

Enhanced IP Resiliency Using Cisco Stateful Network Address Translation

Stateful NAT (SNAT) is a Cisco IOS Software feature allowing two or more network address translators to function as a translation group. A backup NAT provides translation services in the event of failure to the active translator. The result is a more resilient IP network.

The goal is to create a more globally resilient IP network. Networked applications are placing increased demands on the core IP network. Users expect continuous access to servers and data regardless of location. Although the mean time between failure (MTBF) of hardware components has increased, failures can and do occur. Administrative activities can also cause downtime. A resilient IP network offers continuous service, despite failures that may occur.

The concept of a highly resilient IP network is not new; however, this paper introduces a highly innovative approach. The intelligent systems approach creates a highly optimized, resilient IP network where individual component features interact and share services among each other. The result is a network that is inherently more intelligent and less labor-intensive in terms of design and management. Cisco IOS[®] Software is evolving into a more intelligent, shared function system that reduces support costs and increases the benefit and return on investment in network equipment and services.

Network Address Translation (NAT) has been a key Cisco IOS feature since its introduction. It has helped to reduce address depletion and promote Internet growth. NAT has been used to permit

interconnection of private networks, regardless of their use of independent addressing schemes, even when these schemes use addresses that conflict. NAT has also been used to effectively hide networks from outside the administrative domain while allowing predetermined connections to occur. NAT fulfills an important role and will likely do so even as IPv6 is deployed.

Therefore, it follows that enhancement can make NAT even more resilient. This will allow application connectivity to continue, unaffected by potential failures to links and routers at the NAT border. Cisco Stateful NAT (SNAT) provides this enhanced capability.

In the world of IP networking, *stateful* is defined as applying a more global context to the task of forwarding a particular datagram. In other words, there is consideration not just where to forward the datagram, but also understanding about the application state with regard to this datagram. With this knowledge, devices can take action so that potential failures will have less impact on the flow and to the application that is transmitting data. Multiple NAT routers that share stateful context can work cooperatively and thus increase service availability.



SNAT Overview

SNAT allows two or more Network Address Translators to function as a translation group. One member of the translation group handles traffic requiring translation of IP address information. Additionally, it informs the backup translator of active flows as they occur. The backup translator can then use information from the active translator to prepare duplicate translation table entries; therefore, if the active translator is hindered by a critical failure, the traffic can rapidly be switched to the backup. The traffic flow continues since the same network address translations are used, and the state of those translations has been previously defined.

Only sessions that are *statically defined* already receive the benefit of redundancy without the need for this feature. In the absence of SNAT, sessions that use *dynamic* NAT mappings would be severed in the event of a critical failure and would have to be reestablished. Stateful NAT enables the maintenance of continuous service for dynamically mapped NAT sessions. The result is a more resilient IP network.

Phased Release

Cisco is releasing SNAT in phases. Phase I provides a subset of the intended function. Application Level Gateway (ALG) support is not included in Phase I, so protocols that imbed IP address data within the payload of the IP packet will not be able to take advantage of the enhanced redundancy provided by SNAT.

Phase II will provide increased ALG and asymmetric routing support in SNAT.

Protocols and applications supported in Phase I are:

Any TCP/UDP traffic that does not carry source or destination addresses in the payload

- Archie
- Finger
- HTTP
- ICMP
- PING
- rcp, rlogin, rsh
- TCP
- Telnet

Please refer to http://www.cisco.com/warp/public/cc/pd/iosw/ioft/ionetn/prodliit/792_pp.htm for a list of ALG protocol support for Cisco IOS NAT. The following protocols and applications are targeted for support in Phase II:

- FTP
- H225, H245
- PPTP/GRE
- NetMeeting Directory (ILS)
- RAS
- SIP (both TCP & UDP based)
- Skinny
- TFTP

Support for other protocols, which are not listed above may be offered in later releases.



There are additional deployment restrictions for SNAT Phase I. It will only function properly when the return traffic path traverses the primary SNAT router. In other words, asymmetrical routing should be prevented. To ensure return traffic follows a single path to the NAT router, the routing path cost must be adjusted or the BGP metric has to be set appropriately. Phase II will allow for asymmetric routing, which will remove the restriction.

Phase II will include additional support for the following:

- Support for outside NAT pools, using the configuration command `ip nat outside source pool`. SNAT Phase I will only permit inside NAT pools.
- Dynamic entries, which are extended out of static definitions.
- Support for `ip nat inside destination`

Platform Support

Stateful NAT will be supported on all platforms running Cisco IOS Software. Platforms that include hardware acceleration for NAT will benefit, since the mechanism for creating NAT table entries is compatible with the hardware acceleration implementation.

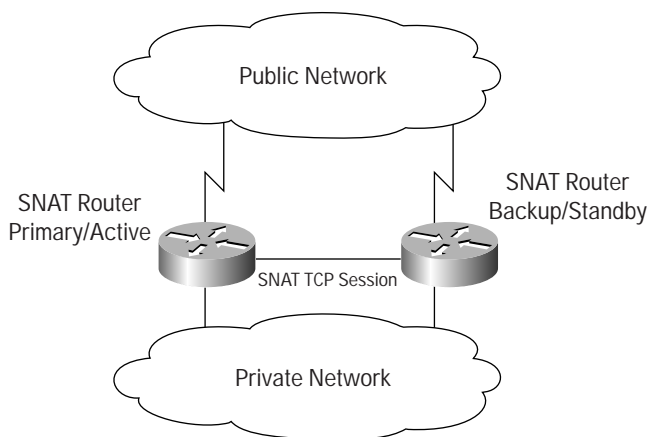
Cisco IOS Software is packaged in feature sets that support specific platforms. To get updated information regarding platform support for this feature, please visit Cisco Feature Navigator at: <http://www.cisco.com/go/fn/>

This application dynamically updates the list of supported platforms as new platform support is added for the feature.

