analysis of Failure Cases in IPv6 Roaming Scenarios", RFC 7445, DOI 10.17487/RFC7445, March 2015, <<http://www.rfc-editor.org/info/rfc7445>>.
- [TS.23402] 3GPP, "Architecture enhancements for non-3GPP accesses", September 2011, <<http://www.3gpp.org/DynaReport/23402.htm>>.

## Authors' Addresses

David Binet  
Orange  
Rennes  
France

EMail: [david.binet@orange.com](mailto:david.binet@orange.com)

Mohamed Boucadair  
Orange  
Rennes 35000  
France

EMail: [mohamed.boucadair@orange.com](mailto:mohamed.boucadair@orange.com)

Ales Vizdal  
Deutsche Telekom AG

EMail: [ales.vizdal@t-mobile.cz](mailto:ales.vizdal@t-mobile.cz)

Gang Chen  
China Mobile

EMail: [phdgang@gmail.com](mailto:phdgang@gmail.com)

Nick Heatley  
EE  
The Point, 37 North Wharf Road,  
London W2 1AG  
U.K

EMail: [nick.heatley@ee.co.uk](mailto:nick.heatley@ee.co.uk)

Ross Chandler  
eircom | meteor  
1HSQ  
St. John's Road  
Dublin 8  
Ireland

EMail: [ross@eircom.net](mailto:ross@eircom.net)

Dave Michaud  
Rogers Communications  
8200 Dixie Rd.  
Brampton, ON L6T 0C1  
Canada

EMail: [dave.michaud@rci.rogers.com](mailto:dave.michaud@rci.rogers.com)

Diego R. Lopez  
Telefonica I+D  
Don Ramon de la Cruz, 82  
Madrid 28006  
Spain

Phone: +34 913 129 041  
EMail: [diego.r.lopez@telefonica.com](mailto:diego.r.lopez@telefonica.com)

Walter Haeffner  
Vodafone D2 GmbH  
Ferdinand-Braun-Platz 1  
Duesseldorf 40549  
DE

EMail: [walter.haeffner@vodafone.com](mailto:walter.haeffner@vodafone.com)

Internet Engineering Task Force  
Internet-Draft  
Intended status: Informational  
Expires: September 11, 2014

G. Chen  
Z. Cao  
China Mobile  
C. Xie  
China Telecom  
D. Binet  
France Telecom-Orange  
March 10, 2014

NAT64 Deployment Options and Experience  
draft-ietf-v6ops-nat64-experience-10

Abstract

This document summarizes NAT64 function deployment scenarios and operational experience. Both NAT64 Carrier Grade NAT (NAT64-CGN) and NAT64 server Front End (NAT64-FE) are considered in this document.

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 September 11, 2014.

Copyright Notice

Copyright (c) 2014 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 . . . . .	2
2. Terminology . . . . .	3
3. NAT64 Networking Experience . . . . .	4
3.1. NAT64-CGN Consideration . . . . .	4
3.1.1. NAT64-CGN Usages . . . . .	4
3.1.2. DNS64 Deployment . . . . .	4
3.1.3. NAT64 Placement . . . . .	5
3.1.4. Co-existence of NAT64 and NAT44 . . . . .	5
3.2. NAT64-FE Consideration . . . . .	6
4. High Availability . . . . .	7
4.1. Redundancy Design . . . . .	7
4.2. Load Balancing . . . . .	9
5. Source Address Transparency . . . . .	9
5.1. Traceability . . . . .	9
5.2. Geo-location . . . . .	10
6. Quality of Experience . . . . .	11
6.1. Service Reachability . . . . .	11
6.2. Resource Reservation . . . . .	12
7. MTU Considerations . . . . .	13
8. ULA Usages . . . . .	14
9. Security Considerations . . . . .	15
10. IANA Considerations . . . . .	15
11. Acknowledgements . . . . .	15
12. Additional Author List . . . . .	16
13. References . . . . .	16
13.1. Normative References . . . . .	16
13.2. Informative References . . . . .	18
Appendix A. Testing Results of Application Behavior . . . . .	20
Authors' Addresses . . . . .	21

## 1. Introduction

IPv6 is the only sustainable solution for numbering nodes on Internet due to the IPv4 depletion. Network operators have to deploy IPv6-only networks in order to meet the needs of the expanding internet without available IPv4 addresses.

Single-stack IPv6 network deployment can simplify networks provisioning, some justification was provided in 464xlat [RFC6877]. IPv6-only connectivity confers some benefits to mobile operators as an example. In the mobile context, IPv6-only usage enables the use of a single IPv6 Packet Data Protocol(PDP) context or Evolved Packet System (EPS) bearer on Long Term Evolution (LTE) networks. This



eliminates significant network costs caused by employing two PDP contexts in some cases, and the need for IPv4 addresses to be assigned to customers. In broadband networks overall, it can allow for the scaling of edge-network growth to be decoupled from IPv4 numbering limitations.

In transition scenarios, some existing networks are likely to be IPv4-only for quite a long time. IPv6 networks and hosts IPv6-only hosts will need to coexist with IPv4 numbered resources. Widespread dual-stack deployments have not materialized at the anticipated rate over the last 10 years, one possible conclusion being that legacy networks will not make the jump quickly. The Internet will include nodes that are dual-stack, nodes that remain IPv4-only, and nodes that can be deployed as IPv6-only nodes. A translation mechanism based on a NAT64[RFC6146] [RFC6145]function is likely to be a key element of Internet connectivity for IPv6-IPv4 interoperability.

[RFC6036] reports at least 30% of operators plan to run some kind of translator (presumably NAT64/DNS64). Advice on NAT64 deployment and operations are therefore of some importance. [RFC6586] documents the implications for IPv6 only networks. This document intends to be specific to NAT64 network planning.

## 2. Terminology

Regarding IPv4/IPv6 translation, [RFC6144] has described a framework for enabling networks to make interworking possible between IPv4 and IPv6 networks. This document has further categorized different NAT64 functions, locations and use-cases. The principle distinction of location is whether the NAT64 is located in a Carrier Grade NAT or server Front End. The terms of NAT-CGN/FE are understood to be a topological distinction indicating different features employed in a NAT64 deployment.

**NAT64 Carrier Grade NAT (NAT64-CGN):** A NAT64-CGN is placed in an ISP network. IPv6 enabled subscribers leverage the NAT64-CGN to access existing IPv4 internet services. The ISP as an administrative entity takes full control of the IPv6 side, but has limited or no control on the IPv4 internet side. NAT64-CGN deployments may have to consider the IPv4 Internet environment and services, and make appropriate configuration choices accordingly.

**NAT64 server Front End (NAT64-FE):** A NAT64-FE is generally a device with NAT64 functionality in a content provider or data center network. It could be for example a traffic load balancer or a firewall. The operator of the NAT64-FE has full control over the IPv4 network within the data center, but only limited influence or control over the external Internet IPv6 network.

### 3. NAT64 Networking Experience

#### 3.1. NAT64-CGN Consideration

##### 3.1.1. NAT64-CGN Usages

Fixed network operators and mobile operators may locate NAT64 translators in access networks or in mobile core networks. It can be built into various devices, including routers, gateways or firewalls in order to connect IPv6 users to the IPv4 Internet. With regard to the numbers of users and the shortage of public IPv4 addresses, stateful NAT64[RFC6146] is more suited to maximize sharing of public IPv4 addresses. The usage of stateless NAT64 can provide better transparency features [I-D.ietf-softwire-stateless-4v6-motivation], but has to be coordinated with A+P[RFC6346] processes as specified in [I-D.ietf-softwire-map-t] in order to address an IPv4 address shortage.

##### 3.1.2. DNS64 Deployment

DNS64[RFC6147] is recommended for use in combination with stateful NAT64, and will likely be an essential part of an IPv6 single-stack network that couples to the IPv4 Internet. 464xlat[RFC6877] can enable access of IPv4 only applications or applications that call IPv4 literal addresses. Using DNS64 will help 464xlat to automatically discover NAT64 prefix through [RFC7050]. Berkeley Internet Name Daemon (BIND) software supports the function. It's important to note that DNS64 generates the synthetic AAAA reply when services only provide A records. Operators should not expect to access IPv4 parts of a dual-stack server using NAT64/DNS64. The traffic is forwarded on IPv6 paths if dual-stack servers are targeted. IPv6 traffic may be routed around rather than going through NAT64. Only the traffic going to IPv4-only service would traverse the NAT64 translator. In some sense, it encourages IPv6 usage and limits NAT translation compared to employing NAT44, where all traffic flows have to be translated. In some cases, NAT64-CGNs may serve double roles, i.e. as a translator and IPv6 forwarder. In mobile networks, NAT64 may be deployed as the default gateway serving all the IPv6 traffic. The traffic heading to a dual-stack server is only forwarded on the NAT64. Therefore, both IPv6 and IPv4 are suggested to be configured on the Internet faced interfaces of NAT64. We tested on Top100 websites (referring to [Alexa] statistics). 43% of websites are connected and forwarded on the NAT64 since those websites have both AAAA and A records. With expansion of IPv6 support, the translation process on NAT64 will likely become less-important over time. It should be noted the DNS64-DNSSEC Interaction[RFC6147] may impact validation of Resource Records retrieved from the the DNS64 process. In particular, DNSSEC

validation will fail when DNS64 synthesizes AAAA records where there is a DNS query with the "DNSSEC OK" (DO) bit set and the "Checking Disabled" (CD) bit set received.

### 3.1.3. NAT64 Placement

All connections to IPv4 services from IPv6-only clients must traverse the NAT64-CGN. It can be advantageous from the vantage-point of troubleshooting and traffic engineering to carry the IPv6 traffic natively for as long as possible within an access network and translate packets only at or near the network egress. NAT64 may be a feature of the Autonomous System (AS) border in fixed networks. It may be deployed in an IP node beyond the Gateway GPRS Support Node (GGSN) or Public Data Network- Gateway (PDN-GW) in mobile networks or directly as part of the gateway itself in some situations. This allows consistent attribution and traceability within the service provider network. It has been observed that the process of correlating log information is problematic from multiple-vendor's equipment due to inconsistent formats of log records. Placing NAT64 in a centralized location may reduce diversity of log format and simplify the network provisioning. Moreover, since NAT64 is only targeted at serving traffic flows from IPv6 to IPv4-only services, the user traffic volume should not be as high as in a NAT44 scenario, and therefore, the gateway's capacity in such location may be less of a concern or a hurdle to deployment. On the other-hand, placement in a centralized fashion would require more strict high availability (HA) design. It would also make geo-location based on IPv4 addresses rather inaccurate as is currently the case for NAT44 CGN already deployed in ISP networks. More considerations or workarounds on HA and traceability could be found at Section 4 and Section 5.

### 3.1.4. Co-existence of NAT64 and NAT44

NAT64 will likely co-exist with NAT44 in a dual-stack network where IPv4 private addresses are allocated to customers. The coexistence has already been observed in mobile networks, in which dual stack mobile phones normally initiate some dual-stack PDN/PDP Type[RFC6459] to query both IPv4/IPv6 address and IPv4 allocated addresses are very often private ones. [RFC6724] always prioritizes IPv6 connections regardless of whether the end-to-end path is native IPv6 or IPv6 translated to IPv4 via NAT64/DNS64. Conversely, Happy Eyeballs[RFC6555] will direct some IP flows across IPv4 paths. The selection of IPv4/IPv6 paths may depend on particular implementation choices or settings on a host-by-host basis, and may differ from an operator's deterministic scheme. Our tests verified that hosts may find themselves switching between IPv4 and IPv6 paths as they access identical service, but at different times [I-D.kaliwoda-sunset4-dual-ipv6-coexist]. Since the topology on each

path is potentially different, it may cause unstable user experience and some degradation of Quality of Experience (QoE) when falling back to the other protocol. It's also difficult for operators to find a solution to make a stable network with optimal resource utilization. In general, it's desirable to figure out the solution that will introduce IPv6/IPv4 translation service to IPv6-only hosts connecting to IPv4 servers while making sure dual-stack hosts to have at least one address family accessible via native service if possible. With the end-to-end native IPv6 environment available, hosts should be upgraded aggressively to migrate in favor of IPv6-only. There are ongoing efforts to detect host connectivity and propose a new DHCPv6 option[I-D.wing-dhc-dns-reconfigure] to convey appropriate configuration information to the hosts.

### 3.2. NAT64-FE Consideration

Some Internet Content Providers (ICPs) may locate NAT64 in front of an Internet Data Center (IDC), for example co-located with a load-balancing function. Load-balancers are employed to connect different IP family domains, and distribute workloads across multiple domains or internal servers. In some cases, IPv4 addresses exhaustion may not be a problem in some IDC's internal networks. IPv6 support for some applications may require some investments and workloads so IPv6 support may not be a priority. The use of NAT64 may be served to support widespread IPv6 adoption on the Internet while maintaining IPv4-only applications access.

Different strategy has been described in [RFC6883] referred to as "inside out" and "outside in". An IDC operator may implement the following practices in the NAT64-FE networking scenario.

- o Some ICPs who already have satisfactory operational experience might adopt single stack IPv6 operation in building data-center networks, servers and applications, as it allows new services delivery without having to integrate consideration of IPv4 NAT and address limitations of IPv4 networks. Stateless NAT64[RFC6145] can be used to provide services for IPv4-only enabled customers. [I-D.anderson-siit-dc] has provided further descriptions and guidelines.
- o ICPs who attempt to offer customers IPv6 support in their application farms at an early stage may likely run proxies load-balancers or translators, which are configured to handle incoming IPv6 flows and proxy them to IPv4 back-end systems. Many load balancers integrate proxy functionality. IPv4 addresses configured in the proxy may be multiplexed like a stateful NAT64 translator. A similar challenge exists once increasingly numerous users in IPv6 Internet access an IPv4 network. High loads on

load-balancers may be apt to cause additional latency, IPv4 pool exhaustion, etc. Therefore, this approach is only reasonable at an early stage. ICPs may employ dual-stack or IPv6 single stack in a further stage, since the native IPv6 is frequently more desirable than any of the transition solutions.

[RFC6144] recommends that AAAA records of load-balancers or application servers can be directly registered in the authoritative DNS servers. In this case, there is no need to deploy DNS64 name-servers. Those AAAA records can point to natively assigned IPv6 addresses or IPv4-converted IPv6 addresses[RFC6052]. Hosts are not aware of the NAT64 translator on communication path. For the testing purpose, operators could employ an independent sub domain e.g. ipv6exp.example.com to identify experimental ipv6 services to users. How to design the FQDN for the IPv6 service is out-of-scope of this document.

#### 4. High Availability

##### 4.1. Redundancy Design

High Availability (HA) is a major requirement for every service and network services. The deployment of redundancy mechanisms is an essential approach to avoid failure and significantly increase the network reliability. It's not only useful to stateful NAT64 cases, but also to stateless NAT64 gateways.

Three redundancy modes are mainly used: cold standby, warm standby and hot standby.

- o Cold standby HA devices do not replicate the NAT64 states from the primary equipment to the backup. Administrators switch on the backup NAT64 only if the primary NAT64 fails. As a result, all existing established sessions through a failed translator will be disconnected. The translated flows will need to be recreated by end-systems. Since the backup NAT64 is manually configured to switch over to active NAT64, it may have unpredictable impacts to the ongoing services.
- o Warm standby is a flavor of the cold standby mode. Backup NAT64 would keep running once the primary NAT64 is working. This makes warm standby less time consuming during the traffic failover. Virtual Router Redundancy Protocol (VRRP)[RFC5798] can be a solution to enable automatic handover in the warm standby. It was tested that the handover takes as maximum as 1 minute if the backup NAT64 needs to take over routing and re-construct the Binding Information Bases (BIBs) for 30 million sessions. In

deployment phase, operators could balance loads on distinct NAT64s devices. Those NAT64s make a warm backup of each other.

- o Hot standby must synchronize the BIBs between the primary NAT64 and backup. When the primary NAT64 fails, backup NAT64 would take over and maintain the state of all existing sessions. The internal hosts don't have to re-connect the external hosts. The handover time has been extremely reduced. Employing Bidirectional Forwarding Detection (BFD) [RFC5880] combined with VRRP, a delay of only 35ms for 30 million sessions handover was observed during testing. Under ideal conditions hotstandby deployments could guarantee the session continuity for every service. In order to timely transmit session states, operators may have to deploy extra transport links between primary NAT64 and distant backup. The scale of synchronization data instance is depending on the particular deployment. For example, If a NAT64-CGN is served for 200,000 users, the average amount of 800, 000 sessions per second is roughly estimated for new created and expired sessions. A physical 10Gbps transport link may have to be deployed for the sync data transmission considering the amount of sync sessions at the peak and capacity redundancy

In general, cold-standby and warm-standby is simpler and less resource intensive, but it requires clients to re-establish sessions when a fail-over occurs. Hot standby increases resource consumption in order to synchronize state, but potentially achieves seamless handover. For stateless NAT64 considerations are simple, because state synchronization is unnecessary. Regarding stateful NAT64, it may be useful to investigate performance tolerance of applications and the traffic characteristics in a particular network. Some testing results are shown in the Appendix A.

Our statistics in a mobile network shown that almost 91.21% of of traffic is accounted by http/https services. These services generally don't require session continuity. Hot-standby does not offer much benefit for those sessions on this point. In fixed networks, HTTP streaming, p2p and online games would be the major traffic beneficiaries of hot-standby replication[Cisco-VNI]. Consideration should be given to the importance of maintaining bindings for those sessions across failover. Operators may also consider the Average Revenue Per User (ARPU) factors to deploy suitable redundancy mode. Warm standby may still be adopted to cover most services while hot standby could be used to upgrade Quality of Experience (QoE) using DNS64 to generate different synthetic responses for limited traffic or destinations. Further considerations are discussed at Section 6.

#### 4.2. Load Balancing

Load balancing is used to accompany redundancy design so that better scalability and resiliency could be achieved. Stateless NAT64s allow asymmetric routing while anycast-based solutions are recommended in [I-D.ietf-softwire-map-deployment]. The deployment of load balancing may make more sense to stateful NAT64s for the sake of single-point failure avoidance. Since the NAT64-CGN and NAT64-FE have distinct facilities, the following lists the considerations for each case.

