

A High-Availability Network for Reducing Network Operation Burden and Supporting Human Mobility

Hiroaki HARAI

In this paper, we address our R&D towards a high convenience and reliable networking. From a user's point of view, this network provides heterogeneous, secure communication and low-delay and stable throughput communication. From an operator's point of view, it provides high-speed, reliable, and easy operation management. We address ID/Locator Split Network HIMALIS and Hierarchical and Automatic Number Allocation HANA, and introduce outcome through these researches.

1 Introduction

The present networks have the following problems: (1) the growth of circuit-scale, and processing complexity for prompt destination address search, and route-convergence time, caused by bloated routing tables; (2) the increase of administration cost due to simultaneous management of more than one network layer; (3) the degradation of availability caused by human error during manual operations in address configuration and name resolution; and (4) the increase of delays in mobile communications and inaccessibility between heterogeneous network protocols. For example, in (1), the number of forwarding table entries at the Internet backbone routers reached 500,000 in 2014. The architecture of TCAM (Ternary Content Addressable Memory) for destination address lookup consumes energy proportional to the number of entries in the forwarding table. The increase in the circuit footprint leads to increase in the power consumption. For (2), different networks such as large-scale IP (Internet Protocol), Ethernet, MPLS (Multi-Protocol Label Switching), optical networks, and OpenFlow at the data centers are integrated in operation and management. (3) Approximately 2% of accidents for telecommunication services reported to the Ministry of Internal Affairs and Communication of Japan comes from human errors^[1]. (4) Mobility causes increase in the end-to-end delay and a single point of failure because even if a host moves to different network, data is passed by the home agent or the local mobility anchor in Mobile IP and Proxy Mobile IP.

In this paper, we introduce our R&D that solves the above-problems and aims the realization of following

features.

- User advantage: This network provides a high convenience and reliable communication environment, namely, heterogeneous, secure communication and low-delay and stable throughput communication based on state-aware automatic route selection.
- Operator advantage: This network provides high-speed, reliable, and easy operation management.

More concretely, our target is effective communication on multiple routes, heterogeneous (i.e., different network protocol) communication support, and management simplicity and automation on a multi-homed network where multiple routes are provided in the two nodes (Fig. 1). We also look at not only usual communication failure such as overload caused by concentrating communication data and equipment failure but also failure recovery from the disasters. Our proposal is to facilitate host mobility, heterogeneous communication, and secure communication by using a new layered architecture HIMALIS (Heterogeneity Inclusion and Mobility Adaptation through Locator Identifier Separation) where information identifier (ID) and location identifier (locator) are separated. Moreover, we provide a reliable, low-cost management network by using HANA (Hierarchical and Automatic Number Allocation) where locator assignment and name server registration are done automatically.

Our basic idea behind in our R&D is to make network structure and management simple and smart. The simplicity and smartness cause improvement of reusability and extendibility of networks. In other words, our developed methods are easily embedded in a new network R&D by other researches and are then widely deployed. We are sure

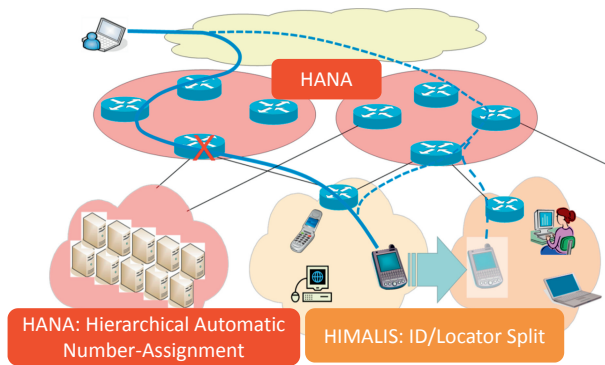


Fig. 1 Network overview for high-convenience, secure, reliable and operation efficient network on a multi-homed network structure

to make designs of HIMALIS and HANA software simple such that application developers and network developers are easily used. Also, we enhance software stability through large-scale verification on the large-scale emulation platform StarBED³ and wide-area deployment on the new-generation network testbed JGN-X.

2 ID/locator separation and automatic locator assignment

This section addresses ID/locator separation architecture HIMALIS and hierarchical automatic number allocation HANA. These are centered in our R&D. HIMALIS is a user oriented technology for facilitating mobility of hosts and increasing the number of communication objects. HANA is a network operator oriented technology for facilitating network equipment operation and decreasing convergence time of routing information of hosts. The network equipment in HIMALIS and HANA are maintained by registries where name or ID is mapped into locator. Additionally, we have the following features.

- Hierarchical locator structure and ID/locator separation contribute to decrease the routing table size and accelerate routing convergence even if the number of operators and hosts is increased.
- It makes a burden of network operator reduce to automate locator assignment and allocation and locator registration to the name registry and to authenticate access to the network and the registries.

