Network Working Group
Request for Comments: 2956
Category: Informational

M. Kaat
SURFnet ExpertiseCentrum bv

October 2000

Overview of 1999 IAB Network Layer Workshop

Status of this Memo

Copyright Notice

Abstract

This document is an overview of a workshop held by the Internet Architecture Board (IAB) on the Internet Network Layer architecture hosted by SURFnet in Utrecht, the Netherlands on 7-9 July 1999. The goal of the workshop was to understand the state of the network layer and its impact on continued growth and usage of the Internet. Different technical scenarios for the (foreseeable) future and the impact of external influences were studied. This report lists the conclusions and recommendations to the Internet Engineering Task Force (IETF) community.

Table of Contents

1. Introduction

From July 7 to July 9, 1999 the Internet Architecture Board (IAB) held a workshop on the architecture of the Internet Network Layer. The Network Layer is usually referred to as the IP layer. The goal of the workshop was to discuss the current state of the Network Layer and the impact various currently deployed or future mechanisms and technologies might have on the continued growth and usage of the Internet.

The most important issues to be discussed were:

- Status of IPv6 deployment and transition issues
- Alternative technical strategies in case IPv6 is not adopted
- Globally unique addresses and 32 bit address depletion
- Global connectivity and reachability
- Fragmentation of the Internet
- End to end transparency and the progressive loss thereof
- End to end security
- Complications of address sharing mechanisms (NAT, RSIP)
- Separation of identification and location in addressing
- Architecture and scaling of the current routing system

The participants looked into several technical scenarios and discussed the feasibility and probability of the deployment of each scenario. Among the scenarios were for example full migration to IPv6, IPv6 deployment only in certain segments of the network, no significant deployment of IPv6 and increased segmentation of the IPv4 address space due to the use of NAT devices.

Based on the discussion of these scenarios several trends and external influences were identified which could have a large impact on the status of the network layer, such as the deployment of wireless network technologies, mobile networked devices and special purpose IP devices.

The following technical issues were identified to be important goals:

- o Deployment of end to end security
- o Deployment of end to end transport
- o Global connectivity and reachability should be maintained
- o It should be easy to deploy new applications
- o It should be easy to connect new hosts and networks to the Internet ("plug and ping")

By the notion "deployment of end to end transport" it is meant that it is a goal to be able to deploy new applications that span from any host to any other host without intermediaries, and this requires transport protocols with similar span (see also [1]).

This document summarizes the conclusions and recommendations made by the workshop. It should be noted that not all participants agreed with all of the statements, and it was not clear whether anyone agreed with all of them. The recommendations made however are based on strong consensus among the participants.

2. Conclusions and Observations

The participants came to a number of conclusions and observations on several of the issues mentioned in section 1. In the following sections 2.1-2.10 these conclusions will be described.

2.1 Transparency

In the discussions transparency was referred to as the original Internet concept of a single universal logical addressing scheme and the mechanisms by which packets may flow from source to destination essentially unaltered [1]. This traditional end to end transparency has been lost in the current Internet, specifically the assumption that IPv4 addresses are globally unique or invariant is no longer true.

There are multiple causes for the loss of transparency, for example the deployment of network address translation devices, the use of private addresses, firewalls and application level gateways, proxies and caches. These mechanisms increase fragmentation of the network layer, which causes problems for many applications on the Internet. It adds up to complexity in applications design and inhibits the deployment of new applications. In particular, it has a severe effect on the deployment of end to end IP security.

Another consequence of fragmentation is the deployment of "split DNS" or "two faced DNS", which means that the correspondence between a given FQDN and an IPv4 address is no longer universal and stable over long periods (see section 2.7).

End to end transparency will probably not be restored due to the fact that some of the mechanisms have an intrinsic value (e.g. firewalls, caches and proxies) and the loss of transparency may be considered by some as a security feature. It was however concluded that end to end transparency is desirable and an important issue to pursue. Transparency is further explored in [1].