- o NAT64-CGN equipment doesn't typically implement load-balancing functions onboard. Therefore, the gateways have to resort to DNS64 or internal host's behavior. Once DNS64 is deployed, the load balancing can be performed by synthesizing AAAA response with different IPv6 prefixes. For the applications not requiring DNS resolver, internal hosts could learn multiple IPv6 prefixes through the approaches defined in[RFC7050] and then select one based on a given prefix selection policy.
- o A dedicated Load Balancer could be deployed at front of a NAT64-FE farm. Load Balancer uses proxy mode to redirect the flows to the appropriate NAT64 instance. Stateful NAT64s require a deterministic pattern to arrange the traffic in order to ensure outbound/inbound flows traverse the identical NAT64. Therefore, static scheduling algorithms, for example source-address based policy, is preferred. A dynamic algorithm, for example Round-Robin, may have impacts on applications seeking session continuity, which described in the Table 1.

#### 5. Source Address Transparency

##### 5.1. Traceability

Traceability is required in many cases such as identifying malicious attacks sources and accounting requirements. Operators are asked to record the NAT64 log information for specific periods of time. In our lab testing, the log information from 200,000 subscribers have been collected from a stateful NAT64 gateway for 60 days. Syslog[RFC5424] has been adopted to transmit log message from NAT64 to a log station. Each log message contains transport protocol, source IPv6 address:port, translated IPv4 address: port and timestamp. It takes almost 125 bytes in ASCII format. It has been verified that the rate of traffic flow is around 72 thousand flows per second and the volume of recorded information reaches up to 42.5 terabytes in the raw format. The volume is 29.07 terabytes in a compact format. At scale, operators have to build up dedicated transport links, storage system and servers for the purpose of managing such logging.

There are also several improvements that can be made to mitigate the issue. For example, stateful NAT64 could configure with bulk port allocation method. Once a subscriber creates the first session, a number of ports are pre-allocated. A bulk allocation message is logged indicating this allocation. Subsequent session creations will use one of the pre-allocated port and hence does not require logging. The log volume in this case may be only one thousandth of dynamic port allocation. Some implementations may adopt static port-range allocations [I-D.donley-behave-deterministic-cgn] which eliminates the need for per-subscriber logging. As a side effect, the IPv4 multiplexing efficiency is decreased regarding to those methods. For example, the utilization ratio of public IPv4 address is dropped approximately to 75% when NAT64 gateway is configured with bulk port allocation (The lab testing allocates each subscriber with 400 ports). In addition, port-range based allocation should also consider port randomization described in [RFC6056]. A trade-off among address multiplexing efficiency, logging storage compression and port allocation complexity should be considered. More discussions could be found in [I-D.chen-sunset4-cgn-port-allocation]. The decision can balance usable IPv4 resources against investments in log systems.

## 5.2. Geo-location

IP addresses are usually used as inputs to geo-location services. The use of address sharing prevents these systems from resolving the location of a host based on IP address alone. Applications that assume such geographic information may not work as intended. The possible solutions listed in [RFC6967] are intended to bridge the gap. However, those solutions can only provide a sub-optimal substitution to solve the problem of host identification, in particular it may not today solve problems with source identification through translation. The following lists current practices to mitigate the issue.

- o Operators who adopt NAT64-FE may leverage the application layer proxies, e.g. X-Forwarded-For (XFF) [I-D.ietf-appsawg-http-forwarded], to convey the IPv6 source address in HTTP headers. Those messages would be passed on to web-servers. The log parsing tools are required to be able to support IPv6 and may lookup Radius servers for the target subscribers based on IPv6 addresses included in XFF HTTP headers. XFF is the de facto standard which has been integrated in most Load Balancers. Therefore, it may be superior to use in a NAT-FE environment. In the downsides, XFF is specific to HTTP. It restricts the usages so that the solution can't be applied to requests made over HTTPs. This makes geo-location problematic for HTTPs based services.



- o The NAT64-CGN equipment may not implement XFF. Geo-location based on shared IPv4 address is rather inaccurate in that case. Operators could subdivide the outside IPv4 address pool so an IPv6 address can be translated depending on their geographical locations. As consequence, location information can be identified from a certain IPv4 address range. [RFC6967] also enumerates several options to reveal the host identifier. Each solution likely has their-own specific usage. For the geo-location systems relying on a Radius database[RFC5580], we have investigated to deliver NAT64 BIBs and Session Table Entries (STEs) to a Radius server[I-D.chen-behave-nat64-radius-extension]. This method could provide geo-location system with an internal IPv6 address to identify each user. It can get along with [RFC5580] to convey original source address through same message bus.

## 6. Quality of Experience

### 6.1. Service Reachability

NAT64 is providing a translation capability between IPv6 and IPv4 end-nodes. In order to provide the reachability between two IP address families, NAT64-CGN has to implement appropriate application aware functions, i.e. Application Layer Gateway (ALG), where address translation is not itself sufficient and security mechanisms do not render it infeasible. Most NAT64-CGNs mainly provide FTP-ALG[RFC6384]. NAT64-FEs may have functional richness on Load Balancer, for example HTTP-ALG, HTTPS-ALG, RTSP-ALG and SMTP-ALG have been supported. Those application protocols exchange IP address and port parameters within control session, for example the "Via" field in a HTTP header, "Transport" field in a RTSP SETUP message and "Received: " header in a SMTP message. ALG functions will detect those fields and make IP address translations. It should be noted that ALGs may impact the performance on a NAT64 box to some extent. ISPs as well as content providers might choose to avoid situations where the imposition of an ALG might be required. At the same time, it is also important to remind customers and application developers that IPv6 end-to-end usage does not require ALG imposition and therefore results in a better overall user experience.

The service reachability is also subject to the IPv6 support in the client side. We tested several kinds of applications as shown in the below table to verify the IPv6 supports. The experiences of some applications are still align with [RFC6586]. For example, we have tested P2P file sharing and streaming applications including eMule v0.50a, Thunder v7.9 and PPS TV v3.2.0. It has been found there are some software issues to support IPv6 at this time. The application software would benefit from 464xlat[RFC6877] until the software adds IPv6 support.. A SIP based voice call has been tested in LTE mobile

environment as specified in [IR.92]. The voice call is failed due to the lack of NAT64 traversal when an IPv6 SIP user agent communicates with an IPv4 SIP user agent. In order to address the failure, Interactive Connectivity Establishment (ICE) described in [RFC5245] is recommended to be supported for the SIP IPv6 transition. [RFC6157] describes both signaling and media layer process, which should be followed. In addition, it may be worth to notice that ICE is not only useful for NAT traversal, but also firewall[RFC6092] traversal in native IPv6 deployment.

Different IPsec modes for VPN services have been tested, including IPsec-AH and IPsec-ESP. It has been testified IPsec-AH can't survive since the destination host detects the IP header changes and invalidate the packets. IPsec-ESP failed in our testing because the NAT64 does not translate IPsec ESP (i.e. protocol 50) packets. It has been suggested that IPsec ESP should succeed if the IPsec client supports NAT-Traversal in the IKE[RFC3947] and uses IPsec ESP over UDP[RFC3948].

Table 1: The tested applications

APPs	Results and Found Issues
Web service	Mostly pass, some failure cases due to IPv4 Literals
Instant Message	Mostly fail, software can't support IPv6
Games	Mostly pass for web-based games; mostly fail for standalone games due to the lack of IPv6 support in software
SIP-VoIP	Fail, due to the lack of NAT64 traversal
IPsec-VPN	Fail, the translated IPsec packets are invalidated
P2P file sharing and streaming	Mostly fail, software can't support IPv6, e.g. eMule, Thunder and PPS TV
FTP	Pass
Email	Pass

## 6.2. Resource Reservation

Session status normally is managed by a static timer. For example, the value of the "established connection idle-timeout" for TCP sessions must not be less than 2 hours 4 minutes[RFC5382] and 5

minutes for UDP sessions[RFC4787]. In some cases, NAT resource maybe significantly consumed by largely inactive users. The NAT translator and other customers would suffer from service degradation due to port consummation by other subscribers using the same NAT64 device. A flexible NAT session control is desirable to resolve the issues. PCP[RFC6887] could be a candidate to provide such capability. A NAT64-CGN should integrate with a PCP server, to allocate available IPv4 address/port resources. Resources could be assigned to PCP clients through PCP MAP/PEER mode. Such ability can be considered to upgrade user experiences, for example assigning different sizes of port ranges for different subscribers. Those mechanisms are also helpful to minimize terminal battery consumption and reduce the number of keep-alive messages to be sent by mobile terminal devices.

Subscribers can also benefit from network reliability. It has been discussed that hot-standby offers satisfactory experience once outage of primary NAT64 is occurred. Operators may rightly be concerned about the considerable investment required for NAT64 equipment relative to low ARPU income. For example, transport links may cost much, because primary NAT64 and backup are normally located at different locations, separated by a relatively large distance. Additional cost has to be assumed to ensure the connectivity quality. However, that may be necessary to some applications, which are delay-sensitive and seek session continuity, for example on-line games and live-streaming. Operators may be able to get added-values from those services by offering first-class services. It can be pre-configured on the gateway to hot-standby modes depending on subscriber's profile. The rest of other sessions can be covered by cold/warm standby.

## 7. MTU Considerations

IPv6 requires that every link in the internet have an Maximum Transmission Unit (MTU) of 1280 octets or greater[RFC2460]. However, in case of NAT64 translation deployment, some IPv4 MTU constrained link will be used in some communication path and originating IPv6 nodes may therefore receive an ICMP Packet Too Big (PTB) message, reporting a Next-Hop MTU less than 1280 bytes. The result would be that IPv6 allows packets to contain a fragmentation header, without the packet being fragmented into multiple pieces. A NAT64 would receive IPv6 packets with fragmentation header in which "M" flag equal to 0 and "Fragment Offset" equal to 0. Those packets likely impact other fragments already queued with the same set of {IPv6 Source Address, IPv6 Destination Address, Fragment Identification}. If the NAT64 box is compliant with [RFC5722], there is risk that all the fragments have to be dropped.

[RFC6946] discusses how this situation could be exploited by an attacker to perform fragmentation-based attacks, and also proposes an improved handling of such packets. It required enhancements on NAT64 gateway implementations to isolate packet's processing. NAT64 should follow the recommendation and take steps to prevent the risks of fragmentation.

Another approach that potentially avoids this issue is to configure IPv4 MTU more than 1260 bytes. It would forbid the occurrence of PTB smaller than 1280 bytes. Such an operational consideration is hard to universally apply to the legacy "IPv4 Internet" NAT64-CGN bridged. However, it's a feasible approach in NAT64-FE cases, since a IPv4 network NAT64-FE connected is rather well-organized and operated by a IDC operator or content provider. Therefore, the MTU of IPv4 network in NAT64-FE case are strongly recommended to set to more than 1260 bytes.

## 8. ULA Usages

Unique Local Addresses (ULAs) are defined in [RFC4193] to be renumbered within a network site for local communications. Operators may use ULAs as NAT64 prefixes to provide site-local IPv6 connectivity. Those ULA prefixes are stripped when the packets going to the IPv4 Internet, therefore ULAs are only valid in the IPv6 site. The use of ULAs could help in identifying the translation traffic. [I-D.ietf-v6ops-ula-usage-recommendations] provides further guidance for the ULAs usages.

We configure ULAs as NAT64 prefixes on a NAT64-CGN. If a host is only assigned with an IPv6 address and connected to NAT64-CGN, when connect to an IPv4 service, it would receive AAAA record generated by the DNS64 with the ULA prefix. A Global Unicast Address (GUA) will be selected as the source address to the ULA destination address. When the host has both IPv4 and IPv6 address, it would initiate both A and AAAA record lookup, then both original A record and DNS64-generated AAAA record would be received. A host, which is compliant with [RFC6724], will never prefer ULA over IPv4. An IPv4 path will be always selected. It may be undesirable because the NAT64-CGN will never be used. Operators may consider to add additional site-specific rows into the default policy table for host address selection in order to steer traffic flows going through NAT64-CGN. However, it involves significant costs to change terminal's behavior. Therefore, operators are not suggested to configure ULAs on a NAT64-CGN.

ULAs can't work when hosts transit the Internet to connect with NAT64. Therefore, ULAs are inapplicable to the case of NAT64-FE.

## 9. Security Considerations

This document presents the deployment experiences of NAT64 in CGN and FE scenarios. In general, RFC 6146[RFC6146] provides TCP-tracking, address-dependent filtering mechanisms to protect NAT64 from Distributed Denial of Service (DDoS). In NAT64-CGN cases, operators also could adopt unicast Reverse Path Forwarding (uRPF)[RFC3704] and black/white-list to enhance the security by specifying access policies. For example, NAT64-CGN should forbid establish NAT64 BIB for incoming IPv6 packets if uRPF in Strict or Loose mode check does not pass or whose source IPv6 address is associated to black-lists.

The stateful NAT64-FE creates state and maps that connection to an internally-facing IPv4 address and port. An attacker can consume the resources of the NAT64-FE device by sending an excessive number of connection attempts. Without a DDoS limitation mechanism, the NAT64-FE is exposed to attacks. Load Balancer is recommended to enable the capabilities of line rate DDOS defense, such as the employment of SYN PROXY-COOKIE. Security domain division is necessary as well in this case. Therefore, Load Balancers could not only serve for optimization of traffic distribution, but also prevent service from quality deterioration due to security attacks.

The DNS64 process will potentially interfere with the DNSSEC functions[RFC4035], since DNS response is modified and DNSSEC intends to prevent such changes. More detailed discussions can be found in [RFC6147].

## 10. IANA Considerations

This memo includes no request to IANA.

## 11. Acknowledgements

The authors would like to thank Jari Arkko, Dan Wing, Remi Despres, Fred Baker, Hui Deng, Iljitsch van Beijnum, Philip Matthews, Randy Bush, Mikael Abrahamsson, Lorenzo Colitti, Sheng Jiang, Nick Heatley, Tim Chown, Gert Doering and Simon Perreault for their helpful comments.

Many thanks to Wesley George, Lee Howard and Satoru Matsushima for their detailed reviews.

The authors especially thank Joel Jaeggli and Ray Hunter for his efforts and contributions on editing which substantially improves the legibility of the document.

Thanks to Cameron Byrne who was an active co-author of some earlier versions of this draft.

## 12. Additional Author List

The following are extended authors who contributed to the effort:

Qiong Sun  
China Telecom  
Room 708, No.118, Xizhimennei Street  
Beijing 100035  
P.R.China  
Phone: +86-10-58552936  
Email: sunqiong@ctbri.com.cn

QiBo Niu  
ZTE  
50,RuanJian Road.  
YuHua District,  
Nan Jing 210012  
P.R.China  
Email: niu.qibo@zte.com.cn

## 13. References

### 13.1. Normative References

- [I-D.ietf-appsawg-http-forwarded]  
Petersson, A. and M. Nilsson, "Forwarded HTTP Extension",  
draft-ietf-appsawg-http-forwarded-10 (work in progress),  
October 2012.
- [RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6  
(IPv6) Specification", RFC 2460, December 1998.
- [RFC3704] Baker, F. and P. Savola, "Ingress Filtering for Multihomed  
Networks", BCP 84, RFC 3704, March 2004.
- [RFC3947] Kivinen, T., Swander, B., Huttunen, A., and V. Volpe,  
"Negotiation of NAT-Traversal in the IKE", RFC 3947,  
January 2005.
- [RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M.  
Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC  
3948, January 2005.

- [RFC4035] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "Protocol Modifications for the DNS Security Extensions", RFC 4035, March 2005.
- [RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, October 2005.
- [RFC4787] Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", BCP 127, RFC 4787, January 2007.
- [RFC5245] Rosenberg, J., "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols", RFC 5245, April 2010.
- [RFC5382] Guha, S., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, "NAT Behavioral Requirements for TCP", BCP 142, RFC 5382, October 2008.
- [RFC5424] Gerhards, R., "The Syslog Protocol", RFC 5424, March 2009.
- [RFC5580] Tschofenig, H., Adrangi, F., Jones, M., Lior, A., and B. Aboba, "Carrying Location Objects in RADIUS and Diameter", RFC 5580, August 2009.
- [RFC5722] Krishnan, S., "Handling of Overlapping IPv6 Fragments", RFC 5722, December 2009.
- [RFC5798] Nadas, S., "Virtual Router Redundancy Protocol (VRRP) Version 3 for IPv4 and IPv6", RFC 5798, March 2010.
- [RFC5880] Katz, D. and D. Ward, "Bidirectional Forwarding Detection (BFD)", RFC 5880, June 2010.
- [RFC6052] Bao, C., Huitema, C., Bagnulo, M., Boucadair, M., and X. Li, "IPv6 Addressing of IPv4/IPv6 Translators", RFC 6052, October 2010.
- [RFC6145] Li, X., Bao, C., and F. Baker, "IP/ICMP Translation Algorithm", RFC 6145, April 2011.
- [RFC6146] Bagnulo, M., Matthews, P., and I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", RFC 6146, April 2011.

- [RFC6147] Bagnulo, M., Sullivan, A., Matthews, P., and I. van Beijnum, "DNS64: DNS Extensions for Network Address Translation from IPv6 Clients to IPv4 Servers", RFC 6147, April 2011.
- [RFC6157] Camarillo, G., El Malki, K., and V. Gurbani, "IPv6 Transition in the Session Initiation Protocol (SIP)", RFC 6157, April 2011.
- [RFC6384] van Beijnum, I., "An FTP Application Layer Gateway (ALG) for IPv6-to-IPv4 Translation", RFC 6384, October 2011.
- [RFC6555] Wing, D. and A. Yourtchenko, "Happy Eyeballs: Success with Dual-Stack Hosts", RFC 6555, April 2012.
- [RFC6724] Thaler, D., Draves, R., Matsumoto, A., and T. Chown, "Default Address Selection for Internet Protocol Version 6 (IPv6)", RFC 6724, September 2012.
- [RFC6887] Wing, D., Cheshire, S., Boucadair, M., Penno, R., and P. Selkirk, "Port Control Protocol (PCP)", RFC 6887, April 2013.
- [RFC6946] Gont, F., "Processing of IPv6 "Atomic" Fragments", RFC 6946, May 2013.
- [RFC7050] Savolainen, T., Korhonen, J., and D. Wing, "Discovery of the IPv6 Prefix Used for IPv6 Address Synthesis", RFC 7050, November 2013.

