# Poster: Evaluating Dual-Stack Content Delivery Within Websites

Florian Dekinder
University of Liège
Belgium
florian.dekinder@uliege.be

Eric Vyncke
CISCO
Belgium
evyncke@cisco.com

Benoit Donnet
University of Liège
Belgium
benoit.donnet@uliege.be

## ABSTRACT

This paper aims at evaluating how content distribution is spread between IPv4 and IPv6 within popular websites. Our measurement tool, based on a headless browser technology, has shown that most websites do not opt for IPv4 or IPv6 alone, but rather for a Mixed infrastructure where both protocols are required. Within this Mixed context, data shows that content delivery through IPv4 is predominant.

## CCS CONCEPTS

• **Networks** → *Network layer protocols*; *Network measurement*.

## KEYWORDS

dual-stack, IPv6, IPv4, content

**ACM Reference Format:**
Florian Dekinder, Eric Vyncke, and Benoit Donnet. 2024. Poster: Evaluating Dual-Stack Content Delivery Within Websites. In *Proc. 20th International Conference on emerging Networking EXperiments and Technologies (CoNEXT '24), December 9–12, 2024, Los Angeles, CA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3680121.3699886

## 1 CONTEXT

Since the early 80's, the Internet Protocol version 4 (IPv4) has been the main protocol supporting packet switching networks and internetworking technology. However, the exponential deployment of the Internet, since the mid-90's, has led to an exhaustion of IPv4 addresses [7], requiring a transition towards its successor, Internet Protocol version 6 (IPv6) [3]. Since then, IPv6 has been more and more adopted [2].

A smooth transition from IPv4 to IPv6 has been made possible thanks to *dual-stack* devices, i.e., devices with network interfaces that can originate and understand both IPv4 and IPv6 packets [9], avoiding so a complete shift from one protocol version to the other. Up to now, studies primarily focus on the adoption criteria, such as addressing and routing [4, 10], and on performance [6, 8]. However, few researches focused on dual-stack content delivery aspects. This is crucial as a significant portion of websites includes content fetched from multiple, distinct servers [1], possibly over different versions of the Internet Protocol, introducing additional delay. It is therefore still unclear which amount of content (e.g., *Image*, *Stylesheet*, *Script*, etc.) is delivered through which version of the Internet Protocol.

This is exactly what we want to investigate in this work. In particular, we make the following contributions: first, we introduce our data collection tool designed to assess and compare the content delivery capabilities of websites across dual-stack servers.[1]. Second, we have deployed the tool over several vantage points and collected data from more than 200,000 analyzable websites extracted and filtered from DomCop list [5]. Third, we propose to classify websites based on their server configurations into three classes: (*i*) Fully IPv4 websites, where all resources are delivered from IPv4-only servers; (*ii*) Fully Dual-Stacked websites, which load all their resources from servers supporting both IPv4 and IPv6; (*iii*) Mixed websites, which retrieve resources from a combination of both server types. Finally, our measurements show that the Mixed class is the largest one, with about 74% of the websites falling into it. Within this class, IPv4 delivery predominates, with 64% of Mixed websites fetching more resources in IPv4 than in IPv6.

## 2 DATA COLLECTION

Our measurement tool takes the URL of a website as input and drives a Google Chrome browser session on top of it. The latter is done using Selenium [11], a headless browser driver technology. It then queries Google's network inspection tool to collect the URLs of all resources fetched from the website along with their categories: *Document*, *Image*, *Media*, *Fetch*, *XHR*, *Font*, *Stylesheet*, *Script*, and *Other*. From these resource URLs, we extract the domains and query the DNS to determine whether the server hosting the resource is dual-stack or IPv4-only. We ran the tool over more than 200,000 websites identified within DomCop [5] domain list. The measurement campaign was launched on March $2^{\text{nd}}$, 2024 and concluded on March $25^{\text{th}}$, 2024.

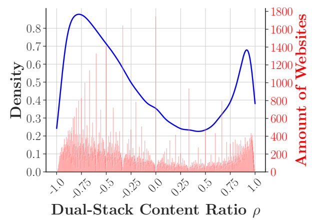
Based on the measurements data, we classify websites into three classes, according to their server configurations: (*i*) Fully IPv4 websites, where all resources are delivered from IPv4-only servers; (*ii*) Fully Dual-Stacked websites, which load all their resources from servers supporting both IPv4 and IPv6; (*iii*) Mixed websites, which retrieve resources from a combination of both server types.

