Internet Engineering Task Force (IETF)
Request for Comments: 7269
Category: Informational
ISSN: 2070-1721

G. Chen
Z. Cao
China Mobile
C. Xie
China Telecom
D. Binet
France Telecom-Orange

June 2014

NAT64 Deployment Options and Experience

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 document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7269.

Table of Contents

1. Introduction

IPv6 is the only sustainable solution for numbering nodes on the 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 network 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 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 function [RFC6145] [RFC6146] 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. Two operation modes (i.e., stateful translation and stateless translation) have been described in Section 3.2 of [RFC6144]. This document describes the usage of those two operation modes and has further categorized different NAT64 functions, locations, and use cases. The principal distinction of location is whether the NAT64 is located in a Carrier-Grade NAT or server Front End. The terms "NAT-CGN" and "NAT-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 it 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. NAT64 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 [MOTIVATION], but it has to be coordinated with Address plus Port (A+P) processes [RFC6346] as specified in [MAP-T] in order to deal with an IPv4 address shortage.

3.1.2. DNS64 Deployment

DNS64 [RFC6147] is recommended for use in combination with stateful NAT64, and it 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 prefixes through [RFC7050]. Berkeley Internet Name Daemon (BIND) software supports that 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 services 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-facing interfaces of NAT64. We tested on the top 100 websites (referring to [Alexa] statistics). 43% of websites are connected and forwarded on 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 that the DNS64-DNSSEC interaction [RFC6147] may impact validation of Resource Records retrieved from the DNS64 process. In particular, DNSSEC validation

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

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 viewpoint 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 Packet 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 vendors' 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 a 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 geolocation based on IPv4 addresses rather inaccurate as is currently the case for NAT44 CGNs already deployed in ISP networks. More considerations or workarounds on HA and traceability can be found in Sections 4 and 5.

3.1.4. Coexistence of NAT64 and NAT44

NAT64 will likely coexist 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 addresses and IPv4-allocated addresses (which 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, a "Happy Eyeballs" [RFC6555] algorithm 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 it 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 services, but at different times [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 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 [CONN-STATUS] 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 address exhaustion may not be a problem in an IDC's internal network. IPv6 support for some applications may require increased investment and workload, so IPv6 support may not be a priority. NAT64 can be used to support widespread IPv6 adoption on the Internet while maintaining access to IPv4-only applications.

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

- 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 to be delivered without having to consider IPv4 NAT or the address limitations of IPv4 networks. Stateless NAT64 [RFC6145] can used to provide services for IPv4-only customers. [SIIT] has provided further descriptions and guidelines.

- ICPs who attempt to offer customers IPv6 support in their application farms at an early stage will likely run proxies, load balancers, or translators that 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 as more users with IPv6 connectivity access IPv4 networks. 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 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 the communication path. For testing purposes, operators could employ an independent subdomain, e.g., ipv6exp.example.com, to identify experimental IPv6 services to users. How to design the Fully Qualified Domain Name (FQDN) for the IPv6 service is outside the scope of this document.

4. High Availability

4.1. Redundancy Design

High Availability (HA) is a major requirement for every service and network service. Deploying redundancy mechanisms is essential to avoiding failure and significantly increasing the network reliability. It’s useful not only to stateful NAT64 cases but also to stateless NAT64 gateways.

Three redundancy modes are mainly used: Cold Standby, Warm Standby, and Hot Standby.

- 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.

- 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. The Virtual Router Redundancy Protocol (VRRP)[RFC5798] can be a solution to enable automatic handover during Warm Standby. During testing, the handover took a maximum of 1 minute if the backup NAT64 had to take over routing and reconstruct the Binding

Information Bases (BIBs) for 30 million sessions. In the deployment phase, operators could balance loads on distinct NAT64 devices. Those NAT64 devices make a warm backup of each other.

- Hot Standby must synchronize the BIBs between the primary NAT64 and backup. When the primary NAT64 fails, the backup NAT64 takes over and maintains the state of all existing sessions. The internal hosts don't have to reconnect the external hosts. The handover time is extremely reduced. During testing that employed Bidirectional Forwarding Detection (BFD) [RFC5880] combined with VRRP, a handover time of only 35 ms for 30 million sessions was observed. Under ideal conditions, Hot Standby deployments could guarantee the session continuity for every service. In order to transmit session states in a timely manner, operators may have to deploy extra transport links between the primary NAT64 and the distant backup. The scale of synchronization of the data instance depends on the particular deployment. For example, if a NAT64-CGN serves 200,000 users, an average amount of 800,000 sessions per second is a rough estimate of the newly created and expired sessions. A physical 10 Gbit/s transport link may have to be deployed for the sync data transmission considering the amount of sync sessions at the peak and the capacity redundancy.

