INTERCONNECT COMMUNICATIONS

Ofcom

MC/159
Report on the Implications of Carrier Grade Network Address Translators

Final Report

Merlin House
Chepstow
NP16 5PB
United Kingdom

Telephone: +44 1291 638400
Facsimile: +44 1291 638401
Email: info@icc-uk.com
Internet: www.icc-uk.com

INTERCONNECT COMMUNICATIONS
Consulting in Communications Regulation and Strategy

Notice

This document is provided in good faith and is based on InterConnect's understanding of the recipient's requirements. InterConnect would be pleased to discuss the contents of this document particularly if the recipient's requirements have in any way changed.

InterConnect is a wholly owned subsidiary of Telcordia Technologies Inc.

InterConnect Communications Ltd
Merlin House
Station Road
Chepstow
NP16 5PB
United Kingdom

Telephone: +44 1291 638400
Facsimile: +44 1291 638401
www.icc-uk.com

Persons to contact in relation to this document:

Brian Aitken
Business Development Executive
DDI: +44 (0) 1291 638426
Fax: +44 (0) 1291 638401
Email: brianaitken@icc-uk.com

Table of Contents

Executive Summary

This report identifies the technical, consumer, policy, privacy and law enforcement issues associated with Carrier Grade Network Address Translation (CGN) as currently implemented in today's Internet.

1.1 Why CGNs?

Addresses are fundamental to the way the Internet works. The Internet's tremendous growth has resulted in the Internet running out of addresses in their current format, IPv4. This development has long been anticipated and a successor format, called IPv6, is ready to be adopted.

However, the transition from IPv4 to IPv6 will take many years. The transition will require upgrades of Internet applications, appliances, services, consumer electronic devices and networks. During this transition network operators will be running networks where IPv4 and IPv6 co-exist. There are a large number of IPv4-to-IPv6 transition mechanisms, all of which still require IPv4 addresses, despite the impending exhaustion of the IPv4 address space.

Therefore**,** it is crucial to find ways to maximize the use of available IPv4 addresses. One tool for conserving IPv4 addresses is called Carrier Grade Network Address Translation, or simply CGN.

1.2 CGNs in the UK

At present, CGN implementation is in very early stages in the UK. A first field trial by the UK ISP, Plusnet, was conducted from January to March this year. CGN has yet to make a significant impact on consumer Internet services. Still, every ISP contacted during the research of this report is aware of CGN technology and is in a variety of stages of research, planning or trials.

Given the slow deployment of IPv6 at this stage of IPv4 exhaustion, CGN is likely to be part of the technology landscape during the transition from IPv4 to IPv6. Also, CGN technology has improved, year-on-year, as experience has been gained with its impacts on applications and end users. Still, implementation of CGN has significant implications for the quality of Internet services provided to UK consumers.

1.3 How CGN Works

CGN is not a single technology but a collection of strategies for sharing addresses among a large pool of Internet consumers. In the past, an Internet user might have had a unique address for every device they connected to the Internet. However, this is no longer true. In fact, the strategy of Network Address Translation (NAT) has been in place for more than a decade. A typical use of NAT is in broadband consumers' home networks where the interface between the service provider's network and the consumer's network translates a pool of private address (used in the home) to a single public address (the public side of the home network).

CGN effectively takes the idea of consumer NAT and extends it into the service provider's network. Rather than each consumer having a single public address, Carrier Grade NAT allows a large number of consumers to share a single address. The goal is to conserve addresses during the period of IPv4-to-IPv6 transition. By sharing a single, public IPv4 address amongst multiple consumers, the demands on scarce IPv4 address space are reduced.

1.4 Technical Implications

This report finds that Internet application developers and consumers will bear most of the impact of CGN implementation:

- Application developers will be faced with multiple CGN implementations and have to find workarounds for each
- Consumers will find that applications they count on no longer function correctly and that troubleshooting those problems becomes more difficult

CGN has been used successfully in mobile networks for some time. This is because mobile networks have features that make them particularly good candidates for CGN. For instance, almost all the connections from a mobile handset are outbound toward the network and the applications that the device can run are usually tightly constrained. Consumer broadband networks are not constrained in this way.

In fact, CGN has been shown to cause deterioration in the quality of Internet access services. While many consumers may continue to have a seamless online service experience after their providers deploy CGN, this report shows that all CGN implementations are likely to cause some online applications to break, degrade or operate in unsatisfactory ways. Further, since there are no standards for CGN in place, CGNs at different access providers will cause services and applications to break in different ways. Consumers may not even know that their network is behind a CGN and would be unaware that the CGN might be contributing to problems in their home networks.

Technically, CGN affects consumers of Internet services in three crucial ways:

1. The number of connections each consumer network can have at any single time is limited
2. The ability to precisely identify any Internet-attached device solely by IP address is no longer available
3. The ability to establish and maintain connections to devices in consumer network becomes much more difficult

While these issues will not affect an Internet consumer who simply checks email or does very basic web browsing, most advanced applications on the Internet depend on rich connectivity between the client and the Internet application. Many advanced, innovative applications either break in the presence of CGN or begin to run in unacceptable ways from the consumer's point of view.

The introduction of CGNs also has the potential to add delay, or latency, in the network. Delay in delivery of information would be especially problematic for applications that are sensitive to delay; for instance, foreign currency exchange, gaming, day trading, or sports betting

While this report documents both lab and field study of the impact of CGNs on applications, the permanent loss of any future innovation is extremely difficult, if not impossible to quantify. However, there clearly is a cost: in both developer burden and lost opportunities.

1.5 Policy Implications

1.5.1 Law Enforcement

Sharing addresses among many users can make it more difficult for Internet Service Providers to respond to legal requests for logged or monitored traffic. Typically, an ISP meets those requirements by capturing, or logging, traffic to-and-from a specific IP address for a specific period of time. If IP addresses are shared, the IP address no longer uniquely identifies the end user. The result is that the ISP will have to gather far more information to respond to legal requests from law enforcement. What becomes troubling is the potential for the amount, and the associated costs, of information to become so enormous under CGN that the ISP is tempted not to log the data, or that fundamental rights of privacy are affected by intrusive data processing.

1.5.2 Competition

CGNs have a potential to affect competition for Internet access, services and applications in the UK:

- **Dominance of existing players**
  The top five ISPs in the UK may benefit from adoption of CGN because they will maintain the advantage they have with dominant, legacy IPv4 networks. As the UK's transition to IPv6 is delayed, new ISPs may have difficulty in securing IP address resources that are needed for their businesses to grow. The potential to delay IPv6 transition and increase the value of IPv4 addresses would possibly limit external competitive pressure on those top five UK ISPs (although they may be subject to intense competition between each other).

- **Greater role for Internet access providers in deciding what applications services their networks will enable**
  Given the difficulties CGNs pose for seamless access to Internet applications and services, there is the possibility for a change in the relationship between ISP and developers of applications. In particular, CGNs may lead to applications developers and service providers needing to enter into relationships with individual ISPs to ensure their services can reach those ISPs' users. Not only does this change the current environment where the online services market is largely independent from the Internet access market, but it also potentially changes the role of ISPs themselves from "mere conduits" to content and application moderators that can be considered liable for the applications they permit through their CGNs.

- **Development of tiered levels of Internet access**
  Given CGNs will cause some applications and services to break, Internet access providers may offer premium levels of Internet access, where subscribers can choose to pay more to be connected via public—possibly static—IPv4 addresses. This has the potential to lead to a tiered Internet with different classes of service for different users. In addition, given that CGN deployment extends the usable life of the IPv4 protocol, possibly inhibiting the deployment of IPv6 in the UK in the medium term, it could cause the Internet to fragment between existing and ongoing deployments of IPv6 in other parts of the world and IPv4-centric networks in the UK. This would further inhibit the competitiveness of UK-based applications and services on the international market.

Section A: Technical Implications

2 Why Are Carrier Grade NATs Required?

The Internet was designed with each host or node being assigned one or more globally unique IPv4 addresses. In the mid-1990s it became apparent that the IPv4 address space was not sufficient for the growing Internet and that it would be exhausted if action was not taken. As a result, addresses began to be shared. Address sharing was implemented using a technique called Network Address Translation (NAT). Today this form of NAT is referred to as NAT44. NAT44 has delayed the exhaustion of the IPv4 address space for over a decade, effectively saving the Internet.

NAT44 has only delayed the exhaustion of the IPv4 address space. Growth in Internet users and new Internet applications and services continues to require an ever increasing number of public IPv4 addresses. NAT44 itself requires a global pool of IPv4 addresses that can be shared. So although NAT44 slows the consumption of IPv4 addresses, it still consumes IPv4 addresses.

Today, ISPs in the UK can only use their existing stock of IPv4 addresses to service customers. If they are able to justify a further allocation from RIPE NCC, they will only receive one final /22, which represents 1,024 addresses.[1] Existing ISPs continue to consume IPv4 addresses to support new customers and new applications and services. Since they cannot obtain more IPv4 addresses, they are forced to use their currently allocated IPv4 addresses more efficiently.

Carrier Grade NAT (CGN), also known as Large Scale NAT[2] (LSN) or NAT444, is a technique that makes it possible to use fewer public IPv4 addresses to support more subscribers. New ISPs will only be able to obtain a /22 IPv4 allocation (1,024 addresses) from RIPE NCC. Even with CGN, this will severely limit the number of subscribers that the ISP can support.

2.1 IPv4 Address Exhaustion

The IPv4 Internet is running out of public IPv4 addresses. Public IPv4 addresses are necessary to communicate between hosts on the global IPv4 Internet. The exhaustion of IPv4 addresses is shown in Table 2.1.

[1] RIPE, *IPv4 Address Allocation and Assignment Policies for the RIPE NCC Service Region*, http://www.ripe.net/ripe/docs/ripe-582, 2013, Section 5.6, accessed 15 April 2013.

[2] "Large Scale NAT" or "LSN" is the most widely accepted nomenclature on today's Internet. "LSN" is considered a more accurate and neutral term than "Carrier Grade NAT" for two main reasons:

1. "Carrier Grade NAT" gives the impression that it is technically different from NAT44 when it is the same
2. "Carrier Grade" implies a better quality of service whereas CGN actually has a negative impact on performance and reliability

This report, however, will use the term "CGN" in conformance with the original Ofcom project brief.

Registry	Exhausted	Comment
IANA Central IPv4 Address Pool	3rd February 2011	
APNIC	19th April 2011	Last /8 policy underway
RIPE NCC	14th September 2012	Last /8 policy underway
ARIN	Predicted 2014	
LACNIC	Predicted 2014	
AFRINIC	Predicted 2020	

Table 2.1 - Exhaustion of IPv4 Addresses in Global and Regional Registries[3]

When RIPE NCC exhausted its IPv4 address allocation, it moved to the "last /8 allocation policy."[4] The /8 (slash eight) refers to the final block of IPv4 addresses that RIPE NCC is able to allocate to its members. The policy for allocating address blocks from this /8 to members is very restrictive and only allows for a single and final allocation of one /22 block to any member that can justify the allocation. A /22 represents 1,024 IPv4 addresses.

To put the IPv4 address consumption in context, it should be noted that prior to exhaustion the consumption rate of IPv4 addresses was substantial.[5] For example, consumption in the APNIC region peaked at 17 /8s (approximately 285 million addresses) per year, or over 23 million addresses per month.

2.2 CGN and IPv6

IPv6 is the only long term solution to the address exhaustion problem. Outside the UK, IPv6 is becoming widely deployed. Now that major sites and service providers such as Google, Facebook, Akamai, Netflix are IPv6-enabled, overseas ISPs are finding that up to 33%[6] of their dual stack subscribers' traffic is IPv6 traffic. This percentage should not be confused with the global percentage of native IPv6 traffic, which is currently much lower.[7]

The delay in deploying IPv6 means that IPv4 is still widely required by IPv4-only services and applications. This is particularly the case in the UK where few ISPs provide a native IPv6 service to their customers. This means that IPv4 must continue to be supported for many years despite the exhaustion of IPv4 addresses. CGN will play an important role in enabling ISPs to provide an IPv4 service to customers during the transition to IPv6.

[3] G. Huston, *IPv4 Address Report*, http://www.potaroo.net/tools/ipv4/index.html, 2013, accessed 17 April 2013.

[4] RIPE, *IPv4 Address Allocation and Assignment Policies for the RIPE NCC Service Region*, http://www.ripe.net/ripe/docs/ripe-582, 2013, Section 5.6, accessed 15 April 2013.

[5] Ibid Figure 28 (a-f).

[6] C. M. Martinez, "Re: Big day for IPv6 - 1% native penetration", *NANOG mailing list*, http://seclists.org/nanog/2012/Nov/570, 2012, accessed 15 April 2013.

[7] A view of this data from April 2013 is available at E.Aben, Networks with IPv6 Over Time – A Short Update, 4 April 2013, https://labs.ripe.net/Members/emileaben/networks-with-ipv6-over-time-a-short-update

2.3 CGN Deployment Scenarios

CGN is not appropriate for all scenarios where there is an IPv4 address shortage. CGN is primarily a tool for ISPs to deploy in their access networks.

CGN is appropriate for:

- **Access networks** – more so for mobile access networks than fixed access networks
- **Basic Internet access** – providing access to fundamental and well-established **Internet protocols** – such as electronic mail and basic World Wide Web services
- **Client access** – that is connections initiated from the subscriber network

CGN is not appropriate for:

- **Carrier networks** – providing Internet backhaul connectivity
- **Content provider networks** – for example, application hosting, cloud services, and web-hosting)
- **Enterprise customers**
- **ISPs providing services to businesses**

CGN is very unlikely to be appropriate for:

- **Subscribers wishing to deploy Internet facing services** – that is, services that can be reached by connections into the subscriber network
- **Subscribers with more advanced networking needs** – for example, web-cams, remote access, and VPNs
- **Subscribers using peer-to-peer applications** – including some multiplayer games

It is important to remember that CGN is not applicable to many types of end-user. Any end-user that requires one or more public IPv4 addresses cannot be placed behind CGN. These include many businesses that require public IPv4 addresses to operate public-facing Internet services.

2.4 CGNs and the Transition to IPv6

IPv6 has a huge address space that is unlikely to ever be exhausted. Therefore, IPv6 does not require NAT or CGN. As a consequence, all of the problems found with NAT and CGN are not found in IPv6 networks.

CGNs can impact the transition to IPv6. Since CGNs modify IP packet headers in transit, they may break some IPv6 transition mechanisms thereby stopping some subscribers from using IPv6. Examples of transition mechanisms that can be broken by CGNs include:

- **6to4**
  6to4 requires a global IPv4 address to function. There are three possible problems with CGNs and 6to4:

  1. 6to4 will not function because the subscriber does not have a global IPv4 address.
  2. 6to4 will attempt to work and will fail because the CGN is configured with global IPv4 addresses on its internal access network. This is unlikely to happen in

modern CGN deployments as long as the correct IPv4 address space is used on the internal side of the CGN. This has been observed in early CGN deployments. It results in performance problems (significant delays on 10s of seconds) or failure to connect to services since the client attempts to use 6to4 with a wrong public address.[8]

3. 6to4 will usually break through CGN. This is because CGN carries IPv6 tunnelled inside IPv4. Therefore there is no transport layer (TCP or UDP) ports to create mappings through the CGN device.

- **Manually configured IPv6 in IPv4 tunnels**
  Manually configured IPv6 in IPv6 tunnels include, for example, subscribers using IPv6 Tunnel Brokers. Such tunnels are likely to fail as they are usually (but not always) configured as IPv6 in IPv4 tunnels resulting in no transport layer (TCP or UDP) ports being available to create mappings through the CGN device.
- **Teredo**
  Teredo is designed to operate through multiple layers of NAT so would be expected to work behind CGNs. However, in testing, it has been found not to work.[9]

Deploying and providing IPv6 to subscribers behind CGNs can be beneficial to ISPs. Since many of the most used websites are now IPv6 enabled, providing subscribers with native IPv6 alongside IPv4 CGN will decrease the load on the CGN device. The more sessions that are carried over IPv6, the fewer sessions have to be supported in the state tables[10] within the CGN device and the fewer sessions have to be logged by the ISP.

All the studies and standards that we have reviewed for this report have stated that deploying IPv6 alongside CGN is beneficial. Some authors have even advocated making it mandatory that subscribers are given at least one public IP address, either IPv4 or IPv6 or both.[11]

2.5 How is CGN in Fixed Line Access Provider Network Different from CGN in Mobile Networks?

CGN has been deployed in many mobile networks. However, there are key differences between the operation of CGN in mobile networks and in fixed line networks. These differences mean that what works in the mobile network environment is not equally workable in the fixed line network environment. In particular:

- A mobile phone is one node. A fixed line subscriber usually has one or more private networks with many devices on them.
- Mobile applications are written to expect CGN in the network and, therefore, they typically use fewer sessions. Mobile devices also seldom host applications that expect, or cater for, inbound connection requests. Traffic in mobile networks is

[8] As an example, in the US some CGN deployments used UK military address space for their internal CGN network. This meant that traffic could not be routed back over the IPv4 Internet to the client's 6to4 router.

[9] *IPv6 – CGN and Teredo Considered Harmful*, 2013, http://thuktun.wordpress.com/2013/02/11/ipv6-cgn-and-teredo-considered-harmful

[10] A state table matches the external IP address and port number combinations being used by the CGN with the internal IP address being used by the provider's customer at any given time so incoming traffic can reach the correct destination behind the CGN.

[11] For instance, see the recommendations section of *Large Scale Network Address Translation*, 2013, http://www.bitag.org/report-lsnat.php

usually asymmetrical with the vast majority of the traffic initiated from the consumer handset or device.

- Mobile nodes do not usually host services. There is no need to cater for inbound connections. The majority of traffic is initiated by the mobile node.
- CGN DDoS issues are limited in mobile networks due to there being fewer instances of DDoS malware on mobile devices. However, this is beginning to be a concern to mobile operators due to the rise in the use of 3G and 4G dongles for laptop and home Internet access. In these scenarios the CGN DDoS issues are the same as those for a fixed network.

For these reasons, CGN deployment in mobile networks has been different from CGN deployments in fixed line networks. As an example, our research identified one UK mobile access provider who found that logging in their CGN network was problematic. “The size of the data is huge,” according to this provider and the cost of storing it too great. This same provider indicated that complex applications still break in mobile networks, giving the example of banking applications that rely on a fixed source IP address.

CGN in mobile networks also have interesting side-effects on mobile devices. In CGN settings the device needs to occasionally issue “keepalive” messages to retain NAT mappings. Since the radio must be restarted to issue these messages, devices connected behind CGNs see shorter battery lifetimes.

2.6 Motivations for Deploying CGNs

The main motivation for deployment of CGNs is the emerging scarcity of IPv4 addresses. ISPs have economic and engineering reasons to extend the life of their IPv4 address pool. With IPv4 address resources becoming scarce, Carrier Grade NATs allow ISPs and other providers of networks with large numbers of connected devices the ability to extend the life of IPv4-only devices and network devices that require IPv4.

In some cases there are also business benefits to be had from delaying the deployment of IPv6. For example, benefits for incumbent large service providers include:

- Extending the life of their substantial investment in IPv4
- Delaying the cost of deploying IPv6

Maximising their advantage over competitors who are unable to obtain IPv4 addresses (their existing stock of IPv4 addresses coupled with CGN will allow them to continue to provide an IPv4 service without having to obtain large numbers of additional addresses).

The implications for the customer are complex. In fact, in some cases the customer will not notice a difference. A customer behind a CGN that limits their Internet experience to basic Web browsing and electronic mail is unlikely to have significant impacts. However, a user of more advanced services (for instance, multiplayer gaming, VoIP, or media streaming) may experience a gradual increase in service and performance problems. For these users, the ISPs’ motivation for deployment is largely invisible to them, but the problems the deployment causes may range from subtle performance issues to entire applications not working.

3 Technical Overview of Carrier Grade NAT

Network Address Translation (NAT) and Carrier Grade NAT are complex technologies. The following sections provide an overview of their operation. There are many forms of NAT and CGN, this section focuses on the main features rather than providing a comprehensive description of every type of NAT and CGN.

This section also describes many of the associated technologies that have been developed to mitigate the limitations of NAT and CGN. Many of these attempt to allow protocols, services and applications that would not normally work behind NAT or CGN.

3.1 How Network Address Translation (NAT) Works

To enable communication between any two points on the public Internet, each host (or node) is uniquely identified by an address: the IP address. Traffic between nodes is then routed through the Internet based on the Destination IP address carried in the IP protocol header. The IP header also contains the Source IP address so that nodes can respond to the source of any traffic. Routing packets between nodes is the primary function of the network layer (layer 3 of the OSI protocol stack).[12] It is intended to be simple, efficient and robust.

IP datagrams can carry many different transport layer protocols with different characteristics and functionality. The most common transport layer protocols are the Internet Control Message Protocol (ICMP), the User Datagram Protocol (UDP) and the Transmission Control Protocol (TCP). Most applications use either TCP or UDP or a combination of both to communicate.

TCP and UDP include source and destination port numbers in their protocol headers. These allow applications to uniquely identify individual sessions between them. The combination of IP address and port number that identifies an individual session is called a “socket”. Some applications may only use one session but others use many. Each session is uniquely identified by a combination of source IP address, destination IP address, source port number, destination port number and protocol number (usually TCP or UDP but it could be another protocol).

Figure 3.1 below illustrates the structure of the IPv4 Internet and access networks prior to the introduction of NAT44. This is a routed network.

[12] The OSI (Open Systems Interconnect) model is a seven-layer representation of how network communications take place. Although it does not map directly to the Internet model, the IP layer in the Internet hourglass model is equivalent to layer 3—the network layer—of the OSI model. For more information, see Information technology – *Open Systems Interconnection – Basic Reference Model: The Basic Model, ISO*/IEC 7498-1:1994, http://standards.iso.org/ittf/PubliclyAvailableStandards/s020269_ISO_IEC_7498-1_1994(E).zip, 1994.

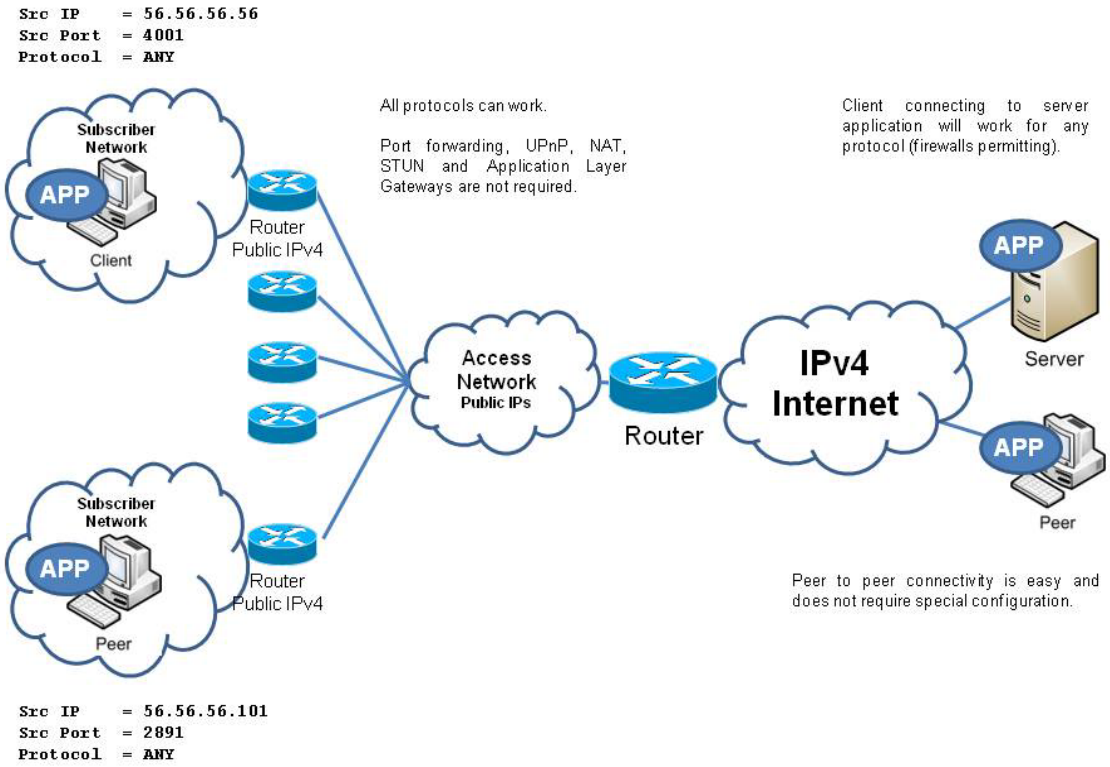

Figure 3.1 - The Internet and Access Networks Prior to the Introduction of NAT44

NAT44 sits at the edge of a private network. The concept of a private network was invented to allow for the deployment of NAT44. The most common private network that deploys NAT44 is a broadband consumer network attached, via the NAT, to an Internet access provider. Within this and all other NAT44 connected networks, private IPv4 addresses are used. These addresses cannot be routed on the global IPv4 Internet. The NAT44 device is configured with at least one global IPv4 address on its Internet-facing interface and at least one private address on its internal interface that is connected to the private network.

