

A Communication Mechanism for Routing Architecture for the Next Generation Internet Based on ID/locator Split

Shaochen Ruan, Meilian Lu, Wendong Wang, Xiangyang Gong and Xiaohu Xu

State Key Lab of Networking & Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China

Network Research Department, Huawei Technologies, Beijing, China

e-mail: ruanshaochen@gmail.com, mllu@bupt.edu.cn, wdwang@bupt.edu.cn, xygong@bupt.edu.cn, xuxh@huawei.com

Abstract—The current Internet architecture problems are becoming increasingly serious, the separation of the over loading of IP address semantics is necessary. RANGI (Routing Architecture for the Next Generation Internet) is a new solution of routing scalability, mobility and multi-homing for the routing architecture. We propose a new Internet communication mechanism through extend RANGI architecture. There are three main contributions: First, we design a communication mechanism which is backward-compatible with legacy host; second, we propose a handshake protocol to provide lightweight security for communications based on CGA verification; third, we propose a ID/locator replacement mechanism in ID layer which enhance the communication efficiency.

Keywords; RANGI; ILMS; RCDB; LD; ID/locator split

1. Introduction

With the rapid expansion of the scale of the Internet and network technology innovation, global BGP routing table entries were undergoing "super-linear" growth which cause increased investment and operating costs for ISP, slow inter-domain routing convergence and stability decreased. IAB's "routing and addressing working group meeting" declare that the major factors which lead to the rapid growth in the global routing table entries are: multi-homing, traffic engineering, provider-independent address and commercial factors (such as mergers and acquisitions)[1]. These factors break the IP address prefix aggregation, resulting in a large number of address prefix which can not be aggregated inject into the global Internet routing table. But the root of the problem is the overloading of IP address semantics: IP address represents both the host identity (the session endpoint at the transport layer), and the host location (the node locator of network topology). Therefore, IAB introduce two name spaces to represent the identity and locator of the node to resolve the overloading of IP address semantics, this is so-called "identifier/locator split". Identifiers are used in the application and transport layers for identifying nodes, and locators are used in the network layer for locating nodes in the network topology [2].

Current solutions for ID/locator split can be divided into two categories: host-based solutions to achieve separation, such as HIP, Shim6, and so on; network-based solutions to achieve separation, such as LISP, GSE, and so on.

HIP is a kind of id/locator split proposal which inserts an the HI (host identity) layer in the middle of the network layer and the transport layer, and use the HI as the identity of the node, IP address is used only for routing. HIP supports mobility and multi-homing effectively, and the security is good, but there are backward-compatibility issues and a lot of certification which reduce the protocol efficiency [3].

LISP uses a part of IP address as host identity called EID, the other part of IP address uses as the routing location called RLOC, through the introduction of ITR (Ingress Tunnel Router) and ETR (Egress Tunnel Router), re-encapsulate the existing IP data packets, in order to achieve the separation of EID and RLOC. LISP does not make any changes to existing host protocol stack, but it has these problems such as poor mobility supporting, and the cost of encapsulation packets reduce the efficiency of communication, and backward-compatibility issues with legacy host[4].

Huawei proposed a new routing and addressing architecture based on the strength of the industry's outstanding technical solutions: RANGI. Compared to HIP and LISP, RANGI not only support mobility and multi-homing, from the introduction of the concept of hierarchical, RANGI also has a reasonable business model and clear trust boundaries, and solve the problem of routing scalability, simplify the site re-addressing and traffic engineering[5]. The proposed RANGI transitional mechanism allows the traditional host to interact with the RANGI host[6], and achieve the incremental deployment in existing Internet, but it introduces a new proxy entity for temporary ID distribution and mapping maintenance, it will not only increase deployment costs, but with the continued expansion of network scale, the processing power of the proxy is also a severe test. Therefore, RANGI also has backward-compatible issues.