2.1 HIMALIS

HIMALIS is a communication network architecture such that a host in a network can securely move to other network of which protocol and area are different^{[2]-[4]}.

Figure 2 shows overview of HIMALIS network. A HIMALIS network consists of a global transit network and multiple edge networks. HIMALIS features in terms of hosts are described as follows.

- Communication is seamlessly continued when a host in an edge network moves to a different edge network.
- The shortest route communication is provided after a host moves to a different edge network (Fig. 3).
- Communication is not interrupted once a host moves to a different edge network when an originally connected edge network fails (Fig. 3).
- A heterogeneous communication is provided between nodes each of which has different network-layer protocol each other.
- Communication is seamlessly continued even when a host in an edge network moves to other edge network of which network protocol is different.
- Multiple connections to two or more different upstream networks, in other words, multi-homing is supported.
- Host authentication function is supported.

More details are explained by using Fig. 2. Each edge network adopts arbitral network protocol such as IPv4 and IPv6. The edge network is connected to the global transit network by using HIMALIS Gateway (HGW) that has a protocol conversion capability. The global transit network uses a common network protocol. Each host is uniquely identified by using globally-unique node identifier (ID). Each HGW includes an authentication agent for a host authentication and local name server for name resolution. The agent and the server are likely to be installed in the edge networks. The transit network has a Domain Name Registry (DNR) and a Host Name Registry (HNR) which are accessible globally. These name registries form hierarchical name resolution architecture like Domain Name System (DNS).

HIMALIS provides each host with a globally unique hostname. The information of the host can be resolved by using name resolution servers. The global hostname is represented by concatenating a locally-unique local hostname and domain name using the concatenation symbol the '#' (e.g., host01#nict.go.jp). When a host connects to an edge network, it is given a local locator (LLoc) and the information for communicating with HGW. Once the access authentication is succeeded, the global locator of HGW (GLoc) is notified to the host. The host can be resolved from outside of the edge network after ID and GLoc of

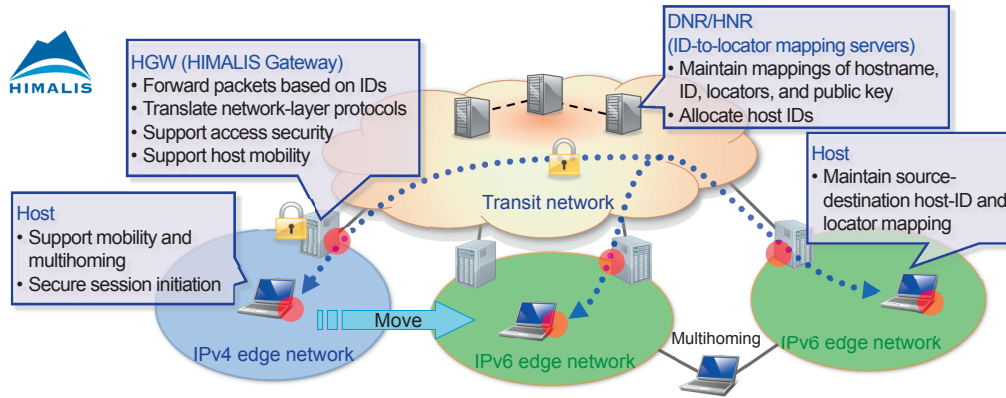
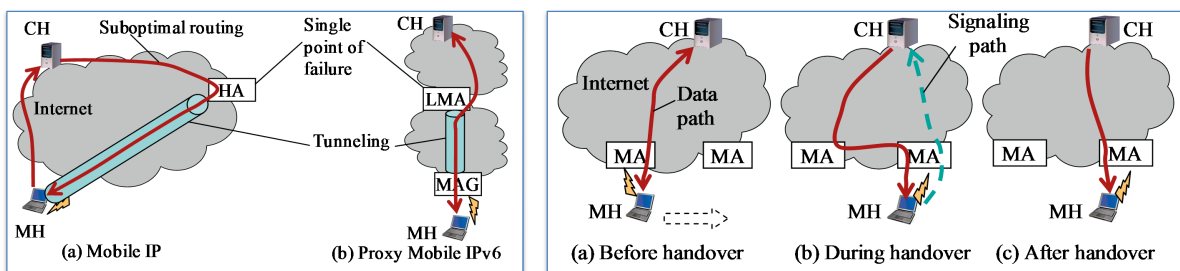


Fig. 2 HIMALIS overview and functions



MH (mobile host) CH (correspondent host) HA (home agent)
 LMA (Local mobility anchor) MAG (mobility access gateway) MA (mobility anchor)

Fig. 3 (Left) Delay increase and single point of failure in the existing mobile communication. (Right) Minimum delay and no single point of failure in the ID/locator split architecture.

