Hierarchical Routing Architecture (HRA)

Xiaohu Xu, Dayong Guo
Huawei Technologies Co.,Ltd.
(xuxh@huawei.com, guoseu@huawei.com)

Abstract

Some recent research activities from IETF and IRTF Routing Research Group (RRG) are to explore a new routing and addressing architecture to meet those challenges that current Internet are facing, especially in scalability. An identifier/locator split idea has been widely recognized as an architectural solution to the routing scalability issue. This paper describes a new routing and addressing architecture, called as Hierarchical Routing Architecture (HRA). HRA is also a kind of id/locator split solution. It introduces a hierarchical and cryptographic host identifier and adopts a hierarchical routing mechanism to support routing across multiple independent address spaces.

Keywords

Identifier, locator, internet architecture, name resolution, hierarchical routing architecture

1. Introduction

Some recent study has shown that the Internet routing table size is growing at a rate which almost exceeds the development speed of the hardware technology. This issue has drawn much attention from both industry and academe. After much discussion following the IAB Routing and Addressing workshop [1] in Amsterdam, a common conclusion is reached that the explosive growth in Internet routing table is mainly caused by widely adoption of multi-homing, traffic engineering and provider-independent address. Multi-homing is becoming a more and more popular phenomenon for cost, performance or redundancy reasons, and traffic engineering is usually deployed for load-balance purpose. These two factors result in prefix de-aggregation. Moreover, more and more enterprises prefer to adopt Provider-Independent (PI) address in order to maintain the freedom of switching between ISPs while avoiding renumbering, since renumbering the IP addresses of a network is usually a hard work. All the above factors led to the explosive growth in Default Free Zone (DFZ) routing table size.

However, the underlying reason for the routing scalability issue is the overlapping semantics of IP address which is used as both locator and identifier. In current Internet, IP address stands for the interface name and the location of a host, which is used by routers to deliver packets to their destinations. Moreover, the transport layer is coupled to the IP address, that is to say, IP address, together with a TCP port, are used as an identification of a TCP connection. The overload of IP address role makes it impossible to renumbering the addresses in a topologically aggregatable way in case of mobility, re-homing.

At present, the IRTF Routing Research Group (RRG) is chartered to explore a new routing and addressing architecture to meet those challenges in scalability, mobility, multi-homing, and inter-domain traffic engineering. An identifier/locator split idea has been widely recognized as an architectural solution to the routing scalability issue. With independent identifier, the locator can be renumbered easily to match changing network topology, which is usually Provider Aggregatable (PA) address, in case of mobility, re-homing or renumbering while avoiding interrupting the continuity of communication.

This paper describes a new routing and addressing architecture, called as Hierarchical Routing Architecture (HRA). HRA is a kind of id/locator split solution. It introduces a hierarchical and cryptographic host identifier

and adopts a hierarchical routing mechanism to support routing across multiple independent address spaces. Within HRA, the Internet routing scalability and stability are improved evidently with adoption of hierarchical routing mechanism. Besides, the scalability issue of flat host identifier in the Host Identity Protocol [5] can be solved with adoption of hierarchical host identifier.

2. Design Goals for a New Architecture

From the start, we design a new architecture according to the following goals:

1) Scalability

The explosive growth in routing table size is mainly caused by multi-homing, traffic-engineering and Provider Independent address, while the underlying reason is the overload of IP address semantics. We want to introduce id/loc split idea to deal with this routing scalability issue. Besides, id/loc split implies a need for mapping distribution system, so we want to design a high-efficient and scalable distributed mapping system with hierarchical Distributed Hash Table (DHT) technology.

2) Stability

In current Internet, the inter-domain route is a flat routing structure, the routing failure in one AS will be flooded in the whole Internet, which results in instability and slow convergence of Inter-Domain Routing (IDR). We want to introduce a hierarchical routing mechanism to improve the stability in the new architecture.