We can summarize the following problems: backward-compatibility issues with legacy hosts; a large number of security encryption and authentication lead to inefficient communication for the lower computing power equipment or the mobile terminal; encapsulated packet take additional load on the network, and may also lead to re-fragment the packet, which would affect communication efficiency. In an attempt to deal with the above problems, this paper proposes a new communication mechanism based on RANGI framework, it not only inherited RANGI advantages, but also extended RANGI framework, which can effectively compatible with legacy host, provides the lightweight security with host communication and provide effective communication mechanism.

2. Introduction of Rangi Framework

RANGI belongs to the host-based ID/locator split solutions, the protocol architecture is illustrated in Figure 1, RANGI inserts a host ID (identity) layer in the middle of the network layer and the transport layer in existing TCP / IP protocol stack. Transport layer session only bind with the ID layer, locator layer is only responsible for routing and forwarding packets. So as to solve the overloading of IP address semantics in existing Internet architecture [5].

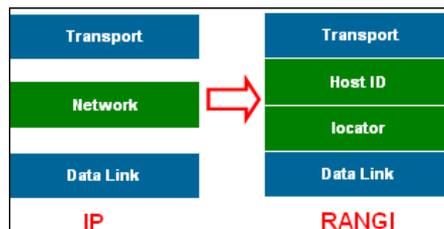


Figure 1. RANGI protocol architecture

2.1 Hierarchical Host ID

Host ID not only has the general meaning of identity uniqueness and global topology independence, in order to overcome the HIP flat ID's lacking of commercial and security model and management deficiencies, RANGI host ID use hierarchical idea and consists of two parts: the first part represent organization affiliation, such as the host ID's national information, regional information and ID management authority where ID belongs to, this part is called AD ID (Administrative Domain ID); The second part is the Local Host ID which is generated by computing a cryptographic one-way hash function from a public key of the ID owner and auxiliary parameters, as shown in Figure 2. 128-bit ID ensures that is compatible with the IPv6 upper application to the maximum extent.

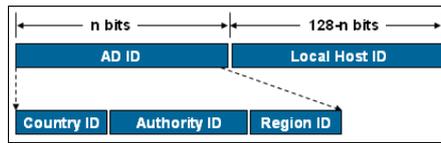


Figure 2. Hierarchical host ID

2.2 Hierarchical Locator

The host locators in RANGI are ordinary IPv6 addresses. Since the IPv4/IPv6 coexistence and transition will last for a long period, in order to reduce the deployment cost of this new routing and addressing architecture, RANGI uses specific IPv4-embedded IPv6 addresses as locators, as depicted in Figure 3. RANGI use the idea of the coexistence of multiple overlapping independent IPv4 address space, network will be divided into several locator domain (LD), each LD will be assigned one or more globally unique domain ID for identification, it is called LD ID; within each LD the node location is identified by IPv4 address space, and routing is based on IPv4 addresses. Different LD can independently allocate addresses and select routing protocols; IPv4 address is unique within the LD. Therefore, we can satisfy the needs of network autonomy, user can change LD very flexible and without the need for extensive re-allocation of IPv4 addresses.



Figure 3. Hierarchical locator

2.3 Hierarchical Routing System

RANGI use hierarchical routing mechanism: LD ID based routing between LDs, and IPv4 addresses based routing within LD. As shown in Figure 4, the packet always forward according to destination LD ID before it arrives at destination LD, because the router within LD only support IPv4 routing, you can use IPv6 over IPv4 tunnel to cross the IPv4 router which does not support the LD routing. BGP connection between border routers exchange LD-level routing information, LD ID can be allocated according to network topology, thus it can be aggregated into LD prefix and has a good scalability.

3. A communication Mechanism for Rangi

Figure 4 is RANGI basic topology, the mapping server includes: DNS for FQDN_to_ID and FQDN_to_IP, and ILMS (ID/locator Mapping System) for ID_to_locator. Two hosts are in different LD, routing is based on destination locator.

