Stateless Mapping and Multiplexing of IPv4 Addresses in Migration to IPv6 Internet

Yuncheng Zhu Maoke Chen Hong Zhang Xing Li
Network Research Center, Tsinghua University, Beijing 100084, P.R.China
{zhu,cmk,neilzh}@ns.6test.edu.cn xing@cernet.edu.cn

***Abstract*—IPv4 addresses will be soon exhausted while more and more users are pouring into the Internet. It is commonly understood that IPv6 is definitely necessary for connecting new users, but lack of resources in current IPv6 world discourages providers from developing their IPv6 market. Address and packet translations are considered necessary for Internet content resources migrating from IPv4 world to IPv6, but up to now there is no one translation approach scalable enough. We understand the scalability is essentially restricted by the states in routing and those in translation, and accordingly we propose a novel architecture, called IVI, where IPv6 subscriber addresses are mapped from IPv4 with a certain rule that keeps IPv6 routing table well aggregated and makes translation stateless at all. Moreover, we introduce multiplex into the mapping to accommodate the trend of IPv4 addresses being more and more precious. Analysis based on real traffic data proves that the proposed approach is feasible to be deployed by an ISP for its customer networks and the potential degradation in end-to-end performance is acceptable.**

I. Introduction

Exhaustion of IPv4 addresses endangers the further development of the Internet. People have to start talking about IPv4 address trading and there have already been some black markets[1]. It is commonly admitted that IPv6 [1] must be deployed for connecting more users and applications in the future, but due to the lack of content resources in IPv6 Internet, providers are very unlike to make the migration for their services even up to current date when vendors have produced high quality IPv6 equipments. Resources and contents concentrating in IPv4 Internet seriously degrade the attractiveness of IPv6, compressing the room for the transition period before the complete exhaustion of the IPv4 address space.

Based on the well-known network effect theory in economics, research on new protocol adoption has concluded that well-designed translators are very crucial to the migration. Lack of cross-protocol accessibility will form an unacceptable gap in the curve of overall usability of the whole Internet [2], which must stop the process of transition. This illuminates that translator-based approaches, rather than dual stack or tunneling with encapsulation, is indispensable for the transition.

Unfortunately, all existing translators, such as the Network Address Translator (NAT) [3] for private-public address translation to the Network Address Translator - Protocol Translator (NAT-PT) [4] for IPv4-IPv6 translation, are not considered suitable for large-scale deployment and often fall into sharp critics [5]. The essential drawback of the legacy translation mechanisms lies in adding extra states for recording address correspondence in the translators.

Such network address and protocol translation mechanisms undermine the end-to-end address transparency, which have been noticed to have severe effect on deployment of end-to-end security, peer-to-peer applications protocols and so forth [6][7]. More severely, states add tight dependence upon certain nodes to end-to-end paths, making the network less robust and disabling routing-based traffic engineering in multi-homed environment.

Indeed, there was a stateless approach among the IPv4/IPv6 transition mechanisms – Stateless IP/ICMP Translation Algorithm (SIIT) [8]. Designed in the early years of IPv6, SIIT fits the requirement of a host just being migrated from IPv4 to IPv6 and willing to have communications with previous peers. SIIT cannot be used in the case where there are not so many IPv4 addresses able to be assigned to the new users, and it is not scalable for global deployment, with the risk of mapped IPv4 routes violating good aggregation nature of the global IPv6 route table. There are also some other approaches, like Network Address and Protocol Translator [9]. However, none of all existing techniques have contributed to mitigate the urgent exhaustion of IPv4 address space.

The idea of this paper is based on the assert that transition will be successful if and only if a scalable translation facility is deployed between two protocol family of the Internet, during the whole period of resource migrating from the old one to the new. Further on, we understand the key to scalability is minimizing states in both routers and translators, and therefore we propose a new, rule-based address mapping scheme. Moreover, it has to be one-to-many rather than one-to-one correspondence because IPv4 addresses that could be used for the stateless mapping will be less and less during the period of the transition. Fortunately, end hosts only use a small part of ports for regular communications. This fact enables a way of mapping one IPv4 address to several IPv6 counterparts without a significant cost. We'll explain the details of the architecture and evaluate its feasibility and potential cost through analysis on real data. For convenience, we call this proposed architecture as IVI[2], throughout the rest parts of the paper.

