



**Technical Brief**

# Carrier-Grade Network Address Translation (CGNAT)

As the supply of IPv4 addresses is rapidly depleting, communication service providers are employing unique solutions to optimize current networks and transition to IPv6. One of the more successful solutions organizations are using to address this issue is carrier-grade network address translation (CGNAT). The F5 BIG-IP system provides unique scalability and flexibility to the CGNAT architecture, enabling service providers to offer more reliable and available service to their subscribers.

**by Jason Haworth**

Senior Director, Worldwide Field Systems Engineering, Carriers and Service Providers



# Contents

<b>Introduction</b>	<b>3</b>
<hr/>	
<b>Current IPv6 Methods</b>	<b>3</b>
<hr/>	
<b>F5 Solutions</b>	<b>4</b>
NAT44/444	4
Connection Limiting	6
IPv6 and NAT64	6
Service-Specific Routing and Accounting	7
<hr/>	
<b>Firewall</b>	<b>8</b>
<hr/>	
<b>Logging</b>	<b>8</b>
<hr/>	
<b>Monitoring</b>	<b>8</b>
<hr/>	
<b>Scaling</b>	<b>9</b>
<hr/>	
<b>iControl</b>	<b>9</b>
<hr/>	
<b>RFC Compliance for NAT-Based Solutions</b>	<b>9</b>
<hr/>	
<b>Conclusion</b>	<b>10</b>
<hr/>	
<b>Resources</b>	<b>11</b>



## Introduction

The pool of IPv4 addresses is swiftly drying up. Companies that provide Internet access are among the first to have to address this depletion as the growth rate of new users and devices is outpacing the available public IPv4 addresses. Eventually, the Internet will all be on IPv6; but as service providers transition their networks to support it, they will need to maintain connectivity to the IPv4 Internet. Service providers and their vendors long wondered when the Internet as a whole would transition to IPv6. Now the transition is upon them and the new question is, which methodologies will be used to make the transition to IPv6? Transitional technologies that can address the needs of both IPv4 and IPv6 are crucial to ensuring the continued reliability of existing data networks, as well as the viability of expanding those networks to account for the massive levels of growth in data access by consumers.

F5® technologies give architects and engineers a reasonable method with which to ensure that their IPv4 networks can provide highly controlled carrier-grade network address translation (CGNAT) functions. The flexibility of these systems allows them to incorporate IPv6 network access with the existing IPv4 Internet.

## Current IPv6 Methods

Today, various vendors have three basic approaches to IPv6: dual stack, tunneling, and CGNAT. Each approach has advantages and disadvantages that service providers must consider when developing an IPv6 strategy.

Almost all vendors support a dual stack (IPv4/IPv6) approach. To accommodate this, service providers' systems must all be able to simultaneously support IPv4 and IPv6. This can lead to massive increases in hardware costs. It is the easiest method for endpoint deployment of new IPv6 terminals for new IPv6 services; however it creates incompatibilities between legacy systems and new IPv6 services. The service provider must develop an EOL (end of life) strategy for legacy systems or face the cost of dual systems/services.

Another way to handle connectivity between IPv6 and IPv4 networks is tunneling. There are two tunneling methods: 6RD and DS-Lite. 6RD is a rapid deployment strategy that deploys IPv6 traffic in encapsulated tunnels on an IPv4 network.<sup>1</sup> 6RD's major advantage is the easy transport of TCP/UDP traffic over IPv4 or IPv6 networks as there is no need for application awareness. Problems with this approach include

1 "RFC 5569 - IPv6 Rapid Deployment on IPv4 Infrastructures (6rd)", n.d., <http://tools.ietf.org/html/rfc5569>.



the CapEx of tunnel-capable consumer premises equipment (CPE), the stress or tunnel density termination, and environmental isolation of IPv4 and IPv6, along with the fact that this method is not a transition to a IPv6 network. DS-Lite transports IPv4 traffic across encapsulated tunnels on an IPv6 network.<sup>2</sup> It has many of the advantages and drawbacks of 6RD, along with increased problems in management and the support of large numbers of IPv6 tunnels.