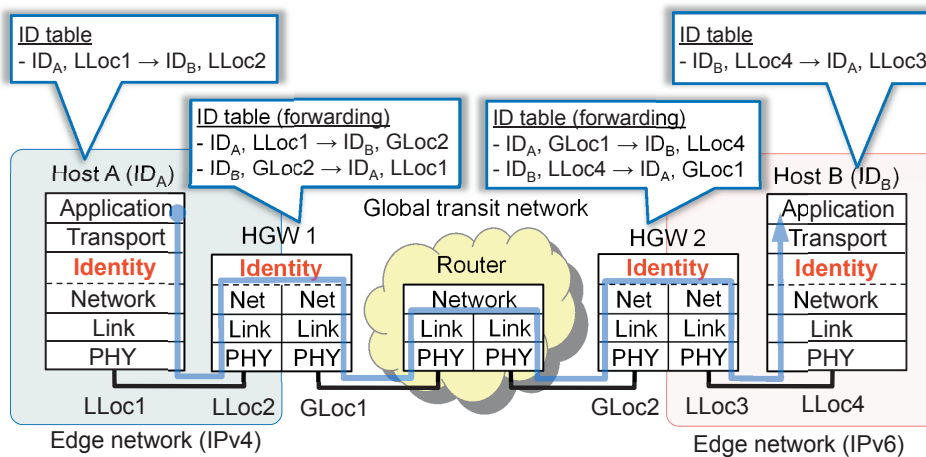


Fig. 4 Communication using the identity layer in HIMALIS

the host are registered to the HNR that maintains the corresponding edge network. When the host communicates with a correspondent host, the host resolves ID and GLoc of the correspondent host by inquiring the global hostname of the correspondent host at HGW. If the HGW cannot resolve the inquiry, the HGW inquires it to DNR and HNR. In more detail, the HGW inquires domain name of the correspondent host to DNR and resolves ID and GLoc of

the HNR that maintains ID and GLoc of the correspondent host. The HGW then inquires local hostname to the HNR and finally resolves ID and GLoc of the correspondent host.

HIMALIS has the identity layer between the transport layer and the network layer. Communication procedure using the identity layer is described with an example depicted in Fig. 4. When host A communicates with host B, host A obtains ID (ID_B) and the global locator (GLoc2) according

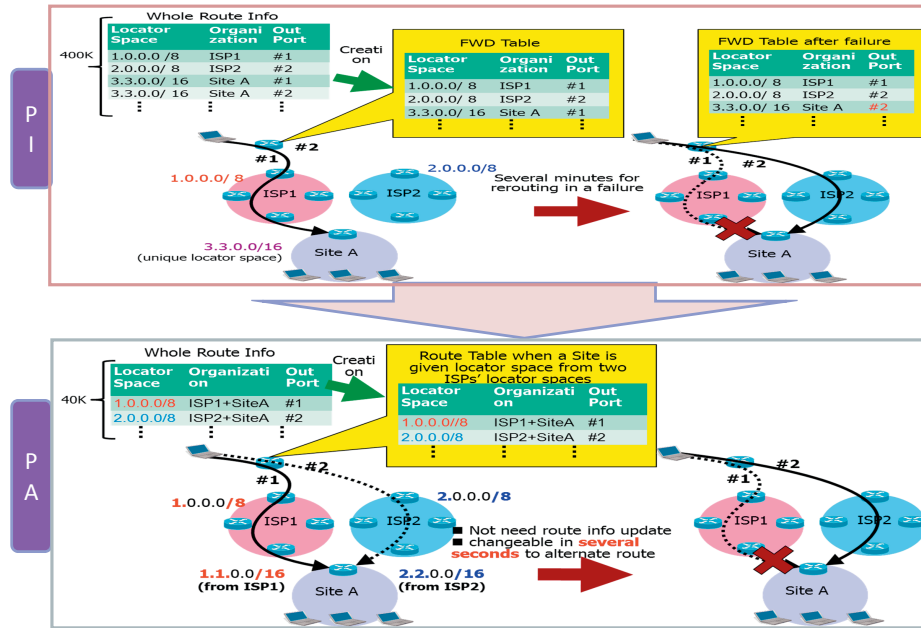


Fig. 5 Advantage of hierarchical multiple locator assignment

to the name resolution procedure. Host A then sends a session initiation request signaling to host B. Host B resolves host A's authentication by using name resolution system and then responds to host A. In the HIMALIS network, for a communication in the transport layer or upper layer, the information of the correspondent hosts has to be registered in the ID tables in the hosts and HGWs. The registration to the ID tables is done only in the case of success of communication initiation procedure for enhancing security in the communication initiation. Thus, HIMALIS uses ID tables in HGWs for communication control between the transit network and edge networks, packet transfer to other HGWs, and mobility. ID tables maintains information of working sessions only. The locator update information resulting from host mobility is notified to HNR and HGW. Thus communication to different hosts is also provided after moving of the hosts.