### 13.2. Informative References

- [Alexa] Alexa, "<http://www.alexa.com/topsites>", April 2013.
- [Cisco-VNI] Cisco, "Cisco Visual Networking Index: Forecast and Methodology, 2012-2017, <http://ciscovni.com/forecast-widget/index.html>", May 2013.
- [I-D.anderson-siit-dc] Anderson, T., "Stateless IP/ICMP Translation in IPv6 Data Centre Environments", draft-anderson-siit-dc-00 (work in progress), November 2012.
- [I-D.chen-behave-nat64-radius-extension] Chen, G. and D. Binet, "Radius Attributes for Stateful NAT64", draft-chen-behave-nat64-radius-extension-00 (work in progress), July 2013.



- [I-D.chen-sunset4-cgn-port-allocation]  
Chen, G., Tsou, T., Donley, C., and T. Taylor, "Analysis of NAT64 Port Allocation Method", draft-chen-sunset4-cgn-port-allocation-03 (work in progress), February 2014.
- [I-D.donley-behave-deterministic-cgn]  
Donley, C., Grundemann, C., Sarawat, V., Sundaresan, K., and O. Vautrin, "Deterministic Address Mapping to Reduce Logging in Carrier Grade NAT Deployments", draft-donley-behave-deterministic-cgn-07 (work in progress), January 2014.
- [I-D.ietf-softwire-map-deployment]  
Qiong, Q., Chen, M., Chen, G., Tsou, T., and S. Perreault, "Mapping of Address and Port (MAP) - Deployment Considerations", draft-ietf-softwire-map-deployment-03 (work in progress), October 2013.
- [I-D.ietf-softwire-map-t]  
Li, X., Bao, C., Dec, W., Troan, O., Matsushima, S., and T. Murakami, "Mapping of Address and Port using Translation (MAP-T)", draft-ietf-softwire-map-t-05 (work in progress), February 2014.
- [I-D.ietf-softwire-stateless-4v6-motivation]  
Boucadair, M., Matsushima, S., Lee, Y., Bonness, O., Borges, I., and G. Chen, "Motivations for Carrier-side Stateless IPv4 over IPv6 Migration Solutions", draft-ietf-softwire-stateless-4v6-motivation-05 (work in progress), November 2012.
- [I-D.ietf-v6ops-ula-usage-recommendations]  
Liu, B. and S. Jiang, "Recommendations of Using Unique Local Addresses", draft-ietf-v6ops-ula-usage-recommendations-02 (work in progress), February 2014.
- [I-D.kaliwoda-sunset4-dual-ipv6-coexist]  
Kaliwoda, A. and D. Binet, "Co-existence of both dual-stack and IPv6-only hosts", draft-kaliwoda-sunset4-dual-ipv6-coexist-01 (work in progress), October 2012.
- [I-D.wing-dhc-dns-reconfigure]  
Patil, P., Boucadair, M., Wing, D., and T. Reddy, "DHCPv6 Dynamic Reconfiguration", draft-wing-dhc-dns-reconfigure-02 (work in progress), September 2013.

- [IR.92] Global System for Mobile Communications Association (GSMA), , "IMS Profile for Voice and SMS Version 7.0", March 2013.
- [RFC6036] Carpenter, B. and S. Jiang, "Emerging Service Provider Scenarios for IPv6 Deployment", RFC 6036, October 2010.
- [RFC6056] Larsen, M. and F. Gont, "Recommendations for Transport-Protocol Port Randomization", BCP 156, RFC 6056, January 2011.
- [RFC6092] Woodyatt, J., "Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service", RFC 6092, January 2011.
- [RFC6144] Baker, F., Li, X., Bao, C., and K. Yin, "Framework for IPv4/IPv6 Translation", RFC 6144, April 2011.
- [RFC6346] Bush, R., "The Address plus Port (A+P) Approach to the IPv4 Address Shortage", RFC 6346, August 2011.
- [RFC6459] Korhonen, J., Soininen, J., Patil, B., Savolainen, T., Bajko, G., and K. Iisakkila, "IPv6 in 3rd Generation Partnership Project (3GPP) Evolved Packet System (EPS)", RFC 6459, January 2012.
- [RFC6586] Arkko, J. and A. Keranen, "Experiences from an IPv6-Only Network", RFC 6586, April 2012.
- [RFC6877] Mawatari, M., Kawashima, M., and C. Byrne, "464XLAT: Combination of Stateful and Stateless Translation", RFC 6877, April 2013.
- [RFC6883] Carpenter, B. and S. Jiang, "IPv6 Guidance for Internet Content Providers and Application Service Providers", RFC 6883, March 2013.
- [RFC6967] Boucadair, M., Touch, J., Levis, P., and R. Penno, "Analysis of Potential Solutions for Revealing a Host Identifier (HOST\_ID) in Shared Address Deployments", RFC 6967, June 2013.

#### Appendix A. Testing Results of Application Behavior

We test several application behaviors in a lab environment to evaluate the impact when a primary NAT64 is out of service. In this testing, participants are asked to connect a IPv6-only WiFi network

using laptops, tablets or mobile phones. NAT64 is deployed as the gateway to connect Internet service. The tested applications are shown in the below table. Cold standby, warm standby and hot standby are taken turn to be tested. The participants may experience service interruption due to the NAT64 handover. Different interruption intervals are tested to gauge application behaviors. The results are illuminated as below.

Table 2: The acceptable delay of applications

APPs	Acceptable Interrupt Recovery	Session Continuity
Web Browse	As maximum as 6s	No
Http streaming	As maximum as 10s(cache)	Yes
Gaming	200ms~400ms	Yes
P2P streaming, file sharing	10~16s	Yes
Instant Message	1 minute	Yes
Mail	30 seconds	No
Downloading	1 minutes	No

## Authors' Addresses

Gang Chen  
 China Mobile  
 Xuanwumenxi Ave. No.32,  
 Xuanwu District,  
 Beijing 100053  
 China

Email: phdgang@gmail.com

Zhen Cao  
China Mobile  
Xuanwumenxi Ave. No.32,  
Xuanwu District,  
Beijing 100053  
China

Email: caozhen@chinamobile.com, zehn.cao@gmail.com

Chongfeng Xie  
China Telecom  
Room 708 No.118, Xizhimenneidajie  
Beijing 100035  
P.R.China

Email: xiechf@ctbri.com.cn

David Binet  
France Telecom-Orange  
Rennes  
35000  
France

Email: david.binet@orange.com

Internet Engineering Task Force  
Internet-Draft  
Intended status: Informational  
Expires: November 3, 2015

B. Liu  
S. Jiang  
Huawei Technologies  
May 2, 2015

Considerations For Using Unique Local Addresses  
draft-ietf-v6ops-ula-usage-recommendations-05

Abstract

This document provides considerations for using IPv6 Unique Local Addresses (ULAs). It identifies cases where ULA addresses are helpful as well as potential problems that their use could introduce, based on an analysis of different ULA usage scenarios.

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 November 3, 2015.

Copyright Notice

Copyright (c) 2015 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 . . . . .	2
2. Requirements Language . . . . .	3
3. Analysis of ULA Features . . . . .	3
3.1. Automatically Generated . . . . .	3
3.2. Globally Unique . . . . .	3
3.3. Independent Address Space . . . . .	3
3.4. Well Known Prefix . . . . .	4
3.5. Stable or Temporary Prefix . . . . .	4
4. Analysis and Operational Considerations of Scenarios Using ULAs . . . . .	4
4.1. Isolated Networks . . . . .	4
4.2. Connected Networks . . . . .	5
4.2.1. ULA-Only Deployment . . . . .	5
4.2.2. ULAs along with PA Addresses . . . . .	7
4.3. IPv4 Co-existence Considerations . . . . .	9
5. General Considerations For Using ULAs . . . . .	10
5.1. Do Not Treat ULA Equal to RFC1918 . . . . .	10
5.2. Using ULAs in a Limited Scope . . . . .	10
6. ULA Usages Considered Helpful . . . . .	10
6.1. Used in Isolated Networks . . . . .	11
6.2. ULA along with PA . . . . .	11
6.3. Some Specific Use Cases . . . . .	11
6.3.1. Special Routing . . . . .	11
6.3.2. Used as NAT64 Prefix . . . . .	11
6.3.3. Used as Identifier . . . . .	12
7. Security Considerations . . . . .	13
8. IANA Considerations . . . . .	13
9. Acknowledgements . . . . .	13
10. References . . . . .	13
10.1. Normative References . . . . .	13
10.2. Informative References . . . . .	14
Authors' Addresses . . . . .	16

## 1. Introduction

Unique Local Addresses (ULAs) are defined in [RFC4193] as provider-independent prefixes that can be used locally, for example, on isolated networks, internal networks, or VPNs. Although ULAs may be treated like addresses of global scope by applications, normally they are not used on the public Internet. ULAs are a possible alternative to site-local addresses (deprecated in [RFC3879]) in some situations, but there are differences between the two address types.

The use of ULAs in various types of networks has been confusing to network operators. This document aims to clarify the advantages and disadvantages of ULAs and how they can be most appropriately used.

## 2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119] when they appear in ALL CAPS. When these words are not in ALL CAPS (such as "should" or "Should"), they have their usual English meanings, and are not to be interpreted as [RFC2119] key words.

## 3. Analysis of ULA Features

### 3.1. Automatically Generated

ULA prefixes can be automatically generated using the algorithms described in [RFC4193]. This feature allows automatic prefix allocation. Thus one can get a network working immediately without applying for prefix(es) from an RIR/LIR (Regional Internet Registry/ Local Internet Registry).

### 3.2. Globally Unique

ULAs are intended to have an extremely low probability of collision. Since multiple networks in which the hosts have been assigned with ULAs may occasionally be merged into one network, this uniqueness is necessary. The randomization of 40 bits in a ULA prefix is considered sufficient enough to ensure a high degree of uniqueness (refer to [RFC4193] Section 3.2.3 for details) and simplifies merging of networks by avoiding the need to renumber overlapping IP address space. Such overlapping was a major drawback to the deployment of private [RFC1918] addresses in IPv4.

Note that, as described in [RFC4864], applications may treat ULAs in practice like global-scope addresses, but address selection algorithms may need to distinguish between ULAs and Global-scope Unicast Addresses (GUAs) to ensure bidirectional communications. As a further note, the default address selection policy table in [RFC6724]) responds to this requirement.

### 3.3. Independent Address Space

ULAs provide internal address independence in IPv6 since they can be used for internal communications even without Internet connectivity. They need no registration, so they can support on-demand usage and do not carry any RIR/LIR burden of documentation or fees.

### 3.4. Well Known Prefix

The prefixes of ULAs are well known thus they are easily identified and filtered.

This feature is convenient for management of security policies and troubleshooting. For example, network administrators can segregate packets containing data which must stay in the internal network by assigning ULAs to internal servers. Externally-destined data can be sent to the Internet or telecommunication network by a separate function, through an appropriate gateway/firewall.

### 3.5. Stable or Temporary Prefix

A ULA prefix can be generated once, at installation time or factory reset, and then possibly never be changed. Alternatively, it can be regenerated regularly, depending on deployment requirements.

## 4. Analysis and Operational Considerations of Scenarios Using ULAs

### 4.1. Isolated Networks

IP is used ubiquitously. Some networks like industrial control bus (e.g. [RS-485], [SCADA], or even non-networked digital interfaces like [MIL-STD-1397] have begun to use IP. In these kinds of networks, the system may lack the ability to communicate with the public networks.

As another example, there may be some networks in which the equipment has the technical capability to connect to the Internet, but is prohibited by administration or just temporarily not connected. These networks may include separate financial networks, lab networks, machine-to-machine (e.g. vehicle networks), sensor networks, or even normal LANs, and can include very large numbers of addresses.

Serious disadvantages and impact on applications due to the use of ambiguous address space have been well documented in [RFC1918]. However, ULA is a straightforward way to assign the IP addresses in the kinds of networks just described, with minimal administrative cost or burden. Also, ULAs fit in multiple subnet scenarios, in which each subnet has its own ULA prefix. For example, when we assign vehicles with ULA addresses, it is then possible to separate in-vehicle embedded networks into different subnets depending on real-time requirements, device types, services and more.

However, each isolated network has the possibility to be connected in the future. Administrators need to consider the following before deciding whether to use ULAs:



- o If the network eventually connects to another isolated or private network, the potential for address collision arises. However, if the ULAs were generated in the standard way, this will not be a big problem.
- o If the network eventually connects to the global Internet, then the operator will need to add a new global prefix and ensure that the address selection policy is properly set up on all interfaces.

If these further considerations are unacceptable for some reason, then the administrator needs to be careful about using ULAs in currently isolated networks.

Operational considerations:

- o Prefix generation: Randomly generated according to the algorithms defined in [RFC4193] or manually assigned. Normally, automatic generation of the prefixes is recommended, following [RFC4193]. If there are some specific reasons that call for manual assignment, administrators have to plan the prefixes carefully to avoid collision.
- o Prefix announcement: In some cases, networks may need to announce prefixes to each other. For example, in vehicle networks with infrastructure-less settings such as Vehicle-to-Vehicle (V2V) communication, prior knowledge of the respective prefixes is unlikely. Hence, a prefix announcement mechanism is needed to enable inter-vehicle communications based on IP. As one possibility, such announcements could rely on extensions to the Router Advertisement message of the Neighbor Discovery Protocol (e.g., [I-D.petrescu-autoconf-ra-based-routing] and [I-D.jhlee-mext-mnpp]).

## 4.2. Connected Networks

### 4.2.1. ULA-Only Deployment

In some situations, hosts and interior interfaces are assigned ULAs and not GUAs, but the network needs to communicate with the outside. Two models can be considered:

- o Using Network Prefix Translation

Network Prefix Translation (NPTv6) [RFC6296] is an experimental specification that provides a stateless one-to-one mapping between internal addresses and external addresses. The specification considers translating ULA prefixes into GUA prefixes as an use case. Although NPTv6 works differently from

traditional stateful NAT/NAPT (which is discouraged in [RFC5902]), it introduces similar additional complexity to applications, which may cause applications to break.

Thus this document does not recommend the use of ULA+NPTv6. Rather, this document considers ULA+PA (Provider Aggregated) as a better approach to connect to the global network when ULAs are expected to be retained. The use of ULA+PA is discussed in detail in Section 4.2.2 below.

- o Using Application-Layer Proxies

The proxies terminate the network-layer connectivity of the hosts and associate separate internal and external connections.

In some environments (e.g., information security sensitive enterprise or government), central control is exercised by allowing the endpoints to connect to the Internet only through a proxy. With IPv4, using private address space with proxies is an effective and common practice for this purpose, and it is natural to pick ULA as its counterpart in IPv6.

Benefits of using ULAs in this scenario:

- o Allowing minimal management burden on address assignment for some specific environments.

Drawbacks:

- o The serious disadvantages and impact on applications imposed by NATs have been well documented in [RFC2993] and [RFC3027]. Although NPTv6 is a mechanism that has fewer architectural problems than a traditional stateful Network Address Translator in an IPv6 environment [RFC6296], it still breaks end-to-end transparency and hence in general is not recommended by the IETF.

Operational considerations:

- o Firewall deployment: [RFC6296] points out that an NPTv6 translator does not have the same security properties as a traditional NAT44, and hence needs be supplemented with a firewall if security at the boundary is an issue. The operator has to decide where to locate the firewall.
  - If the firewall is located outside the NPTv6 translator, then filtering is based on the translated GUA prefixes, and when the internal ULA prefixes are renumbered, the filtering rules do not need to be changed. However, when the GUA prefixes of the

NPTv6 are renumbered, the filtering rules need to be updated accordingly.).

- If the firewall is located inside the NPTv6 translator, the filtering is then based on the ULA prefixes, and the rules need to be updated correspondingly. There is no need to update when the NPTv6 GUA prefixes are renumbered.

#### 4.2.2. ULAs along with PA Addresses

Two classes of network might need to use ULA with PA (Provider Aggregated) addresses:

- o Home network. Home networks are normally assigned with one or more globally routed PA prefixes to connect to the uplink of an ISP. In addition, they may need internal routed networking even when the ISP link is down. Then ULA is a proper tool to fit the requirement. [RFC7084] requires the CPE to support ULA. Note: ULAs provide more benefit for multiple-segment home networks; for home networks containing only one segment, link-local addresses are better alternatives.
- o Enterprise network. An enterprise network is usually a managed network with one or more PA prefixes or with a PI prefix, all of which are globally routed. The ULA can be used to improve internal connectivity and make it more resilient, or to isolate certain functions like OAM for servers.

Benefits of Using ULAs in this scenario:

- o Separated local communication plane: for either home networks or enterprise networks, the main purpose of using ULAs along with PA addresses is to provide a logically local routing plane separated from the global routing plane. The benefit is to ensure stable and specific local communication regardless of the ISP uplink failure. This benefit is especially meaningful for the home network or for private OAM function in an enterprise.
- o Renumbering: in some special cases such as renumbering, enterprise administrators may want to avoid the need to renumber their internal-only, private nodes when they have to renumber the PA addresses of the rest of the network because they are changing ISPs, because the ISP has restructured its address allocations, or for some other reason. In these situations, ULA is an effective tool for addressing internal-only nodes. Even public nodes can benefit from ULA for renumbering, on their internal interfaces. When renumbering, as [RFC4192] suggests, old prefixes continue to be valid until the new prefix(es) is(are) stable. In the process

of adding new prefix(es) and deprecating old prefix(es), it is not easy to keep local communication disentangled from global routing plane change. If we use ULAs for local communication, the separated local routing plane can isolate the effects of global routing change.

Drawbacks:

- o Operational Complexity: there are some arguments that in practice the use of ULA+PA creates additional operational complexity. This is not a ULA-specific problem; the multiple-addresses-per-interface is an important feature of IPv6 protocol. Nevertheless, running multiple prefixes needs more operational consideration than running a single one.