The NAT44 device appears to nodes on the private network as a router that forwards their traffic to the rest of the Internet. Unlike a router, the NAT44 device not only forwards the datagrams, it also modifies them. The NAT44 device changes internal private addresses to its external public address. Since the external address is one address and there may be many nodes on the private network, the NAT44 device has to have a way of uniquely identifying which internal node traffic is destined for when its receives a packet from the Internet. To do this, the NAT44 device has to go beyond the network layer protocol headers (IP headers) and modify the unique identifiers (port numbers) in the transport layer (typically UDP or TCP). It keeps a record of the mapping between internal IP + internal port and external IP + external port for *every* session and every transport layer protocol (typically TDP and UDP).

Figure 3.2. below shows an IPv4 access network where NAT44 has been introduced to conserve IPv4 addresses.

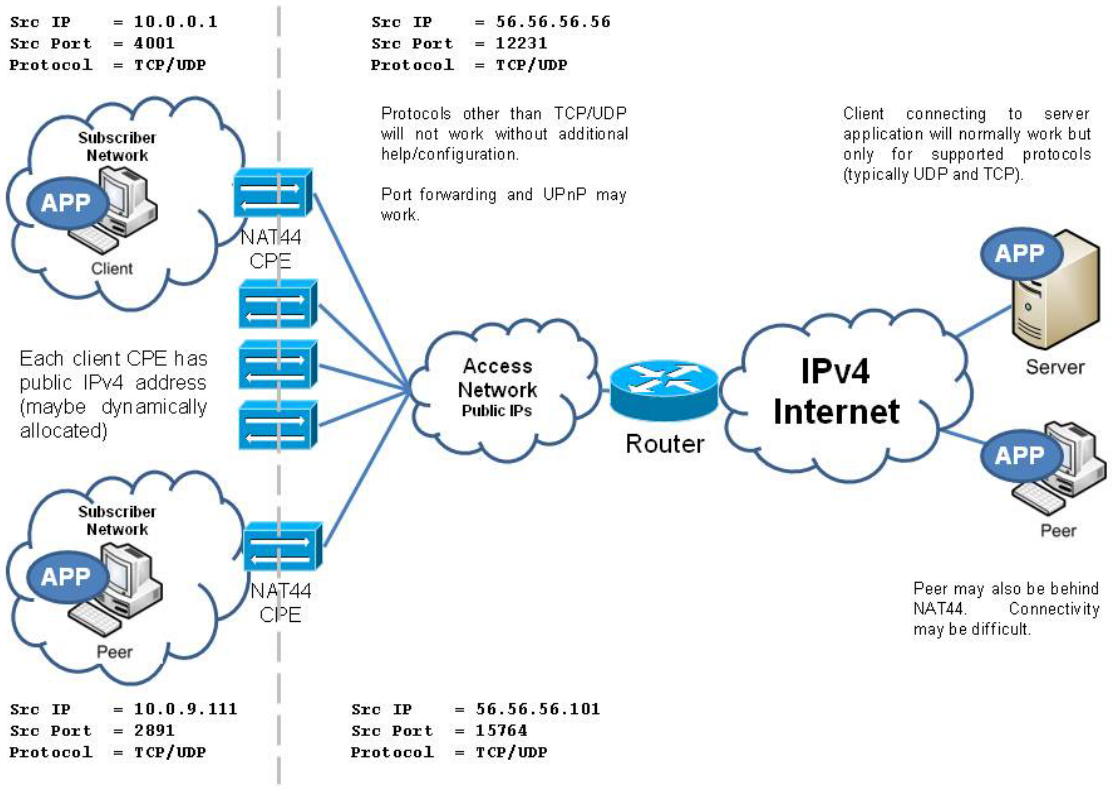

Figure 3.2 - The Internet and Access Networks with the Introduction of NAT44

As packets pass through the NAT44, device IP and transport layer headers are modified so that they have valid ports and addresses for the internal and external networks.

This form of NAT44 requires sessions to use transport layer protocols that have port numbers (typically UDP or TCP). Furthermore, it also requires that higher layer protocols do not reference IP addresses or port numbers. Unfortunately, many transport layer protocols do not have port numbers and many application layer protocols do reference IP addresses and port numbers. Consequently NAT44 does not work for all IP traffic.

3.2 How CGN Works

CGN uses another layer of NAT44 placed in the carrier access network rather than at the edge of the subscriber's network. This is why it is sometimes referred to as NAT444.

In CGN, subscriber datagrams are modified more than once. They are modified at the subscriber edge and in the carrier's access network. In both cases the IP addresses and ports are mapped between internal and external values. The subscriber's traffic uses three different source addresses as the traffic moves from the subscriber's internal network, to the ISP's access network and finally to the global IPv4 Internet. This moves the public IPv4 address from the subscriber's NAT device to the carrier's CGN device.

The purpose of the double NAT architecture is to allow carriers to share a single IPv4 address amongst many subscribers. The operation of CGN is shown in Figure 3.3.

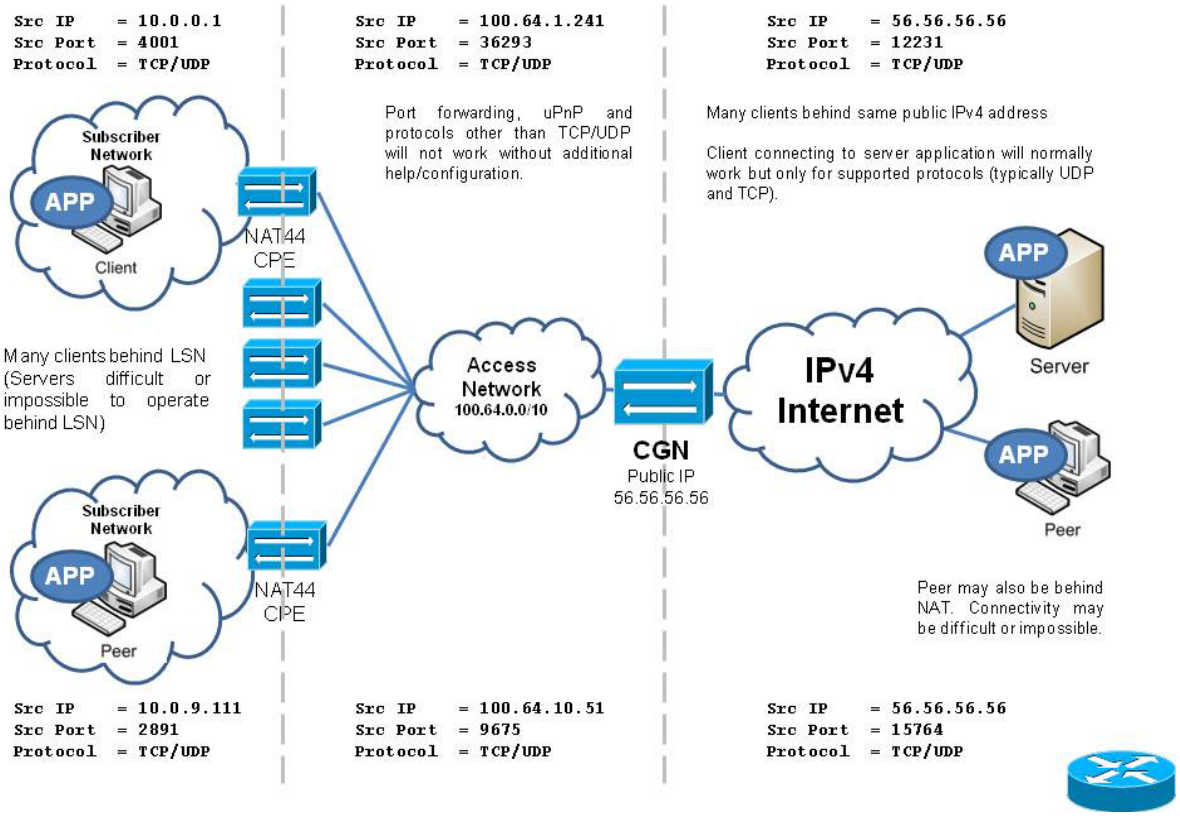

Figure 3.3 - The Internet and Access Networks with the Introduction of CGN

3.3 Associated Technologies

Both NAT and CGN have an impact on the traffic that traverses them. Many protocols and the services or applications that use them are broken by NAT. There has been significant effort put into making protocols that would otherwise not work through NAT work. Whole business have been built upon applications that can traverse NAT (for example Skype which would have not even been necessary in a NAT free Internet).

Many of the techniques that have been used to overcome the limitations of NAT are broken by CGN. As a result significant effort is being put into ways of extending these techniques so that they can work with CGNs. This section looks at some of these techniques. It should be noted that many NAT devices include proprietary functionality to overcome specific problems. For example, some NAT devices even include NAT traversal support for specific games.

3.3.1 Universal Plug and Play (UPnP)

UPnP[13] is a set of standards that define an architecture which enables devices to seamlessly connect to each other in local networks. UPnP includes a solution for NAT traversal called UPnP Internet Gateway Device Protocol (UPnP-IGD).

UPnP-IGD can be used by applications to open up ports in a subscribers firewall and configure port forwarding. This can mitigate some NAT44 traversal issues especially where services are running behind NAT44 and need to be accessed from the global Internet.

UPnP-IGD has a number of weaknesses, including security issues.[14]

[13] UPnP standards can be found at the UPnP Forum, http://www.upnp.org.

By default, UPnP-IGD does not work behind CGN. Standards work is in progress to develop a mechanism that allows CPEs to convert between UPnP on their internal interface and the Port Control Protocol (PCP) on their external interface.[15] In this way, if the CGN device supports PCP then UPnP-IGD may be able to work behind it.

UPnP may be used in IPv6 networks although it is not required to configure NAT traversal.

3.3.2 NAT Port Mapping Protocol (NAT-PMP)

NAT-PMP is an alternative to UPnP-IGD produced by Apple and defined in an IETF draft.[16]

3.3.3 Session Traversal Utilities for NAT (STUN)

Session Traversal Utilities for NAT (STUN)[17] is a protocol that helps other protocols manage the problems of NAT traversal. STUN can be used to:

- Determine the public IPv4 address and transport layer port allocated to a session by a NAT device.
- Check connectivity between two endpoints.
- Act as a keepalive to maintain NAT mappings.

STUN provides a set of tools used to find information that can be used by NAT traversal techniques, it does not provide NAT traversal itself. In order to operate, STUN requires accessible STUN servers on the public Internet.

STUN is not required in IPv6 networks.

3.3.4 Traversal Using Relays around NAT (TURN): Relay Extensions to Session Traversal Utilities for NAT (STUN)

Traversal Using Relays around NAT (TURN)[18] is used when peers located behind NAT devices cannot communicate directly and are therefore forced to communicate through a relay device. TURN was designed to be used with Interactive Connectivity Establishment (ICE) but can be used on its own.

TURN is not required in IPv6 networks.

---

[14] There are many references regarding the security weaknesses in UPnP. For example, see CERT, *Vulnerability Note VU#357851: UPnP requests accepted over router WAN interfaces*, http://www.kb.cert.org/vuls/id/357851, 2012, accessed 15 April 2013.

[15] M. Baugher, E. Nedellec, M. Saaranen & B. Stark, *Internet Draft: IPv6 Services for UPnP Residential Networks*, draft-bnss-v6ops-upnp-01.txt, http://tools.ietf.org/html/draft-bnss-v6ops-upnp-01, 2009, accessed 15 April 2013.

[16] S. Cheshire & M. Krochmal, *Internet Draft: NAT Port Mapping Protocol (NAT-PMP)*, draft-cheshire-nat-pmp-07, http://tools.ietf.org/html/draft-cheshire-nat-pmp-072 2013, accessed 15 April 2013.

[17] J. Rosenberg, R. Mahy, P, Matthews & D. Wing, *Session Traversal Utilities for NAT (STUN)*, RFC 5389, https://tools.ietf.org/html/rfc5389, 2008, accessed 15 April 2013.

[18] , R. Mahy, P. Matthews & J. Rosenberg, *Traversal Using Relays around NAT (TURN): Relay Extensions to Session Traversal Utilities for NAT (STUN)*, RFC 5766, https://tools.ietf.org/html/rfc5766, 2010, accessed 15 April 2013.

3.3.5 Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols

Interactive Connectivity Establishment (ICE)[19] is a protocol designed to enable UDP based multimedia sessions to traverse NAT. ICE uses STUN and TURN.ICE is specific to offer/answer protocols and was designed to assist in situations where devices behind a NAT have to register with a server in the public Internet. An example of this is proxy registration on a SIP server by a device behind a broadband CPE NAT.[20]

ICE is not required in IPv6 networks.

3.3.6 Proprietary Techniques

There are many proprietary techniques in existence to make it possible to contact peers and services located by NAT. These are beyond the scope of this report. However it is worth noting that this could be a problem for CGN implementations as there is no way of guaranteeing that CGN will not break these techniques.

3.3.7 Port Control Protocol (PCP)

The Port Control Protocol (PCP) is designed to provide a standard protocol for controlling mappings between external IP addresses and ports and internal IP addresses and ports. It is specifically designed to work in the context of CGNs. PCP can also be used to reduce the need for NAT-keepalives.[21]

For PCP to work, the client application, subscriber CPE NAT and CGN device must all be PCP aware and configured to allow PCP to configure port mappings.

In the context of CGNs, PCP is the only realistic way for a subscriber to configure port forwarding from the ISPs CGN device to a service in their network. During interviews with ISPs and CGN experts, the view was repeatedly expressed that most ISPs will be reluctant to allow their customers to configure port forwarding on their CGNs. Despite this, much work is being done on PCP by ISPs in Europe, for example, Orange in France and Telecom Italia.[22]

PCP can be used with IPv6 but it is usually not required.

3.3.8 Application Layer Gateways

Application Layer Gateways (ALGs) are necessary when an application layer protocol has features that depend upon the IP address at the application layer. These protocols carry IP addresses at the application layer, which is outside of the IP header. Since most NAT devices only modify the IP header and transport layer headers, they may not detect and translate IP

[19] J. Rosenberg, *Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols*, RFC 6336, http://tools.ietf.org/html/rfc5245, 2010, accessed 15 April 2013.

[20] For a long description of how ICE interacts with VoIP clients, see J. Rosenberg, *Interactive Connectivity Establishment*, http://www.internetsociety.org/articles/interactive-connectivity-establishment, 2006.

[21] See Section 4.1.2 for an example of the potential implication of CGN keepalives on devices in mobile networks.

[22] For example, see France Telecom's report on work in this area: M. Ait Abdesselam, M. Boucadair, A. Hasnaoui & J. Queiroz, *Internet Draft: PCP NAT64 Experiments*, draft-boucadair-pcp-nat64-experiments-00, http://tools.ietf.org/html/draft-boucadair-pcp-nat64-experiments-00, 2012.

addresses at the application layer. Even if they were to do so, the IP address may be embedded in the application layer protocol in such a way that it would still fail to work correctly. Examples of such application layer protocols include:

- File Transfer Protocol (FTP)[23]
- H.323[24]
- Session Initiation Protocol (SIP)[25]
- Domain Name System (DNS)

For some of these protocols there is no easy way to translate datagrams so that they can traverse a NAT device. A classic example is FTP which carries destination and port information in the payload of the packet. Since a NAT is going to change only information in the header, the header and payload information fail to match and the application breaks.

Application Layer Gateways attempt to solve this problem by doing deep packet inspection of the packets as they go by and, for the protocols that the ALG is designed to support, make the adjustments needed to the network layer address information found in the payload of the packet.

A challenge for designers of NATs, and CGNs in particular, is that as new protocols emerge and new services are designed for the Internet, they may require the development of new ALGs. In the absence of an ALG, a new technology may not operate correctly in the CGN environment. Research carried out for this paper also showed that the languages and APIs used for building ALGs for CGNs were varied. One ISP, trailing CGN in the UK in 2013, found that they had to develop custom rule sets for certain applications. This adds overhead and cost to the deployment of the CGN. In addition, in many cases, the languages used for building ALG rulesets can be very arcane and, therefore, difficult and expensive to find staff that can write the required code.

[23] The File Transfer Protocol (FTP) is a IETF standard that allows the transfer of files between nodes on the Internet. It is a fundamental building block of many applications and tools that need to move large quantities of data between nodes connected to the Internet. For more information see: J. Postel, J. Reynolds, *File Transfer Protocol*, http://tools.ietf.org/html/rfc959, 1985.

[24] H.323 is an International Telecommunications Union T Sector (ITU-T) standard for the protocols that provide audio-visual communication sessions on any packet-based network, including the Internet. For more information, see: ITU, *H.323: Packet-based multimedia communications systems*, http://www.itu.int/rec/T-REC-H.323/en, 2008.

[25] SIP is an IETF-defined protocol associated with voice and video communications over the Internet. For more information, see: J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley & E. Schooler, *SIP: Session Initiation Protocol, RFC 2361*, https://tools.ietf.org/html/rfc2361, 2002.

4 Technical Implications of Using Carrier Grade NAT

This section describes the technical implications of using CGNs. Before proceeding to consider the implications of CGN it is worth briefly summarising the implications of using traditional NAT44. These are well documented in many places and many are summarised in the informational RFC 2993.[26]

4.1 Session Limitations

Internet based applications and services commonly use one or more TCP or UDP sessions to communicate with a client. These sessions are uniquely identified by a combination of five pieces of information:

- Source address
- Destination addresses
- Transport layer protocol (typically TCP or UDP)
- Source port number
- Destination port number

This set of information is called a socket. (Note that this can be different when a protocol other than TCP or UDP is in use.)

4.1.1 Limitations on the Number of Sessions

There are many situations where a client node will use multiple sessions for one application. The reasons for this include:

- **Application performance**
  Many sessions working in parallel potentially to multiple servers can improve throughput. For example, BitTorrent and many modern web browsers.
- **Application functionality**
  Different parts of the application use different sessions for different functions. For example, AJAX applications such as Google Maps and many webmail applications.
- **Command and control**
  A session or sessions can be used for control purposes whilst another or other sessions are used to carry data. For example, File Transfer Protocol (FTP).

TCP and UDP have limits on the number of ports they can use: 2^{16} or 65,536 ports. Not all ports are available as some are reserved for specific uses. These reserved ports are called “well-known ports”. Web browsing, for example, uses the well-known port 80, the HTTP port. Even with the well-known ports removed from the pool available for a CGN deployment, there are still a large number of ports available, which is more than adequate for most applications.

Behind traditional NAT, the available source ports are shared amongst all clients behind the same NATted public IP address. For most consumer subscribers this is more than sufficient. In

[26] T. Hain, *Architectural Implications of NAT*, RFC2993, https://tools.ietf.org/html/rfc2993, 2000, accessed 15 April 2013.

some enterprise scenarios where there are many devices behind the NAT, this may become a problem.

In the case of CGN, the source port addresses are shared amongst more than one subscriber. In this case the CGN sets a maximum on the number of sessions a subscriber may use. Depending on the number of subscribers that are sharing a single public IP address on the CGN device, this may result in the subscriber being limited to tens, hundreds or thousands of addresses.

Scenario	Maximum Number of Sessions (TCP)	Comment
Routed	65,536 **per node**	Not all ports are available for client use
Subscriber CPE with NAT44	65,536 **per subscriber**	Not all ports are available for client use
CGN	Maximum set by carrier. Depends on CGN implementation and configuration.	
Maximum 10 subscribers per IP Maximum 100 subscribers per IP Maximum 1000 subscribers per IP	6,536 **per subscriber** 653 **per subscriber** 65 **per subscriber**	Not all ports are available for client use. For example, well known ports.

Table 4.1 - Maximum Number of TCP Sessions Illustrative Estimates

Table 4.1 is an illustration of how the maximum number of sessions available per subscriber is affected by the deployment of CGN. As noted earlier in this section, not all 65,536 ports are available for client ports as some ports are reserved. For the purposes of capacity planning, some estimates use the figure of 30,000 maximum ports per IP address.[27,28] In Table 4.1, we have only considered TCP sessions, there may be an equivalent number of UDP sessions available. The CGN examples illustrate the maximum number of sessions that an ISP may set for a subscriber.

Exhausting the maximum number of sessions per subscriber in a CGN or the total maximum number of sessions available in a CGN device will cause any further sessions to fail. In both cases, the user experience can vary from, best case poor performance to worst case application failure. In between, users may experience intermittent failures which users and support staff would find very hard or impossible to diagnose. This is already the case in mobile networks where some applications (for example Skype) may work some of the time and fail on other occasions for no apparent reason.

4.1.2 Limitations on the Lifetimes of Sessions

The lifetime of sessions has an impact on the amount of state that a CGN device is required to store. Some sessions are short-lived, others can last indefinitely[29]. A CGN must keep state for a session as long as that session is required. Applications operating through NAT or CGN

[27] K. Sundarasan, *CGN Architectures & Impacts*, TXv6TF Summit, http://www.txv6tf.org/wp-content/uploads/2011/04/KarthikS_CGNArch_Sep142011.pdf, 2011, accessed 15 April 2013.

[28] C. Donley, L. Howard, V. Kuarsingh, J. Berg & J. Doshi, *Internet Draft: Assessing the Impact of Carrier-Grade NAT on Network Applications*, draft-donley-nat444-impacts-03, http://tools.ietf.org/html/draft-donley-nat444-impacts-03, 2011, accessed 15 April 2013.

[29] In principle, TCP sockets can last forever.

devices are often modified and therefore send extra—normally unnecessary—datagrams called "keepalives" to maintain sessions when there is no application data to be sent.

Even web applications can keep sessions open for minutes at a time. Figure 4.2. shows a web page accessed through Firefox and Internet Explorer (IE). Both browsers use many sessions to access the web page. Furthermore, these sessions have lifetimes in this example of a couple of minutes.

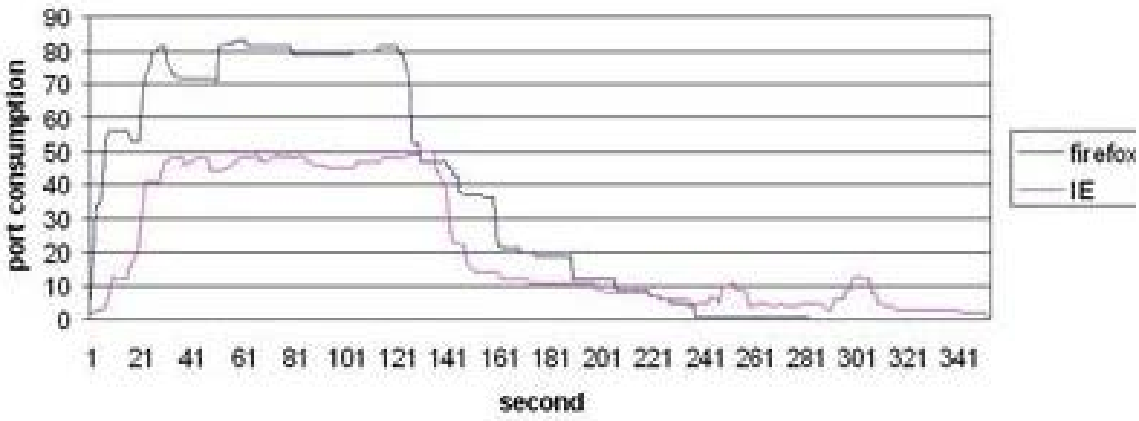

Figure 4.1 - Port Consumption Comparison Firefox and IE to http://www.youku.com[30]

This makes it very difficult for an ISP to predict the instantaneous peak number of sessions that their subscribers will need in order to have a reliable service. Estimating the peak number of sessions is essential when capacity planning in order to choose the number of CGN devices an ISP requires and the efficiency with which IPv4 addresses can be shared.