Stateful NAT Protocol

Stateful NAT uses TCP to communicate NAT table updates between the primary and backup NAT routers. See Figure 1: SNAT Functional Diagram. Once configured for SNAT, the TCP session is established between the SNAT peer routers and is used to transmit messages that communicate updates to the NAT tables and maintain session state.

Figure 1
SNAT Functional Diagram





The distributed NAT protocol will ensure that dynamic NAT entries created at the primary or active NAT are duplicated consistently on the backup or standby NAT router. This prepares the backup NAT to take over in the event of a critical failure.

The distributed NAT protocol defines a set of messages that are exchanged between NAT routers:

- **Add Message**—Sent to the peer NAT router whenever traffic flow dictates that a dynamic entry be created locally. The action creates an entry at the recipient's database, based on information in the message (see also Mapping ID).
- **Delete Message**—Sent to the peer when a dynamic-entry is deleted from the local database. The action deletes the corresponding entry at the recipient's database. In SNAT Phase II, the Delete message will be extended to include three types of delete operations:
 - **Forced-Delete**: The recipient will delete the entry.
 - **Delete-Query**: Upon entry-timeout, the Active/Primary that timed out the entry will send this to query the other router as to whether it has received packets later than the NAT router, which is actually running the timer on the entry. In other words, the query permits adjustment of the timer so an entry is not prematurely deleted due to asymmetric flow of traffic.
 - **Delete-Response**: This is sent in response to the Delete-Query. A time-to-restart value is included to adjust the timer on the entry at the Active/Primary that is handling the timers for this entry. Value of 0 in the time-to-restart field will indicate that the recipient has not received packets for this flow later than the Active/Primary.
- **Dump-Request Message**—Sent whenever the router comes up asking for the snapshot of the NAT database from the peer NAT router.
- **Dump-Reply Message**—This is sent in response to the Dump-Request. The message will include the previously learnt dynamic-entries from the router that issued the dump-request plus the dynamic-entries created locally (see also Mapping ID).
- **Update Message**—Distributes application specific information (valid only in SNAT Phase II).
- **Sync Message**—Informs the peer of the local SNAT ID-number. After the SNAT TCP sessions are set up, each router sends the SYNC message. This informs every NAT peer router about the configured SNAT ID-number at the peer.

A consistent set of NAT entries is maintained through the exchange of the aforementioned messages. When a SNAT router fails or reloads, it will request a dump of the current NAT entries from the currently active SNA router upon restoration, and will assume its role in the SNAT group.

Mapping ID

command is used to specify whether or not the local SNAT router will distribute a particular set of locally created entries to a peer SNAT router.

The logic used for distributing the entries created locally to the peer is as follows:

Each dynamically created entry inherits a mapping-id number based on the configuration setting at the point of creation. This comes from the mapping defined on the NAT rule. For example, entries created using rule `ip nat inside source route-map rm-101 pool SNATPOOL1 mapping-id 10 overload` will have id 10 associated with them.



For each Stateful NAT router, a mapping list may also be defined using the command `mapping-id` within the stateful NAT configuration as shown below:

```
ip nat Stateful id 1
    redundancy SNATHSRP
        mapping-id 10
        mapping-id 11
```

Multiple `mapping-id` statements can be used to form a mapping list. The list specifies which of the entries will be forwarded to peers in that group. It provides a way to specify that entries from particular NAT rules should be forwarded.

Show Commands

Use the command `sh ip snat` to get status information about the SNAT processes. In Example 1: Show `ip snat` you can see an example of a router that is configured for IP Redundancy mode and is currently in STANDBY due to the corresponding HSRP group being in STANDBY state. This is because the tracked interface (FastEthernet 0/1) is down.

Example 1 Show ip snat

```
cheney#sh ip snat distributed verbose

Stateful NAT Connected Peers
SNAT: Mode IP-REDUNDANCY :: STANDBY
  : State READY
  : Local Address 10.88.194.17
  : Local NAT id 1
  : Peer Address 10.88.194.18
  : Peer NAT id 2
  : Mapping List 10
  : InMsgs 5210, OutMsgs 5212, tcb 0x82C2DC28, listener 0x826D1790

cheney#sh stand
FastEthernet0/0 - Group 0
  State is Standby
    4 state changes, last state change 1d04h
  Virtual IP address is 10.88.194.20
  Active virtual MAC address is 0000.0c07.ac00
    Local virtual MAC address is 0000.0c07.ac00 (default)
  Hello time 3 sec, hold time 10 sec
    Next hello sent in 0.372 secs
  Preemption enabled, delay min 20 secs
  Active router is 10.88.194.18, priority 100 (expires in 9.796 sec)
  Standby router is local
  Priority 95 (configured 105)
    Track interface FastEthernet0/1 state Down decrement 10
  IP redundancy name is "SNATHSRP" (cfgd)

cheney#sh ip int brie
Interface          IP-Address      OK? Method Status      Protocol
FastEthernet0/0    10.88.194.17    YES NVRAM  up         up
FastEthernet0/1    10.88.161.6     YES NVRAM  up         down
Loopback0          10.88.194.5     YES NVRAM  up         up
Virtual-Access1    unassigned      YES unset  up         up
cheney#
```



This illustrates how SNAT works together with HSRP to achieve improved redundancy.

The current NAT entries can be displayed using the command `sh ip nat translation`. Additional information is shown when the verbose option is included.

NAT entries have been extended to include information about which of the SNAT routers created them, and which router is responsible for the state and timing of that particular entry. The combination of the entry id-number and the SNAT router id-number make each entry unique within the group.

In Example 2: Show IP NAT Translation Entries you can see that SNAT router “cheney” has two entries numbered 3 and 4 that have “left” values counting down from 00:00:35. These entries are “timing-out”. The SNAT router that created the entry is responsible for timing the entry. All three entries are duplicated on a “standby” SNAT router “capefear1” and are flagged, created-by-remote. This indicates that this router is “backup” for these entries as they are not being timed locally.

