A Layer 2 IP sharing device based on the NPT mechanism

Wen-Sung Chen Wen-Kang Jia
Information Management Department
National Kaohsiung First University *Ming-Chuan University*
wschen@ccms.nkfust.edu.tw s2756124@ss24.mcu.edu.tw

Abstract*-In this article we introduce a design which named Network Port Translation (NPT) that can substitute the function of Network Address Translation (NAT) and can get better performance in translation. The different concept between NPT and NAT is that NPT worked in Layer 2 and NAT worked in Layer 3. Because the design of NPT is "Translation on Demand (ToD)" to the Service Port Number and do not execute any action in translation to the IP Address, the packet will forward directly and save a lot of processing time. NPT can improve the performance of translation significantly which include translation latency, packet throughput and reduce the loading of the device. At last, NPT also can support a high level network security environment which includes all internal and external networks by controlling the access action from client to client.*

Keywords: IP Sharer, Network address Translation (NAT), Network Port Translation (NPT), Translation on Demand (ToD), Network Performance.

1. Introduction

The Internet used Transmission Control Protocol/ Internet Protocol (TCP/IP) to link all the hosts that can exchange or share information between each user in the world.

In IPv4 protocol, it used 32 bits length to address all the network devices. The device can be computers, printers, routers, exchangers, gateway or other networking devices which can be identified to the source or terminal in the Internet. Because the technology developing of the Internet is very fast, the number of 2^{32} addresses that specific defined in the beginning is not enough to assign to all network devices. It is difficult to make all network devices has its unique global IP address. In normal case, enterprise can't get enough global IP address compared to the number of network devices which the enterprise has. In the other hand, having unique global IP address means can be accessed by any source in the Internet. The enterprise to avoid the situation such like security loss or palsied in move which caused by Internet hacker, the simplest way is using an IP address shared gateway to separate the Internet and Intranet. The device translates the internal IP address to Internet global IP address and the working principle is based on the NAT mechanism.

The next generation of addressed method which called IPv6 will have 2^{128} addresses space. But it will need several years at least to upgrade the routers and related devices to support the IPv6 standard. The Internet users will use the IP shared device to save the quantity of global IP address in the transition time. It means that NAT cab extend the life of IPv4. Because the wideband technology used in the Internet, the Internet traffic is increasing very fast. The NAT devices that we used before will make the bottleneck of performance in the Internet. The goal of this paper is trying to design a layer 2 NPT device for improving the performance of transmission in the Internet.

2. Literature review

2.1. Traditional NAT

The NAT is just a generic name and it also includes extended application of Network Address Port Translation (NAPT) [6]. The NAT is defined in RFC-1631 standard originally and only translated the IP address in the original design. Its goal is to hide the private IP address for the Internet security reason. This kind of IP address translation was defined to the Basic (Traditional or Static) NAT mode [4] until RFC-2663 standard appeared and defined the NAPT in IETF RFC-2993 additionally. It is a translation method which translates the port number and IP address in the same time. Using this mode, NAT can share the external IP address to the inter-multiple private IP address was defined in RFC-1918. It can use by enterprise at will and no need to apply for assignment from the Global IP Registry. The enterprise must to avoid the packet or the Routing Information from the Private IP address flow out to the external network. It means that local hosts can only link to the local hosts in the internal and the external network can not see those internal hosts. Those private IP address can repeat used by different

units and reach the goal of saving IP address. If the private IP address needs to access the information in the external network, it must use the NAP mechanism to share the IP address [2]. The framework of NAT network link is showed in Figure (1). It shows that the internal network user by way of the NAT translation process in IP shared device to link the server in the external network.

