Does IPv6 Improve the Scalability of the Internet?

Philippe Owezarski

LAAS-CNRS, 7 Avenue du Colonel Roche
31077 Toulouse cedex 4, France
owe@laas.fr

Absract. The Internet is growing very fast since 10 to 20 years, following an exponential increase. Some scalability issues start to arise in the Internet. A well known one is related to IPv4 addresses exhaustion, that should make the Internet growth stop. Because an access to the Internet is a very strong need for many people, the Internet growth continues thanks to some additional mechanisms as NAT for example. An other important scalability issue, not well known by most of Internet services users concerns the explosion of routing tables that are growing very fast (their size went from 15,000 to 150,0000 entries during the 6 last years), then limiting the Internet performance and scalability by increasing the routing table lookup time, and then reducing routing performances. IPv6 has been designed to cope with such scalability issues (addresses exhaustion and routing table explosion). This paper proposes a monitoring study of some BGP routing tables to analyze the reasons of this huge growth of the number of entries in routing tables. This paper then gives quantitative analysis of the reasons why all routing tables prefixes cannot be aggregated, speaking then of the consequences of NAT, multi-homing, load balancing, broken addresses hierarchy, etc. on routing tables sizes. This paper also presents some of the threats for IPv6 whose deployment in the Internet is so slow, and this point is analyzed in relation with the current strong scalability issues of the Internet.

1. Introduction

The Internet is growing in an exponential way, so following the Moore law[1]. This increase is pushed by the number of new comers that are getting connected to the Internet. This increase is getting more and more important as the Internet is getting more and more popular in new "Internet coming countries", especially in Asia as China, India or Korea. In particular, these new countries are some of the ones that have the largest populations, so increasing again the need for new addresses. The Internet increase is also pushed by new multimedia communication technologies, and related services, as 3rd generation mobiles (3G mobiles) as UMTS[2]. The issues of such an increase are mainly two folds:

1. The amount of IP addresses (meaning the current IPv4 addresses) is (almost) exhausted, and new comers cannot get new address spaces, or at least address spaces that are not large enough to give a native routable IP address to every new host or device that has to be connected to the Internet;

---

[1] Initially, the Moore law was stated for the increase of the power of processors, but it also perfectly fits the increase of the Internet

[2] UMTS: Universal Mobile Telecommunications System

2. The size of BGP[3] [14] routing tables is exploding, meaning that their increase is so large that it creates a big QoS issue: routers are then spending too much time for routing table lookup, then increasing routers and end-to-end delays, and then decreasing the global performance on all connections and of the global network. Just to give an idea about BGP routing table growth, their size went, in average, from 15,000 to 150,0000 entries during the 6 last years.

In order to solve the two issues, expressed just above, IPv6[4] has been proposed by the IETF[5]. At the beginning, IPv6 has been designed for:

1. Providing an increased number of addresses going from 2^{32} to 2^{128}. This change in the number of addresses was supposed to solve the problem of IP addresses exhaustion (providing at least 5 IP addresses by square foot on Earth);
2. Stopping, and even better, reducing the explosion of routing tables. This property can be achieved thanks to a new dynamic mechanism for IPv6 addresses allocation that depends on the real location on Earth of the computer or device to connect to the Internet. This is the auto-configuration mechanism of IPv6. Then, associated to a static process of IPv6 prefixes allocation, with larger enough spaces between adjacent prefixes, IPv6 is supposed to provide a fully hierarchical IPv6 addresses structure, then facilitating the search of IPv6 addresses or prefixes in BGP routing tables.

But IPv6, today, is very far from being widely deployed. One of the reasons is mainly related to the arrival of NAT[6] [11] [12] that provides an alternative solution to IPv6. Of course, NAT is a “dirty” solution that breaks the end-to-end model of the Internet, and in particular not suited for some new kinds of applications as peer-to-peer applications for instance [7]. But NAT has also the strong advantage of being a fast, cheap and available without delay solution. Chinese people that need IP addresses cannot wait for IPv6 to be deployed. So to connect to the Internet they are obliged to chose such a solution, even if it does not provide them with some services and creates some problems with some applications, or for security. The second question about IPv6, not yet answered, and that delays the wide deployment of IPv6, concerns the ability of IPv6 to improve the scalability of the Internet, and in particular if IPv6 is able to reduce the BGP table sizes that are reaching a limit over which the global performance of the Internet is going to decrease. This is one important requirement for Internet carriers.