2.2 NAT, Application Level Gateways & Firewalls

The previous section indicated that the deployment of NAT (Network Address Translation), Application Level Gateways and firewalls causes loss of network transparency. Each of them is incompatible with certain applications because they interfere with the assumption of end to end transparency. NAT especially complicates setting up servers, peer to peer communications and "always-on" hosts as the endpoint identifiers, i.e. IP addresses, used to set up connections are globally ambiguous and not stable (see [2]).

NAT, application level gateways and firewalls however are being increasingly widely deployed as there are also advantages to each, either real or perceived. Increased deployment causes a further decline of network transparency and this inhibits the deployment of new applications. Many new applications will require specialized Application Level Gateways (ALGs) to be added to NAT devices, before those applications will work correctly when running through a NAT device. However, some applications cannot operate effectively with NAT even with an ALG.

2.3 Identification and Addressing

In the original IPv4 network architecture hosts are globally, permanently and uniquely identified by an IPv4 address. Such an IP address is used for identification of the node as well as for locating the node on the network. IPv4 in fact mingles the semantics of node identity with the mechanism used to deliver packets to the node. The deployment of mechanisms that separate the network into multiple address spaces breaks the assumption that a host can be uniquely identified by a single IP address. Besides that, hosts may wish to move to a different location in the network but keep their identity the same. The lack of differentiation between the identity and the location of a host leads to a number of problems in the current architecture.

Several technologies at this moment use tunneling techniques to overcome the problem or cannot be deployed in the case of separate address spaces. If a node could have some sort of a unique identifier or endpoint name this would help in solving a number of problems.

It was concluded that it may be desirable on theoretical grounds to separate the node identity from the node locator. This is especially true for IPsec, since IP addresses are used (in transport mode) as identifiers which are cryptographically protected and hence MUST remain unchanged during transport. However, such a separation of identity and location will not be available as a near-term solution, and will probably require changes to transport level protocols. However, the current specification of IPsec does allow to use some other identifier than an IP address.

2.4 Observations on Address Space

There is a significant risk that a single 32 bit global address space is insufficient for foreseeable needs or desires. The participants' opinions about the time scale over which new IPv4 addresses will still be available for assignment ranged from 2 to 20 years. However, there is no doubt that at the present time, users cannot obtain as much IPv4 address space as they desire. This is partly a result of the current stewardship policies of the Regional Internet Registries (RIRs).

It was concluded that it ought to be possible for anybody to have global addresses when required or desired. The absence of this inhibits the deployment of some types of applications. It should however be noted that there will always be administrative boundaries, firewalls and intranets, because of the need for security and the implementation of policies. NAT is seen as a significant complication on these boundaries. It is often perceived as a security feature because people are confusing NATs with firewalls.

2.5 Routing Issues

A number of concerns were raised regarding the scaling of the current routing system. With current technology, the number of prefixes that can be used is limited by the time taken for the routing algorithm to converge, rather than by memory size, lookup time, or some other factor. The limit is unknown, but there is some speculation, of extremely unclear validity, that it is on the order of a few hundred thousand prefixes. Besides the computational load of calculating routing tables, the time it takes to distribute routing updates across the network, the robustness and security of the current routing system are also important issues. The only known addressing

scheme which produces scalable routing mechanisms depends on topologically aggregated addresses, which requires that sites renumber when their position in the global topology changes. Renumbering remains operationally difficult and expensive ([3], [4]). It is not clear whether the deployment of IPv6 would solve the current routing problems, but it should do so if it makes renumbering easier.

At least one backbone operator has concerns about the convergence time of internetwork-wide routing during a failover. This operator believes that current convergence times are on the order of half a minute, and possibly getting worse. Others in the routing community did not believe that the convergence times are a current issue. Some, who believe that real-time applications (e.g. telephony) require sub-second convergence, are concerned about the implications of convergence times of a half minute on such applications.

Further research is needed on routing mechanisms that might help palliate the current entropy in the routing tables, and can help reduce the convergence time of routing computations.

