Study into Websites Sharing Internet Protocol Addresses

Final Report

Prepared for:

Prepared by:

CSMG
Descartes House
8 Gate Street
London WC2A 3HP
United Kingdom
www.csmg-global.com

26 April 2012

ABOUT TMNG GLOBAL

www.tmng.com

TMNG Global (NASDAQ: TMNG) is a leading provider of professional services to the converging communications industry. Its companies, TMNG, CSMG, and Cartesian, and its base of over 500 consultants, have provided strategy, management, and technical consulting, as well as products and services, to more than 1,200 communications service providers, entertainment, media, and technology companies and financial services firms worldwide. The company is headquartered in Overland Park, Kansas, with offices in Boston, London, New Jersey, and Washington, D.C.

ABOUT CSMG

www.csmg-global.com

CSMG, a division of TMNG Global, is a leading strategy consultancy that focuses on the communications, digital media, and technology sectors. CSMG consultants combine a deep understanding of the global communications industry with rigorous analytic techniques to assist their clients in outmanoeuvring the competition. The organization prides itself on understanding the complex technology and financial chain that links the digital economy. CSMG serves its international client base through its offices in Boston and London.

CSMG Boston, Two Financial Center, 60 South Street, Suite 820, Boston, Massachusetts, 02111 USA
Telephone +1 617 999.1000 • Facsimile +1 617 999.1470
CSMG London, Descartes House, 8 Gate Street, London, WC2A 3HP, UK
Telephone +44 20 7643 5550 • Facsimile +44 020 7643 5555

Contact Information	
Michael Dargue *Principal* +44 20 7643 5477 michael.dargue@csmg-global.com	**Peter Wells** *Director of Consultancy* +44 20 7643 5563 peter.wells@cartesian.com

ACKNOWLEDGEMENTS

CSMG would like to acknowledge and thank Benjamin Edelman, Assistant Professor at Harvard Business School for his support in this study.

Table of Contents

List of Figures

1. EXECUTIVE SUMMARY

Introduction

1.1. Ofcom engaged CSMG to estimate the spread of websites that are sharing IP addresses. Specifically, Ofcom asked CSMG to provide an update to a study by Edelman (2003) which estimated that 87% of websites shared IP addresses within active COM, NET and ORG websites[1].

1.2. IP address sharing is of interest, because it makes website blocking technically complex. It is common for multiple websites to share a single IP address. This means that the blocking of an IP address in response to a single unlawful website could impact lawful websites hosted by the same IP address.

Approach

1.3. Our analysis considered a total of 113.3 million distinct websites from four major top-level domains (TLDs) typically used by UK organisations: COM, NET, ORG and UK.

1.4. In order to obtain data for our analysis, we looked up the IP addresses associated with the www host of all known domain names as of 17 February 2012. These domain names were extracted from "zone files" which we obtained from the relevant registry organisations, namely VeriSign for COM and NET, and the Public Interest Registry for ORG. For the UK TLD, Nominet did not provide a zone file, but supported our analysis.

Results and Conclusion

1.5. IP address sharing is prevalent across the three generic top-level domains (gTLDs) studied, namely COM, NET and ORG, and has significantly increased in usage since 2002. For the COM, NET and ORG top-level domains, 97% of websites reside on IP addresses shared with other websites, compared to 87% in 2002.

1.6. IP address sharing is similarly prevalent across the UK TLD. 97% of all website instances reside on shared IP addresses, the same proportion for COM, NET and ORG. Each IP address on average hosts 17.3 website instances for the UK TLD, compared to 20.2 for COM, NET and ORG.

1.7. The extent of IP address sharing remains high, even if we exclude the impact of "domain squatters", where firms hold large number of websites and often 'park' these at a common IP address.

[1] Web Sites Sharing IP Addresses: Prevalence and Significance, Benjamin Edelman, Berkman Center for Internet and Society, Harvard Law School, September 2003
http://cyber.law.harvard.edu/people/edelman/ip-sharing/

2. INTRODUCTION

Objectives and Scope of CSMG's Study

2.1. Ofcom engaged CSMG to estimate the extent of IP address sharing between websites and to provide a written report detailing findings.

2.2. Firstly, CSMG was asked to provide an estimation of the extent of IP addresses that are shared by websites within the COM, ORG and NET top-level domains.

2.3. Secondly, CSMG was asked to provide an estimation of UK-based websites sharing under the UK country-code domains.

2.4. Thirdly, CSMG was asked to provide an update to the Edelman (2003) study including the following elements:

- The methodology used to estimate extent of web sharing IP addresses
- Raw data sources used to conduct the review
- A written report detailing findings and significant/noteworthy observations

2.5. The study was conducted by CSMG in collaboration with its sister company, Cartesian.

Context

2.6. Governments worldwide recognise the challenge of managing content on the Internet that has a detrimental impact on society. In response, governments have attempted to block access to illegal or inappropriate content using a variety of technical means.

2.7. The UK Digital Economy Act 2010 contains provisions for blocking websites dedicated to copyright infringement. The Secretary of State subsequently asked Ofcom to consider the feasibility of technical solutions. Ofcom's response[2] reviewed various blocking techniques which Internet Service Providers (ISPs) could implement and highlighted potential technical barriers and concerns.

2.8. As covered in Ofcom's response, one approach to blocking websites is to block the IP addresses of infringing websites. This would be achieved by requesting Internet Service Providers (ISPs) to configure their network infrastructure to prevent access to the IP addresses in question.

2.9. IP address sharing – also known as IP sharing, virtual hosting or name-based hosting – makes website blocking technically complex. Where multiple websites share a single IP address, the blocking of an IP address in response to a single infringing website would impact non-infringing websites hosted by the same IP address.

[2] "Site Blocking" to reduce online copyright infringement, A review of sections 17 and 18 of the Digital Economy Act, Ofcom, 27 May 2011
http://stakeholders.ofcom.org.uk/binaries/internet/site-blocking.pdf

Summary of Edelman (2003)

2.10. Edelman (2003) provided an analysis on the prevalence of IP address sharing and discussed the potential impact of IP-based filtering on the Internet.

2.11. Edelman's analysis considered 20.1 million distinct websites from three gTLDs – COM, NET and ORG – as of December 2002. The analysis uses raw data on IP addresses associated with the default website of each domain. Edelman defined default websites as those prefixed with www. For example, the default website within the ofcom.org.uk domain is www.ofcom.org.uk.

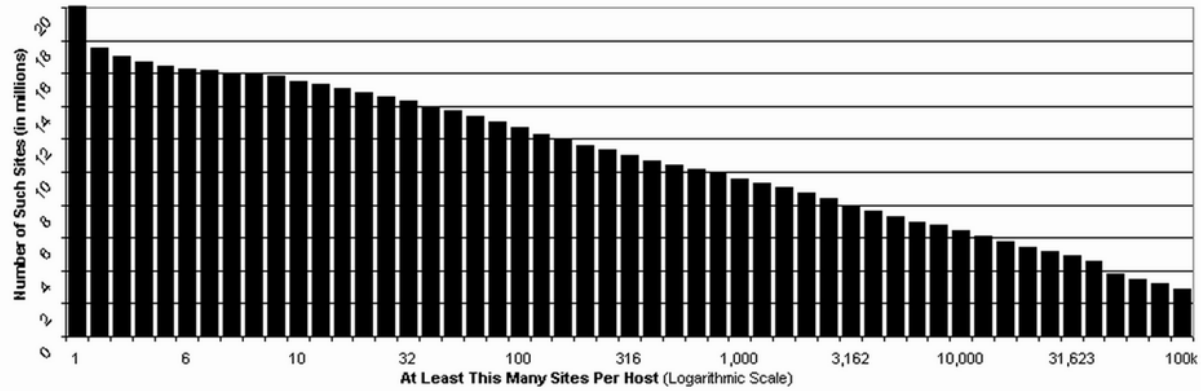
2.12. Edelman's analysis indicated that 87% of websites were hosted on shared IP addresses and therefore concluded that the use of IP filtering was likely to lead to over-blocking whereby acceptable websites hosted alongside websites deemed unacceptable on the same IP addresses would be blocked as well.

2.13. Edelman's key findings are as follows:

Figure 1: Statistics on Websites Sharing IP Addresses, Dec 2002

	Number of Websites (to nearest 100,000)	Proportion of Total Websites Considered
Total distinct websites considered	20.1m	100%
Distinct websites residing on unshared hosts	2.5m	13%
Distinct websites residing on shared hosts, hosting 2 websites or more	17.6m	87%
Distinct websites residing on shared hosts, hosting 5 websites or more	16.5m	82%
Distinct websites residing on shared hosts, hosting 20 websites or more	15.1m	75%
Distinct websites residing on shared hosts, hosting 50 websites or more	14.0m	70%
Number of distinct websites on most shared host	1.0m	5%

Figure 2: COM, ORG and NET Domains: Number of Websites on IP Addresses Sharing at Least a Specific Number of Websites, Dec 2002



3. APPROACH

Design Principles

3.1. To obtain a list of domain names for evaluation, we sought access to "zone files" which we obtained from the relevant registry organisations, namely VeriSign for COM and NET, and the Public Interest Registry for ORG. For the UK TLD, Nominet did not provide a zone file, but supported our analysis as required. We provide a description of how Nominet prepared the data in Section 3.20.