The third approach is NAT (network address translation), which is supported by vendors such as F5, Juniper, and Cisco (dependent on requirements for NAT44 or NAT64). NAT can be implemented to support address translation between IPv4 networks (NAT44) or between IPv6 and IPv4 networks (NAT64). When NAT is deployed to support millions of transactions per second in a distributed architecture, it's referred to as carrier-grade NAT (CGNAT). CGNAT has no effect on legacy endpoint systems, providing seamless access for any terminal to any service, whether IPv4 or IPv6. However this approach does require application awareness and requires administrators to address the issue of NAT termination density.

NAT44 installations have many references and require minimal changes to service providers' legacy systems. But if unmodified, a NAT44 deployment would be a never-ending expansion, delay IPv6 implementation, and risk future application breaks.

NAT64 enables predictable expansion by providing a managed transition to IPv6 with minimal changes to existing systems. This is a relatively new solution and application breaks may occur.

## F5 Solutions

### NAT44/444

Both NAT44 and NAT444 are essentially trying to solve the address depletion problem in the same fashion: by masking several non-publicly routable IP addresses<sup>3</sup> behind a smaller number of publicly routable IP addresses.

NAT44 functions often have problems with IP address affinity because addresses are distributed across multiple private side hosts. But the F5® iRules® scripting language enables developers to easily write code into their products.

<sup>2</sup> "RFC 6333 - Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", n.d., <http://tools.ietf.org/html/rfc6333>.

<sup>3</sup> "RFC 1918 - Address Allocation for Private Internets", n.d., <http://tools.ietf.org/html/rfc1918>.

## Technical Brief

### Carrier-Grade Network Address Translation (CGNAT)



This is an example of how administrators can use iRules to assign an IP address based on the third and fourth octet in their address:

```
when CLIENT_ACCEPTED {snat 12.0.[getfield [IP::client_addr] "." 2].  
[getfield [IP::client_addr] "." 3]}
```

While this NAT technique is proven and has been widely used, the double NAT functions of NAT444 on a network can have trouble transporting providers because they need Application Layer Gateway (ALG) functions that require certain applications be Layer 3 (IP) and/or Layer 4 (for TCP, UDP, SCTP) aware. Without this awareness some applications can break within the transport system, which generates customer care calls. To prevent this, administrators can implement the F5 BIG-IP® system as a Service Delivery Controller (SDC) to look at protocol-specific information and to change data in both the headers and payload. The inherent ALG functions on a BIG-IP product as an SDC can be rapidly implemented by combining existing profile information for well-known protocols such as HTTP, SIP, and FTP., with F5's powerful iRules functionality. Customer ALG functions, which can take other vendors months to release in coding, cycles require custom development and support.

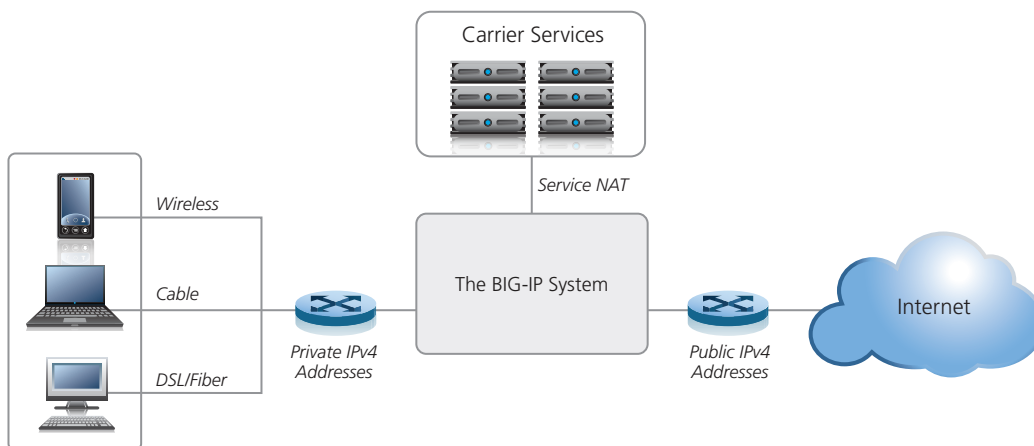


Figure 1: The BIG-IP system supporting IPv4 CGNAT

Additional considerations should be made for dynamic versus static NAT assignments. Based on the registration methodology for a particular user (for example, PDP Context, BRAS Auth, Simple AAA), that user can be assigned according to their specific registration device (for example, MSISDN, IMEI, MAC, Serial Number) and have an IP mapping to a specific policy function that can bind that user to a NAT pool or a static NAT IP mapping function. That NAT mapping function can exist within the same system, and while the user's RFC 1918 address assignment might need to change, their registration on the system can be used to determine whether