In mobile networks, the characteristics of CGN sessions are different from the characteristics of sessions in CGN deployments in fixed networks. To reduce data costs and conserve mobile phone battery life, mobile data sessions are typically shorter lived than those on fixed networks, where Internet connectivity may be on 24 hours a day every day.[31] To cater for the way mobile device users access data in shorter bursts than their counterparts on fixed networks, mobile data applications are often designed to have very short lived sessions. As a result, CGNs in mobile networks have a much higher session churn rate than in fixed line deployments.

For example, where the www.youku.com data session illustrated in Figure 4.1 requires around 80 ports to be used for two minutes in a fixed network environment, in a mobile data session, those ports would only be used for perhaps 30 seconds, leaving those ports available to other users for the remaining one minute 30 seconds. In other words, where in a two-minute interval, a fixed network operator could serve only one user of an 80-port consuming data session, a mobile network operator could accommodate four users. The shortened lifetimes of data sessions in the mobile environment, therefore, allows mobile operators to run their CGNs at a much higher capacity than is realistic for fixed line scenarios.

[30] *Application behaviors in interms [sic] of port/session consumptions on NAT*, http://opensourceaplusp.weebly.com/implementations--experiments-results.html, n.d., accessed 15 April 2013.

[31] It is desirable to keep sessions short and to power down the radio. However, when sessions have long lasting mappings in the CGN, those mappings must be maintained—and so keepalives are required. As a result, packets (keepalives) must be transmitted, hence the radio transmitter has to be powered up, and this is what affects battery life in the handset.

4.2 Protocol Limitations

Traditional NAT maps between internal and external combinations of addresses and ports, changing the datagram headers in the process. This mapping can only occur with IP datagrams that contain transport layer protocols that have port numbers. Usually NAT devices provide support for the two main transport layer protocols: TCP and UDP. However, there are many other transport layer protocols that can be carried by IP datagrams, many of which do not have the concept of port numbers. Some of these other transport layer protocols can be mapped by NAT boxes and some cannot. Significant effort has been put into getting some protocols to work through NAT devices that otherwise would not work.

Examples of protocols that may be problematic when NAT is involved include:

- IPsec (see Section 4.2.1)
- IP tunnels and IPv6 transition mechanisms (see Section 4.7)
- Real-Time Transport Protocol (RTP)[32]
- Internet Control Message Protocol (ICMP)[33]

There are numerous other examples of application specific protocols, but many of these are rarely required in subscriber networks.

4.2.1 IPsec

Internet Protocol network layer security can be provided by IPsec (Internet Protocol Security). IPsec provides security at the network layer, that is it secures the Internet Protocol and its payload (all higher layer protocols, including application data). It does this through the introduction of two new headers that authenticate the source of traffic:

- **Authentication Header (AH)**
  Protects against modification of the IP headers and payload.
- **Encapsulating Security Header (ESP)**
  Protects against the modification of payload and encrypts the payload.

These headers can be used separately or in combination. In combination, the AH header authenticates the packet and ensures that the IP header is not modified and the ESP header's primary purpose is to encrypt the contents of the packet.

IPsec can also use additional protocols such as the Internet Key Exchange (IKE) protocol to manage the configuration of Security Associations (SAs) and session keys.

IPsec is used in one of two modes; tunnel or transport mode. In tunnel mode traffic is tunnelled and the tunnel is secured by IPsec. In transport mode, traffic is not tunnelled and the each individual datagram is secured directly using IPsec headers.

[32] RTP is an IETF-defined protocol that standardises the packet format for delivering audio and video over IP-based networks. For more information, see: H. Schulzrinne, S. Casner, R. Frederick & V. Jacobson, *RTP: A Transport Protocol for Real-Time Applications*, RFC 3550, http://tools.ietf.org/html/rfc3550, 2003.

[33] ICMP is an early IETF-defined protocol that is used to send certain types of error messages when communication cannot be established between two Internet hosts. For more information, see: J. Postel, Internet Control Message Protocol, RFC 792, http://tools.ietf.org/html/rfc792, 1981.

IPsec headers appear between the network layer (the IP header) and the transport layer (for example, TCP or UDP). These headers do not contain port numbers.

There are a number of problems when using IPsec behind NAT44.[34] The major problems are summarised below:

- The ESP header encrypts the transport layer so transport layer ports are not visible to the NAT device for the purposes of mapping or translation. Therefore there is no way to differentiate between multiple ESP secured sessions behind NAT.
- The AH header authenticates the IP header, so any changes to the IP header (such as IP address mapping) or changes to transport layer header (such as port mapping[35]) will cause authentication to fail.
- IKE cannot traverse NAT without the help of IKE NAT traversal functionality.[36]

These problems are the main reason that IPv4 IPsec VPNs use tunnel mode and are usually terminated at the edge of a network before traversing any NAT devices. Securing traffic from within a network behind NAT44 using IPsec is often problematic and sometimes impossible.

In the case of traditional NAT, it is commonly the case that one IPsec session (possibly a VPN) will work behind NAT but any further sessions cannot be distinguished from the first and will fail. In the case of CGN, IPsec might work for one subscriber behind a CGN's external address but other subscribers would not be able to use IPsec. This means that for all practical purposes IPsec is not viable behind CGN.

IPsec is used not only for VPNs, but also in other products and applications. As a consequence, all applications that use IPSec would fail behind NAT. For example, Vodafone's SureSignal mobile femtocell uses IPsec to secure connections back to Vodafone's network. This would not be possible behind CGN.[37]

4.3 Packet Header Modification in Access Network Rather than at Subscriber Edge

The use of CGN changes the nature of an access network from a transparent network that carriers a subscriber's traffic unchanged to one where the subscriber's datagrams are modified in transit.[38]

Currently subscribers have control over how they modify their datagrams at their network edge. CGN moves control of packet modification to the ISP.

---

[34] B. Aboba & W. Dixon, *IPsec-NAT Compatibility Requirements*, RFC 3715, http://tools.ietf.org/html/rfc3715, 2004, accessed 15 April 2013.

[35] Port mapping is essential to NAT so a private IP address can be linked to the public IP address and port number combination used to identify the private host when communicating with the public IPv4 Internet.

[36] T. Kivinen, B. Swander, A. Huttunen & V. Volpe, *Negotiation of NAT-Traversal in the IKE*, RFC 3947, http://tools.ietf.org/html/rfc3947, 2005, accessed 15 April 2013.

[37] See the *SureSignal Network Setup Guide*, which documents the requirement to have a DHCP provided address that can support IPsec at: http://www.vodafone.co.uk/cs/groups/public/documents/webcontent/vftst061123.pdf

[38] Historically, modifications to packets in transit are associated with man-in-the-middle attacks, where an attacker impersonates a trusted source to gain access to data that is not meant to be shared with the attacker.

4.4 Multiple Users Sharing the Same IPv4 Address

Sharing public IPv4 addresses across multiple subscribers is the primary purpose of CGN. Sharing public IPv4 addresses has implications for:

- Applications such as financial applications that track the source IPv4 addresses for security reasons
- IP reputation where one subscriber carries out abuse using an IPv4 address that is shared with other innocent subscribers
- Subscribers running services such as remote access to webcams, PVRs and other devices

4.5 Port Forwarding Limitations

The Internet was designed to be bi-directional where traffic can flow equally in either direction. This means that sessions can be initiated from inside a subscriber's network out to the global Internet and from the global Internet into the subscriber's network.

This is important because many services and applications depend upon the ability to connect into a subscribers network. Examples include:

- **Peer-to-peer gaming**
  Where connections can be initiated from other peers in a multiplayer game to the subscriber's network.
- **VoIP services**
  Where the VoIP service can receive incoming calls and/or make direct calls between subscribers.
- **Personal Video Recorders (PVRs)**
  Where subscribers can remotely access and control their PVR from any location on the Internet.
- **Home security**
  Where subscribers can access security IP webcams and other security devices remotely.
- **Home automation**
  Where subscribers can view and control appliances in the home remotely. For example, control of central heating or security systems.
- **Mobile femtocells**
  For example Vodafone's SureSignal which provides subscribers mobile coverage in areas outside the reach of Vodafone's mobile network using a small base station (femtocell) connected to the subscriber's broadband connection.

In NAT44 environments, to make applications and services such as the examples above work, a number of techniques have been developed to traverse the NAT device. These techniques are necessary because devices in the subscriber's network are hidden behind the NAT device. The only device that is reachable on the global Internet is the NAT device but only it has a public IPv4 address.

In order to reach a service or device inside a subscriber's network, a method of traversing the NAT device must be used. The most basic method is to manually configure the NAT device to forward traffic to a specific port on its public IPv4 address to an internal IPv4 address and port

number. This is called Port Forwarding. Port forwarding only works for protocols that have port numbers such as UDP and TCP.

Manually configuring a NAT device to forward ports can be complicated for the average user. To reduce or remove the need for a user to configure port forwarding a number of techniques have been invented to automatically traverse NAT or help with NAT traversal. These include:

- Universal Plug and Play (UPnP)
- Session Traversal Utilities for NAT (STUN)
- Interactive Connectivity Establishment (ICE)
- Traversal Using Relays around NAT (TURN)
- Various Proprietary Techniques

All of these techniques are impacted by CGN. The CGN, which is controlled and managed by the Internet Service Provider, replaces the control the subscriber had, in a non-CGN environment, over their public IPv4 address and capability to configure port forwarding. With the loss of this control, subscribers may no longer be able to use the above techniques to ensure uninterrupted access to services such as peer-to-peer gaming and home security.

Many of the parties that were interviewed for this report expressed the opinion that ISPs would be reluctant to allow their customers to manually configure port forwarding through their CGN devices. Despite this we did find examples of ISPs who would consider providing this service to their customers. Two solutions were mentioned:

- Providing customers with a web page based configuration
- Using techniques to extend UPnP to work behind CGNs (in conjunction with PCP)[39]

An important point to note with automated solutions using UPnP and/or PCP is that they require upgrades or replacements to equipment at customers' premises.

4.6 Public, Private (RFC 1918) or Shared (RFC 6598) Address Space

When deploying CGN, ISPs must decide on the address space to use in their access network, that is between the subscriber edge and the CGN device. Historically both public IPv4 address space and private IPv4 address space have been used. However, public and private address space have a number of problems. These are discussed in RFC 6598[40] and summarised here:

- **Legitimate public IPv4 address space**
  This negates the purpose of CGNs, that is to preserve public IPv4 address space.
- **Illegitimate public IPv4 address space**
  That is address space not owned by the ISP ("squat space"). This will make it impossible for subscribers to route to the network where the squatted address space is legitimately deployed. Also, this approach may force protocols such as 6to4, into a failure mode leading to connectivity issues and failures.

[39] M. Boucadair, R. Penno, D. Wing & F. Dupont, *Universal Plug and Play (UPnP) Internet Gateway Device (IGD)-Port Control Protocol (PCP) Interworking Function*, draft-bpw-pcp-upnp-igd-interworking-02, http://tools.ietf.org/html/draft-bpw-pcp-upnp-igd-interworking-02, 2011, accessed 15 April 2013.

[40] J. Weil, V. Kuarsingh, C. Donley, C. Liljenstolpe & M. Azinger, *IANA-Reserved IPv4 Prefix for Shared Address Space*, RFC6598, http://tools.ietf.org/html/rfc6598, 2012, accessed 15 April 2013.

- **Private IPv4 address space**
  There are two main problems with using private IPv4 address ranges, defined in RFC 1918,[41] in the access network. Some Customer Premise Equipment (CPE) NAT devices will not operate if the external interface has a RFC 1918 address. If a subscriber allocates the same RFC 1918 address space to their internal network as the ISP is using on the access network NAT is likely to fail.

For these reasons IANA reserved IPv4 address space specifically for use within CGN access networks. This space is called the Shared Address Space and has the IPv4 address prefix 100.64.0.0/10.[42]

This allocation is not without its own problems. Specifically because it is a /10 prefix it is much smaller than a /8 obtained by other means (for example, using the private address space 10.0.0.0/8). The size of the prefix has an impact on the maximum size of a network that can be placed behind a single CGN. For large ISPs this will mean that networks will have to be split into regions using overlapping Shared Address space and each having their own CGNs.

4.7 Impact on the IPv6 Transition

CGNs do not have to have any negative impact on the IPv6 transition. If CGNs are deployed by the ISP as a part of a dual-stack access network then native IPv6 connectivity can be provided to subscribers along with IPv4 connectivity through CGNs. This approach has the additional advantage in that it provides subscribers with a NAT free path to the IPv6 Internet.

However, CGNs can have a negative impact on the transition to IPv6 where the ISP does not provide a native IPv6 service to their subscribers. In an environment without CGNs, a subscriber could use IPv6 transition techniques such as 6to4, Teredo and Tunnel Brokers to access the global IPv6 Internet. When CGNs are deployed, techniques 6to4,[43,44] Teredo and Tunnel Brokers are likely to stop functioning when they hit a CGN interface.[45] This means that the only possible methods by which subscribers can gain access to the IPv6 Internet are prevented from functioning by CGNs.

4.8 Service Performance and Reliability

In an access network that does not have CGNs deployed, traffic reaches the Internet through routers. Usually, these have multiple paths to the Internet providing a high degree of availability and performance.

---

[41] Y. Rekhter, B. Moskowitz, D. Karrenberg, J. de Groot & E. Lear, *Address Allocation for Private Internets*, RFC 1918, http://tools.ietf.org/html/rfc1918, 1996, accessed 15 April 2013.

[42] IANA, *IANA IPv4 Special Purpose Address Registry*, http://www.iana.org/assignments/iana-ipv4-special-registry/iana-ipv4-special-registry.xml, 2013, accessed 15 April 2013.

[43] B. Carpenter, *Advisory Guidelines for 6to4 Deployment*, RFC 6343, http://tools.ietf.org/html/rfc6343, 2011, accessed 15 April 2013.

[44] Some work has been carried out to provide a mechanism by which 6to4 could work behind CGN. For example, see: V. Kuarsingh, Y. Lee & O. Vautrin, *6to4 Provider Managed Tunnels*, RFC 6732, http://tools.ietf.org/html/rfc6732, 2012, accessed 15 April 2013.

[45] C. Donley, L. Howard, V. Kuarsingh, J. Berg & J. Doshi, *Assessing the Impact of Carrier-Grade NAT on Network Applications*, draft-donley-nat444-impacts-05, http://tools.ietf.org/html/draft-donley-nat444-impacts-05, 2012, accessed 15 April 2013.

In the case where CGNs have been deployed in the access network, performance and reliability may be affected. In particular, CGN devices may:

- Be a single point of failure
- Be a choke point
- Be vulnerable to resource exhaustion

These effects can be mitigated to some extent.

The single-point of failure problem can be addressed through the deployment of CGN devices with high availability functionality. Since CGNs, unlike routers, maintain state, true high availability must include the ability to keep the state necessary to ensure that subscriber connections are not lost.

The choke point problem can be addressed by deploying CGN devices with sufficient capacity for all the subscribers located behind it. Modern CGN devices can support tens of thousands of subscribers. It is also possible to locate the CGN in a variety of places in the access network and is some scenarios this limits the impact of having a choke point in the network.

4.9 Logging

Identifying the source of Internet traffic is a legal requirement for the lawful interception of traffic and for other law enforcement activities. It is also necessary to be able to identify the source of network abuse. The process of recording network information, such as the source of Internet traffic, is called "logging". The standard mechanism for logging network information is called Syslog.

In an Internet without CGN, the source of traffic can easily be identified using the source IPv4 address.[46] When a subscriber is using NAT44, the source IP address will at least identify the source subscriber if not the source node within the subscriber's network.

In access networks where dynamic IP addresses are provided to subscribers, the mapping of the dynamic IP address to the subscriber has to be logged in order to identify the source of traffic.

Logging becomes significantly more complicated when CGN is involved. In CGN, the IP address is not sufficient to identify the source of traffic. This is because IP addresses are shared across many subscribers. In order to identify a particular subscriber behind CGN, it is necessary for the service provider to be able to map the subscriber's internal source IP address and internal source port to the CGN's external source IP address and external source port *for every session* used by the subscriber.

Worse, for law enforcement agencies and/or others to be able to request the identity of the subscriber, they must be able to provide the service provider with the source IP address and *source port* and *the specific time* an incident took place.

[46] Assuming, of course, that source has not been "spoofed".

4.9.1 Logging Requirements

The requirements for identifying the source of Internet traffic are shown in Table 4.2.

Access Network	Service Provider Logging Requirements	Logging Requirements at Destination
Routed	None (fixed record of allocation)	Source IP address
Subscriber CPE with NAT44	None (fixed record of allocation)	Source IP address
Subscriber CPE with NAT44 (Dynamic public IP)	Subscriber dynamic IP address (typically changes daily)	Source IP address + date and timestamp
CGN	Per session: • Date and time • Internal IP address (may be dynamic) • Internal source port • External CGN source IP address • External CGN source port number	Per session: • Date and time • Source IP address • Source port number

Table 4.2 - Comparison of Logging Requirements

In lab testing by CableLabs,[47] it was found that a typical CGN log message was approximately 150 bytes long. A typical household in the US was found to have an average of 33,000 sessions per day. For an ISP with one million subscribers, this will generate approximately 150 terabytes of log data per month or 1.8 petabytes per year (1,800,000,000,000,000 bytes). The logging would require approximately 23 Mbps of bandwidth between the CGN devices and the logging servers.[48] In our conversations with one provider, the ISP indicated that "it was impractical to trace back sessions to subscribers." They indicated that they had no intention of trying.

This volume of data is substantial and has the following implications:

- **Infrastructure**
  The ISP will require significant infrastructure to log and store this data over the regulatory periods required. (Datacentres, redundancy, backups, suitable querying or indexing toolset/software disk drives, power etc).
- **Delays**
  Retrieving the data is likely to be delayed by the time taken to search the logs for the specific records.
- **Additional requirements for trace-back requests**
  Content providers and others will be required to provide the ISP with additional information to be able to determine the correct log records. Specifically in addition to the source IP address, they must also provide the source port number at a specific time. This is a new logging requirement -- currently many only log the source IP address.

[47] C. Donley, C. Grundemann, V. Sarawat, K. Sundaresan & O. Vautrin, *Deterministic Address Mapping to Reduce Logging in Carrier Grade NAT Deployments*, draft-donley-behave-deterministic-cgn-05, http://tools.ietf.org/html/draft-donley-behave-deterministic-cgn-05, 2013, retrieved 15 April 2013.

[48] The CableLabs study (see Footnote 32 above) is extremely useful, but they did not test some important configurations. For example, the study did not test two consoles behind NAT, two behind CGN and two behind different ISPs CGNs. Also, the study's use of "pass" in their tables could be misleading. Some entries show Xbox as a pass when we know that it won't work under certain circumstances (see their update in the same report).

Experience of logging and trace back in mobile networks is that it has been too difficult and too costly to implement.[49] In addition, their legal experts have raised concerns about the potential privacy issues of the level of logging detail required for trace back through CGN.

4.9.2 Reducing the Logging Requirements

Logging requirements can be reduced using a number of techniques. One such proposal is Deterministic CGN.[50] Deterministic CGN can reduce the logging requirements by "orders of magnitude". It is very difficult to quantify this claim because the level of the reduction in the logging volume is dependent on implementation and configuration choices and will therefore vary significantly between deployments.

Deterministic CGN has been implemented in some current CGNs. One down-side to deterministic CGN is that it reduces the efficiency of CGN and thereby reduces the number of subscribers who can be aggregated behind CGN.

4.9.3 Standard Logging Formats for CGN Events

Another issue that has been addressed through standardisation work is the specification of a standard log format.[51] This would greatly simplify the communication of log information between the parties involved.

4.10 Cost Implications

There have been several attempts to quantify the cost implications of CGN. However, besides CAPEX and OPEX, there are other categories of costs association with CGN implementation:

- **Support Costs**
  One mobile operator told us that there was no way that they could directly measure these costs, but that they knew that CGN does increase support calls. Another software vendor indicated that CGN has a significant impact on them including increased support costs, increased logging costs and infrastructure costs as services need to use a central server rather than remain peer-to-peer.
- **Logging Costs**
  These can be reduced by using deterministic NAT, but the actual costs are difficult to quantify as implementation dependent.
- **Security Costs**
  No matter what location in the network the provider chooses to implement CGN, it becomes another device that needs to be secured and protected against DDOS and other attacks.
- **CAPEX/OPEX**
  This is CGN device costs + maintenance and management. If an ISP needs CGN to provide IPv4 service continuity then this cost is simply unavoidable. The need for

[49] This statement is based on conversations with mobile operators who wish to remain anonymous.

[50] C. Donley, C. Grundemann, V. Sarawat, K. Sundaresan & O. Vautrin, *Deterministic Address Mapping to Reduce Logging in Carrier Grade NAT Deployments*, draft-donley-behave-deterministic-cgn-05, http://tools.ietf.org/html/draft-donley-behave-deterministic-cgn-05, 2013, retrieved 15 April 2013.

[51] S. Sivakumar & R. Penno, *PFIX Information Elements for logging NAT Events*, draft-sivakumar-behave-nat-logging-06, http://tools.ietf.org/html/draft-sivakumar-behave-nat-logging-06, 2013, accessed 15 April 2013.

CGN capacity can be reduced if native IPv6 is deployed to customers at the same time. This can off-load sessions to IPv6. For example, Google is IPv6 enabled so all Google traffic would use IPv6 if it was available and these sessions would not take up capacity in the CGNs.

- **Lost Customers**
  It seems likely that some users who have applications and services that are affected by CGN will contact the ISP for support. However, others may take matters a step further and switch providers. Because ISPs will experience IPv4 exhaustion at different times, it may happen that certain markets have a mix of CGN providers and ISPs that don't use CGNs. In a market such as this, it is certainly possible that the first ISP to start deploying CGN technology will lose customers to the ISPs that do not yet deploy CGN.

Lee Howard of Time-Warner Cable has published an extensive analysis of potential CAPEX/OPEX costs.[52] Cisco also provides a proprietary tool to potential customers to help calculate the cost of deploying CGN.

[52] L. Howard, *Internet Access Pricing in a Post-IPv4 Runout World*, http://www.asgard.org/images/pricing_v1.3.docx, 2012, retrieved 13 April 2013.

5 Implications of CGN for Internet Consumers

5.1 Why Do Things Break Behind CGNs?

Address sharing, by itself, is not bad. However, the address sharing implemented in CGN impacts Internet consumers in four critical ways:

1. The number of connections to the Internet available to each user is limited
2. The ability to connect, and maintain the connection, with end user devices and services becomes much more difficult
3. The ability to identify a device in an end-user network using only the source IP address is completely lost
4. Assumptions about the topology of the network (for instance, assumptions about the relationship between addresses and security) are no longer valid

One of the features of CGN—that it is supposed to be transparent to an Internet consumer—is actually the source of many of the challenges of deploying CGNs. Indeed, in cases where the user is simply reading email or browsing the web, the transparency of CGN is a success. However, as Internet consumers use more sophisticated applications and deploy services within their own networks, the transparency that is a virtue of CGN becomes the source of many of its problems.

Our research for this paper shows that some ISPs are being reasonably transparent about their trials and plans for CGN deployment. This is helpful because it is useful for Internet consumers to be aware of whether or not they are behind a CGN. Trouble-shooting problems related to services and applications becomes more difficult if the consumer is unaware of how the network topology is affecting the Internet services they are trying to use.