Operational considerations:

- o Default Routing: connectivity may be broken if ULAs are used as default route. When using RIO (Route Information Option) in [RFC4191], specific routes can be added without a default route, thus avoiding bad user experience due to timeouts on ICMPv6 redirects. This behavior was well documented in [RFC7084] as rule ULA-5 "An IPv6 CE router MUST NOT advertise itself as a default router with a Router Lifetime greater than zero whenever all of its configured and delegated prefixes are ULA prefixes." and along with rule L-3 "An IPv6 CE router MUST advertise itself as a router for the delegated prefix(es) (and ULA prefix if configured to provide ULA addressing) using the "Route Information Option" specified in Section 2.3 of [RFC4191]. This advertisement is independent of having or not having IPv6 connectivity on the WAN interface.". However, it needs to be noticed that current OSes don't all support [RFC4191].
- o SLAAC/DHCPv6 co-existing: Since SLAAC and DHCPv6 might be enabled in one network simultaneously; the administrators need to carefully plan how to assign ULA and PA prefixes in accordance with the two mechanisms. The administrators need to know the current issue of the SLAAC/DHCPv6 interaction (please refer to [I-D.ietf-v6ops-dhcpv6-slaac-problem] for details).
- o Address selection: As mentioned in [RFC5220], there is a possibility that the longest matching rule will not be able to choose the correct address between ULAs and global unicast addresses for correct intra-site and extra-site communication. [RFC6724] claims that a site-specific policy entry can be used to cause ULAs within a site to be preferred over global addresses.

- o DNS relevant: if administrators choose not to do reverse DNS delegation inside of their local control of ULA prefixes, a significant amount of information about the ULA population may leak to the outside world. Because reverse queries will be made and naturally routed to the global reverse tree, so external parties will be exposed to the existence of a population of ULA addresses. [ULA-IN-WILD] provides more detailed situations on this issue. Administrators may need a split DNS to separate the queries from internal and external for ULA entries and GUA entries.

#### 4.3. IPv4 Co-existence Considerations

Generally, this document does not consider IPv4 to be in scope. But regarding ULA, there is a special case needs to be recognized, which is described in Section 3.2.2 of [RFC5220]. When an enterprise has IPv4 Internet connectivity but does not yet have IPv6 Internet connectivity, and the enterprise wants to provide site-local IPv6 connectivity, a ULA is the best choice for site-local IPv6 connectivity. Each employee host will have both an IPv4 global or private address and a ULA. Here, when this host tries to connect to an outside node that has registered both A and AAAA records in the DNS, the host will choose AAAA as the destination address and the ULA for the source address according to the IPv6 preference of the default policy table defined in the old address selection standard [RFC3484]. This will clearly result in a connection failure. The new address selection standard [RFC6724] has corrected this behavior by preferring IPv4 than ULAs in the default policy table. However, there are still lots of hosts using the old standard [RFC3484], thus this could be an issue in real networks.

Happy Eyeballs [RFC6555] solves this connection failure problem, but unwanted timeouts will obviously lower the user experience. One possible approach to eliminating the timeouts is to deprecate the IPv6 default route and simply configure a scoped route on hosts (in the context of this document, only configure the ULA prefix routes). Another alternative is to configure IPv4 preference on the hosts, and not include DNS A records but only AAAA records for the internal nodes in the internal DNS server. Then outside nodes have both A and AAAA records and can be connected through IPv4 as default and internal nodes can always connect through IPv6. But since IPv6 preference is default, changing the default in all nodes is not suitable at scale.

## 5. General Considerations For Using ULAs

### 5.1. Do Not Treat ULA Equal to RFC1918

ULA and [RFC1918] are similar in some aspects. The most obvious one is as described in Section 3.1.3 that ULA provides an internal address independence capability in IPv6 that is similar to how [RFC1918] is commonly used. ULA allows administrators to configure the internal network of each platform the same way it is configured in IPv4. Many organizations have security policies and architectures based around the local-only routing of [RFC1918] addresses and those policies may directly map to ULA [RFC4864].

But this does not mean that ULA is equal to an IPv6 version of [RFC1918] deployment. [RFC1918] usually combines with NAT/NAPT for global connectivity. But it is not necessary to combine ULAs with any kind of NAT. Operators can use ULA for local communications along with global addresses for global communications (see Section 4.2.2). This is a big advantage brought by default support of multiple-addresses-per-interface feature in IPv6. (People may still have a requirement for NAT with ULA, this is discussed in Section 4.2.1. But people also need to keep in mind that ULA is not intentionally designed for this kind of use case.)

Another important difference is the ability to merge two ULA networks without renumbering (because of the uniqueness), which is a big advantage over [RFC1918].

### 5.2. Using ULAs in a Limited Scope

A ULA is by definition a prefix that is never advertised outside a given domain, and is used within that domain by agreement of those networked by the domain.

So when using ULAs in a network, the administrators need to clearly set the scope of the ULAs and configure ACLs on relevant border routers to block them out of the scope. And if internal DNS is enabled, the administrators might also need to use internal-only DNS names for ULAs and might need to split the DNS so that the internal DNS server includes records that are not presented in the external DNS server.

## 6. ULA Usages Considered Helpful

### 6.1. Used in Isolated Networks

As analyzed in Section 4.1, ULA is very suitable for isolated networks. Especially when there are subnets in the isolated network, ULA is a reasonable choice.

### 6.2. ULA along with PA

As described in Section 4.2.2, using ULAs along with PA addresses to provide a logically separated local plane can benefit OAM functions and renumbering.

### 6.3. Some Specific Use Cases

Along with the general scenarios, this section provides some specific use cases that could benefit from using ULA.

#### 6.3.1. Special Routing

For various reasons the administrators may want to have private routing be controlled and separated from other routing. For example, in the business-to-business case described in [I-D.baker-v6ops-b2b-private-routing], two companies might want to use direct connectivity that only connects stated machines, such as a silicon foundry with client engineers that use it. A ULA provides a simple way to assign prefixes that would be used in accordance with an agreement between the parties.

#### 6.3.2. Used as NAT64 Prefix

The NAT64 PREF64 is just a group of local fake addresses for the DNS64 to point traffic to a NAT64. Using a ULA prefix as the PREF64 easily ensures that only local systems can use the translation resources of the NAT64 system since the ULA is not intended to be globally routable. The ULA helps clearly identify traffic that is locally contained and destined to a NAT64. Using ULA for PREF64 is deployed and it is an operational model.

But there is an issue needs to be noted. The NAT64 standard [RFC6146] specifies that the PREF64 should align with [RFC6052], in which the IPv4-Embedded IPv6 Address format was specified. If we pick a /48 for NAT64, it happens to be a standard 48/ part of ULA (7bit ULA well-known prefix+ 1 "L" bit + 40bit Global ID). Then the 40bit of ULA is not violated by being filled with part of the 32bit IPv4 address. This is important, because the 40bit assures the uniqueness of ULA. If the prefix is shorter than /48, the 40bit would be violated, and this could cause conformance issues. But it is considered that the most common use case will be a /96 PREF64, or

even /64 will be used. So it seems this issue is not common in current practice.

It is most common that ULA PREF64 will be deployed on a single internal network, where the clients and the NAT64 share a common internal network. ULA will not be effective as PREF64 when the access network must use an Internet transit to receive the translation service of a NAT64 since the ULA will not route across the Internet.

According to the default address selection table specified in [RFC6724], the host would always prefer IPv4 over ULA. This could be a problem in NAT64-CGN scenario as analyzed in Section 8 of [RFC7269]. So administrators need to add additional site-specific address selection rules to the default table to steer traffic flows going through NAT64-CGN. However, updating the default policy tables in all hosts involves significant management cost. This may be possible in an enterprise (using a group policy object, or other configuration mechanisms), but it is not suitable at scale for home networks.

#### 6.3.3. Used as Identifier

ULAs could be self-generated and easily grabbed from the standard IPv6 stack. And ULAs don't need to be changed as the GUA prefixes do. So they are very suitable to be used as identifiers by the up layer applications. And since ULA is not intended to be globally routed, it is not harmful to the routing system.

Such kind of benefit has been utilized in real implementations. For example, in [RFC6281], the protocol BTMM (Back To My Mac) needs to assign a topology-independent identifier to each client host according to the following considerations:

- o TCP connections between two end hosts wish to survive in network changes.
- o Sometimes one needs a constant identifier to be associated with a key so that the Security Association can survive the location changes.

It needs to be noticed again that in theory ULA has the possibility of collision. However, the probability is desirably small enough and can be ignored in most cases when ULAs are used as identifiers.



## 7. Security Considerations

Security considerations regarding ULAs, in general, please refer to the ULA specification [RFC4193]. Also refer to [RFC4864], which shows how ULAs help with local network protection.

As mentioned in Section 4.2.2, when using NPTv6, the administrators need to know where the firewall is located to set proper filtering rules.

Also as mentioned in Section 4.2.2, if administrators choose not to do reverse DNS delegation inside their local control of ULA prefixes, a significant amount of information about the ULA population may leak to the outside world.

## 8. IANA Considerations

This memo has no actions for IANA.

## 9. Acknowledgements

Many valuable comments were received in the IETF v6ops WG mail list, especially from Cameron Byrne, Fred Baker, Brian Carpenter, Lee Howard, Victor Kuarsingh, Alexandru Petrescu, Mikael Abrahamsson, Tim Chown, Jen Linkova, Christopher Palmer Jong-Hyouk Lee, Mark Andrews, Lorenzo Colitti, Ted Lemon, Joel Jaeggli, David Farmer, Doug Barton, Owen DeLong, Gert Doering, Bill Jouris, Bill Cervený, Dave Thaler, Nick Hilliard, Jan Zorz, Randy Bush, Anders Brandt, , Sofiane Imadali and Wesley George.

Some test of using ULA in the lab was done by our research partner BNRC-BUPT (Broad Network Research Centre in Beijing University of Posts and Telecommunications). Thanks for the work of Prof. Xiangyang Gong and student Dengjia Xu.

Tom Taylor did a language review and revision thought the whole document. The authors appreciate a lot for his help.

This document was produced using the xml2rfc tool [RFC2629] (initially prepared using 2-Word-v2.0.template.dot.).

## 10. References

### 10.1. Normative References

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

[RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629, June 1999.

[RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, October 2005.

## 10.2. Informative References

[I-D.baker-v6ops-b2b-private-routing]  
Baker, F., "Business to Business Private Routing", draft-baker-v6ops-b2b-private-routing-00 (work in progress), July 2007.

[I-D.ietf-v6ops-dhcpv6-slaac-problem]  
Liu, B., Jiang, S., Bonica, R., Gong, X., and W. Wang, "DHCPv6/SLAAC Address Configuration Interaction Problem Statement", draft-ietf-v6ops-dhcpv6-slaac-problem-03 (work in progress), October 2014.

[I-D.jhlee-mext-mnpp]  
Tsukada, M., Ernst, T., and J. Lee, "Mobile Network Prefix Provisioning", draft-jhlee-mext-mnpp-00 (work in progress), October 2009.

[I-D.petrescu-autoconf-ra-based-routing]  
Petrescu, A., Janneteau, C., Demailly, N., and S. Imadali, "Router Advertisements for Routing between Moving Networks", draft-petrescu-autoconf-ra-based-routing-05 (work in progress), July 2014.

[MIL-STD-1397]  
"Military Standard, Input/Output Interfaces, Standard Digital Data, Navy Systems (MIL-STD-1397B), 3 March 1989".

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

[RFC2993] Hain, T., "Architectural Implications of NAT", RFC 2993, November 2000.

[RFC3027] Holdrege, M. and P. Srisuresh, "Protocol Complications with the IP Network Address Translator", RFC 3027, January 2001.

[RFC3484] Draves, R., "Default Address Selection for Internet Protocol version 6 (IPv6)", RFC 3484, February 2003.

- [RFC3879] Huitema, C. and B. Carpenter, "Deprecating Site Local Addresses", RFC 3879, September 2004.
- [RFC4191] Draves, R. and D. Thaler, "Default Router Preferences and More-Specific Routes", RFC 4191, November 2005.
- [RFC4192] Baker, F., Lear, E., and R. Droms, "Procedures for Renumbering an IPv6 Network without a Flag Day", RFC 4192, September 2005.
- [RFC4864] Van de Velde, G., Hain, T., Droms, R., Carpenter, B., and E. Klein, "Local Network Protection for IPv6", RFC 4864, May 2007.
- [RFC5220] Matsumoto, A., Fujisaki, T., Hiromi, R., and K. Kanayama, "Problem Statement for Default Address Selection in Multi-Prefix Environments: Operational Issues of RFC 3484 Default Rules", RFC 5220, July 2008.
- [RFC5902] Thaler, D., Zhang, L., and G. Lebovitz, "IAB Thoughts on IPv6 Network Address Translation", RFC 5902, July 2010.
- [RFC6052] Bao, C., Huitema, C., Bagnulo, M., Boucadair, M., and X. Li, "IPv6 Addressing of IPv4/IPv6 Translators", RFC 6052, October 2010.
- [RFC6146] Bagnulo, M., Matthews, P., and I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", RFC 6146, April 2011.
- [RFC6281] Cheshire, S., Zhu, Z., Wakikawa, R., and L. Zhang, "Understanding Apple's Back to My Mac (BTMM) Service", RFC 6281, June 2011.
- [RFC6296] Wasserman, M. and F. Baker, "IPv6-to-IPv6 Network Prefix Translation", RFC 6296, June 2011.
- [RFC6555] Wing, D. and A. Yourtchenko, "Happy Eyeballs: Success with Dual-Stack Hosts", RFC 6555, April 2012.
- [RFC6724] Thaler, D., Draves, R., Matsumoto, A., and T. Chown, "Default Address Selection for Internet Protocol Version 6 (IPv6)", RFC 6724, September 2012.
- [RFC7084] Singh, H., Beebe, W., Donley, C., and B. Stark, "Basic Requirements for IPv6 Customer Edge Routers", RFC 7084, November 2013.

- [RFC7269] Chen, G., Cao, Z., Xie, C., and D. Binet, "NAT64 Deployment Options and Experience", RFC 7269, June 2014.
- [RS-485] "Electronic Industries Association (1983). Electrical Characteristics of Generators and Receivers for Use in Balanced Multipoint Systems. EIA Standard RS-485.".
- [SCADA] "Boyer, Stuart A. (2010). SCADA Supervisory Control and Data Acquisition. USA: ISA - International Society of Automation.".
- [ULA-IN-WILD]  
"G. Michaelson, "conference.apnic.net/data/36/apnic-36-ula\_1377495768.pdf"".

## Authors' Addresses

Bing Liu  
Huawei Technologies  
Q14, Huawei Campus, No.156 Beiqing Road  
Hai-Dian District, Beijing, 100095  
P.R. China

Email: leo.liubing@huawei.com

Sheng Jiang  
Huawei Technologies  
Q14, Huawei Campus, No.156 Beiqing Road  
Hai-Dian District, Beijing, 100095  
P.R. China

Email: jiangsheng@huawei.com

Network Working Group  
Internet Draft  
Intended status: Proposed Standard  
Expires: April 24, 2014

B. Liu  
Huawei Technologies  
R. Bonica  
Juniper Networks  
X. Gong  
W. Wang  
BUPT University  
October 21, 2013

DHCPv6/SLAAC Address Configuration Interaction Problem Statement  
draft-liu-bonica-v6ops-dhcpv6-slaac-problem-00.txt

#### 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 April 24, 2014.

#### 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.

#### Abstract

This document analyzes the host behavior of DHCPv6/SLAAC interaction issue. It reviews the standard definition of the host behaviors and provides the test results of current mainstream implementations. Some potential operational gaps of the interaction are also described.

## Table of Contents

1. Introduction .....	3
2. Host Behavior of DHCPv6/SLAAC Interaction .....	3
2.1. Relevant RA Flags Defined in Standards .....	4
2.1.1. A (Autonomous) Flag .....	4
2.1.2. M (Managed) Flag .....	4
2.1.3. O (Otherconfig) Flag .....	4
2.2. Behaviors of Current Implementations .....	5
2.2.1. A flag .....	5
2.2.2. M flag .....	5
2.2.3. O flag .....	6
3. Possible Operational Gaps of DHCPv6/SLAAC Interaction .....	6
3.1. Renumbering .....	6
3.2. Cold Start Problems .....	6
3.3. Strong Management .....	7
4. Conclusions .....	7
5. Security Considerations .....	7
6. IANA Considerations .....	7
7. References .....	7
7.1. Normative References .....	7
7.2. Informative References .....	8
8. Acknowledgments .....	8
Appendix A. Test Details of Host Behaviors .....	9
A.1 Host Initialing Behavior .....	10
A.2 Host Transition Behavior .....	11
A.3 Host Stateful/Stateless DHCPv6 Behavior .....	11
Authors' Addresses .....	12

## 1. Introduction

In IPv6, both of the DHCPv6 [RFC3315] and Neighbor Discovery [RFC4861] protocols can provide automatic IP address configuration for the hosts. They are known as stateful address auto-configuration and SLAAC (stateless address auto-configuration)[RFC4862], and are suitable for different scenarios respectively. Sometimes the two address configuration modes may be both available in one network.

In ND protocol, there is a M (ManagedFlag) flag defined in RA message, indicating the hosts there is DHCPv6 service in the network if the flag is set. And there is an O "OtherConfigFlag", if set, indicating configure information other than addresses (e.g. DNS, Route .etc) is available through DHCPv6 configuration. Moreover, there's another A (Autonomous) flag defined in ND, which indicating the hosts to do SLAAC, may also influent the behavior of hosts.

So with the A/M/O flags, the two separated address configuration modes are somehow correlated. But for some reason, the ND protocol didn't define the flags as prescriptive but only advisory. This ambiguous definition may vary the behavior of hosts when interpreting the flags. In section 2, we provided a brief test result to identify different host operating systems have taken different approaches. This would add additional complexity for both the hosts and the network management.