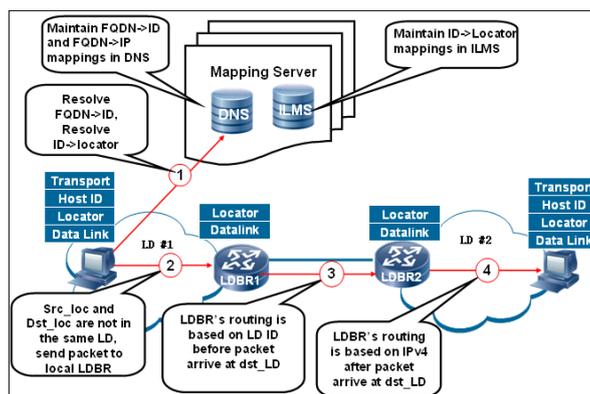


Figure 4. RANGI basic topology

It can be seen from the topology: RANGI communication mechanism involves ID generation and registration before communication, DNS extensions, ID_to_Locator mapping, connection establishment and send or receive data packets. The following describes these several aspects above to introduce RANGI communication mechanism in detail.

3.1 ID Generation and Registration

Hosts which enable RANGI protocol must generate the host ID and register to its own ILMS before communication, so ID can be managed and realize ID/locator split communication. The host can achieve ID in variety of ways, such as online applications, or similar to the distribution of SIM cards from ISP etc. We adopt the way as follows: host apply for AD ID from ICANN alike authority, then call the CGA algorithm to generate a complete ID and register to ILMS, ILMS execute ID verification and collision detection, if the ID is not valid, host ID will be regenerated and re-register until it reach the maximum number of ID conflicts.

The ID name space can be divided, as shown in Figure 5: ID namespace prefix 0x2f (ID tag), local host ID for the hash part, the rest is according to standard e.164 coding. Such as 0x86 for China, 0x0010 for Beijing, the last four byte of region ID is reserved as 0x0000.

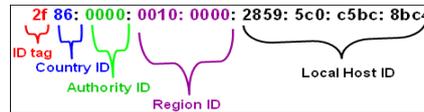


Figure 5. ID example

If we see the IPv6 address prefix in CGA as AD ID, the CGA input parameters (64-bit IPv6 prefix, RSA public key and security parameters) can just match the n-bit AD ID, public key and security parameters in RANGI framework, it only need to set the value of n to 64. This AD ID and Local Host ID finally assemble into a 128-bit host ID[7].

3.2 DNS Extensions

RANGI use the AAAA RR (Resource Records) of traditional DNS to support FQDN_to_ID mapping resolution, but does not distinguish the meaning of the AAAA records: does it store host ID or traditional IPv6 address?

In fact, only the smaller expansion of traditional DNS can support backward compatibility for RANGI host and legacy IPv6 host. That is to say, based on the AAAA RR of traditional DNS which supports FQDN_to_IP mappings, we add a new ID types RR rather than AAAA records defined in RANGI to support FQDN_to_ID mappings, and constructing a new query type to query the ID types RR [7].

In RANGI communication mechanism, there are two types of queries:

- ID query: RANGI-aware host sends such a request by default, if peer host is also a RANGI-aware host and register the new FQDN_to_ID RR to the DNS server, DNS server will return the peer host ID; if peer host does not register the new FQDN_to_ID RR to the DNS server, it demonstrate that the peer host is traditional host, this time DNS server returns the IP address of the peer host.
- IP query: This is the traditional DNS query request, there is no difference with the existing mechanism, and DNS returns the IP address of the peer host.

3.3 ID/locator Mapping Mechanism

ID/locator split implies a need for storing and distributing the mappings from host IDs to locators. Since ID does not support the global scope routing, in order to forward the packet to the remote domain, we need map ID to its corresponding locator.

The implementation of RANGI ID/locator mapping system is based on reverse DNS; it provides the function of mapping information registration, query and update. We need to define a new type of resource record to store the ID_to_locator mapping. In addition, A 128-bit host ID will be transformed to a FQDN format string, such as “Country ID.Authority ID.Region ID.Hash value” is transformed into a FQDN-format string as “Hash value. Region ID. Authority ID. Country ID”, thus we can reuse DNS mechanism to support ID / Locator mapping[8].