The workshop discussed global routing in a hypothetical scenario with no distinguished root global address space. Nobody had an idea how to make such a system work. There is currently no well-defined proposal for a new routing system that could solve such a problem.

For IPv6 routing in particular, the GSE/8+8 proposal and IPNG WG analysis of this proposal ([5]) are still being examined by the IESG. There is no consensus in the workshop whether this proposal could be made deployable.

2.6 Observations on Mobility

Mobility and roaming require a globally unique identifier. This does not have to be an IP address. Mobile nodes must have a widely usable identifier for their location on the network, which is an issue if private IP addresses are used or the IP address is ambiguous (see also section 2.3). Currently tunnels are used to route traffic to a mobile node. Another option would be to maintain state information at intermediate points in the network if changes are made to the packets. This however reduces the flexibility and it breaks the end to end model of the network. Keeping state in the network is usually considered a bad thing. Tunnels on the other hand reduce the MTU size. Mobility was not discussed in detail as a separate IAB workshop is planned on this topic.

2.7 DNS issues

If IPv6 is widely deployed, the current line of thinking is that site renumbering will be significantly more frequent than today. This will have an impact on DNS updates. It is not clear what the scale of DNS updates might be, but in the most aggressive models it could be millions a day. Deployment of the A6 record type which is defined to map a domain name to an IPv6 address, with the provision for indirection for leading prefix bits, could make this possible ([6]).

Another issue is the security aspect of frequent updates, as they would have to been done dynamically. Unless we have fully secured DNS, it could increase security risks. Cached TTL values might introduce problems as the cached records of renumbered hosts will not be updated in time. This will become especially a problem if rapid renumbering is needed.

Another already mentioned issue is the deployment of split DNS (see section 2.1). This concept is widely used in the Intranet model, where the DNS provides different information to inside and outside queries. This does not necessarily depend on whether private addresses are used on the inside, as firewalls and policies may also make this desirable. The use of split DNS seems inevitable as Intranets will remain widely deployed. But operating a split DNS raises a lot of management and administrative issues. As a work around, a DNS Application Level Gateway ([7]) (perhaps as an extension to a NAT device) may be deployed, which intercepts DNS messages and modifies the contents to provide the appropriate answers. This has the disadvantage that it interferes with the use of DNSSEC ([8]).

The deployment of split DNS, or more generally the existence of separate name spaces, makes the use of Fully Qualified Domain Names (FQDNs) as endpoint identifiers more complex.

2.8 NAT and RSIP

Realm-Specific IP (RSIP), a mechanism for use with IPv4, is a work item of the IETF NAT WG. It is intended as an alternative (or as a complement) to network address translation (NAT) for IPv4, but other uses are possible (for example, allowing end to end traffic across firewalls). It is similar to NAT, in that it allows sharing a small number of external IPv4 addresses among a number of hosts in a local address domain (called a 'realm'). However, it differs from NAT in that the hosts know that different externally-visible IPv4 addresses are being used to refer to them outside their local realm, and they

know what their temporary external address is. The addresses and other information are obtained from an RSIP server, and the packets are tunneled across the first routing realm ([9], [10]).

The difference between NAT and RSIP - that an RSIP client is aware of the fact that it uses an IP address from another address space, while with NAT, neither endpoint is aware that the addresses in the packets are being translated - is significant. Unlike NAT, RSIP has the potential to work with protocols that require IP addresses to remain unmodified between the source and destination. For example, whereas NAT gateways preclude the use of IPsec across them, RSIP servers can allow it [11].

The addition of RSIP to NATs may allow them to support some applications that cannot work with traditional NAT ([12]), but it does require that hosts be modified to act as RSIP clients. It requires changes to the host's TCP/IP stack, any layer-three protocol that needs to be made RSIP-aware will have to be modified (e.g. ICMP) and certain applications may have to be changed. The exact changes needed to host or application software are not quite well known at this moment and further research into RSIP is required.

