v6ops
Internet-Draft
Intended status: Informational
Expires: August 4, 2013

G. Chen
Z. Cao
China Mobile
C. Byrne
T-Mobile USA
C. Xie
China Telecom
D. Binet
France Telecom

January 31, 2013

NAT64 Deployment Considerations
draft-ietf-v6ops-nat64-experience-01

Abstract

This document summarizes NAT64 deployment scenarios and operational experience with stateful NAT64-CGN(NAT64 Carrier Grade NATs) and NAT64-FE (NAT64 server Front End).

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 August 4, 2013.

Table of Contents

1. Introduction 4
2. Terminology 5
3. NAT64-CGN Deployment Experiences 5
3.1. NAT64-CGN Networking 5
3.2. High Availability Considerations 6
3.3. Traceability 6
3.4. Quality of Experience 7
3.5. NAT64-CGN Load Balancer 8
3.6. MTU Consideration 8
4. NAT64-FE Deployment Experiences 9
4.1. NAT64-FE Networking 9
4.2. Source Address Traceability 10
4.3. DNS Resolving 10
4.4. Load Balancer 11
4.5. MTU Consideration 11
4.6. Anti-DDoS/SYN Flood 11
5. Security Considerations 11
6. IANA Considerations 12
7. Acknowledgements 12
8. Additional Author List 12
9. References 13
9.1. Normative References 13
9.2. Informative References 13
Authors' Addresses 15

1. Introduction

With IANA's global IPv4 address pool was exhausted, IPv6 is the only sustainable solution for numbering nodes on the Internet. Network operators have to deploy IPv6-only networks in order to meet the numbering needs of the expanding internet without available IPv4 addresses.

As IPv6 deployment continues, IPv6 networks and hosts will need to coexist with IPv4 numbered resources. The Internet will include nodes that are dual-stack, nodes that remain IPv4-only, and nodes that can be deployed as IPv6-only nodes.

Single stack IPv6 network deployment can simplify the network provisioning. Some justifications have been described in [I-D.ietf-v6ops-464xlat]. IPv6-only networks confer some benefits to mobile operators employing them. In the mobile context, it enables the use of a single IPv6 PDP(Packet Data Protocol), which eliminates significant network cost caused by doubling the PDP count on a mass of legacy mobile terminals. In broadband networks overall, it can allow for the scaling of edge-network growth decoupled from IPv4 numbering limitations.

In a transition scenario, an existing network may rely on the IPv4 stack for a long time. There is also the troublesome trend of access network providers squatting on IPv4 address space that they do not own. Allowing for interconnection between IPv4-only nodes and IPv6-only nodes is a critical capability. 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. A translation mechanism based on a NAT64[RFC6146] function will be a key element of the internet infrastructure supporting such legacy networks.

[RFC6036] reported at least 30% operators plan to run some kind of translator (presumably NAT64/DNS64). Advice on NAT64 deployment and operation is therefore of some importance. [RFC6586] documented the implications for ipv6 only networks. This document intends to be specific to NAT64 network planning.

In regards to IPv4/IPv6 translation, [RFC6144] has described a framework of enabling networks to make interworking possible between IPv4-only and IPv6-only networks. This document has further categorized different NAT64 location and use case. The principle distinction of location is if the NAT64 is located in a NAT64-CGN (Carrier Grade NATs) or NAT64-FE (server Front End).

2. Terminology

The terms of NAT-CGN/FE are understood to be a topological distinction indicating different features employed in a NAT64 deployment.

NAT64-CGN: A NAT64-CGN (Carrier Grade NATs) is placed in an ISP network. IPv6 only subscribers leverage the NAT64-CGN to access existing IPv4 internet services. The ISP as an administrative entity takes full control on the IPv6 side, but has limited or no control on the IPv4 side. ISP's should attempt to accommodate the behavior of IPv4 networks and services.

NAT64-FE: A NAT64-FE (server Front End) is generally a traffic load balancer with NAT64 functionalities in a ICP network.