to map them externally to specific IPs. This could be a static one-to-one mapping, a static many-to-one mapping, or a fully dynamic mapping.

## Connection Limiting

Many vendors believe that connection limiting should be a function of assigning a range of static ports for all ephemeral connections, and that when these ports are used up, additional ports can be assigned. But this method limits operators' ability to provide overflow queues, and makes it difficult to track users since overflow pools have to be permitted outside the pre-assigned port allocations and need to be tracked for every connection that goes into the overflow queue. Quite often the connection limitations implemented are related to the variance in port assignments. With an F5 SDC, users can be assigned a small block of ephemeral ports to use. A bit counter watches the incremental connections being used, and decrements the bit bucket as sessions start and end for a given source address. This minimizes the effects of carving out large blocks of IPs and pinning a smaller, fixed number of clients to each external IP address.

Connection limiting can also trigger an event, such as an HTTP redirect back to a user, to allow that user to purchase more connection space on the system. This would increase revenue generated by that user and protect the infrastructure from abusive users.

## IPv6 and NAT64

F5 uses NAT64 and DNS64 to enable service providers to transition seamlessly to IPv6. DNS64 is the ability to translate DNS responses in either IPv6 or IPv4 format.<sup>4</sup> One significant benefit of DNS64 functions is that they ensure that administrators don't have to make additional DNS infrastructure changes.

This is an example of an iRule for NAT64:

```
when CLIENT_ACCEPTED {
  translate address enable
  set dest [string range [IP::addr IP::local_addr]mask ::ffff:ffff] 0 end]
  log local0. "node $dest"
  node $dest}
```

<sup>4</sup> "RFC 6147 - DNS64: DNS Extensions for Network Address Translation from IPv6 Clients to IPv4 Servers", n.d., <http://tools.ietf.org/html/rfc6147>.