3) Build-in security

Today's Internet infrastructure lacks in embedded security mechanism. DDoS and other attacks are threatening the Internet security environment. This has encumbered the healthy development of e-commerce. So we want to bring in a cryptographic host identifier in the new architecture, which has build-in security feature.

4) Huge address space

In the future, there will be a lot of mobile hosts and wireless sensors, which implies a huge demand for addresses. So the new Internet architecture should provide huge locator and identifier spaces for future purpose.

5) Deployability

While adding more benefits into the new Internet architecture, we should not neglect the cost. So we want to reuse IPv4 address and eliminate the necessity of adopting IPv6 for huge address space. We also want to ease the management of a global mapping system with hierarchical host identifier and hierarchical Distributed Hash Table (DHT) technology.

3. Architecture Description

3.1. Host Identifier Namespace

In the Host Identity Protocol [5], each host will have a globally unique Host Identifier (HI) and a corresponding Host Identity Tag (HIT). HI is the public key of an asymmetric key-pair. HIT is a 128-bit datum created by taking a cryptographic hash over the HI, which is a flat label without any semantics. In most cases, it's the HIT that plays the role of host identifier due to its benefits of fixed-length and independence of the cryptographic algorithms used.

HRA borrows some idea from the HIP, however, it introduces a 128-bit hierarchical host identifier shown in Figure 1, which is composed of an Administrative Domain (AD) ID and a hash value of AD ID and the public key. The AD ID is a hierarchical label with embedded organizational affiliation and global uniqueness. The purpose of the hierarchical host identifier within HRA is to ease the management of a global identifier namespace and improve the lookup efficiency in mapping system.

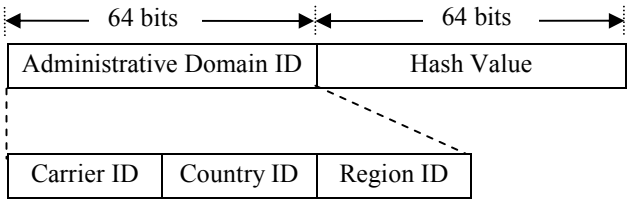

Figure 1: Hierarchical Host ID Format

In fact, the generation of the hierarchical host identifier is much similar to Cryptographically Generated Addresses

(CGA) [6]. In CGA, The process of generating a new CGA takes three input values: a 64-bit subnet prefix, the public key of the address owner as a DER-encoded ASN.1 structure of the type SubjectPublicKeyInfo, and the security parameter Sec, which is an unsigned three-bit integer. In HRA, the process of generating a new HI takes three input values: a 64-bit AD ID, the public key and the security parameter Sec. The difference between CGA and HRA is in the semantics of the 64-bit string.

To ease the deployment of this new architecture, we can adopt IPv6 address as the host identifier at the initial phase of deployment.

Of course, the flat HIT in HIP can still be used as host identifier within HRA.

3.2. Host Locator Namespace

HRA does not require globally unique IP address (also called as locator). Multiple independent address spaces (also called as locator domains) could coexist within HRA. Each locator domain (LD) may deploy an independent address space, that is to say, different LDs may deploy different networking technologies, in particular IPv4, IPv6, global and private address spaces, or difference LDs can deploy overlapped address spaces. Each LD has a globally unique ID, which is a hierarchical label as shown in Figure 2. In nature, a combination of LD ID and locator is a new globally unique locator.

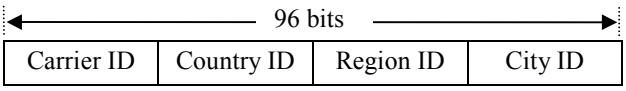

Figure 2: Hierarchical LD ID Format

3.3. ID/Locator Mapping Resolution

ID/locator split implies a need of storing and distributing the mapping of identifier and locator.

Within HRA, the mapping of host name and HI is stored in DNS, while the mapping of HI, LD ID and Locator is stored in DHT, so the host will need two-step query to get the HI, LD ID and locator of the destination host.

In contrast with flat HI, the mapping lookup efficiency can be improved evidently by using the hierarchical HI in a hierarchical DHT system [11].