2.2 HANA

HANA assigns and configures locators to all network equipment such as routers and switches in a network^{[5][6]}. HANA hierarchically and automatically assigns locators to a multi-hierarchical network. Multi-homed networks are given multiple locators to the routers in the network. HANA has the following features. Figure 5 illustrates advantage of assigning multiple locators to a single node hierarchically.

- Automatic number assignment reduces address configuration time and reliefs the network administrator

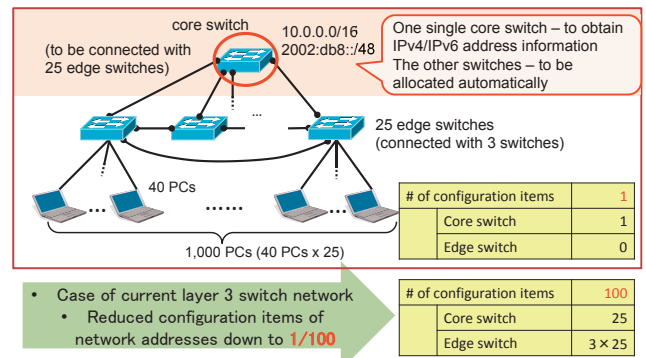


Fig. 6 Efficiency of HANA's automatic locator assignment

from human-error in configuration. In a network consisting of 1,000 hosts, HANA reduces address-allocation burden down to 1/100 (Fig. 6)^[7].

- Multi-locator assignment provides redundant routes natively. This makes network reliable and tolerable to failure and congestion of the network.
- Hierarchical locator assignment leads to routing table aggregation. Concretely, if HANA applies to the Internet backbone, the FIB (Forwarding Information Base) size, which is more than 500,000 at the time of 2015, is decreased to several tens of thousands^[8]. This is attractive although the required locator space, in other words, the number of required locators, is increased ten times because a node is given multiple locators.

Routers, switches, and hosts are assigned locators from a HANA server. When a network equipment joins

the network, it seeks neighbor nodes by using link-local multicast and establishes peer. In the peer network, messages needed to locator assignment are exchanged by using unique names (e.g., string of MAC address and FQDN (Fully Qualified Domain Name)) as identifiers.

HANA servers are likely to be installed hierarchically. The developed set of the HANA servers provides a hierarchical locator assignment. In HANA, a locator is divided into three portions called prefix, midfix, and suffix. A locator of host is determined by concatenating the 3 values. The value of each portion is assigned independently as shown below.

Prefix portion is periodically advertised from network upstream to downstream via PrefixInfo messages and distributed to all equipment in the downstream networks. In the case of Fig. 7, where prefix values are assigned from multiple upstream networks, each host of the network is

given multiple prefixes. This means the host has multiple routes natively. Midfix portion is allocated a unique one from the HANA server in the network. Suffix portion is decided by equipment own.

In the example of Fig. 7, a PrefixInfo message including prefix 1.1/16 from upstream ISP1 and that including prefix 2.2/16 from upstream ISP2 are distributed to ISP3. On the other hand, in ISP3, a HANA server periodically advertises MidfixOffer messages. The nodes receiving the MidfixOffer send midfix-value-request (MidfixReq) messages. When the HANA servers receives the MidfixReq message, it allocates a unique midfix value (among midfix space, i.e., front 17th bit to 24th bit) by using MidfixAck messages. Self-assigned suffix value (i.e., front 25th bit to 32th bit) is sent to HANA servers via MidfixReq messages. In doing so, the HANA servers maintain the latest FQDN and locator mapping of all hosts of which midfix are assigned by the HANA servers. By interworking DNS to the automatic update function of FQDN-locator mapping provided by HANA, DNS records can be updated automatically.

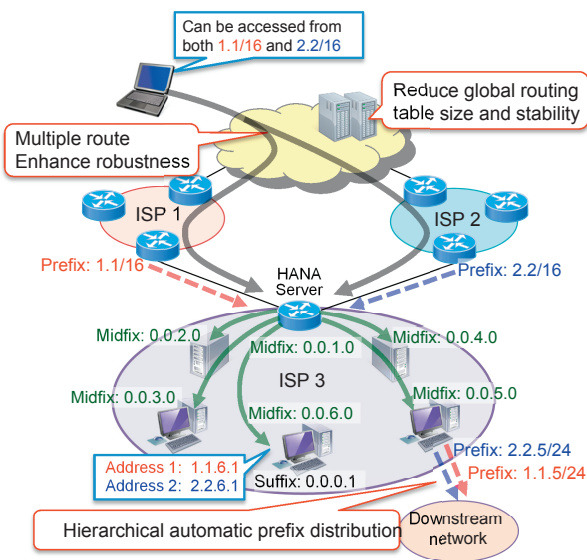


Fig. 7 Overview of automatic locator assignment and multiple route provisioning

3 Outcome

- We developed basic software of HIMALIS^{[4][9]-[11]}. We have implemented the identity layer by extending TCP/IP kernel (see Fig. 8). In spite of overhead caused by identity-layer insertion, we have verified that transmission throughput performance at hosts and latency-performance at the HIMALIS gateway are almost the same as those of existing TCP/IP based communication on a GbE-link network (see Fig. 9). We developed middleware for communicating existing TCP/IP applications on HIMALIS environment^[12].
- Security support. We designed and implemented an integrated security mechanism in HIMALIS, where