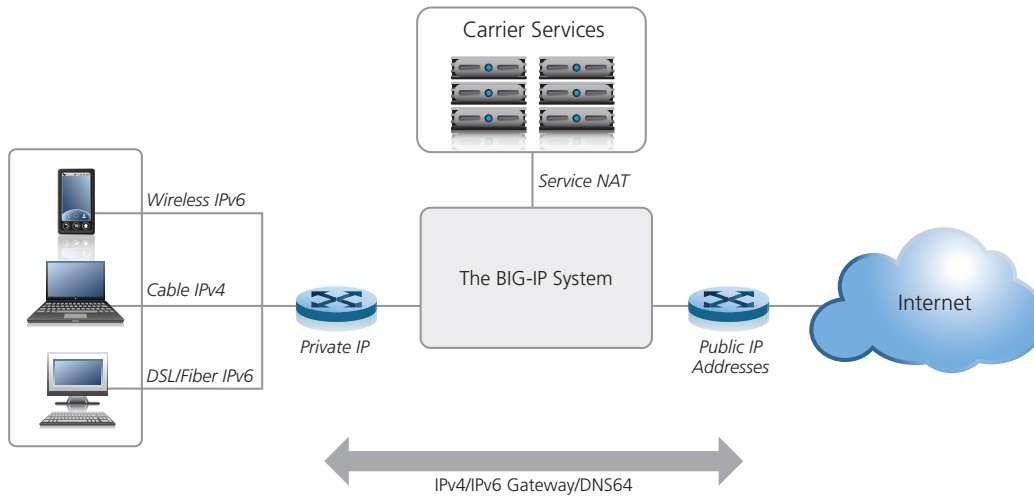


Figure 2: The BIG-IP system supporting IPv4 and IPv6 CGNAT simultaneously

## Service-Specific Routing and Accounting

Some services may need to be allowed through the system and could need different paths for billing and port allocation concerns; this is known as service-specific routing and accounting, or traffic steering. These different paths could require different source NAT addresses, different next hop logic, or could be destined for more advanced content manipulation. F5 SDCs offer content inspection, interpretation, and modification for IP information, and transport protocols, as well as application layer-specific information. For example, if a provider wanted to allow their customers to access a popular social media outlet or search engine, it could set up several proxies that have their own IP scheme, and handle those connections as well as private side connections to the content provider. The BIG-IP system could proxy those connections to the providers without connection limitations using a different IP scheme from normal traffic; using TCP sockets for rapid content delivery; and transforming critical application information to enhance the user experience.

## Firewall

F5 BIG-IP® Local Traffic Manager™ (LTM) is a default deny device. It can be used to perform basic L2–L4 firewall capabilities, including logging functions, to satisfy bi-directional firewall requirements. Additional capabilities can be implemented using BIG-IP® Application Security Manager™ (ASM), Protocol Security Module™ (PSM), and Access Policy Manager™ (APM).



## Logging

The BIG-IP system includes a highly customizable high-speed logging (HSL) engine that could track every connection if necessary. But logging can be configured for a particular class of user or a large group of users to ensure that logging data is minimized. Most providers, however, are more concerned with logging for tracking purposes, for example, a start/stop function for a source IP being assigned an IP for connecting. Additional information could include connection quota overages, warnings of connection overages, upstream or downstream connections lost, and RST thresholds. Logging an event is simple and can be configured at multiple points. Several vendors have log messages that are close to 1500 bytes per message, where a session start log message with the source IP, source port, destination IP, destination port, and a time stamp can be under 100 bytes from the HSL engine in the BIG-IP system. This is a significant savings in data storage requirements for logging.

## Monitoring

The ability to monitor and manage connections based on the health of connected devices is an important aspect of using an F5 SDC. With its advanced monitoring functions, the SDC can make traffic steering decisions based on multiple site state health decisions.

This can help the BIG-IP system intelligently direct traffic to a site that can continue to provide the client with service while alerting the operations team of the outage. An example of this would be if the BIG-IP system were monitoring the path from itself to several different sites on the Internet. When the BIG-IP system sees that a certain number of these sites are no longer sending back valid data (perhaps there are not 200 OK messages sent in response to HTTP requests), the BIG-IP system can communicate with the providers' routing cloud and remove that network segment as an available next hop user to connect to the Internet. This monitoring capability allows data centers to fail over and continue to provide service, while also sending out alerts to the operations center regarding the detected failure and subsequent actions taken to restore service.

There are several other monitoring options available in the BIG-IP system such as selective service proxying to other access gateways, or directing users to a site down page.



## Scaling

F5 SDCs can scale on a single platform to 72 Gbps, 2.4 million connections per second, and 48 million concurrent connections. In addition, BIG-IP products can be clustered to provide approximately eight times this level of throughput, even for advanced L7 functionality. F5 products' ability to scale is unmatched in the industry.

## iControl

F5 iControl® is the first open API that enables applications to work in concert with the underlying network based on true software integration.

Using SOAP/XML to ensure open communication between dissimilar systems, iControl helps service providers realize new levels of automation and configuration management efficiencies.

Whether monitoring network-level traffic statistics, automating network configuration and management, or facilitating next-generation service-oriented architectures, iControl gives organizations the power and flexibility to ensure that the network and applications work together for increased reliability, security, and performance. In this way, iControl can help reduce the cost of managing complex environments.

## RFC Compliance for NAT-Based Solutions

While F5 is committed to RFC compliance, it isn't the only driving factor in ensuring communications work effectively across the network. Quite often one vendor's interpretation of an RFC differs from another's. F5 products are unique in their programmability, in that they give organizations the ability to comply with RFC operations for interoperability between vendors.