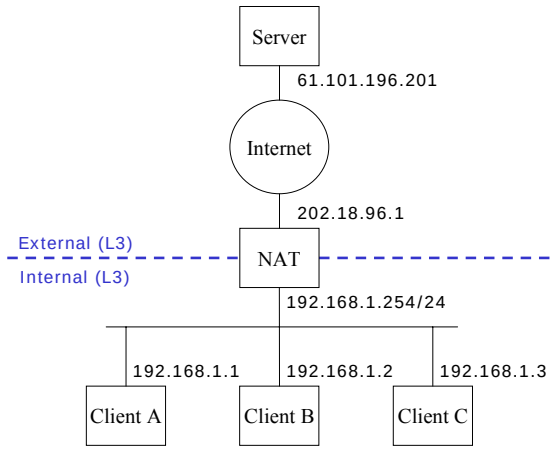


Figure (1) The Architecture of NAT Network

2.2. Proxy ARP

The Proxy ARP is usually used for answering some network host ARP request behind it on the router, and then passes on the packet for this host. Comparing with Routing, the hosts lying on different network segment are the place that broadcast can't be reached. By using proxy ARP, it can make the different network segment looks like in the same network domain without separate the subnet. It is a cheating method by way of replacing distinguished the subnet through the router. Because the source host will think that the destination host of the communication is the router itself, but in fact, the destination host is in another network segment [9].

The proxy ARP is used in the environment of remote connection network like dialing. For example, there is a remote access server (RAS) in local network and offer other hosts to dial in and connect to the local network. When local host want to find the carried ARP request from RAS because its broadcast packet can't reach the destination and then the RAS replace the remote host and answer its ARP request. At the same time, RAS also transit the packets between the remote hosts and the local host at the same time.

The other very useful occasion of proxy ARP is solving the situation of IP insufficiently with separated sub network. Such as a Class C IP network, when it subnets into 32 network segment (subnet mask 255.255.255.248) and use it, each subnet will only have five (Gateway IP address excluded) available IP addresses, there is about 37.5% IP address waste. If we will divide into the subnet mask 255.255.255.252 under such a small range, two IP address taken up by the router excluded , there is only one IP can be used and will waste the address location space up to 75% and it seems meaningless. Using proxy ARP can avoid this problem. Commercial ADSL network, Using PPPoE to simulate dial-up to assign the Dynamic IP address to subscriber is also using the concept of proxy ARP.

3. Research and Design

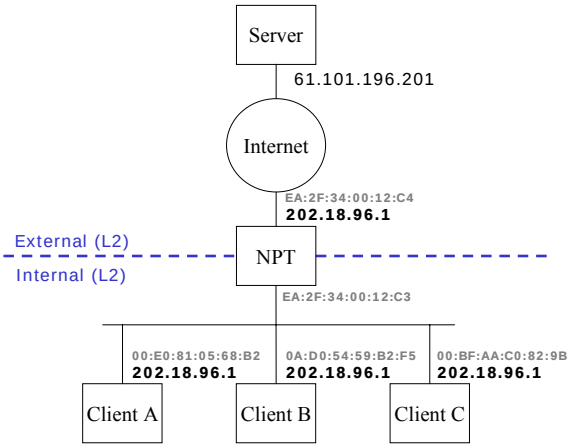
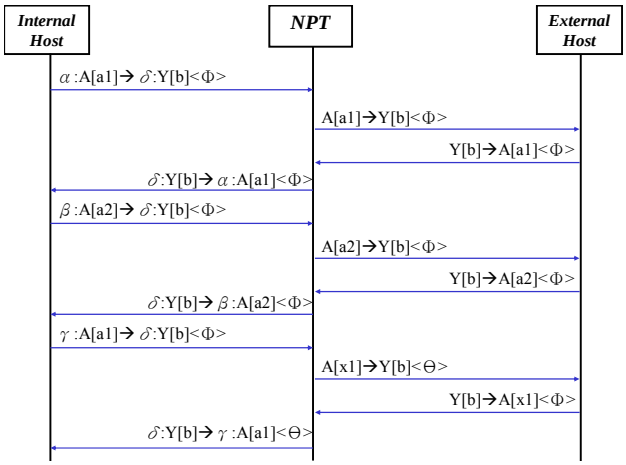