Example 2 Show IP NAT Translation Entries

```
cheney#sh ip nat tr
Pro Inside global      Inside local      Outside local      Outside global
tcp 11.1.1.3:1173      10.88.194.22:1173 88.1.88.8:10115    88.1.88.8:10115
tcp 11.1.1.3:1174      10.88.194.22:1174 88.1.88.8:10115    88.1.88.8:10115
tcp 11.1.1.3:1175      10.88.194.22:1175 88.1.88.8:47950    88.1.88.8:47950

cheney#sh ip nat tr ver
Pro Inside global      Inside local      Outside local      Outside global
tcp 11.1.1.3:1173      10.88.194.22:1173 88.1.88.8:10115    88.1.88.8:10115
    create 00:00:24, use 00:00:24, left 00:00:35, Map-Id(In): 1,
    flags: extended, timing-out, use_count: 0 nat_id: 1 nat_entry_num: 3
nat_mapping_id: 10

tcp 11.1.1.3:1174      10.88.194.22:1174 88.1.88.8:10115    88.1.88.8:10115
    create 00:00:24, use 00:00:24, left 00:00:35, Map-Id(In): 1,
    flags: extended, timing-out, use_count: 0 nat_id: 1 nat_entry_num: 4 nat_mapping_id:
10

tcp 11.1.1.3:1175      10.88.194.22:1175 88.1.88.8:47950    88.1.88.8:47950
    create 00:00:24, use 00:00:00, left 1d00h, Map-Id(In): 1,
    flags: extended, use_count: 0 nat_id: 1 nat_entry_num: 5 nat_mapping_id: 10

capefear1#sh ip nat tr
Pro Inside global      Inside local      Outside local      Outside global
tcp 11.1.1.3:1173      10.88.194.22:1173 88.1.88.8:10115    88.1.88.8:10115
tcp 11.1.1.3:1174      10.88.194.22:1174 88.1.88.8:10115    88.1.88.8:10115
tcp 11.1.1.3:1175      10.88.194.22:1175 88.1.88.8:47950    88.1.88.8:47950

capefear1#sh ip nat tr ver
Pro Inside global      Inside local      Outside local      Outside global
tcp 11.1.1.3:1173      10.88.194.22:1173 88.1.88.8:10115    88.1.88.8:10115
    create 00:00:47, use 00:00:47, Map-Id(In): 2,
    flags: extended, created-by-remote, use_count: 0

tcp 11.1.1.3:1174      10.88.194.22:1174 88.1.88.8:10115    88.1.88.8:10115
    create 00:00:46, use 00:00:46,
    flags: extended, created-by-remote, use_count: 0

tcp 11.1.1.3:1175      10.88.194.22:1175 88.1.88.8:47950    88.1.88.8:47950
    create 00:00:45, use 00:00:45,
    flags: extended, created-by-remote, use_count: 0
```



Configuration

Configuration for Stateful NAT is the same as regular NAT, but there are some simple additional commands. The first step in defining Stateful NAT is to determine the method of redundancy. Stateful NAT can be configured to work with Hot Standby Router Protocol (HSRP) by leveraging the IP Redundancy API built into Cisco IOS Software. When HSRP mode is set, the primary and backup NAT routers are elected according to the HSRP standby state. Alternatively, Stateful NAT can be manually defined as primary or backup. A sample configuration is shown in Figure 2: Stateful NAT Example.

Figure 2
Stateful NAT Example

CHENEY	CAPEFEAR1
<pre>ip nat Stateful id 1 redundancy SNATHSRP mapping-id 10</pre>	<pre>ip nat Stateful id 2 redundancy SNATHSRP mapping-id 10</pre>

As you can see, the two routers, CHENEY and CAPEFEAR1 form a NAT group. They are designated members of the group by coding the command:

```
ip nat stateful id <id-number>
```

Note: Note the ID is different for the each router. Each SNAT router should have a unique ID number.

See Mapping ID for more information on how the mapping-id is used.

The dynamic NAT pools may be configured on the primary or active router only. Or the pool definitions can be configured identically on both SNAT peer routers. If the pools are only configured on the primary or active SNAT router, then the peer will not be able to create new entries. This is even the case when it has taken over the NAT function. Therefore, it is recommended that you always code identical dynamic NAT configurations for peer SNAT routers.

SNAT Primary/Backup

There are two configuration modes for Stateful NAT: Primary/Backup mode and HSRP mode. Primary/Backup mode allows explicit configuration of the primary SNAT router and the backup SNAT router. Each router is defined explicitly, and the IP address of the peer router is specified. See Figure 3: Primary/Backup Example.

Figure 3
Primary/Backup Example

PRIMARY	BACKUP
<pre>ip nat Stateful id 1 primary 10.88.194.17 peer 10.88.194.18 mapping-id 10</pre>	<pre>ip nat Stateful id 2 backup 10.88.194.18 peer 10.88.194.17 mapping-id 10</pre>



The `primary` command identifies an interface and IP address that the primary SNAT will use as the source for communicating with the backup SNAT router (for sending SNAT protocol messages). Likewise, the `backup` command does the same for the backup SNAT router. The `peer` command defines the destination IP address to use for communicating with the peer.

SNAT Interaction With HSRP

SNAT can be configured to interoperate with HSRP, using an IP Redundancy API within Cisco IOS Software. The “Active” and “Standby” routers are determined from the API and do not need to be explicitly defined. An example of configuration using IP Redundancy mode is depicted in Figure 2: Stateful NAT Example. Merely coding `redundancy SNATHSRP` causes SNAT to make use of the IP Redundancy API. The name is the same as that used in the command `standby name SNATHSRP`.