5.2 What Applications Usually Work Behind CGN?

Some applications usually work behind CGN; however, it is important to understand that *any* application can fail to work behind CGN due to the number of variables at play. For example, if a subscriber's maximum allocation of sessions in the CGN device has been reached, then applications will be unable to obtain new sessions and will fail.

It is important to define what is meant in this context by "work". Some measurements of applications "working" behind CGN have not considered problems such as the following:

- **Intermittent failures** – appears to work some of the time
- **Performance issues** – application "works" but in real world performance may be unacceptable
- **Features fail** – often applications will hide the failure of features from end-users so this is often not at all obvious
- **CGN relationship fail** – Failures only occur when a number of clients are behind the same or different CGNs

These four cases are difficult to measure and detect. It is also almost impossible for the end-users or the application provider to detect that the problems are a result of CGN in the network. Furthermore, many applications behind CGN will experience reduced performance and intermittent reliability. These applications may appear to work some of the time, making it more difficult to diagnose the source of the problem when the applications stop working.

ISPs will want to avoid service problems but as they cannot control how their subscriber uses the Internet and how the Content Providers provide their service, they cannot guarantee a reliable service behind CGN.[53]

Even so, there is an expectation that basic Internet services will work behind CGN. Basic Internet services are the simplest forms of Internet service. They include basic web-browsing and email. These applications have the characteristics that they:

- Use few sessions
- Sessions are initiated at the subscriber end
- Do not require peer-to-peer functionality
- Use TCP or UDP at the transport layer

Usually applications that meet these criteria are likely to work behind CGN. As mentioned above, this is not guaranteed as session exhaustion can cause any application to fail. Some examples of the number of concurrent sessions required by widely used Internet sites are shown in Table 5.1.

Web Page	Number of Concurrent Sessions
No operation	5 to 10
Yahoo Home Page	10 to 20
Google Image Search	30 to 60
Google Maps	20 to 50
Nico Nico Douga	50 to 80
OCN Photo Friend	170 to 200
iTunes	230 to 270
iGoogle	80 to 100
Rakuten	50 to 60
Amazon	90
HMV	100
YouTube	90
BitTorrent	700[54] (Typically hundreds of sessions)

Table 5.1 - Examples of Number of Concurrent Sessions for Web Applications[55]

These figures are typical of modern web-sites that use advanced techniques such as Asynchronous JavaScript and Extensible Mark-up Language (AJAX) to provide innovative consumer experiences. Any of these applications may fail behind CGN if a subscriber's

[53] In fact, in the United States, Verizon allows its customers to directly opt-out of being behind a CGN while acknowledging problems that CGNs create for certain applications. See: Verizon, *What is CGN - and How to opt-out*, http://www22.verizon.com/support/residential/internet/highspeed/networking/troubleshooting/portforwarding/123897.htm, n.d., accessed 15 April 2013.

[54] *Application behaviors in interms [sic] of port/session consumptions on NAT*, http://opensourceaplusp.weebly.com/implementations--experiments-results.html, n.d., accessed 15 April 2013.

[55] S. Miyakawa, *From IPv4 only to v4/v6 Dual Stack,* IETF72 IAB Technical Plenary, Dublin, 27 July - 1 August 2008.

maximum number of sessions are exhausted or become exhausted due to the application's requirements.

Take, for example, the popular application Google Maps. Google Maps typically requires thirty to fifty ports to function correctly. Less than this and it will not be able to fully display a map. This is illustrated in the screenshots[56] below for a range of connections.

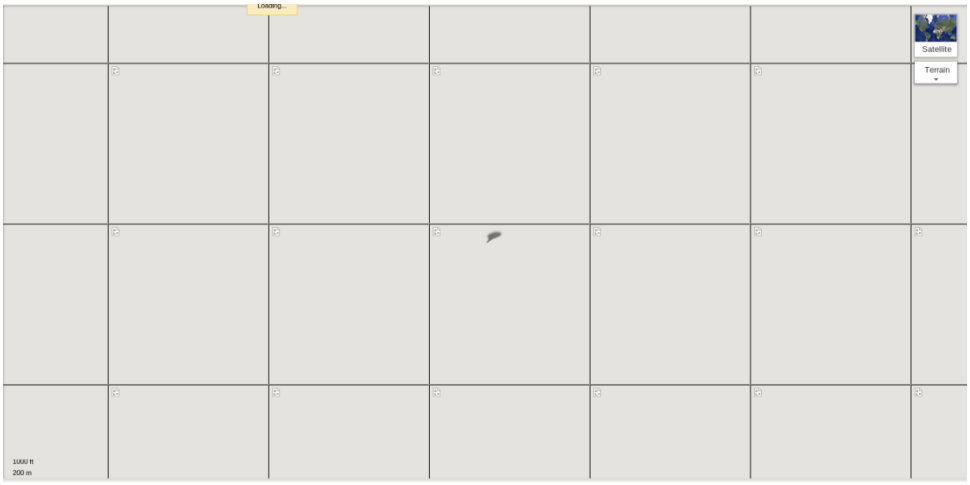

Figure 5.2 - Google Maps Limited to 5 Connections

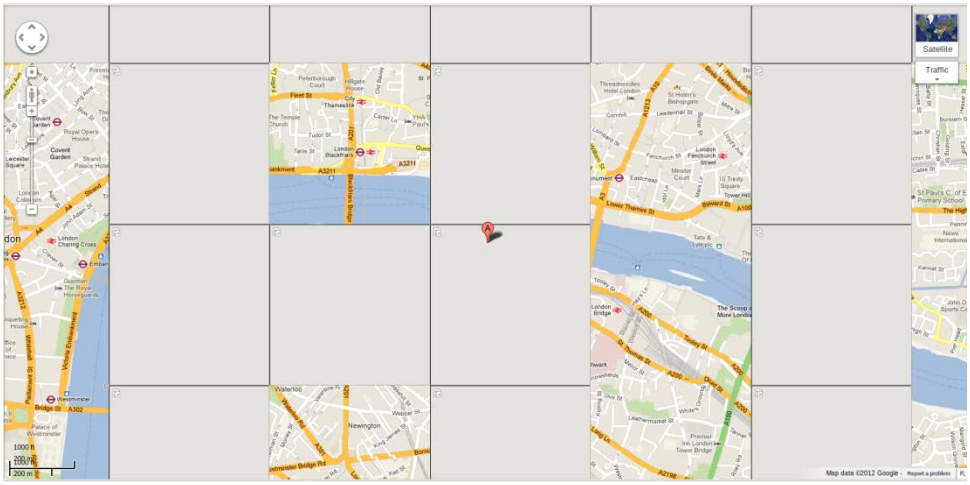

Figure 5.3 - Google Maps Limited to 10 Connections

[56] Screenshots provided by Erion Ltd.

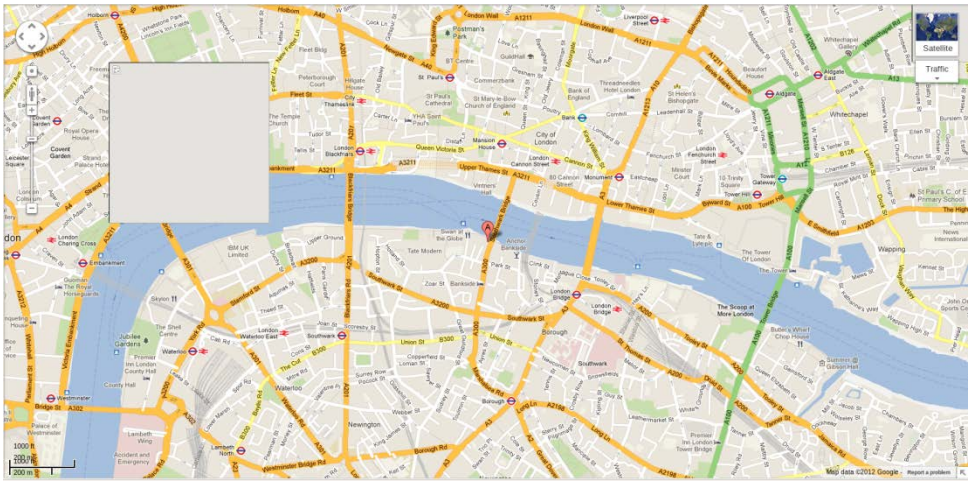

Figure 5.4 - Google Maps Limited to 20 Connections

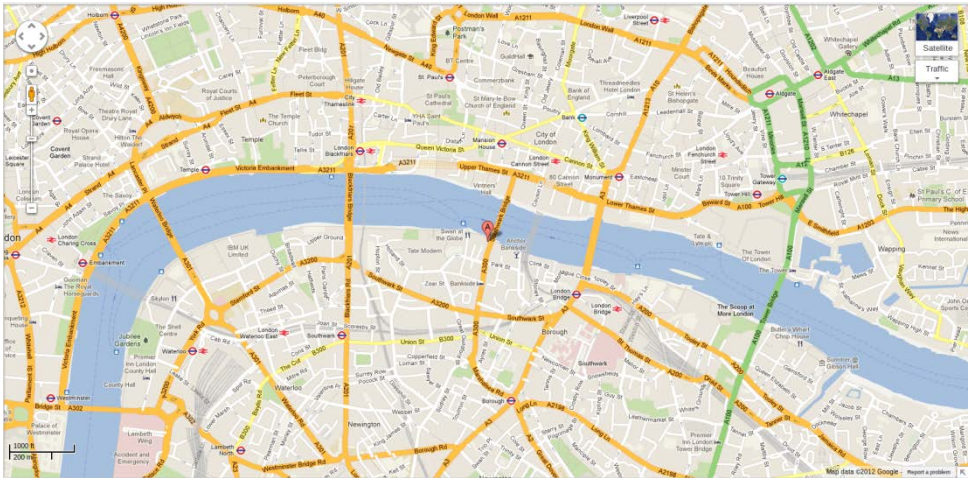

Figure 5.5 - Google Maps Limited to 30 Connections

5.2.1 Web Browsing

Web browsing is one of the basic Internet services that is expected to work behind CGN. However, *any* web browsing will fail if the maximum number of subscriber sessions is reached in the CGN device.

A common misconception is that only one session is required in the case of accessing a simple web page. The reason for this misconception is that historically a web-browser would fetch all the elements of a web page in sequence, one after the other, using just one session (assuming all the elements such as images were stored on the same server). Today, modern web browsers, usually fetch elements in parallel in order to improve performance. This is the case with widely used web-browsers such as Firefox and Internet Explorer (IE). Firefox will fail to display some web pages if enough sessions cannot be opened in parallel. The use of multiple sessions is shown in Figure 5.6 which was used earlier in this report to demonstrate session lifetimes and is repeated here for convenience.

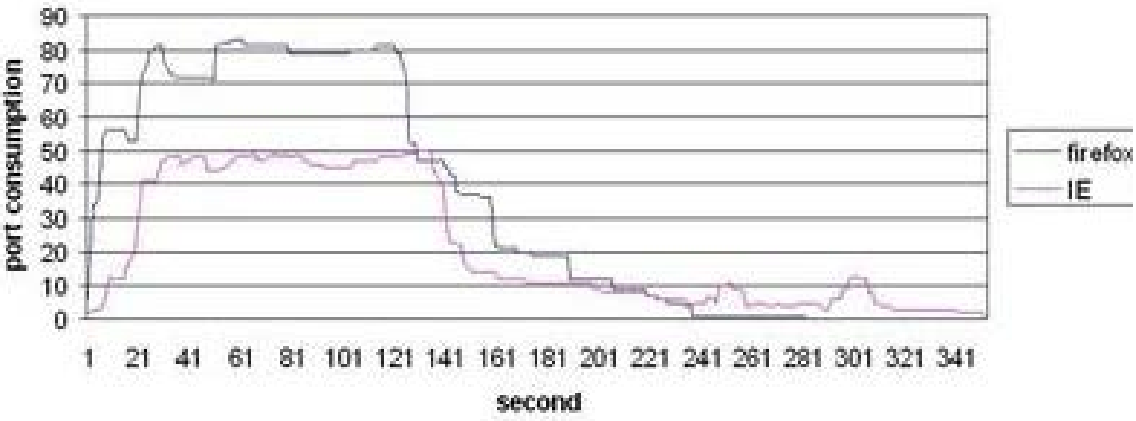

Figure 5.6 - Port Consumption Comparison Firefox and IE to http://www.youku.com[57]

Furthermore, many modern web-sites use techniques that themselves use multiple sessions. These techniques are used to create function-rich and innovative web-applications.

Examples include:

- Mapping web-sites such as Google Maps and Bing Maps
- Webmail such as Outlook, Gmail, and thousands of offerings from service providers
- Application service providers

These sites are built on technologies that use HTTP features that allow a single web page to open up multiple sessions to a web-server to improve performance and provide rich user experience. The leading technology in this area is AJAX[58] (Asynchronous JavaScript and XML).

If insufficient sessions are available to a subscriber, innovative and feature-rich applications may fail to function correctly or even at all. The Google Maps example earlier in this section demonstrates this.

5.2.2 Electronic Mail

Electronic mail, as opposed to webmail, requires a client email application. Common email clients include Microsoft Outlook and Mozilla Thunderbird. Email clients use application layer protocols such as SMTP (Simple Mail Transfer Protocol), POP (Post Office Protocol) and IMAP (Internet Message Access Protocol) to send and retrieve emails from mail servers.

Unlike web-browsers, where it is becoming increasingly common for many sessions to be consumed by a single web page, email clients typically only require a small number of sessions per email account. As such, email clients will only fail behind CGN when available sessions are exhausted.

[57] *Application behaviors in interms [sic] of port/session consumptions on NAT*, http://opensourceaplusp.weebly.com/implementations--experiments-results.html, n.d., accessed 15 April 2013.

[58] J. J. Garrett, *Ajax: A New Approach to Web Applications*, http://www.adaptivepath.com/ideas/ajax-new-approach-web-applications, 2005, accessed 15 April 2013.

5.2.3 Skype

Skype is designed to function behind NAT44. It uses a proprietary form of STUN to allow incoming connections into the subscriber's network even though they are located behind NAT. The success of Skype is a direct consequence of the widespread deployment of NAT44. Without NAT, Skype would not have been necessary as other Open VoIP protocols that existed before Skype would have worked just as well.

CGN presents a challenge for Skype. Skype has to use a range of different methods to traverse NAT depending on the location of the Skype clients:[59]

- When one client is behind NAT, Skype uses connection reversal where the node behind NAT initiates the connection
- With both clients behind NAT, Skype uses its own proprietary form of NAT traversal which is similar to STUN[60]
- If this fails, Skype relays the media session via a Skype super node. Relaying the media has performance implications and the service is likely to degrade or even fail

For the STUN-like approach to work, the Skype sessions from a particular client must all use the same IPv4 address. This often breaks the STUN approach in mobile networks where sessions for the same subscriber are not guaranteed to use the same IPv4 address. However, in a fixed line CGN environment Skype may be able to use its STUN-like approach.

When Skype falls back to relaying the media session via a Skype super node, there are significant implications for performance and reliability. The super node may be located some distance from one or both of the clients and so latency could be orders of magnitude worse than for a peer-to-peer session. Worse, bandwidth will be more limited due to the indirect and possibly long path the traffic must take. For these reasons, fall back to sessions via a super node will reduce performance and may cause intermittent failures to session.

To add to the user's difficulties, it will not be apparent that the cause is CGN. Users will assume that it is Skype at fault. Even if they suspect the carrier network is at fault there is no mechanism for fault finding that will enable them to determine that the problem is caused by CGN. The intermittent nature of failures make the fault finding process even more difficult and convoluted. Most users will simply give up attempt to use Skype.

At the present time, there are many reports of intermittent failures[61] with Skype on mobile networks. These can be found on the support forums of the mobile operators and on Skype's own support forum.[62] Indeed, such is the magnitude of these problems that Skype created a mailing list thread asking their users to record information about their problems so Skype could

[59] S Guha, N. Daswani & R. Jain, *An Experimental Study of the Skype Peer-to-Peer VoIP System*, http://research.microsoft.com/en-us/um/people/saikat/pub/iptps06-skype, n.d., accessed 15 April 2013.

[60] J. Rosenberg, R. Mahy, P, Matthews & D. Wing, *Session Traversal Utilities for NAT (STUN)*, RFC 5389, https://tools.ietf.org/html/rfc5389, 2008, accessed 15 April 2013.

[61] Intermittent failures resulting from problems with state management or address/port mapping are difficult to diagnose for many applications, not just Skype. While, the CableLabs study on NAT444 impacts (see Footnote 32) simplifies the discussion by saying that the protocol "works," the situation is often much more subtle. Intermittent failures in Skype and other protocols have yet to be publicly tested outside the lab environment.

[62] For instance, in Skype's Android Issue Report Forum, linked at: http://community.skype.com/t5/Android/How-to-create-a-good-Android-issue-report/m-p/211384/highlight/true#M4943

attempt to determine the cause. Because of the lack of visibility of CGN and the extreme difficulty of attributing such problems to CGN, it is impossible to say that these reports are caused by CGN.

It has been claimed that Skype works behind CGN. In particular, the investigation by CableLabs demonstrated that Skype works. However, this investigation in the lab is unlikely to observe the intermittent problems described above. As a consequence, it would be better if CGN issues with Skype could be measured in production networks such as mobile networks. As far as we are aware no one has done this and we understand that the information necessary to determine if CGN is the root cause of such problems is not recorded by the operators.

The model of NAT traversal in Skype is similar to that used in many other arenas including game consoles and peer-to-peer applications. Therefore these will also experience the same problems as Skype.

5.2.4 Instant Messaging

We have found that the chat, presence and control sessions of some Instant Messaging clients need to come from the same public source address. If the sessions don't come from the same address, the server will reject them. Behind a CGN this is not guaranteed. For example, when the AOL Instant Messenger (AIM) client is first started, it authenticates with an AOL server to establish the session. Then, when a user starts a chat window, a new session is started. If the chat session originates from a source address that is different from the authentication address, the AIM server immediately rejects the chat session because it is not perceived to be a properly authenticated session.

5.2.5 Facebook

Facebook is now IPv6 enabled. In CGN environments which also have IPv6 deployed, Facebook would not be impeded by the CGN device but would instead use IPv6.

Facebook's MobileVOIP service depends upon third party VoIP brands for administrative support and upon a small, Facebook-designed application that runs on Android, iOS, and several other mobile operating systems. The MobileVOIP app is not yet IPv6 enabled. For a call made from a wireless LAN connection behind a CGN, the impact is unknown.

5.2.6 Twitter

Twitter is not yet available via IPv6. Web based access to Twitter is subject to the limitations of sessions behind the CGN, however Twitter turns out to not have an especially high rate of concurrent session use. On the web, Twitter's profile is more like Google Maps than Facebook.

However, Twitter apps on iOS and Android are far greater consumers of sessions because these apps appear to provide live updates of search results and groups of followers in addition to the base stream of short messages. Behind a CGN, Twitter users are only likely to see occasional degradation of performance. This degradation would be due to latency in message arrival and, in much less likely cases, service disruption due to exhaustion of session limits.

5.2.7 Banking Applications

During research for this paper several examples of banking applications were found (for smartphones and tablets, as well as traditional browsers) that require all connections from a given host (SSL or not) to come from the same IP address. If the CGN is unable to be

configured to ensure that the same external address is used for all sessions originating from the same internal host, then these banking applications all fail.

5.2.8 Spotify

The internal architecture of Spotify relies on peers being able to seed music content once authenticated connections have been made with a central server. Rather than streaming content from a central source, content is streamed from peers in the network. Finding the peers is a proprietary task, but if the peers appear to change IP addresses and ports, the stream fails. It is also worth noting that Spotify requires communication with a specific IPv4 address to establish sessions.

5.2.9 Apple FaceTime and Google Talk

Apple's conferencing tool, FaceTime and Google Talk have been noted to require that not only must all connections from a given host come from the same IP address, but also that the same port must be used. This is a far more strict requirement than even the banking applications seen in Section 5.2.7 above. This means that these applications will fail if the CGN does not configure the sessions to ensure the same IP address and port is used for the connections at all times.

5.2.10 BitTorrent

Some significant testing of the behaviour of BitTorrent in CGN environments was done in 2012.[63] The result is that most features of BitTorrent will work in the presence of port forwarding (for instance, PCP), but these features fail when machines have the same IP address and port forwarding is not enabled. This turns out to be true in cases where two machines, both behind a CGN, were attempting to share files.

Even if port forwarding is turned on, there are still significant limitations for BitTorrent. In one case the BitTorrent configuration on the serving (seeding) peer does not permit sending the same file to more than one port on the same IP address. The second limitation is when a client attempts to gather a file located on several seeders that share a common IP address. In these cases the client can only connect to one seeding server at a time, dramatically limiting the performance of the peer-to-peer architecture.

The crucial feature of port forwarding on CGNs (for instance, see Section 2.3.7) is that during interviews conducted for this report, no ISP expressed either a commitment or an interest in allowing consumers to configure their own port forwarding on a service providers' CGN. In that context, BitTorrent is an example of an application that will be significantly impacted by implementations of CGN.

5.2.11 Understanding CGN By-Pass Part of the Strategy

In our interviews with service providers outside of the UK, we found that part of their strategy when deploying CGNs is to also deploy IPv6 at the same time. By doing this they have been able to:

- Reduce the amount of traffic that has to traverse the CGN devices thereby reducing the load on the CGN devices

[63] M. Boucadair, T. Zheng, R. Ng Tung, X. Deng & J. Queiroz, *Internet Draft: Behavior of BitTorrent service in PCP-enabled networks with Address Sharing*, draft-boucadair-pcp-bittorrent-00, http://tools.ietf.org/html/draft-boucadair-pcp-bittorrent-00, 2012, accessed 15 April 2013.

- Improve the service provided to their customers. The IPv6 traffic is unimpeded by NAT44 or CGNs

As noted in Section 2.2 earlier, some service providers have found that 33% of their dual stack customer's traffic is now IPv6.

5.2.12 Services Within the ISP's Access Network

External providers of services may experience problems if their customers are located behind CGNs. In contrast to this, ISPs will be able to provide the same services in the access network behind the CGN device without any of the problems that external service providers face. This means that ISPs will have an advantage over external service providers.

An example of this is the provision of VoIP services. Customers behind CGN may find that VoIP services that are not provided by their ISP either do not work, fail intermittently or perform badly. Therefore they are likely to be forced to use the services provided by their ISP if they wish to have an acceptable service.

5.3 What Applications Break?

Implementation of CGN in a network fundamentally changes the topology of a network. That change has implications for applications that expect the network to support traditional, end-to-end connections between nodes. Address sharing via a CGN will have implications for a broad range of applications.

- Applications that need to connect into an Internet consumer's network will need to make sure that the needed ports are forwarded properly by the CGN. This is not necessarily how CGNs are configured.
- Applications that, because of their design, carry IP addressing or port information in the payload portion of the IP packet. Because translation of IP addresses and ports takes place on the IP header, additional software is required to scan the internal content of these protocols and make the translations in the payload. When the payload is encrypted it can be nearly impossible to make this translation.
- Applications that require fixed ports are a problem because this allows for no differentiation between nodes behind the CGN.
- Some applications do not even use a port number as part of the protocol. These are a problem because they expect to use only the IP address to reach the endpoints of the network connection.
- Applications that explicitly prevent multiple connections from the same source address. These applications often use the IP address as part of the support for providing security and prohibit connections from the same IP address.
- Applications that are not based on TCP or UDP for their transport protocols. Almost every CGN strategy in use supports only TCP, UDP and ICMP as the underlying transport protocols. As a result, applications that use more recent or specialised transports will fail in the presence of a CGN.

Traditional NATs, located at Internet consumer premises, have for some time had Application Layer Gateways that allow the NAT to perform the packet inspection needed to behave correctly in the presence of protocols with IP addresses and ports inside the packet payload. This works well in today's traditional network environment because the Internet consumer has control of the device and has the flexibility to adapt it to his or her needs. For CGNs, the situation changes. When an Internet consumer is behind a CGN they will be dependent on the ALGs and services provided by the ISP.

While it is tempting to start to list applications that either work properly or fail under CGN, it may be better to consider categories of applications and services. Recently there has been a steadily increasing amount of research into what applications and services work and which fail partially or completely.