[1]See C. Marsan, "Will there be an IP address black market?", http://www.networkworld.com/news/2008/021408-ipv4-iana-conrad.html

[2]"IVI" is a combination of the two protocol version numbers, where IV stands for IPv4 and VI stands for IPv6.

Authorized licensed use limited to: UCLA Library. Downloaded on August 04,2023 at 18:46:48 UTC from IEEE Xplore. Restrictions apply.
978-1-4244-2324-8/08/$25.00 © 2008 IEEE.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2008 proceedings.
Juniper Exhibit 1023, page 1

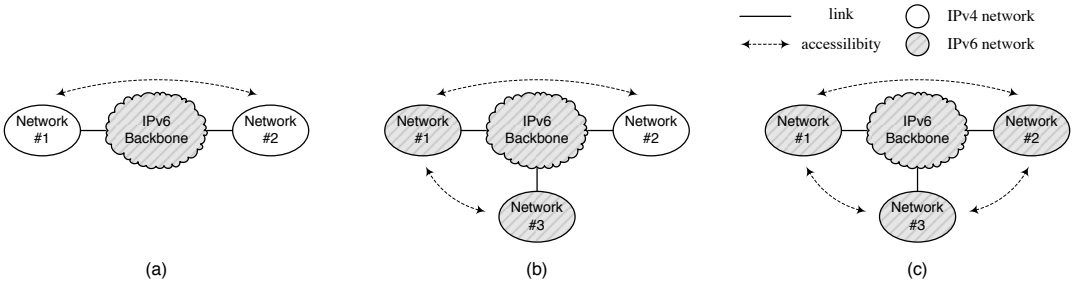

Fig. 1. The process of a reasonable transition.

We'll firstly describe a transition model that enabled by our idea of stateless multiplexing IPv4 addresses for IPv6 hosts and then, with real traffic observation, propose the idea of limiting port usage to support the multiplexing. In Sect III, we illustrate the major principles and components of the IVI architecture. Then, in Sect. IV, advantages and trade-offs of the architecture are evaluated in analyzing real traffic data, followed by the concluding remarks.

II. Transition Model and Basic Idea

Our heuristic model for the IPv4/IPv6 transition comes from a simple understanding that people will be happy to change their service protocol if the change doesn't hurt existing communication ability but extend it.

A. Outline of the Transition Process

As ever mentioned, economic modeling of new protocol adoption has concluded that either dual stack or encapsulation cannot, but only translation can complete the task of transition. Based on this understanding, it is a reasonable assumption that the transition can be done firstly in backbone infrastructure and then in applications and the application-related entities: customer networks and end hosts. The scenario is depicted in Fig. 1(a), where the core network is IPv6-ready. If the translation facilities have ever deployed at the boundary of the IPv6-ready core and IPv4 customer networks, it is pretty good to run the core without IPv4 at all. Actually, some other research has shown that the IPv6-only core is advantageous to IPv4 or dual-stack core even for current customer networks, in the term of path diversity in multi-homed environment [10].

With the help of the translators, some networks (e.g. "network #1" in Fig. 1(b)) may migrate into IPv6 first. Translators enable network #1 having access to and being accessible from other networks that are still not upgraded to IPv6, and any host in network #1 definitely has accessibility to new IPv6 peer ends. In this second stage, stateless address mapping keeps the temporal uniqueness of IP address, i.e. from IPv4-only to dual-stack and then IPv6-only, a host in network #1 is always identified as a unique peer at the view of an IPv4 host in, for example, network #2, and the latter does not need to be aware of its change of protocol stack. This temporal uniqueness ensures the ever existing accesses in IPv4 Internet will not be lost.