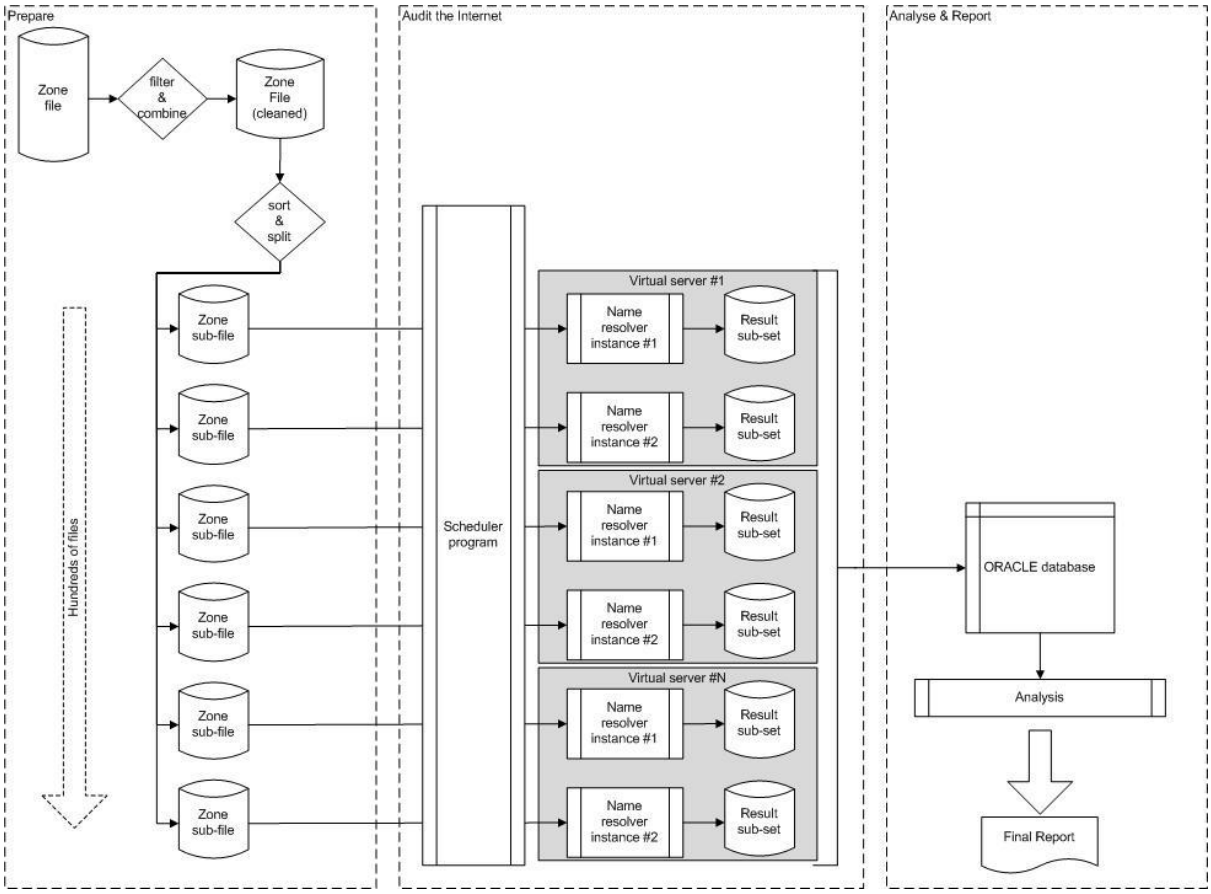
3.2. The Domain Name System has a hierarchical structure. The TLD zone files list the domains within each TLD and map these to their authoritative name servers. The authoritative name servers in turn have their own zone files which, for each domain, map hostnames to IP addresses.

3.3. For more information on how name servers work and the Domain Name System (DNS) more generally, refer to the Annex in Section 6.

3.4. Given the large number of domains (in the order of 100 million), it was necessary to divide the zone files into manageable segments for concurrent processing on multiple servers.

3.5. The diagram below shows the stages of processing, from preparation of the input data, through the actual gathering of IP data, to the analysis stage.

Figure 3: Data Preparation, Audit and Analysis



3.6. The scope of the data was specified as follows:

- Only COM, NET, ORG and UK top-level domains were considered
- The analysis was conducted only against hosts with the www prefix
- When multiple IP addresses were returned by the lookup for a single www host, each IP address was recorded as an individual entry in the results file

Data Cleansing

3.7. Before querying the name servers for IP addresses, the zone files were cleansed according to the following steps:

- Retrieve the compressed COM, ORG and NET zone files from the TLD registries
- Verify the integrity of files using the checksum information from the registries
- Uncompress the zone files
- Cleanse the data by
  - Extracting the “NS“ records for each domain

  - Removing the “DNSSEC” records
  - Removing records with no name server specified
  - Append .COM, .NET, and .ORG suffix to name server records where missing
- Filter out duplicate entries
- Re-format the zone files by condensing multiple records into one per domain

  From

  <domain>,<name server1>

  <domain>,<name server2>

  <domain>,<name server3>

  . . .

  <domain>,<name serverN>

  To

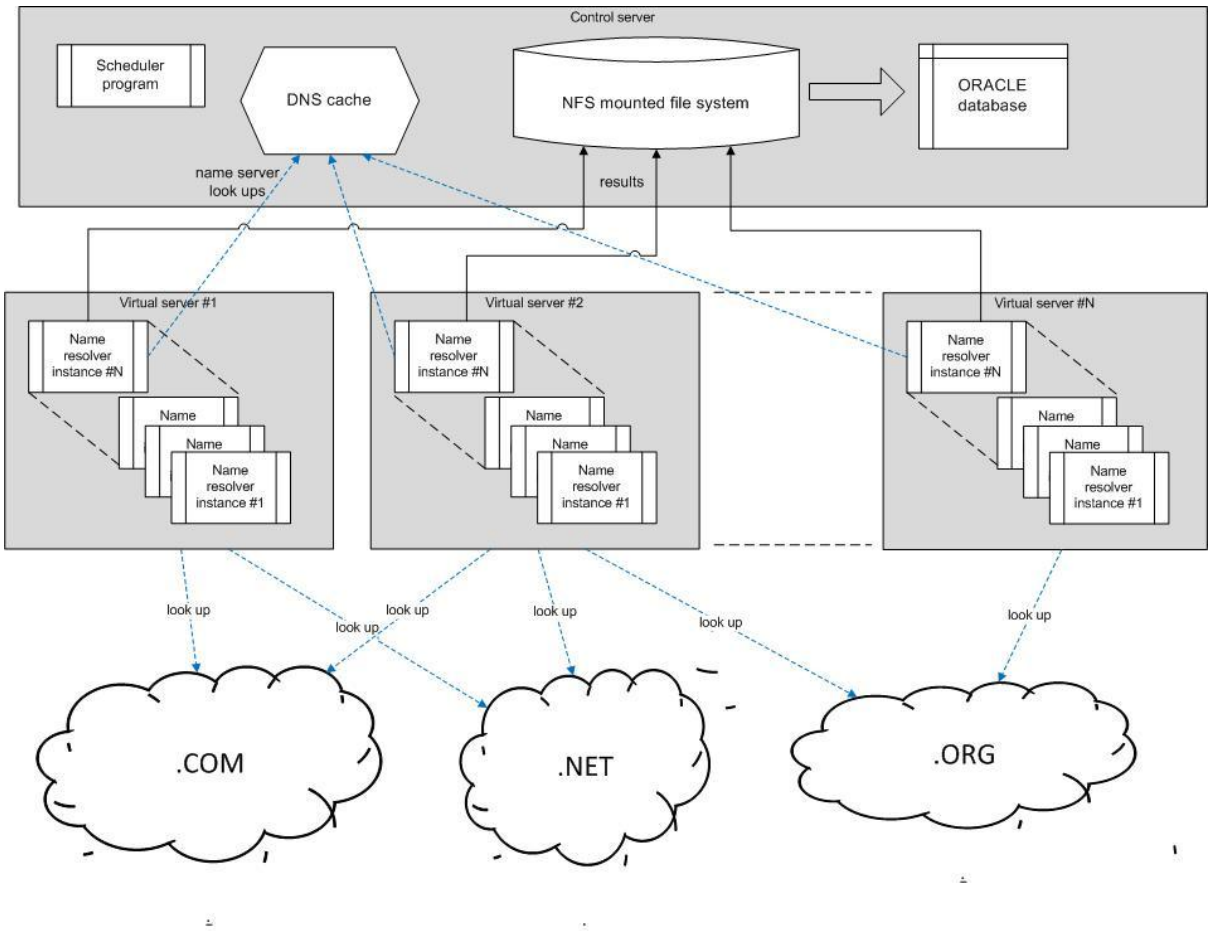
  <domain>,<name server N>, . . . ,<name server 1>
- Split the files into manageable segments for the multi-threaded look program
- Upload files to server infrastructure

Infrastructure

3.8. Temporary servers were used for the project. One of the servers was configured with a local DNS cache in order to minimise repetitive requests to name servers over the Internet.