Verification

The status of the SNAT configuration can be examined by using the commands `sh ip snat distributed verbose` and `sh ip snat peer <ip-address>`. The TCP connection between the peer routers can be seen using the command `sh tcp brief`. See Figure 4: Show IP SNAT commands.

Figure 4
Show IP SNAT commands

```
cheney#sh ip snat distributed verbose

Stateful NAT Connected Peers

SNAT: Mode IP-REDUNDANCY :: ACTIVE
      : State READY
      : Local Address 10.88.194.17
      : Local NAT id 1
      : Peer Address 10.88.194.18
      : Peer NAT id 2
      : Mapping List 10
      : InMsgs 3246, OutMsgs 3247, tcb 0x82BF8BFC, listener 0x0

cheney#sh tcp brie
TCB      Local Address      Foreign Address      (state)
82BF8BFC 10.88.194.17.11001      10.88.194.18.15555  ESTAB

capefear1#sh ip snat distributed verbose

Stateful NAT Connected Peers

SNAT: Mode IP-REDUNDANCY :: STANDBY
      : State READY
      : Local Address 10.88.194.18
      : Local NAT id 2
      : Peer Address 10.88.194.17
      : Peer NAT id 1
      : Mapping List 10
      : InMsgs 3249, OutMsgs 3248, tcb 0x82BB4D0C, listener 0x826D1790

capefear1#sh tcp brief
TCB      Local Address      Foreign Address      (state)
82BB4D0C 10.88.194.18.15555      10.88.194.17.11001  ESTAB
```



Debug

SNAT can perform debugs to examine the message exchange and operation of the SNAT function. Example 3: Debug IP SNAT shows what happens at the backup SNAT router when a clear IP NAT translation * command is entered at the primary SNAT router. Delete messages are sent to the backup and processed.

Example 3 Debug IP SNAT

```
capefear1#debug ip snat ?
<1-99>   Access list
detailed Detailed Stateful NAT events
<cr>
capefear1#debug ip snat det
IP SNAT detailed debugging is on
capefear1#

Jul 18 10:21:15 EDT: SNAT (Send): Enqueuing SYNC Message for Router-Id 2
Jul 18 10:21:15 EDT: SNAT(write2net):ver 1, id 2, opcode 1, len 52
Jul 18 10:21:15 EDT: SNAT(write2net): 10.88.194.18 <--> 10.88.194.17 Writing message
#7303

Jul 18 10:21:33 EDT: SNAT (readfromnet): SYNC Message of length 52 read from
Router-id 1
Jul 18 10:21:33 EDT: SNAT (readfromnet): Enqueuing SYNC Message msg to readQ
Jul 18 10:21:33 EDT: SNAT(dbm): Process IPNAT_SNAT_PROCESS_MESSAGE EVENT
Jul 18 10:21:33 EDT: SNAT (Receive): Processed SYNC Message from Router-Id 1
Jul 18 10:21:33 EDT: SNAT (Receive): Sync message processed from Router-Id 1

Jul 18 10:21:35 EDT: SNAT (Send): Enqueuing SYNC Message for Router-Id 2
Jul 18 10:21:35 EDT: SNAT(write2net):ver 1, id 2, opcode 1, len 52
Jul 18 10:21:35 EDT: SNAT(write2net): 10.88.194.18 <--> 10.88.194.17 Writing message
#7304

Jul 18 10:21:44 EDT: SNAT (readfromnet): DELETE Message of length 52 read from
Router-id 1
Jul 18 10:21:44 EDT: SNAT (readfromnet): Enqueuing DELETE Message msg to readQ
Jul 18 10:21:44 EDT: SNAT(dbm): Process IPNAT_SNAT_PROCESS_MESSAGE EVENT
Jul 18 10:21:44 EDT: SNAT (Receive): Processed DELETE Message from Router-Id 1
Jul 18 10:21:44 EDT: SNAT (Receive): Entry deleted Nat-Id 298 Router-Id 1
Jul 18 10:21:45 EDT: SNAT (readfromnet): DELETE Message of length 52 read from
Router-id 1

*** SOME MESSAGES DELETED FOR BREVITY ***

Jul 18 10:21:50 EDT: SNAT (readfromnet): Enqueuing DELETE Message msg to readQ
Jul 18 10:21:50 EDT: SNAT(dbm): Process IPNAT_SNAT_PROCESS_MESSAGE EVENT
Jul 18 10:21:50 EDT: SNAT (Receive): Processed DELETE Message from Router-Id 1
Jul 18 10:21:50 EDT: SNAT (Receive): Entry deleted Nat-Id 304 Router-Id 1
Jul 18 10:21:51 EDT: SNAT (readfromnet): DELETE Message of length 52 read from
Router-id 1
Jul 18 10:21:51 EDT: SNAT (readfromnet): Enqueuing DELETE Message msg to readQ
Jul 18 10:21:51 EDT: SNAT(dbm): Process IPNAT_SNAT_PROCESS_MESSAGE EVENT
Jul 18 10:21:51 EDT: SNAT (Receive): Processed DELETE Message from Router-Id 1
Jul 18 10:21:51 EDT: SNAT (Receive): Entry deleted Nat-Id 305 Router-Id 1
Jul 18 10:21:52 EDT: SNAT (readfromnet): DELETE Message of length 52 read from
Router-id 1
```