Network #2 cannot have access to network #3, a natively established pure IPv6 network, if the network #3 doesn't have originally used IPv4 address and its provider assign no IPv4 but only IPv6 addresses to it. However, as long as network #2 is migrated to IPv6, as is depicted in Fig.1(c), it may also do the same as network #1 has ever done: keeps communication with the remaining IPv4 world, and extends communication in new, pure IPv6 world as well.

However, the temporal uniqueness faces a challenge of scale. When network #1 transited from IPv4 to IPv6, its host number must be increased, and every new host may require connectivity to IPv4 resources. Therefore, a one-to-one mapping at the address translation doesn't fit the needs. NAT-PT is essentially a one-to-many mapping mechanism but it loses statelessness completely. We'd like to make a tradeoff that uniqueness requirement is relaxed to some extent but the statelessness is rigorously maintained. The key to achieve such a tradeoff is multiplexing precious IPv4 addresses by limiting port usage for end hosts.

B. Idea of Multiplexing Address by Limiting Ports

Intuitively, a host only use a small part of port numbers to connect with others. Therefore, if we limit the port that one host can use, then multiple hosts in an IPv6 network can share a same IPv4 address in the stateless translation and partition the port space corresponding to it without overlap. Note that the method of informing a host which ports it can use must be also stateless. It is reasonable to code both the shared IPv4 address and the exclusive port range into an IPv6 address for a host ever migrated. For example, if each host is offered a port range of 256 ports, one IPv4 address could be multiplexed to 256 IPv6 hosts. In fact, some mainstream operating systems have limited the maximum number of concurrent TCP connections of a host.[3]

This intuition could be verified by an observation of the real traffic. Fig. 2 is made through a one-hour traffic collected at the edge of a campus network[4]. It can be inferred from the concurrent flow number distribution that offering an end host 256 ports could satisfy more than 99% of all user sessions. If

[3]For example, Windows XP allow up to 100 concurrent connections and up to 10 unestablished connections as default. See http://support.microsoft.com/kb/158474.

[4]Data source: http://www.net-glyph.org/data-platform.shtml.gb

Authorized licensed use limited to: UCLA Library. Downloaded on August 04,2023 at 18:46:48 UTC from IEEE Xplore. Restrictions apply.

978-1-4244-2324-8/08/$25.00 © 2008 IEEE.
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2008 proceedings.
Juniper Exhibit 1023, page 2

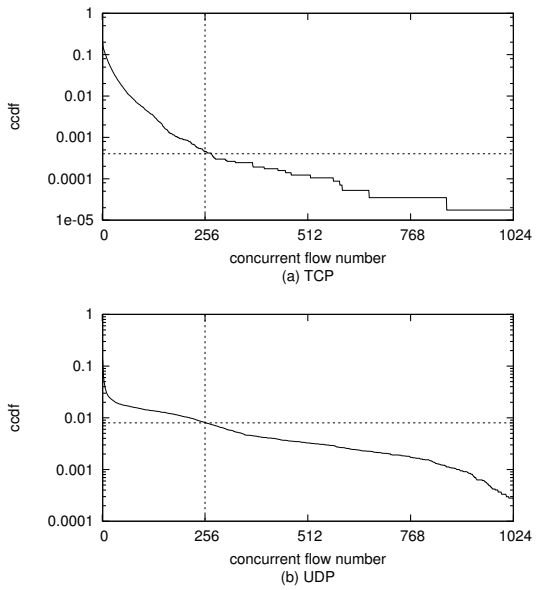

Fig. 2. The complementary cumulative distribution of end hosts with different maximum concurrent flow number.

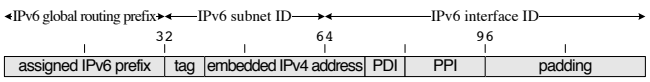

Fig. 3. The address mapping format of IVI.

we focus the observation on TCP only, then the satisfactory rate will increase to 99.95%.

III. The Architecture

Stateless address mapping with multiplexing is the essential component of the proposed IVI architecture. Here we'll firstly introduce the mapping rule and then describe how translators and end host need to change themselves to accommodate the mapping. Throughout this paper, it is assumed that an IPv6 provider who would like to adopt this approach holds at least one IPv6 prefix not smaller than /32, and it would like to devote 1/256 of the address space (e.g. a /40 block of the provider prefix) to be applied for the address mapping.