3.9. The infrastructure used for the data gathering exercise is shown below.

Figure 4: Data Gathering Infrastructure



‘Stale data’ and Errors

3.10. The data contained in the zone files represents a constantly changing environment. Consequently, by the time the data is acquired, cleansed and passed through the lookup program, a proportion of it became ‘stale’ (i.e. out of date); these are domains which have either been removed, or had their authoritative name servers changed.

3.11. The zone files for the COM, NET and ORG top-level domains, were downloaded separately on 17th February 2012. Results from the queries of all domains in each TLD (i.e. ‘first run’) were available after several weeks.

Figure 5: Name Server Query Dates (First Run)[3]

Top-Level Domain (TLD)	Date of Zone File Download	End Date of Query of All Domains in TLD	Duration of Query of All Domains In TLD
COM	17/02/2012	05/03/2012	5 days
NET	17/02/2012	01/03/2012	1 day
ORG	17/02/2012	28/02/2012	1 day

3.12. www hostnames which did not resolve on the first run against their authoritative name server, were re-queried with a full recursive look-upon 6th and 7th March 2012. Of the 20.3 million hostnames which did not resolve on the first run and were subsequently re-queried, 8.6 million resolved and 11.7 million remained as errors.

Figure 6: Re-query of Domains Unresolved on First Run (i.e. Second Run)

Top-Level Domain (TLD)	Domains Which Successfully Resolved on the First Run	Domains Which Subsequently Resolved on the Second Run	Domains Which Remained as Errors on the Second Run
COM	91,833,143	7,189,779	8,834,944
NET	12,723,418	838,469	1,835,480
ORG	8,779,512	589,660	1,028,850

3.13. In addition, a statistically insignificant number DNS lookups did not return correctly-formed IP addresses, mostly due to lookup errors. 1871 domains for the COM, NET and ORG top-level domains were excluded for this reason.

Final Data: COM, NET and ORG

3.14. For the COM, NET and ORG top-level domains, the final data for analysis comprised 6.0 million distinct IP addresses and 113.3 million distinct websites, excluding all errors.

3.15. In addition to multiple websites sharing single IP addresses, there are also single websites that are hosted at multiple IP addresses. The latter is common for websites which need multiple servers to manage large volumes of internet traffic or seek greater resilience. We use the term “website instances” to refer to unique occurrences of the same website. For

[3] Between successfully downloading the zone files and processing the “first run” of queries, we used the time to set up data collection environments and run performance tests.

example, if a single website is hosted by 5 IP addresses, this would be counted as 5 "website instances"[4].

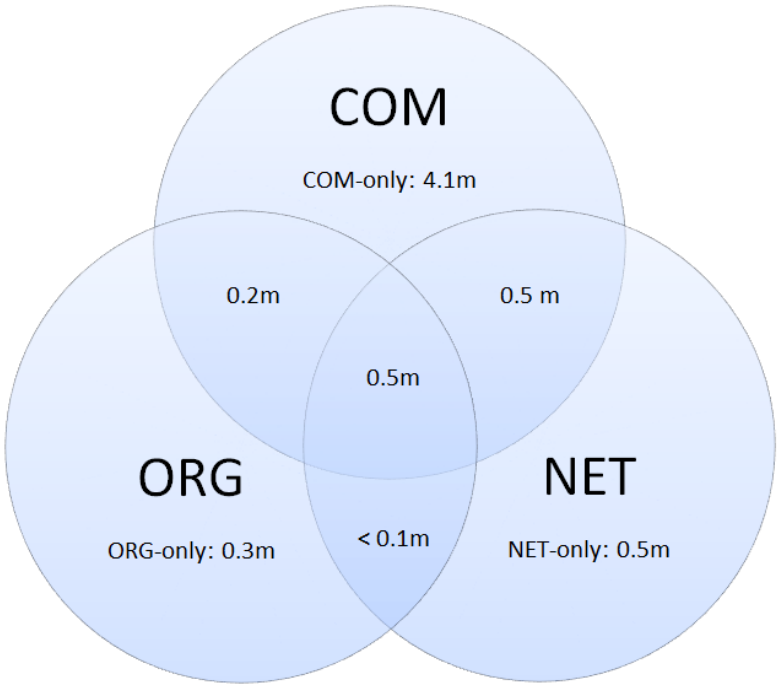
3.16. Each website on average has 1.08 IP addresses. That is, the total number of "website instances" is 8% higher than the total number of distinct websites.

3.17. Thus, due to some websites hosted on multiple IP addresses, our estimates of the extent of IP address sharing are based on "website instances".

3.18. It should also be noted that our estimates are likely to understate the extent of IP address sharing, as our analysis does not fully explore the extent to which websites within other TLDs share common IP addresses. In particular, we consider only four TLDs in total; adding additional top-level domains such EU and INFO would further increase the proportion of IP addresses hosting two or more web sites.

3.19. The Venn diagram below illustrates the extent of IP addresses hosting website instances from multiple top-level domains (across COM, NET and ORG). Out of 6.0 million distinct IP addresses considered, 1.2 million IP addresses or 20% (i.e. sum of intersections) each hosted website instances from more than one TLD.

Figure 7: IP Addresses Hosting Website Instances in COM, NET, ORG, and Combinations



[4] Edelman (2003) used first responses to each of the DNS lookups which yielded multiple responses, meaning Edelman made the simplifying assumption that each website was hosted by one IP address only.

Approach and Final Data: UK

3.20. Similar to the approach used for COM, NET and ORG data, Nominet queried the name server listed for each domain in the UK TLD zone file. However, for any queries which failed to return a response, Nominet assumed that all domains sharing the same authoritative server would not resolve, and thus did not seek to resolve those domain names.

3.21. As a result, Nominet's approach was prone to the impact of 'stale data' and temporary server errors. Of 10.3 million distinct websites, only 57% resolved successfully. Although Nominet's approach provides an incomplete sample, we assume it is representative of the UK TLD, since our analysis demonstrates that the distribution of websites for the UK TLD in terms of IP sharing is broadly similar to COM, NET and ORG.

3.22. The final data for analysis for the UK top-level domain comprised 0.3 million distinct IP addresses and 5.9 million distinct websites. Thus, the estimated ratio of websites to IP addresses for the UK top-level domain is 13:1, compared to 19:1 for the COM, NET and ORG top-level domains.

3.23. As expected, there are occurrences of websites which are hosted at multiple IP addresses. Each website on average has 1.09 IP addresses. That is, the total number of "website instances" is 9% higher than the total number of distinct websites.