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

Our statistics in a mobile network shown that almost 91.21% 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) when deploying a suitable redundancy mode. Warm Standby may still be adopted to cover most services, while Hot Standby could be used to upgrade the Quality of Experience (QoE) and 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 can be achieved. Stateless NAT64s allow asymmetric routing, while anycast-based solutions are recommended in [MAP-DEPLOY]. The deployment of load balancing may make more sense to stateful NAT64s for the sake of avoiding single-point failures. Since the NAT64-CGN and NAT64-FE have distinct facilities, the following lists the considerations for each case.

- NAT64-CGN normally doesn't implement load-balancing functions; they may be implemented in other dedicated equipment. Therefore, the gateways have to resort to DNS64 or an internal host's behavior. Once DNS64 is deployed, the load balancing can be performed by synthesizing the AAAA response with different IPv6 prefixes. For the applications not requiring a 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.

- A dedicated load balancer could be deployed at the front of a NAT64-FE farm. The load balancer could use 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, a source-address-based policy, is preferred. A dynamic algorithm, for example, Round-Robin, may have impacts on applications seeking session continuity, which are described in Table 1.

5. Source-Address Transparency

5.1. Traceability

Traceability is required in many cases, such as meeting accounting requirements and identifying the sources of malicious attacks. 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 was collected from a stateful NAT64 gateway for 60 days. Syslog [RFC5424] has been adopted to transmit log messages from NAT64 to a log station. Each log message contains the 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 72,000 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 systems, 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 be configured with the 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 ports and hence do not require logging. The log volume in this case may be only one thousandth of that of dynamic port allocation. Some implementations may adopt static port-range allocations [DET-CGN] that eliminate the need for per-subscriber logging. As a side effect of those methods, the IPv4 multiplexing efficiency is decreased. For example, the utilization ratio of public IPv4 addresses drops to approximately 75% when the NAT64 gateway is configured with bulk port allocation. (The lab testing allocates each subscriber with 400 ports.) In addition, port-range-based allocation should consider port randomization as described in [RFC6056]. The trade-off among address multiplexing efficiency, logging storage compression, and port allocation complexity should be considered. More discussions can be found in [PORT-ALLOC]. The decision can balance usable IPv4 resources against investments in log systems.

5.2. Geolocation

IP addresses are usually used as inputs to geolocation 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 suboptimal substitution to solve the problem of host identification; in particular, it may not solve today’s problems with source identification through translation. The following lists current practices to mitigate the issue.

- Operators who adopt NAT64-FE may leverage the application-layer proxies, e.g., X-Forwarded-For (XFF) [RFC7239], 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 that has been integrated in most load balancers. Therefore, it may be superior to use in a NAT-FE environment. On the downside, XFF is specific to HTTP. It

restricts usage so that the solution can't be applied to requests made over HTTPS. This makes geolocation problematic for HTTPS-based services.

- The NAT64-CGN equipment may not implement XFF. Geolocation based on shared IPv4 addresses is rather inaccurate in that case. Operators could subdivide the outside IPv4 address pool so an IPv6 address can be translated depending on the IPv6 subscriber's geographical locations. As a 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 its own specific usage. For the geolocation systems relying on a RADIUS database [RFC5580], we have investigated delivering NAT64 BIBs and Session Table Entries (STEs) to a RADIUS server [NAT64-RADIUS]. This method could provide a geolocation system with an internal IPv6 address to identify each user. It can be paired with [RFC5580] to convey the original source address through the 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 reachability between two IP address families, NAT64-CGN has to implement appropriate application-aware functions, i.e., Application Layer Gateways (ALGs), where address translation is not sufficient and security mechanisms do not render the functions infeasible. Most NAT64-CGNs mainly provide FTP-ALG [RFC6384]. NAT64-FEs may have functional richness on the 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 a control session, for example, using the "Via" field in a HTTP header, "Transport" field in an RTSP SETUP message, or "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 support. The experiences of some applications are still aligned with [RFC6586]. For example, we 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 with the support of 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 the LTE mobile environment as specified in [IR.92]. The voice call 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) as described in [RFC5245] is recommended to be supported for the SIP IPv6 transition. [RFC6157] describes both signaling and the media-layer process, which should be followed. In addition, it is worth noting that ICE is not only useful for NAT traversal, but also for firewall [RFC6092] traversal in a native IPv6 deployment.