A. Address Mapping

Fig. 3 described the stateless address mapping rule for the IVI architecture, where a block of IPv4 customer network address is mapped into an IPv6 block that its provider holds. Each segment defined in the address format is explained as follows:

- *Assigned IPv6 prefix:* the 32-bit IPv6 prefix held by the network.
- *Tag:* a characteristic octet indicating that the following 32 bits are an IPv4 address embedded for the purpose of mapping and multiplexing.
- *Embedded IPv4 address:* an IPv4 address assigned to a couple of hosts within the customer network.
- *Port Divisor Indicator (PDI):* an octet to indicate, in the power of 2, how many IPv6 addresses sharing a same IPv4 address. In fact, only the lower 4-bit is significant, so the legal value of PDI is one from 0x00 to 0x0F.
- *Port Prefix Indicator (PPI):* a 16-bit field to indicate the port range a specific IPv6 host can use as a port prefix.
- *Padding:* reserved currently, filled with zero.

TABLE I
An Example of Address Mapping Table

IPv4 prefix	IPv6 prefix	PDI
`100.1.0.128/30`	`2001:ab:ff00::/40`	0x00
`100.1.0.0/24`	`2001:ab:ff00::/40`	0x08
`100.2.0.0/24`	`2001:cd:ff00::/40`	0x08
`0.0.0.0/0`	`2001:cd:ff00::/40`	0x00

For example, 2001:ab:ff64:100:208:d00:: is an IPv6 address conforming to this format, where 2001:ab::/32 is held by the provider and the prefix 2001:ab:ff00::/40 is used to deploy IVI. The embedded IPv4 address of this IPv6 address is 100.1.0.2. The PDI, 0x08, indicates that the multiplex ratio is 1:256, i.e. 256 ports are available for this host; while the PPI, 0x0D00, tells the these 256 port numbers are those from 3328 to 3583 (0x0D00 to 0x0DFF).

Note that PDI, which indicates the multiplexing ratio, is not a fixed parameter. Thus, IVI is flexible in its deployment. Specially, if the Internet Service Provider (ISP) is more concerned with deployment, PDI could be set to zero (consequently, PPI will also be zero), in which way IVI could fall back into a one-to-one mapping scheme and end hosts could remain totally unchanged.

With the address mapping rule commonly known within the network, IPv4 source and destination addresses in a packet can be translated into their IPv6 version without any explicitly stored states.

SIIT ever used the stateless address mapping defined in [11], with a not very successful practice due to it potentially messing the global IPv6 routing table. Our mapping is a prefix-specific approach, which keeps the provider's IPv6 address space well aggregated. Furthermore, when the transition completely ends, the mapped IPv6 address is also a native, globally aggregated unicast address, which can be used further without renumbering.

B. Edge Router

The equipments at the edges of an IPv6 network and the IPv4 networks are called edge routers. Edge routers perform the functionality of mapping addresses for packets and translate routing information from IPv4 to IPv6. Edge routers could be the only dual-stack nodes of the IPv6 network. As field translation has been much discussed elsewhere (like in [8], [12], etc.), we only focus on the address manipulation rules.

Every edge router maintains a local mapping table for the dedicated IPv6 prefix and PDI for every IPv4 prefix. This mapping table is organized similar to a routing table in that it also uses the longest match when looking up an entry for a specific IPv4 address. Table I shows an example.

When a packet is received from IPv6 network, the edge router will compare PDI and PPI in the IP header with

Authorized licensed use limited to: UCLA Library. Downloaded on August 04,2023 at 18:46:48 UTC from IEEE Xplore. Restrictions apply.
978-1-4244-2324-8/08/$25.00 © 2008 IEEE.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2008 proceedings.
Juniper Exhibit 1023, page 3

the port number used in the transportation layer header to validate the port number used is adhere to the rule. And then the embedded IPv4 addresses are extracted to form a new IPv4 header. For instance, an "IPv6 address, port" pair [2001:ab:ff64:100:0208::]:80 will be translated into "IPv4 address, port" pair 100.1.0.2:80.

