第３部

Overlaying and Slicing with Internet's End-to-End Discipline based on the Practices IP version 6 Deployment in Japan

Hiroshi ESAKI、Ryo NAKAMURA

Abstract

This paper discuss and analyze the IPv6 deployment in Japan, from the view point of large scale multiple-stack layer 3 network development and deployment, focusing on the future network development. Since IPv6 network does not have compatibility with IPv4 network, it is considered the dual-stack operation is mandatory. However, when we analyze the IPv6 deployment in Japan, we realized that the integration of multiple single-stack networks using a tunneling with "locator" function works well, both in wired and wireless infrastructures. Also, the paper discuss the Internet is going to third wave with IoT and entering from CPS to Cyber-Twin and Cyber-First. To come up with this situation, the paper proposes the system design and implementation should be based on "Internet-by-Design", which preserve the key features the Internet. Finally, the practical examples of system design and implementation based on the "Internet-by-Design". These are smart building/campus to integrate different IoT systems and the "locator" and "identifier" separation via "tunneling" in IP layer for large scale multiple-stack layer 3 network development.

I　INTRODUCTION

The research, development and deployment of IPv6, which was called as IPng (IP next generation), has long history and practical experiences to us, regarding the technical and operational knowledge, which we can refer to for the future new system design, implementation and operation. The acceleration of IPv6 deployment may achieved by the introduction of IPv6 service by many hyper giants, such as Google, FaceBook or Apple. Although it would be said that it is because they looks future, there would be other reason why they started to use IPv6 in their service. It would be because of huge number of servers accommodated in their back yard. When the number of servers were small, the system using either private or global IPv4 address is fine. But, since the global IPv4 address at ARIN is exhausted in September 2015 and the number of physical and virtual servers, they need, is significantly increased, it seems that they must use IPv6 in their large/huge scale server clusters in their data centers.

Also, since their service includes and needs a lot of sliced virtual networks, the increase of number of IP addresses consumed by their serves must be accelerated.

Recently, the Internet seems to enter into the third wave, which is "IP for Everything" or "IoT, Internet of Things". The first wave with Web and the second wave with SNS was "IP for Everyone". This means that we need highly scalable Internet, which can accommodate huge number of end-nodes. Therefore, we need full deployment of IPv6 and high quality cyber security counter measures.

First, this paper discusses analyze the IPv6 deployment in Japan from the view point of large scale multiple-stack layer 3 network development and deployment. Second, the paper discusses the system design and implementation toward the third wave of the Internet and leads that it should be based on "Internet-by-Design", which preserves the key features the Internet. Finally, the practical examples of system design and implementation based on the "Internet-by-Design".

II IPV6 RESEARCH AND DEPLOPMENT IN WIDE PROJECT

WIDE (Widely Integrated Distributed Environment) Project, www.wide.ad.jp, is R&D consortium related with the Internet technologies established in 1988. WIDE project has established an IPv6 working group in 1994, according to the IETF decision on SIPP as a base of IPv6.

A. Works, the WIDE Project has achieved

In 1998, KAME Project[1] and TAHI Project[2] has been established to deliver open source of IPv6 protocol stack for BSD platform and it's conformance testing environment[3]. In 2000, USAGI (UniserSAl playgGound for IPv6) project[4], which has delivered IPv6 protocol stack for Linux platform has been established. Also, IPv6 Promotion Council[5] has been established in 2000 and Task Force on IPv4 address Exhaustion Japan[6] has been established in 2008, so as to identify the technical and business issues to deploy IPv6 in all the stake holders.

B. Experiences on IPv6 network operation

WIDE project has implemented and operated their own R&D live-testbed by themselves across the sites of member organizations, called "WIDE Internet" since it has been established. Since 1995, the WIDE 6bone has been operated and has internetworked with 6bone-JP and global 6bone in 1998. Figure 1 and figure 2 show the WIDE 6bone topology in 1995 and in 1997, respectively. These 6bone networks were overlay networks over IPv4using tunneling or single stack networks provided by layer 2 links.

Since 1999, IPv4 and IPv6 dual-stack environment has been implemented and operated at WIDE project camp networks (i.e.,LAN). In these camp networks, external connectivity was IPv6-only or IPv4/IPv6 multi-links (i.e., WAN). Based on these experiences, we has challenged to nation-wide broadband R&D IPv6 network development since 2000. In the first stage, we have developed sliced networks. using ATM and MPLS. The sliced networks were 6bone, Mbone and Qbone. Figure 3 shows the overview of DV(Digital Video) multicasting using PIM-SM over it's layer 2 network. This means that PC routers had installed over layer 2 network, since 6bone and Mbone had not been integrated, i.e., this network was not overlay network over IPv6 network, since there were no commercial routers with IPv6 PIM-SM.

