Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 1 September 2025

C. Xie
J. Sun
China Telecom
X. Li
CERNET Center/Tsinghua University
G. Han
Indirection Network Inc.

28 February 2025

EVPN Route Types and Procedures for EVN6
draft-xie-bess-evpn-extension-evn6-02

Abstract

EVN6 is a mechanism designed to carry Ethernet virtual networks, providing Ethernet connectivity to customer sites dispersed on public IPv6 networks. At the data layer, EVN6 directly places the Ethernet frames in the payload of IPv6 packet, and dynamically generates the IPv6 addresses of the IPv6 header using host MAC addresses and other information, then sends them into IPv6 network for transmission. This document proposes extensions to EVPN for EVN6, including two new route types and related procedures.

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 https://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 1 September 2025.

Copyright Notice

Table of Contents

1. Introduction

EVN6 [I-D.xls-intarea-evn6]is a mechanism designed to carry Ethernet virtual networks, providing Ethernet connectivity between customer sites dispersed on public IPv6 networks. When receiving Ethernet frame to be transmitted by host of local site, ingress PE of IPv6 network will map the host MAC addresses, virtual network identity (VEI), and IPv6 mapping prefixes to their appropriate positions of the source and destination IPv6 address, so the address is unique for each host. After the outer source and destination IPv6 addresses are ready, the data frame is directly encapsulated in IPv6 packet and transmitted to the correct destination site through the IPv6 network. Since the outer address contains the MAC address of the host, tunnel in EVN6 is specifically for the communication between two hosts located at different sites, rather than the same tunnel being shared by multiple host pairs, furthermore, tunnel generated in this way is non-explicit since network managers does not need to pre configure tunnels between any customer sites.

Dynamic generation of EVN6 virtual network instances over IPv6 network needs the support of control layer. For the encapsulation of Ethernet frame, the data in MAC-VRF should be available, so the ingress PE device can find out the mapping prefix, VEI of the host identified by the MAC address from the local MAC-VRF. In small-scale networks, exchanging MAC/Ref6 mapping of each host between PE nodes can be achieved through manual configuration. But in large-scale networks, this procedure should be automatic, which requires support from the control layer. For dynamic creation of EVN6 network instances, this document introduces the extensions of EVPN [RFC7432], including two newly defined route types and the related processes. This is a companion document of EVN6[I-D.xls-intarea-evn6].

2. Specification of Requirements

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. Terminology

Broadcast Domain: In a bridged network, the broadcast domain corresponds to a Virtual LAN (VLAN), where a VLAN is typically represented by a single VLAN ID (VID) but can be represented by several VIDs where Shared VLAN Learning (SVL) is used per 802.1Q.

EVN6: Ethernet Virtual Network over IPv6.

MAC-VRF: A Virtual Routing and Forwarding table for Media Access Control (MAC) addresses on a PE.

VEI: Virtual Ethernet Identifier, it is 32-bits in length to identify each EVN6 instance.

4. Advantages of Using EVPN to Create EVN6

The EVN6 tunnel can be dynamically created using BGP EVPN as the Control Plane. EVPN is a universal Control Plane protocol that can be combined with various Data Plane technologies (including MPLS, SRv6, VxLAN, etc.) to achieve a complete forwarding and control separation SDN solution. So far EVPN has five route types based on MP-BGP NLRI extension, as follows:

Type 1. Ethernet Auto-Discovery Route

Type 2. MAC/IP Advertisement Route

Type 3. Inclusive Multicast Ethernet Tag Rout

Type 4. Ethernet Segment Route

Type 5. IP Prefix Route

As the control layer of EVN6, EVPN has the following advantages:

(1) EVPN can automatically establish EVN6 Tunnel, thereby reducing the complexity of network operation and improving network scalability.

(2) EVPN can automatically announce IP, MAC, VEI, and host routing information, effectively reducing BUM flooding traffic.