This report has referenced the excellent Cablelabs report[64] in other places in this document, but we have also noted that the lab work conducted seems to have limitations in both breadth and in the network topology being tested. Our research has also noted the work of Chris Grundemann[65] in summarising some of the typical applications where CGN causes problems. In addition, research prior to this report evaluated the soon to be completed trial of CGN at Plusnet.[66] The trial should be welcomed for its transparency and commitment to accurate findings.

5.4 What Services Break?

5.4.1 Web Browsing

CGNs are built to support traditional and heavily used applications such as web browsing and electronic mail. In fact, in both field trials and lab studies, simple connections between a web browser and a web server are unaffected by the insertion of a CGN into the network.

However, contemporary web browsing is very different to the style of browsing that was common ten years ago. To render a web page, a browser almost never interacts with a single server. Instead, a browser establishes connections to many servers to request the resources needed to render a page. In addition, many web pages keep sessions between the client and server alive so that material on the page can constantly be refreshed. Finally, web pages localise content through geolocation.

It is the new style of browsing that is most at risk from the injection of CGNs into the path between server and browser. Exhausting the available connections at the CGN will degrade or cause the page not to load. Being unable to identify an Internet consumer's location will cause local customisation to fail.

It is certainly true that most web browsing behind a CGN will continue to work, there are occasions where the contemporary model for interactions between browsers and servers will break down in the presence of CGNs. Section 4.4 provides a classic example of this problem.

5.4.2 AJAX Applications

AJAX applications on the Web are the foundation of many modern websites. In section 4.2, Google Maps was shown to open 70 parallel connections at a time. iTunes store has been shown to open as many as 300 parallel connections. These are existing applications that are in widespread use. New applications, which depend on many connections at a time, may be just over the horizon.

---

[64] C. Donley, L. Howard, V. Kuarsingh, J. Berg & J. Doshi, *Internet-Draft: Assessing the Impact of Carrier-Grade NAT on Network Applications*, draft-donley-nat444-impacts-05, http://tools.ietf.org/html/draft-donley-nat444-impacts-05, 2012, accessed 15 April 2013.

[65] C. Grundemann, *NAT444 (CGN/LSN) and What it Breaks*, http://chrisgrundemann.com/index.php/2011/nat444-cgn-lsn-breaks, 2011, accessed 15 April 2013.

[66] *plusnet CG NAT Community Forum*, http://community.plus.net/forum/index.php/board,72.0.html, 2013, accessed 15 April 2013.

CGN's fundamental purpose is to share a single address between many users using that address's 64,000 port numbers. If we take iTunes as an example, 64,000 ports shared amongst users using iTunes results in about 200 customers being able to share a public IPv4 address. Since restricting the number of available ports impacts the utility and performance of the AJAX application, conservative approaches to the port range may apply. If they do, the number of consumers who can share a single IPv4 address begins to drop.

In addition, AJAX services usually do not allow connections from the same client to span multiple CGN public side addresses. When a port pool becomes exhausted on the CGN, the resulting connections on other, shared IPv4 addresses will be refused and the application will fail.

5.4.3 Media Libraries in the Home, PVRs and Other Home Resources

An emerging class of applications makes access to home networks from the public Internet essential. In this class of applications, a device, appliance or service that resides in an Internet consumer's network is controlled from an authenticated device or web page on the public Internet. This growing class of applications requires that an external client be able to initiate a connection to a device in a (typically) home network.

Some examples of this class of application include:

- **Webcams or monitoring devices** where a person using a web page or an application on a smartphone can access in realtime the images provided by cameras in a residential network. Often marketed by security or child safety companies, these devices usually provide authenticated access to an application that controls the camera and displays the images.
- **Personal Video Recorders** (PVR) where an application can control and sometime play back recordings of subscriber-based video. Usually marketed by the company providing the subscription service, the applications run on smartphones, tablets and even web pages and connect to the residential video recorder to set up, manage and sometimes play recordings.
- **Home security systems** where a person is able to control security alarms and other devices managed by a home security system. These systems are usually based on a central server which mediates the connection between the owners' smartphone or other device and the security system.
- **Home media libraries**. Many software tools now allow a person to connect to and then stream their personal media libraries to devices that are remote from the home network.
- **Appliance control** where a person can control "non-computer" devices connected in the home network from devices that connect to them remotely.
- **Remote/time-shifting of television** where, from a remote device, a user can authenticate and then watch their television signal as if there were in the house.

Common to this class of applications is the need to support authenticated external access to devices and services in the consumer network.

This is a challenge for CGNs because the traditional bindings on addresses and port numbers is made on connections coming from the consumer's network, going through the local NAT and then out through the CGN to the Internet. Establishing bindings from connections that are begun as external connections is impossible in a traditional way.

We have already seen that one approach to solving this problem is to use the emerging standard PCP (see Section 3.3.7). However, during interviews conducted for this report, no ISP expressed either a commitment or an interest in allowing consumers to configure their own port forwarding on a service providers' CGN.

The result is that consumers who fit certain profiles—such as those who access consumer networks remotely—might not fit as good candidates for CGN service. That leads to consideration of differentiated services for ISPs, which is a market question considered in Section 10 of this report.

5.4.4 IPsec

The problems with IPsec and CGN have been described in Section 4.2.1 of this report. When a subscriber is located behind NAT44, IPsec may be function successfully either with manual configuration of port and protocol forwarding or using IPsec NAT traversal techniques. Even so it will be limited. For example, one VPN terminating in the subscriber's network may work but a second session may fail or cause the first session to fail.

The operation of IPsec behind the CGN device is even more difficult. For scenarios using the IPsec Authentication Header (AH) that cannot tolerate modifications to the IP header, sessions will fail. AH headers are commonly used in most VPN configurations. This means that VPN sessions terminating on the subscriber's NAT device or in the subscriber's network will fail. Scenarios using the ESP header result in the transport layer ports being encrypted. This means that the CGN device cannot multiplex sessions based on port mappings. The only information available to the CGN device is the protocol number, which for ESP is protocol 50. This is the same for *all* ESP secured sessions for *all* subscribers sharing the same CGN IPv4 address. This means that the CGN cannot differentiate between sessions and would therefore be unable to handle more than one session between all the users sharing the same IPv4 address on the CGN.

In addition to these problems, the CGN device would need to be able to handle NAT traversal for IKE. This is not currently possible: NAT traversal works only if the NAT devices along the route between the two IPSec gateways maintain the same address mapping since the last IKE negotiation. This prevents IPSec from first sharing enough information to authenticate each end of the connection and then, second, from establishing the secure channel between the two endpoints.

5.4.5 SIP Clients

SIP Clients behind a CGN require multiple connnections to come from a common address. As an example, if a SIP client is sending Real-Time Transport Protocol (RTP) and Real-Time Control Protocol packets, each end expects that they come from the same IP address, even after they have emerged from the network provider's side of the CGN. The only way to avoid this is to have the two endpoints negotiate an alternative scheme beforehand.

If RTP and RTCP addresses are different, the receiving endpoint simply drops the packets that do not match the original packet's IP address.

5.4.6 6to4, Teredo and Tunnel Brokers

Consumers who are accessing the global IPv6 Internet using transition techniques such as 6to4, Teredo and Tunnel Brokers will find that these are unlikely to work behind CGN. As noted in Section 3.4, all these techniques are impacted by CGNs. To recap:

- 6to4 requires a public IPv4 address, shared addresses or private addresses used in a CGB access network will not work
- Teredo has been shown in testing to not work behind CGNs
- Tunnel brokers typically require a public IPv4 address as an end-point

5.4.7 XBOX and PlayStation (P2P gaming)

A number of organisations interviewed for this report have tested different games consoles and games behind CGN. All of them reported problems with specific games behind CGNs. One game console manufacturer reported that in their testing of twenty games more than half were broken by CGN.

Internet consumers who have purchased games for their console may find that some games suddenly cease to work if their ISP deploys CGN.

5.4.8 Geo-location

Geo-location is the process of approximating the location of a subscriber from their IP address. Geo-location is used by many applications to provide users with information relevant to their location. Location may also be useful for emergency services and for law-enforcement agencies.

Geo-location on the Internet is not the same as location services on mobile devices where built in GPS receivers are used to provide an accurate location for applications.

An example of content providers who use geo-location is Yahoo. They have a number of products that use geo-location. These are illustrated by examples in Figure 5.7.

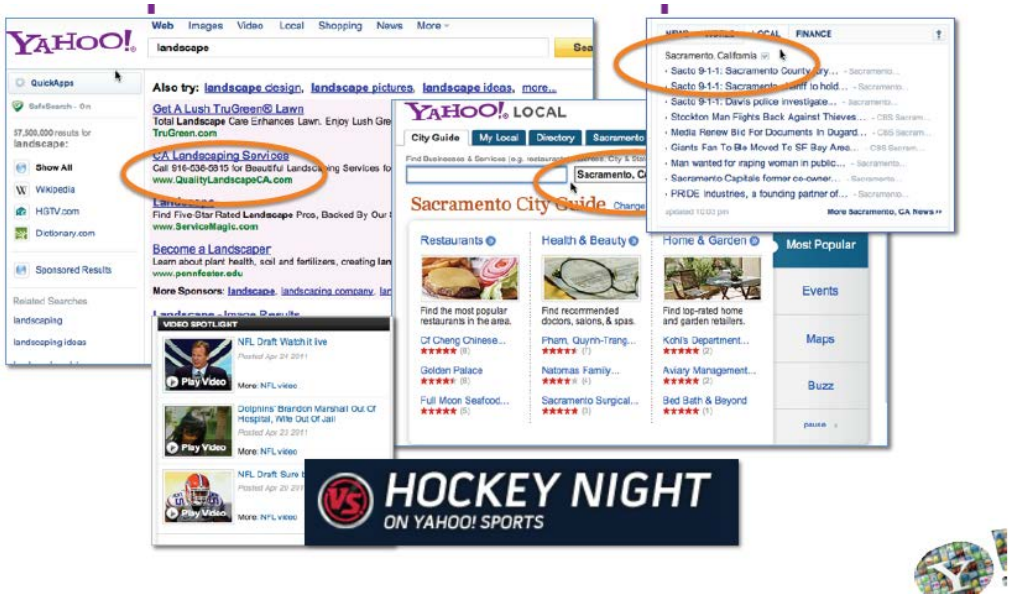

Figure 5.7 - Examples of Geo-Location in Yahoo!'s Products

When using geo-Location without CGN, it is usually possible to determine the city or town that a subscriber is in. Sometimes better resolution is possible. When CGN is introduced, the resolution that is lost depends on how large an area of subscribers is placed behind the CGN device.

This is illustrated in Figure 5.8 and Figure 5.9 which show best case and worst case examples of the impact of CGNs on geo-location.[67]

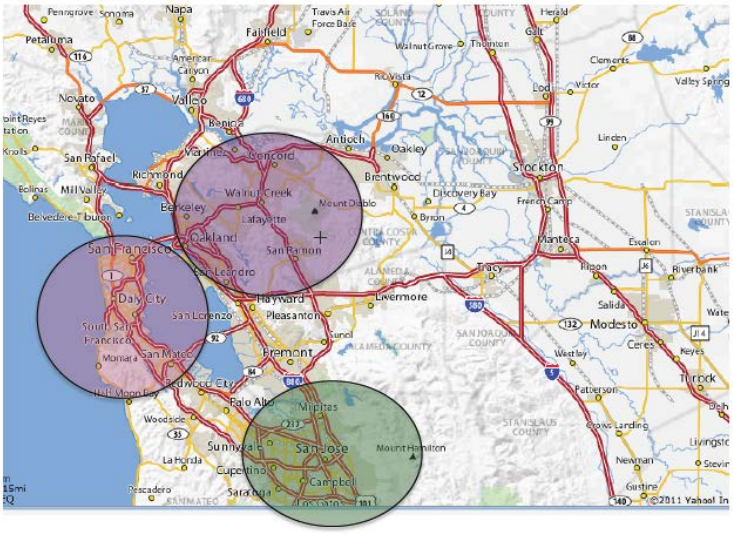

Figure 5.8 - Impact of CGN on Geo-Location Best Case (Yahoo!)

[67] J. Fesler, *World IPv6 Day Recap*, http://meetings.apnic.net/__data/assets/file/0003/38298/fesler_yahoo_2011_post_ipv6_day_20min.pptx.pdf, 2011, accessed 15 April 2013.

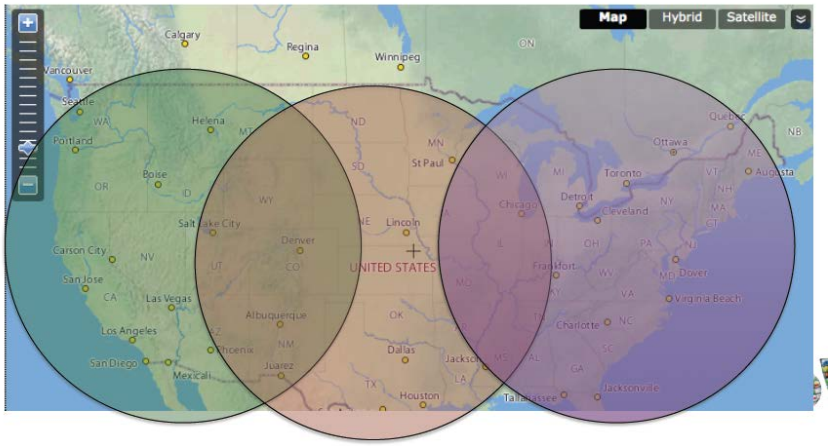

Figure 5.9 - Impact of CGN on Geo-Location Worst Case (Yahoo!)

Geo-location can also used to determine the location of VoIP callers to the emergency number. This is a legal requirement in some countries.[68,69]

5.4.9 Geo-proximity

Geo-proximity is used to determine how near two nodes are. Usually nearness is measured in terms of latency, which is the time taken for traffic to travel between the two nodes.
Many applications and services are sensitive to latency, for example:

- VoIP
- Video Conferencing
- Peer to Peer Gaming
- Peer to Peer Applications

VoIP is well known to be sensitive to latency issues. Research has established that round-trip latency less than 150 ms is not immediately noticeable, be that latency greater than 300 ms is unacceptable[70] (Cisco states that latency higher than just 150 ms is unacceptable[71]). CGNs increase latency in two ways; through the latency in the CGN device and by moving the egress point in the network to the location of the CGN device which may be distant from the subscriber (as was shown in the Geo-Location section above).

In games consoles and peer-to-peer games, latency is very important. In multi-player live action games, when a user moves their character they expect the other players to see the change immediately. If there is a delay the user may become annoyed or disoriented. Excessive delays may make the game unplayable.

For these reasons, games developers and games console manufacturers go to a lot of effort to reduce latency. One techniques that they use is geo-proximity. Typically they use techniques

[68] European Commission, *EU Rules on 112*, http://ec.europa.eu/digital-agenda/en/eu-rules-112, n.d., accessed 15 April 2013.

[69] US Department of Transport, *Next Generation 9-1-1*, http://www.its.dot.gov/NG911, 2012, accessed 15 April 2013.

[70] T. Wallingford, *Switching to VoIP*, O'Reilly Media, Sebastapol (California), 2005, p. 193.

[71] Cisco, *Enabling VoIP: Data Considerations and Evolution of Transmission Network Design*, http://www.cisco.com/en/US/prod/collateral/optical/ps5724/ps2006/prod_white_paper0900aecd803faf8f_ps2001_Products_White_Paper.html, 2006, accessed 15 April 2013.

that measure the nearness of players in terms of latency. Players with low latency between each other can then be paired together. Next, games developers use peer-to-peer sessions between the players to keep the inter player traffic latency to a minimum.

When CGN is deployed, there are two problems for this geo-proximity scenario:

1. **Latency is increased**

   Users will be topologically further apart as they will have to traverse a CGN device that may be geographically distant from them. Since CGN will aggregate large numbers of users the geographical area that a single CGN device covers can be significant. So the latency is increased through a combination of CGN device latency and the increased topological distances.

2. **Traffic may have to pass through an intermediate server**

   CGN may also break peer-to-peer connectivity between users causing connections to fall back to going via an intermediate server. This will have a significant impact on latency and bandwidth. For many gaming applications the impact of going via a server would introduce unacceptable performance limitations. Also, an additional impact of sessions having to go via a server is an increase in capacity required in those servers. This increases capital and on-going costs for the service provider and has an indirect environmental cost. Another potential problem for the service provider is at the point that the carrier enables CGN there could be an unplanned for and substantial increase in load on their servers.

5.5 Could ISPs Use CGN to Create Walled Gardens?

A “walled garden” is a service that provides easy access to services and resources of the ISP, but limited, controlled or no access to other content. This is in contrast to traditional ISPs who provide open access to all content on the public Internet. Since a CGN acts as a “chokepoint” between the Internet consumer and the public Internet, in theory it would be possible for the ISP to control what services were provided to the consumer.

However it goes beyond simply controlling access to resources. An ISP could use the CGN as a revenue option. For instance, the ISP could allocate a small number of ports to each individual consumer. If the consumer bumped up against the limit, they could be redirected to a web page that encouraged them to purchase more connection space on the CGN. Thus, without adding any content or value to the public Internet, the ISP could use the CGN to make transit a greater part of their revenue stream.[72]

[72] One vendor of CGN equipment advocates exactly this, saying:

> “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 the user and protect the infrastructure from abusive users.”

For more, see: J. Haworth, *Carrier-Grade Network Address Translation (CGNAT)*, http://www.f5.com/pdf/white-papers/cgnat-wp.pdf, 2011, accessed 15 April 2013.

5.6 Summary of the Technical Implications of CGN

Table 5.10 below provides a summary of the technical impact of CGN for consumers of Internet access.

	Routed Network	NAT44	CGN	IPv6
Client Web Connection	Works	Works	Works	Works
Running Servers (Web/Email/IM/Calendar/VPN etc)	Works	Requires additional configuration and is limited in scope. May require UPnP or port forwarding. Will not support mulitple servers on the same port. Will only work for TCP/UDP and a limited subset of other protocols.	Likely to be difficult or impossible. May be possible with cooperation from service provider and use of PCP/UPnP gateway. Indications are that service providers will be reluctant to do this.	Works
Provides Customer with Public IP Address	Yes IPv4 every node end to end (Optionally ISP may provide IPv6)	Yes IPv4 shared (Optionally ISP may provide IPv6)	No IPv4 (Optionally ISP may provide IPv6)	Yes
Web AJAX Content	Works	Works	Will be limited by number of sessions allocated per subscriber by ISP and the number of sessions already in use by subscriber (e.g. Google Maps)	Works
VPN	Works	Often works due to use of NAT-traversal techniques and support in many CPEs. Has limitations. For example, IPsec transport mode is often not possible.	Unlikely to be possible. In future limited support may be possible if there is widespread deployment of PCP and other techniques.	Works
SIP	Works	Requires SIP Proxy	May work if SIP ALG is provided.	Works
H.323	Works	Requires H.323 Proxy	Unlikely to work.	Works
FTP	Works	Usually works due to FTP Application Layer Gateway (ALG) support in CPE or passive mode.	Works due to passive mode	Works
Transport Layer Protocols other than UDP and TCP	Works	May require additional functionality or configuration depending on the specific protocol.	May be difficult or impossible.	Works
IPv6 Transition Mechanisms	Works	Teredo client behind NAT and 6to4/6rd on CPE will work.	Teredo may work (although testing has shown otherwise). Other techniques will not work without special configuration. DS-Lite will work as it includes support for LSN.	Not required

	Routed Network	NAT44	CGN	IPv6
Remote Access to Home Devices (for example webcams)	Works	Requires additional configuration and is limited in scope. May need UPnP or port forwarding. Will not support multiple servers on the same port. Will only work for TCP/UDP and a limited subset of other protocols.	Currently likely to be impossible. May be possible with cooperation from service provider. In long term future may be possible if there is widespread deployment of PCP and other techniques.	Works
Peer to Peer Applications	Works	Requires additional configuration and is limited in scope. May need UPnP or port forwarding. Will not support multiple servers on the same port. Will only work for TCP/UDP and a limited subset of other protocols. May also require STUN, TURN and ICE.	Difficult and will require additional configuration or techniques. Will vary from application to application. STUN may fail to work and fall-back to TURN may be necessary forcing the use of an intermediate relay server.	Works
Vulnerable to State Attacks	No state	Yes, affects one customer	Yes, may affect large numbers of customers at once	No state
DDoS Mitigation	Works. Mitigation is per end-point addresses	Affects whole of network behind NAT.	Affects multiple customers including those from which the attack is not originating	Mitigation is per /64, that is per customer or IPv6 subnet.
IP Blacklists	Affects individually blacklisted hosts only	Affects all nodes behind NAT if NAT public address is blacklisted	Affects multiple customers if one subscriber is blacklisted.	Works (there are challenges for IPv6 blacklists but that is beyond the scope of this report)
Logging to meet legal requirements	Direct trace back to source IP address. Minimal or no logging required.	Direct trace back to subscriber network not individual node behind subscriber NAT. Minimal logging required. For example if the subscribers address is assigned dynamically.	Significant logging required. Logging must also be accurately time stamped in order to trace external address and port combination back to subscriber network. Substantial volumes of logging data. The volume of logging data may be reduced using deterministic methods of port assignment.	Minimal or no logging required. Traceable from IP address only back to subnet and/or node.

	Routed Network	NAT44	CGN	IPv6
Logging Disk Space	None (IP address is fixed for each node on network)	Minimal logging. IP address changes for NAT boxes that do not have fixed IP addresses. Log daily changes (if address changes most CPEs are on 24x7 and renew their leases). Annual logging requirements for a million customers could be stored on one harddrive)	Significant logging. Can be reduced by deterministic CGN, but will still be greater than for none CGN scenarios. (Annual logging requirements for a million customers will require 100s or 1000s of Terabyte drives - will require management systems, datacenter space and power).	None
Future Applications and Services	Unhindered by network limitations	Limited by NAT	Significantly limited by CGN or may be made impossible.	Unlimited
Complexity	Routing/Forwarding	Routing/Forwarding Address Translation Port Mapping Application Layer Gateways for some Protocols Proxies for other protocols Port Forwarding uPnP (with associated security issues) Single point of failure	Routing/Forwarding Address Translation Port Mapping Application Layer Gateways for some Protocols Proxies for other protocols (difficult to implement so may not be possible, they will also need to be placed in the access network) Port Forwarding not an option without cooperation from the access network UPnP (with associated security issues) Will also require additional support for PCP etc in order to function. Single point of failure	Routing / Forwarding no need for complexity of NAT CGN, ALGs, Proxies etc...
Impact on Internet of Things	None	Some impact will require techniques to allow external access to devices.	Significant impact. May make external access impossible.	None
Impact on Future Web Applications	None	Minimal impact.	Significant impact. Will limit any applications that use multiple sessions which is common in most advanced web applications. (for example AJAX)	None
Impact on Internet Services and Application Developers	None	Some impact. Developers may need to work round different types of NAT. This has a cost and hinders development. It is unknown how many applications and services have never been deployed due to difficulties traversing NAT. However, it is known that this happens.	Significant impact. May require substantial development effort to make all but the most simple applications and services work reliably across CGN.	None

Table 5.10 - Summary of Implications of CGN for Internet Consumers

6 Implications of CGN for Internet Service Providers

Some ISPs will be forced to deploy CGNs in order to address their shortage of IPv4 addresses. This will allow them to continue to provide an IPv4 service to their customers.

Unfortunately the IPv4 service provided to an ISP's customers using CGNs is not equivalent to one provided without CGNs. This will present some issues for ISPs. However, many of the negative impacts of CGN affect not the ISP but the subscriber and/or the application and service providers. Furthermore, in many cases, the subscriber and/or application and service providers will not attribute the problems to CGNs, especially since they may not even know that ISP has deployed CGNs in their network. As a result much of the negative effects of CGNs on an ISP's customers and third parties will not have a negative impact on the ISP.

6.1 Managing Impact on Users

The impact of CGN on an ISPs customers is described in Section 5. The ISP can take action to mitigate some of the impacts upon users but not all.