This draft reviews the standard definition of the above mentioned flags, and provides a test result of several major desktop operating systems' behavior. And then identifies potential requirement/gaps of DHCPv6/SLAAC interaction.

## 2. Host Behavior of DHCPv6/SLAAC Interaction

In this section, we analyzed A/M/O flags definition, and provide the test result of host behavior of interpreting these flags in mainstream operating systems implementations.

Please note that, A flag has no direct relationship with DHCPv6, but it is somewhat correlated with M/O flags.

## 2.1. Relevant RA Flags Defined in Standards

### 2.1.1. A (Autonomous) Flag

In ND Prefix Information Option, the autonomous address-configuration flag (A flag). When set indicates that this prefix can be used for stateless address configuration as specified in SLAAC.

For the host behavior, there is an explicit rule in the SLAAC specification [RFC4862]: "If the Autonomous flag is not set, silently ignore the Prefix Information option."

But when A flag is set, the SLAAC protocol didn't provide a prescriptive definition.

### 2.1.2. M (Managed) Flag

In earlier SLAAC specification [RFC2462], the host behavior of interpreting M flag is as below:

"On receipt of a valid Router Advertisement, a host copies the value of the advertisement's M bit into ManagedFlag. If the value of ManagedFlag changes from FALSE to TRUE, and the host is not already running the stateful address autoconfiguration protocol, the host should invoke the stateful address auto-configuration protocol, requesting both address information and other information. If the value of the ManagedFlag changes from TRUE to FALSE, the host should continue running the stateful address auto-configuration, i.e., the change in the value of the ManagedFlag has no effect. If the value of the flag stays unchanged, no special action takes place. In particular, a host MUST NOT reinvoke stateful address configuration if it is already participating in the stateful protocol as a result of an earlier advertisement."

But in the updated SLAAC specification [RFC4862], the relative description was removed, the reason was "considering the maturity of implementations and operational experiences. ManagedFlag and OtherConfigFlag were removed accordingly. (Note that this change does not mean the use of these flags is deprecated.)"

### 2.1.3. O (Otherconfig) Flag

As mentioned above, the situation of O flag is similar with M. In earlier SLAAC [RFC2462], the host behavior is clear:

"If the value of OtherConfigFlag changes from FALSE to TRUE, the host should invoke the stateful autoconfiguration protocol, requesting



information (excluding addresses if ManagedFlag is set to FALSE). If the value of the OtherConfigFlag changes from TRUE to FALSE, the host should continue running the stateful address autoconfiguration protocol, i.e., the change in the value of OtherConfigFlag has no effect. If the value of the flag stays unchanged, no special action takes place. In particular, a host MUST NOT reinvoke stateful configuration if it is already participating in the stateful protocol as a result of an earlier advertisement."

And there's another description of the relationship of M and O flags in [RFC2462]:

"In addition, when the value of the ManagedFlag is TRUE, the value of OtherConfigFlag is implicitly TRUE as well. It is not a valid configuration for a host to use stateful address autoconfiguration to request addresses only, without also accepting other configuration information."

## 2.2. Behaviors of Current Implementations

We did tests of current mainstream desktop/mobile operating systems on the behaviors; please refer to the appendix for details. This section illustrates the important results of the tests.

### 2.2.1. A flag

A flag is a switch to control whether to do SLAAC, and it is independent with M/O flags, in another word, A is independent with DHCPv6.

At the non-SLAAC-config state (either non-configured or DHCPv6-configured only), all the OSes acted the same with A flag, if A set, they all configured SLAAC, it is obvious and reasonable. But when SLAAC-configured, and A changed from 1 to 0, the behaviors varied, some deprecated SLAAC while some ignored the RA messages.

### 2.2.2. M flag

M is a key flag to interact ND/DHCPv6, but the host behaviors on M flag were quite different.

In our test, one OS treats the flag as instruction, it even released DHCPv6 session when M=0. But the other two just treat the flag as advisory, when SLAAC was done, it won't care about M=1, and M=0 won't cause operation for the already configured DHCPv6 addresses. Moreover, the two OSes even would not initiate DHCPv6 session until they

receives RA messages with M=1, this behavior has an implication that DHCPv6 somehow depends on ND.

Please refer to [I-D.liu-6renum-dhcpv6-slaac-switching] for more details.

### 2.2.3. O flag

In our tests, when M flag is set, the O flag is implicitly set as well; in another word, the hosts would not initial stateful DHCPv6 and stateless DHCPv6 respectively. This is a reasonable behavior.

But the O flag is not independent from A flag in some OSes. In our test, there are two OSes won't initiate stateless DHCPv6 when A flag is not set, that is to say, it is not applicable to have a ''stateless DHCPv6 only'' configuration state for some operating systems; it is also not applicable for these two OSes to switch between stateful DHCPv6 and stateless DHCPv6 (according to O flag changing from 0 to 1 or verse vice).

## 3. Possible Operational Gaps of DHCPv6/SLAAC Interaction

According to the abovementioned tests, there are possible operational issues as the following.

### 3.1. Renumbering

During IPv6 renumbering, the SLAAC-configured hosts can reconfigure IP addresses by receiving ND Router Advertisement (RA) messages containing new prefix information. The DHCPv6-configured hosts can reconfigure addresses by initialing RENEW sessions when the current addresses' lease time is expired or receiving the reconfiguration messages initialed by the DHCPv6 servers.

The above mechanisms have an implicit assumption that SLAAC-configured hosts will remain SLAAC while DHCPv6-managed hosts will remain DHCPv6-managed. But in some situations, SLAAC-configured hosts may need to switch to DHCPv6-managed, or verse vice. In [RFC6879], it described several renumbering scenarios in enterprise network for this requirement; for example, the network may split, merge, relocate or reorganize. But due to current implementations, this requirement is not applicable and has been identified as a gap in [RFC7010].

### 3.2. Cold Start Problems

If all nodes, or many nodes, restart at the same time after a power cut, the results might not consistent.

### 3.3. Strong Management

Since the host behavior of address configuration is somehow uncontrolled by the network side, it might cause gaps to the networks that need strong management (for example, the enterprise networks and the ISP CPE networks). Examples are:

- the network wants the hosts to do DHCPv6-only configuration, it is not applicable for some operating systems due to current implementation unless manually configure the hosts to DHCPv6-only model
- the hosts have been SLAAC-configured, then the network need the hosts to do DHCPv6 simultaneously (e.g. for multihoming)
- the network wants the hosts to do statelss DHCPV6-only; for example, the hosts are configured with self-generated addresses (e.g. ULA), and they also need to contact the DHCPv6 server for info-configuration

### 4. Conclusions

- The host behavior of SLAAC/DHCPv6 interaction is ambiguous in standard.
- The implementations have been varied on this issue. In [RFC4862] it is said "Removed the text regarding the M and O flags, considering the maturity of implementations and operational experiences." The description seems not true anymore.
- It is foreseeable that the un-uniformed host behavior can cause operational gaps, e.g. in renumbering and strong management.

### 5. Security Considerations

No more security considerations than the Neighbor Discovery protocol [RFC4861].

### 6. IANA Considerations

None.

### 7. References

#### 7.1. Normative References

- [RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, September 2007.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, September 2007.

## 7.2. Informative References

[RFC2462] Thomson, S. and T. Narten, "IPv6 Stateless Address Autoconfiguration", RFC 2462, December 1998.

[RFC3315] R. Droms, Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.

[RFC3736] Droms, R., "Stateless Dynamic Host Configuration Protocol (DHCP) Service for IPv6", RFC 3736, April 2004.

[RFC5887] Carpenter, B., Atkinson, R., and H. Flinck, "Renumbering Still Needs Work", RFC 5887, May 2010.

[RFC7010] Liu, B., Jiang, S., Carpenter, B., Venaas, S., and W. George, "IPv6 Site Renumbering Gap Analysis", RFC 7010, September 2013.

[RFC6879] Jiang, S., Liu, B., and B. Carpenter, "IPv6 Enterprise Network Renumbering Scenarios, Considerations, and Methods", RFC 6879, February 2013.

## 8. Acknowledgments

The test was done by our research partner BNRC-BUPT (Broad Network Research Centre in Beijing University of Posts and Telecommunications). Thanks for the hard efficient work of student Xudong Shi and Longyun Yuan.

Valuable comment was received from Sheng Jiang and Brian E Carpenter to improve the draft.

This document was prepared using 2-Word-v2.0.template.dot.

## Appendix A. Test Details of Host Behaviors

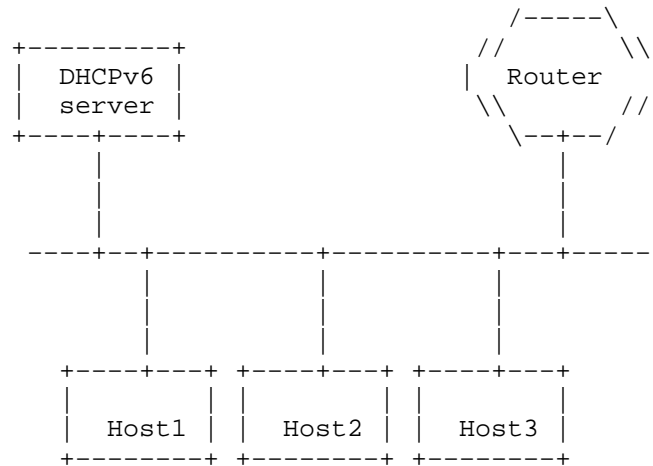


Figure 1 Test Environment

The 5 elements were all created in Vmware in one computer, for ease of operation.

- Router quagga 0.99-19 soft router installed on Ubuntu 11.04 virtual host
- DHCPv6 Server: dibbler-server installed on Ubuntu 11.04 virtual host
- Host A Window 7 Virtual Host
- Host B Ubuntu 12.10 Virtual Host
- Host C Mac OS X v10.7 Virtual Host

Another test was done dedicated for the mobile phone operating systems. The environment is similar (not in VMware, all are real PC and mobile phones):

- Router quagga 0.99-17 soft router installed on Ubuntu 12.10
  - DHCPv6 Server: dibbler-server installed on Ubuntu 12.10
  - Host D Android 4.0.4 (kernel: 3.0.16-gfa98030; device: HTC Incredible S)
  - Host E IOS 6.1.3 (model: iPod Touch 4)
- (Note: The tested Android version didn't support DHCPv6 well, so the following results don't include Android.)
- #### A.1 Host Initialing Behavior

Host from non-configured to configured, we tested different A/M/O combinations in each OS platform. The states are enumerated as the following, 3 operation systems respectively:

- o Window 7/Apple IOS
  - A=0&M=0&O=0, non-config
  - A=1&M=0&O=0, SLAAC only
  - A=1&M=0&O=1, SLAAC + Stateless DHCPv6
  - A=1&M=1&O=0, SLAAC + DHCPv6
  - A=1&M=1&O=1, SLAAC + DHCPv6
  - A=0&M=1&O=0, DHCPv6 only (A=0 or Non-PIO)
  - A=0&M=1&O=1, DHCPv6 only (A=0 or Non-PIO)
  - A=0&M=0&O=1, Stateless DHCPv6 only
- o Linux/MAC OS X
  - A=0&M=0&O=0, non-config
  - A=1&M=0&O=0, SLAAC only
  - A=1&M=0&O=1, SLAAC + Stateless DHCPv6
  - A=1&M=1&O=0, SLAAC + DHCPv6
  - A=1&M=1&O=1, SLAAC + DHCPv6
  - A=0&M=1&O=0, DHCPv6 only (A=0 or Non-PIO)
  - A=0&M=1&O=1, DHCPv6 only (A=0 or Non-PIO)
  - A=0&M=0&O=1, non-config

As showed above, Linux and MAC OSX acted the same way, but differated from Windows 7 and Apple IOS. The only difference is when A=0&M=0&O=1, Windows 7/Apple IOS did stateless DHCPv6 while Linux/MAC OSX did nothing.

Result summary:

- A is interpreted as prescript in each OS at the initial state
- M is interpreted as prescript in each OS at the initial state
- O is interpreted as prescript in Windows 7
- A and M are independent in each OS at the initial state
- A and O are not totally independent in Linux and Mac, A=1 is required for O=1 triggering DHCPv6 info-request

- M and O are not totally independent in each OS. M=1 has the implication O=1

#### A.2 Host Transition Behavior

- o SLAAC-only host receiving A=0&M=1
  - Window 7 would deprecate SLAAC and initiate DHCPv6
  - Linux/MAC/IOS would keep SLAAC and don't initiate DHCPv6 unless SLAAC is expired and no continuous RA
- o DHCPv6-only host receiving A=1&M=0
  - Window 7 would release DHCPv6 and do SLAAC
  - Linux/MAC/IOS would keep DHCPv6 and do SLAAC

When the host has been configured, either by SLAAC or DHCPv6, the operating systems interpreting the M flag quite differently. Windows 7 treats the flag as instruction, it even released DHCPv6 session when M=0. Linux and OS X were likely to treat the flag as advisory, when SLAAC was done, it won't care about M=1, and M=0 won't cause operation for the already configured DHCPv6 addresses.

Please refer to [I-D.liu-6renum-dhcpv6-slaac-switching] for more details.

#### A.3 Host Stateful/Stateless DHCPv6 Behavior

- o Stateless DHCPv6-configured host receiving M=1 (while keeping O=1)
  - Window 7 would initiate stateful DHCPv6, configuring address as well as re-configuring other information
  - Linux/MAC/IOS no action
- o Statefull DHCPv6-configured host receiving M=0 (while keeping O=1)
  - Window 7 would release all DHCPv6 configurations including address and other information, and initiate stateless DHCPv6
  - Linux/MAC/IOS no action

Authors' Addresses

Bing Liu  
Q14-4-A Building  
Huawei Technologies Co., Ltd  
Zhong-Guan-Cun Environment Protection Park, No.156 Beijing Rd.  
Hai-Dian District, Beijing  
P.R. China

Email: leo.liubing@huawei.com

Ron Bonica  
Juniper Networks  
Sterling, Virginia 20164  
USA

Email: rbonica@juniper.net

Xiangyang Gong  
No.3 Teaching Building  
Beijing University of Posts and Telecommunications (BUPT)  
No.10 Xi-Tu-Cheng Rd.  
Hai-Dian District, Beijing  
P.R. China

Email: xygong@bupt.edu.cn

Wendong Wang  
No.3 Teaching Building  
Beijing University of Posts and Telecommunications (BUPT)  
No.10 Xi-Tu-Cheng Rd.  
Hai-Dian District, Beijing  
P.R. China

Email: wdwang@bupt.edu.cn





Internet Engineering Task Force  
Internet-Draft  
Intended status: Best Current Practice  
Expires: April 24, 2014

A. Moreiras  
E. Cordeiro  
R. Santos  
NIC.br  
A. Servin  
LACNIC  
A. Acosta  
Universidad Nueva Esparta  
October 21, 2013

IPv6 Address Prefixes Reserved for Documentation  
draft-moreiras-v6ops-rfc3849bis-01

Abstract

[RFC3849] specified an IPv6 prefix to be used in documentation, in order to reduce the likelihood of conflict and confusion when relating examples of deployed systems. This prefix was reserved to be used in examples in RFCs, books, documentation, and the like. It became widely accepted and used.

Although the IPv6 documentation prefix proved to be very useful, a /32 prefix is not enough to be used to document some kinds of IPv6 deployments, such as large ISP deployments, transition techniques, and other useful examples that require longer prefixes. This document defines the allocation of a new global unicast (GUA) block and a new unique local (ULA) block, to expand the range of documentation blocks. It also updates [RFC3849].

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 April 24, 2014.

## Copyright Notice

Copyright (c) 2013 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 . . . . .	2
2. Terminology . . . . .	3
3. Problem Statement . . . . .	3
3.1. Didactic Usage . . . . .	3
3.2. Test Networks . . . . .	4
3.3. Visual Identification and Address Filtering . . . . .	4
4. IPv6 Documentation Prefixes . . . . .	4
5. Operational Implications . . . . .	4
6. Security Considerations . . . . .	5
7. IANA Considerations . . . . .	5
8. References . . . . .	5
8.1. Normative References . . . . .	5
8.2. Informative References . . . . .	5
Authors' Addresses . . . . .	6

## 1. Introduction

This document describes the IPv6 address blocks provided to be used in documentation. These blocks SHOULD be used to describe network topologies, transition techniques or other systems, in RFCs, books, videos, and documentation in general. They also MAY be used in didactic laboratories, which aim to teach IPv6 or network principles.

The first block was reserved in [RFC3849], from the address space of the Asia Pacific (APNIC) regional addressing community. Other documentation ranges have been defined in the IETF, such as example domain names described in [RFC2606], and IPv4 documentation-only address blocks described in [RFC5737]. The IPv4 ranges reserved in [RFC1918] for private use are also used in documentation, as well as the Autonomous System numbers reserved in [RFC6996].

Although the address block defined in [RFC3849] was within the range of a conventional allocation size for an Internet Service Provider, and it was expected that it could accurately match deployment scenarios, there are some situations that can't be represented accordingly with a prefix of 32 bits, such as: transition techniques, peering between multiple ISPs, IPv6 address plan for multi-regional ISPs, and others.

This situation leads to the same problem that [RFC3849] tried to address. Some documentation material, particularly some didactic material and laboratories, today is using IPv6 prefixes drawn from address blocks already allocated or assigned. A similar situation with IPv4 addresses caused problems in production environments, because of address and routing conflicts with other services.

This document reserves an additional larger IPv6 block for documentation, avoiding such problems. It does not obsolete the current IPv6 documentation block 2001:0db8::/32, since it is widely deployed. Nonetheless, it updates the current practice and specifies one larger IPv6 block, for the same use.

## 2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

## 3. Problem Statement

This document proposes a solution for the situations where the length of the current documentation prefix for IPv6 (2001:db8::/32) is not enough to represent some addressing scenarios and also proposes a specific documentation block for ULA.

### 3.1. Didactic Usage

In some didactic laboratories and materials, people are using other prefixes from Global Address space when they need networks bigger than /32. For example, if you have a lab setup where each group represents an Autonomous System (AS) the ideal situation is that each group receives a block of the same size of the smallest allocation, it means a /32 for each group. A typical scenery is to have 8 to 16 groups, each one with your own /32, that requires a /28. This scenery is used by the authors of this document in IPv6 courses in Latin America.