F5 supports the following RFCs:

**RFC 4787:** F. Audet and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP," BCP 127, RFC 4787, January 2007.

**RFC 5382:** S. Guha, K. Biswas, B. Ford, S. Sivakumar, and P. Srisuresh, "NAT Behavioral Requirements for TCP," BCP 142, RFC 5382, October 2008.



**RFC 5508:** P. Srisuresh, B. Ford, S. Sivakumar, and S. Guha, “NAT Behavioral Requirements for ICMP,” BCP 148, RFC 5508, April 2009.

F5 supports the following RFCs with exception:

**RFC 5597:** R. Denis-Courmont, “Network Address Translation (NAT) Behavioral Requirements for the Datagram Congestion Control Protocol,” BCP 150, RFC 5597, September 2009.

**RFC 5135:** D. Wing and T. Eckert, “IP Multicast Requirements for a Network Address Translator (NAT) and a Network Address Port Translator (NAPT),” BCP 135, February 2008.

F5 iRules enables administrators to adjust the application/packet flow as needed to support potentially unsupported RFCs.

## Conclusion

IPv6 is here—but most content for subscribers is still on the IPv4 Internet. Service providers must be prepared to rapidly adopt IPv6. The challenge they face is determining what method to use to transition and interoperate between IPv4 and IPv6 networks. CGNAT, coupled with DNS64, provides the comprehensive connectivity to content that subscribers demand. CGNAT capabilities in the BIG-IP system provide the service provider with an application-aware, high-performance, flexible, and highly scalable solution to IPv4 to IPv6 transition and interoperability. With the flexibility provided by F5 technologies, service providers can be prepared for today’s problems as well as tomorrow’s challenges.

## Resources

The following resources include additional information about the architecture and design of CGNAT implementations.

### Traffic steering

<http://www.f5.com/news-press-events/web-media/webcasts/traffic-steering.html>

### NAPT

<http://www.ietf.org/rfc/rfc2766.txt>

### NAT64

<http://datatracker.ietf.org/doc/draft-ietf-behave-v6v4-xlate-stateful/>



## **XLATE**

<https://datatracker.ietf.org/doc/draft-ietf-behave-v6v4-xlate/>

## **NAT66**

<http://tools.ietf.org/html/draft-mrw-behave-nat66-02>

## **RFC 2473**

A. Conta and S. Deering, "Generic Packet Tunneling in IPv6 Specification," RFC 2473, December 1998.

## **RFC 2663**

P. Srisuresh and M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations," RFC 2663, August 1999.

## **RFC 4787**

F. Audet and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP," BCP 127, RFC 4787, January 2007.

## **RFC 5382**

S. Guha, K. Biswas, B. Ford, S. Sivakumar, and P. Srisuresh, "NAT Behavioral Requirements for TCP," BCP 142, RFC 5382, October 2008.

## **RFC 5508**

P. Srisuresh, B. Ford, S. Sivakumar, and S. Guha, "NAT Behavioral Requirements for ICMP," BCP 148, RFC 5508, April 2009.

## **IPv4 Exhaustion**

<http://www.potaroo.net/tools/ipv4/index.html>

## **IANA Allocation**

<http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xhtml>

## **NAT444**

<http://tools.ietf.org/html/draft-shirasaki-nat444-isp-shared-addr-03>

## **INC-CGN**

<http://tools.ietf.org/html/draft-ietf-v6ops-incremental-cgn-01>

## **CGN**

<https://wiki.tools.ietf.org/html/draft-nishitani-cgn-04>

**Technical Brief**

Carrier-Grade Network Address Translation (CGNAT)

**F5 Networks, Inc.** 401 Elliott Avenue West, Seattle, WA 98119 888-882-4447 [www.f5.com](http://www.f5.com)

---

F5 Networks, Inc.  
Corporate Headquarters  
[info@f5.com](mailto:info@f5.com)

F5 Networks  
Asia-Pacific  
[apainfo@f5.com](mailto:apainfo@f5.com)

F5 Networks Ltd.  
Europe/Middle-East/Africa  
[emeainfo@f5.com](mailto:emeainfo@f5.com)

F5 Networks  
Japan K.K.  
[f5j-info@f5.com](mailto:f5j-info@f5.com)