This will be particularly important for ISPs when their customers are unable to use applications and services which their competitors' customers can use. Unless this is carefully managed, ISPs may find that they lose customers to competitors. Whilst application and service providers will attempt to overcome the limitations of CGN, this may only be possible if the ISP deploys CGNs in a specific way.

In particular ISPs may:

- Inform their customers that they are deploying CGNs
- Provide a mechanism for customers to configure port forwarding through CGNs
- Configure their CGNs to ease the operation of NAT traversal techniques such as STUN
- Provide contact details for application and service providers who wish to work with them to resolve suspected CGN problems
- Implement support procedures for detecting and addressing CGN problems
- Provide their customers with an IPv6 service that can act as a CGN by pass

ISPs may wish to set different charges for a service provided through CGN to reflect the degradation in service. This may be a way by which an ISP can retain customers who would otherwise consider moving to an ISP that has not deployed CGN.

It is in the interests of ISPs to only put behind CGN those users that they know only require "basic services" and to avoid putting "advanced users" behind CGN. This way they will minimise the potential support problems and customers' dissatisfaction. ISPs may find this difficult to do if they have no way of determining which users are "basic" and which are "advanced".

6.2 Security

There are a number of security issues relating to the deployment of CGNs in an ISP's access network.

6.2.1 CGNs and Distributed Denial of Service (DDoS) Attacks

A recent survey carried out by Arbor Networks showed that ISPs around the world are already experiencing Distributed Denial of Service (DDoS) attacks against CGN devices.

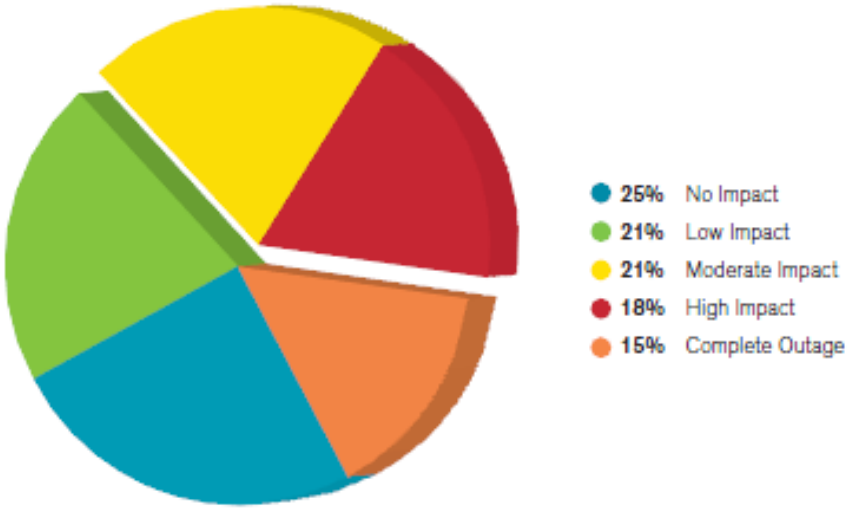

Figure 6.1 - Impact of Attacks Against NAT (CGN) Infrastructure[73]

The problem with CGN (and any NAT) devices is that they have to maintain a record of the state of mappings between internal ports and addresses and external ports and addresses. Any software or device that has to maintain state is vulnerable to DDoS that exhaust the resources used to hold and maintain the state.

In the survey, 31.1% of respondents knew that they had had an attack against their CGN during the previous 12 months and 14.9% did not know if they had.[74] Of the 31.1% who detected an attack, over 55% were significantly impacted by the attack.

CGN devices are vulnerable to DDoS attacks in a number of ways. The CGN device can be a target in both directions (incoming and outgoing). Also, the CGN device can be impacted due to DDoS attacks that traverse the CGN device. The main problem occurs with outgoing attacks. For example, a common attack such as a UDP flood can quickly fill up the state tables in a CGN device. One or more subscribers that are infected with DDoS malware could initiate such an attack.

Mobile operators have had experience using CGNs for many years. However, they do not have the same level of malware on mobile phones as is found on laptops and PCs. As a result, mobile operators have not yet experienced significant problems with the filling up of state tables in CGNs. This is beginning to change due to the increasing use of 3G and 4G dongles where the attached PC or laptop is much more likely to be infected with DDoS malware than is a mobile phone.

Mitigating these attacks can be split into two cases; mitigating inbound attacks and mitigating outbound attacks. Inbound attacks are mitigated in the same way as any inbound DDoS attack typically using mitigation systems such as those provided by Arbor Networks.

[73] Arbor Networks, *Worldwide Infrastructure Security Report*, http://www.arbornetworks.com/research/infrastructure-security-report, 2012, accessed 15 April 2013.

[74] Additional data provided by Arbor Networks.

The main problem are the outbound attacks which are much more likely to result in a serious impact on the CGN device. Outbound attacks are harder to mitigate. Currently, ISPs rarely monitor or mitigate outbound DDoS attacks originating within their own networks. With the deployment of CGN, it will become necessary for ISPs to take action to protect their CGN devices against such attacks. To do so, operators will have to either monitor at the customer edge or generate more detailed flow information in the CGN device itself.

At the present time, it is unclear what will happen. However Arbor's recent survey indicates that it is a problem that will need to be addressed.

6.2.2 IP Reputation

The reputation of an IP address is used by security products and servers to determine whether to trust traffic originating from the IP address. If an IP address has a sufficiently low reputation, traffic originating from it may be blocked.

There are many providers and users of IP reputation information.[75] Typically they measure the quantity and type of attacks originating from IP addresses and then assign them a rating. The rating is then used by security products, services and applications to filter traffic.

Mail servers often check IP reputation lists and use this information as a factor when deciding if email should be treated as Spam or not.

With CGN, a group of subscribers will share the same IP address. If one or more of this group of subscribers had been compromised with malware then their IP addresses reputation could easily be impacted. The result of a negative IP reputation would not just affect one subscriber it would affect all subscribers behind the CGN device that are sharing the same IP address. Some aspects of a negative IP reputation are unlikely to affect users behind CGN. For example, many subscribers will not send their emails directly from their email client to the destination mail server. Instead they will usually send their emails through a mail server. As a result the reputation of their shared CGN IP address will not be a factor[76] as the destination server will only see the IP address of the subscribers mail server not mail client.

6.3 Implications of CGN for Application and Service Providers

Application and Service providers will find that their applications may not work globally. In cases where CGN is deployed, they may find that the performance of their applications and services may deteriorate and even that they may fail completely.

Content providers have no control over the access network their customers use. Therefore they will be forced to attempt to work around CGN (which may be impossible) or to advise potential customers that their service will only work in non-CGN access networks. This has already happened in current NAT environments where some services may not work and this is included

[75] For example:

- WatchGuard http://www.reputationauthority.org/about.php
- Symantec http://ipremoval.sms.symantec.com/lookup
- McAfee http://www.mcafee.com/threat-intelligence/ip/spam-senders.aspx
- Barracuda Networks http://www.barracudacentral.org/lookups

[76] Unless, that is, their mail server takes their IP address reputation into account when they submit mail. This is unlikely as the subscriber should be using some form of authentication when submitting email.

in service contracts. For example, some VoIP services stipulate the use of static IPs and specific NAT configurations.[77]

Traditional NAT has often required the rewriting of applications to work around its limitations. Skype is a common example of this as it exists purely as a solution to the NAT traversal problem. Significant effort (time and money) has been extended to enabling services and applications to work behind traditional NAT. In some cases, it has not been economically viable to deploy applications because of the difficultly of overcoming the restrictions of NAT. CGN will significantly compound this problem.

Within some organisations there are whole teams dedicated to the issue of NAT traversal. For example, in Microsoft several years ago, there were five teams dedicated to addressing the issues of NAT traversal.

A side effect that was raised by some of the parties interviewed for this report is that the effort put into mitigating CGN issues will take away resources from implementing and deploying IPv6 in applications and services.

Providers whose applications and services are sensitive to CGN can mitigate problems by ensuring that their applications and services support IPv6. If they do, then in CGN environments where the ISP has also deployed IPv6 the applications and services will be able to avoid CGN and any associated problems.

6.4 Specific Implications for Games Consoles and Gaming in General

In gaming consoles and multiplayer games, peer-to-peer is used for a number of strategic reasons, including:

- **Highest possible performance**
  Direct peer-to-peer connections reduce latency. Direct path decreases the possibility of traffic traversing network bottlenecks which could reduce bandwidth.
- **Provider does not have to pay for bandwidth**
  Indirect connections mean that traffic would have to traverse the game or console provider's network. With potentially millions of users this would result in significant bandwidth costs.
- **Provider does not have to deploy excess server capacity**
  Direct peer-to-peer connections reduce the server capacity requirements at the provider.
- **Impact of CGN can be sudden and without warning**
  When an ISPs deploys a CGN on its network, it may place many (10,000s or 100,000s) of customers behind CGN at once. An ISP is unlikely to be able to warn everyone of this prior to the event. This could reduce in a sudden and dramatic increase in support calls to the games console or game vendor as features of their products fail to work behind CGN.

For these reasons the impact of CGN upon games consoles and games can be significant. It should be noted that in certain scenarios the results will be complete failure. This is already the

[77] An example of this is the VoIP service provided by AAISP, which stipulates NAT requirements in its terms of service.

case with NAT44 where some implementations will cause games consoles and certain games to fail. It is not the case that the games console and game providers can work around all forms of CGN, they cannot.

In a recent draft information RFC[78], it was reported that Microsoft's Xbox has been modified so that it is less likely to fail behind CGNs. The draft RFC includes this update:

> *"Update: in December, 2011, Microsoft released an update for Xbox. While we did not conduct thorough testing using the new release, preliminary testing indicates that Xboxes that upgraded to the latest version can play head-to-head behind a CGN, at least for some games."*[79]

Notice the final words; "at least for some games". The conclusion that can be drawn from this is that for other games the Xbox will not work behind CGNs.

It is also worth noting that in the same draft RFC, Playstation 3 is listed as "pass". However, our research revealed that this is over simplistic and there are circumstances in which Playstations will fail behind CGNs.

6.5 Application and Service Features

CGN can have an impact on specific application and service features. Creating a exhaustive list of features that may be impacted by CGN is impossible. This is because these features are often:

- application or service specific
- innovative and unique

To illustrate the kinds of features that can be impacted by CGN we have included a number of examples below.

6.5.1 Multiplayer Gaming and Geo-proximity

Geo-proximity is discussed in Section 5.4.9 of this report. Some games and games consoles use geo-proximity measurements to determine which users are located geographically close to each other. This allows them to restrict users that play each other, to those users that have the lowest latency between each other. This is extremely important to the quality of the gaming experience.

Note too that this is only beneficial if the users are able to also communicate directly that is in a peer-to-peer configuration. CGN may also make this impossible.

6.5.2 Peer-to-Peer VoIP

VoIP communications can take place peer-to-peer or through a server that is connected to by both parties in a call. Peer-to-peer has technical benefits such as lower latency and little need to provision bandwidth for VoIP calls on a central server or servers. However, in some cases, peer-to-peer is necessary to avoid the application provider being forced to operate as a

[78] C. Donley, L. Howard, V. Kuarsingh, J. Berg & J. Doshi, *Internet Draft: Assessing the Impact of Carrier-Grade NAT on Network Applications*, draft-donley-nat444-impacts-06, http://tools.ietf.org/html/draft-donley-nat444-impacts-06, 2013, accessed 15 April 2013.

[79] ibid.

telecommunications provider and thereby adhere to national or international telecommunications laws. In a CGN environment, peer-to-peer communication may be impossible forcing traffic to traverse a server. In the example given here using a server is not an option so the server would fail.

6.6 Support and Management

Application and service providers face particular difficulties when it comes to supporting customers behind CGN. Firstly, they usually have no visibility that their customer is behind CGN. Then they have no access to the information that they could use to diagnose and address problems.

ISPs are not providing application and service providers with the tools that they need to diagnose and resolve CGN related issues.

Application and service providers expect there to be a significant increase in support calls as a result of the deployment of CGN. In most cases they do not expect to be able to determine that the cause is CGN.

6.7 Logging

Application providers, content providers and service providers will all have an increased logging requirements as a result of the deployment of CGN in their user's ISP's access network. They will now need to record more detail in order to provide information for law-enforcement agencies.

Previously, if requested, all they would need to supply would be the IPv4 address of the client at a specific time. With CGNs in the traffic path, they will need to record the IPv4 address and port number at an accurate time. Since they do not know which clients are behind CGNs, they will have to log this information for all their traffic.

Accurate timing will be critical for both the ISP and the application providers. If there are timing mismatches then a different ISP subscriber could be using the IPv4 address than the one who needs to be traced. It is conceivable that innocent parties could be blamed for acts they have not carried out.

All web-servers,[80] application-servers and service providers will need to add port numbers to their logging systems. This may require software modifications and/or upgrades.

6.8 Security

Some applications and services use the source IP address as a factor when determining the authenticity of a client. Since the source IP address can be spoofed, it is not on its own sufficient as an authentication mechanism. Despite this, some applications still limit connects from multiple users from one IP address. Examples include some banking applications that will fail if too many users try to login using the same IP address or will fail if different sessions use different source IP addresses as in mobile networks CGN implementations.

[80] In the Apache web server this can be achieved using the %{remote}p log format. See: Apache, *Apache Module mod_log_config*, http://httpd.apache.org/docs/2.2/mod/mod_log_config.html#formats, 2013.

As far as the authors of this report have been able to determine, no such functionality currently exists within Microsoft's IIS web server.

It is expected that applications that use the source IP address as a factor in security techniques will have to be modified to take account of CGNs in the network.

Section B: Policy Implications

7 Implications of CGN for Public Policy

The following section provides an analysis of public policy implications associated with adoption of CGN. It begins with a literature review and provides a critical analysis of two popular schools of thought on network theory, and the state of thinking as to whether or not IP addresses are property. In each case, the relevance of the issues to CGN is assessed. Next, is an evaluation of the impact of CGN on competition, in particular the impact on potential barriers to entry, slowing innovation and the competitive position of the UK versus world markets. The impact of CGN on privacy, security and intellectual property enforcement is considered next, followed by an assessment of the impact on the Internet industry (both at the ISP and application provider levels) and consumers of widespread CGN deployment. The section ends with conclusions.

Much of the analysis is inevitably speculative given that deployment of CGNs is still in its infancy. Where available, evidence is drawn from relevant comparators. Where conclusions are hypothetical, this is stated. However, we are able to draw conclusions on the basis that CGN highlights and exemplifies well-worn issues in the Internet policy arena, such as network neutrality, the tension between data retention and privacy, and the role of Internet service providers as proxies for law enforcement.

7.1 Methodology

In preparing the policy review, the authors have reviewed relevant academic and industry literature (see bibliography), and have interviewed industry experts.[81]

There is some industry literature on CGN (see Annex A), but for the most part CGN has escaped academic analysis. The technical review indicates that CGN raises familiar policy issues on which there is substantial academic and industry debate. Therefore, it is appropriate to review the generalised literature and highlight areas relevant to CGN.

7.2 Issues Arising from the Technical Review

The technical review of this study highlights the following issues that have policy implications:

- UK incumbent ISPs currently hold sufficient IPv4 stocks to maintain service
- A grey market in IPv4 addresses is emerging
- A lack of technical standards risks ad hoc implementation of CGN
- CGN creates large, opaque networks behind single IPv4 address nodes
- Large CGN networks limit the number of concurrent sessions available to users (see Section 4.1) requiring proactive traffic management by the network provider
- CGN brings increased complexity and cost in keeping accurate traffic data attendant with CGN. Experience from the mobile environment suggests that some data are simply not kept at all.
- CGN will impact application and service providers, and may lead to impaired services or increased costs for consumers and business users

[81] See Section 16, Acknowledgements.

8 Literature Review

CGNs are, like many network management techniques, the prerogative of the networks that own the infrastructure they operate over. The neoliberal philosophy dominant in much of the developed world suggests that private industry will find the most efficient and stable ways to ensure their businesses continue to flourish. However, as shown earlier in this report (see Section 4), CGNs operating on private infrastructure can have impacts outside the boundaries of the networks that have chosen to implement them - such as security complications and data privacy implications - which fall into the sphere of public goods management.

The *Communications Act 2003* and *Digital Economy Act 2010* both make provision for the regulation by Ofcom of "electronic communications networks", therefore providing Ofcom with the option of mitigating any negative ramifications of CGN deployment in the UK.

While there is little specific literature on the policy implications of CGN deployment, the changes to the structure of the Internet's basic transport system inherent in CGN deployment bring in elements of network theory, network neutrality and frameworks for understanding how to manage the complex public/private structure of the Internet, on which there is extensive academic and industry thinking.

The rest of this section reviews relevant elements of that literature.

8.1 The Effects of Network Formation

Network formation theory explains how the members of a network can collectively leverage benefits greater than those they could on their own. It is the cooperation between members of the network, and an common set of operating principles, that allows this to occur:

> *"Like the railroad system or the electric power grid, the Internet is a collection of independent networks that coordinate their actions, forming what appears to be a seamless collective. This structure allows all users, application creators, and content providers to leverage the full power of the global inter-network. The Internet fosters innovation by eliminating transaction costs, enabling new services to emerge…. Common networks facilitate innovation independent of the infrastructure platform, which can create significantly more value than the network itself… [A] company such as Amazon.com need not worry about how its customers access the network. It can deploy new services and features without making special arrangements with network operators."*[82]

However, as this report has demonstrated in Section 3.3.8 above, deployment of CGNs may change the ease with which new services can be deployed. Network formation theory, which models 'what happens as networks add and remove connections', is useful in assessing how changes in some network nodes, such as the implementation of CGNs, may have an effect on the characteristics of the network as a whole.[83] This is primarily the case in networks like the Internet, where growth of connections is not linear:

[82] K. Werback, "The Centripetal Network - How the Internet Holds Itself Together, and the Forces Tearing It Apart", *UC Davis Law Review*, 2008, vol. 42, pp. 347-9.

[83] Ibid p. 346.

> *"Network effects tend to produce powerful hubs because new nodes express 'preferential attachment' to the most-connected nodes in the existing network. The best connected nodes become even more dominant as the network grows."*[84]

The result of this asymmetrical node formation is that the most connected hubs "see the opportunity to become more proprietary, and those outside the hubs worry that the hubs will dominate them".[85] For example, when in 2012 Twitter's API dropped support for RSS standards, that single change had an impact on numerous third party applications that used the API, like FriendFeed, requiring them to modify their applications.[86]

Despite the significant influence dominant hubs can have on all participants in the Internet, the idea of "dumb network at the centre; intelligence at the edges" remains key to the way Internet technology has developed. The addition of CGNs to a "dumb network centre" further disrupts the "end-to-end" principle.

8.2 The End-to-End Principle

For many, the interoperability and simplicity of the IP layer are key reasons for the Internet's growth and its continued ability to foster innovation over the past 15 years.

TCP/IP was created to enable diverse, heterogeneous and incompatible networks to interconnect.[87] This new interconnectivity permitted two hosts on separate networks to communicate seamlessly, with no need to be aware of, or make provisions for, any differences between the lower levels of the protocol stack of each host's network. Over time, the ability for end points on the Internet to communicate seamlessly became known as the "end-to-end principle". Associated with the end-to-end principle is the "simplicity principle": while there are many choices that can be made at the lower levels of the protocol stack (wifi vs wired, etc) and at the higher levels of the protocol stack (ftp vs http), the network layer's simplicity—sending packets from source to destination and that is all—is the Internet's greatest virtue. Simplicity enables "smart" edges of the network to engage in application layer innovation that can pass unimpeded over a "dumb" network core.[88]

While "end-to-end" has been one of the driving principles of the Internet, operational reality has been more complex. NATs interrupt the unimpeded passing of packets from end to end, as do their more complex cousins, CGNs. Other forms of interruption of the principle include firewalls, deep packet inspection, and technical measures to block or filter content.

While there are many operational ways that disrupt the end-to-end principle, for more than a decade, the IETF chose not to include NAT-related technologies as part of its standards development. This has resulted in non-standardised forms of address translation, and the continued development of other Internet technologies with specifications that assume end-to-

[84] K. Werbach, "Connections – Beyond Universal Service in the Digital Age", *Journal on Telecommunications & High Technology Law,* 2009, vol. 7, p. 84.

[85] K. Werback, "The Centripetal Network - How the Internet Holds Itself Together, and the Forces Tearing It Apart", *UC Davis Law Review*, 2008, vol. 42, p. 346.

[86] C. Warren, *New Twitter API Drops Support for RSS, Puts Limits on Third-Party Clients*, http://mashable.com/2012/09/05/twitter-api-rss, 2012, accessed 15 Aril 2013.

[87] J. Naughton, *A brief history of the future*, Phoenix, London, 1999, pp. 162-163.

[88] R. Bush & D. Meyer, *Some Internet Architectural Guidelines and Philosophy*, RFC 3439, http://www.ietf.org/rfc/rfc3439.txt, 2002, accessed 15 April 2013.

end connectivity is the Internet's default operational environment. The outcome of this commitment to the end-to-end ideal has been the non-standardised development of protocols that can bridge NATs and other end-to-end disrupting technologies.

More recently, however, the IETF has recognised the need to address complications that NATs, and CGNs, cause. IETF currently has a Behavior Engineering for Hindrance Avoidance ("BEHAVE") Working Group that is looking for ways for both IPv4 and IPv4/IPv6 NATs to function in more uniform ways.[89]

8.3 Network Neutrality

The network neutrality debate provides insight into how other forms of network management, namely traffic management, can affect consumer rights and competition. By adding complexity to the dumb centre of a network, and removing the ability to have unimpeded end-to-end connections, CGN can have a significant effect on services that rely on end-to-end connections, as well as on the consumers who use those services.

There is extensive, diverse, and ongoing debate about network neutrality. Adopting the terminology of Werback[90] the literature reveals three schools of thought:

1. **Openists**

   Openists, including the well-known Lawrence Lessig, advocate the maintenance of a "dumb network that does not differentiate between different types of traffic… They support a policy based on open access, in which Internet infrastructure and applications cannot be bundled using either technical or business mechanisms. They argue that vertical integration harms consumers, that most innovation comes from application providers".[91] Whilst the literature does not mention CGNs, CGN deployment clearly places the Internet service provider in the position of causing changes to the network layer that affect many applications.

2. **Deregulationists**

   Deregulationists believe that the development of smarter networks—and vertical integration between layers of the Internet hourglass model—will support a diversified offering of applications. Deregulationists believe that, in a world of deregulation, Internet service providers are in the best position to determine the most beneficial approach for their users and that regulation of network management techniques will reduce ISPs' desire to invest in their networks. Instead of arguing for network neutrality, Christopher Yoo argues for network diversity, "an environment in which networks make different choices about architecture, pricing and service".[92] Viewed through this frame, CGNs would be an example of a diversified, smarter network.

3. **Nondiscriminationists**

[89] IETF, *Behavior Engineering for Hindrance Avoidance (behave) – Charter*, https://datatracker.ietf.org/wg/behave/charter, n.d., accessed 15 April 2013.

[90] S. Jordan, "Implications of Internet Architecture on Net Neutrality", *ACM Transactions on Internet Technology*, 2009, vol. 9, no. 2.

[91] Ibid.

[92] K. Werback, "The Centripetal Network - How the Internet Holds Itself Together, and the Forces Tearing It Apart", *UC Davis Law Review*, 2008, vol. 42, p. 375.

Nondiscriminationists, including Tim Wu, view net neutrality as more complex than either the openists or the deregulationists. The nondiscriminationist, Jon Peha, suggests there can be both good and bad uses of traffic management.[93] Using Peha's concepts, a good use could be prioritising realtime video traffic, while a bad use could be filtering traffic to enable discriminatory pricing schemes. Tim Wu writes:

> *"The concept of a total ban on network discrimination is counterproductive. Rather we need to distinguish between forbidden grounds of discrimination, those that distort secondary markets, and permissible grounds, those necessary to network administration and harm to the network".*[94]

A nondiscriminationist view would be that deploying a technology like CGN, may be beneficial on the basis that it leverages a scarce resource—IPv4 addresses—and that the focus should be on mitigation or elimination of potential negative effects on other parts of the network, including secondary markets.