When a packet is received from IPv4 network, the edge router will look up the mapping table to decide suitable IPv6 prefixes and PDI for source and destination addresses, and will use corresponding PDI and port number to calculate the value of PPI. And then it can embed IPv4 addresses with IPv6 prefixes, PDI and PPI into IPv6 addresses. For instance, 100.1.0.2:3447 will be translated into [2001:ab:ff64:100:208:d00::]:3447.

C. End Host

Once IVI is deployed, there will be few impacts on end hosts except every end host could only use a limited range of port number. As we ever explained, this is a indispensable tradeoff. Since most client application could bind any port that is available in the system, applications on end hosts are expected to run normally without any modification.

As mentioned above, if the ISP is mostly concerned with the deployment, it could deploy IVI with multiplexing ratio set to 1:1 so that there is no change required on end hosts.

On the other hand, if the multiplexing ratio is not 1:1, operating systems on end hosts are required to recognize the address allocated and to offer applications only available ports. Two sorts of solutions, both of which could be delivered by patch or plug-in provided by ISP, are described here.

A fundamental solution is to modify the system call related to bind() in the socket library of the operating systems, so that it will refuse the application from binding an invalid port according to its IPv6 address.

We could have a work-around as well. A simple toolkit that could automatically occupy (i.e. listen to) all ports which are not actually available according to the addressing scheme could be developed and deployed. This work-around is easy and platform-independent.

IV. Evaluation

Port limitation is a big concern that impacts the feasibility of IVI architecture, and therefore we carry out a series of simulations. The actual user behaviors are modeled through network traffic measurement. We passively collect one-hour traffic passing through the edge of a campus network[5], which has been assigned a /16 IPv4 prefix, with about 57,000 hosts online simultaneously. Through collected traffic, we profile every single flow of each end host using the methodology specified in [13], with the timeout parameter set to 60 seconds influenced by [14]. Because we focus on the hosts inside the network, we classify every flow according to the tuple "local address, local port". Therefore, each flow in our evaluation occupies a single port on end hosts.

[5]Data are collected on the edge of Tsinghua University Network, from 2007-07-07 15:00 UTC+8 through 2007-07-07 15:59 UTC+8.

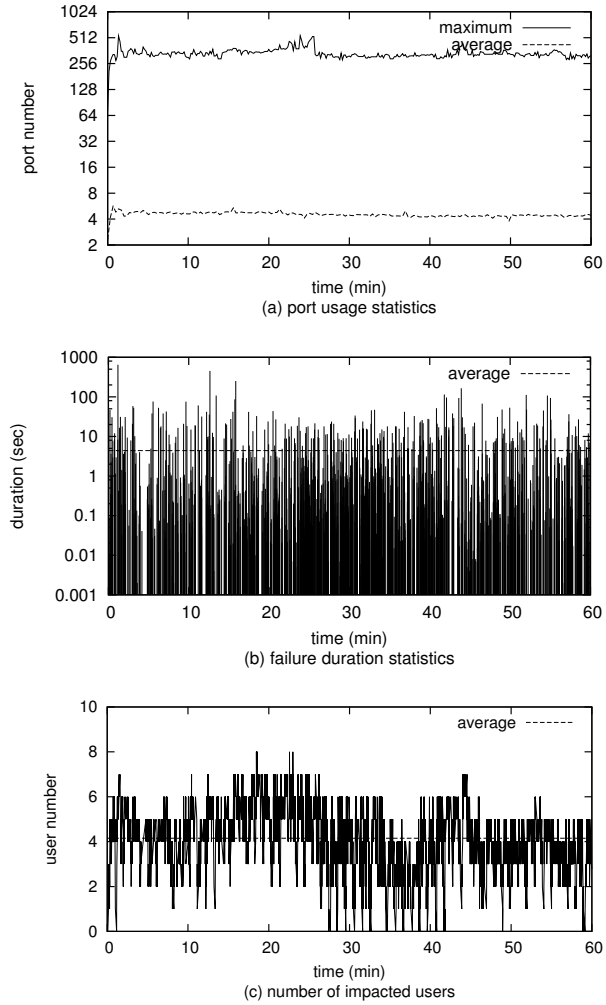

Fig. 4. Statistic results on port assignment failure.