This paper aims to answer this last question by analyzing the reasons of such an increase of BGP routing tables with IPv4, and then, evaluating (theoretically) if IPv6 is able to provide an improvement for the Internet performance / scalability? Of course, analyzing the BGP tables supposed that the point of view chosen in this paper is the one of an Internet carrier or service provider that has to operate and manage a transport network. This study is based on the monitoring of some routers, especially analyzing some BGP routing tables publicly available. Tables that have been analyzed

[3] BGP: Border Gateway Protocol
[4] For an introduction to IPv6, readers can refer to [6] and [10]
[5] IETF: Internet Engineering Task Force (http://www.ietf.org)
[6] NAT: Network Address Translation

are from Telstra[7], RIPE[8], and some public networks as MAE-EAST and MAE-WEST[9,10].

The following of the paper is built as follows: section 2 presents an analysis of the increase of routing table sizes on a very long period (1989-2001). Section 3 tries to explain the reasons of such a behavior. In particular, it analyzes the reasons that make routing table size increase and the ones that make them decrease. Then based on real tables analysis (the ones coming from IP networks monitoring), this section analyzes the prefixes in BGP tables and explains the reasons why the address space is so segmented. Then section 4, based on the results of BGP tables analysis, presents the main threats for IPv6 not to be deployed in the Internet. Finally, section 5 concludes the paper on our opinion about the need to deploy or not IPv6.

2. The Internet Scalability Issue

As presented before, the average size of BGP tables for Internet backbone routers went from 15,000 to 150,0000 entries during the 6 last years. Figure 1 depicts the evolution of this increase, while Figure 2 depicts the evolution of the number of AS[11] in the Internet. The 2 curves have to be analyzed at the same time to provide an explanation of what happened or is happening. This analysis has been made on Telstra traces. The curve of Figure 1 has 3 main periods:

1. From 1989 to 1994, there is an exponential growth of the BGP tables. Compared to the evolution of the number of AS during this period, it is clear that BGP tables are evolving the same way as AS number. The BGP tables growth is then fully due to the increase of the Internet (because of the arrival of the web).
2. From 1994 to 1998 there was only a linear increase of BGP tables. Compared to the increase of AS numbers that was still increasing in an exponential way, it is clear that the scalability of the Internet as well as its performances were getting better and better: BGP table sizes were growing more slowly than AS numbers. The reason of such an improvement of Internet scalability is related to the deployment of CIDR[12] (CIDR will be presented in part 3.1).
3. Finally, since 1998 there is a resume of the exponential growth of routing table sizes, and compared to the evolution of the number of AS during the same period, the routing tables size is increasing much faster than the number of AS, what corresponds to a strong decrease of the Internet scalability and performance. This is this last exponential increase that will be analyzed in the following to find out its reasons, and we will evaluate if IPv6 can stop such an increase to make it – at most – follow the curve of the increase of the number of AS.

---

[7] TELSTRA web page: http://www.telstra.com/

[8] RIPE NCC web page: http://www.ripe.net/

[9] BGP tables for MAE-EAST and MAE-WEST are available at the following URL: http://nitrous.digex.net/

[10] Tables of prefixes for MAE-WEST as well as BGP tables are available at the following URL: http://www.rsng.net/rs-views/mae-west/

[11] AS: Autonomous System. It consists of a single Internet domain managed by a single entity.