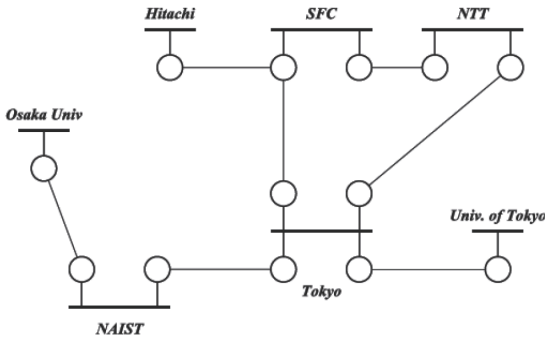


Fig.1 WIDE 6bone in 1995

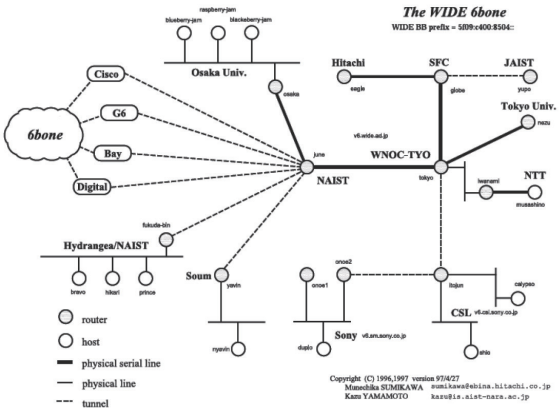


Fig.2 WIDE 6bone in 1997

＊1 http://www.kame.net/
＊2 http://www.tahi.org/
＊3 TAHI project leaded to"IPv6 Ready Logo Program", run by IPv6 Forum.
＊4 http://www.linux-ipv6.org/
＊5 http://www.v6pc.jp/en/index.phtml
＊6 http://www.kokatsu.jp/blog/ipv4/en/

Since 2004, we started IPv6 and IPv4 dual-stack operation in the backbone. Then, we has experienced and realized large operational overhead due to dual-stack operation, from the view points of trouble shooting, system reconfiguration, and of implementation of cyber security countermeasures.

Based on our dual-stack operation in the WIDE Internet, we decided to implement and operate IPv6-only network at WIDE camp in 2012, so as to realize technical issues for IPv6 only single stack operation. The best practices in this experimental operation had been reported in [32].

Our implementation and operational experiences have shared with professional commercial operators in Japan (and IETF community), for their network design and operation. For example, as we discuss in the following sections, the single protocol stack network over single stack network is the direction that most of emerging networks adopts. These are the view points of trouble shooting, reconfiguration or cyber security operation.

III IPV6 DEPLOYMENT IN JAPAN

This section describes the current status of IPv6 deployment in Japan, focusing on network providers and some wide area private IPv6 network development.

A. Wired Public Provider

Figure 4 shows the development status of IPv6 users in NGN (Next Generation Network) subscribers/customers, which is the largest broadband access network operated by NTT group. It was 0.8% (67,000 subscribers) in December 2012, but it is **30.5% (5,797,000 subscribers) in March 2017.**

KDDI, who is ranked as third with 26.78% of deployment ratio as of February 20, 2017 at IPv6 Launch site[7], completed 100% of IPv6 deployment in it's optical fiber access network in December of 2012. Since KDDI is now preparing to provide IPv6 connectivity as a default service in their cellular phone network since around summer of 2017, the number of IPv6 deployment ratio will be further improved in 2017.

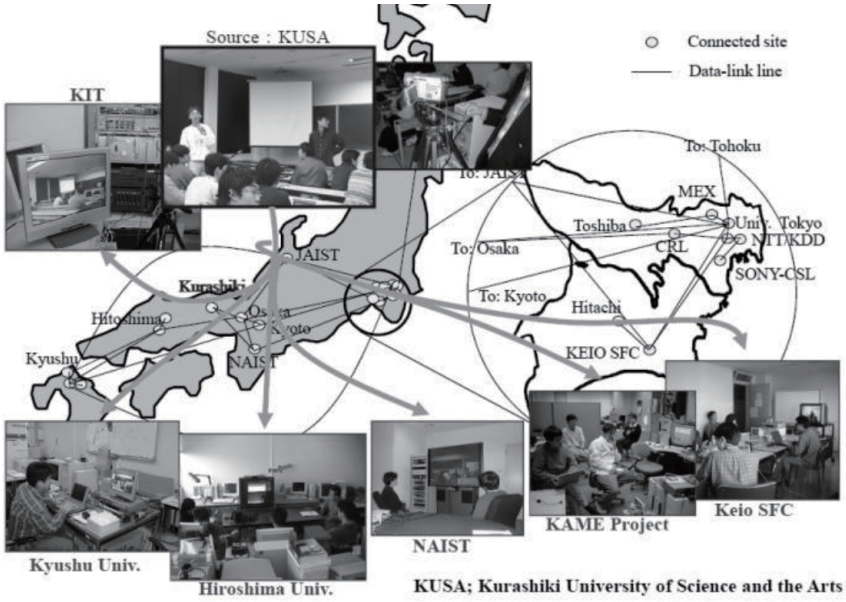


Fig.3 Digital Video multicast with PIM-SM over JB 6bone

＊7 http://www.worldipv6launch.org/measurements/

The other wired ISPs are also progressing IPv6 service deployment, as a default service to all the subscriber/customer. When ISP provides IPv6 service over NGN access network, two methods has been defined by Japanese government. One is IPoE and the other is PPPOE. [33, 34]