To support the operation of EVN6, the following new route types, i.e. Type 12 route and Type 13 route are defined,

(1) Type 12. EVN6 Auto-Discovery Route (tentative name)

(2) Type 13. MAC/IPv6 Advertisement Route (tentative name)

5. Route Types for EVN6

5.1. EVPN Type 12 Route

EVPN Type 12 Route is used to advertise VEI and IPv6 mapping prefix allocated by PE for virtual network between PEs, to establish a head end replication list, which is used for automatic discovery of PE and dynamic establishment of EVN6 tunnel. If the peer PE is reachable by IPv6 Mapping prefix routing, an EVN6 tunnel is established to the peer. Meanwhile, if the VEI of the other end is the same as that of the local end, a head end replication table is created for subsequent BUM message forwarding.

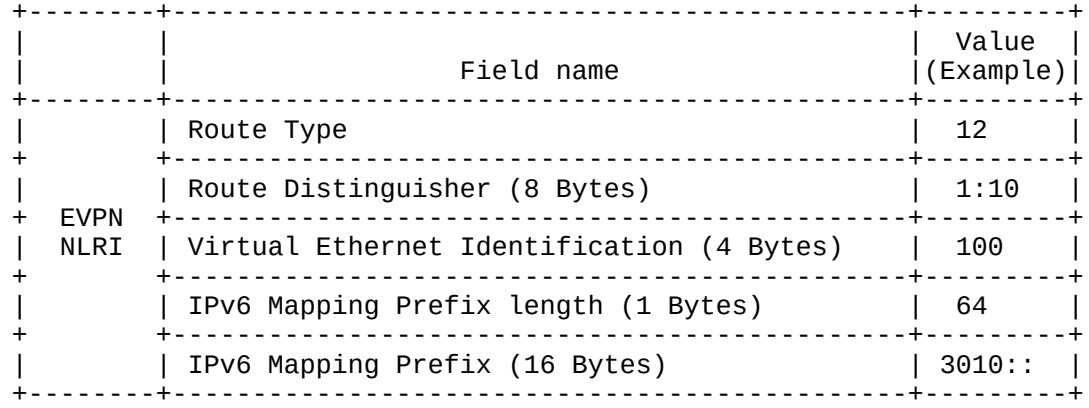

Figure 1: Type 12 Route NLRI Format Definition

The fields of Type 12 Route NLRI are illustrated as follows:

Route Distinguisher: This field is the RD (Route Distinguisher) value set under the EVPN instance.

Virtual Ethernet Identification: This field represents the layer-2 VEI of the EVN6 tunnel carried by this route, which is the identifier of the virtual Ethernet private network.

IPv6 Mapping prefix length: This field represents the length of the IPv6 address mapping prefix for the local Site carried by this route.

IPv6 Mapping prefix: This field represents the IPv6 address mapping prefix of the local Site carried by this route.

5.2. EVPN Type 13 Route

EVPN Type 13 route is used to advertise the MAC address and IPv6 address of the host between PE peers, and to dynamically establish control plane entries at both ends of the tunnel.