Figure (2) the Architecture of NPT Network

The structural design of the NPT network is show in Figure (2), NPT IP shared device is Layer2 network equipment which works in the Bridge Mode. The device itself does not take up any IP address. The Default Gateway of hosts in the internal network also points to the routing equipment that the network exported to and do not point to NPT IP shared device. NPT IP shared device does not translate any IP address. Because the internal network host has already used the external unique IP address directly, every host covers and uses the same unique IP address repeatedly, the inside network can't be differentiated by IP address. All packets will be changed to the (Ethernet, IEEE 802.3) 48 bits MAC address or the wireless network (Wireless LAN, IEEE 802.11) for the only differentiating way. The original thought of this design considers that traditional NAPT IP of structure shared device must check their source IP and translation procedure, and recalculate the packet and check the checksum again. It will waste a large amount of operation efficiency. It may become the bottleneck of packet flow in the high loading network. The IP shared device under the NPT also has the same procedure of checking in each packet, but only a small amount of packets need to translate its source port number and the source IP will never need to translate. The majority packet does not need translating and recalculating the packet checksum again. Reducing those operations will improve its efficiency for those NPT IP devices or the routers enabled by NPT features. In

high loading network, the NPT IP shared device can perform similar efficiency with the switch, and to eliminate the bottleneck of packet flow.

This kind of design (just Translating on Demand; ToD) for port number, its translation's probability, will depend on the probability of the following conditions: Two hosts in the internal network linking to the same server in the external network at the same time, and using the same service port (Destination Port) for the external server, and two hosts of the random selected source port also the same.



α: Local Source MAC Address (Client A)00:E0:81:05:68:B2
β: Local Source MAC Address (Client B) 0A:D0:54:59:B2:F5
γ: Local Source MAC Address (Client C) 00:BF:AA:C0:82:9B
δ: Local Dest. MAC Address (NPT) EA:2F:34:00:12:C3
ε:Global Dest. MAC Address (NPT) EA:2F:34:00:12:C3
A: Local/Global Source IP Address (Client A/Client B/Client C/NPT) 202.18.96.1
Y: Global Destination IP Address (Server) 61.101.196.21
a: TCP/UDP Source Port Number (Internal)
b: TCP/UDP Destination Port Number (Internal)
x: TCP/UDP Source Port Number (External) --Random Generation by TCP/IP Socket
Φ: CRC-32 CheckSum
Θ: CRC-32 CheckSum (Regeneration by NPT)

Figure (3) the packet flow of NPT process

In above case, it will translate the source port and select a random free port again and write down on NAT-Table, for the incoming packet to do the correct translation. The probability of "source port collision" is less than 1/216 when two computers link to the same external server in theory. This probability will increase when network hosts increase or frequently used the specific server. If there are the upper limits reaching 256 hosts in the inside segment, the collision probability is less than 1/256. (It is an ideal value, and we do not consider that all machines open the same source port to transmit its packets. In other words, there is 99% of the session/packet doesn't need translation and keep its transmission transparent to the NPT devices. All the IP packets just change the source MAC address in DLC header when passing to NPT device.

The Figure (3) states the packet flow of NPT process. When the first packet is sent from the host of inside network MAC address α (the IP address A in the outside network is valid too), passing through NPT device, the packet header have not been changed at all, and conveying it to the destination host Y directly. Similarly, the packet that the destination host response is also not translated either, passing through NPT and get back to host α (A) directly. When this session still keeps active, a host β (use the same IP address A) set up the connection with destination host Y of the external networks too. The source port number that β used is selected at random and set to a2. It can be differentiated with the connection using the a1 source port number before by writing down the inside record in NPT. Setting up the second record in the NPT table, and transmitting this packet to the host Y directly still don't need any translation too. When the host γ comes from inside and try to set up the third sessions with outside host Y, NPT will checked the inside two records, and if find it conflict with the first record It is unable to reach correct host γ while the packet return from the third sessions, and reached the wrong host α. In order to solve this problem, it must replace the source port number conflicting as x1, and it could avoid collision and pass on to correct host γ when the packets transmit back. It is similar to the traditional NAPT, but it does not need to change the IP address. Because the port number of the source has already changed, the value of IP header checksum of the packets will need to recalculate again.