(1) IPoE (IP over Ethernet [33])

Currently, we have three operating VNE (Virtual Network Enabler) providers and other three providers have registered for future VNE providers. VNE has global IPv6 address to allocate to subscribers and transit the IPv6 packet to any destination of IPv6 global internet. VNE is a kind of aggregator, i.e., VNE provides IPv6 address prefix(s) from his IPv6 address pool to it's partner ISPs. The partner ISP of VNE runs it's router to accommodate it's subscribers in NGN. Here, the IPv6 packets destined to the interface (both among any NGN's closed IPv6 address and any VNE's open IPv6 address) in the NGN access network can be transferred within the NGN without being transferred to the corresponding ISP's router.

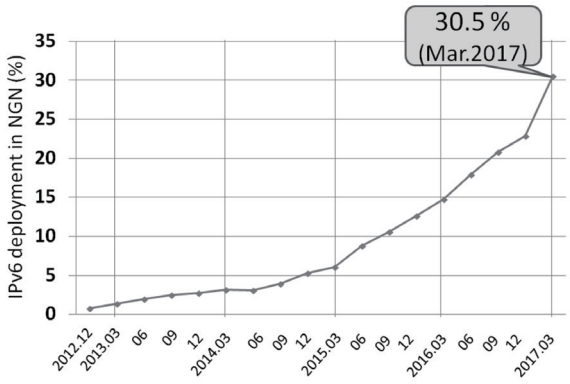
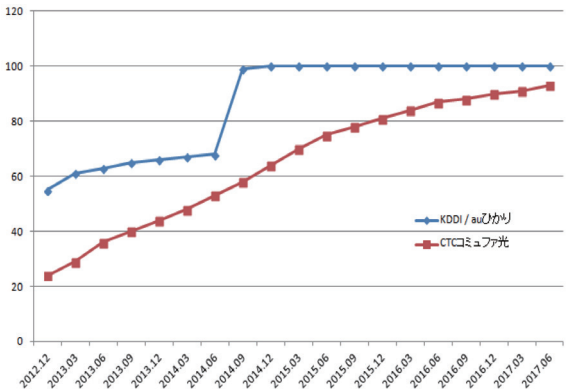


Fig.4 IPv6 deployment in NGN



(2) PPPoE (Point to Point Protocol over Ethernet [34])

Most of IPv4 service provided by ISPs in Japan over NTT group's access network, including NGN, uses PPPoE.

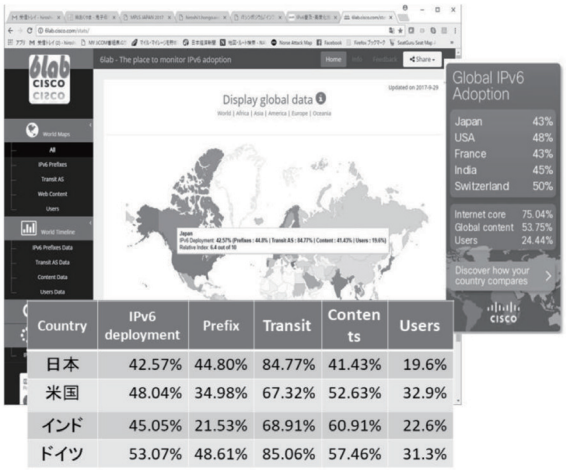
rank	name	ASN	IPv6
1	KDDI	2516	36.85 %
2	SoftBank BB	17676	27.54 %
3	OCN / plala	4713	23.14 %
4	So-net	2527	37.01 %
5	BIGLOBE	2518	32.65 %
6	CTC	18126	45.70 %
7	TOKAI	10010	24.62 %
8	IIJ	2497	11.7 %
9	NTT DoCoMo	9605	2.02 %
10	@nifty	2510	10.62 %
11	iTSCOM	9365	10.83 %
12	Sony Global Sol.	9619	99.63 %

rank	name	ASN	IPv6
13	Star cat	17529	15.23 %
14	Bit-Drive	9600	12.03 %
15	VECTANT	2519	0.97 %
16	K-Opticom	17511	0.39 %
17	SINET	2907	1.66 %
18	TDNC	9354	2.08 %
19	SuperCSI	2506	42.78 %
20	Keio Univ.	38635	29.33 %
21	FreeBit	4691,10013	0.48 %
22	J:Com	9824	0.04 %
23	UCOM	17506	0.06 %