8.3.1 Regulation of Networks

Werback articulates the positive role that can be played by regulators of the Internet ecosystem: catalysing network formation and moderating the forces that push towards excessive concentrations of power.[95] The concept of light touch regulation on preventing distortion of competition echoes the Communications Act. In this way, rather than mandating a blanket openist form of network neutrality, or adopting the deregulationist assumption that the market will always self-correct itself, regulators can ensure that both the needs of the market and consumers are weighed up when the deployment of technologies like CGN are being considered by market players. In the area of Internet regulation, the concept of the semicommons, described below, is a particularly useful framework for understanding the intersection between market needs and public goods ideals.

8.4 Internet as a Semicommons

Semicommons theory[96] seeks to explain why and how hybrid public/private systems like the Internet work so well. The theory also describes why such hybrid systems tend to experience tensions resulting from the concurrent management of private goods and common goods. The Internet is described by many analysists as a public good: the end-to-end principle enables anyone who has a connection to the Internet to create a website, application or even a new protocol to run on top of the IP-based network layer.

At the same time, given that the infrastructure of the Internet is mostly privately owned, the owners have rights to run their piece of the infrastructure as they see fit. Semicommons theory aims to identify the right balance between these two systems with competing sets of priorities (public vs private). Semicommons theory has much in common with the nondiscriminationist

[93] J. M. Peha, "The benefits and risks of mandating network neutrality, and the quest for a balanced policy", *Proceedings of the 34th Telecommunications Policy Research Conference*, http://repository.cmu.edu/epp/27, 2006, accessed 15 April 2013.

[94] T. Wu, 2003, "Network Neutrality, Broadband Discrimination", *Journal on Telecommunication and High Technology Law*, 2003, vol. 2, p. 170.

[95] K. Werback, "The Centripetal Network - How the Internet Holds Itself Together, and the Forces Tearing It Apart", *UC Davis Law Review*, 2008, vol. 42, p. 410.

[96] J. Grimmelmann, "The Internet is a Semicommons", *Fordham Law Review*, 2009, vol. 78.

view of network neutrality; rather than prioritise either the public value of the Internet or the private rights of the network owners, it recognises the value of both, employing "sharing rules and boundary-setting to keep the whole thing functioning".[97]

James Grimmelmann has applied semicommons theory to the Internet.[98] Particularly relevant to the policy implications of CGNs is Grimmelmann's recognition that the Internet's private sphere includes not only the connectivity providers, but also the end user. Both types of owners can exercise their rights by using their property in ways that best suit them. In the case of CGNs, many consumers choose not to buy IPv6-enabled routers and other equipment. This in turn affects ISPs' investment decisions: an ISP may have considered IPv6, but if their customer base won't buy IPv6-capable devices, then the ISP must consider non-native-IPv6 options for its network. Equally, lack of customer interest in IPv6 affects equipment manufacturers, who continue to produce many IPv4-only devices. The reluctance of the owners of private goods to invest in IPv6, an investment that would be in the interests of the common good, is a classic example of the "tragedy of the commons": everyone wants to get maximum use out of the goods, but nobody wants to pay for maintenance, a situation Tim Berners-Lee equates to a "utopian commune with no structure, which doesn't work because nobody actually takes out the garbage."[99]

The semicommons framework is helpful in emphasising the complexity of a network environment like the Internet. So while it is obvious to an observer that IPv6 adoption is slow, the causation is less obvious. Numerous actors (including consumers) are contributing. Semicommons also fits with the UK communications framework in advocating regulatory intervention as a last resort, only where there is actual harm to public values such as security or competition, or a risk of a "tragedy of the commons".

[97] Ibid, p. 2818.

[98] Ibid.

[99] T. Wu, *The Master Switch*: *The Rise and Fall of Information Empires*, Vintage, New York, 2010, p. 283.

9 Impact of CGNs on Competition

The following section uses concepts and findings from the technical analysis and literature review to assess the potential impact of widespread CGN deployment on competition in the market for Internet access, applications and services. It uses EU competition law as the framework for review, but it is important to note that the analysis is hypothetical. No widespread CGN deployment has taken place in the UK or in other comparator markets. The analysis makes no findings in competition law—for example, market analysis, dominance or abuse—it simply points to possible *effects of CGN deployment* that may affect existing (or new) services, and the impact on ISPs, Internet application and services providers, and Internet users including consumers.

9.1 Current Situation

At present, CGN implementation is in early stages and has yet to make a significant impact on the market. The assumptions are the other market conditions remain the same, in particular:

- There are high levels of UK Internet access
- Established UK ISPs have existing pools of IPv4 addresses to draw from[100]
- Uptake of IPv6 in the UK continues to be sluggish

There are a number of ISPs providing Internet access services in the UK. In 2013, the top 10 have a subscriber base of 22.3 million, of which the top four have over 19 million subscribers.[101]

Barriers to entry into the Internet connectivity market (whether fixed or mobile) are high, requiring substantial physical infrastructure, spectrum allocation, overcoming the challenges associated with obtaining IP address allocations from RIPE (even pre-IPv4 scarcity), setting up and paying for both peering and transit from upstream providers. New entrants are likely to have less buying power, and therefore pay higher prices for transit than established providers.[102]

The Internet is now "at the heart of how many people communicate" in the UK.[103] On every headline metric, whether Internet take up or broadband take up (fixed or mobile), the trends are upward. For example:

- The UK has a home Internet access rate of 80%, with over 64% of 65-75 year olds now online
- 41% of the UK's smartphone users consider themselves to have a "high level of addiction" to their smartphones
- On average, each UK household has three different types of Internet- enabled device[104]

[100] Ofcom, *The Communications Market 2012*, http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr12/CMR_UK_2012.pdf, 2012, section 4.1

[101] *ISP Review*, www.ispreview.co.uk/review/top10.php, accessed 28 February 2013.

[102] For an explanation of the economics of peering and transit see: R. van der Berg, *How the 'Net works: an introduction to peering and transit*, http://arstechnica.com/features/2008/09/peering-and-transit/2/, 2008.

[103] Ofcom, *The Communications Market 2012*, http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr12/CMR_UK_2012.pdf, 2012, section 4.1.

In short, consumers are increasingly reliant on the Internet in general, and on sophisticated, session-hungry applications in particular.

9.2 Likely Impact of CGN Deployment

There are a number of scenarios in which CGN deployment could have an impact on competition, both within ISP and Internet services/applications markets. The next section looks at the likely impact of CGN deployment within each of those sectors and how existing suppliers, new entrants and consumers may feel the effects.

9.3 Impact of CGN on Internet Access Markets

9.3.1 Benefits of CGN deployment

The deployment of CGN may benefit existing ISPs with large IPv4 stocks, by delaying the need to invest in IPv6 and increasing the number of subscribers served by each IPv4 address. CGN enables a seamless continuity of service despite IPv4 scarcity and exhaustion. Some providers explicitly see CGN as a staging post, and a significant driver for adopting IPv6.[105]

Consumers in the short term may also benefit from the retention of IPv4, since it will delay or avoid the need for them to invest in IPv6 compatible hardware

9.3.2 Impact on New Entrants and Innovation

New market entrants in the UK ISP market face already high barriers to entry. The enhanced strategic advantage given to those with IPv4 addresses may deter market entry particularly from innovative new players.

As IPv4 addresses become more difficult to obtain, those with existing stocks of IPv4 will be at a relative competitive advantage over competitors, in a way that is unrelated to the quality of their value proposition to customers. CGN enables those with existing IPv4 addresses to increase their customer base and size of network, an opportunity not open to those without access to IPv4 addresses.

9.3.3 Impact on Consumers and Business Users

Assuming that Internet penetration in the UK is unlikely to grow exponentially, is the lack of IPv4 addresses or CGN deployment really going to inhibit competition? Potentially, yes, because IP addresses do not map to subscribers in a ratio of 1:1. Increasingly, individual consumers will have multiple devices with Internet access, running applications or services that require numerous concurrent sessions. For example, over half of UK households have three or more Internet-enabled devices, and games consoles are more popular than laptops or PCs amongst DE households.[106] Looking to the near future, sensor technologies (such as burglar alarms, fire detection and remote monitoring of home appliances and infrastructure) will demand further IP capacity as well as providing opportunities for innovation in the ISP and application markets.

---

[104] Ibid.

[105] For example, interview with BSkyB.

[106] Ofcom, *The Communications Market 2012*, http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr12/CMR_UK_2012.pdf, 2012, section 4.2.3

1. **Consumers**
   Consumers are increasingly reliant on sophisticated services, which the technical analysis in Section 5.6 shows are likely to work less well on CGN. Therefore, it can be anticipated that a two tier charging system would emerge.

   Internet consumers have long had the option of choosing to pay more for a fixed IP address or less for a dynamically assigned IP address. Consumers might also choose to pay more for a non-CGN connection after they begin to experience service degradation for applications they had, pre-CGN, taken for granted.

   Deployment of CGN brings opportunities for more fine grained market segmentation by existing ISPs (an opportunity not available to those without access to IPv4 addresses) by offering basic services for those suited to a CGN and premium services for those with more sophisticated demands.

   While greater choice may benefit consumers, concerns may arise, if customers are, in reality, forced to pay more to achieve reliable performance of Internet applications that would function seamlessly in non-CGN environments.

2. **Business customers**
   Business customers rely on high levels of up-time for their Internet-based services. They are therefore likely to pay for fixed IP addresses to ensure that their services remain available to their customers. Due to the nature of NATs (where outside connections cannot initiate a connection to a location behind a NAT), businesses already need to choose this more expensive fixed IP address option to ensure their services are available to potential customers.

   Where CGNs may complicate matters is with the growth of telecommuters and home-based businesses. Businesses may find that telecommuters are not able to access business-specific applications such as virtual private networks behind CGNs, increasing the demand for fixed IP address connections, potentially leading to inflated prices for the few IPv4 addresses not being deployed behind CGNs.

   With businesses increasingly needing to obtain public addresses to guarantee quality of service, a prolonged life of the limited IPv4 market will be under immense pressure from both operators of CGNs wishing to expand the address pool available to their CGNs as well as from consumers wishing to use public IP addresses. This is likely to lead to higher prices for Internet access.

9.3.4 Impact on Existing, Smaller ISPs

Any distortion arising from CGN adoption may well be corrected via a competitive response from providers offering cheaper connection prices and more reliable service levels for applications through adoption of IPv6.

However, network formation theory predicts that networks show preferential attachment to powerfully connected network hubs. In the context of CGNs, where many of the networks that are considering deploying CGNs are well-connected nodes on the Internet (BT, Virgin Media, Sky Broadband, Talk Talk), and are therefore attractive to others to connect with. Deploying CGNs on large hubs in the Internet's fabric will have a significant impact on smaller nodes that connect to those hubs, even if smaller nodes choose not to deploy CGNs. In particular, less-connected nodes may find less reason to deploy IPv6 when the larger hubs they are connected to prefer to extend the life of their IPv4 resources. Instead, the CGN practice spreads from dominant hubs out to the smaller nodes connected to it.

The outcome is a prolonging of the IPv4 market. Due to an ever-expanding global Internet, however, the market's limited IPv4 resources will become increasingly more desirable, and therefore more expensive to obtain, particularly for smaller ISPs.

9.3.5 Impact on "Mere Conduit" Role of ISPs

By employing CGNs that can break applications, ISPs then need to make network management decisions on a case-by-case basis as new applications emerge. The implication here is that CGNs enable ISPs to operate their networks ways that move them further away from being "mere conduits" (providers of "dumb network centres"[107]) towards more active participants in decisions about what services and applications should and shouldn't be accessible to their customers.

Currently, applications in a non-NATted environment operate in a "permissionless" environment, regardless of bandwidth consumption or any other factor. In a CGN world, ISPs may make decisions on whether or not to allow applications through their CGNs based on factors such as bandwidth use, or contractual relationship with the application provider. For example, an ISP could decide that peer-to-peer applications similar to BitTorrent that break by default across the CGN are, due to the illegal nature of many files shared using the application, not worth manually configuring the CGN network to permit. If an ISP, on the other hand, did manually configure their CGN gateway to permit an application that enabled illegal activity, the ISP may have actively enabled that activity to exist. This potentially erodes ISPs' mere conduit status and risks incurring liability as an Internet intermediary.[108]

9.3.6 Summary

In summary, whilst CGN deployment will benefit existing ISPs with large IPv4 stocks, network effects and scarcity of IPv4 addresses would be accentuated by CGN deployment, creating additional challenges for new entrants to the UK ISP market. In the medium to long term, this scenario could adversely impact consumers by reducing choice of ISP and potentially enabling existing providers to raise prices.

In addition, the need for ISPs using CGNs to make decisions about which online services should be allowed through their CGN gateways creates a potential change in role for ISPs away from mere conduits of Internet access towards a role in which the ISP is seen as liable for the content and applications they permit on their networks. More about the interaction between CGN-based ISPs and online services is discussed in the section directly below.

9.4 Internet Applications and Services

Deployment of CGN is likely to have a far-reaching impact on the wider Internet application and services industry. As suggested by network formation theory (see Section 9.1), if the larger ISPs choose to deploy CGNs, it will have a significant impact on the developers and users of online services.

The technical section of this report has described how CGN impairs performance and breaks many Internet applications. Unless CGNs are standardised, there will continue to be problems for application developers wishing to develop services that can be reached by anyone on the

[107] See Section 8.3, Network Neutrality.

[108] For more information on discussion on the role of Internet intermediaries, see: OECD, *The Economic and Social Role of Internet Intermediaries*, http://www.oecd.org/internet/ieconomy/44949023.pdf, 2010.

Internet.[109] With non-standardised CGNs introducing a range of different behaviours that cannot be anticipated by application developers, many applications will need to rely on ISPs adjusting their CGNs to allow the service to operate on their network on a case-by-case basis.

There are parallels in the deployment of Internationalised Domain Name (IDN), where some browser vendors operate "white lists" of specific domains whose policies the vendor has approved. Only white listed IDNs are supported in such browsers, and some domain name registries point to this practice as a factor inhibiting wider uptake of IDN.[110]

Such case-by-case decision-making provides opportunities for those deploying large CGN networks to inhibit competition by failing to enable new applications working on their CGN in a way that would not have been possible in an IPv6 or non-CGN IPv4 environment. As such, amounts to a fundamental change in the Internet's architecture, moving away from the current "innovation without permission", to a position where ISPs become gatekeepers standing between application or service providers and consumers.

The need to enable applications and services on a case-by-case basis also creates additional cost and complexity for both application/service providers and ISPs, who may need to support a permissions system to allow access (where none is currently needed). Additional costs or liability for ISPs may arise from the erosion of the separation between content and carriage. In a permission-based system, it will be difficult to sustain mere conduit immunity, as noted in Section 9.3.5 above.

Finally, is important to note that CGNs will not stop innovation entirely or may provide opportunities for innovation. For example, as described in Section 5.2.3, Skype's development was inspired by the need to replace VoIP services that break over NAT deployments.

9.4.1 Potential for Discrimination

There is evidence that CGN adoption can break services (for example, location services, Google Maps, peer-to-peer gaming as discussed in Sections 5.4 and 6.4). At the same time, the CGN provider's own services, which do not have pass through the CGN translation process, will work well within the ISP's own network. CGN is thus analogous to the creation of a closed system. While the Internet has many examples of closed systems at the application layer of the OSI model, the creation of closed systems at the network layer undermines the basic utility of the Internet (that is, dumb networks connecting heterogeneous systems).

The literature offers examples of discriminatory behaviour occurring within closed systems. For example, following an incident in 2005, the US Federal Communications Commission found that an Internet Service Provider, Madison River Telephone Company LCC "was cutting off access to Vonage and other VoIP services by blocking certain IP ports". Apple's application store in 2008 rejected applications such as Google Voice on the grounds that it "duplicates features that come with the phone".[111] The relevance of such examples to this study is that CGN offers a single provider greater level of control over, and access to, its network for applications than a provider with non-CGN architecture.

[109] Even if CGNs are standardised, legacy non-standardised CGNs will, in all probability, remain in use for many years.

[110] *EURid/UNESCO World Report on IDN Deployment 2012*, http://www.eurid.eu/files/publ/insights_2012_idnreport.pdf, 2012, pp. 23-24, accessed 15 April 2013.

[111] S. D. Meinrath, J. W. Losey & V.W. Pickard, "Digital Feudalism – Enclosures and Erasures from Digital Rights Management to the Digital Divide", *CommLaw Conspectus*, 2011, vol. 19, pp. 423-479.

9.4.2 Relationship Changes Between ISPs and Internet Application and Service Providers

Internet application and service providers will seek ways to enable their products to reach customers behind CGN networks when CGN traversal techniques do not work. Two ways that are likely to emerge are variations of establishing closer relationships between ISPs and the providers of online applications and services:

1. **Paying for access to CGN networks**

   Application providers experience increased costs through having to change their products to include CGN traversal techniques. This is problematic to achieve due to the non-standard nature of CGN deployments, and may lead to innovative applications or services never coming to market—a loss which is hard to quantify, but may nevertheless have significant impact on the UK's long term economic competitiveness.

   Net neutrality literature shows that application providers already feel the need to develop contractual relationships with ISPs in order to guarantee the operation of their application on particular networks.[112] Such contracts seek to avoid their applications being on the losing side of any traffic management. In traditional Internet networks, applications "just work", without the need for the network provider's permission. CGNs may break many of today's applications, requiring the ISP to manually configure its CGN gateway to allow specific applications to traverse the network. This may give rise to a growth of payment by application developers for access to networks. Clearly, it is foreseeable that in some scenarios ISPs would refuse such permission or charge a premium. By omitting to manually configure their gateway to support applications for those without contracts, ISPs will be able to block those applications from accessing the ISP's subscriber base.

2. **Hosting mirrored services behind CGN gateways**

   Large online service providers such as Google distribute mirrors of their services to locations around the world partly in response to calls by ISPs for content providers to pay for the increasingly large expenses incurred by ISPs when their customers use bandwidth-heavy services that travel international routes. Such distributed content and services are often currently housed at Internet Exchange Points. However, with the advent of CGNs, and the existing practice of housing mirrors of online services now well established, larger online service providers may develop agreements with larger operators of CGNs to house mirrors of their services behind the CGN gateway to enable customers within the CGN to access their services.

9.4.3 Impact on New Entrants and Innovation

If the majority of large ISPs are using CGNs, the requirement that new applications and services be "allowed" separately by each ISP could result in innovation being stifled.

Harm would arise for new entrants where the bargaining power of ISPs is greater than the application provider. For established application providers, such as Google (mapping), ISPs are more likely to make adjustments to enable their customers to access them than they are for new applications or new market entrants. For new entrants, or services, however, the deployment of CGN is likely to increase the bargaining power of ISPs by giving them a new

[112] S. Jordan, "Implications of Internet Architecture on Net Neutrality", *ACM Transactions on Internet Technology*, 2009, vol. 9, no. 2, p. 5:12.

"gatekeeper" role—giving permission for suppliers to access their subscriber networks—which does not exist in non-CGN environments.

As network formation theory has shown (see the literature review section above), the fewer connections that a node—or in this case an application—has, the less likely it is that future connections will be made to it. CGNs operated on large networks may prevent applications from reaching the critical mass of users that would make the enterprise viable. This would limit the financial value of a new application for its developers, and also limit its value to customers and any third party applications that may have been able to tap into the market potential offered by a large consumer base.[113]

9.4.4 Impact on Consumers and Business Users

For consumers, the impact of CGN would be felt through a reduction in choice, lack of access to innovative services or applications, and degradation in services. Even if ISPs are transparent in the information they give consumers about the likely effects of CGN on existing applications or services, the technical nature of the information and its inherent complexity may also serve to limit choice, or pose difficulties for consumers in distinguishing between services.

Ofcom's recent paper, *A Review of Consumer Information Remedies*,[114] addresses the problem of providing suitable information to help consumers make informed choices, providing a number of ways that information can be made more accessible. It notes, however, that inherently complex products and services may not benefit from increased information available for consumers, and that some providers of such complex services may, in fact, benefit from consumer confusion created by the "option swamp".[115]

9.5 UK versus International Competitors

Deployment of IPv6 is well advanced in other countries and territories outside of the UK. For example, the Chinese government is pushing for IPv6 deployment within the country. India, too, with its rapidly expanding Internet market, is looking to the bigger address space of IPv6 to solve its Internet needs. With over two billion citizens in these two countries alone, what happens with IPv6 deployment in China and India will affect international markets.

While developing applications and services for a CGN-heavy environment in the well-developed markets in the UK and USA in the short-term may produce financial benefits, in the longer term, there is a risk that UK online service providers or application developers won't be able to take advantage of IPv6-enabled markets that, ultimately, are the long-term Internet future.[116]

---

[113] There are a large number of third party applications that leverage the benefits of successful services such as Facebook, Twitter, Pinterest. On top of such applications, businesses in sectors outside the Internet industry use such services to develop their own consumer bases.

[114] Ofcom, *A Review of Consumer Information Remedies*, http://stakeholders.ofcom.org.uk/binaries/research/research-publications/information-remedies.pdf, 2013, accessed 22 April 2013.

[115] Ibid, p.43.

[116] Not only is IPv6 the long-term solution to the needs of an ever-expanding Internet, but given IPv6 is, like many new technologies, more easily deployed in developing markets that don't have complex legacy infrastructure to interact with—such as the rapid deployment of mobile telephony in Africa, where fixed line telephony infrastructure was never widely deployed—the UK online services sector will probably find a greater market for their services in the IPv6-based developing world than in domestic and other developed world markets.

Such a situation would potentially inhibit the global competitiveness of the UK in the medium term, or even lead to a fragmentation in the transport layer of the Internet along country or regional lines, inhibiting global communications.

10 Privacy, Security and Intellectual Property

This section considers the impact of widespread CGN adoption on privacy, security and intellectual property. As with the analysis of impact on competition above, it is our view that the policy issues arising from CGN are not new in themselves, but are closely related to, even accentuated versions of, those relating to IPv4 scarcity and NATs.

10.1 Privacy

10.1.1 Subscriber Logging and Traceback

Subscriber logging and traceback are important enablers of law enforcement and those enforcing private law rights (such as intellectual property infringement) in identifying potential perpetrators or infringers. It is well understood that such techniques risk infractions against individuals' fundamental rights of privacy and cause data protection issues, and European case law articulates the need for balance and proportionality.

Recent European cases *SABAM v Scarlet Extended*[117] and *SABAM v Netlog*[118] emphasise that intellectual property rights are not absolute, but need to be balanced with fundamental rights. In the *Netlog* case, the ECJ reversed an injunction intended to avoid copyright infringement, on the basis that the first instance court order was tantamount to imposing on Netlog general obligations to monitor all customers, at exclusively its own cost, for an unlimited period.

The technical analysis [in Section 4.9] describes the increased complexity inherent in providing subscriber logging and traceback within CGN, requiring the storage of more detailed parameters than in an IPv4 or IPv6 environment. By inference from the SABAM cases, the more detailed and potentially intrusive such data collection is, the more likely it is to run up against fundamental rights enshrined in privacy and data protection laws.

Furthermore, the opacity of the networks behind CGN mean that "innocent bystander" subscribers with no connection to allegedly illegal or unlawful activity are more likely to have their personal data processed by ISPs in response to third party enquiries, in order to identify the likely source of bad activity. This could raise concerns on the fundamental data protection principles such as proportionality, and fair processing.

10.2 Security

10.2.1 Increased Vulnerability of CGN Supernodes

The Internet's basic design recalls the vision of Paul Baran, a network engineer at RAND, of a "survivable network" which could withstand attack:

[117] SABAM v Scarlet Extended, (ECJ C70/10, 2011), http://curia.europa.eu/juris/liste.jsf?language=en&num=C-70/10, accessed 15 April 2013.

[118] SABAM v Netlog (ECJ C-360/10, 2012), http://curia.europa.eu/juris/liste.jsf?language=en&num=C-360/10, accessed 15 April 2013.