3.4 RANGI Connection Establishment

We designed a three-way handshake protocol for RANGI connection establishment before data communication. Because CGA algorithm is used to generate ID, we should just use the CGA authentication mechanism to authenticate the identity of the peer host, thus two hosts will establish a trust relationship. Considering that ID layer is between network layer and transport layer, the handshake message packet does not require transport layer treatment, but rather direct access to IP layer, there is no need to create a transport

layer and higher layer information structure, so we select raw socket to load RANGI handshake message packets and transform it as IPv6 extension header, packet format is shown in Figure 6.

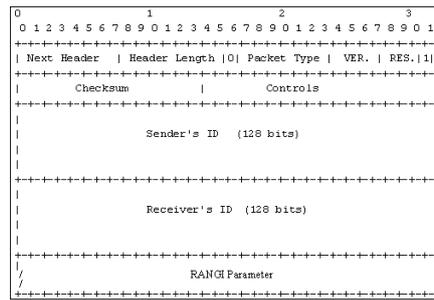


Figure 6. RANGI packet format

RANGI parameters in a RANGI handshake message, packets are stored like the TLV format and the message can be divided into three types:

- connection_request: its RANGI Parameter field contains the CGA parameter structure.
- connection_confirm: its RANGI Parameter field contains the CGA parameter structure and random SQN (SeQuence Number).
- connection_ACK: its RANGI Parameter field only contains the confirmation number ACK, the value is SQN+1.

In order to manage the state of the RANGI connection establishment and implement the mapping between ID and locator, we create a RCDB (RANGI Connection DataBase), which implements the RANGI connection establishment state machine and maintain the ID_to_locator mappings which belong to the current active user, but this user must have passed the authentication. Mapping information maintained by RCDB is always effective and will not be deleted until the user disconnects network or shuts down the connection. The connection state information is stored in the form of CSR (Connection State Record), it includes: our ID, peer ID, our locator, peer locator and the current state of connection.

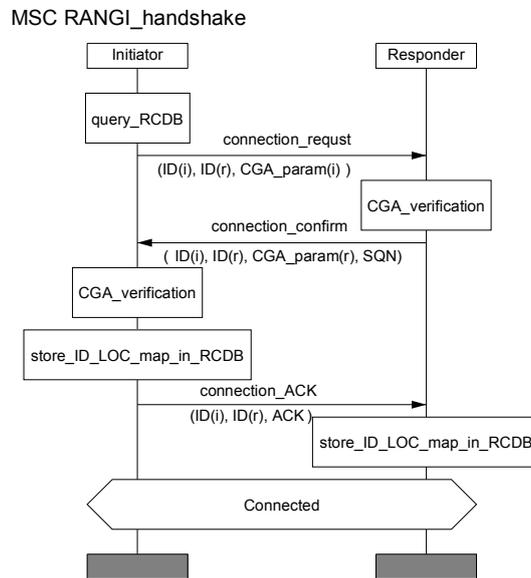


Figure 7. RANGI connection establishment process

As shown in Figure 7, in the RANGI connection establishment process, the initiator first query RCDB to check whether there is an active existing CSR in RCDB, if not then initiate the handshake process; after connection_request message arrives at responder, the responder use CGA verification to authenticate the identity of initiator, if the verification succeed then respond with connection_confirm message; when connection_confirm arrives at the initiator, we use the same verification process, if succeed then create a CSR in RCDB at the initiating side which stores both sides ID and locator mappings, and respond with

connection_ACK message; the responder will compare ACK with SQN+1 when it receives connection_ACK, if succeed then create a CSR in RCDB at the responding side, thus RANGI connection is established.

RANGI connection establishment state machine is shown in Figure 8, it can be divided into four states: idle, conf_sent, req_sent and established. The state transition is based on the events depicted in the diagram.