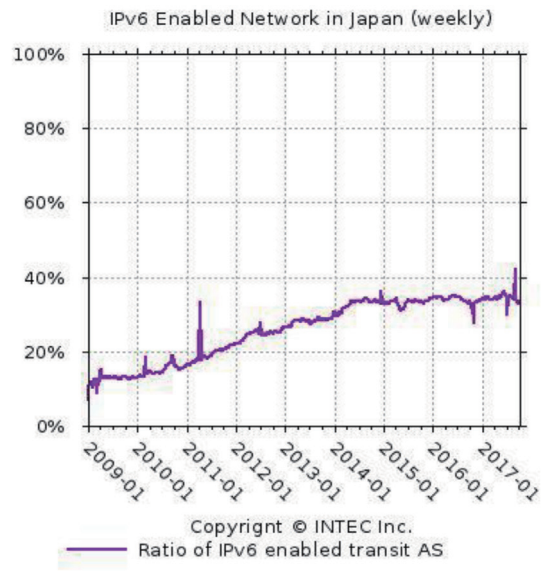
In the NGN, the closed IPv6 address allocated to NTT East and West, which is worked as “locator” discussed in section V-C, is used for the transmission of IPv4 packet[8] by packet encapsulation (IPv4 over IPv6). For IPv6 service provided by ISPs, the same architecture and protocol stack is applied to. In this case, it is global/open IPv6 packet over global/closed IPv6 platform. At the subscriber side, the home router must establish two PPP sessions over NGN’s closed IPv6 network. One is for IPv4, and the other is for IPv6.

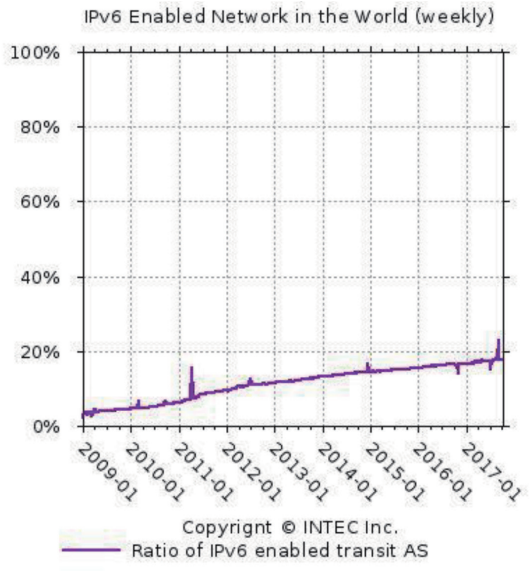
When NTT East and West started their NGN service, NGN platform was a wide-area “closed” IPv6 network for the transmission of global IPv4 packets, i.e., IPv6 closed address corresponds to “locator”[35]. With the PPPoE, the function of NGN platform is the same as IPv4, i.e., closed IPv6 address in the NGN is the “locator”, whose function is the transportation of global IPv6 packets between the routers interconnecting to global IPv6 networks and end-nodes in NGN. On the contrary, with the IPoE, the function of NGN platform is a part of global IPv6 networks. The open/global IPv6 address allocated to end-node is both for “identifier”, discussed in section V-C, of end-node and for “locator” to reach to the end-node. Here, SoftBank group is migrating IPv4 PPPoE service over NGN platform to IPv4 tunneling over IPoE. With this migration, VNE’s open/global IPv6 network in NGN run by BBIX works as “locator” for the transmission of IPv4 packet over open/global IPv6 network by IP packet encapsulation.

This architectural consideration, regarding “locator” and “identifier”, is related with the discussion in section V, where we discuss the system design in the emerging future IoT and cloud networks.



Country	IPv6 deployment	Prefix	Transit	Contents	Users
日本	42.57%	44.80%	84.77%	41.43%	19.6%
米国	48.04%	34.98%	67.32%	52.63%	32.9%
インド	45.05%	21.53%	68.91%	60.91%	22.6%
ドイツ	53.07%	48.61%	85.06%	57.46%	31.3%





＊8 IPv4 address in the IPv4 packet, encapsulated in IPv6 packet, is “identifier” discussed in section V-C.

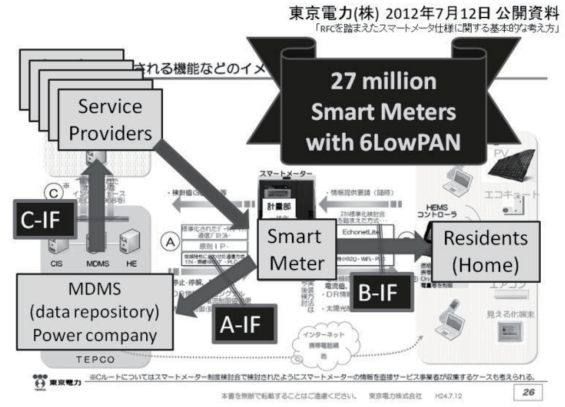
B. Wireless Public Provider

Japan has three major cellular phone network providers, which are NTT DoCoMo, KDDI/au and SoftBank. It has been committed by these three major cellular phone network providers so that they will start the IPv6 and IPv4 dual-stack service as a default since around summer of 2017. This means that it is expected that the IPv6 deployment measure provided by IPv6 Launch site for these three Japanese service providers will be soon and significantly improved since around summer of 2017.