3.4. Inter-LD Routing Protocol

Within HRA, LDs are connected via Locator Domain Border Routers (LDBR). A LDBR has at least one locator in each LD to which it is connected, and these locators have only LD-scope meanings and uniqueness. The adjacent LDBRs exchange LD reachability information with an inter-LD routing protocol. BGP can be extended with a new address family to fill this need. Besides, we can also design a new link-state protocol or distance-vector protocol as a inter-LD routing protocol from scratch. The LD ID can be aggregated into LD prefix provided some distance-vector protocol is deployed as inter-LD routing protocol.

3.5. Packet Format

Generally, the LD ID and HI of the source host and the destination host should be contained in the packet, whereas the locator of the source host and the destination host is optional. The purpose of carrying the destination host locator in the packet is to keep the LDBR of the destination LD from performing mapping lookup, that is to say, once the packet arrived at the LDBR of the destination LD, the LDBR just needs to replace the destination IP address with the destination host locator. During the transmission, the HI and LD ID fields usually remain unchanged, whereas the destination IP address and the source IP address in the IP header will be continuously rewritten by each-hop LDBR along the path to the destination.

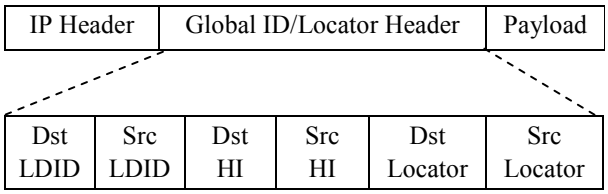

Figure 3: Packet Format

In IPv6 packets, the global ID/Locator header could be carried as an extension header, while in IPv4 packets, the global ID/Locator header could be carried as a new-type payload.

3.6. Packet Forwarding Behavior

3.6.1. Host Behavior

Generally, a source host will firstly obtain the locator and the LD ID information of a destination host from a distributed mapping system before initializing a communication with the destination host. If the LD ID of the destination host is the same as its own, the source host will encapsulate the packet with the destination IP address being filled with the destination host locator and send it out, otherwise, the source host encapsulate the packet with the destination IP address in the IP header being filled with one of its LDBR locator and send it out.

The Host can get its local LDBR information in one of the following options: 1) LDBR information is contained in the Router Advertisement extension; 2) LDBR information is carried in the Dynamic Host Configuration Protocol (DHCP) extended option; 3) LDBRs provide a well-known anycast address for hosts to access.

3.6.2. LDBR Behavior

Except for exchanging the LD reachability information with each other, an LDBR can receive those packets with the destination being one of its locators, and forward those packets on basis of the destination LD ID and locator within those packets. Besides, an LDBR can also do some source LD validation, similar to the source IP address validation mechanism in current Internet.

3.6.3. Non-LDBR Router Behavior

There is no additional requirement on the Non-LDBR routers. These routers just forward the received packets according to the destination IP address.

3.6.4. Packet Forwarding Procedure

HRA introduces a hierarchical routing mechanism which is composed of LD-based routing and prefix-based routing. The former is used for inter-LD routing while the latter is used for intra-LD routing.

Let's illustrate the forwarding procedure with Figure 4.