Using flow profile data, we first quantify the limitation of IVI – port assignment failure, and then compare state storage and minimum IPv4 address required in our approach with NAT-based ones and SIIT. Throughout our evaluation, the multiplexing ratio is set to 1:256, which means each host could use 256 ports at most.

A. Port Assignment Failure

We give a group of statistics concerning the cases when hosts attempt to use more ports but quota has been reached.

Fig.4(a) illustrates the average and maximum port number of all hosts in every second. The average number is about 6, which is quite low compared to the maximum number, which is shifting between 256 and 512 during most of the time. Thus, although assigning 256 ports to each host cannot satisfy the maximal requirement of some hosts, but it is already quite ample for most hosts in average.

Looking further into those hosts not satisfied by the 256 ports, we plot the duration of every occurrence when a host consumes more than 256 ports in Fig. 4(b). The result suggests that even if a host wants to consume more ports than available,

Authorized licensed use limited to: UCLA Library. Downloaded on August 04,2023 at 18:46:48 UTC from IEEE Xplore. Restrictions apply.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2008 proceedings.
Juniper Exhibit 1023, page 4

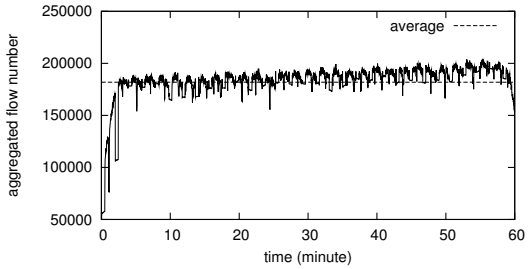

Fig. 5. Aggregated flow number against the lapse of time.

TABLE II
COMPARISON BETWEEN IVI AND ITS COUNTERPARTS

Scheme	Avg. State Mem. (KB)	Max. State Mem. (KB)	Min. Address
NAT	2,547	2,860	4
NAT-PT	4,366	4,905	4
SIIT	0	0	57,088
IVI	0	0	223

the duration of such relative shortage is commonly less than 10 seconds, or about 4 seconds on average. That is to say, at the peak moment, connection may suffer 4 seconds average delay to be made.

The number of hosts consumes more than 256 ports simultaneously is plotted in Fig. 4(c). The maximum number of hosts never exceed 10, with the average number is less than 5. That is to say, among about 57,000 hosts online, only less than 10 hosts have the demand to establish more than 256 flows. We confirmed that the hosts, which are top 5 in long-duration, high-occurrence of using more than 256 ports, are in fact FTP servers which reside inside the campus network. This inspires us to tune the multiplexing ratio for dedicated server with lower PDI values.

B. Comparison with Counterparts

Aside from the novelties of IVI ever stated in the Section I, especially its support for a smooth transition, we still would like to present here a quantitative comparison with existing schemes, focusing on the state storage and minimum IPv4 address requirements.

The state storage of NAT-based schemes is decided by the aggregated flow number, as shown in Fig. 5. For traditional NAT, each entry will contain a quintuple "global IPv4 address, global port, local IPv4 address, local port, timeout", consuming at least 14 bytes of memory; while for NAT-PT, each entry will contain a similar quintuple which occupies 24 bytes. The computed result is shown in Tab. II.

It is clear that we eliminate the address binding tables which have about 40,000 entries and thus take up several megabytes. In NAT-based solutions, the translation device need to look up such tables every time a packet passes through it, and thus becomes the bottle-neck in both performance and scalability in the network.

It is also notable that NAT-based solutions really have higher IPv4 address utilization but IVI gains the statelessness, while SIIT consumes 256 times of addresses to gain the same.