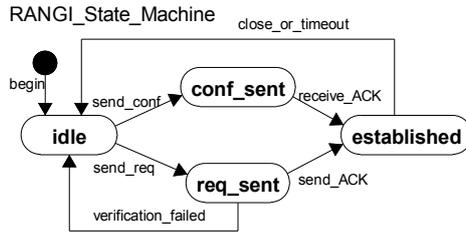


Figure 8. RANGI state machine

3.5 Send and Receive Packets

It is required to classify packets in the communication to achieve the backward-compatibility features, so as to separate the traditional TCP/IP communication and RANGI communication.

As shown in Figure 9, the packets down from the upper layer arrives at ID layer.

- It needs to determine whether the address in the packet is an ID based on the ID tag part of the address field, so as to determine whether to execute RANGI communication.
- If it is ID, it needs to determine whether the packet belongs to an established RANGI connection, using the source and destination ID as searching keywords.
- If it does not establish RANGI connection, then enter the module of RANGI connection establish and send connection_request.
- If RANGI connection has been established, then enter the module of RANGI basic communication and find the active existing CSR in RCDB, source and destination ID in the IPv6 header are replaced by the corresponding source and destination locator, and then send the packet to the network layer.
- If it does not need RANGI communication, just send the packets directly to the network layer to execute the traditional TCP/IP communications.

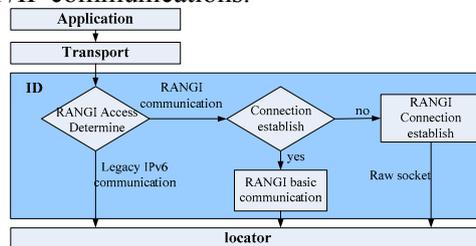


Figure 9. Packet sending procedure

As depicted in Figure 10, the network layer receives the packet:

- At first, it needs to determine what the protocol is, if the packet is RANGI connection establishment packet which has RANGI header, then enter the part of RANGI connection establish;
- If it is TCP or UDP then determine whether it need RANGI communication, if it dose then enter the module of RANGI basic communication and find the active existing CSR in RCDB, using the source locator and destination locator as searching keywords to query ID_to_locator mappings.
- Source locator and destination locator in IPv6 header are replaced by source ID and destination ID.
- If it dose not need RANGI communication then execute the traditional TCP/IP communications.

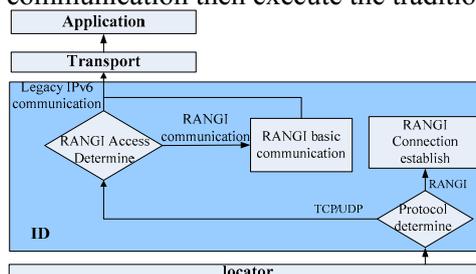


Figure 10. Packet receiving procedure

In summary, we determine the need for mapping replacement according to the ID tag field at sender's ID layer, and then determine whether a CSR has in RCDB at receiver's ID layer to need backward mapping replacement. Therefore, ID/location split communication is realized by the host side.

4. Analysis

4.1 Backward Compatibility

RANGI protocol mechanism can support the communication between RANGI-aware host and legacy host by extending the traditional DNS and classifying packets. Here, we call a host using RANGI protocol RANGI_A and another host without RANGI Legacy_B.

Scenario One: RANGI_A sends request to Legacy_B for data communication.

1) RANGI_A queries DNS for ID resolution through the FQDN of legacy_B. The DNS returns an IP address due to the non-existence of the FQDN_to_ID type DNS RR. The ID layer of RANGI_A will first check the ID tag. RANGI_A will ignore the ID layer if the ID tag is an IP address and communicate with Legacy_B through traditional TCP/IP protocol.

2) When the RANGI_A only knows the 128bits-ID instead of the domain name of Legacy_B, the RANGI_A will also forbid the RANGI protocol by checking the ID tag whether it is an ID or an IP address in ID layer.