Both NAT and RSIP assume that the Internet retains a core of global address space with a coherent DNS. There is no fully prepared model for NAT or RSIP without such a core; therefore NAT and RSIP face an uncertain future whenever the IPv4 address space is finally exhausted (see section 2.4). Thus it is also a widely held view that in the longer term the complications caused by the lack of globally unique addresses, in both NAT and RSIP, might be a serious handicap ([1]).

If optimistic assumptions are made about RSIP (it is still being defined and a number of features have not been implemented yet), the combination of NAT and RSIP seems to work in most cases. Whether RSIP introduces specific new problems, as well as removing some of the NAT issues, remains to be determined.

Both NAT and RSIP may have trouble with the future killer application, especially when this needs QoS features, security and/or multicast. And if it needs peer to peer communication (i.e. there would be no clear distinction between a server and a client) or assumes "always-on" systems, this would probably be complex with both NAT and RSIP (see also section 2.2).

2.9 NAT, RSIP and IPv6

Assuming IPv6 is going to be widely deployed, network address translation techniques could play an important role in the transition process from IPv4 to IPv6 ([13]). The impact of adding RSIP support

to hosts is not quite clear at this moment, but it is less than adding IPv6 support since most applications probably don't need to be changed. And RSIP needs no changes to the routing infrastructure, but techniques such as automatic tunneling ([14]) and 6to4 ([15]) would also allow IPv6 traffic to be passed over the existing IPv4 routing infrastructure. While RSIP is principally a tool for extending the life of IPv4, it is not a roadblock for the transition to IPv6. The development of RSIP is behind that of IPv6, and more study into RSIP is required to determine what the issues with RSIP might be.

2.10 Observations on IPv6

An important issue in the workshop was whether the deployment of IPv6 is feasible and probable. It was concluded that the transition to IPv6 is plausible modulo certain issues. For example applications need to be ported to IPv6, and production protocol stacks and production IPv6 routers should be released. The core protocols are finished, but other standards need to be pushed forward (e.g. MIBs). A search through all RFCs for dependencies on IPv4 should be made, as was done for the Y2K problem, and if problems are found they must be resolved. As there are serious costs in implementing IPv6 code, good business arguments are needed to promote IPv6.

One important question was whether IPv6 could help solve the current problems in the routing system and make the Internet scale better. It was concluded that "automatic" renumbering is really important when prefixes are to be changed periodically to get the addressing topology and routing optimized. This also means that any IP layer and configuration dependencies in protocols and applications will have to be removed ([3]). One example that was mentioned is the use of IP addresses in the PKI (IKE). There might also be security issues with "automatic" renumbering as DNS records have to be updated dynamically (see also section 2.7).

Realistically, because of the dependencies mentioned, IPv6 renumbering cannot be truly automatic or instantaneous, but it has the potential to be much simpler operationally than IPv4 renumbering, and this is critical to market and ISP acceptance of IPv6.

Another issue is whether existing TCP connections (using the old address(es)) should be maintained across renumbering. This would make things much more complex and it is foreseen that old and new addresses would normally overlap for a long time.

There was no consensus on how often renumbering would take place or how automatic it can be in practice; there is not much experience with renumbering (maybe only for small sites).

3. Recommendations

3.1 Recommendation on Namespace

The workshop recommends the IAB to appoint a panel to make specific recommendations to the IETF about:

i) whether we should encourage more parts of the stack to adopt a namespace for end to end interactions, so that a) NAT works 'better', and b) we have a little more independence between the internetwork and transport and above layers;
ii) if so, whether we should have a single system-wide namespace for this function, or whether it makes more sense to allow various subsystems to chose the namespace that makes sense for them;
iii) and also, what namespace(s) [depending on the output of the point above] that ought to be.

3.2 Recommendations on RSIP

RSIP is an interesting idea, but it needs further refinement and study. It does not break the end to end network model in the same way as NAT, because an RSIP host has explicit knowledge of its temporary global address. Therefore, RSIP could solve some of the issues with NAT. However, it is premature to recommend it as a mainstream direction at this time.

It is recommended that the IETF should actively work on RSIP, develop the details and study the issues.