3. NAT64-CGN Deployment Experiences

A NAT64-CGN deployment scenario is depicted in Figure 1

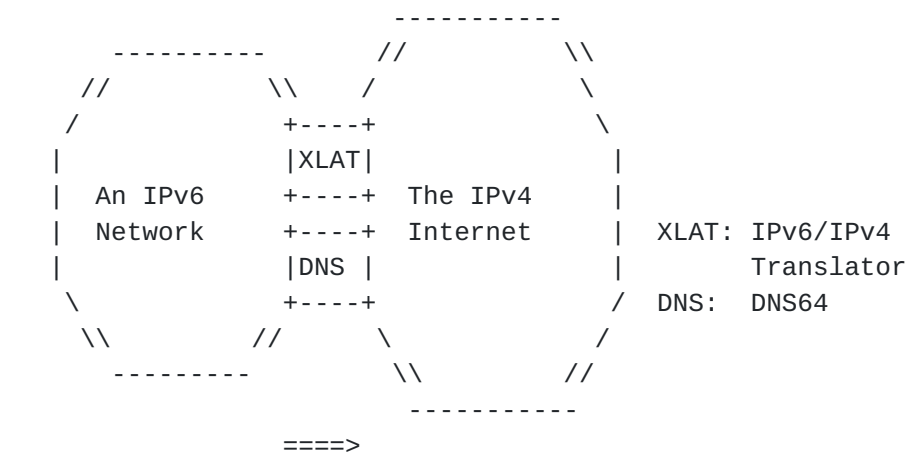

Figure 1: NAT64-CGN Scenario: IPv6 Network to IPv4 Internet

3.1. NAT64-CGN Networking

The NAT64-CGN use case is employed to connect IPv6-only users to the IPv4 Internet. The NAT64 gateway performs protocol translation from an IPv6 packet header to an IPv4 packet header and vice versa according to the stateful NAT64 [RFC6146]. Address translation maps IPv6 addresses to IPv4 addresses and vice versa for return traffic.

All connections to the IPv4 Internet from IPv6-only clients must traverse the NAT64-CGN. It is 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 translates only at or near the network egress. In many service provider networks, NAT64 is considered feature of the AS boarder. This allows consistent attribution and traceability within the service provider network. Meaning, within one network, the packet only has one source. As the packet leaves the network destine for another network, the packet source may be translated as needed.

In mobile networks, various possibilities can be envisaged to deploy the NAT64 function. Whichever option is selected, the NAT64 function will be deployed beyond the GGSN (Gateway GPRS Support Node) or PDN-GW (Public Data Network-Gateway), i.e. first IP node in currently deployed mobile architectures.

In a given implementation, NAT64 functionality can be provided by either a dedicated device or an multifunction gateway with integrated NAT64 functionality. If NAT64 is integrated into an existing node, capacities of existing nods can be potentially limited by the new functionality, e.g. maximum concurrent connections. In a mobile context, the NAT64 function likely be implemented in a firewall, which is the first hop routed from GGSN/PGW.

3.2. High Availability Considerations

High Availability (HA) is a major requirement for every service.

Two mechanisms are typically used to achieve high availability, i.e. cold-standby and hot-standby. Cold-standby systems have synchronized configuration and mechanism to failover traffic between the hot and cold systems such as VRRP [RFC5798] . Unlike hot-standby, cold-standby does not synchronize NAT64 session state. This makes cold-standby less resource intensive and generally simpler, but it requires clients to re-establish sessions when a fail-over occurs. Hot-standby has all the features of cold-standby but must also synchronize the binding information base (BIB). Given that short lived sessions account for most of the bindings, hot-standby does not offer much benefit for those sessions. Consideration should be given to the importance (or lack thereof) of maintaining bindings for long lived sessions across failovers.

3.3. Traceability

Traceablility is required in many cases such as identifying malicious attacks sources and accounting requirements. NAT64 devices are required to log events like creation and deletion of translations and information about the occupied resources. There are two different demands for traceability,i.e. online or offline.

- Regarding the Online requirements, XFF (X-Forwarded-For) [I-D.ietf-appsawg-http-forwarded]would be a candidate, it appends IPv6 address of subscribers to HTTP headers which is passed on to WEB servers, and the querier server can lookup radius servers for the target subscribers based on IPv6 addresses included in XFF HTTP headers. X-Forwarded-For is specific to HTTP, requires the use of an application aware gateway, cannot in general be applied to requests made over HTTPs and cannot be assumed to be preserved end-to-end as it may be overwritten by other application-aware proxies such as load balancers.