[12] CIDR: Classless Inter-Domain Routing

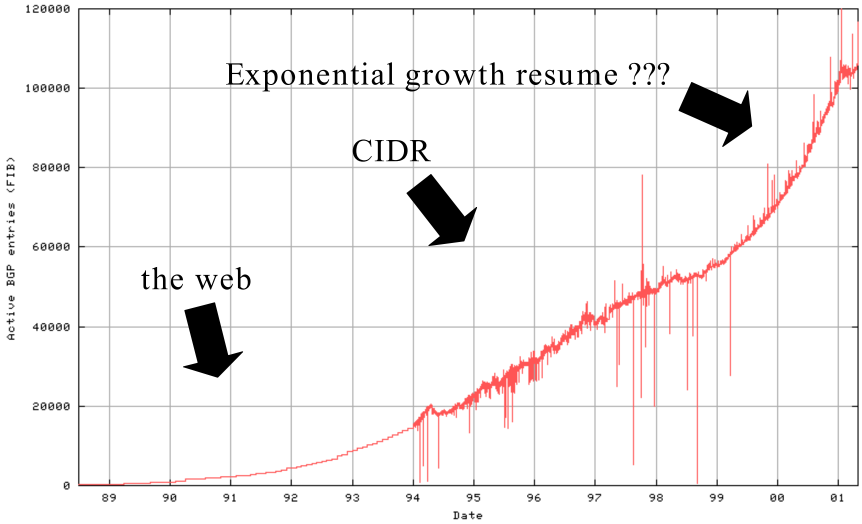

Figure 1. BGP tables growth (1989-2001)

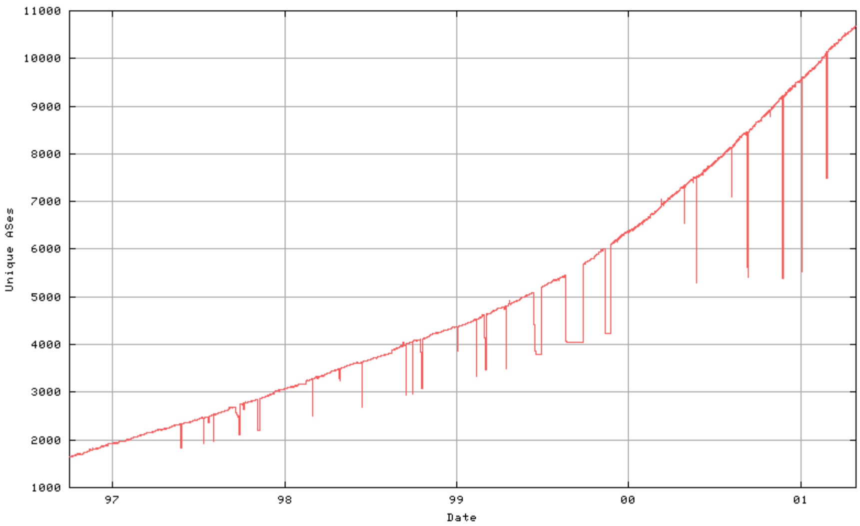

Figure 2. AS number growth

3. Analysis of the Evolution of Routing Table Sizes

3.1. What Helps to Reduce Routing Table Size and Improves Scalability

This section starts a deep analysis of the reasons why since 1998 the routing table are increasing so rapidly? For that, the paper starts with the possible solutions for reducing routing tables. In particular, it is not clear why CIDR that has so efficient results before 1998 does not provide a suited or efficient enough solution today. CIDR was in 1994 the answer of the IETF to the exponential increase of routing tables that started to limit the Internet scalability. Before CIDR, IP(v4) addresses were managed using classes (A, B or C) [13] statically managed. Such classes were responsible of so much wastes of IP addresses because each time a single IP address is needed, a whole class had to be allocated, even for an AS consisting of very few hosts. With CIDR, it is now (almost) possible to allocate to any system administrator the exact number of IP addresses he requires[13]. And to fight BGP table increase, it is now possible to aggregate prefixes of adjacent addresses (as shown on Figure 3 for carrier #1) to limit the number of entries in routing tables, and then increase the routing performances. As well, it helps to reduce the load of the BGP protocol by reducing the number of routes (prefixes) to advertise.

But even if CIDR helped and is still helping to reduce routing tables sizes, it is now not sufficient. One solution found by carriers to limit routing tables is to force routers to aggregate some addresses segments even if they are not exactly adjacent, so creating some small addresses segments that cannot be reached, also called “black holes”. Figure 3 depicts one example where carrier #2 aggregates two segments not exactly adjacent, even if the addresses in between are then advertised by routers of this carrier, these addresses being not accessible by these routers. These addresses in between are possibly not reachable from some parts of the Internet. Of course, such a way of increasing scalability and performances (in term of routing speed) decreases the reliability of the IP service. But the speed of routers can be so much decreased by too large routing tables that it is now essential for carriers to increase this routing speed, and then to find the right trade off between routing speed and reliability.