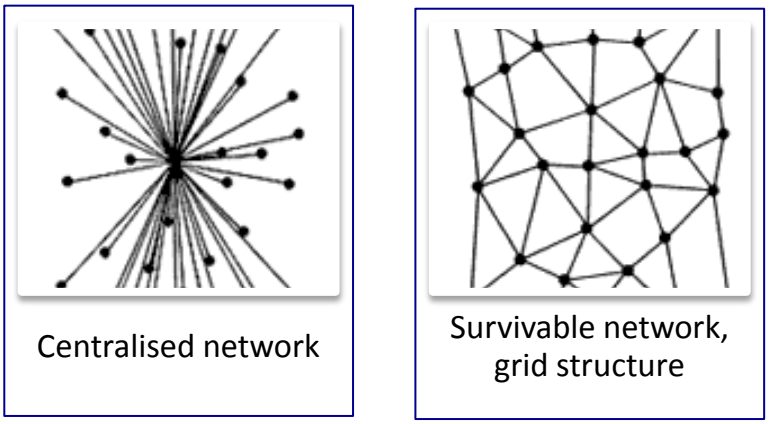

Figure 10.1 - Centralised and Grid Networks

Figure 10.1 shows a centralised network, typical of traditional communications networks such as telecommunications or broadcast media. A direct hit on the central node eliminates the entire network's connectivity. The Internet's innovation was to replace centralised networks with a grid through which packets would travel using whatever paths were available.

> *"There are significant benefits to the Internet's federated structure. Common networks facilitate innovation independent of the infrastructure platform, which can create significantly more value than the network itself."*[119]

Adoption of CGN alters the network's structure by creating large private networks behind a single or few IPv4 addresses. In effect, these become super-nodes which are more like the centralised design of traditional communication networks (figure 10.1). CGN therefore introduces the vulnerabilities associated with centralised communication systems including single points of failure, or possible capacity bottlenecks, and representing a single "attack point" for those with malicious intent.[120] Defending vulnerable attack points is likely to give rise to additional costs for ISPs through the deployment of additional mitigation infrastructure and associated systems support.

10.2.2 Standard Security Features Do Not Work on NAT and CGN

Security at the level of the Internet Protocol are provided by IPSec. The technical analysis indicates that IPsec is not viable behind CGN, which has a knock on effect on other products or applications that rely on IPSec, for example Vodafone's SureSignal mobile femotcell.

10.2.3 Changes to the Way Online Services Use IP Addresses for Security Screening

As shown in section 6 of the technical analysis, banking and other applications often rely on fixed source IP addresses for security reasons. This function, however, breaks in CGN. The impact of such breakage could be increased security vulnerability for consumers and banking customers or, more likely, denial of access to consumers' accounts through online banking.

---

[119] K. Werback, "The Centripetal Network - How the Internet Holds Itself Together, and the Forces Tearing It Apart", *UC Davis Law Review*, 2008, vol. 42, p. 348.

[120] Broadband Internet Technical Advisory Group, *Implications of Large Scale Network Address Translation (NAT)*, www.bitag.org/documents/BITAG_TWG_Report-Large_Scale_NAT.pdf, 2012, p 11, accessed 15 April 2013.

In the medium term, should CGN persist, it would lead to costs for banks (and therefore their customers) in developing solutions that work around these issues. The technical analysis shows that overcoming difficulties imposed by NAT or CGN is not technically trivial, and cannot be guaranteed.

Recent press reports of computer failure by banks in the RBS group indicate that when customers are denied access to their accounts, it becomes a national news story causing reputational damage to the banks, even if the problem (in this scenario, CGN) has been caused by a third party (the ISP of the customer).

This is a further example of the way in which CGN removes one of the main benefits of the Internet, the simplicity of the transport layer in the "Internet hourglass".

10.2.4 Difficulties for Law Enforcement and Lawful Interception through Increased Opacity and Complexity of IP Tracing

IP addresses are important to law enforcement in identifying the source of traffic, or (through reverse IP look up) identifying an individual subscriber. For years, ISPs have used dynamic IP address allocation which necessitates additional logging on the part of the ISP to be able to identify subscribers. With CGN, as the technical analysis describes, the process becomes more involved, and evidence from the mobile industry is that it is not feasible to keep records with the required granularity to identify subscribers with confidence. In the context of mobile providers, the problem is masked because of the availability of other useful data for law enforcement. However, transposed to fixed line, where the only relevant data is subscriber identity, CGN is likely to accentuate difficulties in lawful interception.

Until recently, the Regional Internet Registries responsible for IP address allocation policies have not recognised IP addresses as transferable from one holder to another. By convention, unused IP addresses were returned to the "free pool" ready for reallocation by the relevant RIR.[121] However, with IPv4 depletion, IPv4 addresses are increasingly being treated more like property, traded and transferred between "owners."

CGN requires IPv4 addresses. It is already apparent that a grey market is emerging for IPv4 addresses. Whilst some in the industry express concerns over this development, another view is that is shows the adaptation of the market to prevent wastage of increasingly scarce resources.

The security issue that arises is that at least for the period when (as currently) the market is ahead of RIR policy, there is no guarantee that trades and transfers will result in the RIR's records being updated. As a result, identifying subscribers or even operators through reverse IP address look-up is likely to be adversely affected, because it is possible that over time the RIR records for IPv4 will become less accurate, less authoritative. Because of the dependency on IPv4 of CGN, and the likelihood that widespread CGN would further drive up prices and increase trade and transfer, this is likely to delay or add further stages to law enforcement and other investigations, similar to current experiences of domain name WHOIS records.[122]

[121] This would primarily happen because the alternative, keeping unused addresses, would still attract maintenance fees from the RIRs, which most ISPs considered financially wasteful. The end result, however, was a system that would ensure that unused IPv4 addresses would be recycled and made available to networks that could use them.

[122] For more on issues for law enforcement arising from poor data accuracy of domain name WHOIS records, see *WHOIS Review Team final report*, http://www.icann.org/en/about/aoc-review/whois/final-report-redline-11may12-en.pdf, 2012. Of particular relevance to this report is the following quote from the report:

The greater complexity of tracing identity through IP addresses in CGN networks will require a higher level of technical skill from law enforcement and others with legitimate reasons for requiring subscriber identification than at present, leading to training costs.

10.2.5 Impact on Anti-Spam Measures

As noted in the technical analysis, there have been reports of anti-spam/anti-abuse measures impacting email clients behind CGN, as a result of mail servers detecting too many sessions from a single IPv4 address.[123]

In the event that an IPv4 address is blocked or blacklisted as a source of spam, the impact on a CGN would be greater, potentially affecting an entire subscriber base. This would increase cost and support load for the ISP, and, as we have seen earlier, damage its IP reputation.

10.2.6 Emergencies and Disasters

Recent experiences of natural disasters (for example, the Haiti and Chile earthquakes, and Japan's most recent tsunami) show that the Internet has been the most reliable and robust of communications networks, even at times when all others have failed. For example, in Haiti, the fact that the Internet still worked even though the earthquake in 2010 had destroyed much of the national communications infrastructure, allowed rescuers to use the Internet to locate and save victims trapped in the rubble of collapsed buildings.[124]

The technical analysis (see Section 6.5.1) indicates that CGN can severely impair the accuracy of geo-location services, which are currently relied upon by emergency services in the course of their work. Moreover, during an emergency, already congested networks will have to pass through an even more throttled gateway (CGNs and their limited IP addresses and ports), possibly making it much harder to take advantage of the benefits that the Internet, which is likely to impair the effectiveness of disaster response.

> *"In 2009-10, ICANN commissioned a study on data accuracy, which was undertaken by the National Opinion Research Council of the University of Chicago (NORC) (the "NORC WHOIS Data Accuracy Study 2009/10"). The study found that only 23% of WHOIS records met the study's criteria for No Failure, and over 20% were categorised as Full Failure or Substantial Failure3. Concerns about the accuracy of WHOIS records were raised in a number of responses to the WHOIS Review Team's public Discussion Paper and in public sessions at four ICANN meetings."*

[123] C. Donley, L. Howard, V. Kuarsingh, J. Berg & J. Doshi, *Internet-Draft: Assessing the Impact of Carrier-Grade NAT on Network Applications*, draft-donley-nat444-impacts-05, http://tools.ietf.org/html/draft-donley-nat444-impacts-05, 2012, accessed 15 April 2013.

[124] See *IGF 2012 Emerging Issues Main Session*, http://www.intgovforum.org/cms/component/content/article/114-preparatory-process/1254-igf-2012-emerging-issues-main-session, 2012:

> *"Haiti was another watershed moment for us because we saw that maps could be combined with other disaster relief tools. This is actually the area image of Port Au Prince in 2009 and this is what you see after the disaster. Responders used maps to plan and choose medical evacuations, locations. It proved to be a very useful tool. Then what Google were doing was that we created a disaster response crisis response team explicitly dedicated to disaster relief. Since then we have responded to about 28 disaster cases including, most recently, Hurricane Sandy which was late last month."*

11 Impact on Internet Industry

11.1 Increased Complexity of Data Retention, Logging, and Identifying Subscriber Sessions

11.1.1 Data Retention Costs for ISPs

As noted in Section 4.9 of the technical report above, to provide law enforcement agencies with the minimum data needed to track illegal activities, operators of CGN networks will need to expand the information they collect about subscriber activity to include the port number. This increased logging requirement will increase the ISPs' data retention costs, but it is assumed that ISPs considering deploying CGNs have weighed up the added cost of logging against the cost savings made by choosing a CGN over moving to IPv6 in the short time.

Less well understood, however, is the widely distributed costs of increased data retention that CGNs will impose on the larger online services and applications sectors. This is explained in detail below.

11.1.2 Data Retention Costs for Online Service and Application Providers

Even if only one CGN is deployed in the world - not necessarily even in the UK - to ensure that any malicious activity can be traced back to its source, *all* online services in the world will have to adapt their own network operational practices to cater for the lack of end-to-end transparency that CGNs create. Clearly, the more CGNs in the world, the more likely it is that other actors will need to adapt their data retention systems to cope. As shown in Section 4.9 above, to provide law enforcement agencies with the minimum data needed to track illegal activities, in addition to recording IP address and time of access for all connections to their services, web services such as Amazon.co.uk will also need to record the port number associated with that IP address and time of access.

The reason that all online services will need to make this change is that a service cannot be sure whether any particular connection is coming from a normal non-CGN network - where only the IP address and time of connection need to be recorded - or from a CGN network - where the port number must also be recorded. If online services were not to adapt their logging processes to include the port number, CGN networks would most likely become a major source of malicious activity, as hackers would be aware that without online services recording the port number in use at the time, there would be no way for the source of the malicious activity to be traced behind the CGN.[125] The more CGNs are deployed, and online services do not record port number for each connection made to their services, then those CGNs will increasingly become attractive locations for hackers and other perpetrators of illegal activities to house or hide their activities.

[125] Without a port number, all the information that an ISP could provide to law enforcement is the total number of users of that IP address at that particular time. Given the large number of port numbers that can be bound with a single IP address at any given moment, if an affected web service only recorded the IP address and not the port number, the best a CGN-based ISP could do would be to provide law enforcement with a list of perhaps 20,000 users or more of that IP address at the specified time.

11.1.3 Existing Data Logging Issues Further Complicated by CGNs

Anecdotal evidence in the industry is that many providers may already be failing to comply with existing data retention obligations. CGN deployment is likely to make that failure more visible, as the number of parameters than need to be stored increase. As documented earlier in this report, mobile providers say that they are unable to keep adequate logs behind CGN.

As discussed above, there are known tensions between data protection and data retention obligations. Decisions of EU Member States potentially threaten the viability of current data retention obligations, decreasing the data available to law enforcement. For example, the cases of Romania, Germany and Austria, and the referral by Austria of data retention questions to the ECJ.[126] The SABAM decisions also raise questions as to whether increasingly intrusive monitoring of entire customer bases can be sustainable.

11.1.4 The Changing Role of the Internet Service Provider

CGN heightens the control by a single provider over every aspect of the private network behind the node. At the same time, the ISP industry has lobbied hard against legislative provisions (for example, in the *Digital Economy Act 2010*) which compel Internet carriers to take more responsibility over content on their networks. The industry has expressed concerns that such obligations will involve complex questions, and assessing the relative rights of their customers on the one hand, as against the rights of third parties or law enforcement on the other which are often finely balanced, and can raise concerns over lack of due process.

The technical analysis describes the increased opacity of CGN networks, and the difficulties associated with identifying the sources of “bad” traffic either into or out of the network. The only party with the relevant knowledge or control in a CGN scenario will be the ISP and this will likely increase the frequency and complexity of requests from law enforcement or third parties, driving up costs for ISPs and potentially exposing them to legal liability arising from such decisions.

[126] It should also be noted that in The Netherlands, logging is explicitly not allowed on privacy grounds.

12 Impact on Consumers

In 2012, a group of ISPs signed a voluntary code of practice in support of the Open Internet. One of the commitments was to provide "greater transparency in instances where certain classes of legal content, applications and/or services are unavailable on a product."[127] The obligation could therefore extend to CGN networks, which impair certain functionality (such as mapping, or geolocation).

However, given not only the small pool of ISPs that most consumers can choose from, but also the fact that consumers can be overwhelmed by too much information, this approach is not likely to be an effective way to have consumers choose between CGN-operating ISPs and IPv6-deploying ISPs. Parallels can be drawn with consent models (for example, privacy policies) where:

> *"there is increasing evidence from behavioural economics that a consent model has significant failings. Very few users have the time or legal training to fully read and understand privacy policies, let alone enforce them."*[128]

As mentioned in Section 9.4.4, Ofcom has recently published a paper on the issue of information provision and its impact on informed consumer choice. While the paper describes many ways that information can be better presented for consumers, it also acknowledges that in complex markets can result in "option swamps" where some market players benefit when confused customers choose "significantly sub-optimal options that are more expensive than they need."[129] In the context of CGN networks, this suggests that not only could customers end up with a sub-optimal service that does not meet their needs, they could also end up choosing a package—such as a public static IPv4 address—that is beyond their needs.

In addition to informed choice issues, the technical analysis indicates that in certain cases, CGN could severely impact the availability of services to consumers, raising the possibility of liability for carriers under consumer protection legislation for rendering substantially different contractual performance to that expected.[130]

[127] Broadband Stakeholder Group, I*SPs launch Open Internet Code of Practice*, http://www.broadbanduk.org/2012/07/25/isps-launch-open-internet-code-of-practice, 2012, accessed 15 April 2013.

[128] I. Brown, C. Marsden, *Regulating Code: Good Governance and Better Regulation in the Information Age*, MIT, Cambridge (MA), 2013, p. 54.

[129] Ofcom, *A Review of Consumer Information Remedies*, http://stakeholders.ofcom.org.uk/binaries/research/research-publications/information-remedies.pdf, 2013, accessed 22 April 2013, p. 43.

[130] *Unfair Contract Terms Act 1977*, section 3(2).

13 Policy Conclusions

To date, CGN has not been widely deployed in the UK or other environments. There is little specific literature on CGN policy issues. Necessarily, conclusions at this stage are based on numerous assumptions and hypothetical scenarios.

The Internet's journey to date has been one of constant innovation and disruption. On one level, CGN may represent another innovation in Internet routing, and criticism from within the established community can be expected, as with any disruptive technology.

That said, there is evidence to support the view that CGN has the potential to create adverse competition, security, and privacy issues for consumers and business customers of carriers, for ISPs themselves. Most significantly, adverse impacts are likely for online service providers who not only rely on carriers to provide direct access to potential customers. Moreover, the Internet industry as a whole will likely bear the financial burden of adding port number logging—while only the ISPs deploying CGNs enjoy potential financial savings associated with those CGNs.

Albeit that there are high levels of Internet penetration in the UK, and that current ISPs have good stocks of IPv4 addresses, literature on network theory and network neutrality points to concerns over CGN's erosion of the end-to-end principle and the harms to competition and consumers that could cause. In particular, widespread CGN deployment is likely to distort the market, at both the transport and application levels, by creating walled gardens closed off to innovative applications or new market entrants, and entrenching market power of popular application providers. Semi-commons theory indicates that such scenarios risk a "tragedy of the commons".

CGN poses risks of distortion through the ability to sustain prices above competitive levels, weakening existing competition, raising barriers to entry not only for would-be ISPs but for application developers, and slowing innovation. Most potential distortions arise from the opacity and complexity of CGN compared with the classic "dumb" Internet networks, and the necessity to have ISPs configure CGN gateways to allow individual applications in many cases. Such distortions would only be exacerbated by network effects.

Asian markets and other emerging economies are implementing IPv6, raising the risk that if the UK pursues a policy of supporting CGN, the UK will become out of step with markets targeted by new applications, resulting in detriment to consumers through lack of access to innovative services, and an adverse impact on the competitiveness of the UK economy.

CGN raise privacy, security and intellectual property concerns. CGN increases the vulnerability of big nodes to attack. Some security protocols and services do not work on CGN. CGN also require more complex and costly logging to identify subscribers, leading to potential extra steps or failure in criminal investigations. There are also concerns over the emerging grey market for IPv4 addresses. As ISPs are required to store ever more detailed data to identify the source of traffic, privacy concerns are more likely to come into conflict with data retention obligations. CGN also adversely affect anti-spam and disaster response.

While implementing CGN may save the cost of investment in IPv6, running CGN is not likely to be a cheaper option for ISPs, given the cost of record keeping, and support for customers and third party application providers. CGN will also increase the likelihood of ISPs being made to make difficult decisions over Internet content, something the industry has to date resisted.

Annexes

Annex A: Recommendations Regarding CGN in the Literature

Standards bodies and industry bodies such as the IETF (Internet Engineering Task Force) and BITAG (Broadband Internet Technical Advisory Group) have carried out a lot of work researching CGNs and developing best practice for the deployment of CGNs. Many of these reports and standards include concrete best practice recommendations. For completeness these have been included here.

Implications of Large Scale Network Address Translation (NAT) report

The Implications of Large Scale Network Address Translation (NAT) report[131] includes the following recommendations in its *Executive Summary* (pp. ii-iii). These are included in full below:

- **Commit to rapid deployment of IPv6.** The best way to mitigate the impacts of LSN is to reach a state where IPv6 is the dominant addressing scheme. BITAG suggests that ISPs deploy IPv6, that equipment manufacturers support IPv6 in their devices, and that applications sensitive to NAT be supported via IPv6 as soon as possible.
- **Address application impacts of LSN.** BITAG suggests that vendors of LSN equipment adhere to existing requirements [Common requirements for Carrier Grade NAT (CGN)] intended to increase the likelihood that applications will function properly in the presence of LSN. BITAG also suggests that ISPs test their LSN implementations and mitigate application issues prior to deployment, and that application developers use LSN work--- arounds or avoid deploying services that do not function properly in the presence of NAT or LSN.
- **Disclose LSN deployment.** To assist with end user troubleshooting, BITAG suggests that ISPs be transparent with respect to the locations and timing of LSN deployment.
- **Provide mechanisms to facilitate LSN traversal to end users.** BITAG suggests that, where feasible, ISPs and equipment vendors support mechanisms to facilitate NAT traversal, including mechanisms for the manual or automatic creation and management of port forwarding rules. Such mechanisms increase the likelihood that applications requiring inbound connections to end users can function across LSN.
- **Provide contact information.** BITAG suggests that ISPs provide a means for application providers to contact them to discuss LSN impacts and possible mitigations.
- **Consider Logging Impacts of Port Allocation.** BITAG suggests that ISPs deploying LSN consider logging and operational impacts when deciding whether to implement a deterministic or dynamic mechanism (or a hybrid of the two) for assigning ports to subscriber sessions.
- **Include Port Number When Logging Activity.** BITAG suggests that Application Providers that maintain a log of user activity include both the IPv4 address and port number in the log. This would ensure that logs accurately reflect the actions of a single ISP customer when IPv4 traffic goes across a LSN.

[131] Broadband Internet Technical Advisory Group, *Implications of Large Scale Network Address Translation (NAT)*, www.bitag.org/documents/BITAG_TWG_Report-Large_Scale_NAT.pdf, 2012, accessed 15 April 2013.

IPv6 Policy White Paper

This white paper reflects the opinions of Alain Fiocco who is the Senior IPv6 Programme Manager at Cisco. The report includes a detailed discussion of the impacts of CGN and the following recommendations (p. 12), the first two of which are particularly relevant to CGN deployments:

- Electronic Communication regulators should put forward policy driving ISP to provide unique and global IP addresses to their subscribers, whether IPv4, IPv6 or both.
- Electronic Communication regulators should mandate transparency on IPv4 address sharing policy and enable reporting of impact on end-users Service Quality.
- Government should set deadlines for all central Government Agencies and Public Services web sites and services to be on IPv6 Internet.
- Incentivise regional and local Government/Public service to be on IPv6 Internet.
- Foster Education programs toward Enterprise and Business groups and associations: about the importance of IPv6 deployment in order to continue to innovate and create sustainable differentiation and competitiveness.

RFC 2993, Architectural Implications of NAT

This RFC[132] contains recommendations in the section *Deployment Guidelines*. These are included verbatim below:

- Given that NAT devices are being deployed at a fairly rapid pace, some guidelines are in order. Most of these cautionary in nature and are designed to make sure that the reader fully understands the implications of the use of NATs in their environment.
- Determine the mechanism for name resolution, and ensure the appropriate answer is given for each address administration. Embedding the DNS server, or a DNS ALG in the NAT device will likely be more manageable than trying to synchronize independent DNS systems across administrations.
- Is the NAT configured for static one to one mappings, or will it dynamically manage them? If dynamic, make sure the TTL of the DNS responses is set to 0, and that the clients pay attention to the don't cache notification.
- Will there be a single NAT device, or parallel with multiple paths? If single, consider the impact of a device failure. If multiple, consider how routing on both sides will insure the packets flow through the same box over the connection lifetime of the applications.
- Examine the applications that will need to traverse the NAT and verify their immunity to address changes. If necessary provide an appropriate ALG or establish a VPN to isolate the application from the NAT.
- Determine need for public toward private connections, variability of destinations on the private side, and potential for simultaneous use of public side port numbers. NAPTs increase administration if these apply.
- Determine if the applications traversing the NAPT or RSIP expect all ports from the public IP address to be the same endpoint. Administrative controls to prevent simultaneous access from multiple private hosts will be required if this is the case.

[132] T. Hain, *Architectural Implications of NAT*, RFC 2993, http://tools.ietf.org/html/rfc2993, 2000, accessed 15 April 2013.

- If there are encrypted payloads, the contents cannot be modified unless the NAT is a security endpoint, acting as a gateway between security realms. This precludes end-to-end confidentiality, as the path between the NAT and endpoint is exposed.
- Determine the path for name resolutions. If hosts on the private side of a NAPT or RSIP server need visibility to each other, a private side DNS server may be required.
- If the environment uses secure DNS records, the DNS/ALG will require access to the source authentication keys for all records to be translated.
- When using VPNs over NATs, identify a clearinghouse for the private side addresses to avoid collisions.
- Assure that applications used both internally and externally avoid embedding names, or use globally unique ones.
- When using RSIP, recognise the scope is limited to individual private network connecting to the public Internet. If other NATs are in the path (including web-server load-balancing devices), the advantage of RSIP (end-to-end address/port pair use) is lost.
- For RSIP, determine the probability of TCP_Time_Wait collisions when subsequent private side hosts attempt to contact a recently disconnected public side service.

Annex B: Acknowledgements

- Darren Anstee (Arbor Networks)
- Randy Bush (IIJ)
- Vint Cerf (Google)
- John Curran (ARIN)
- Owen DeLong (Hurricane Electric)
- Ian Dickinson (BSkyB)
- Kelly Dorset (PlusNet)
- Alain Fiocco (Cisco)
- Mat Ford (Internet Society)
- Grant Forsyth (BSkyB)
- Tony Hain (Hain Global Consulting)
- Malcolm Hutty (LINX)
- Adrian Kennard (AAISP)
- Erik Kline (Google)
- Juliet Kramer (EE)
- Andreas Muller (Technische Universitat Munchen)
- Tom Perrine (Sony Computer Entertainment - Playstation)
- Ben Shaw (BSkyB)
- Robert Sleigh (EE)
- Barbara Stark (AT&T)