- Some potential solutions to online traceability are explore in [I-D.ietf-intarea-nat-reveal-analysis].

- A NAT64-CGN could also deliver NAT64 sessions (BIB and STE) to a Radius server by extension of the radius protocol. Such an extension is an alternative solution for online traceability, particularly high performance would be required on Radius servers in order to achieve this.

- For off-line traceability, syslog might be a good choice. [RFC6269] indicates address sharing solutions generally need to record and store information for specific periods of time. A stateful NAT64 is supposed to manage one mapping per session. A large volume of logs poses a challenge for storage and processing. In order to mitigate the issue, [I-D.donley-behave-deterministic-cgn]is proposed to pre-allocated a group of ports for each specific IPv6 host. A trade-off among address multiplexing efficiency, port randomization security[RFC6056] and logging storage compression should be considered during the planning. A hybrid mode combining deterministic and dynamic port assignment was recommended regarding the uncertainty of user traffic.

3.4. Quality of Experience

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 ALGs where address translation is not itself sufficient and security mechanisms do not render it infeasible. e.g. FTP-ALG[RFC6384], RSTP-ALG, H.323-ALG,etc. 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.

Session status normally is managed by a static life-cycle. 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[I-D.ietf-pcp-base] 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 an ability should also be considered to upgrade user experiences, e.g. assigning different sizes of port ranges for different subscribers. Such a mechanism is also helpful to minimize terminal battery consumption and reduce the number of keepalive messages to be sent by terminal devices.

3.5. NAT64-CGN Load Balancer

Load balancers are an essential tool to avoid the issue of single points of failure and add additional scale. It is potentially important to employ load-balancing considering that deployment of multiple NAT64 devices. Load balancers are required to achieve some service continuity and scale for customers. [I-D.zhang-behave-nat64-load-balancing] discusses several ways of achieving NAT64 load balancing, including anycast based policy and prefix64 selection based policy, either implemented via DNS64[RFC6147] or Prefix64 assignments. Since DNS64 is normally co-located with NAT64 in some scenarios, it could be leveraged to perform the load balance. For traffic which does not require a DNS resolution, prefix64 assignment based on[I-D.ietf-behave-nat64-learn-analysis] could be adopted.

3.6. MTU Consideration

IPv6 requires that every link in the internet have an 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 message, reporting a Next-Hop MTU less than 1280. The result would be that IPv6 allows packets to contain a fragmentation header, without the packet being fragmented into multiple pieces. [I-D.ietf-6man-ipv6-atomic-fragments] 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 protocol level, which might imply potential upgrade/modifications on behaviors to deployed nodes. Another approach that potentially avoids this issue is to configure IPv4 MTU>=1260. It would forbid the occurrence of PTB<1280. However, such an operational consideration is hard to

universally apply to the legacy "IPv4 Internet".

4. NAT64-FE Deployment Experiences

The NAT64-FE Scenario is depicted in Figure 2

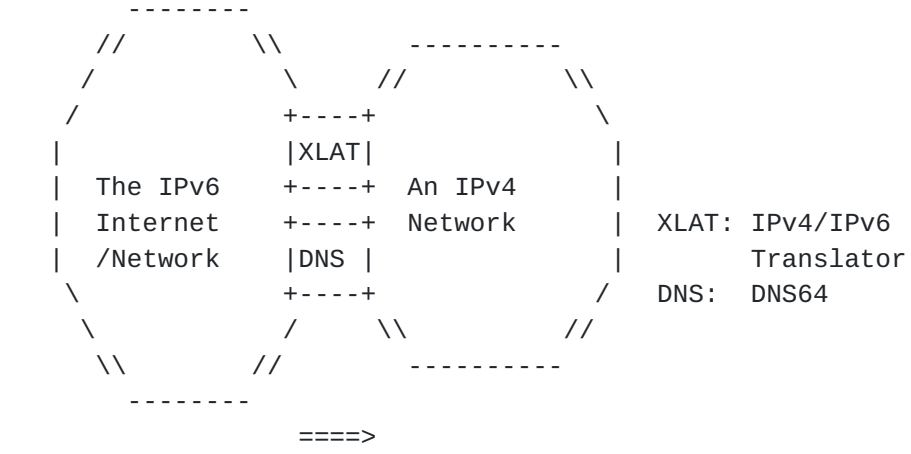

Figure 2: NAT64-FE Scenario: IPv6 Internet/Network to IPv4 Network

4.1. NAT64-FE Networking

There are plenty of practices to equip load balancer with NAT64 at front of servers. Two different cases appeared in the NAT64-FE networking.