Example 3 Debug IP SNAT

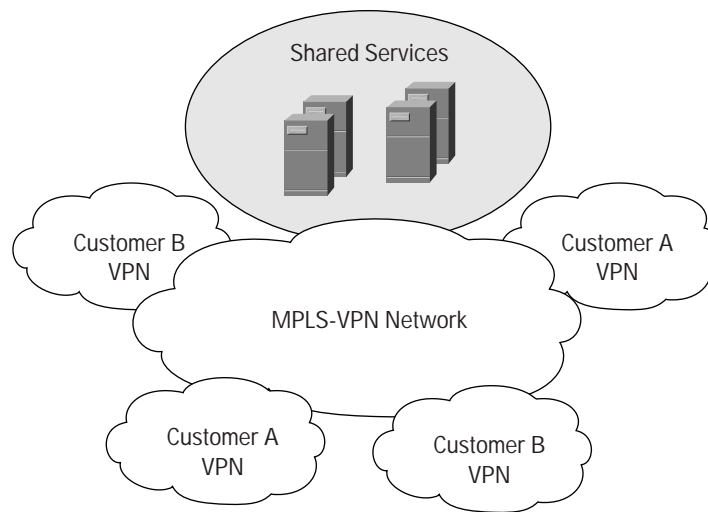
```
Jul 18 10:21:52 EDT: SNAT (readfromnet): Enqueuing DELETE Message msg to readQ
Jul 18 10:21:52 EDT: SNAT(dbm): Process IPNAT_SNAT_PROCESS_MESSAGE EVENT
Jul 18 10:21:52 EDT: SNAT (Receive): Processed DELETE Message from Router-Id 1
Jul 18 10:21:52 EDT: SNAT: Router-Id 1 deleted from snat hash
Jul 18 10:21:52 EDT: SNAT (Receive): Entry deleted Nat-Id 306 Router-Id 1
Jul 18 10:21:53 EDT: SNAT (readfromnet): SYNC Message of length 52 read from
Router-id 1
Jul 18 10:21:53 EDT: SNAT (readfromnet): Enqueuing SYNC Message msg to readQ
Jul 18 10:21:53 EDT: SNAT(dbm): Process IPNAT_SNAT_PROCESS_MESSAGE EVENT
Jul 18 10:21:53 EDT: SNAT (Receive): Processed SYNC Message from Router-Id 1
Jul 18 10:21:53 EDT: SNAT (Receive): Sync message processed from Router-Id 1

Jul 18 10:21:55 EDT: SNAT (Send): Enqueuing SYNC Message for Router-Id 2
Jul 18 10:21:55 EDT: SNAT(write2net):ver 1, id 2, opcode 1, len 52
Jul 18 10:21:55 EDT: SNAT(write2net): 10.88.194.18 <--> 10.88.194.17 Writing message
#7305
```

Deployment Example

The network as shown in Figure 6: SNAT Test Network was configured within a test lab to illustrate SNAT benefits. This design was chosen because it is common to deploy NAT at the Enterprise edge to a Service Provider network. In this case, a shared server is located within a Service Provider MPLS network that provides services to customers of the MPLS-VPN service. Service providers are poised to offer large-scale shared application services. These services will be very attractive to medium and large Enterprise customers, who prefer to outsource much of their communications and IT functions and concentrate on their core business. Service providers will be able to offer services at lower costs by scaling them to support many MPLS-VPN customers. (See Figure 5: Shared Service Deployment).

Figure 5
Shared Service Deployment



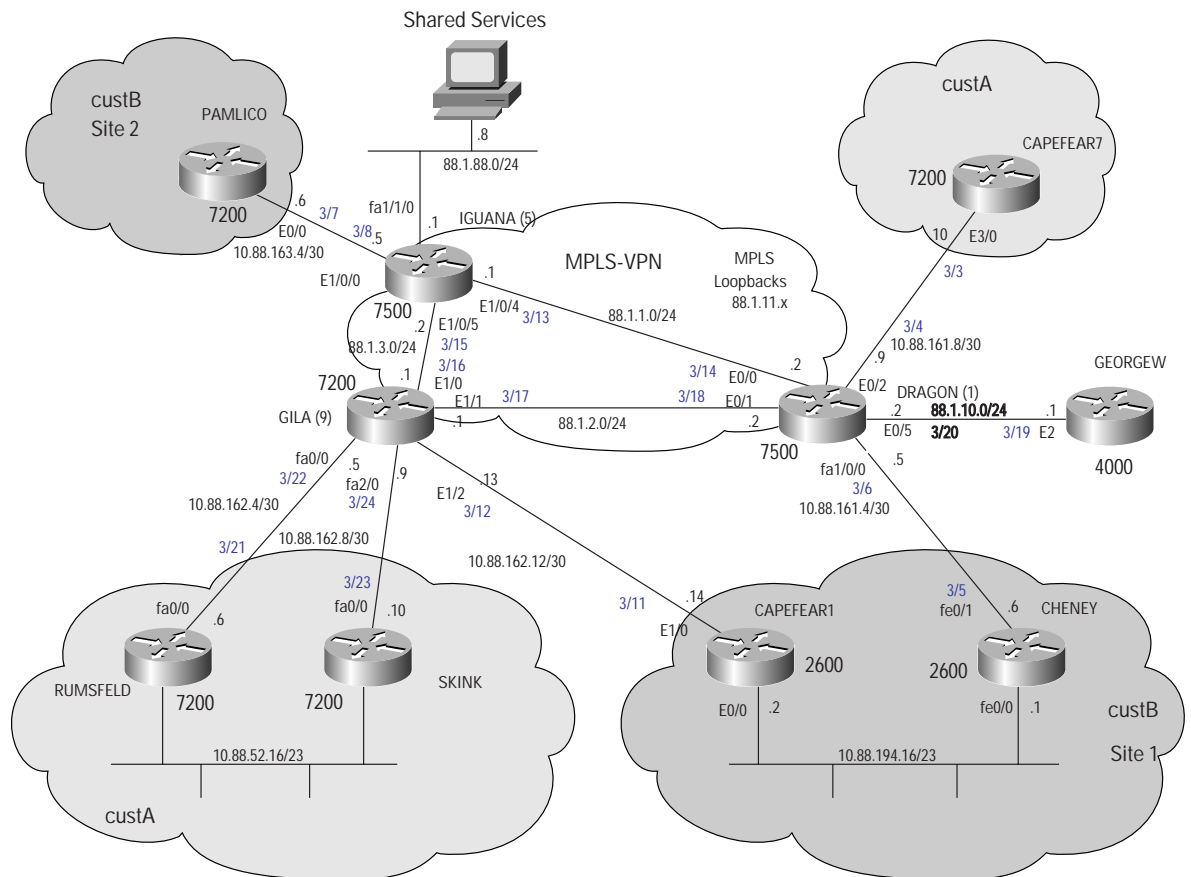


To illustrate how SNAT can enable the continuation of application data flow in the event of a critical failure, a simple test scenario was created.