Different IPsec modes for VPN services have been tested, including IPsec Authentication Header (AH) and IPsec Encapsulating Security Payload (ESP). It has been shown that IPsec AH fails because the destination host detects the IP header changes and invalidates 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 would succeed if the IPsec client supports NAT traversal in the Internet Key Exchange Protocol (IKE) [RFC3947] and uses IPsec ESP over UDP [RFC3948].

Table 1: The Tested Applications

Application	Results and Issues Found
Web service	Mostly pass; some failures 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" must not be less than 2 hours 4 minutes [RFC5382] for TCP sessions and 5 minutes for UDP sessions [RFC4787]. In some cases, NAT resources may be significantly consumed by largely inactive users. The NAT and other customers would suffer from service degradation due to port consumption by other subscribers using the same NAT64 device. A flexible NAT session control is desirable to resolve these issues. The Port Control Protocol (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. Doing so might improve user experiences, for example, by 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 sent by mobile terminal devices.

Subscribers can also benefit from network reliability. It has been discussed that Hot Standby offers a satisfactory experience after outage of the primary NAT64 has occurred. Operators may rightly be concerned about the considerable investment required for NAT64 equipment relative to low ARPU. For example, transport links may be expensive, because the primary NAT64 and the backup are normally located at different locations, separated by a relatively large distance. Additional cost would be incurred to ensure the connectivity quality. However, that may be necessary to applications that are delay-sensitive and seek session continuity, for example, online games and live streaming. Operators may be able to get added value from those services by offering first-class services. The service sessions can be pre-configured on the gateway to Hot Standby mode depending on the subscriber’s profile. The rest of the sessions can be covered by Cold or Warm Standby.

7. MTU Considerations

IPv6 requires that every link in the Internet have a Maximum Transmission Unit (MTU) of 1280 octets or greater [RFC2460]. However, if NAT64 translation is deployed, some IPv4 MTU constrained link will be used in a communication path and the 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 a fragmentation header in which the "M" flag is set to 0 and the "Fragment Offset" is set 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 a risk that all the fragments will have to be dropped.

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

Another approach that potentially avoids this issue is to configure the IPv4 MTU to more than 1260 bytes. This would prevent getting a PTB message for an MTU smaller than 1280 bytes. Such an operational consideration is hard to universally apply to the legacy "IPv4 Internet" that is bridged by NAT64-CGNs. However, it’s a feasible approach in NAT64-FE cases, since an IPv4 network NAT64-FE is rather well-organized and operated by an IDC operator or content provider. Therefore, the MTU of an IPv4 network in NAT64-FE case is strongly recommended to be 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 go to the IPv4 Internet; therefore, ULAs are only valid in the IPv6 site. The use of ULAs could help in identifying the translation traffic. [ULA-USAGE] provides further guidance on using ULAs.

We configure ULAs as NAT64 prefixes on a NAT64-CGN. If a host is assigned with only an IPv6 address and connected to a NAT64-CGN, when it connects to an IPv4 service, it would receive a 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 addresses, it would initiate both A and AAAA record lookup, then both the original A record and DNS64-generated AAAA record would be received. A host that is compliant with [RFC6724] will never prefer a ULA over an IPv4 address. An IPv4 path will always be selected. It may be undesirable because the NAT64-CGN will never be used. Operators may consider adding additional site-specific rows into the default policy table for host address selection in order to steer traffic flows through the NAT64-CGN. However, it involves significant costs to change a terminal’s behavior. Therefore, it is not suggested that operators configure ULAs on a NAT64-CGN.

ULAs can’t work when hosts transit the Internet to connect with NAT64. Therefore, ULAs are not applicable 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 could also adopt unicast Reverse Path Forwarding (uRPF) [RFC3704] and blacklisting and whitelisting to enhance security by specifying access policies. For example, NAT64-CGN should forbid establishing NAT64 BIB for incoming IPv6 packets if they do not pass the uRPF check in Strict or Loose mode or if their source IPv6 address is blacklisted.