- Some content providers who has a wealth of IPv6 experiences consider IPv6-only strategy to serve customers since it allows new services delivery without having to integrate consideration of IPv4 NAT and address limitations of IPv4 networks. Whereas they have to provide some IPv4 service continuity to their customers, stateless IP/ICMP Translation (SIIT) [RFC6145]has been used to continue provide services for IPv4 subscribers.

- ICPs who have insufficient IPv6 incentive likely prefer short-term alternatives to provide IPv6 connectivity due to the widespread impact of supporting IPv6 within a ICP environment. A stateful NAT64 has been located at front of servers. It could simultaneously facilitate the IPv6 accessibility and conservation of IPv4 servers. [I-D.ietf-v6ops-icp-guidance]has described the cases, in which an HTTP proxy can readily be configured to handle incoming connections over IPv6 and to proxy them to a server over IPv4.

For first case, [I-D.anderson-siit-dc]has provided further descriptions and guidelines. This document is addressed to second

case. An administrator of the IPv4 network needs to be cautious and aware of the operational issues in the case since the native IPv6 is always more desirable than transition solutions.

One potential challenge is stateful NAT64-FE facing IPv6 Internet, in which a significant number of IPv6 users may initiate connections. When increasingly numerous users in IPv6 Internet access an IPv4 network, scalability concerns(e.g. additional latency, a single point of failure, IPv4 pool exhaustion, etc) are apt to be applied. For a given off-the-shelf NAT64-FE, those challenges should be seriously assessed. Potential issues should be properly identified.

Following subsections described several considerations to stateful NAT64-FE case. For operators who seek a clear precedent for operating reliable IPv6-only services, it should be well noted that the usage is problematic.

4.2. Source Address Traceability

IP addresses are usually used as input to geo-location services. The use of address sharing will prevent 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 at section 3.3 intended to bridge the gap. However, the analysis reveals those solutions can't be a optimal substitution to solve the problem of host identification, in particular it does not today mitigate problems with source identification through translation. That makes NAT64-FE usage becoming a unappealing approach, if customers require source address tracking.

For the operators, who already deployed NAT64-FE approach, the source address of the request is obscured without the source address mapping information previously obtained. It's superior to present mapping information directly to applications. Some application layer proxies e.g. XFF (X-Forwarded-For) , can convey this information in-band. Another approach is to ask application coordinating the information with NAT logging. But that is not sufficient, since the applications itself wants to know the original source address from an application message bus. The logging information may be used by administrators offline to inspect use behavior and preference analysis, and accurate advertisement delivery.

4.3. DNS Resolving

In the case of NAT64-FE, it is recommended to follow the recommendations in [RFC6144]. There is no need for the DNS to synthesize AAAA from A records, since static AAAA records can be

registered in the authoritative DNS for a given domain to represent these IPv4-only hosts. How to design the FQDN for the IPv6 service is out-of-scope of this document.

4.4. Load Balancer

Load balancing on NAT64-FE has a couple of considerations. If dictated by scale or availability requirements traffic should be balanced among multiple NAT64-CE devices. One point to be noted is that synthetic AAAA records may be added directly in authoritative DNS. load balancing based on DNS64 synthetic resource records may not work in those cases. Secondly, NAT64-FE could also serve as the load balancer for IPv4 backend servers. There are also some ways of load balance for the cases, where load balancer is placed in front of NAT64-FE.

4.5. MTU Consideration

As compared to the MTU consideration in NAT64-CGN, the MTU of IPv4 network are strongly recommended to set to more than 1260. Since a IPv4 network is normally operated by a particular administrative entity, it should take steps to prevent the risk of fragmentation discussed in [I-D.ietf-6man-ipv6-atomic-fragments].