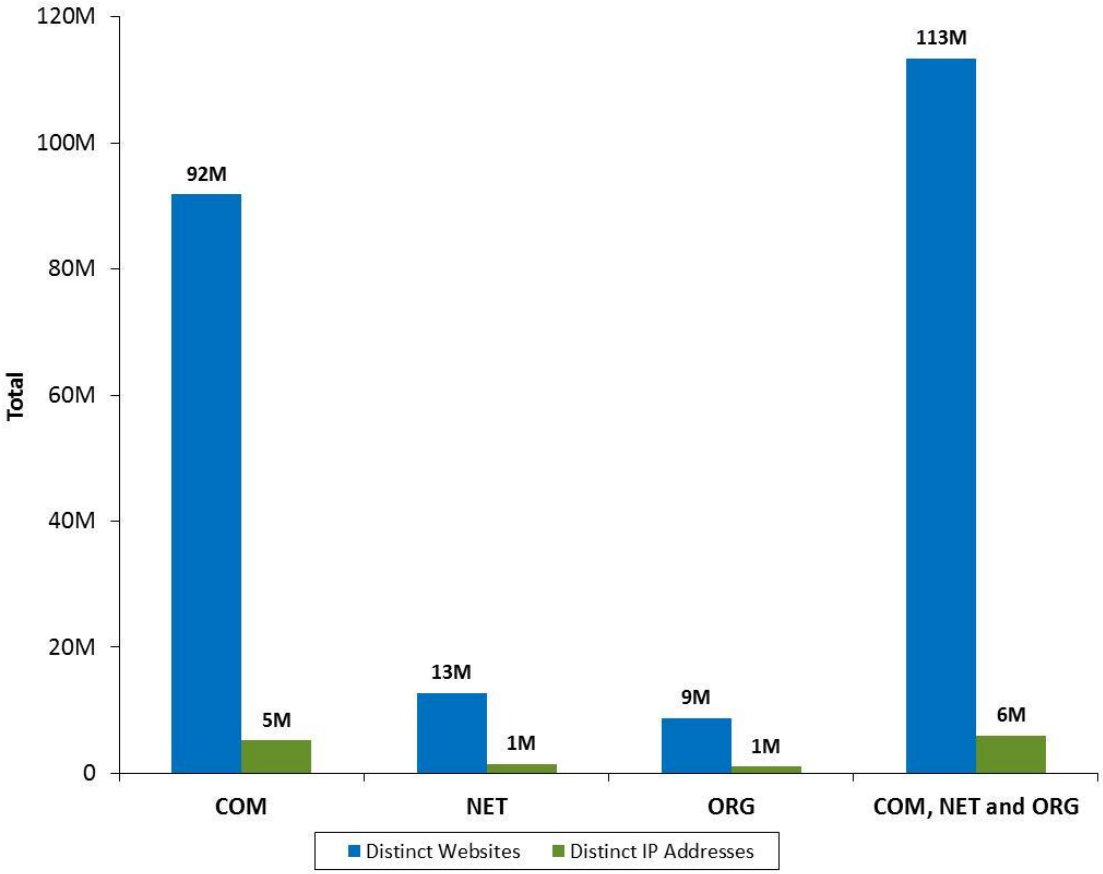
4. RESULTS

4.1. In this section, we review results for the 3 major TLDs, namely COM, NET and ORG, and make a comparison with Edelman (2003). We then review results separately for the UK TLD.

COM, NET and ORG Top-level Domains

4.2. As described, we found 6.0 million distinct IP addresses and 113.3 million distinct websites. The fact that websites significantly outnumber IP addresses immediately suggests significant sharing of IP addresses between websites. The gap between the number of IP addresses and distinct websites can be observed across the three gTLDs, as shown by the diagram below. The difference is particularly pronounced for COM domain names.

Figure 8: Total Number of Distinct Websites and IP Addresses

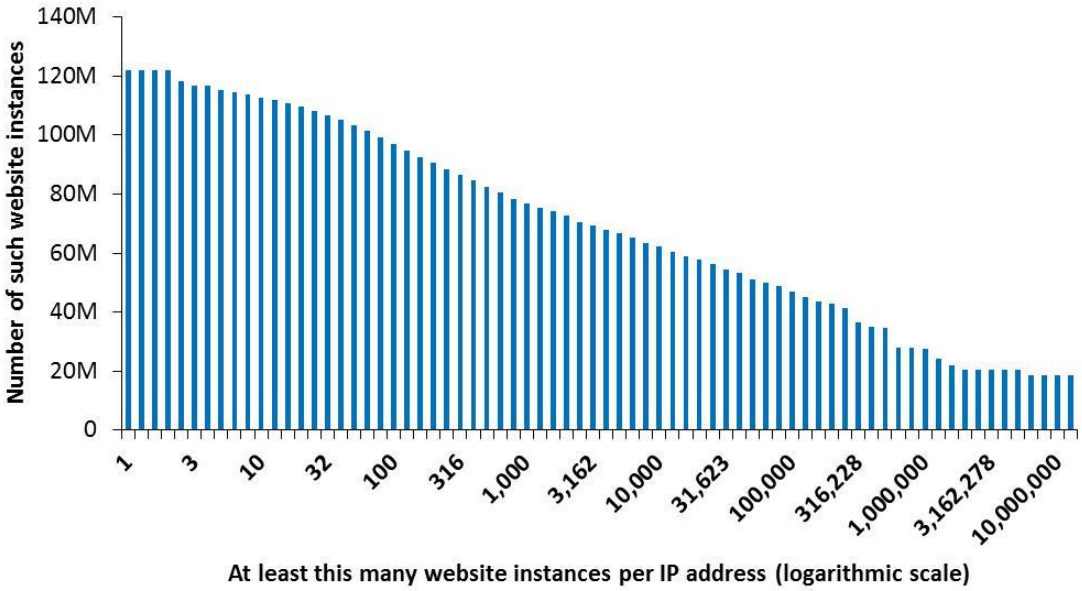


4.3. Our results confirm that IP address sharing is prevalent across these three gTLDs. For the COM, NET and ORG top-level domains, only 3% of website instances reside on IP addresses which are not shared with other websites. This means 97% of all website instances reside on shared IP addresses. Each IP address on average hosts 20.2 website instances.

4.4. In NET and ORG, the proportions of website instances on unshared IP addresses are 7% each, slightly higher compared to 3% for COM. Nevertheless, the vast majority of NET and ORG websites are hosted on shared IP addresses. Each IP address on average hosts 18.9 website instances for COM, compared to 9.4 and 8.9 for NET and ORG respectively.

4.5. The diagram below shows that there are many occurrences of websites sharing IP addresses in both small and large groups. For example, 5.0% of website instances are hosted on IP addresses hosting between 2 and 10 website instances inclusive. More than 50% of website instances are hosted on IP addresses hosting at least 11,580 website instances.

Figure 9: COM, NET and ORG: Number of Website Instances on IP Addresses Sharing at Least a Specific Number of Website Instances



4.6. We note, however, that the results are skewed by IP addresses which host significantly large volumes of websites. This can be observed if we view the results in terms of IP addresses instead of websites. For the COM, NET and ORG top-level domains, the 3% of website instances which reside on unshared IP addresses are served by 64% of distinct IP addresses.

Figure 10: COM, NET and ORG: Number of IP Addresses Sharing a Specific Number of Website Instances

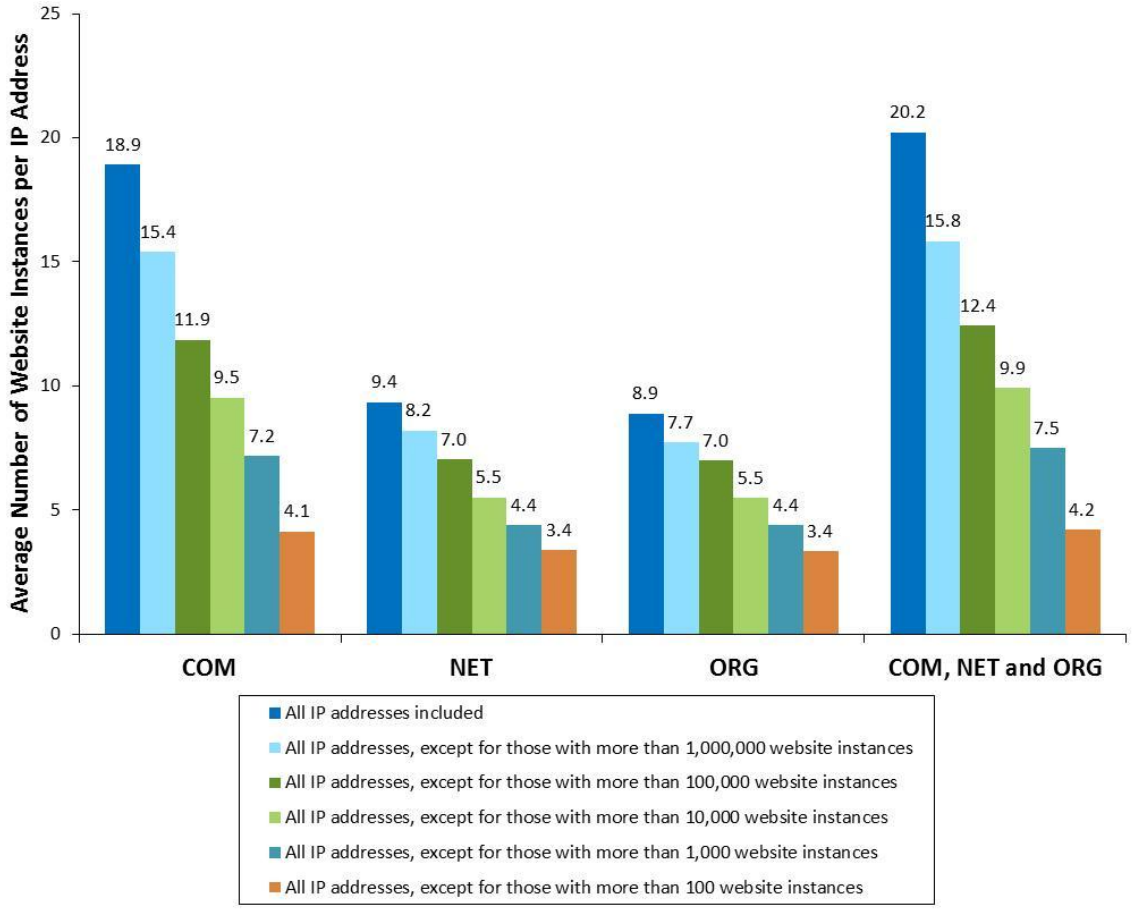
Website Instances per IP Address	Number of IP Addresses	% of Total Distinct IP Addresses
1	3,838,205	63.6%
2	682,984	11.3%
3	319,734	5.3%
4	187,505	3.1%
5	128,946	2.1%
6	95,466	1.6%
7	70,920	1.2%
8	56,653	0.9%
9	46,993	0.8%
10	39,590	0.7%
11-100	478,716	7.9%
>100	87,975	1.5%

4.7. IP addresses which host significantly large numbers of websites are typically examples of the "domain squatting" practice whereby a single organisation occupies multiple unused domains (i.e. websites) for the purpose of resale at a profit. For instance, an invented domain could resolve to a holding page, which serves that domain as well as hundreds of others.