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. The load balancer is recommended to enable the capabilities for line-rate DDOS defense, such as the employment of SYN proxy/cookie. In this case, division of the security domain is necessary as well. Therefore, load balancers could not only optimize the 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 the DNS response is modified and DNSSEC intends to prevent such changes. More detailed discussions can be found in [RFC6147].

10. 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 their efforts and contributions on editing, which substantially improved the readability of the document.

Thanks to Cameron Byrne who was an active coauthor of some earlier draft versions of this document.

11. Contributors

The following individuals contributed extensively 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

12. References

12.1. Normative References

[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.

[RFC7239] Petersson, A. and M. Nilsson, "Forwarded HTTP Extension", RFC 7239, June 2014.

12.2. Informative References

[Alexa] Alexa, "Top 500 Global Sites", April 2013, <http://www.alexa.com/topsites>.

[COEXIST] Kaliwoda, A. and D. Binet, "Co-existence of both dual-stack and IPv6-only hosts", Work in Progress, October 2012.

[CONN-STATUS] Patil, P., Boucadair, M., Wing, D., and T. Reddy, "IP Connectivity Status Notifications for DHCPv6", Work in Progress, February 2014.

[Cisco-VNI] Cisco, "Cisco VNI Global Mobile Data Traffic Forecast, 2012-2018", February 2014, <http://ciscovni.com/forecast-widget/index.html>.

[DET-CGN] Donley, C., Grundemann, C., Sarawat, V., Sundaresan, K., and O. Vautrin, "Deterministic Address Mapping to Reduce Logging in Carrier Grade NAT Deployments", Work in Progress, January 2014.

[IR.92] Global System for Mobile Communications Association (GSMA), "IMS Profile for Voice and SMS Version 7.0", March 2013.

[MAP-DEPLOY] Qiong, Q., Chen, M., Chen, G., Tsou, T., and S. Perreault, "Mapping of Address and Port (MAP) - Deployment Considerations", Work in Progress, April 2014.

[MAP-T] Li, X., Bao, C., Dec, W., Troan, O., Matsushima, S., and T. Murakami, "Mapping of Address and Port using Translation (MAP-T)", Work in Progress, February 2014.

[MOTIVATION] Boucadair, M., Matsushima, S., Lee, Y., Bonness, O., Borges, I., and G. Chen, "Motivations for Carrier-side Stateless IPv4 over IPv6 Migration Solutions", Work in Progress, November 2012.

[NAT64-RADIUS] Chen, G. and D. Binet, "Radius Attributes for Stateful NAT64", Work in Progress, July 2013.

[PORT-ALLOC] Chen, G., Tsou, T., Donley, C., and T. Taylor, "Analysis of NAT64 Port Allocation Methods for Shared IPv4 Addresses", Work in Progress, April 2014.

[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.

[SIIT] Anderson, T., "Stateless IP/ICMP Translation in IPv6 Data Centre Environments", Work in Progress, November 2012.

[ULA-USAGE] Liu, B. and S. Jiang, "Recommendations of Using Unique Local Addresses", Work in Progress, February 2014.

Appendix A. Test Results for Application Behavior

We tested several application behaviors in a lab environment to evaluate the impact when a primary NAT64 is out of service. In this testing, participants were asked to connect an IPv6-only WiFi network using laptops, tablets, or mobile phones. NAT64 was deployed as the gateway to provide Internet service. The tested applications are shown in the table below. Cold Standby, Warm Standby, and Hot Standby were each tested. The participants may have experienced service interruption due to the NAT64 handover. Different interruption intervals were tested to gauge application behaviors. The results are shown below.

Table 2: The Acceptable Delay of Applications

Application	Acceptable Interrupt Recovery	Session Continuity
Web browsing	Maximum of 6 s	No
HTTP streaming	Maximum of 10 s (cache)	Yes
Games	200-400 ms	Yes
P2P file sharing and streaming	10-16 s	Yes
Instant Message	1 minute	Yes
Email	30 s	No
Downloading	1 minute	No

Authors' Addresses

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

EMail: chengang@chinamobile.com, phdgang@gmail.com

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

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

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

EMail: xiechf@ctbri.com.cn

David Binet
France Telecom-Orange
Rennes
35000
France

EMail: david.binet@orange.com