4.6. Anti-DDoS/SYN Flood

For every incoming new connection from the IPv6 Internet, 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. With service provisioning, attacks have the potential could also deteriorate service quality. One consideration in internet content providers is place a L3 load balancer with capable of line rate DDOS defense, such as the employment of SYN PROXY-COOKIE. Security domain division is necessary in this case. Load Balancers could not only serve for optimization of traffic distribution, but also serve as a DOS mitigation device.

5. Security Considerations

This document presents the deployment experiences of NAT64 in CGN and FE scenario, some security considerations are described in detail regarding to specific NAT64 mode in section 3 and 4. In general, RFC 6146[RFC6146] provides TCP-tracking, address-dependent filtering mechanisms to protect NAT64 from DDOS. In NAT64-CGN cases, ISP also

could adopt uRPF 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 (Strict or Loose mode) check does not pass or whose source IPv6 address is associated to black-lists.

6. IANA Considerations

This memo includes no request to IANA.

7. Acknowledgements

The authors would like to thank Jari Arkko, Dan Wing, Remi Despres, Fred Baker, Hui Deng, Lee Howard, Iljitsch van Beijnum and Philip Matthews for their helpful comments. Many thanks to Wesley George and Satoru Matsushima for their reviews.

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

8. 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

9. References

9.1. Normative References

[I-D.ietf-pcp-base]
Wing, D., Cheshire, S., Boucadair, M., Penno, R., and P. Selkirk, "Port Control Protocol (PCP)", draft-ietf-pcp-base-29 (work in progress), November 2012.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC5798] Nadas, S., "Virtual Router Redundancy Protocol (VRRP) Version 3 for IPv4 and IPv6", RFC 5798, March 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.

[RFC6384] van Beijnum, I., "An FTP Application Layer Gateway (ALG) for IPv6-to-IPv4 Translation", RFC 6384, October 2011.

9.2. Informative References

[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.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-05 (work in progress), January 2013.

[I-D.ietf-6man-ipv6-atomic-fragments]
Gont, F., "Processing of IPv6 "atomic" fragments", draft-ietf-6man-ipv6-atomic-fragments-03 (work in progress), December 2012.

[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.

[I-D.ietf-behave-nat64-learn-analysis]
Korhonen, J. and T. Savolainen, "Analysis of solution proposals for hosts to learn NAT64 prefix", draft-ietf-behave-nat64-learn-analysis-03 (work in progress), March 2012.

[I-D.ietf-intarea-nat-reveal-analysis]
Boucadair, M., Touch, J., Levis, P., and R. Penno, "Analysis of Solution Candidates to Reveal a Host Identifier (HOST_ID) in Shared Address Deployments", draft-ietf-intarea-nat-reveal-analysis-04 (work in progress), August 2012.

[I-D.ietf-v6ops-464xlat]
Mawatari, M., Kawashima, M., and C. Byrne, "464XLAT: Combination of Stateful and Stateless Translation", draft-ietf-v6ops-464xlat-09 (work in progress), January 2013.

[I-D.ietf-v6ops-icp-guidance]
Carpenter, B. and S. Jiang, "IPv6 Guidance for Internet Content and Application Service Providers", draft-ietf-v6ops-icp-guidance-05 (work in progress), January 2013.

[I-D.zhang-behave-nat64-load-balancing]
Zhang, D., Xu, X., and M. Boucadair, "Considerations on NAT64 Load-Balancing", draft-zhang-behave-nat64-load-balancing-03 (work in progress), July 2011.

[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.

[RFC6144] Baker, F., Li, X., Bao, C., and K. Yin, "Framework for IPv4/IPv6 Translation", RFC 6144, April 2011.

[RFC6269] Ford, M., Boucadair, M., Durand, A., Levis, P., and P. Roberts, "Issues with IP Address Sharing", RFC 6269, June 2011.

[RFC6586] Arkko, J. and A. Keranen, "Experiences from an IPv6-Only Network", RFC 6586, April 2012.

Authors' Addresses

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

Email: phdgang@gmail.com

Zhen Cao
China Mobile
53A,Xibianmennei Ave.,
Xuanwu District,
Beijing 100053
China

Email: caozhen@chinamobile.com

Cameron Byrne
T-Mobile USA
Bellevue
Washington 98105
USA

Email: cameron.byrne@t-mobile.com

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

Email: xiechf@ctbri.com.cn

David Binet
France Telecom
Rennes
35000
France

Email: david.binet@orange.com