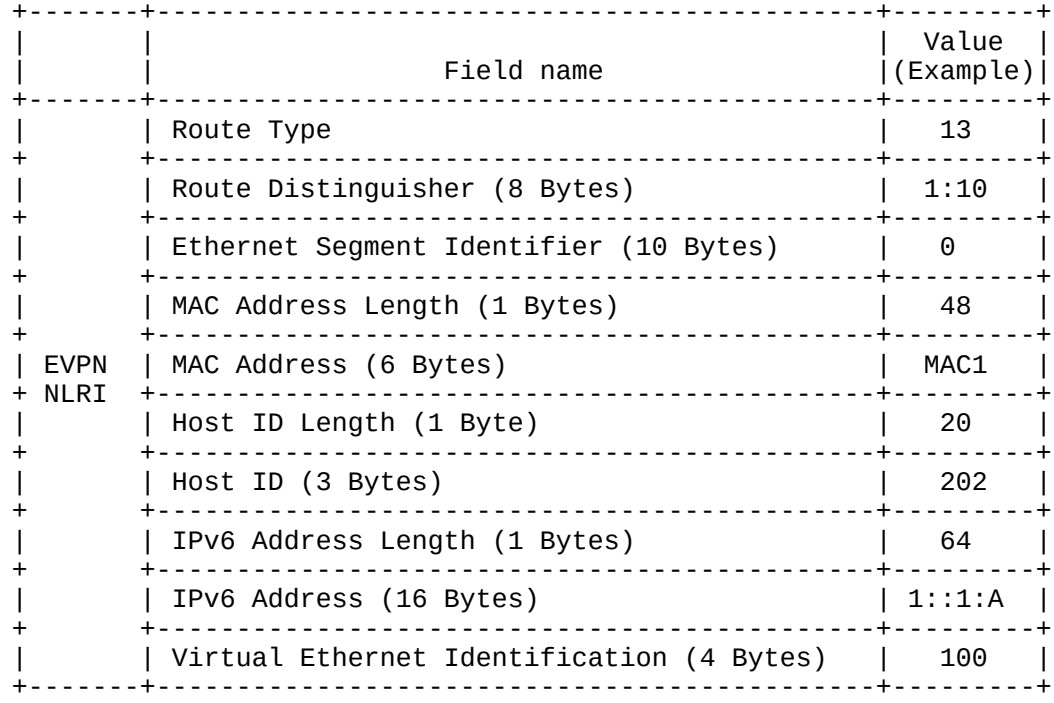

Figure 2: Type 13 Route NLRI Format Definition

The fields of Type 13 Route NLRI are illustrated as follows:

Route Distinguisher: This field is the RD (Route Distinguisher) value set under the EVPN instance.

Ethernet Segment Identifier (field to be determined): This field is the unique identifier defined for the connection between the PE and a certain CE. ESI manual configuration, two PE configurations with the same ESI in CE multi attribution scenarios. ESI of 0 indicates that CE to PE is a single return.

MAC Address Length: This field represents the length of the host MAC address carried by this route.

MAC Address: This field is the host MAC address carried by this route.

Host ID Length: This field represents the length of the host ID in bits carried by this route.

Host ID : This field is the host ID carried by this route.

IPv6 Address Length: This field represents the mask length of the host’s IPv6 address carried by this route.

IPv6 Address: This field is the host IPv6 address carried by this route.

Virtual Ethernet Identification: This field represents the layer-2 VEI of the EVN6 tunnel carried by this route, which is the identifier of the virtual Ethernet private network.

6. Procedures

To dynamically create EVN6 instance, EVPN utilizes extended Type 12 and Type 13 routes to convey control plane information. This approach requires configuring BGP EVPN neighbors on PE devices and associating them with the configuration information of EVN6 instances.

The following scenario is used to explain the procedure: H1 and H2 are located in different sites belonging to the same virtual network, with each site connected to PE1 and PE2 respectively. Communication between H1 and H2 is carried out through an EVN6 instance built over the IPv6 network. Creation of EVN6 instance using EVPN involves the following three processes.

6.1. EVN6 Tunnel Setup for BUM Traffic

Broadcast, unknown-unicast and multicast (BUM) traffic refers to that kind of network traffic that will be forwarded to multiple destinations or that cannot be addressed to the intended destination only. The MAC addresses of BUM traffic have the following characteristics:

Broadcast traffic - MAC address all FF;

Multicast traffic - the lowest bit of the first byte of the MAC address is 1;

Unknown Unicast - Unicast MAC address.

The EVN6 tunnel for BUM traffic is established by using Type 12 route.