1. Given a managed service within an SP MPLS-VPN, configure NAT at the enterprise edge to the Service Provider network.
2. Establish a TCP connection (Telnet) from the client within the customer VPN to the managed service host.
3. Force a failure of the link upstream from the active Network Address Translator.
4. Observe continuation of the session and data transfer between the TCP application and the client.
5. Record the time it takes for traffic to resume over the backup path.
6. Restore the link.
7. Record the time it takes for traffic to resume over the primary link.
8. Force a reload of the primary, active SNAT router.
9. Record the time it takes for traffic to resume over the backup path and then back to the primary following the reload.

In addition to the test scenario as described, we will also examine the issues related to the time it takes to resume traffic forwarding across the backup path.

Figure 6
SNAT Test Network





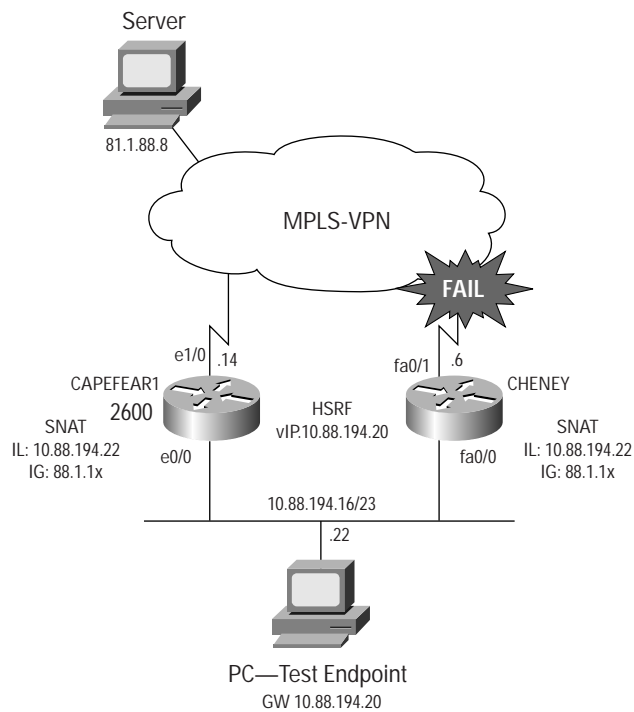
In the diagram shown in Figure 6: SNAT Test Network, we see two MPLS-VPN customers: CustA and CustB. A server exists within the MPLS-VPN service provider network connected at the PE router labeled, IGUANA. This test scenario involves the CustB site in the lower right labeled Site 1. There are two Cisco 2600 Series Routers labeled CAPEFEAR1 and CHENEY. These two routers are configured for SNAT.

Figure 7: SNAT Test Scenario illustrates this test. A PC attached to the LAN in CustB Site 1 will connect via Telnet to the shared server at IP Address 88.1.88.8. To simulate continuous transactions, a script was run to transmit a screen of data from the server to the PC every second.

During the continuous transfer of data traffic, a link failure was forced between the active NAT router and the MPLS-VPN PE router. Specifically, the cable connecting CHENEY to the PE router DRAGON was disconnected.

The Telnet sessions and data transmission along with the associated NAT function continued over the redundant path using router CAPEFEAR1 as the NAT. No loss of connection was observed.

Figure 7
SNAT Test Scenario



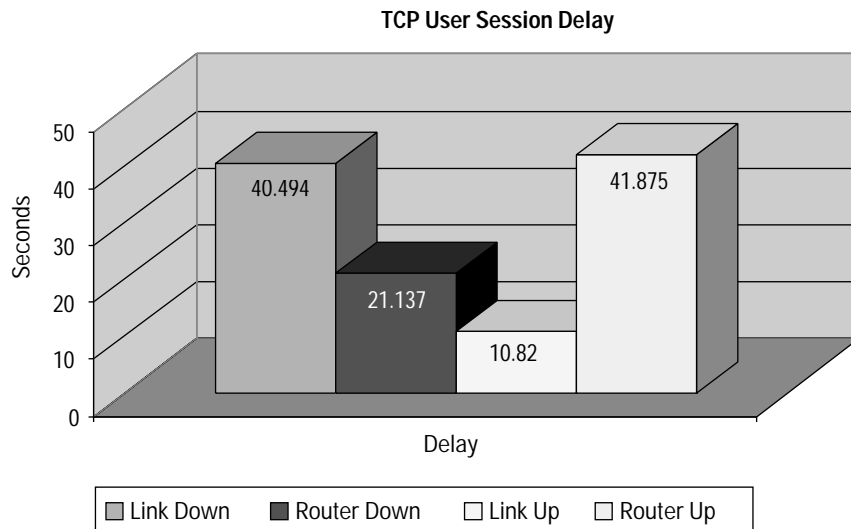
The cable was then reconnected and traffic resumed over the primary SNAT router CHENEY.

Next, CHENEY was forced to reload and traffic was automatically redirected to the standby SNAT and back after CHENEY came back online.

A Sniffer placed on the LAN segment observed traffic flow to the PC. The impact of the link failure was determined by the effect a user might see in terms of the data traffic. The trace determined the time of the last good frame received from the server for the TCP session. It can also determine the time when the data traffic resumed.. The user delay is the difference between these times. These results are shown in Figure 8: User Delay.