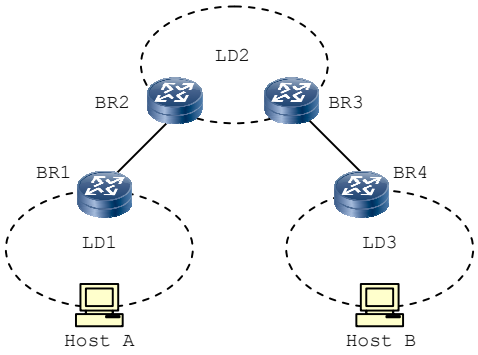

Figure 4: Topology Example

In Figure 4, host A will get the HI, LD ID and locator of host B before sending packets to host B, since the LD ID of host B is not the same with its own, host A will send the packets out with destination IP address being filled with the locator of one of it's LDBR, BR1, and source IP address being filled with its own locator, the LD ID and locator fields will also be filled in. When the packets arrive at the BR1, BR1 will lookup the LD routing table on basis of the destination LD ID of the packets, since the next-hop LDBR of the matching LD routing entry is BR2, BR1 will rewrite the destination IP address and source IP address in the IP header with the locators of BR2 and its own respectively, which are the IP addresses of the interconnected interfaces of BR1 and BR2. When BR2 receives the packets, it will also lookup the LD routing table, since the next-hop LDBR of the matching entry is BR3, BR2 will rewrite destination IP address and source IP address in the IP header with the locators of BR3 and its own respectively, which are routable in LD2. When BR3 receives the packets, it will forward these packets to BR4 according to the LD routing table. When the packets arrive at BR4, BR4 will lookup the LD routing table and find the next-hop for the matching entry is itself BR4, which means the packet has arrived at the destination LD, BR4 will subsequently lookup the corresponding prefix routing table of the destination LD and fill the destination IP address and the source IP address respectively with the destination host locator and one of its locators, which is routable in LD3 and send them out. In the end, the packets will be forwarded to host B by the internal routers within LD3 according to the destination IP address.

To some extent, HRA lifts the routing granularity of the current Internet one level, that is, the LD within HRA looks

like IP subnet, the LDBR within HRA looks like IP router, and the locator within HRA looks like MAC address.

3.7. Mobility and Multihoming

3.7.1. Host Mobility and Multihoming

When hosts change their attachments during to mobility or re-homing, they should register to the mapping system with their new location information and notice the correspondent host their current location.

3.7.2. Site Multihoming

Hosts within a multi-homed site usually have more than one locator, and these locators can be obtained either from different LDs or from the same LD. There should be no overlapping among the locator blocks allocated from the different LDs. Once overlapping occured, there should be some mechanism to assure that the same locators from different LDs to be allocated to one host. Otherwise, packets can not be forwarded correctly to the proper host since routing within LD is only based on the locator. In accordance with the uni-cast Reverse Path Forwarding (uRPF) policy implanted in Internet Service Provider (ISP) network, the site edge router should be able to forward the outgoing packets according to the source LD ID in those packets.

The mapping of HI, LD ID and locator will show the multi-homing status of the host.

3.7.3. Network Mobility

To support network mobility, we still need to introduce some NEMO [7] like mechanism into the HRA. Since there can be multiple LDs with independent address spaces coexisting and the locator has only LD-scope uniqueness within HRA, so we need to do some extension to the current NEMO mechanism. The home agent within HRA will maintain the mapping between the globally unique home address and the globally unique foreign address. The globally unique address means the combination of the LD ID and the locator. The mobile router should update its location information, including the current foreign LD which the mobile router is currently located in, and the corresponding foreign locator that the mobile router has obtained from the foreign LD, to its home agent, as soon as its attachment changed.

The home agent can act either as LDBR or as host within HRA, in the former case, the home agent should just receive LD routing information and forward the packets with destination of the mobile network to the proper LDBR, in the latter case, the home agent just forwards the packet with destination of the mobile network to one of its local LDBR，which may not be the optimal LDBR.

Like the current NEMO mechanism, packets are forwarded via the home agent to the mobile router, and network mobility event is transparent to those hosts within the mobile network.

3.8. Traffic-engineering

3.8.1. Host-controlled Traffic-engineering

Hosts can select one of its LD ID and corresponding locator pairs when they send packets out. The site edge router can forward the outgoing packets to the proper upstream LD according to the source LD ID of the packets.

3.8.2. Site-controlled Traffic-engineering

In order to realize site-controlled traffic-engineering for the multi-homed site network, the site edge router can adjust the upstream LD by rewriting the source LD ID in the outgoing packets according to traffic-engineering policy.