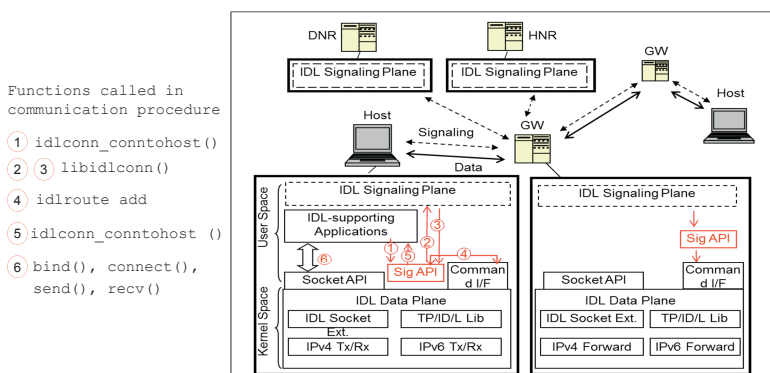


Fig. 8 Overview of software architecture

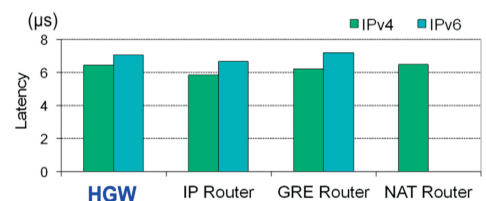


Fig. 9 Forwarding performance at intermediate nodes

mapping between host ID and locator is securely registered in host name registry and host is authenticated at the time of network access^[4]. In order to have reliability of the communication in reality, we implemented software for securely registering and deleting information of hosts and securely obtaining information of corresponding hosts via hierarchical trusted authentication mechanism. Here, HNR is the trusted anchor point of the hierarchical trusted authentication for secure registration, communication, and mobility of hosts. The proposed integrated security method enables secure communication in the network where the hosts frequently change their access networks.

- Seamless handover and resiliency. In heterogeneous communications between hosts each of which uses different network protocol each other, we demonstrated periodical changes of networks to which hosts connect. On condition that a host can connect multiple networks simultaneously, we verified

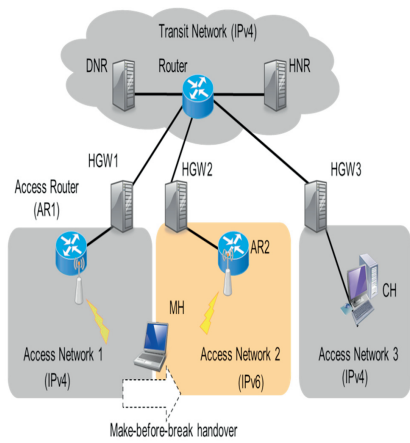


Fig. 10 An experimental network

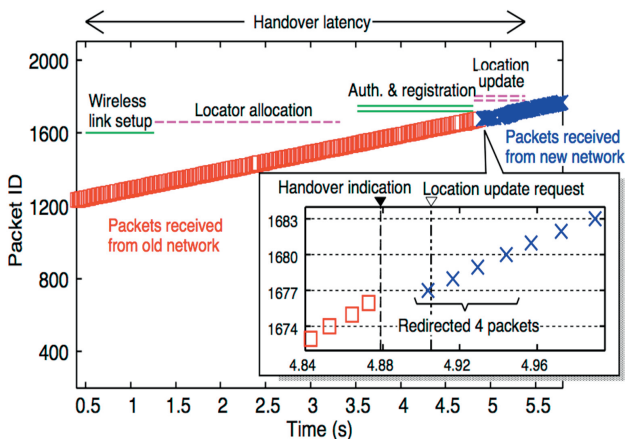


Fig. 11 Host mobility in 4 seconds with no data loss

the host successfully changes network for access in 4 seconds without data loss (see Figs. 10 and 11)^[10].