Scenario Two: Legacy_B sends request to RANGI_A for data communication.

1) On the Legacy_B side, Legacy_B first queries DNS for IP resolution through the FQDN of RANGI_A. The DNS returns the IP address of RANGI_A according to the traditional DNS mechanism. Then Legacy_B will send packets to RANGI_A using its IP address. On RANGI_A side, RANGI_A will first check the connection state of RANGI in RCDB by locator. If there is no CSR about the RANGI connection, RANGI_A will communicate with Legacy_B through the traditional TCP/IP stack and forbid the RANGI protocol.

2) If Legacy_B only knows the IP address of RANGI_A, it will use traditional TCP/IP protocol as described in part 1).

To sum up, compared to the HIP and LISP protocol, RANGI has some advantages on the backward compatibility and it can realize incremental deployment.

4.2 Security

RANGI communication mechanism effectively uses the address encryption generating mechanism and the address authentication mechanism according to the standard CGA algorithm, and provides CGA-based authentication mechanism between the two communication sides. If successfully passed the verification in RANGI connection establishment process, the verification program can determine the CGA public key parameter is the address of the owner, and is reliable. Two communication sides setup a trust relationship for data transmission, it can identify malicious network entities and traffic.

4.3 Communication Efficiency

After the introduction of ID layer, It seems that the header of RANGI protocol should be included in each packet, which contains both source ID and destination ID, then the data packet can be correctly sent to the communication endpoint. In fact, in RANGI communication mechanism, it only sends packet with RANGI header when deal with ID layer related functions. This is because: first, the length of RANGI header is relatively long, we can find there are at least 40 bytes besides RANGI parameter field in the RANGI handshake message packet format. If each IP packet encapsulates an RANGI header, it will bring additional network load, and even may lead to the packet- fragmentation because of MTU limits. Secondly, after RANGI connection established, the two communication sides have stored each other's ID/locator mappings in their respective RCDB, as a result the RANGI connection implicitly includes both source ID and destination ID of the packet, and the mapping replacement between ID and locator is based on RANGI connection state maintained by RCDB.

5. Conclusion

Based on the analysis of the existing ID/locator split solutions, in this paper we propose a RANGI communication mechanism based on RANGI framework: we inherit RANGI framework advantages and extend it: define the ID name structure and describes the ID generation and registration; extend the traditional DNS and execute ID/locator mapping replacement by host side which makes RANGI host compatible with legacy IPv6 hosts in to realize incremental deployment; RANGI connection establish a trust relationship for data transmission between the two sides, ID generation and authentication are consistent to provide lightweight security for communications based on CGA algorithm.

In the future, we will implement the prototype system, and compare with other ID/locator split solutions and evaluate the optimal next-generation Internet architecture solution.

6. References

- [1] T. Li, "Design Goals for Scalable Internet Routing," draft-irtf-rrg-design-goals-01, July 2007.
- [2] D. Meyer, L. Zhang, and K. Fall. "Report from the IAB Workshop on Routing and Addressing," draft-iab-raws-report-01(work in progress), February 2007.
- [3] R. Moskowitz and P. Nikander, "Host Identity Protocol (HIP) Architecture," RFC 4423, May 2006.
- [4] D. Farinacci, V. Fuller and D. Meyer, "Locator/ID Separation Protocol (LISP)," draft-farinacci-lisp-11(work in progress), December 2008.
- [5] X. Xu, "Routing Architecture for the Next Generation Internet (RANGI)," draft-xu-rangi-04, August 2010.
- [6] X. Xu, "Transition Mechanisms for Routing Architecturefor the Next Generation Internet (RANGI)," draft-xu-rangi-proxy-01, July 2009.
- [7] T. Aura, "Cryptographically Generated Addresses (CGA)," RFC3972, Mar 2005.
- [8] S. Thomson and C. Huitema, "DNS Extensions to support IP version 6," RFC 3596, October 2003.