The other important improvement of Internet scalability is due to the new technologies of routers. New routers working at line speed integrates more and more advanced technologies as optical components, faster memories, new switch fabrics, etc. The increase of routers technologies helps to limit the decrease of performances due to the increase of routing tables. Also, routers builders designed and developed

[13] This was true as long as the addresses space was not exhausted. Today, it is very difficult to ask for new addresses and to get them adjacent to the segment of addresses that we already got from the organisms in charge of IP addresses allocations: ARIN (American Registry for Internet Numbers), RIPE (Réseaux IP Européens), and APNIC (Asia Pacific Network Information Center)

new techniques for forwarding and routing packets. This is the case for CISCO[14] for instance that developed new mixed routing / switching techniques. These solutions were NetFlow [3] few years ago, and CEF (CISCO Express Forwarding) nowadays. It consists in considering packets as part of a flow (mainly TCP flows in the Internet) instead of considering every packet independently. Then, it is just needed to route the first packet of every flow, to establish a switched path from the ingress to the egress port for this flow, and then to switch all the following packets of this flow. Such kinds of mechanisms avoid to look in routing tables for every packet, so reducing the impact of the increase of routing tables sizes. Also, the way routing tables are organized can improve the performances of research algorithms. For instance, JUNIPER[15] designed a tree based organization of addresses, with bounded depth. Currently, they manage to find any IPv4 address in 16 steps at most. It then makes the process for searching an address in routing tables very fast, with very low variability. Scalability and performances of the Internet are thus well improved.

3.2. What Makes Routing Table Size Increase

Even if a lot of advances have been made since 1994 to improve the Internet scalability, it is not sufficient to stop the current huge growth of routing table sizes. The first reason is certainly due to the growing number of NAT servers used by network / system administrators that do not have enough IPv4 addresses for all their machines. NAT [11] [12] is used to make address translation. It means that a whole computer network behind a NAT server will be addressed by only 1 or 2 public and routable IP addresses. The NAT server is then in charge to deliver packets to the right receiving machines addressed by a private IP address, valid only on this network and not known outside. [5] presents some of NAT issues. But dealing with BGP behavior and routing tables, it means that the increase of NAT servers in the Internet makes the number of 31 bits long prefixes increase very fast in routing tables. Prefixes are then longer and longer, more and more numerous, and less and less easy to aggregate. It makes routing table size increase very fast.

Another large issue making routing tables grow is due to multi-homing. In fact, for improving the availability of their Internet access, to improve dependability and optimize QoS, system administrators, more and more, use several accesses to the Internet, as depicted on figure 3 where a LAN is connected on both carrier #1 (or ISP #1) and carrier #2 (or ISP #2). The issue with such a behavior is that it makes the hierarchical distribution of addresses difficult to maintain as every ISP allocates addresses belonging to its addressing plan. This LAN then belongs to 2 AS, and the initial tree based structure of IP addresses is broken. It then increases routing tables sizes.

The last issue that makes routing tables grow is related to the un-constancy of users that often change from carriers / ISP with their own address space. The problem is

---

[14] CISCO web page: http://www.cisco.com
[15] JUNIPER web page: http://www.juniper.net

that the first time, the address space has been allocated to respect the tree based hierarchical structure of the Internet addressing plan. But when people are changing from carrier or ISP, the hierarchical structure is broken what makes routing tables grow. In fact, the aggregation mechanisms that were working are no more working. The number of entries in routing table increases as well as the length of prefixes, making the research of addresses in routing table more complex and slower.