On the design of the NPT IP sharing device, it must ensure that the hosts MAC addresses in the internal and external network are separated effectively. There is only one unique MAC address exists for the internal and external network. When the inside host broadcasts the packet to look for the outside host with ARP Request, the ARP Proxy will replace α of the source MAC address to ε of the NPT external MAC address. When the response packet from outside host getting back to the inside network, it needs to change the source MAC address of the outside host to the MAC address δ of the NPT inside, and replace the source MAC address ε as α. The original inside inquirer host will use the internal MAC address of NPT device to do the connection with hosts in the external network.. In addition, the ARP broadcast packet of the inside network just need to make above-mentioned packet flow and ToD process when passing on to outside network. Unless we agree that the hosts in the internal network can be communicate with each other, it needs not to translate to the other physical port of the internal network.

4. Performance Analysis

4.1. The Probability of "ToD" Simulation

This study uses System Dynamics simulation tool -- Vensim™ to set up NPT operational emulation model, in order to simulate the probability of the packet needs to be Translating. There three controllable decision parameters in this design: the number of internal network host (SP), the number of the external network host (DIP), and the number of external network host (DP). The simulation situation is using single IP address to pass the NPT device in order to the source port that is produced for 100 internal network hosts at random. Then 10 kinds of service types (DP) in 100 external network hosts (DIP) to set up TCP/UDP session at random (suppose that each host will offer all service types), and keeps the activity state alive. If we have 65,000 times of simulation, it represents that will set up [the inside network host*65,000] session record in NPT Table. But the NAT Table of the general commercial network equipment is usually designed for only 2,048-4,096 connects limitation. It means that only keep the same number of session record stored space. In fact, if the session record is too big, it will cause the searching time longer. It is relatively impractical in practice.

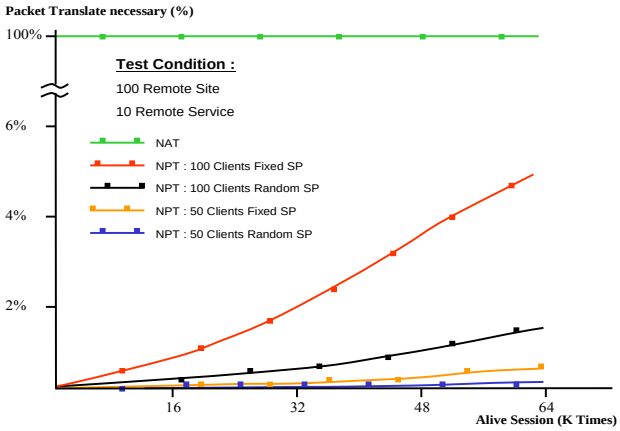


Figure (4) the probability of NPT ToD

This research sets the number of simulation is 10,000 times, when the inside number of network host sets to be 50, there are 500,000 times connection simultaneously. It has been very enough at the extreme state of this kind of pressure testing scale. The simulation result is in Figure (4). At the end of simulation, only about the sessions of 26,730 times (about 5%) need to be translated, it shows that the efficiency of translating is promoted apparently.

But this simulation has two restrictions: First, the sessions that are translated after collision, has not been considered the probability that collision with the another sessions that are set up newly in a moment takes place again, this part of probability should be very low, so this research do not consider this situation; Second, on most TCP/IP Protocol Stack design in various OS platform, its source port number don't generate by random, taking Windows as an example, it increases its port number from Port 2000. If inside network all host computers start the machine at one time and do external networking, the collision rate will rise very fast. This situation will lead to this simulation distorted, but it will still reach in unanimity while reaching the capacity limit.