Some IPv6 instructors use of ULA addresses when they have to represent networks bigger than /32, but it generates confusion on

students that are struggling to understand the differences between IPv4 and IPv6. Some other instructors use non allocated or already allocated prefixes and it may lead to operational problems in the event of an example being inadvertently copied into a production environment.

The same problems may occur to ULA being used for teaching purposes, as the [RFC3849] does not define an ULA for documentation.

### 3.2. Test Networks

Some transition techniques from IPv4 to IPv6 uses the IPv4 addresses embedded in the IPv6 addresses, for example, 6rd [RFC5969], 464XLAT [RFC6877] and MAP-E [I-D.ietf-softwire-map]. Using only a /32 to test those network may generate prefixes bigger than /64 that will conflict with SLAAC mechanism, as described in [RFC6052].

New protocols development may also need test networks larger than a single /32, specially when making a large functional test to check the new protocol behavior on a big network trying to emulate a production environment.

### 3.3. Visual Identification and Address Filtering

It is important that the documentation blocks and addresses can be easily identified, specially to avoid that those address to generate problems in production networks. A simple visual identification also avoid that people will use allocated or unallocated addresses to test or teach IPv6, just because those addresses would be easier to remember.

Books and documentation that includes IPv6 addresses would have a standard to follow and that will serve their needs.

Having a large defined block for documentation will also help filtering test and documentation addresses that may leak into production networks.

## 4. IPv6 Documentation Prefixes

The blocks provided for use in documentation are: 2001:0db8::/32 (v6-TEST-NET-1), UUUU:U000::/20 [Note to RFC Editor: this address range is to be added before publication] (v6-TEST-NET-2) and FCUU:UUUU:UUU0::/44 [Note to RFC Editor: this address range is to be added before publication] (v6-TEST-NET-3).

## 5. Operational Implications

Addresses within the v6-TEST-NET-1, v6-TEST-NET-2 and v6-TEST-NET-3 SHOULD NOT appear on the public Internet and are used without any coordination with IANA or an Internet Regional Registry (RIR). Network operators SHOULD add these address blocks to the list of non-routable address spaces, and if packet filters are deployed, then this address block SHOULD be added to packet filters.

These blocks are not for local use, and the filters may be used in both local and public contexts.

## 6. Security Considerations

There are no new security considerations pertaining to this document.

## 7. IANA Considerations

IANA recorded the allocation of the IPv6 global unicast address prefix v6-TEST-NET-1 as a documentation-only prefix in the IPv6 address registry.

IANA is asked to record the allocation of v6-TEST-NET-2 prefix, within the range reserved for Global IPv6 addresses, for use as an additional documentation-only prefix, in the IPv6 address registry.

IANA is asked to record the reservation of v6-TEST-NET-3 prefix, within the range reserved for Unique Local IPv6 addresses, for use as an additional documentation-only prefix, in the IPv6 address registry.

No end party is to be assigned any of these address blocks.

## 8. References

### 8.1. Normative References

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

### 8.2. Informative References

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

[RFC2606] Eastlake, D. and A. Panitz, "Reserved Top Level DNS Names", BCP 32, RFC 2606, June 1999.

- [RFC3849] Huston, G., Lord, A., and P. Smith, "IPv6 Address Prefix Reserved for Documentation", RFC 3849, July 2004.
- [RFC5737] Arkko, J., Cotton, M., and L. Vegoda, "IPv4 Address Blocks Reserved for Documentation", RFC 5737, January 2010.
- [RFC6996] Mitchell, J., "Autonomous System (AS) Reservation for Private Use", BCP 6, RFC 6996, July 2013.
- [RFC5969] Townsley, W. and O. Troan, "IPv6 Rapid Deployment on IPv4 Infrastructures (6rd) -- Protocol Specification", RFC 5969, August 2010.
- [RFC6877] Mawatari, M., Kawashima, M., and C. Byrne, "464XLAT: Combination of Stateful and Stateless Translation", RFC 6877, April 2013.
- [RFC6052] Bao, C., Huitema, C., Bagnulo, M., Boucadair, M., and X. Li, "IPv6 Addressing of IPv4/IPv6 Translators", RFC 6052, October 2010.
- [I-D.ietf-softwire-map]  
Troan, O., Dec, W., Li, X., Bao, C., Matsushima, S., Murakami, T., and T. Taylor, "Mapping of Address and Port with Encapsulation (MAP)", draft-ietf-softwire-map-08 (work in progress), August 2013.

#### Authors' Addresses

Antonio Marcos Moreiras  
NIC.br  
Av das Nacoes Unidas 11541 7o andar  
Sao Paulo, SP 04578-000  
Brazil

Phone: +55 11 5509 3553  
Email: moreiras@nic.br

Edwin Cordeiro  
NIC.br  
Av das Nacoes Unidas 11541 7o andar  
Sao Paulo, SP 04578-000  
Brazil

Phone: +55 11 5509 3537  
Email: ecordeiro@nic.br

Rodrigo Santos  
NIC.br  
Av das Nacoes Unidas 11541 7o andar  
Sao Paulo, SP 04578-000  
Brazil

Phone: +55 11 5509 3537  
Email: rsantos@nic.br

Arturo Servin  
LACNIC  
Rambla Republica de Mexico 6125  
Montevideo 11300  
Uruguay

Phone: +598 2604 2222  
Email: aservin@lacnic.net

Alejandro Acosta  
Universidad Nueva Esparta  
Avenida Sur 7  
Caracas, Los Naranjos del Cafetal CP 1081  
Venezuela

Email: aacosta@rocketmail.com



IPv6 Operations Working Group (v6ops)  
INTERNET-DRAFT  
Intended status: Proposed Informational

H. Rafiee  
C. Meinel  
Hasso Plattner Institute  
E. Nordmark  
Arista Networks  
October 21, 2013

Expires: April 21, 2014

Interface ID lifetime Algorithms  
<draft-rafiiee-v6ops-iid-lifetime-00.txt>

## Abstract

This document introduces a framework, i.e., an application layer based lifetime [applicationiid] to enable applications maintaining their users' privacy as well as controlling the number of Interface IDs (IIDs) per network adapter. It will also explain different approaches that can be used for maintaining the lifetime of an IID. We also compare this framework to the other available mechanisms. This document also explains when to remove deprecated IP addresses from the a network interface.

## 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 Expires: February 10, 2014.

## Copyright Notice

Copyright (c) 2013 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
1.1. Problem Statement . . . . .	3
2. Conventions used in this document . . . . .	3
3. Terminology . . . . .	4
4. Generation of an Interface ID . . . . .	4
5. Lifetime Explained in RFC 4941 . . . . .	4
6. Connection Based Lifetime (layer-4) . . . . .	4
7. Application Layer based Lifetime for the Interface ID (IID) .	5
7.1. Configuring the Default values . . . . .	6
7.1.1. Deprecated Interface ID . . . . .	7
7.1.2. Configuring Default values . . . . .	7
7.2. Receiving more than one RA message . . . . .	7
7.3. Automate the process for setting the lifetime . . . . .	7
8. Threat Analysis of Application Layer based lifetime . . . . .	8
8.1. Location based tracking . . . . .	8
8.2. Obtaining confidential data . . . . .	8
9. Security Considerations . . . . .	8
10. IANA Considerations . . . . .	8
11. Conclusions . . . . .	9
12. References . . . . .	9
12.1. Normative . . . . .	9
12.2. Informative . . . . .	9
Authors' Addresses . . . . .	11

## 1. Introduction

The primary use of Network Address Translation (NAT) (RFC 4787), in IPv4, is to address IPv4 exhaustion address space. NAT made the companies and places to use other approaches in order to collect more information about their users for advertisement or other purposes. Browsers' cookie is one example of these mechanisms. To address this problem and protect user's privacy, some Internet browsers offer the use of "incognito mode" or "private browsing". In this mode, the browser does not store any cookies or it will remove all user's information as soon as the user closes its browser.

Today, IPv6 large address space, reduces the need of using NAT. IPv6 also solves the problem of end-to-end communication. This is because the IP addresses used by nodes are globally valid and can be used to directly connect to other nodes via the Internet. This means that it is easier for advertisement companies to distinguish their users only by their unique IP addresses. This is also gives this ability to attackers to obtain user's information by using simple approaches such as creating a fake website and use it as trap to find the user's IP addresses.

One possible solution might be the use of temporary Interface ID (IID). It might be the future capability of the browsers, in "incognito mode", not only to prevent storing user's information but also generate a new IID for user's connection. However, this might be a good solution to maintain user's privacy but it might be problematic, especially, if many applications wants to generate their IIDs themselves and start a connection. There will be no control over the number of valid IIDs. To address this issue, we introduce a framework, i.e., an application layer based privacy that can enable the applications to request for IIDs without involving in detail of IID generation (network layer). This document also introduces different mechanisms that may be used in order to maintain the lifetime of an Interface ID (IID) used in IPv6 networks. It then explain the deficiencies exist in these mechanisms.

### 1.1. Problem Statement

Increase the difficulty of correlating a user's activities by using different IIDs for different applications, without negatively impacting the robustness of the applications. Since there is some network overhead associated with each host having lots of IIDs at the same time, the mechanism needs to limit the number of IIDs that are in use at any given time.

## 2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 [RFC2119].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying RFC 2119 significance.

In this document the use of || indicates the concatenation of the values on either side of the sign.

### 3. Terminology

- application : An application is a process in a system, which includes all other dependent processes. Usually when an application, such as a Firefox browser, is opened in a system, there are many other processes that are called into play by this application. The meaning of application, as used in this document, is to consider the application, itself, including all of its dependent processes.

### 4. Generation of an Interface ID

This document assumes that the node generates its temporary IID by using any available algorithm as explained in [RFC4941], [StableAddresses], [RFC3972], [ra-privacy], [ssas], etc.

### 5. Lifetime Explained in RFC 4941

There are some variables in play for maintaining the lifetime of an IID. There should be no more than one valid IID per network interface, at any one time. The preferred lifetime for this IID is one day and the maximum lifetime is one week. One drawback to using this lifetime IID is that the length of time the IIDs remain in use after expiration of the lifetime ( deprecated IIDs) is left solely up to the implementers. This means that it is possible for a node to cut its connections to other nodes after the expiration of the maximum lifetime for that IID. This act could possibly cause problems for any applications that are still using that ID.

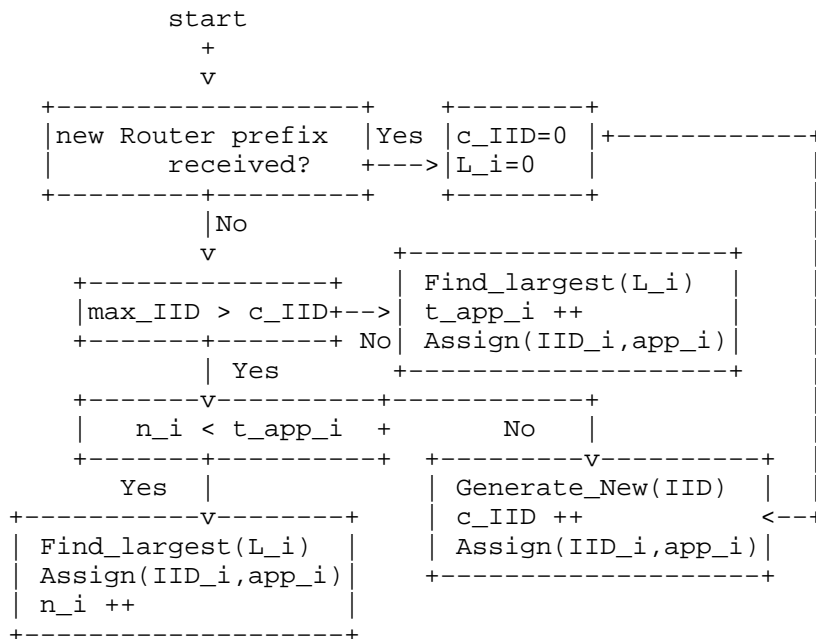
### 6. Connection Based Lifetime (layer-4)

One possible way of maintaining the lifetime of an IID is connection based, i.e., as long as the connection is active, the node keeps this IID. The drawback to using this approach is that this may prove

problematic for ftp and other applications that use different layers and open different connections.

## 7. Application Layer based Lifetime for the Interface ID (IID)

Another possible way of maintaining the lifetime of an IID is application layer based. Figure 1 depicts the algorithm used to determine the lifetime of an IID. The use of this algorithm minimizes the chance of an attacker being able to obtain a user's private information.



An IID is assigned to a group of applications and this IID is valid for the length of time that its lifetime is valid. After that time expires the state of this IID is changed to deprecated. This deprecated IID can not be assigned to new applications, but it will remain valid for as long as any application is using it. The lifetime that the deprecated IID should have is explained in section 6.2.

- app\_i is a new application that has been initiated by the node
- t\_app is the maximum number of applications per IID
- L is the maximum lifetime of an IID
- max\_IID is the maximum number of valid IIDs

- c\_IID is the current number of IIDs
- IID\_i is a specific IID
- t\_app\_i is the total number of applications allowed per specific IID
- n\_i is the current number of applications for a specific IID
- L\_i is the current lifetime of a specific IID

When a node wants to start a new application, it first checks to see whether or not it has also received a new router advertisement message. In the case where a new router advertisement message is received, the node sets the total lifetime for the current valid IID to zero and resets the c\_IID to zero. In this case, all of the current IIDs associated with this node MUST have expired and the node MUST generate, and use, new IIDs for any upcoming applications. But the node can still use an expired IID as long as the current applications using them are active. (Please refer to section 7.1.1 for more detail)

If the node does not receive a new router advertisement message, then it checks to see whether or not the current number of IIDs is less than the maximum number of IIDs allowed. If the condition is met, the node then checks to see whether or not there are any IIDs where the current number of applications, n\_i, is less than the total number of applications, t\_app\_i. If the condition is met the node SHOULD sort these IIDs based on their current lifetime, L\_i, in descending order, and then assign app\_i to the IID with the higher L\_i. The current number of applications, n\_i, for this specific IID is then incremented. If n\_i is equal to t\_app\_i, then the node generates a new IID and assigns this application to this new IID.

In cases where the current number of IIDs is equal to the max\_IID, and n\_i is equal to t\_app, then the node will be unable to generate a new IID for the application nor will it be able to assign a current IID to the application. In this case the node SHOULD find the IID with the longest lifetime and then increase the total number of applications, t\_app\_i, that can be assigned to it. The node then assigns this IID to the new application. The advantage to using this algorithm, with regard to the IID lifetime, is that it allows for the control of the number of valid IIDs while, at the same time, allowing users to keep their current application layer connections. This results in user satisfaction.

#### 7.1. Configuring the Default values

If the implementation wants to implement this mechanism, they SHOULD consider a mechanism to allow applications send their IID request to

the framework.

#### 7.1.1. Deprecated Interface ID

A Deprecated IID is an IID that is no longer valid but can still be used by any current applications that are running. It is RECOMMENDED to remove deprecated IIDs when the node's Operating System (OS) reboots or when there is no application using it for 5 minutes (no sockets). The maximum period of time that a deprecated IID can be valid is one month. Implementers should consider a way of giving users a means of setting a new default value.

#### 7.1.2. Configuring Default values

- $t_{app} = 20$
- $L = 10$  days
- $max\_IID = 10$
- $t_{app\_i} = t_{app}$  at the start of the application layer lifetime IID and SHOULD incremented this if the maximum number of IIDs is equal to current number of IIDs, as explained in the algorithm.

#### 7.2. Receiving more than one RA message

If a node receives an RA message it will assign an IID to the application or to the group of applications as described above. If, after a short period of time, another RA message is received, but with a different prefix, and if this router prefix is not in his list of the routers in his cache, then the node SHOULD send a Router Solicitation (RS) message. If it received more than one RA message with different prefixes then it SHOULD assume that these routers are in the same network. The maximum number of valid IIDs per router prefix is calculated using the following formula:

$Max\ IID\ per\ router\ prefix = max\_IID / (number\ of\ routers)$ . This mitigates the problem of multicast groups in the network as the maximum number of IIDs will never increase, regardless of number of routers in the network.

#### 7.3. Automate the process for setting the lifetime

The implementations MIGHT consider an option where, when RA messages are being processed, the RA message can be used to update the

lifetime for all the addresses that are generated using this approach. This will eliminate the need for the manual step, used during installation, to set the default value for the lifetime (based on network policy) for any future IIDs generated using this approach. The format for this lifetime value will be the same as that explained in section 5.3.1 RFC 3971. The "type" SHOULD be set to the next sequential number available in the SeND options, i.e., 15. When use is made of the lifetime option, this field SHOULD be added to the ICMPv6 option for RA messages.

## 8. Threat Analysis of Application Layer based lifetime

### 8.1. Location based tracking

Similar to the mechanism explained in RFC 4941 and [Ra-privacy], the attacker might not have enough time to track the node. This is because the IID is valid for only a short period of time and will change when a new RA message is received. Since the IIDs are valid for a short period of time, storing them using trap websites also will not help the attackers because it will be changed after a while.

### 8.2. Obtaining confidential data

If for any reason one of the applications in use on this node exposed the users' real IP address to an attacker, the attacker might not be able to obtain other confidential information related to other user applications on this node. This is because this IID is only used for certain purposes and for certain applications. This IID is also valid for only a short period of time, and so, the attacker might not have enough time to obtain confidential information about this node.%h2

## 9. Security Considerations

As is explained in the Privacy Extension document, the same approaches are used to maintain security, such as using Secure Neighbor Discovery (SeND)(RFC 3971) or using a monitoring system which would inform the administrator of the status of the network and of any suspended activities in the network.

## 10. IANA Considerations



No IANA actions are needed for this document.

## 11. Conclusions

This document explained different mechanisms used for maintaining the lifetime of an IID. It introduced an application layer based lifetime as a solution for the lifetime used for temporary IIDs in order to satisfy users needs and to not force them to cut their connections or to not open a new connection with an application using a new IID that could prove problematic for the application.