4.8. "Domain squatting" has an upward impact on estimates of IP address sharing. To understand this, we can exclude IP addresses which host large numbers of websites. This should be seen as an approximate method, as it excludes IP addresses used in the web hosting industry to host large numbers of small websites. It also excludes IP addresses used to host large number of websites which are "under construction" (not yet operational), and IP addresses which are used for forwarding/redirecting to another destination. For the latter, the blocking of IP addresses used for forwarding/redirecting would block access to the destination IP addresses and thus lead to the actual content being inaccessible by domain name.

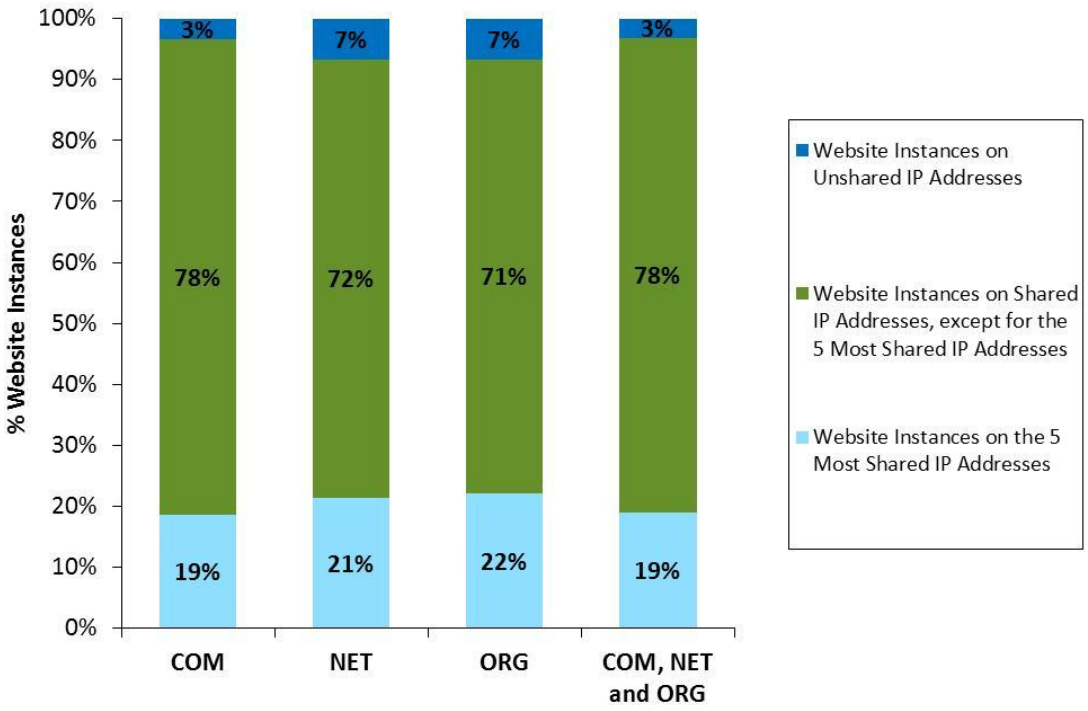
4.9. As shown below, discarding the largest IP addresses reduces the perceived extent of IP address sharing, but the estimates remain large. For the COM, NET and ORG top-level domains, if we discard IP addresses hosting more than 1,000 website instances, the average number of website instances sharing a single IP address reduces from 20.2 to 7.5. The proportion of website instances hosted on IP addresses with two or more website instances reduces from 97% to 92%.

Figure 11: Impact of IP Addresses Hosting Large Numbers of Website Instances



4.10. For the COM, NET and ORG top-level domains, the 5 IP addresses which host the most website instances contribute 19% of all website instances. This proportion is similar across COM, NET and ORG individually, as shown below.

Figure 12: Website Instances on the 5 Most Shared IP Addresses



4.11. The top 3 most-shared IP addresses are owned by Go Daddy[5], the largest Internet domain registrar. These websites are likely to be small in size or awaiting sale or development. The single most-shared IP address (68.178.232.100) hosts 12.8 million website instances comprising 9.9 million COM website instances, 1.7 million NET and 1.2 million ORG.

4.12. For the results divided into individual TLD categories and in table form, refer to the Annex in Section 7.

Comparison with Edelman (2003)

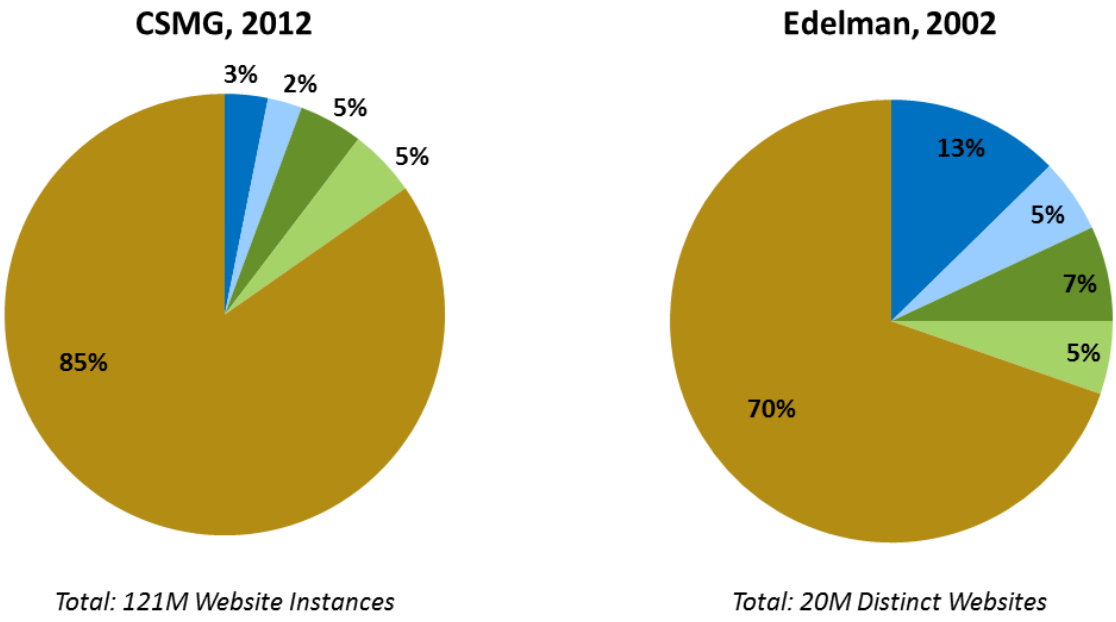
4.13. Edelman (2003) uses data from December 2002. Our analysis (entitled CSMG) uses more recent data from February 2012. Edelman (2003) used first responses to each of the DNS lookups which yielded multiple responses, meaning Edelman made the simplifying assumption that each website was hosted by one IP address only. Edelman's results therefore refer to "distinct websites" instead of "website instances".

4.14. Since 2002, the Internet has grown from 20 million distinct websites to 113 million distinct websites for the COM, NET and ORG top-level domains. The common practice of IP address sharing has persisted and become more prevalent. In 2002, Edelman estimated that 13% of websites resided on unshared IP addresses for COM, NET and ORG. In 2012, from our

[5] Go Daddy manages over 52 million domain names (www.godaddy.com, 5 April 2012)

analysis, we estimate that this has fallen to 3%. More websites also reside on IP addresses hosting a greater number of websites. For example, we estimate that in 2012, 85% of website instances reside on IP addresses with more than 50 website instances, compared to 70% in 2002.