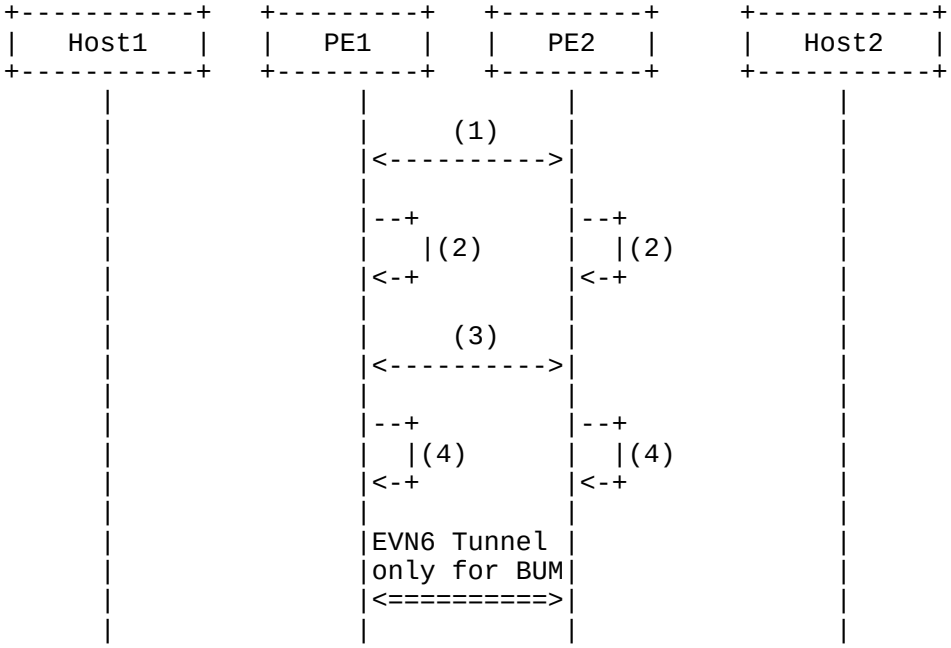

Figure 3: Process of establishing EVN6 tunnel for BUM traffic

The process is described as follows:

(1) Firstly, BGP EVPN peer is established between PE1 and PE2.

(2) L2 broadcast domain (i.e.BD)is created on PE1 and PE2 respectively, and configure associated layer 2 VEIs under the broadcast domain. An EVPN instance is created in the L2 broadcast domain and configure the RD(Route Distinguisher), Export Route Target (ERT), and import Route Target (IRT) of the local EVPN instance.

(3) After configuring the local EVN6, PE1 and PE2 will generate BGP EVPN routes and send them to the other end, which carry the Export Route Target of the local EVPN instance and the Type 12 route.

(4) After receiving the BGP EVPN route from the other end, PE1 and PE2 first check the Export Route Target of the EVPN instance carried by the route. If it is equal to the Import Route Target of the local EVPN instance, they receive the route. Otherwise, they discard the route. After receiving the route, PE1 and PE2 will obtain the IPv6 Mapping Address and Layer 2 VEI carried in it. If the IPv6 Mapping Address of the other end is reachable by

the Layer 3 route, an EVN6 tunnel will be established to the other end. Meanwhile, the local end will create a VEI based header replication table entry (refer to the static EVN6 creation table) for subsequent BUM packet forwarding.

BUM frames are forwarded using head end replication. When the BUM frame enters the EVN6 tunnel, the ingress end will encapsulate the frame in EVN6 according to the head end replication list and send the packet to all the egress ends in the head end replication list. This method ensures that the BUM frame can be correctly forwarded to the destination device through IPv6 network.

6.2. Host Entry Generation in MAC-VRF

The general EVN6 tunnel is designed for unicast Ethernet frame forwarding. The prerequisite for establishing a universal EVN6 tunnel is to generate Host MAC mapping entry in MAC-VRF table, which is achieved by dynamically exchanging Type 13 route.