## 12. References

### 12.1. Normative References

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

[RFC4941] Narten, T., Draves, R., Krishnan, S., "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 4941, September 2007.

[RFC3972] Aura, T., "Cryptographically Generated Addresses (CGA)", RFC 3972, March 2005.

### 12.2. Informative References

[ssas] Rafiee, H., Meinel, C., "A Simple Secure Addressing Scheme for IPv6 AutoConfiguration", Work In Progress, <http://tools.ietf.org/html/draft-rafiiee-6man-ssas>, July, 2013

[StableAddresses] Gont, F., "A method for Generating Stable Privacy-Enhanced Addresses with IPv6 Stateless Address Autoconfiguration (SLAAC)", Work In Progress, <http://tools.ietf.org/html/draft-ietf-6man-stable-privacy-addresses>, August 2013

[ra-privacy] Rafiee, H., Meinel, C., "Router Advertisement based privacy extension in IPv6 autoconfiguration", Work In Progress, <http://tools.ietf.org/html/draft-rafiiee-6man-ra-privacy>, July, 2013

[applicationiid] Rafiee, H., Meinel, C., "Privacy and Security in IPv6 Networks: Challenges and Possible Solution". The 6th International Conference on Security of Information and Networks, ACM, November 2013



Authors' Addresses

Hosnieh Rafiee  
Hasso-Plattner-Institute  
Prof.-Dr.-Helmert-Str. 2-3  
Potsdam, Germany  
Phone: +49 (0)331-5509-546  
Email: ietf@rozanak.com

Erik Nordmark  
Arista Networks  
Santa Clara, CA  
USA  
Email: nordmark@acm.org

Dr. Christoph Meinel  
(Professor)  
Hasso-Plattner-Institute  
Prof.-Dr.-Helmert-Str. 2-3  
Potsdam, Germany  
Email: meinel@hpi.uni-potsdam.de



Internet Engineering Task Force  
Internet-Draft  
Updates: 6204 (if approved)  
Intended status: Standards Track  
Expires: January 31, 2014

M. Smith  
IMOT  
July 30, 2013

IPv6 CE Device DHCPv6 Option Transparency  
draft-smith-v6ops-ce-dhcpv6-transparency-00

Abstract

[RFC6204] defines basic requirements for IPv6 Customer Edge Routers, to suit residential or small office IPv6 deployments. It describes a WAN interface DHCPv6 client and LAN interface DHCPv6 server implementation model. This model constrains the set of DHCPv6 options that can be conveyed between the upstream service provider and the hosts downstream of the CE device, to those supported by both the CE device's DHCPv6 client and server. To support further current and future DHCPv6 options, this memo instead proposes a DHCPv6 option "transparent" implementation model for CE devices, primarily using DHCPv6 message relaying.

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 January 31, 2014.

Copyright Notice

Copyright (c) 2013 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 . . . . .	2
1.1. Requirements Language . . . . .	4
2. Internet Transparency and Application Configuration . . . . .	4
3. LAN Interface(s) DHCPv6 Relay Agent . . . . .	5
4. LAN Interface(s) DHCPv6 Hybrid Server/Relay . . . . .	5
5. Security Considerations . . . . .	8
5.1. Client Information Disclosure . . . . .	8
5.2. Additional State . . . . .	9
6. Acknowledgements . . . . .	9
7. Change Log [RFC Editor please remove] . . . . .	9
8. References . . . . .	9
8.1. Normative References . . . . .	9
8.2. Informative References . . . . .	9
Author's Address . . . . .	10

## 1. Introduction

[RFC6204] defines basic requirements for IPv6 Customer Edge Routers, to suit residential or small office IPv6 deployments.

For these devices, the model of operation for DHCPv6 is to operate as a client on the CE device's WAN interface and as a server on the CE device's LAN interface(s).

Operating as a client on the WAN interface, DHCPv6 is used for two purposes. Firstly, the DHCPv6 client is used to acquire configuration parameters useful or necessary to the CE device itself. Examples of these parameters are an IPv6 address for the WAN interface, DNS server information, and Simple Network Time Server information. Secondly, the DHCPv6 client is used to acquire configuration parameters that are either essential or useful for the operation of the downstream hosts, such as a delegated IPv6 prefix and DNS domain information.

Operating as a server on the LAN interface(s), DHCPv6 is used for purposes such as stateful IPv6 address assignment (if supported by the CE device and requested by the hosts), providing hosts with DNS resolver and DNS domain information, and possibly being able to provide other DHCPv6 options such as SIP server addresses.

This client and server model of operation of DHCPv6 on the CE device constrains the use of DHCPv6 options between the downstream host(s) and the upstream service provider. The DHCPv6 options that a downstream host can request, and that an upstream service provider can supply, is limited to the set of options supported by both the CE device's WAN DHCPv6 client and LAN DHCPv6 server. This set of supported DHCPv6 options is significantly limited compared to the current set of available DHCPv6 options [IANA-DHCPV6-OPTIONS]. Additionally, it cannot accommodate DHCPv6 options that may be defined in the future. This means the CE device is not DHCPv6 option "transparent".

The main consequence of a lack of DHCPv6 option transparency is that users or operators of the hosts downstream of the CE device will have to resort to manual configuration of application and host parameters, for information that could be provided by the CE device's unsupported DHCPv6 options. Manual configuration is time consuming, error prone, static, not scalable and most importantly, not user friendly.

In markets where service provider customers can supply their own CE device, the CE device supported DHCPv6 options may vary significantly across the customers' devices. This could create a perverse incentive for the service provider to not use DHCPv6 options at all, other than the minimum necessary, and instead have customers use manual configuration. The motive to do so would be for the service provider to have a single and consistent method of configuration, simplifying customer fault troubleshooting, despite the other drawbacks of manual configuration described previously. CE device DHCPv6 option transparency would avoid creating this incentive for manual configuration.

A DHCPv6 Relay Agent [RFC3315] is DHCPv6 option transparent. It does not constrain the set of options that can be used between a downstream DHCPv6 client and an upstream DHCPv6 server, as its operation does not depend on being able to interpret the options conveyed between the client and the server. Consequently a DHCPv6 relay agent inherently supports all current and future DHCPv6 options.

To overcome a CE device's lack of DHCPv6 option transparency, this memo proposes the use of a DHCPv6 relay agent for processing of stateless DHCPv6 requests, and a hybrid mode of operation for stateful DHCPv6 requests, where the IPv6 addressing related options are processed locally, and other options requested by hosts are acquired from a DHCPv6 server via a relayed, synthesised Information-request.

### 1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

## 2. Internet Transparency and Application Configuration

A variety of RFCs discuss Internet transparency, and its benefits [RFC1958] [RFC3724] [RFC4924]. Internet transparency essentially means that the Internet is transparent to the applications using it, such that deploying a new application only requires installing the application on the hosts which will be executing it. The devices (routers) within the Internet remain "oblivious" to the new application's protocols, and consequently immediately facilitate its use. The benefits of this model are significant; having to make fundamental changes or upgrades to one or more routers' operation to support a new application would likely be disruptive to the devices currently using the routers. All routers that may potentially carry the new application's traffic would need to be upgraded. Furthermore, the software or firmware upgrades required to support the application may not currently or may never be available from the router vendor. These network upgrades would significantly delay, or perhaps make impossible the deployment of new applications. Application awareness in the Internet limits its flexibility and generality.

This Internet transparency model should also be followed by protocols and the deployment models used to provide application parameters. If a device within the network is going to participate in an application configuration protocol, it should do so in a manner that is transparent to the application(s) configuration.

The DHCPv6 protocol supports this transparency model. DHCPv6 clients and servers, located on the hosts at the edge of the network, are option and application parameter aware. A DHCPv6 relay, located within the network, does not need to understand the DHCPv6 options the clients and servers use to be able to relay them. When a new DHCPv6 configured application is to be deployed, only the DHCPv6 clients and servers involved in configuring the application need to be upgraded. Widespread upgrades of DHCPv6 relays within the Internet are not required.

The DHCPv6 deployment model described in [RFC6204] does not provide application configuration transparency, as it uses the combination of a DHCPv6 client and server, rather than a DHCPv6 relay, to convey service provider DHCPv6 options to hosts downstream of the CE device. The DHCPv6 options that can be conveyed "across" the CE device are



limited to those understood by both the co-located DHCPv6 client and server on the CE device. A DHCPv6 relay deployment model would be better for the types of CE devices described in [RFC6204], and devices located within the Internet in general.

The addition of DNS related options to RAs [RFC6106] is also venturing down the path of violating network application configuration transparency. Subsequent to these options being added to RAs, there has been at least one proposal to add a further application related option to RAs that is already present in DHCPv6. Following a model of using RAs to configure applications will result in network located constraints on application deployment, as described previously. Proposals for further application configuration options in RAs should be resisted. RA options should be limited to configuring network layer parameters, relevant to a single link or subnet, with DHCPv6 being used to configure higher layer transport and application layer parameters. This will preserve application configuration transparency in the Internet.

### 3. LAN Interface(s) DHCPv6 Relay Agent

When SLAAC [RFC4862] is being used for LAN interface IPv6 address autoconfiguration, it is likely that stateless DHCPv6 [RFC3736] will be used by hosts to acquire other application-oriented configuration parameters.

Instead of using a DHCPv6 server to answer hosts' DHCPv6 Information-request messages, a CE device uses DHCPv6 relay functionality [RFC3315] to relay the Information-request message out of its WAN interface, towards the service provider's DHCPv6 server infrastructure. The service provider's DHCPv6 service infrastructure will either not respond, or provide values for some or all of the options requested by the DHCPv6 client.

(I wonder if it would be useful to be able to supply one or more LAN interface relay agent target addresses during the DHCPv6-PD transaction on the WAN interface, via a currently undefined DHCPv6 option. This would allow the service provider to use different DHCPv6 server infrastructure to answer these LAN interface originated DHCPv6 Information-requests verses what they use for the CE device's WAN interface DHCPv6-PD transaction. This could provide the benefit of both minimising the DHCPv6 configuration complexity on the SP router, and as relaying or answering of DHCPv6 is going to be a router control plane function, will reduce the router control plane load.)

### 4. LAN Interface(s) DHCPv6 Hybrid Server/Relay

To support stateful IPv6 address assignment on the LAN interface(s) (perhaps in addition to stateless IPv6 address assignment [RFC4862]), while also remaining DHCPv6 option transparent, a hybrid mode of DHCPv6 operation is necessary. This hybrid mode involves locally processing the IPv6 address assignment related DHCPv6 options, while acquiring other option information from the upstream service provider's DHCPv6 infrastructure, using DHCPv6 relay functionality. To the LAN attached hosts, the DHCPv6 Hybrid Server/Relay appears to be a local stateful DHCPv6 server. The hybrid operation is transparent to the DHCPv6 clients.

In this section, the following terminology is used:

- o DHCPv6 upstream server - the DHCPv6 server and related DHCPv6 infrastructure located within the upstream service provider's network.
- o DHCPv6 hybrid server - the DHCPv6 server that is available on the CE device's LAN interface(s).
- o DHCPv6 hybrid transaction - the transaction where IPv6 addressing related DHCPv6 options are processed locally by the server, while other option information is acquired from a DHCPv6 upstream server.
- o DHCPv6 client - an IPv6 host attached to the CE device's LAN interface, which uses the services of the DHCPv6 hybrid server.

The DHCPv6 hybrid transaction occurs when the DHCPv6 hybrid server is preparing a Reply response to a client generated DHCPv6 Solicit (when Rapid Commit is in use), Request, Renew or Rebind message. At this point in the transaction, if the message from the DHCPv6 client contains an Option Request Option, the DHCPv6 hybrid server synthesises a DHCPv6 Information-request message [RFC3736] on behalf of the DHCPv6 client. The synthesised Information-request is constructed using the permitted subset of DHCPv6 options received from the DHCPv6 client. Specifically, the IA options and other non-relevant options are excluded. An Information Refresh option [RFC4242] code is added to the Option Request Option in the Information-request, to acquire option refresh time information [RFC4076].

The synthesised Information-request is then relayed to the DHCPv6 upstream server using standard DHCPv6 relay functionality [RFC3315]. To detect lack of a response to the relayed Information-request, the DHCPv6 hybrid server also starts a count down to zero timer, with an initial value of INF\_TIMEOUT [RFC3315].

If the DHCPv6 hybrid server does not receive a Reply before the count down timer reaches zero, the DHCPv6 hybrid server abandons its knowledge of both the DHCPv6 client's original DHCPv6 message, and any state related to the synthesised Information-request. Recovery from the lack of response from the DHCPv6 upstream server is left to the DHCPv6 client; it will eventually retransmit its Solicit, Request, Renew or Rebind message, restarting the DHCPv6 hybrid transaction, or will give up completely on the whole stateful DHCPv6 transaction.

While the synthesised Information-request is being relayed (i.e., within INF\_TIMEOUT from when the Information-request was sent), the DHCPv6 client may timeout its original message and therefore retransmit it. These retransmissions are to be silently dropped. They can be recognised by the DHCPv6 client's reuse of the same DHCPv6 transaction ID.

If the DHCPv6 hybrid server receives a Reply to the relayed Information-request, it uses the enclosed information to populate the set of options that will be sent to the DHCPv6 client. The received options the DHCPv6 hybrid server provides to the DHCPv6 client are the options that fall within the set the DHCPv6 client originally requested using an Option Request Option. Other options received in the relayed Reply are not sent to the DHCPv6 client. They may be used by the DHCPv6 hybrid server for other purposes.

If an Information Refresh option is received in the relayed Reply, the DHCPv6 hybrid server should use the supplied refresh time to cause the DHCPv6 client to refresh its option information at the appropriate time. There are three possible ways this can be achieved, with the first being most preferred.

Firstly, if the DHCPv6 client has indicated Reconfiguration support, the DHCPv6 hybrid server should send a Reconfigure message to the DHCPv6 client after the refresh time interval, as per the procedure in [RFC3315]. This should cause the DHCPv6 client to initiate a Renew/Reply transaction. When responding to the Renew message, the DHCPv6 hybrid server performs the DHCPv6 hybrid transaction to acquire up-to-date option values.

If the DHCPv6 client does not support Reconfiguration, one of two mechanisms can be used, depending on whether the IPv6 addresses provided to DHCPv6 clients are non-temporary or temporary addresses.

For non-temporary addresses, supplied using IA\_NA options, the T1 time interval [RFC3315] should be set to the refresh time interval, if it is lower than the DHCPv6 hybrid server's normal T1 time interval. This will cause the client to initiate a Renew/Reply

transaction at time T1. The DHCPv6 hybrid server then performs the DHCPv6 hybrid transaction when preparing the Reply. The T2 value is derived from the T1 value, as per the advice in [RFC3315].

For temporary addresses, supplied using IA\_TA options, the refresh time interval is used to set the preferred lifetime values of all addresses supplied using the IA Address options. This should cause the client to initiate a Renew/Reply transaction when any of the temporary addresses becomes deprecated, at which time the DHCPv6 hybrid server performs the DHCPv6 hybrid transaction when preparing the Reply.

If the DHCPv6 client does not support Reconfiguration, and the DHCPv6 hybrid server supplies both temporary and non-temporary addresses, then the non-temporary address method for causing clients to refresh their option information should be used.

## 5. Security Considerations

### 5.1. Client Information Disclosure

In the existing CE device DHCPv6 WAN client/LAN server model, messages sent by DHCPv6 clients to the LAN DHCPv6 server are processed locally. This means that any client supplied options that provide client specific attributes, such as the Client Identification Option, the Vendor Class Option or the Vendor-specific Information Option, are not sent to the upstream provider.

In the DHCPv6 relay models presented in this memo, client supplied information is now provided by the CE device to the upstream service provider.

This is not a new risk to DHCPv6 clients. Unless a DHCPv6 client uses DHCPv6 authentication, there is always a possibility that the client will be supplying information about itself to possibly an unknown and perhaps untrusted operator of an available DHCPv6 server.

If a client wishes to avoid classification or unique identification, it should avoid supplying options which disclose client specific attributes. A client may choose to supply these client specific attributes only when DHCPv6 authentication is being used and the DHCPv6 server is known and trusted.

A client needs to consider the benefits and drawbacks of supplying client specific information. If it supplies this information, it may receive client specific option values, resulting in useful service or application benefits. If it does not supply this information, it is likely to receive a default and possibly a less beneficial service.

A client must provide a Client Identifier Option within its DHCPv6 messages, containing a DHCP Unique Identifier (DUID) [RFC3315]. DUIDs formed using [RFC3315] methods contain client specific attributes. To minimise this attribute disclosure, a DHCPv6 client should use the UUID-based form of DUID (DUID-UUID) [RFC6355], with a version 4 UUID [RFC4122] created using a truly random or pseudo-random number. This should uniquely identify the client to the DHCPv6 server, while avoiding providing any other client attribute information.

## 5.2. Additional State

When performing the DHCPv6 Hybrid Server/Relay method, additional state is created while the CE device's DHCPv6 server is relaying the synthesised DHCPv6 Information-request. The amount of additional state is proportional to the number of client DHCPv6 requests for which Information-requests are synthesised and relayed. This state could be a target for a state capacity exhaustion attack (more generally, a resource exhaustion attack) from a malicious or misbehaving DHCPv6 client. Existing methods used to protect a DHCPv6 server from state capacity exhaustion attacks should also be used to protect this additional state.

## 6. Acknowledgements

Review and comments were provided by YOUR NAME HERE!

This memo was prepared using the xml2rfc tool.

## 7. Change Log [RFC Editor please remove]

draft-smith-v6ops-ce-dhcpv6-transparency-00, initial version,  
2013-07-30

## 8. References

### 8.1. Normative References

[IANA-DHCPV6-OPTIONS]

Internet Assigned Numbers Authority, "DHCP Option Codes",  
2013, <<http://www.iana.org/assignments/dhcpv6-parameters/dhcpv6-parameters.xhtml#dhcpv6-parameters-2>>.