- Packaging HIMALIS software modules^[13]. Our HIMALIS software is configured as a package for dissemination. The package can be installed easily via GUI operation without solid knowledge of the protocol. The HIMALIS gateway in the installed HIMALIS edge network is connected to the Internet. For dissemination, we have installed and operated a set of registries like HNR and DNR in NICT. Thus the researchers outside NICT can use a HIMALIS testbed via Internet by installing the package to their networks (Fig. 12)^[14].
- We developed a mobile wireless sensor network where ID-based communication of HIMALIS is used for secure and seamless mobile communication. The sensors in the network are capable of 6LowPAN that has a feature of low-communication overhead. The network has high convenience and operability in the following points. At first, data generated at sensors

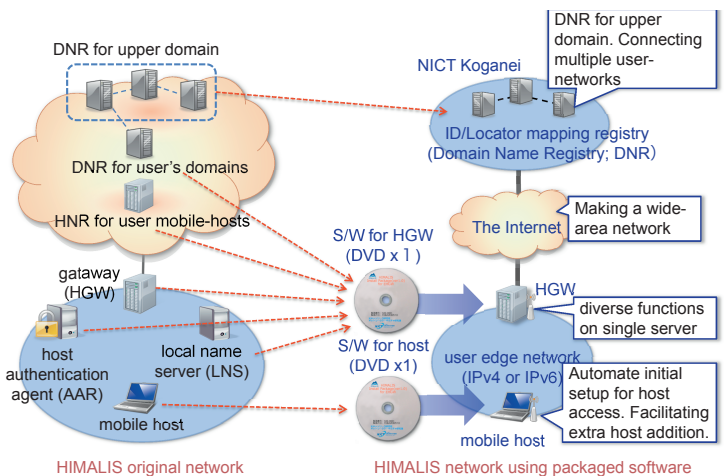


Fig. 12 HIMALIS testbed and use cases

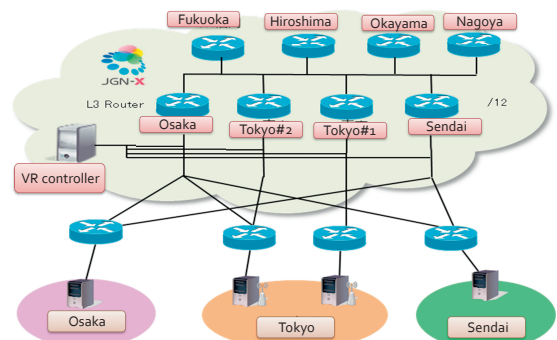


Fig. 13 A multi-homed HANA network built on JGN-X

can be sent to a sensor-data storage server (i.e., a sink server) that can communicate via IPv4 and IPv6 communication. Sensors can move to different 6LoWPAN networks from the home network. A remote server in IP networks can configure parameters of sensors^{[15]-[17]}. See 7-6 of this special issue in more detail.

- We proposed Hierarchical Automatic Number Assignment HANA for multi-homed networks that have different access lines to the different provider networks. Using JGN-X, we built a network that provides multiple routes to an end-to-end communication for proper hierarchical address assignment. A pair of two end nodes using a route can communicate in a different route even if a part of communication channels and node equipment fails (Fig. 13)^[18]. We also developed a link-state routing protocol HQLIP that is capable of 100 times faster route convergence compared to OSPF in an ideal condition by cooperation of HANA^[19]. We developed a network visualization system^[20]. By installing the system, we facilitated fault detections that are necessity feature for communication service. The HANA-based network built on JGN-X can allocate address spaces to the access networks where HIMALIS and

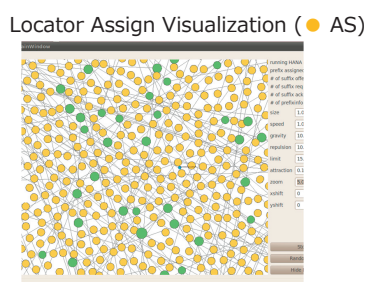


Fig. 14 Visualization of 10,000-AS verification

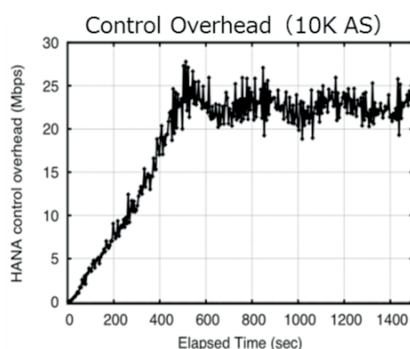


Fig. 15 Control overhead in HANA (25 Mbps control packets flows in 10,000 ASes)

a wireless mesh network with information sharing system NerveNet^{[21]-[23]}.