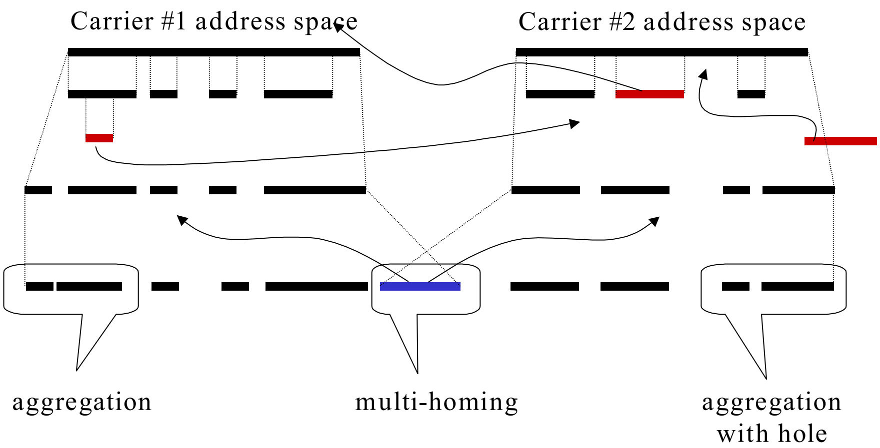

Figure 3. BGP aggregation with CIDR / Multihoming / Black holes

3.3. Analyses of Prefixes in Routing Tables

Finally, after having analyzed several reasons of the decrease and increase of routing table size, a qualitative analysis of all prefixes in some actual BGP tables has been performed. It is supposed to quantify what are the main reasons of the fragmentation of the IP addressing space. The results we got are the following. Address space fragmentation is due to:

- Multi-homing: between 20 and 30 % (depending on the traces) of BGP tables entries are due to multi-homing.
- Failure to aggregate: between 15 and 20 % of prefixes cannot be aggregated as they do not respect any hierarchical structure. This can be due to people having moved from one ISP to another, or to address allocations not respecting the Internet hierarchical addressing. This problem is also favored by the presence of NAT servers and /31 prefixes difficult to aggregate with other shorter prefixes.
- Load balancing: between 20 and 25 % of prefixes are advertised on several links, then multiplying the number of entries for routing tables. This is mainly the case in core networks and core routers. By definition, the Internet has to be

meshed, the more being certainly the best as it improves network availability, dependability and QoS by avoiding to take advantage of congested links. To dynamically solve these kinds of conflicts between several available path, routing protocols integrate load balancing mechanisms. The problem here is that for increasing the dynamic features of traffic engineering processes, it makes routing tables grow, then reducing the Internet scalability and performances.

Finally, this is an average of 75 % of the IPv4 address space that is fragmented. It means that the size of BGP tables could be divided by 4, thus facilitating and speeding-up the research of addresses in routing tables.

4. Threats for IPv6

Recalling that IPv6 has been designed to solve the Internet scalability issues, this part analyzes if IPv6 can provide a suited solution for solving the actual current issues described in section 3. It also indicates all the threats for IPv6 that can delay its global deployment in the Internet.

a) NAT

Even if NAT is a "dirty" solution having many drawbacks [5] (especially dealing with security or peer-to-peer applications[16]) – and in particular the one of making routing tables grow because of many /31 prefixes very difficult to aggregate with other prefixes – it has the strong advantage of being available, easy to deploy, and very cheap. It provides a solution for anybody to be connected to the Internet, even if he does not have enough public addresses. It then solves the problem of IPv4 addresses exhaustion. Even if NAT does not provide any solution against routing tables growth, it does not have only bad sides. For example, it can be useful to improve the Internet performances by contributing to the splitting of connections. In fact, it is then possible to change the parameters of the used transport protocol to adapt to the characteristics of every network crossed with, each time, a dedicated protocol configuration. Such a principle has been described in [2] that also demonstrates the benefits of such an approach.

b) Routers Technologies and Techniques Advances

As described in 3.1, new routers technologies and techniques make routers faster and they are more and more able to cope with very large routing tables. With such new devices, the Internet scalability – in term of performance – will not be an issue anymore. Even if NAT does not help for routing table size, the advances in routers should make things better. Finally, NAT and recent advances in routers could replace IPv6 whose effects do not seem to be fully suited to current problems has shown in section 3.