The protocol stack of LTE system is open/global IP (=”identifier”) over closed/private IPv4 (=”locator”) platform from all the end-terminal to very small number of boarder gateway router to the Internet. The network topology is very simple so that routing of closed/private IPv4 platform can be also very simple. This means that the function of closed/private IPv4 network is “locator”, which is for open/global IP packet transmission between the boarder gateway router(s) and end-terminals in their access networks. This can be realized that a physically and logically large scale single layer 3 segment defined by “identifier” is built on a kind of large scale data-link, which is built by closed/private “locator” IPv4 address space.

C. Smart Meter Network Development in Japan

The development and deployment of smart meter infrastructure is one of top agenda in Japan to build a smart city. All the major electrical power utility companies in Japan have been working on the smart meter system for residential home’s energy management and control. Because of large number of smart meters in their subscriber region, many utility companies adopt IPv6. Here, two typical and interesting systems are picked up. One is smart meter system based on 6LowPAN[36] by TEPCO(Tokyo Electric Power Co.Inc.), and the other system is based on IPv6 LTE by KEPCO(Kyushu Electric Power Co.Inc.). When we observe these two networks, it is obviously the smart meter system based on IPv6 LTE works well and provided higher communication quality, than those of the system based on 6LowPAN.



This experience may show that, for large scale wide-area IoT system, IPv6 LTE system works fine. Also, we have consideration on cyber security measures for both systems. Since the IPv6 LTE system has enough experiences and implementation for higher cyber security quality compared to 6LowPAN system, IPv6 LTE system is also better from the cyber security point of view.

IV WHERE THE INTERNET IS GOING ?

A. IoT as the Third Wave of the Internet

The Internet is now entering into the third wave, which is IoT, after the first wave (Web) and the second wave (SNS). The first and the second waves are “IP for Everyone”. The third wave is “IP for Everything”, which is commonly called as IoT (Internet of Things). In the IoT, we may think all the non-IT/ICT industries will be built in and connected to the Internet. However, due to various reasons, most of IoT business players love “Silo” or “Stove-and-Pipe” system and business structure, so as to lock-on the customers to their proprietary system. So, the following three points would be the current risk for the third wave of the Internet.

1. Fragmentation of the Internet by IoT
2. Lack of "Trust", such as cyber security
3. Lack of interoperability, especially in the cloud platform

They, IoT and legacy industries' people, may frequently say that their system does not need enough cyber security counter measures, since their system does not connect with the Internet and {sometimes} their system does not use open technology but use proprietary technologies. This is really the serious risk. Or, even they have the IP connectivity, they may not have upper layer interoperability.

This is especially for cloud system or for SDN/NFV, in these days.
This means that we must design and implement the emerging system based on; (1) connected to the Internet is premise, (2) Security-by-Design for trustworthy, and (3) Interoperability-by-Design.

B. Toward "Cyber-First"

Now, what system features will the third wave introduce to our society? The third wave may have the following three sub-waves, but each sub-wave has huge social impacts.

1. CPS; Cyber supports Physical, i.e., "Physical-First"
2. "Cyber-Twin"; Cyber emulates/copies Physical
3. "Cyber-First"; Cyber designs Physical

We could realize that SDN(Software Defined Network) or SDI(Software Defined Infrastructure) is an implementation of "Cyber-First" in the networks. In the "Cyber-First", the physical system is designed based on full simulation and evaluation in the Cyber system. Especially, since virtualization of hardware platform become reality with sufficient performance and quality, cyber objects have been getting independency from the physical objects. This leads that we can design the system in cyber space with cyber objects (including virtual physical objects), then allocate them on physical objects and easy to migrate cyber objects on physical objects.

V ARCHITECTURAL DERECTION WITH INTERNET-BY-DESIGN

A. Internet-by-Design

The following eight features can identified as the architectural and operational features of the Internet. And, we should design, implement and operate the emerging system based on these features, as discussed below. This paper calls this as "Internet-by-Design".

1. Global

Nation or government is one of multi-stake holder, and the technologies and rule/regulation must be based on global consistency, think global and act local. The following is the joint declaration[9] by G7 ICT Summit held at Ise-Shima, Japan in 2016.