Figure 13: Website Instances Sharing IP Addresses for COM, NET and ORG: 2012 vs. 2002



Note: 'Distinct websites' were used instead of 'website instances' in Edelman (2003)

Figure 14: Comparison of Edelman (2003) and CSMG (2012)

	Edelman (2003)	CSMG (2012)
Date of dataset	Dec 2002	Feb 2012
Total distinct websites instances considered	20m	122m
Website instances residing on unshared hosts	13%	3%
Website instances residing on shared hosts, hosting 2 website instances or more	87%	97%
Website instances residing on shared hosts, hosting 5 website instances or more	82%	94%
Website instances residing on shared hosts, hosting 20 website instances or more	75%	90%
Website instances residing on shared hosts, hosting 50 website instances or more	70%	85%
Number of website instances on the most shared host (as % of total distinct website instances)	5%	10%

UK Top-level Domain

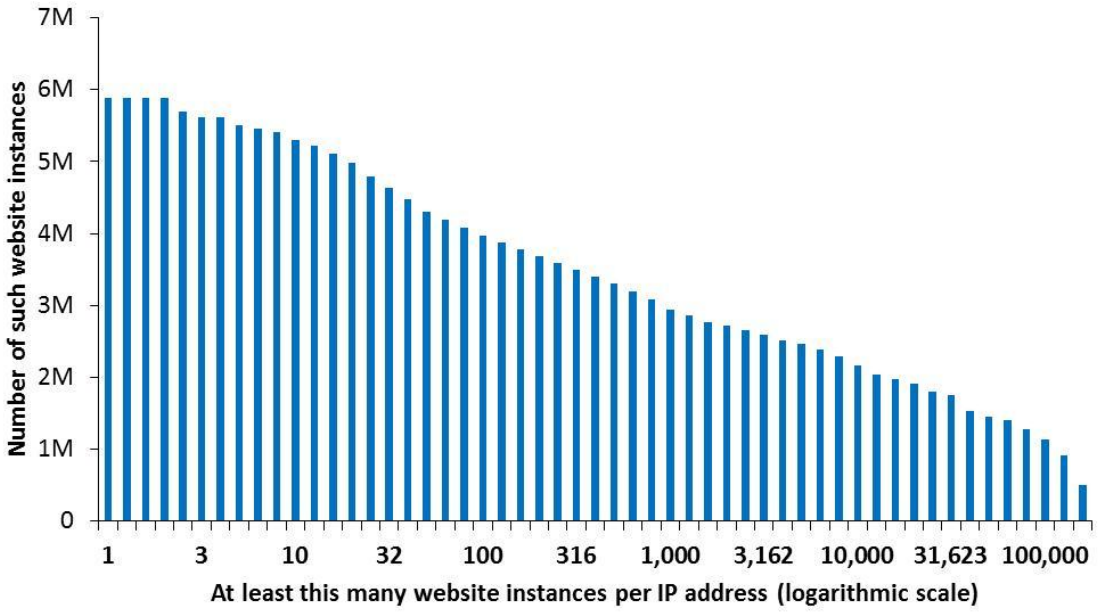
4.15. The results for the UK top-level domain are broadly similar to COM, NET and ORG, demonstrating that IP address sharing is prevalent.

4.16. As described, we found 0.3 million distinct IP addresses and 5.9 million distinct websites. The ratio of websites to IP addresses for the UK top-level domain is therefore 13:1, compared to 19:1 for the COM, NET and ORG top-level domains. The lower ratio immediately suggests that IP address sharing between websites is less significant for the UK top-level domain.

4.17. Our results confirm that IP address sharing is prevalent for the UK TLD; 97% of all website instances reside on shared IP addresses, the same proportion as for COM, NET and ORG. Each IP address on average hosts 17.3 website instances, compared to 20.2 for COM, NET and ORG. This confirms that IP address sharing is slightly less pronounced for the UK TLD.

4.18. The diagram below shows that there are many occurrences of websites sharing IP addresses in both small and large groups. For example, 7.5% of website instances are hosted on IP addresses hosting between 2 and 10 website instances inclusive. More than 50% of website instances are hosted on IP addresses hosting at least 1,036 website instances.

Figure 15: UK: Number of Website Instances on IP Addresses Sharing at Least a Specific Number of Website Instances



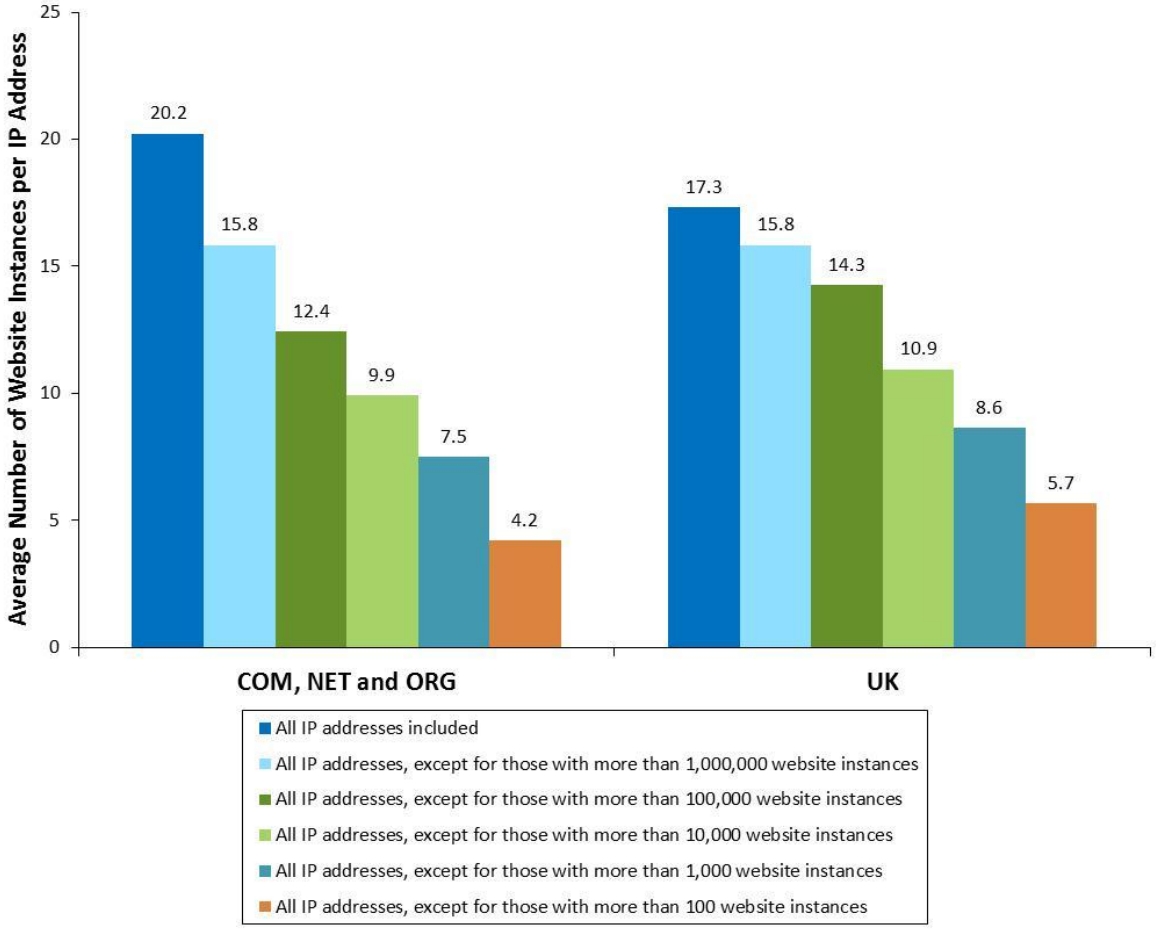
4.19. The results are skewed by IP addresses which host significantly large volumes of websites. This can be observed if we view the results in terms of IP addresses instead of websites. The 3% of website instances in the UK TLD which reside on unshared IP addresses are served by 53% of distinct IP addresses.

Figure 16: UK: Number of IP Addresses Sharing a Specific Number of Websites

Websites Instances per IP Address	Number of IP Addresses	% of Total Distinct IP Addresses
1	178,090	52.5%
2	40,920	12.1%
3	20,210	6.0%
4	12,907	3.8%
5	9,152	2.7%
6	7,173	2.1%
7	5,773	1.7%
8	4,913	1.4%
9	4,281	1.3%
10	3,910	1.2%
11-100	47,997	14.1%
>100	4,202	1.2%

4.20. Discarding the largest IP addresses reduces the perceived extent of IP address sharing, but the estimates remain large. If we discard IP addresses hosting more than 1,000 website instances, the average number of website instances sharing a single IP address reduces from 17.3 to 8.6. The proportion of website instances hosted on IP addresses with two or more website instances reduces from 97% to 94%.

Figure 17: Impact of IP Addresses Hosting Large Numbers of Website Instances



4.21. The 5 IP addresses which host the most website instances contribute 14% of all website instances. This proportion is lower than each of the top 5 IP addresses relevant to COM, NET and ORG which contribute 19%, 21% and 22% to websites instances respectively.