[16] One of the advantage of IPv6 is to be cleaner than NAT, and to maintain the end-to-end IP service, then allowing the use of peer-to-peer applications, for instance, that are very popular applications since few years

c) 3G Mobiles and the Ubiquitous Internet

The arrival of 3G mobiles (as UMTS) is often presented as the main reason for deploying IPv6 pretty fast, first because UMTS service providers stated that the growth of UMTS mobiles will be similar to the one of cellular phone. Today this assumption does not seem to be correct. At least, the predictive planning will be strongly delayed. Another reason going in the direction of IPv6 is related to the mobility aspect of UMTS devices and users. And then the auto-configuration mechanism of IPv6 that dynamically allocates IPv6 addresses to users depending on their geographical situation on Earth seem to provide a good solution for mobile users.

But the problem then is to know if the IPv6 auto-configuration mechanism can be fast enough to cope with such mobility, especially if the predictions about 3G mobiles are right and if billions of such devices are available one day. In such a case, the mobile network will have to multiply the number of terrestrial equipments – in particular the number of SGSN[17] – to be able to handle all users and all traffic. In such a case, the move from one cell to another that infers the advertising of new routes between equipments in the mobile network and the update of all routing table in SGSN can take up to 45 seconds. Such a duration is of course not suited for interactive, continuous and real-time services or applications. This is mainly a problem of routing / naming tables stability. It could be even worse if the 3G mobile service is extended, one day, between several 3G mobiles operators, thus allowing users to get a seamless service when traveling from one country to another, or more simply, from one area covered by an operator to an other area cover by another operator. In this case, the different mobile networks will be interconnected by the classical wire Internet, and mobility capabilities will be limited by the BGP tables stability. For instance, the current measurements shows that a change in a routing table takes around 5 minutes to be advertised in all routers of a tier 1 network and more than 30 minutes in the global Internet [8] [9]. At the end of this evolution process is the ubiquitous Internet where the Internet service will be provided seamlessly between any kinds of communication infrastructures, ISP and carriers all around the world. The recent “Open Mobile Alliance” consortium [15] between (almost) all the wireless devices vendors to elaborate a set of standards allowing the interoperation of devices of different brands is a big step toward the ubiquitous Internet. In such cases, IPv6 is then not necessarily better suited to provide services requested by 3G mobiles – or several steps forward the ubiquitous Internet – than IPv4. With billions of 3G and or wireless mobile devices connected to the Internet, a huge growth of routing tables – internally to the 3G mobile network as well as in the global Internet – will arise. This growth should be so important that routers will not be able to compute them fast enough to at least maintain the current scalability and performance level. In such a case, IPv6 that can contribute to the explosion of 3G mobiles by allowing the allocation to everybody of native and routable IP addresses, will also necessarily lead to a dramatic routing table increase that can possibly make the global, possibly ubiquitous, Internet stop.

---

[17] SGSN: Serving GPRS Support Node

d) Migration Process

Another important reason why IPv6 deployment is more and more delayed is related to the complexity and the cost of IPv4 to IPv6 migration, with a risk for a more or less long period to get bad services corresponding to a period where IPv6 services can be of second importance or priority (during the migration period, IPv6 traffic can be tunneled in the IPv4 Internet networks, for example, with software routers on some points of the network). As well, dynamic IPv6 address allocation mechanisms are not yet fully defined, and if you make a choice today there is a big probability that the choice you made will not be the one the IETF or IPv6 forum will do. There is then a risk to redo parts of the migration process many times depending on IPv6 evolutions.

e) Multi-homing and Users Un-constancy

The impact of multi-homing and un-constancy of users on routing tables, and then on scalability, has been presented in part 3.2. Up to very recently, there was no solution to "fight" this issue with IPv6. The first proposal for IPv6 multi-homing was issued in October 2001 [BLA01]. The problem is that it is just an evolution of [1] for multi-homing with IPv4, that proved to have a quite limited positive impact on routing tables growth (cf. Figure 1). In fact, the objectives of this solution is quite limited. It just uses tunneling techniques (not native IPv6). So, it is not clear whether this approach can provide an acceptable solution to the general multi-homing issue? Past experiences with IPv4 seem to prove the opposite. IPv6 seems then not to be able yet to solve the current scalability issues of the Internet, what was initially one of its 2 main objectives. This is certainly one of the main reasons for Tier 1 carriers not to deploy IPv6 yet given the amount of extra entries in routing tables due to such a characteristic of the Internet topology.