4.2. Performance Simulations

The NPT device will also set up one set of NPT mapping Table dynamically during the process of translating, abbreviated as NPT Table. The purpose and functionality are the same as NAT Table, but the data structure of the table is different. When passing on, it must check the packet of associated session records that NPT Table have first. If yes, it needs to judge those old session record for its no collision or collision state. If not, it will pass on the sessions that already had and state the collision status after assign a source port randomly and record it in NPT Table. When packet passing on from outside to inside, NPT devices also perform the same checking procedures, it needs to check the content and judges whether it is translated before. If yes, it use the source port and MAC address for originally state before then pass on. If not, it will refuse this packet pass on and drop it. All of these translations in MAC layer (Layer 2) will be taken place while sending the packet to the interface of the network. It will not deal with the binding problem of the IP layer and t the MAC layer. We can consider this design that combine ARP Table and NPT Table as one directly. Refer the Figure (5), the NAT Table record needs pieces of 18 Bytes at least to show the local source IP address (32 bits), destination IP address (32 bits), local source port number (16 bits), destination port number (16 bits), global source IP address (32 bits), and the global source port number (16 bits). It also needs the ARP Table to show the correspondence of local IP address (32 bits) and MAC address (48 bits) in addition.

Comparing with the design of NPT Table, records of the NPT table needs only 16 Bytes to show local MAC address (48 bits), destination IP address (32 bits), local source port number (16 bits), global port number (16 bits), and global source port number (16 bits). The global source port number that used will be need this address field in the collision condition but most of the time does not conflict (can insert 0x00 values). In addition, NPT device can lookup NPT Table and find MAC address for transmission the packets directly. It doesn't needs to analysis the MAC layer of network addresses of ARP Table and encapsulation act. In fact, the ARP Table will be only responsible for the analysis of the part of the external network and no need deal with the internal network. It seemed that the data structure, the search efficiency of NPT Table, and the translation efficiency all will show better performance than NAT way.

Following we discuss the performance of NPT device platform. We compare the performance of NPT device with NAT device using a one MIPS CPU processor. NAT program codes are executed and assume the search time and CPU time of NAT table will be influenced by the memory capacity of the NAT device. The more entry records in the table, the longer search time and CPU time to take up. The ARP Table lookup's procedure also has the same condition. In addition, each packet checksum recalculating also increase latency and consume more CPU power. Again, comparing with NPT device under the same condition, we suppose NPT Table consume same search time with NAT Table. But the searching time of the ARP table can be totally neglected, we use the 5% packets that need to recalculate for this comparison by packet delay time and taken up CPU utilization ratio.