3.3 Recommendations on IPv6

3.3.1

The current model of TLA-based addressing and routing should be actively pursued. However, straightforward site renumbering using TLA addresses is really needed, should be as nearly automatic as possible, and should be shown to be real and credible by the IPv6 community.

3.3.2

Network address translation techniques, in addition to their immediate use in pure IPv4 environments, should also be viewed as part of the starting point for migration to IPv6. Also RSIP, if successful, can be a starting point for IPv6 transition.

While the basic concepts of the IPv4 specific mechanisms NAT and RSIP are also being used in elements of the proposed migration path to IPv6 (in NAT-PT for NAT, and SIIT and AIIH for RSIP), NAT and RSIP for IPv4 are not directly part of a documented transition path to IPv6.

The exact implications, for transition to IPv6, of having NAT and RSIP for IPv4 deployed, are not well understood. Strategies for transition to IPv6, for use in IPv4 domains using NAT and RSIP for IPv4, should be worked out and documented by the IETF.

3.3.3

The draft analysis of the 8+8/GSE proposal should be evaluated by the IESG and accepted or rejected, without disturbing ongoing IPv6 deployment work. The IESG should use broad expertise, including liaison with the endpoint namespace panel (see section 3.1) in their evaluation.

3.4 Recommendations on IPsec

It is urgent that we implement and deploy IPsec using some other identifier than 32-bit IP addresses (see section 2.3). The current IPsec specifications support the use of several different Identity types (e.g. Domain Name, User@Domain Name). The IETF should promote implementation and deployment of non-address Identities with IPsec. We strongly urge the IETF to completely deprecate the use of the binary 32-bit IP addresses within IPsec, except in certain very limited circumstances, such as router to router tunnels; in particular any IP address dependencies should be eliminated from ISAKMP and IKE.

Ubiquitous deployment of the Secure DNS Extensions ([8]) should be strongly encouraged to facilitate widespread deployment of IPsec (including IKE) without address-based Identity types.

3.5 Recommendations on DNS

Operational stability of DNS is paramount, especially during a transition of the network layer, and both IPv6 and some network address translation techniques place a heavier burden on DNS. It is therefore recommended to the IETF that, except for those changes that are already in progress and will support easier renumbering of networks and improved security, no fundamental changes or additions to the DNS be made for the foreseeable future.

In order to encourage widespread deployment of IPsec, rapid deployment of DNSSEC is recommended to the operational community.

3.6 Recommendations on Routing

The only known addressing scheme which produces scalable routing mechanisms depends on topologically aggregated addresses, which requires that sites renumber when their position in the global topology changes. Thus recommendation 3.3.1 is vital for routing IPv6.

Although the same argument applies to IPv4, the installed base is simply too large and the PIER working group showed that little can be done to improve renumbering procedures for IPv4. However, NAT and/or RSIP may help.

In the absence of a new addressing model to replace topological aggregation, and of clear and substantial demand from the user community for a new routing architecture (i.e. path-selection mechanism) there is no reason to start work on standards for a "next generation" routing system in the IETF. Therefore, we recommend that work should continue in the IRTF Routing Research Group.

3.7 Recommendations on Application layer and APIs

Most current APIs such as sockets are an obstacle to migration to a new network layer of any kind, since they expose network layer internal details such as addresses.

It is therefore recommended, as originally recommended in RFC 1900 [3], that IETF protocols, and third-party applications, avoid any explicit awareness of IP addresses, when efficient operation of the protocol or application is feasible in the absence of such awareness. Some applications and services may continue to need to be aware of IP addresses. Until we once again have a uniform address space for the Internet, such applications and services will necessarily have limited deployability, and/or require ALG support in NATs.

Also we recommend an effort in the IETF to generalize APIs to offer abstraction from all network layer dependencies, perhaps as a side-effect of the namespace study of section 3.1.

4. Security Considerations

The workshop did not address security as a separate topic, but the role of firewalls, and the desirability of end to end deployment of IPsec, were underlying assumptions. Specific recommendations on security are covered in sections 3.4 and 3.5.