Finally, IPv6 appears as a very suited solution addressing all the Internet issues of the early 1990's, but 10 years later it seems that it does not bridge the new gaps with the new current issues of the Internet (as multi-homing for example).

5. Conclusion

This paper presented an evaluation of IPv6 mechanisms compared to the current Internet problematic, in particular the one of the scalability issues of the global Internet. More specifically, this paper focuses on the analysis of BGP routing table whose size has a strong impact on performance, QoS, and scalability of Internet services.

It then appeared that IPv6 is pushed by 3 main kinds of people:

- 3G mobile operators that need a huge number of addresses without breaking the end-to-end IP model;
- New Internet coming countries, especially in Asia that need a huge amount of IP addresses because of their huge population (and that were not involved in the first rounds for IPv4 addresses allocations);
- New applications as peer-to-peer that require an end-to-end communication model at IP level. Such applications are not able to cross NAT servers as they

are not able to build the tree of users if some of them have non routable addresses.

But this paper also showed that IPv6 cannot address the problem of the current huge growth of routing tables, and is then not able to provide a solution for the Internet scalability. In addition, core network carriers, mainly located in the US, almost do not have IPv4 addresses limitations issues, as they where involved in the first rounds for IPv4 addresses allocation. They do not suffer from IPv4 addresses exhaustion. Finally, these kinds of issues concern mainly end users and not carriers that are most of the time at the beginning of evolutions in the Internet in terms of communication infrastructures.
Finally, this situation seems to show that, at least, the Internet backbone, mainly located in the US, will remain in IPv4 for a very long time. The IPv4 to IPv6 migration process will be pushed by Asian countries, recently coming to the Internet, and countries involved in 3G mobiles design and deployment (Japan, western Europe). The US that got 75% of the IPv4 address space, and that are not really interested by nationwide 3G mobile services (because of the demographic distribution of the country) will continue to promote IPv4 for a long time. Between end users that are promoting the arrival of IPv6, and large Tier 1 carriers that have more interest in continuing using IPv4, the IPv4 to IPv6 migration process should last for decades. IPv6 is then not very well engaged because of some issues of the current Internet context. It can finally be the new enhancements in routers technologies and techniques – that make them less sensitive to routing table sizes – that can save IPv6, by solving the scalability issues.

References

1. T. Bates, Y. Rekhter, "Scalable Support for Multi-homed Multi-provider Connectivity", Request for Comments N° 2260, January 1998
2. P. Berthou, T. Gayraud, P. Owezarski, M. Diaz, "Multimedia Multi-Networking: a New Concept", To be published in Annals of Telecommunications, 2002
3. CISCO corporation, "NetFlow Services Solutions Guide", http://www.cisco.com/univercd/cc/td/doc/-cisintwk/intsolns/netflsol/
4. J. Hagino, H. Snyder, "IPv6 Multihoming Support at Site Exit Routers", Request for Comments N°3178, October 2001
5. M. Holdrege, P. Srisuresh, "Protocol Complications with the IP Network Address Translator", RFC 3027, January 2001
6. C. Huitema, "IPv6 - The New Internet Protocol", Prentice Hall, October 1997
7. C. Huitema, "Deploying IPv6", Conference on Deploying IPv6 Networks, Paris, France, November 20th – 23rd, 2001
8. C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian, "Delayed Internet Routing Convergence", Sigcomm 2000
9. C. Labovitz, A. Ahuja, "The Impact of Internet Policy Topology on Delayed Routing Convergence", Infocom 2001
10. P. Loshin, "IPv6 Clearly Explained", Morgan Kaufmann publisher, January 1999

11. P. Srisuresh, M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations", RFC 2663, August 1999
12. P. Srisuresh, K. Egevang, "Traditional IP Network Address Translator (Traditional NAT) ", RFC 3022, January 2001
13. W.R. Stevens, "TCP/IP Illustrated, Volume 1: The Protocols", Addison-Wesley, 1994
14. J.W. Stewart, "BGP4: Inter-Domain Routing in the Internet", Addison Wesley Longman, Inc., December 1998
15. The Wall Street Journal, "Open Mobile Alliance hopes to succeed where WAP failed", June 17th, 2002