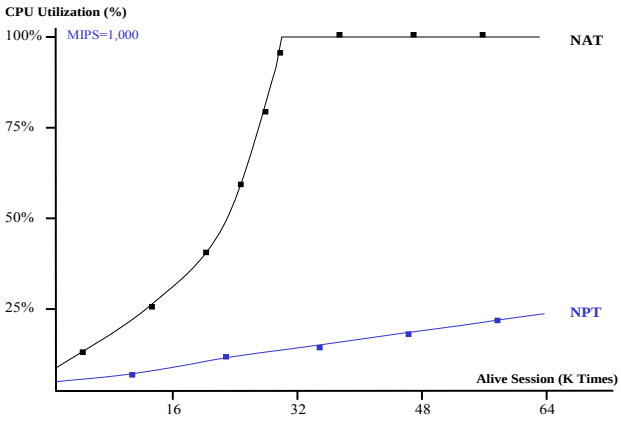


Figure (5) the CPU utilization comparison between NAT and NPT

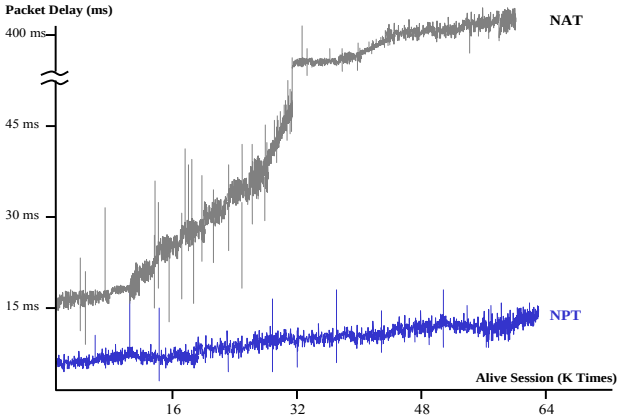


Figure (6) the Forwarding delay comparison between NAT and NPT

The simulation results are summary in the Figure (5) and Figure (6). Those two figures show while setting up to the connection reaching 30,000 times, CPU utilization rate of NAT has already been fully loaded, at the same time the forwarding delay time increases too. But NPT, while setting up to the connection reaching 64,000 times, CPU of NPT utilization rate only about 25%, forwarding delay time that packet passing on can keep up about 10ms; But please notify those assumption data simulated in the model aiming at the observation the performance difference between NAT and NPT devices. Data itself seems no too much meaningful.

5. Conclusion

The NPT mechanism can not only be realized in the single function of IP sharing device but also can be combined and built it into the routers, switchs, firewalls, IA, home gateways, and WLAN access points etc. These products will enable an NPT optional added function that we can use.

The environment and opportunity for NPT-enabled devices are very broad. It will especially suit for concerning with inside invade or paralyzed attack, like public library, network coffee shop, computer classroom, or hotspot areas which support network service with the wireless network. Because all hosts of internal network adopt the same IP address, they can not set up the connection as usual and will break the threatening that comes from the internal network effectively. We can say that: Traditional NAT/NAPT mechanism is a procedure that protects the external network accessing to the internal network normally. However, the NPT mechanism is a procedure that both protect the external network accessing to internal network, and internal network accessing to internal network at the same time normally. This design that can break the attack coming from other users is the first contribution for the network service safety on WiFi environment.

If we consider that this design allows normal connection within internal network hosts, it must act with Proxy ARP that simulates and maintains true MAC address and Masquerade pseudo IP that disguise among the host. And this extended function of passing on the packet can also be enhanced into NPT mechanism. If we want put this design into commercial products, this extended part should be a point of follow-up study [10].

Any host of the internal network will continuous receive the ARP broadcast packet that send from any other sites when link with Hub/Switch, and analyze it's content for monitoring the same IP address appears on the same network domain. Once detecting, the operating system will show the "Duplicate IP Warning" message that IP repeated state on the master station. The internal network hosts adopt the design of the same IP address like NPT mechanism will present this IP conflict question in this structure. Although this design uses MAC address for separating effectively and not to influence the normal usage, it is still a problem that must be solved in follow-up study. There are several feasible solutions like disabling this warning message at the end user's operation system and

combining physical network design of physical port in NPT devices in order to collocate ARP mechanism with broadcast of packet that strain others through the "port-base VLAN" function. Via this function, client in each VLAN can not detect and examine the ARP broadcast. The design diagram is as Figure (7).

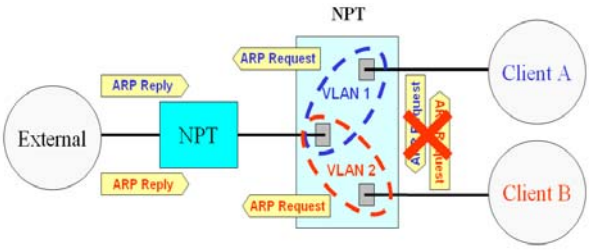


Figure (7) solve the "Duplicate IP Warning" by the "port-base VLAN" design

How to assign duplicate IP address on IP/MAC address management issue in the internal network? There are two potential ways, one is establishing all hosts manually as the same IP address, the other is using special designed DHCP server to send duplicate IP address. The latter seems to be more friendly method to user. How to realize? It also deserves further study.