- We verified large-scale automatic address allocation by building a 10,000-node network (i.e., one-fourth size of the Internet) and applying HANA on large-scale emulation environment StarBED³^[5]. Figure 14 shows an example of visualization result of address allocation. Figure 15 shows a measured value of control overhead traffic by HANA. We observe that 25 Mbps control message flows among 10,000 networks. This is equivalent to several kbps for each network. Thus we conclude that HANA does not have negative impact the network. Till now, we have succeeded hierarchical and automatic address allocation to a 46,000-node network, namely, we have done the Internet-scale verification.
- We developed a HANA-capable 48×10 Gbps layer 3 switch. This is a rack-mountable size and the height is 1U (Fig. 16)^{[7][24]}. This switch leads the number of address configurations down to 1/100 for a 1,000 server-scale network. These L3 switches are being installed in the NICT enterprise network and an experimental network. Through the daily operation verification, we are modifying software for solving functional errors and stabilize control and management system of the HANA network.
- We combined HANA into SDN. We developed a unified IP address allocation for LAN switches and terminals for an enterprise network management. The developed system includes auto-generation function of maintenance management table for network administrator^[25]. This enterprise network reliefs network administrators from human-error, reduces operational times steeply, and enables existing

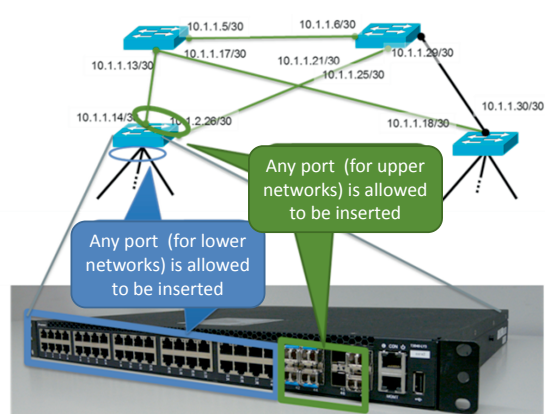


Fig. 16 Developed HANA-capable layer 3 switch and an example of address assignment

We have designed and implemented them as a form of the software and verified function, performance, and scalability by using JGN-X and StarBED. Currently, we are trying practical use of the software and collaborating application developers who have interest to our software. We pursue dissemination of enabling technologies and applied technology of the HIMALIS and HANA.

Acknowledgment

This study has been conducted with Dr. Kenji FUJIKAWA, Dr. Ved Prasad KAFLE, Dr. Yusuke FUKUSHIMA, Dr. Ruidong LI, Dr. Abu Hena Al MUKTADIR, Mr. Tomoji TOMURO, Mr. Yasunaga KOBARI, Mr. Hiroaki ABE, Dr. Hajime TAZAKI (currently the University of Tokyo), Dr. Sugang XU, and Dr. Yong JIN (currently Tokyo Institute of Technology). The author also thanks to collaborators in academia, industry, and NICT.