Figure 8
User Delay



The user delay seen is primarily the result of route convergence, in addition to the TCP retransmission timeout. In our test scenario, the site where SNAT is deployed is connected to an MPLS VPN, and BGP propagates route information. The server that generates the traffic is located within a separate VPN, which is attached to a different PE router. Routing information is being imported and exported between the customer VPN and the shared server VPN. The delays that affect traffic flow on a particular TCP session are largely due to the delay in waiting for the routing to converge for the return traffic from the server to the PC.

Only one set of parameters was adjusted at the primary Stateful NAT router. Tests showed that increasing the delay for the primary HSRP router to take over following an outage resulted in better overall results (less and more consistent delay for the TCP user traffic). This was accomplished with the following configuration on the primary (CHENEY) router.

```
standby preempt delay minimum 35 reload 45
```

In our tests, the effect of this statement was to force HSRP to delay the transfer from standby to active for thirty-five seconds after the link was restored and for forty-five seconds after a reload. This permitted traffic to continue over the backup router while route convergence occurred. This minimized the delay to the TCP data flow.

Complete configuration for the two stateful NAT routers are shown in the section Configurations.

Related Technologies

NAT for MPLS-VPNs

In the previous example, NAT was deployed on the Enterprise edge. While this is certainly an acceptable design and quite common, it may be more appropriate for the Service Provider to handle the NAT function. After all, an Enterprise customer that chose to purchase application services or outsource some portion of his processing workload would likely want to take advantage of NAT services, if possible.



Cisco NAT for MPLS-VPNs extends NAT so that Service Providers can establish the translation function within an MPLS network. This is the subject of a separate paper.

Conclusion

Cisco continues to enhance core features to provide increased benefit in terms of productivity gained from deployment of a more resilient IP network. Stateful NAT can provide higher availability to applications that use NAT services. It is expected that IP networks will use NAT in the foreseeable future. You can expect Cisco to lead the development of more robust and automated features relative to NAT, which will lower administrative costs and increase the return on investment in network technology.

Configurations:

NAT Router Cheney

```
cheney#sh run
Building configuration...

Current configuration : 3441 bytes
!
! Last configuration change at 14:25:29 EDT Fri Aug 2 2002
! NVRAM config last updated at 14:26:16 EDT Fri Aug 2 2002
!
version 12.2
service timestamps debug datetime localtime show-timezone
service timestamps log datetime localtime show-timezone
no service password-encryption
!
hostname cheney
!
boot system flash
logging buffered 32000 debugging
enable password *****
!
clock timezone EST -5
clock summer-time EDT recurring
ip subnet-zero
!
!
ip ftp username cisco
ip ftp password cisco
ip domain name ibd.cisco.com
ip name-server 172.18.60.179
!
voice call carrier capacity active
!
mta receive maximum-recipients 0
!
interface Loopback0
 ip address 10.88.194.5 255.255.255.252
!
interface Loopback1
 ip address 11.1.2.1 255.255.255.0
!
```



```
interface FastEthernet0/0
ip address 10.88.194.17 255.255.255.240
ip nat inside
no ip mroute-cache
speed 100
full-duplex
standby ip 10.88.194.20
standby timers 1 3
standby priority 105
standby preempt delay minimum 35 reload 45
standby name SNATHSRP
standby track FastEthernet0/1
!
interface FastEthernet0/1
ip address 10.88.161.6 255.255.255.252
ip nat outside
no ip mroute-cache
keepalive 2
speed 100
full-duplex
!
router bgp 65011
no synchronization
bgp log-neighbor-changes
redistribute connected
redistribute static
neighbor 10.88.161.5 remote-as 65002
neighbor 10.88.161.5 route-map SetMetOut out
no auto-summary
!
ip nat Stateful id 1
    redundancy SNATHSRP
    mapping-id 10
ip nat pool SNATPOOL1 11.1.1.1 11.1.1.9 prefix-length 24
ip nat inside source route-map rm-101 pool SNATPOOL1 mapping-id 10 overload
ip classless
ip route 11.1.1.0 255.255.255.0 Null0
no ip http server
ip pim bidir-enable
!
!
logging 172.18.60.179
access-list 1 permit 172.18.60.179
access-list 10 permit 10.88.194.22
access-list 22 permit 88.1.88.8
access-list 101 permit ip 10.88.194.16 0.0.0.15 11.0.0.0 0.255.255.255
access-list 101 permit ip 10.88.194.16 0.0.0.15 88.1.88.0 0.0.0.255
!
route-map rm-101 permit 10
    match ip address 101
!
route-map SetMetOut permit 10
    set metric 100
!
tftp-server system:vfiles
snmp-server engineID local 0000000902000030193E78C0
snmp-server community **** RO
```



```
snmp-server community **** RW 1
snmp-server trap-source Loopback0
snmp-server packetsize 4096
snmp-server location IBD Pineview 2
snmp-server contact ibd-tme@cisco.com
snmp-server system-shutdown
snmp-server enable traps snmp authentication linkdown linkup coldstart warmstart
snmp-server enable traps tty
snmp-server enable traps isdn call-information
snmp-server enable traps isdn layer2
snmp-server enable traps hsrp
snmp-server enable traps config
snmp-server enable traps entity
snmp-server enable traps envmon
snmp-server enable traps bgp
snmp-server enable traps rsvp
snmp-server enable traps frame-relay
snmp-server enable traps frame-relay subif
snmp-server enable traps rtr
snmp-server enable traps syslog
snmp-server enable traps dlsw
snmp-server enable traps dial
snmp-server enable traps voice poor-qov
snmp-server host 172.18.60.179 public
snmp-server manager
bridge 1 protocol ieee
bridge 2 protocol ieee
call rsvp-sync
!
mgcp profile default
!
dial-peer cor custom
!
line con 0
  exec-timeout 0 0
  password
  login
line aux 0
line vty 0 4
  password login
  length 0
!
exception core-file cheney-core compress
exception protocol ftp
exception dump 172.18.60.179
ntp clock-period 17180142
ntp peer 10.88.208.5
!
end
```