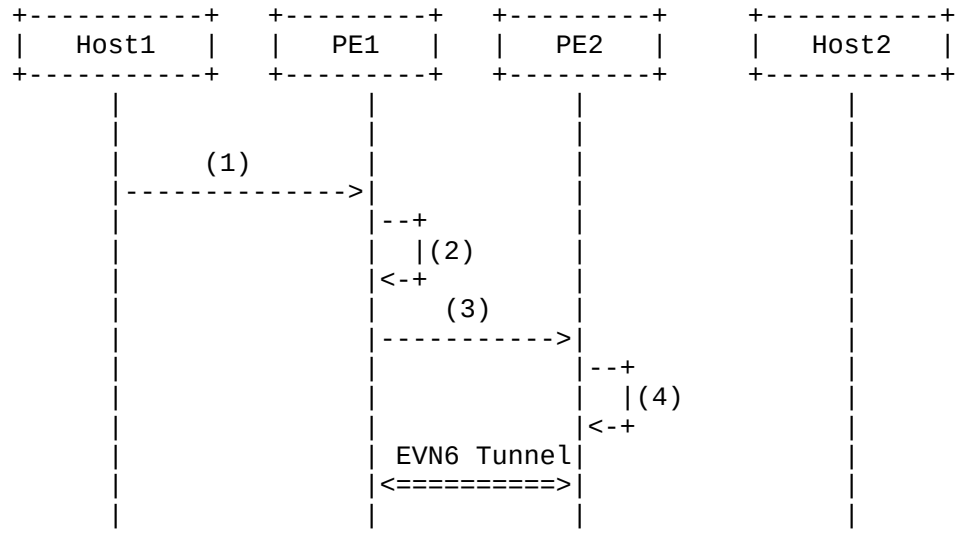

Figure 4: Process of PE2 learning the host MAC address and Host ID of Host1

The process of PE2 learning the host MAC address of Host1 is described as follows:

(1) Host1 initiates communicates with PE1 for the first time.

(2) PE1 learns the corresponding relationship between the MAC address, BD (L2 Broadcast Domain) Identifier, and packet input interface of Host1 through the IPv6 ND protocol, and generates a MAC entry for Host1 in the local MAC table.

(3) Meanwhile, PE1 generates a BGP EVPN route based on Host1's IPv6 ND information and sends it to peer PE2, which carries the outbound direction VPN Target of the local EVPN instance, the next hop attribute of the route, and the newly defined Type 13 route.

(4) After receiving the BGP EVPN route sent by PE1, PE2 first checks the outgoing direction VPN Target of the EVPN instance carried by the route. If it is equal to the incoming direction VPN Target of the local EVPN instance, it receives the route. Otherwise, it discards the route. After receiving the route, PE2 obtains the corresponding relationship between Host1's MAC address, Host ID, VEI, and IPv6 Mapping address on PE1, and generates a MAC entry for Host1 in the local MAC-VRF table.

The process of PE1 learning the host MAC of Host2 is the same as the above process, and will not be repeated here.

6.3. Ethernet Frame Transmission

This process is about hosts sending Ethernet frames to each other, EVN6 will be used when the frames pass though IPv6 network. When Host2 and Host1 are located at different sites, they first need to learn each other's MAC addresses, and then they can transmit unicast Ethernet frame to each other through the EVN6 virtual network.