References

- MIC Press, "Situations of Accident of Telecommunication Services (FY2013)," Sept. 1, 2014. (in Japanese)
- Ved P. Kafle and Masugi Inoue, "HIMALIS: Heterogeneity inclusion and mobility adaptation through locator ID separation in new generation networks," *IEICE Transactions on Communications*, Vol.E93-B No.3, pp.478–489, March 2010.
- Ved P. Kafle, Hideki Otsuki, and Masugi Inoue, "An ID/locator split architecture for future networks," *IEEE Communications Magazine*, Vol.48 No.2, pp.138–144, Feb. 2010.
- Ved P. Kafle, Ruidong Li, Daisuke Inoue, and Hiroaki Harai, "Design and Implementation of Security for HIMALIS Architecture of Future Networks," *IEICE Transactions on Information and System* 2013, Vol.E96-D, No.2, pp.226–237, Feb. 2013.
- Kenji Fujikawa, Hajime Tazaki, and Hiroaki Harai, "Inter-AS Locator Allocation of Hierarchical Automatic Number Allocation in a 10,000-AS Network," *Proc. SAINT 2012*, July 2012.
- Kenji Fujikawa, Hiroaki Harai, and Masataka Ohta, "The Basic Procedures of Hierarchical Automatic Locator Number Allocation Protocol HANA," *Proc. Asia Workshop on Future Internet Technologies (AWFIT 2011)*, pp.124–131, Oct. 2011.
- NICT Press, "The World's First Implementation of Automatic Network Address Configuration Technique in Layer3 Switch," <http://www.nict.go.jp/>, June 3, 2014. (Official Version is Japanese).
- Yang Song, Lixin Gao, and Kenji Fujikawa, "Resilient Routing under Hierarchical Automatic Addressing," *IEEE Globecom 2011*, Dec. 2011.
- Ved P. Kafle, Hajime Tazaki, Tomoji Tomuro, Yasunaga Kobari, and Hiroaki Harai, "Implementation and Evaluation of ID/Locator Split-based HIMALIS Network Protocol Stack," *IEICE Technical Report (NS2012-153)*, Jan. 2013.
- Ved P. Kafle, Yusuke Fukushima, and Hiroaki Harai, "ID/Locator Split-Based Distributed Mobility Management Mechanism," *Springer Wireless Personal Communications*, Vol.76, Issue 4, pp.693–712, June 2014.
- V. P. Kafle, Y. Fukushima, and H. Harai, "New Mobility Paradigm with ID/Locator Split in the Future Internet," *IEEE Mobiworld Workshop 2014* (collocated with IEEE CCNC 2014), pp.365–371, Jan. 2014. (invited)
- Yusuke Fukushima, Ved P. Kafle, and Hiroaki Harai, "A middleware enabling ID-based communication for legacy socket applications." *IEICE Technical Report (IN2013-85)*, pp.49–54, Oct. 2013. (in Japanese).
- Kenji Fujikawa and Yusuke Fukushima, Ved P. Kafle, and Hiroaki Harai, "Development of Easy Install Package Software for ID/Locator Split architecture HIMALIS." *IEICE Technical Report (IN2013-206)*, pp.367–372, March 2014. (in Japanese).
- Yusuke Fukushima, Kenji Fujikawa, Ved P. Kafle, Hiroaki Abe, Yasunaga Kobari, and Hiroaki Harai, "Construction and Use of ID/Locator Split Architecture HIMALIS Testbed Network" *IEICE Technical Report (IN2014-80)*, pp.35–40, Oct. 2014. (in Japanese).
- Ved P. Kafle, Yusuke Fukushima, and Hiroaki Harai, "Dynamic Mobile Sensor Network Platform for ID-based Communication," *ITU Kaleidoscope 2014*, June 2014. [Received Best Paper Award].
- Ved P. Kafle, Yusuke Fukushima, and Hiroaki Harai, "Design and Implementation of Dynamic Mobile Sensor Network Platform for ID-Based Communications," *IEEE Communications Magazine Standards Supplement*, Vol.53, No.3, pp.48–57, March 2015.
- Yusuke Fukushima, Ved P. Kafle, and Hiroaki Harai, "ID-based Communication for M2M in Next-Generation Mobile Networks" *IEICE Technical Report (RCS2014-250)*, pp.177–182, Dec. 2014. (in Japanese).
- Kenji Fukukawa, Yasunaga Kobari, and Hiroaki Harai, "Construction of Mini Data Center and Wide Area Network using Hierarchical/Automatic Locator Number Allocation Protocol HANA " *IEICE Technical Report (IA2012-1)*, pp.1–5, June 2012. (in Japanese)
- Kenji Fujikawa, Motoyuki Ohmori, and Hiroaki Harai, "Stable Renumbering in Link-State Routing Protocol Network," *IEICE Technical Report (IA2014-44)*, pp.57–62, Nov. 2014.
- Kenji Fujikawa, Tsunemasa Hayashi, Hiroki Nakano, Yusuke Fukushima, and Hiroaki Harai, "A Management and Visualization System of Hierarchical/Automatic Locator Number Allocation Protocol HANA" *IEICE Technical Report (ICM2013-25)*, pp.13–18, Nov. 2013. (in Japanese).
- Masugi Inoue, Masaaki Ohnishi, Chao Peng, Ruidong Li, and Yasunori Owada, "NerveNet: A Regional Platform Network for Context-Aware Services with Sensors and Actuators," *IEICE Transactions on Communications*, Vol.E94-B, No.3, pp.618–629, March 2011. (invited)
- Yusuke Fukushima, Kenji Fujikawa, Ved P. Kafle, Hajime Tazaki, Yong Jin, and Hiroaki Harai, "Design and Implementation of ID/Locator Split Architecture for Improving Availability" *IEICE Technical Report (NS2013-12)*, pp.7–12, May 2013. (in Japanese).
- NICT Press, "Success in the World's First Automated Networking for Wide Area Network," <http://www.nict.go.jp/>, June 7, 2012. (Official version is Japanese)
- Kenji Fujikawa and Hiroaki Harai, "L3 Switch Implementation of Automatic Locator Number Allocation Protocol HANA" *IEICE Technical Report (NS2014-123)*, pp.109–114, Oct. 2014. (in Japanese)
- Kenji Fujikawa and Hiroaki Harai, "Applying HANA to SDN," *IEICE Network Virtualization Workshop*, July 2015. (in Japanese)
- Yong Jin, Kenji Fujikawa, Hiroaki Harai, and Masataka Ohta, "Secure Glue A Cache and Zone Transfer Considering Automatic Renumbering," accepted for publication in the proceedings of *IEEE COMPSAC 2015 (The 39th Annual International Computers, Software & Applications Conference)*, July 2015.
- S. Xu, K. Fujikawa, and H. Harai, "An Inter-AS Address Space (Re) Allocation Planning Scheme in Hierarchical and Automatic Number Allocation," *IEEE Globecom 2013*, Dec. 2013.



Hiroaki HARAI, Ph.D.

Director of Network Architecture Laboratory,
Photonic Network Research Institute/
Research Manager, New Generation Network
Laboratory, Network Research Headquarters
Network Architecture, Optical Network