References

[1] Carpenter, B., "Internet Transparency", RFC 2775, February 2000.

[2] Hain, T., "Architectural Implications of NAT", Work in Progress.

[3] Carpenter, B. and Y. Rekhter, "Renumbering Needs Work", RFC 1900, February 1996.

[4] Ferguson, P and H. Berkowitz, "Network Renumbering Overview: Why would I want it and what is it anyway?", RFC 2071, January 1997.

[5] M. Crawford, A. Mankin, T. Narten, J.W. Stewart, III, L. Zhang, "Separating Identifiers and Locators in Addresses: An Analysis of the GSE Proposal for IPv6", Work in Progress.

[6] Crawford, M., and C. Huitema, "DNS Extensions to Support IPv6 Address Aggregation and Renumbering", RFC 2874, July 2000.

[7] Srisuresh, P., Tsirtsis, G., Akkiraju, P. and A. Heffernan, "DNS extensions to Network Address Translators (DNS_ALG)", RFC 2694, September 1999.

[8] Eastlake, D., "Domain Name System Security Extensions", RFC 2535, March 1999.

[9] M. Borella, D. Grabelsky, J. Lo, K. Tuniguchi "Realm Specific IP: Protocol Specification", Work in Progress.

[10] M. Borella, J. Lo, D. Grabelsky, G. Montenegro "Realm Specific IP: Framework", Work in Progress.

[11] G. Montenegro, M. Borella, "RSIP Support for End-to-end IPsec", Work in Progress.

[12] M. Holdrege, P. Srisuresh, "Protocol Complications with the IP Network Address Translator", Work in Progress.

[13] Tsirtsis, G. and P. Srisuresh, "Network Address Translation - Protocol Translation (NAT-PT)", RFC 2766, February 2000.

[14] Gilligan, R. and E. Nordmark, "Transition Mechanisms for IPv6 Hosts and Routers", RFC 2893, August 2000.

[15] B. Carpenter, K. Moore, "Connection of IPv6 Domains via IPv4 Clouds", Work in Progress.

Appendix A. Participants

Harald Alvestrand	harald@alvestrand.no
Ran Atkinson	rja@corp.home.net
Rob Austein	sra@hactrn.net
Steve Bellovin	smb@research.att.com
Randy Bush	randy@psg.com
Brian E Carpenter	brian@hursley.ibm.com
Vint Cerf	vcerf@MCI.NET
Noel Chiappa	jnc@lcs.mit.edu
Matt Crawford	crawdad@fnal.gov
Robert Elz	kre@munnari.OZ.AU
Tony Hain	tonyhain@microsoft.com
Matt Holdrege	matt@ipverse.com
Erik Huizer	huizer@cs.utwente.nl
Geoff Huston	gih@telstra.net
Van Jacobson	van@cisco.com
Marijke Kaat	Marijke.Kaat@surfnet.nl
Daniel Karrenberg	Daniel.Karrenberg@ripe.net
John Klensin	klensin@jck.com
Peter Lothberg	roll@Stupi.SE
Olivier H. Martin	Olivier.Martin@cern.ch
Gabriel Montenegro	gab@sun.com
Keith Moore	moore@cs.utk.edu
Robert (Bob) Moskowitz	rgm@htt-consult.com
Philip J. Nesser II	pjnesser@nesser.com
Kathleen Nichols	kmn@cisco.com
Erik Nordmark	nordmark@eng.sun.com
Dave Oran	oran@cisco.com
Yakov Rekhter	yakov@cisco.com
Bill Sommerfeld	sommerfeld@alum.mit.edu
Bert Wijnen	wijnen@vnet.ibm.com
Lixia Zhang	lixia@cs.ucla.edu

Author’s Address

Marijke Kaat
SURFnet ExpertiseCentrum bv
P.O. Box 19115
3501 DC Utrecht
The Netherlands

Phone: +31 30 230 5305
Fax: +31 30 230 5329
EMail: Marijke.Kaat@surfnet.nl

Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.