> *We strongly support an accessible, open interoperable, reliable and secure cyberspace as one essential foundation for economic growth and prosperity.*
>
> *We must commit to facilitate the free flow of information to ensure openness, transparency and freedom of the Internet, and fair and equal access to the cyber space for all actors of digital economy while respecting privacy and data protection, as well as cyber security.*

2. Unique system on the Earth

Connected to the Internet is the given premise. This means we need "Security-by-Design" and "Privacy-by-Design". Also, we must avoid silo (or stove & pipe) model, and should implement horizontal cooperation model, such as data-centric or open-data model, as shown in figure 5.

Especially in the IoT and Big-data system;
(i) User can access/use the lawful data with the same way.
(ii) User can connect/put the sensor, which does not harm the network, with their choice, with the same way.

＊9 http://www.mofa.go.jp/files/000160266.pdf

(iii) User can provide service using the data in data repository.

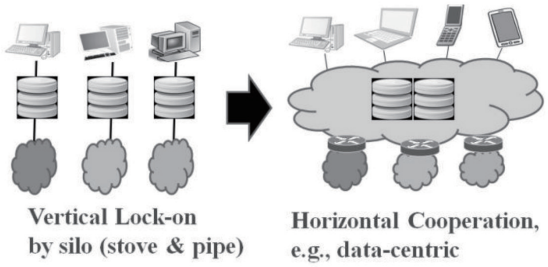


Fig.5 System models

3. Provision of alternatives

We should not too optimize the systems, intentionally, so as to preserve the chance to put alternative solution.

4. Respects running code and system

Proposal without practical/actual implementation and operation can not believe for real deployment.

5. Best effort

We should not have some particular quality object, in general, in order to maintain the efforts to improve the quality and functionalities, i.e., avoiding spoil. Also, we serious earthquake in Japan on 3.11 (in 2017) prove that, since daily service is “best-effort” and work with abnormal situation in the Internet, the Internet could continue, even with low quality, their services even in abnormal and emergency situation.

6. Transparency and end-to-end principle

Since the network is transparent and simple (sometimes said as stupid), (1) we can share data and knowledge without any filtering, (2) we can solve complicated issues by the “End-node (including server)”, not by “network (=switch or router)”. The later feature is of essential for the proposed direction discussed in Section V-C.

7. Social eco-system

The Internet is built and operated by “One for All, All for One”, i.e., eco-system. In order to run the eco-system, the capability of interaction and cooperation is mandatory. This is “interoperability”. Even if we do not need interoperability today, we should have interoperability with other system for the future opportunity of cooperation. And, therefore, we need cyber security implementation.

8. Independency, autonomous and distributed

While maintaining the interoperability, we must intentionally preserve the diversity for technologies and system operation, so as to ensure the provision of alternative. This is related with the survivability of the system in the future.

The following subsections show practical example of system design and implementation based on the ”Internet-by-Design” described above.

B. Smart Building/Campus based on IEEE 1888 [37]

Smart building or smart campus is an application of IoT with LAN and MAN platform for non-IT/ICT industry. The devices and networks used in buildings or in campus have (1) large product lifecycle, which would be decade(s) years, (2) isolated silo system, i.e., exclusive vertical lock-on, with

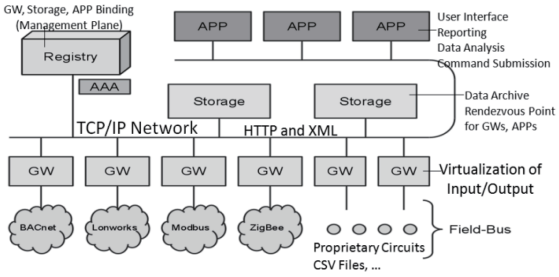


Fig.6 Architectural overview of IEEE1888

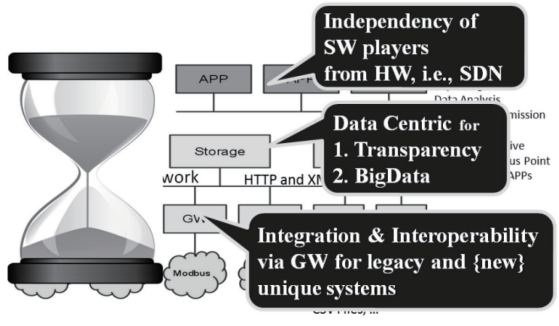


Fig.7 Three layer structure of IEEE1888

non-open proprietary technologies and (3) no cyber security consideration. In order to integrate and upgrade the system, we standardized IEEE1888, which is open facility network architecture and protocol based on internet protocol platform, in 2011 and approved by ISO/IEC as ISO/IEC/IEEE 18880 in 2015 [38]. The architecture overview of IEEE1888 is shown in figure 6 and 7.