In the internal network environment which need high loading connection to outside network, NPT mechanism can reduce the time of checking NAT for each packet, IP header replacement, IP/TCP header checksum recalculation, even the time of checking ARP Table that can be omitted too when network device executing the network address translation. The result is very apparent on promoting the performance of network sharing device packet forwarding rate and reducing the CPU loading. It will be valuable to the reference on router, switch, firewall and loading balancer devices. These features are based on the foundations of NPT mechanism with simple data structure and with lower CPU loading. It is the second contribution of this NPT mechanism design.

There are no quite difference between NAT and NPT mechanism on the functionality of IP sharing. Based on NAT Variation Definition by RFC-3489[13], depends on customers' need, the NPT may implement as "Port Restricted Cone Mode" or "Symmetric NAT Mode". But it is insignificance if NPT support on the "Full Cone Mode" and "Restricted Cone Mode". Because these two modes are not considered to use in the IP sharing design originally.

But NPT can solve restriction at some specific protocol which described in RFC-3027. Such as IPSec [7][11], the reason it can not work normally under NAT structure is because of the IP checksum be altered in the packet and it seems to be solved in this structure[4]. The IPSec can operate normally under NPT structure due to its very low probability of the port collision. Even though it still has very small emergence probability and can not avoid up to 100%. Regarding to the restriction on FTP PORT Mode or H.323 with VoIP application caused by external hosts attempting to set up the new TCP connection via new port number[7]. It is the fundamental limitation and NPT is also unable to solve. Moreover, some protocol such as SNMP, RSVP and H.323 also have problems[7], the reason that can not operate normally is because the payload has the IP address information field and can't solve efficiently when passed through NAT[7][8]. This problem can solve efficiently in the NPT mechanism. This is the third contribution of this paper and its continuous refinement can be regarded as the important direction of follow-up study.

This paper does not intend to strengthen network application of IPv4 or delay the retirement schedule of IPv4. But in the coming potential transition period that upgraded to new developing of the IPv6 network, there are the mutual conversion demands between two protocols. The result of this research is an interesting direction that we can expand what kind of application it can use to the thought of translating between IPv4 and IPv6 (NAT-PT)[5].

References

[1] K. Egevang, and P. Francis, "The IP Network Address Translator (NAT)", *IETF RFC1631*, May 1994

[2] Y. Rekhter, B. Moskowitz, D. Karrenberg, G.J. Groot, , and E. Lear, "Address Allocation for Private Internets", *IETF RFC1918*, February 1996.

[3] S. Kent and R. Atkinson, "IP Authentication Header", *IETF RFC2402*, November 1998.

[4] P. Srisuresh and M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations", *IETF RFC2663*, August 1999.

[5] G.Tsirtsis and P. Srisuresh, "Network Address Translation-Protocol Translation (NAT-PT)", *IETF RFC2766*, February 2000.

[6] P. Srisuresh and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", *IETF RFC3022*, January 2001

[7] M. Holdrege and P. Srisuresh, "Protocol Complications with the IP Network Address Translator", *IETF RFC3027*, January 2001

[8] D. Senie, "Network Address Translator (NAT) - Friendly Application Design Guidelines", *IETF RFC3235*, January 2002

[9] J. Postel, "Multi-LAN Address Resolution", *IETF RFC925*, October 1984

[10] S. Carl-Mitchell and J. S. Quarterman, "Using ARP to Implement Transparent Subnet Gateways", *IETF RFC1027*, October 1984

[11] Lisa Phifer, "The Trouble with NAT", *The Internet Protocol Journal*, December 2000

[12] J. Rosenberg, J. Weinberger, C. Huitema and R. Mahy, "STUN - Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)", *IETF RFC3489*, March 2003