4.22. For the results in table form and additional details, refer to the Annex in Section 7.

5. CONCLUSION

5.1. IP address sharing is prevalent across the four major TLDs typically used by UK organisations, namely COM, NET, ORG and UK, and has significantly increased in usage since 2002. As described, for the COM, NET and ORG top-level domains, 97% of websites reside on IP addresses shared with other websites, compared to 87% in 2002. This has important implications for the use of IP address blocking to prevent access to unlawful Internet content. In particular, it is likely that some unlawful websites and lawful websites may share the same IP address.

5.2. Our analysis shows that websites are increasingly being hosted by IP addresses serving large numbers of website instances. For example, for the COM, NET and ORG top-level domains, more than 50% of website instances are now estimated to be on IP addresses hosting at least 11,580 website instances. For the UK TLD, more than 50% of website instances are hosted on IP addresses hosting at least 1,036 website instances. This is lower, mainly due to significantly fewer websites in the UK TLD space.

5.3. For the COM, NET and ORG top-level domains, the 3% of website instances which reside on unshared IP addresses are served by 64% of distinct IP addresses. This implies that for any given IP address, if the IP address hosts at least one web site, there is a 64% chance that it will only host a single website. We have not explored any qualitative differences between domains which are hosted on single IP addresses versus shared.

5.4. Excluding IP addresses which host large numbers of websites – IP addresses which are likely to be used for "domain squatting" or parking – reduces the prevalence of IP address sharing. However, we demonstrate that estimates IP address sharing remains prevalent even after that exclusion. For the COM, NET and ORG top-level domains, if we discard IP addresses hosting more than 1,000 sites, the average number of website instances sharing a single IP address is 7.5. For the UK TLD, the same average is 8.6.

ANNEX

6. Domain Name System (DNS) Annex

6.1. The Domain Name System (DNS) is a critical component of the global internet and hence UK electronic communication services. DNS provides a means of translating a human-readable domain name into a machine-readable IP address, a process termed "resolving" an address. DNS is fundamental to the operation of the World Wide Web and many other internet services.

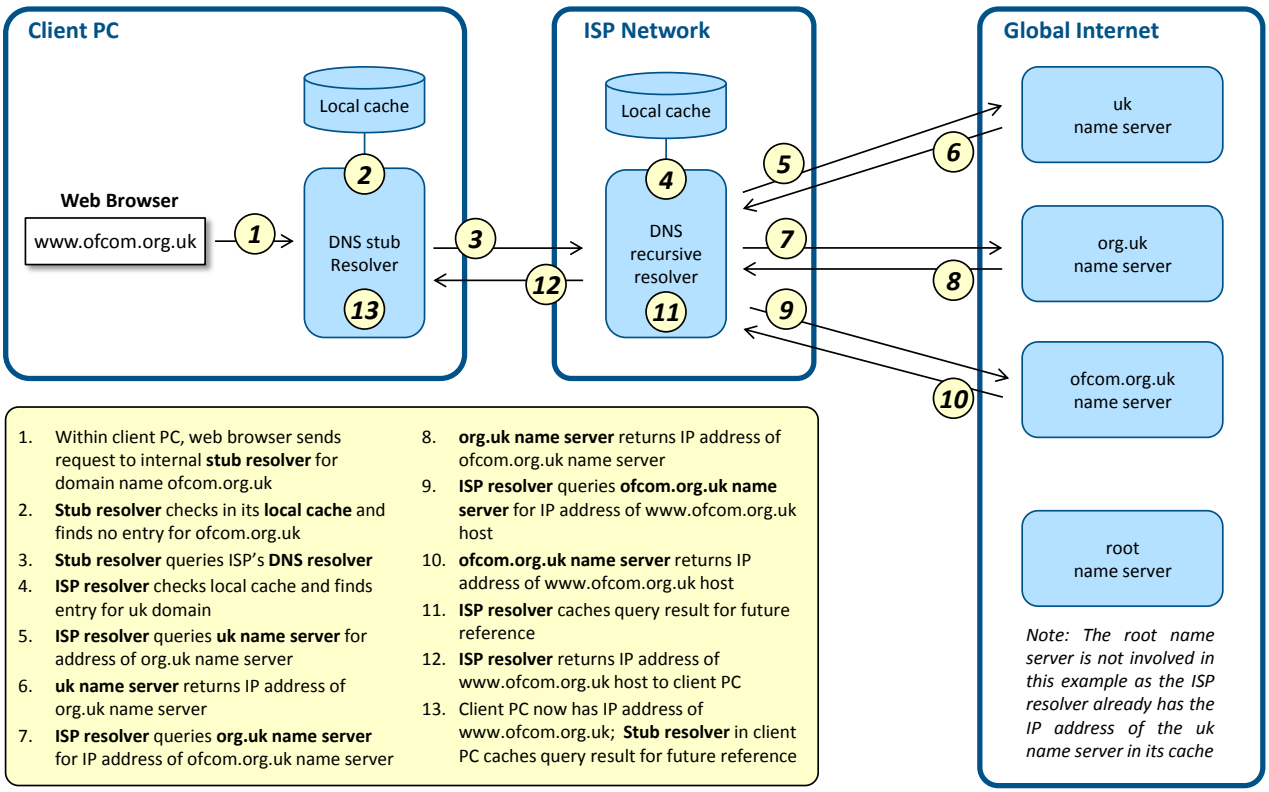
6.2. DNS is a globally distributed system which is organised as a hierarchy of interconnected servers.

6.3. Domain names are also organised hierarchically. For example: "ofcom.org.uk" is a subdomain of "org.uk" which itself is a subdomain of the top-level domain (TLD) for our country, "uk". DNS enables delegation of sub-zones such that one name server can hold the authoritative record for the uk TLD, another can hold the record for org.uk and a third to hold the record for the ofcom.org.uk domain. The architecture enables scalability, whilst permitting delegated control down to an individual domain or host.

6.4. For each domain, at least one name server will be designated as the authoritative name server, i.e. holding the definitive record for that domain.

6.5. To improve efficiency and scalability, the name resolution process employs a combination of recursive lookups and caching (temporary storage of answers to previous queries). The figure below illustrates process of DNS query and response, highlighting the roles of the servers involved.

Figure 18: DNS Query and Response Process



A sample DNS query and response for the www.ofcom.org.uk host is shown below. This figure also illustrates some of the additional information stored within the DNS records besides the IP address.

Figure 19: Sample DNS Query and Response

```
bash$ dig www.ofcom.org.uk any showsearch

; <<>> DiG 9.2.4 <<>> www.ofcom.org.uk any showsearch
;; global options:  printcmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 57574
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 4, ADDITIONAL: 0

;; QUESTION SECTION:
;www.ofcom.org.uk.              IN      ANY

;; ANSWER SECTION:
www.ofcom.org.uk.       300     IN      A       194.33.160.25

;; AUTHORITY SECTION:
ofcom.org.uk.           21599   IN      NS      ns2.netnames.net.
ofcom.org.uk.           21599   IN      NS      ns5.netnames.net.
ofcom.org.uk.           21599   IN      NS      ns6.netnames.net.
ofcom.org.uk.           21599   IN      NS      ns1.netnames.net.

;; Query time: 25 msec
;; SERVER: 172.17.202.13#53(172.17.202.13)
;; WHEN: Tue Apr 10 14:35:50 2012
;; MSG SIZE  rcvd: 134

;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NXDOMAIN, id: 33058
;; flags: qr rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 0

;; QUESTION SECTION:
;showsearch.                    IN      A

;; AUTHORITY SECTION:
.                       10800   IN      SOA     a.root-servers.net. nstld.verisign-
grs.com. 2012041000 1800 900 604800 86400

;; Query time: 108 msec
;; SERVER: 172.17.202.13#53(172.17.202.13)
;; WHEN: Tue Apr 10 14:35:50 2012
;; MSG SIZE  rcvd: 103
```

7. TABLES OF RESULTS

Figure 20: Website Instances Hosted by IP Addresses by Top-level Domain Category

Website Instances Hosted by Each IP Address by TLD Category	Number of Distinct IP Addresses
COM only	4,054,416
NET only	456,830
ORG only	318,092
NET and COM only	460,829
NET and ORG only	34,514
ORG and COM only	203,244
COM, NET and ORG	505,762
Total Distinct IP Addresses	6,033,687

Figure 21: IP Addresses Hosting Multiple Website Instances (Absolute Values)

	CSMG (Feb 2012 Dataset)					Edelman (Dec 2002 Dataset) [6]
	COM	**NET**	**ORG**	**COM, NET and ORG**	**UK**	**COM, NET and ORG**
Distinct Websites (Not Website Instances)	91, 833,143	12,723,418	8,779,512	113,336,073	5,406,278	20,113,430
Website Instances	98,910,794	13,635,797	9,437,686	121,984,277	5,878,151	Unknown
Website Instances on Unshared IP Addresses	3,380,908	906,163	636,249	3,838,205	178,090	2,545,326
Website Instances on Shared IP Addresses with 2 or more Website Instances	95,529,886	12,729,634	8,801,437	118,146,072	5,700,061	17,568,104
Website Instances on Shared IP Addresses with 5 or more Website Instances	92,979,041	11,891,694	8,130,318	115,070,882	5,505,963	16,482,601
Website Instances on Shared IP Addresses with 20 or more Website Instances	88,052,990	10,275,110	6,964,564	109,351,635	4,946,222	15,087,978
Website Instances on Shared IP Addresses with 50 or more Website Instances	82,751,921	9,353,255	6,334,959	103,337,626	4,310,702	14,013,627
Website Instances on the Single Most Shared IP Address	9,887,369	1,665,392	1,235,115	12,787,876	183,916	970,412
Website Instances on the 5 Most Shared IP Addresses	18,387,802	2,923,178	2,094,146	23,139,317	806,315	2,320,048

[6] Edelman (2003) used first responses to each of the DNS lookups which yielded multiple responses, meaning Edelman made the simplifying assumption that each website was hosted by one IP address only. Edelman's results therefore refer to "distinct websites" instead of "website instances".