As shown in figure 8, IEEE1888 has three layer structure. The bottom layer accommodates various legacy local bus and emerging bus via GW(Gate Way), i.e., allowing alternative or new technologies and sub-systems. And, GW is the boundary between TCP/IP based system and non-TCP/IP system. The middle layer is shared common data repository accommodating various bus in the bottom layer, and providing all the data transparently to any application in top layer. The top layer is application players, who do not need to care about hardware difference, since the abstracted virtual interface and object definition is provided to them via the common API among objects in the three layers. Also, IEEE1888.3 defines cyber security function, i.e., authentication and encryption.

Using the IEEE1888, the cloud based wide-area smart building and campus have been implemented and been in operation. For example in the University of Tokyo, more than 20% energy saving for campus level and 30% energy saving for R&E building have been achieved, with multi-vendor environment (more than 20 vendors), while continuous introduction of new functions (software in the top layer) and components (hardware in the bottom layer).

The IEEE1888 with IEEE1888.3 will be introduced to new data center run by Sakura Internet Inc., www.sakura.ad.jp, in the summer of 2017. The reason of the introduction of IEEE1888 for facility management and control in their data center is (1) cyber security for facilities in the data center, and (2) integration of separated facility systems and (3) application development, such as a big-data analysis with artificial intelligence technologies, “by themselves”.

C. Separation of Idetifire and Locator

Separation of identifier and locator has been discussed, e.g., LISP as RFC 6830 [35], for long time [39, 40]. And, there has been a lot of discussion and standardized architectures and protocols for packet encapsulation and for network slicing such as VPN(Virtual Private Networking). For example, MPLS can be realized as an encapsulation and aggregation shim layer inserted between IP (layer 3) and data-link (layer 2), so as to provide a “locator” function using MPLS header, which is generated from the combination of information in any layer .

We have discussed IPv6 service deployment in Section II. Since IPv4 and IPv6 do have different address spaces, the dual-stack operation has large operational overhead especially in the large scale networks. Therefore, we could realize that, in the large scale carrier-grade networks discussed in Section II, the single-stack operation with overlay single-stack slicing are applied to. Underlay IP network works as a “locator” for overlay IP network and as for both “locator” and “identifier” for itself. For example, in the SoftBank’s BBIX VNE system, the NGN’s global/open IPv6 address to communicate with IPv6 node in the IPv6 global network is used both for “identifier” and “locator”, and the NGN’s global/open IPv6 address is used for “locator” for the transmission of IPv4 packet to the entry point to IPv6 global network.

Now, we propose and proposed the separation of “identifier” and “locator” in IP layer with the optimization method to achieve sufficient packet throughput at the end-node [41] [42],. The proposed architecture and implementation does not require any modification nor introduction of new functionality to expensive routers, which is not easy to replace nor modify. On the other hand, since the proposed mechanism can be achieved only by the end-nodes, the deployment difficulty of the proposed system should be far less. This could be said that “identifier” function and “locator function” in the IP layer is separated by “tunnel”. The detailed implementation is described in [41, 42].

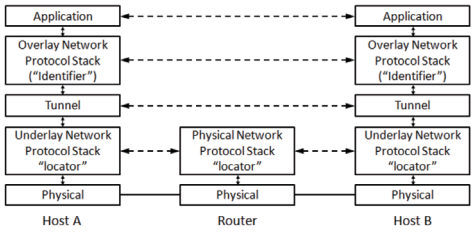


Fig.8 Separation of "identifier" and "locator" via "tunnel"

Finally, the internetworking between the different address families needs an address translator. We should identify the translator function is installed in the end-node, rather than in network node, i.e.., router. In order to reduce the processing overhead of translation function, translator node should be located at of edge as possible. Then, we can use single-stack networks, even we have multiple single-stack slicing networks.

VI CONCLUSION

This paper discuss and analyze the IPv6 deployment in Japan, from the view point of large scale multiple-stack layer 3 network development and deployment, focusing on the future network development from CPS to Cyber-Twin and Cyber-First. Then, we realized that the integration of multiple single-stack networks using a tunneling with "locator" function works well, both in wired and wireless infrastructures. Toward the emerging future IoT networks, the paper proposes that the system design and implementation should be based on "Internet-by-Design". Finally, the practical examples of system design are shown; one is smart building/campus to integrate different IoT systems and the other is "locator" and "identifier" separation via "tunneling" in IP layer for large scale multiple-stack layer 3 network development.

Appendix

REPORTING THE CURRENT STATUS OF WORLD WIDE IPV6 DEPLOYMENT AND PROGRESS TO ITOJUN
18:30-20:30, 27th November 2017

It has been ten years since Junichiro Hagino known as "itojun" left us. He foresaw the potential possibilities of IPv6 from the beginning, promoted standardization by serving as a chair of the v6ops (IPv6 Operations) working group at IETF, and as an active member of the IAB (Internet Architecture Board). He made great efforts to help people all over the world advance towards the same goal, leading not only technical topics, but also the direction of the community as a whole. At the same time, he implemented the IPv6 reference code on his own as a core member of the KAME project, and actively integrated the results into the BSD families, creating the foundation for current most IPv6 implementations.