For example, the source host, which is located in a site multi-homed to LD X and LD Y, choose the LD X as the source LD and send the packets out, the site edge router can rewrite the source LD ID of the outgoing packets as LD Y and forward them to the LD Y based on the source LD, when the destination host receive these packets, it will choose the locator of the source's which is corresponding to the source LD ID in the received packets as the destination locator in the replying packets. When the source host receives the above packets, it will recognize the address rewriting and use that address from then on.

4. Related Work

A lot of proposals have been put forward in IETF and other academic organizations to meet the challenges that current Internet architecture are facing, especially in scalability, i.e. Nimrod [8], ENCAPS[9], HLP[10], HIP[5], Node ID Internetworking Architecture (NIIA) [13], Internet Indirection Infrastructure (I3) [14], Layered Naming Architecture (LNA)[15], Routing on Flat Label (ROFL)[16]. Due to space limitations, this paper can only briefly compare these proposals to the HRA.

Nimrod, ENCAPS and HLP are of hierarchical routing approaches. Especially, ENCAP and HLP adopt AD-based routing mechanism, while Nimrod introduces a new hierarchical locator namespace. To some extent, they are also some kind of id/locator split ideas, but they reuse IP address as identifier.

HIP, NIIA, I3, LNA and ROFL are of id/locator split approaches and use a cryptographic string as the host identifier, which is a flat label. In contrast, HRA adopts a hierarchical label as the host identifier.

In the following, we will emphasis on the comparison between ENCAPS and HRA, NIIA and HRA respectively since they are more similar.

Node ID Internetworking Architecture (NIIA) [13] is an architecture that can work across multiple heterogeneous address domains, and support routing based on both locators within domains and NIDs or default routes between domains.

In respect of multiple locator domains coexistence, HRA looks similar to NIIA. The main differences from NIIA are:

1) Within the NIIA, there should be a stable core LD, and all the other LDs should be connected to the core LD directly or indirectly. Most of the traffic will go across the core LD. Within HRA, there is no limitation on the topology, that is to say, those LDs within HRA can be connected in mesh.

2) Within the NIIA, as the network topology is tree-based, there seems no need to run a LD-based routing protocol. Besides, the NR use host-based routing mechanism which means a potential scalability issues if a LD contains a lot of hosts. Within HRA, LDBRs exchange LD reachability information and support LD-based routing mechanism.

3) Within the NIIA, the existence and characteristics of connectivity between two locator domains, especially the edge locator domains, may change dynamically on relatively short timescales, due to routing changes, mobility or multi-homing events. LD mobility triggers host within the mobility LD to update the registration, especially when the CER is changed, that's to say, the LD mobility is not fully transparent to the host. Within HRA, the connectivity between locator domains is relatively stable and the mobility of partial network in LD still depends on the NEMO [7] like mechanism, and network mobility is transparent to those hosts within the mobile network.

ENCAPS [9] is a kind of inter-domain routing mechanism with routing on Autonomous Domains (AD). In respect of two-level routing architecture, HRA looks more like ENCAPS. The main differences between HRA and ENCAPS are:

1) ENCAPS doesn't introduce an independent host identifier namespace to hide the heterogeneity of different address spaces and so it can not support the co-existence of multiple independent address spaces.

2) ENCAPS adopts reserved IPv4 address for Autonomous Domains (AD) address and the AD address is directly used as tunnel destination address, which should be routable for internal routers within AD, whereas HRA uses the next-hop LDBR locator as the IP destination address in the IP header, and the IP address in the IP header will be rewritten by each LDBR along the path to the destination, which looks more like the usage of the MAC addresses between routers.

3) ENCAPS assumes that the current IP-Addresses can remain globally unique for a long time, and since the address space in each AD is not independent, ENCAPS is helpless in dealing with the depletion of IPv4 address space. On the contrary, the combination of LD ID and locator (if each locator domain adopts the same network address technology, such as IPv4) within HRA will form a new global locator namespace, which eliminates the necessity of adopting IPv6 for providing huge addresses.

5. Conclusion

Firstly, within HRA, only the LD-based routing information will be exchanged between LDs and the prefix-based routing information is just maintained within each LD. In this way, the routing table size in each DFZ router will be reduced greatly. That is to say, the routing scalability issue will be solved with HRA.