V. CONCLUSION

Economic models have revealed that the inter-protocol accessibility is very important for the adoption of a new protocol and for a successful transition model. To have inter-protocol accessibility, deploying translation facility in large scale is unavoidable. However, ever existing approaches are not counted as scalable due to either involving huge amount of states or violating the address uniqueness and routing aggregation during the transition process. Our approach combines the statelessness and the temporal address uniqueness, via a prefix-specific address mapping mechanism. Furthermore, considering the trend of IPv4 address exhaustion, we introduce multiplexing into the proposed approach, to gain enough time for the migration of Internet resources and applications from IPv4 to IPv6.

Limiting the port number that an individual host can use causes a tradeoff between the address utilization and the end-host performance. Our quantitative evaluations provide positive support for accepting the tradeoff. We offer ISP with the flexibility of tuning its multiplex ratio according to its reality. Indeed, along with the IPv4 addresses becoming more and more precious, the expected multiplex ratio will be higher but we believe this is also the process that more resources is migrating from the IPv4 world to the IPv6 one and therefore the necessity of having the translation is degrading.

REFERENCES

[1] S. Deering and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification," IETF RFC2460, Dec. 1998.
[2] D. Joseph, N. Shetty, J. Chuang, and I. Stoica, "Modeling the Adoption of new Network Architectures," in *Proceedings of 3rd International Conference on emerging Networking EXperiments and Technologies, CoNEXT '07*, New York, United States, 2007.
[3] P. Srisuresh and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)," IETF RFC3022, Jan. 2001.
[4] G. Tsirtsis and P. Srisuresh, "Network Address Translation - Protocol Translation (NAT-PT)," IETF RFC2766, Feb. 2000.
[5] C. Aoun and E. Davies, "Reasons to Move the Network Address Translator - Protocol Translator (NAT-PT) to Historic Status," IETF RFC4966, July 2007.
[6] B. Carpenter, "Internet Transparency," IETF RFC2775, Feb. 2000.
[7] L. Phifer, "The Trouble with NAT," *The Internet Protocol Journal*, vol. 3, no. 4, pp. 2–13, 2000.
[8] E. Nordmark, "Stateless IP/ICMP Translation Algorithm (SIIT)," IETF RFC2765, Feb. 2000.
[9] M. Fiuczynski, V. Lam, and B. Bershad, "The Design and Implementation of an IPv6/IPv4 Network Address and Protocol Translator," in *Proceedings of the USENIX Annual Technical Conference, USENIX '98*, New Orleans, United States, 1998, pp. 201–212.
[10] H. Zhang and M. Chen, "Forming an IPv6-only Core for Today's Internet," in *Proceedings of ACM SIGCOMM Workshops, 2007*, Kyoto, Japan, 2007, pp. 379–384.
[11] R. Hinden and S. Deering, "IP Version 6 Addressing Architecture," IETF RFC4291, Feb. 2006.
[12] M. Chen, X. Li, A. Li, and Y. Cui, "Forwarding IPv4 Traffics in Pure IPv6 Backbone with Stateless Address Mapping," in *Proceedings of 10th IEEE/IFIP Network Operations and Management Symposium, NOMS 2006*, Vancouver, Canada, 2006, pp. 260–270.
[13] K. Claffy, H. Braun, and G. Polyzos, "A Parameterizable Methodology for Internet Traffic Flow Profiling," *IEEE Journal of Selected Areas in Communications*, vol. 13, no. 8, pp. 1481–1494, 1995.
[14] C. Fraleigh, S. Moon, B. Lyles, C. Cotton, M. Khan, D. Moll, R. Rockell, T. Seely, and C. Diot, "Packet-Level Traffic Measurements from the Sprint IP Backbone," *IEEE Network*, vol. 17, no. 6, pp. 6–16, 2003.

Authorized licensed use limited to: UCLA Library. Downloaded on August 04,2023 at 18:46:48 UTC from IEEE Xplore. Restrictions apply.

978-1-4244-2324-8/08/$25.00 © 2008 IEEE.
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2008 proceedings.
Juniper Exhibit 1023, page 5