NAT Router Capefear1

```
capefear1#sh run
Building configuration...
```

```
Current configuration : 2650 bytes
```



```
!  
! Last configuration change at 14:26:10 EDT Fri Aug 2 2002  
! NVRAM config last updated at 14:26:12 EDT Fri Aug 2 2002  
!  
version 12.2  
no parser cache  
service timestamps debug datetime localtime show-timezone  
service timestamps log datetime localtime show-timezone  
no service password-encryption  
!  
hostname capefear1  
!  
logging buffered 32000 debugging  
enable password *****  
!  
clock timezone EST -5  
clock summer-time EDT recurring  
ip subnet-zero  
!  
voice call carrier capacity active  
!  
mta receive maximum-recipients 0  
!  
interface Loopback0  
 ip address 10.88.194.1 255.255.255.252  
!  
interface Ethernet0/0  
 ip address 10.88.194.18 255.255.255.240  
 no ip proxy-arp  
 ip nat inside  
 no ip mroute-cache  
 half-duplex  
 standby ip 10.88.194.20  
 standby timers 1 3  
 standby preempt  
 standby name SNATHSRP  
 standby track Ethernet1/0  
!  
interface Serial0/0  
 no ip address  
 no ip mroute-cache  
 no keepalive  
 shutdown  
!  
interface TokenRing0/0  
 no ip address  
 no ip mroute-cache  
 shutdown  
 ring-speed 16  
!  
interface Serial0/1  
 no ip address  
 shutdown  
!  
interface Serial0/2  
 no ip address  
 no ip mroute-cache
```



```
no keepalive
shutdown
!
interface Ethernet1/0
 ip address 10.88.162.14 255.255.255.252
 ip nat outside
 half-duplex
!
interface Ethernet1/1
 no ip address
 shutdown
 half-duplex
!
interface Ethernet1/2
 no ip address
 shutdown
 half-duplex
!
interface Ethernet1/3
 no ip address
 shutdown
 half-duplex
!
router bgp 65011
 no synchronization
 bgp log-neighbor-changes
 redistribute connected
 redistribute static
 neighbor 10.88.162.13 remote-as 65002
 neighbor 10.88.162.13 route-map SetMetOut out
 no auto-summary
!
ip nat Stateful id 2
    redundancy SNATHSRP
    mapping-id 10
ip nat pool SNATPOOL1 11.1.1.1 11.1.1.9 prefix-length 24
ip nat inside source route-map rm-101 pool SNATPOOL1 mapping-id 10 overload
ip classless
ip route 11.1.1.0 255.255.255.0 Null0 250
no ip http server
ip pim bidir-enable
!
access-list 10 permit 10.88.194.22
access-list 101 permit ip 10.88.194.16 0.0.0.15 11.0.0.0 0.255.255.255
access-list 101 permit ip 10.88.194.16 0.0.0.15 88.1.88.0 0.0.0.255
!
route-map rm-101 permit 10
 match ip address 101
!
route-map SetMetOut permit 10
 set metric 200
!
snmp-server engineID local 00000009020000505008B780
snmp-server community ***** RW
snmp-server community ***** RO
snmp-server packetsize 4096
snmp-server enable traps tty
```



```
snmp-server manager
call rsvp-sync
!
mgcp profile default
!
dial-peer cor custom
!
line con 0
  exec-timeout 0 0
line aux 0
  exec-timeout 0 0
line vty 0 4
  password
  login
line vty 5 15
  login
!
ntp clock-period 17208253
ntp peer 10.88.208.5
!
end
```

Note: For the final test the configuration was modified to use “overload”:

```
ip nat inside source route-map rm-101 pool SNATPOOL1 mapping-id 10 overload
```

Not all displays were done with this same configuration.

References

Cisco High Availability Initiatives

The High Costs of Network Downtime

http://newsroom.cisco.com/dlls/innovators/Core_IP/high_avail_initiatives.html

Cisco Globally Resilient IP

<http://www.cisco.com/go/grip/>



Corporate Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters
Cisco Systems International BV
Haarlerbergpark
Haarlerbergweg 13-19
1101 CH Amsterdam
The Netherlands
www-europe.cisco.com
Tel: 31 0 20 357 1000
Fax: 31 0 20 357 1100

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-7660
Fax: 408 527-0883

Asia Pacific Headquarters
Cisco Systems, Inc.
Capital Tower
168 Robinson Road
#22-01 to #29-01
Singapore 068912
www.cisco.com
Tel: +65 6317 7777
Fax: +65 6317 7799

Cisco Systems has more than 200 offices in the following countries and regions. Addresses, phone numbers, and fax numbers are listed on the

Cisco Web site at www.cisco.com/go/offices

Argentina • Australia • Austria • Belgium • Brazil • Bulgaria • Canada • Chile • China PRC • Colombia • Costa Rica • Croatia
Czech Republic • Denmark • Dubai, UAE • Finland • France • Germany • Greece • Hong Kong SAR • Hungary • India • Indonesia • Ireland
Israel • Italy • Japan • Korea • Luxembourg • Malaysia • Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland
Portugal • Puerto Rico • Romania • Russia • Saudi Arabia • Scotland • Singapore • Slovakia • Slovenia • South Africa • Spain • Sweden
Switzerland • Taiwan • Thailand • Turkey • Ukraine • United Kingdom • United States • Venezuela • Vietnam • Zimbabwe

All contents are Copyright © 1992-2003 Cisco Systems, Inc. All rights reserved. Cisco, Cisco Systems, Cisco IOS, and the Cisco Systems logo are registered trademarks or trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and certain other countries.

All other trademarks mentioned in this document or Web site are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company.
(0303R) 202925.F/ETMG 05/03