Secondly, prefix-based route change or route churn in one LD will not be flooded to another LD, which greatly improves the route stability.

Thirdly, provided that each locator domain adopts an independent IPv4 address space, a combination of LD ID and locator will become a new globally unique locator in nature, which eliminates the necessity of adopting IPv6 for providing huge addresses. Besides, most of the routers except LDBRs do not need to be upgraded, in respect of the forwarding plane.

Lastly, with adoption of the hierarchical HI, the lookup efficiency for id/loc mapping is improved further and maintain and management of the global HI namespace becomes more practical.

6. Acknowledgements

The authors would like to thank Louise Burness, Philip Eardley and Marcelo Bagnulo Braun for their insightful comments. We also thank other reviewers for their useful comments that helped improve the paper.

7. References

[1] D. Meyer, L. Zhang, and K. Fall. "Report from the IAB Workshop on Routing and Addressing". Internet draft, draft-iab-raws-report-01.txt, work in progress, February 2007.

[2] T. Li, "Design Goals for Scalable Internet Routing", draft-irtf-rrg-design-goals-01, July 2007.

[3] N. Chiappa, "Endpoints and Endpoint Names: A Proposed Enhancement to the Internet Architecture", URL http://ana.lcs.mit.edu/~jnc//tech/endpoints.txt, 1999.

[4] B. Carpenter, "Architectural Principles of the Internet", RFC1958, June 1996.

[5] R. Moskowitz and P. Nikander, "Host Identity Protocol (HIP) Architecture", RFC 4423, May 2006.

[6] T. Aura, "Cryptographically Generated Addresses (CGA)", RFC3972, Mar 2005.

[7] V. Devarapalli, R. Wakikawa, A. Petrescu and P. Thubert "Network Mobility (NEMO) Basic Support Protocol", RFC 3963, January 2005.

[8] I. Castineyra, N. Chiappa and M. Steenstrup, "The Nimrod Routing Architecture", RFC 1992, August 1996.

[9] R. Hinden, "New Scheme for Internet Routing and Addressing (ENCAPS) for IPNG", RFC 1995, June 1996.

[10] L. Subramanian, M. Caesar, C.T. Ee, M. Handley, M. Mao, S. Shenker and I. Stoica, "HLP: A Next Generation Inter-domain Routing Protocol", SIGCOMM'05, August 2005, Philadelphia, Pennsylvania, USA.

[11] L. Garces-Erice, E. Biersack, P. Felber, K. Ross, and G. Urvoy-Keller, "Hierarchical Peer-to-peer Systems" In Proc. Euro-Par 2003, Klagenfurt,Austria, 2003.

[12] M. O'Dell, "GSE-An Alternative Addressing Architecture for IPv6", Internet-Draft, Feb 1997.

[13] B. Ahlgren, J. Arkko, L. Eggert and J. Rajahalme,"A Node Identity Internetworking Architecture", 9th IEEE Global Internet Symposium, Barcelona, Spain, April 2006.

[14] Ion Stoica, Daniel Adkins, Shelley Zhuang, Scott Shenker and Sonesh Surana, "Internet Indirection Infrastructure", Proc. ACM SIGCOMM, Pittsburgh, PA, USA, August 2002.

[15] Hari Balakrishnan, Karthik Lakshminarayanan, Sylvia Ratnasamy,Scott Shenker, Ion Stoica and Michael Walfish, "A Layered Naming Architecture for the Internet", Proc. ACM SIGCOMM, Portland, Oregon, USA, August 30 - September 3, 2004.

[16] M. Caesar, T. Condie, J. Kannan, K. Lakshminarayanan, I. Stoica and S. Shenker, "ROFL: Routing on Flat Labels", SIGCOMM'06, September 2006, Pisa, Italy.

[17] H. Soliman, C. Castelluccia, K. El Malki, L. Bellier, "Hierarchical Mobile IPv6 Mobility Management (HMIPv6) ", RFC4140, August 2005.