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

### 8.2. Informative References

- [RFC1958] Carpenter, B., "Architectural Principles of the Internet", RFC 1958, June 1996.
- [RFC3315] Droms, R., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.
- [RFC3724] Kempf, J., Austein, R., IAB, "The Rise of the Middle and the Future of End-to-End: Reflections on the Evolution of the Internet Architecture", RFC 3724, March 2004.
- [RFC3736] Droms, R., "Stateless Dynamic Host Configuration Protocol (DHCP) Service for IPv6", RFC 3736, April 2004.
- [RFC4076] Chown, T., Venaas, S., and A. Vijayabhaskar, "Renumbering Requirements for Stateless Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 4076, May 2005.
- [RFC4122] Leach, P., Mealling, M., and R. Salz, "A Universally Unique Identifier (UUID) URN Namespace", RFC 4122, July 2005.
- [RFC4242] Venaas, S., Chown, T., and B. Volz, "Information Refresh Time Option for Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 4242, November 2005.
- [RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, September 2007.
- [RFC4924] Aboba, B. and E. Davies, "Reflections on Internet Transparency", RFC 4924, July 2007.
- [RFC6106] Jeong, J., Park, S., Beloeil, L., and S. Madanapalli, "IPv6 Router Advertisement Options for DNS Configuration", RFC 6106, November 2010.
- [RFC6204] Singh, H., Beebee, W., Donley, C., Stark, B., and O. Troan, "Basic Requirements for IPv6 Customer Edge Routers", RFC 6204, April 2011.
- [RFC6355] Narten, T. and J. Johnson, "Definition of the UUID-Based DHCPv6 Unique Identifier (DUID-UUID)", RFC 6355, August 2011.

Author's Address

Mark Smith  
In My Own Time  
PO BOX 521  
HEIDELBERG, VIC 3084  
AU

Email: markzzzsmith@yahoo.com.au

Network Working Group  
Internet-Draft  
Intended status: Informational  
Expires: April 24, 2014

C. Xie  
Q. Sun  
China Telecom  
C. Zhou  
Huawei Technologies  
October 21, 2013

Address Management for IPv6 Transition  
draft-sun-v6ops-openv6-address-pool-management-00

Abstract

This document proposes a mechanism to manage the address pools centrally. Different transition mechanisms can require the address pools on-demand. Therefore, carriers does not need to configure the address pools one by one manually and it also help to use the address pools more efficiently.

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 April 24, 2014.

Copyright Notice

Copyright (c) 2013 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. Terminology . . . . .	3
3. Overall Procedure . . . . .	3
3.1. Initial Address Pool Configuration . . . . .	4
3.2. Address Pool Status Report . . . . .	6
3.3. Address Pool Status Query . . . . .	7
3.4. Run Out of Address . . . . .	7
3.5. Address Pool Release . . . . .	7
4. Deployment Consideration . . . . .	9
5. Security Considerations . . . . .	9
6. Acknowledgements . . . . .	9
7. References . . . . .	9
7.1. Normative References . . . . .	9
7.2. Informative References . . . . .	9
Authors' Addresses . . . . .	9

## 1. Introduction

Network address migration to IPv6 is ongoing or upcoming throughout the world due to the lack of IPv4 addresses. When carriers are facing with address shortage problem, the remaining IPv4 address pools are usually quite scattered. It is quite complicated for a carrier to manage scattered address pools in many transition devices. The situation will become even worse when multiple transition mechanisms in the same device need to be configured with different address pools. Besides, since the occupation of the address pools may vary during different transition periods, (e.g. at the early stage of IPv6 transition, IPv4 traffic will normally occupy a great portion of the total traffic, while in the later stage of IPv6 transition, IPv4 traffic will decrease and the amount of IPv4 address pools will decrease accordingly.

In this document, we propose a mechanism to manage the address pools centrally. Different transition mechanisms can require the address pools on-demand. For example, when one transition mechanism is running out of the current address pools, it may request a additional address pool. It can also release the address pools that it is not using anymore. In this way, carriers does not need to configure the address pools one by one manually and it also help to use the address pools more efficiently.

## 2. Terminology

The following terms are used in this document:

## 3. Overall Procedure

This mechanism consists of two components: Address Management Server(AMS) and Transition Device (TD). AMS is responsible for address pool management while the TD will do traditional transition mechanisms, e.g. DS-Lite , NAT64, etc. The overall procedure is as follows:

- o Operators will configure remaining address pools centrally in the Address Management Server. There are multiple address pools which can be configured centrally. The AMS will then divide the address pools with addressing unit (AU) which will be allocated to transition device by default.

- o Transition Device will initiate AddressPool request to the AMS. It can carry its desired size of address pool the request, or just use a default value. The address pool size in the TD's request is only used as a hint. The actual size of the address pool is totally determined by AMS. It will also carry the TD's identification and the type of transition mechanism.
- o AMS lookups the remaining address pool in its local database. It will then allocate a set of address pools to the TD. Each address pool has a related lifetime.
- o TD receives the AddressPool reply and use them for the transition mechanisms.
- o If the lifetime of the address pool is going to expire, the TD should issue an AddressPoolRenew request to extend the lifetime.
- o The AddressPoolReport module keeps monitoring and reports the current usage of the current address pools for each transition mechanism. if one transition mechanism is running out of address pools, it can renew the AddressPoolRequest for a new one. It can also release an existing address pool if the that address pool has not been used for a long time.
- o When the status of AMS is lost or the AMS needs the status information of certain applications, the AMS may actively query the TD for the status information.

The following sub-section describes the detailed procedures of the address pool management.

### 3.1. Initial Address Pool Configuration

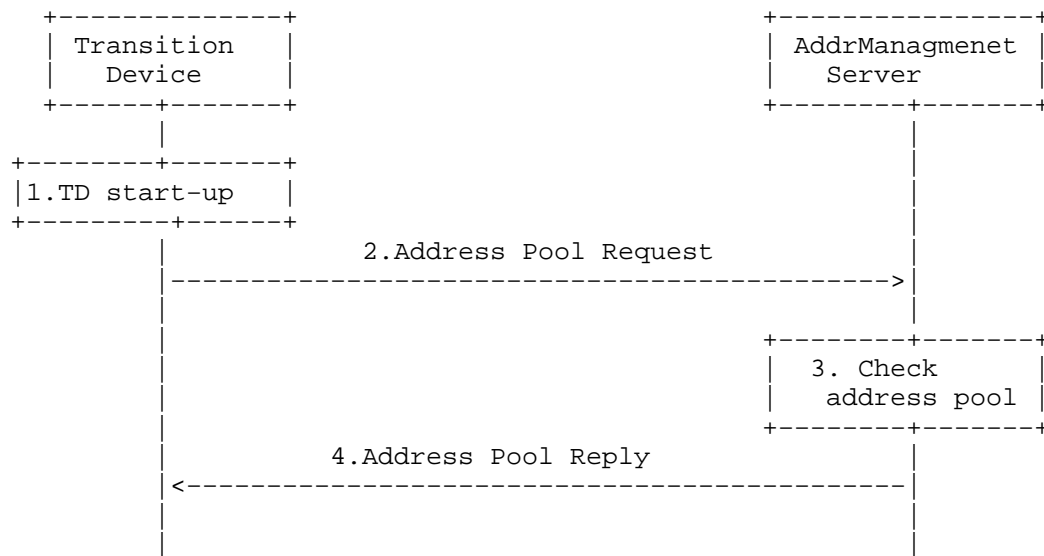


Figure 1: Initial Address Pool Configuration

Figure 1 illustrates the initial address pool configuration procedure:

1. The TD checks whether there is already address pool configured in the local site when it starts up. if no, it means the initial start-up or the address pool has been released. if yes, the address pool could be used directly.
2. The TD will initiate Address Pool request to the AMS. It can carry its desired size of address pool in the request, or just use a default value. The address pool size in the TD's request is only used as a hint. The actual size of the address pool is totally determined by AMS. It will also carry the TD's identification, the type of transition mechanism and the indication of port allocation support.
3. The AMS determines the address pool allocated for the TD based on the parameters received.
4. The AMS sends the Address Pool Reply to the TD. It will also distribute the routing entry of the address pool automatically. In particular, if the newly received address pool can be aggregated to an existing one, the routing should be aggregated accordingly.

### 3.2. Address Pool Status Report

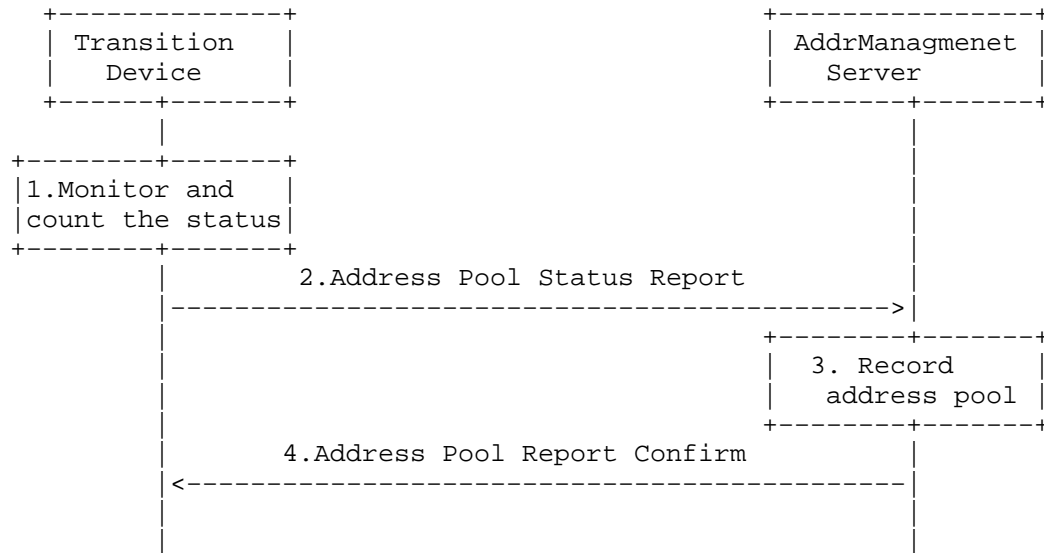


Figure 2: Address Pool Status Report

Figure 2 illustrates the active address pool status report procedure:

1. The TD will monitor and count the usage status of the local address pool. The TD counts the address usage status in one month, one week and one day, which includes the local address, address usage ratio (peak and average values), and the port usage ratio (peak and average values).
2. The TD reports the address pool usage status to the AMS. for example, it will report the address usage status in one day, which contains the IP address, NAT44, address list: 30.14.44.0/28, peak address value 14, average address usage ratio 90%, TCP port usage ratio 20%, UDP port usage ratio 30% and etc.
3. The AMS records the status and compares with the existing address information to determine whether additional address pool is needed.
4. The AMS will confirm the address pool status report request to the TD. It will keep sending the address pool status report request to the AMS if no confirm message is received.

### 3.3. Address Pool Status Query

When the status of AMS is lost or the AMS needs the status information of the TDs, the AMS may actively query the TD for the status information, as shown in step 1 of Figure 3. The following steps 2,3,4,5 are the same as the Address Pool Status Report procedure.

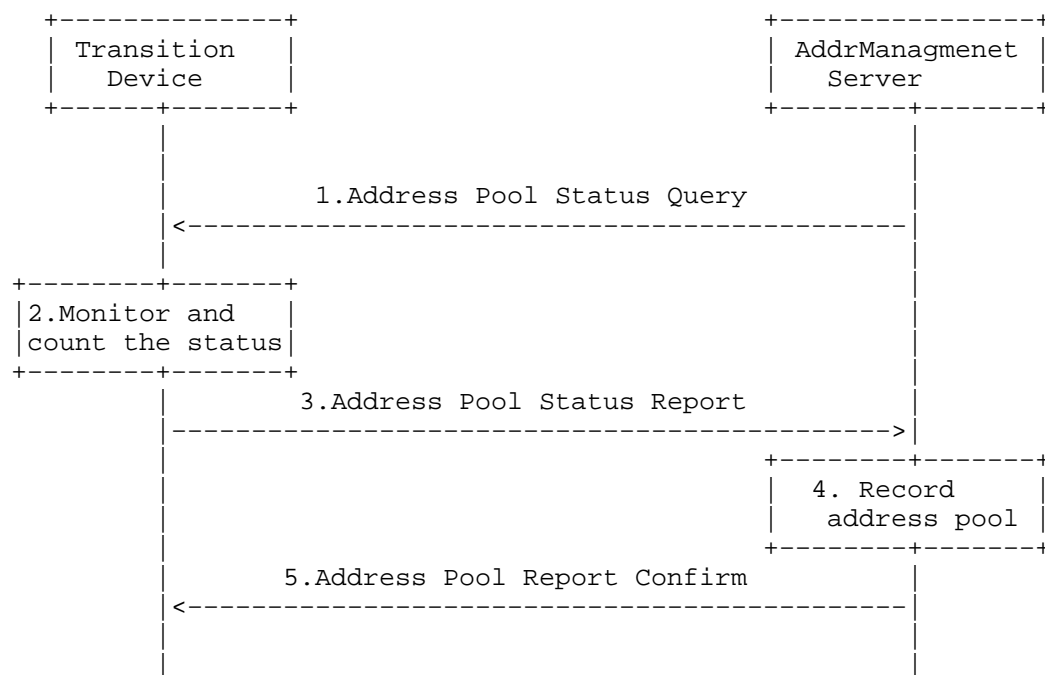


Figure 3: Address Pool Status Query

### 3.4. Run Out of Address

When the TD used up the addresses allocated, it will renew the address pool request to the AMS for additional address pool. The procedure is the same as the initial address pool request.

### 3.5. Address Pool Release

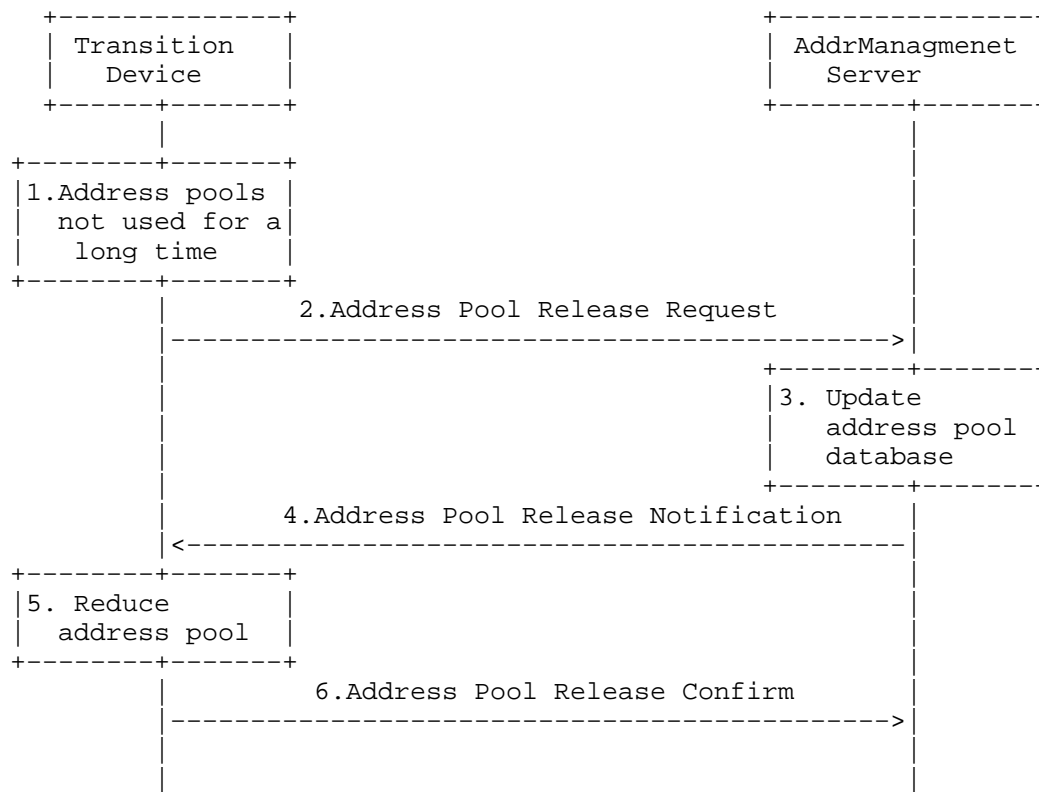


Figure 4: Address Pool Release

Figure 4 illustrates the address pool release procedure:

1. The counting module in the TD checks that there are addresses not used for a long time;
2. The TD sends the address pool release request to the AMS to ask the release of those addresses;
3. The AMS updates the local address pool information to add the new addressed released.
4. The AMS notifies the TD that the addresses have been release successfully;
5. The TD will update the local address pool. if no Address Pool Release Notification is received, the TD will repeat step 2;

6. The TD confirms with the AMS that the address pool has been released successfully.

#### 4. Deployment Consideration

In actual network, the AMS can be deployed centrally, e.g. centralized in one province. One AMS can take management on the TDs of the whole area. The requests between ASM and TD would not be too frequent and the lifetime of the address pool can be relatively long.

#### 5. Security Considerations

#### 6. Acknowledgements

N/A.

#### 7. References

##### 7.1. Normative References

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

##### 7.2. Informative References

- [RFC6674] Brockners, F., Gundavelli, S., Speicher, S., and D. Ward, "Gateway-Initiated Dual-Stack Lite Deployment", RFC 6674, July 2012.
- [RFC6888] Perreault, S., Yamagata, I., Miyakawa, S., Nakagawa, A., and H. Ashida, "Common Requirements for Carrier-Grade NATs (CGNs)", BCP 127, RFC 6888, April 2013.

#### Authors' Addresses

Chongfeng Xie  
China Telecom  
No.118 Xizhimennei street, Xicheng District  
Beijing 100035  
P.R. China

Email: xiechf@ctbri.com.cn



Qiong Sun  
China Telecom  
No.118 Xizhimennei street, Xicheng District  
Beijing 100035  
P.R. China

Email: [sunqiong@ctbri.com.cn](mailto:sunqiong@ctbri.com.cn)

Cathy Zhou  
Huawei Technologies  
Bantian, Longgang District  
Shenzhen 518129  
P.R. China

Email: [cathy.zhou@huawei.com](mailto:cathy.zhou@huawei.com)