Figure 22: IP Addresses Hosting Multiple Website Instances (% of Total Distinct Website Instances)

	CSMG (Feb 2012 Dataset)					Edelman (Dec 2002 Dataset)[7]
	COM	NET	ORG	COM, NET and ORG	UK	COM, NET and ORG
Website Instances	100%	100%	100%	100%	100%	100%
Website Instances on Unshared IP Addresses	3%	7%	7%	3%	3%	13%
Website Instances on Shared IP Addresses with 2 or more websites	97%	93%	93%	97%	97%	87%
Website Instances on Shared IP Addresses with 5 or more websites	94%	87%	86%	94%	94%	82%
Website Instances on Shared IP Addresses with 20 or more websites	89%	75%	74%	90%	84%	75%
Website Instances on Shared IP Addresses with 50 or more websites	84%	69%	67%	85%	73%	70%
Website Instances on the Single Most Shared IP Address	10%	12%	13%	10%	3%	5%
Websites on the 5 Most Shared IP Addresses	19%	21%	22%	19%	14%	12%

[7] As described, Edelman's results refer to "distinct websites" instead of "website instances".

Figure 23: Average Number of Website Instances Hosted per IP Address

	COM	NET	ORG	COM, NET and ORG	UK
All IP addresses included	18.9	9.4	8.9	20.2	17.3
All IP addresses, except for those with more than 1,000,000 website instances	15.4	8.2	7.7	15.8	14.3
All IP addresses, except for those with more than 100,000 website instances	11.9	7.0	7.0	12.4	14.3
All IP addresses, except for those with more than 10,000 website instances	9.5	5.5	5.5	9.9	10.9
All IP addresses, except for those with more than 1,000 website instances	7.2	4.4	4.4	7.5	8.6
All IP addresses, except for those with more than 100 website instances	4.1	3.4	3.4	4.2	5.7
All IP addresses, except for those with more than 10 website instances	1.7	1.9	1.9	1.8	2.0

Figure 24: Top 5 Most Shared IP Addresses

Rank	Top 5 Most Shared IP Addresses and Their Associated Number of Website Instances				
	COM	**NET**	**ORG**	**COM, NET and ORG**	**UK**
1	68.178.232.100 Go Daddy	68.178.232.100 Go Daddy	68.178.232.100 Go Daddy	68.178.232.100 Go Daddy	81.21.76.62 Webfusion Internet Solutions
	9,887,369 website instances	1,665,392 website instances	1,235,115 website instances	12,787,876 website instances	183,916 website instances
2	64.202.189.170 Go Daddy	64.202.189.170 Go Daddy	64.202.189.170 Go Daddy	64.202.189.170 Go Daddy	194.72.108.26 BT
	4,581,448 website instances	679,476 website instances	495,799 website instances	5,756,723 website instances	181,436 website instances
3	68.178.232.99 Go Daddy	208.91.197.27 Confluence Networks Inc	68.178.232.99 Go Daddy	68.178.232.99 Go Daddy	85.233.160.22 Namesco Hosting
	1,489,284 website instances	215,031 website instances	153,965 website instances	1,856,798 website instances	147,135 website instances
4	208.73.210.29 Oversee.net	68.178.232.99 Go Daddy	208.91.197.27 Confluence Networks Inc	208.73.210.29 Oversee.net	85.233.160.24 Namesco Hosting
	1,299,975 website instances	213,549 website instances	120,422 website instances	1,453,796 website instances	146,915 website instances
5	208.87.35.104 Secure Hosting Ltd	64.95.64.195 Internap Network Services	209.85.225.121 Google Inc	208.87.35.104 Secure Hosting Ltd	85.233.160.23 Namesco Hosting
	1,129,726 website instances	149,730 website instances	88,845 website instances	1,284,124 website instances	146,913 website instances

Figure 25: IP Addresses Hosting a Specific Number of Websites (Absolute Values and Percentage)

Number of Websites Hosted per IP Address	Number of IP Addresses									
	COM	NET	ORG	COM, NET and ORG	UK	COM %	NET %	ORG %	COM, NET and ORG %	UK %
1	3,380,908	906,163	636,249	3,838,205	178,090	64.7%	62.2%	59.9%	63.6%	52.5%
2	574,492	184,977	143,349	682,984	40,920	11.0%	12.7%	13.5%	11.3%	12.1%
3	258,831	85,630	69,939	319,734	20,210	5.0%	5.9%	6.6%	5.3%	6.0%
4	156,342	52,774	43,651	187,505	12,907	3.0%	3.6%	4.1%	3.1%	3.8%
5	107,469	36,300	30,228	128,946	9,152	2.1%	2.5%	2.8%	2.1%	2.7%
6	77,951	27,462	22,291	95,466	7,173	1.5%	1.9%	2.1%	1.6%	2.1%
7	59,249	21,455	16,818	70,920	5,773	1.1%	1.5%	1.6%	1.2%	1.7%
8	47,726	17,239	13,060	56,653	4,913	0.9%	1.2%	1.2%	0.9%	1.4%
9	39,194	14,089	10,585	46,993	4,281	0.8%	1.0%	1.0%	0.8%	1.3%
10	33,278	11,923	8,413	39,590	3,910	0.6%	0.8%	0.8%	0.7%	1.2%
11-100	418,247	93,126	62,074	478,716	47,997	8.0%	6.4%	5.8%	7.9%	14.1%
>100	70,564	6,797	4,955	87,975	4,202	1.4%	0.5%	0.5%	1.5%	1.2%
Total IP Addresses	5,224,251	1,457,935	1,061,612	6,033,687	339,528	100.0%	100.0%	100.0%	100.0%	100.0%

Figure 26: Average Number of IP Addresses per Website (not Website Instance)

	COM	NET	ORG	COM, NET and ORG	UK
Average IP addresses per website	1.08	1.07	1.07	1.08	1.09
Average IP addresses per website, excluding websites hosted by 1 IP address	3.73	3.64	3.91	3.73	2.78

Figure 27: Websites Hosted by a Specific Number of IP Addresses (Absolute Values and Percentage)

Number of IP Addresses Hosting Each Website	Number of Websites (not Website Instances)									
	COM	NET	ORG	COM, NET and ORG	UK	COM %	NET %	ORG %	COM, NET and ORG %	UK %
1	89,238,568	12,378,064	8,553,162	110,169,794	5,141,899	97.2%	97.3%	97.4%	97.2%	95.1%
2	1,020,001	155,088	87,456	1,262,545	85,484	1.1%	1.2%	1.0%	1.1%	1.6%
3	179,426	27,566	18,352	225,344	161,285	0.2%	0.2%	0.2%	0.2%	3.0%
4	664,680	56,877	36,048	757,605	12,398	0.7%	0.4%	0.4%	0.7%	0.2%
5	38,772	2,474	776	42,022	228	0.0%	0.0%	0.0%	0.0%	0.0%
6	674,552	101,948	82,551	859,051	4,468	0.7%	0.8%	0.9%	0.8%	0.1%
7	2,276	484	256	3,016	448	0.0%	0.0%	0.0%	0.0%	0.0%
8	4,261	299	393	4,953	20	0.0%	0.0%	0.0%	0.0%	0.0%
9	622	128	90	840	2	0.0%	0.0%	0.0%	0.0%	0.0%
10	667	72	118	857	4	0.0%	0.0%	0.0%	0.0%	0.0%
>10	1,344	418	450	10,046	42	0.0%	0.0%	0.0%	0.0%	0.0%
Total Websites	91,833,143	12,723,418	8,779,512	113,336,073	5,406,278	100.0%	100.0%	100.0%	100.0%	100.0%