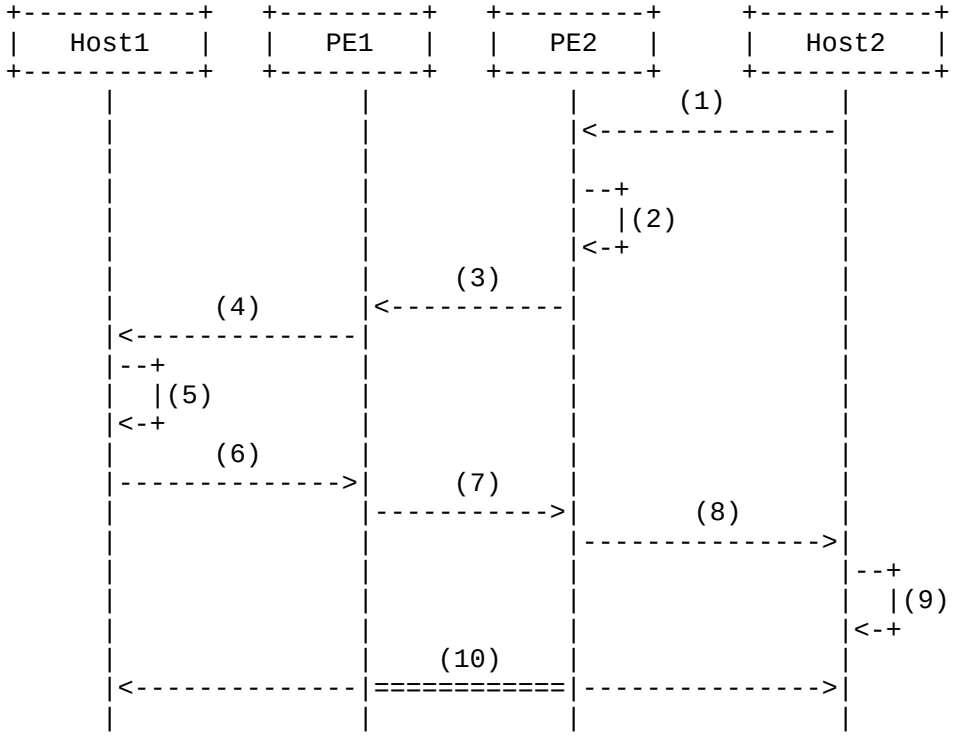

Figure 5: Process of Ethernet Frame Transmission

The process is described as follows:

(1) When Host2 communicates with Host1 for the first time, it needs to request the MAC address of Host1. It first sends an IPv6 NS request message with the destination IP being Host1-IP, this packet is encapuslation into a framework with its destination MAC being a Solicited-Node multicast address.

(2) By default, PE2 will broadcast on this network segment upon receiving the request. In order to reduce broadcast packets, the IPv6 ND suppression function can be enabled on PE2 at this time. In this way, when PE2 receives the IPv6 NS request packet, it first checks whether there is a local MAC address of Host1 based on the destination IPv6 address. If there is one match, it replaces the destination MAC with the MAC address of Host1, converts the multicast packet of the IPv6 NS request into a unicast packet.

(3) Then PE encapsulates it through the EVN6 tunnel and sends it to PE1.

(4) After receiving the packet sent from PE2, PE1 removes the packet header and forwards the frame to Host1.

(5) After receiving the IPv6 NS request contained in the frame, Host1 learns the MAC address of Host2 and responds with IPv6 NA in unicast form.

(6-7-8-9) After receiving the response message, Host2 learned the MAC address of Host1. At this point, Host1 and Host2 learn each other’s MAC addresses.

(10) Host1 and Host2 use unicast communication within EVN6 instance. The process of transmitting Ethernet frames is described in section 4.2 of [I-D.xls-intarea-evn6].

7. Security Considerations

TBD.

8. IANA Considerations

With this document IANA is requested to allocate codes for the "EVN6 Auto-Discovery Route" and "MAC/IPv6 Advertisement Route" in "EVPN Route Types" registry.

The two codes above use this document as the reference.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <https://www.rfc-editor.org/info/rfc4760>.

[RFC7153] Rosen, E. and Y. Rekhter, "IANA Registries for BGP Extended Communities", RFC 7153, DOI 10.17487/RFC7153, March 2014, <https://www.rfc-editor.org/info/rfc7153>.

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <https://www.rfc-editor.org/info/rfc7432>.

9.2. Informative References

[I-D.xls-intarea-evn6]
Xie, C., Sun, J., Li, X., Bao, C., and M. Smith, "EVN6: Mapping of Ethernet Virtual Network to IPv6 Underlay for Transmission", Work in Progress, Internet-Draft, draft-xls-intarea-evn6-03, 21 February 2025, <https://datatracker.ietf.org/doc/html/draft-xls-intarea-evn6-03>.

[IANA-EVPN-Route-Types]
IANA, "EVPN Route Types", <https://www.iana.org/assignments/evpn/evpn.xhtml>.

Authors' Addresses

Chongfeng Xie
China Telecom
Beiqijia Town, Changping District
Beijing
102209
China
Email: xiechf@chinatelecom.cn

Jibin Sun
China Telecom
Beiqijia Town, Changping District
Beijing
102209
China
Email: sunjb@chinatelecom.cn

Xing Li
CERNET Center/Tsinghua University
Shuangqing Road No.30, Haidian District
Beijing
100084
China
Email: xing@cernet.edu.cn

Guoliang Han
Indirection Network Inc.
Email: guoliang.han@indirectionnet.com