IPv6 received the support of many equipment vendors and Internet operators, and is deployed on the global network. IPv6 support of client devices has also expanded, and general users are moving to the world of IPv6 naturally without noticing it.

In this event, we would like to report to itojun, who loved IPv6, on the achievements of the work we took over after he left, and our hopes for future development together.

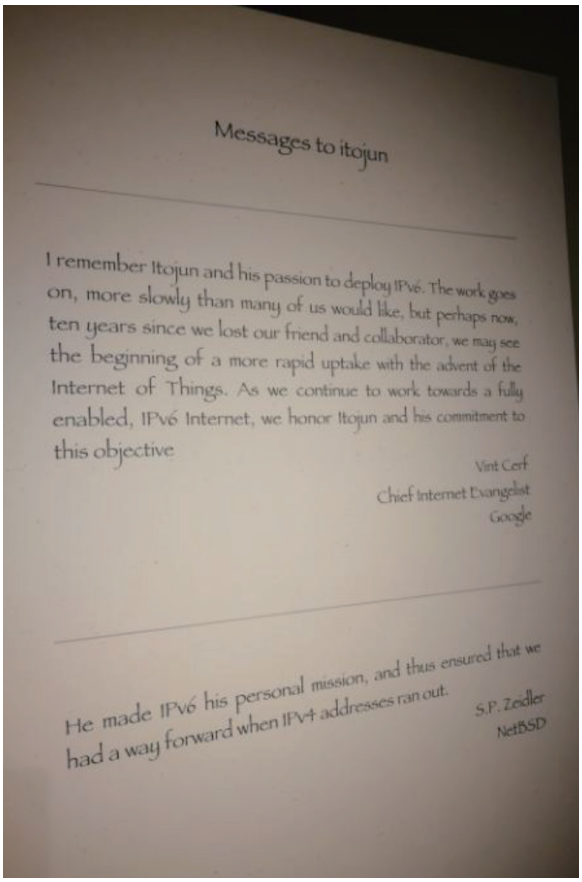
Organizers

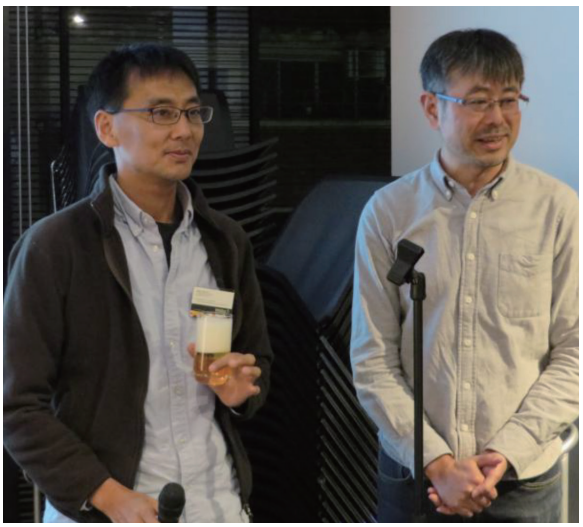
Chair :

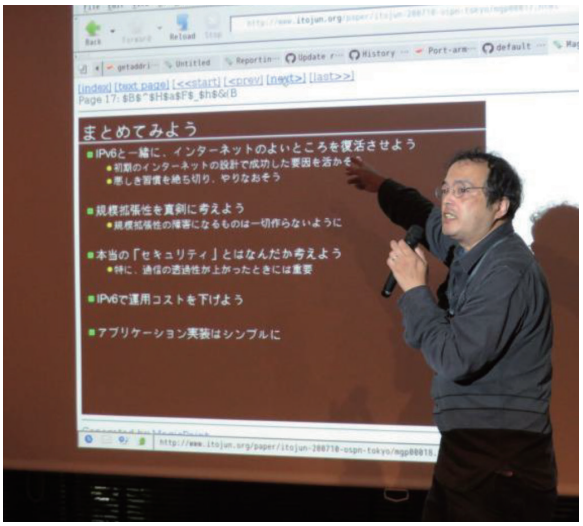
Jun Murai

Organizers (in alphabetical order) :

Hiroshi Esaki, Hiroshi Fujiwara, Shigeki Goto, Tatsuya Jinmei, Osamu Nakamura, Koichi Suzuki, Kazumasa Utashiro, Kazuhiko Yamamoto, Tatsuya Yamashita






まとめてみよう
IPv6と一緒に，インターネットのよいところを復活させよう
初期のインターネットの設計で成功した要因を活かそう
悪しき習慣を絶ち切り，やりなおそう
規模拡張性を真剣に考えよう
規模拡張性の障害になるものは一切作らないように
本当の「セキュリティ」とはなんだか考えよう
特に，通信の透過性が上がったときには重要
IPv6で運用コストを下げよう
アプリケーション実装はシンプルに