## 3 PRELIMINARY RESULTS

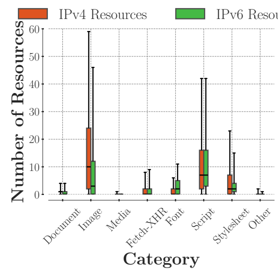
From the collected dataset, we observed that the Mixed class was the most predominant, with 74% of websites falling into it, followed by Fully IPv4 (~14%) and Fully Dual-Stacked (~12%). As a result, websites are less likely to utilize a single protocol exclusively

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
*CoNEXT '24, December 9–12, 2024, Los Angeles, CA, USA.*
© 2024 Association for Computing Machinery.
ACM ISBN 979-8-4007-1108-4/24/12…$15.00
https://doi.org/10.1145/3680121.3699886

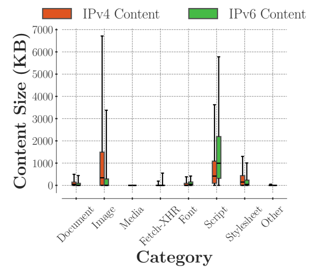
[1]Our source code and data is available at https://gitlab.uliege.be/Florian.Dekinder/dissecting-dual-stack-web-content



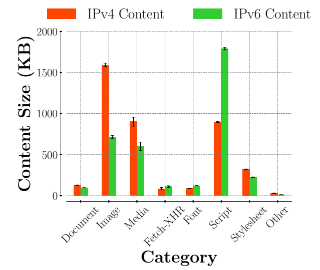

**(a) PDF of a normalized ratio ($\rho$) prioritizing more resources loaded over IPv4 on negatives values, and more resources loaded over IPv6 on positive values.**




**(b) Distribution of resources loaded over IPv4 and IPv6 for the different categories.**




**(c) IPv4 content size and IPv6 content size distribution over each category.**




**(d) IPv4 mean content size and IPv6 mean content size over each category. 95% confidence intervals are used.**

**Figure 1: Dual-Stack content distribution within Mixed websites. In the boxplots, the bottom and top whiskers represent the 5$^{th}$ and 95$^{th}$ percentiles, respectively. Outliers are not shown in the boxplots.**

and are instead more inclined to employ both in a Mixed configuration. Furthermore, we see that IPv4 remains a crucial component of the internet infrastructure, despite the ongoing transition to IPv6.

The Mixed websites are those which, when loaded with the default behavior of a dual-stack client (i.e., to prefer IPv6), load a certain portion of their resources in IPv6 and the other in IPv4. Fig. 1a presents the probability density function of a normalized ratio calculated for each so-called Mixed website. We express the ratio as

$$\rho = \frac{\texttt{IPv6}\_resources - \texttt{IPv4}\_resources}{total\_number}.$$

where $\rho \in ]-1,1[$. The limits of the interval (-1 and 1) are rejected as they represent the case of Fully IPv4 and Fully Dual-Stacked websites respectively. A negative $\rho$ means the website prefers to load more resources over IPv4, while a positive $\rho$ means the website prefers IPv6. A null value means that the website equally loads resources in IPv4 and IPv6.

Fig. 1a shows that 64% of the data points fall into negative values, indicating a higher number of resources are loaded over IPv4. Conversely, 35% of the points are positive, showcasing a preference for IPv6, with the remaining 1% balancing out at zero, indicating an equal distribution of resources across both protocols.

Fig. 1 delves deeper into the category of content loaded over each protocol. The box plots in Fig. 1b demonstrate that most categories predominantly load over IPv4. *Image* stands out the most by loading only 42% of the images over IPv6 when considering the 95$^{th}$ percentile. However, exceptions are observed with *Fetch & XHR* asynchronous requests and *Font*, which are more frequently loaded over IPv6. Even though the number of *Script* resources appears comparable between IPv4 and IPv6, a significant difference emerges when examining the raw –meaning after decoding– size of these resources. Fig. 1c reveals that scripts loaded over IPv6 tend to be larger and more complex, with sizes reaching up to 8,000 KB in the 95$^{th}$ percentile, while reaching only the half in IPv4 content.

The analysis of average values (Fig. 1d) across these resource categories further enhances our understanding of protocol prioritization by showing that *Media* files also predominantly use IPv4.

# 4 CONCLUSION AND FURTHER WORK

This work addresses a gap in the literature on the actual delivery of IPv4 and IPv6 content by websites. We developed a measurement tool to determine IP capabilities and categorize resources fetched by over 200,000 websites. Our findings reveal three distinct classes: 14% of websites use only IPv4, 12% fully adopt IPv6, and the remaining 74% are in hybrid configurations, favoring IPv4 64% of the time.

We found that *Image* category is more frequently loaded via IPv4, while *Script* is distributed equally across both protocols, but with larger scripts served over IPv6. Further studies could explore why IPv4 still prevails, potentially through a longitudinal study or by analyzing server environments for older hardware that does not support IPv6.

# ACKNOWLEDGMENTS

This work has been supported the CyberExcellence project funded by the Walloon Region, under number 2110186.

# REFERENCES

[1] M. Butkiewicz, H. V. Madhyastha, and V. Sekar. 2011. Understanding Website Complexity: Measurements, Metrics, and Implications. In *Proc. ACM Internet Measurement Conference (IMC).*

[2] J. Czyz, M. Allman, J. Zhang, S. Iekel-Johnson, E. Osterweil, and M. Bailey. 2014. Measuring IPv6 adoption. In *Proc. ACM SIGCOMM.*

[3] S. Deering and R. Hinden. 1998. *Internet Protocol, Version 6 (IPv6).* RFC 2460. Internet Engineering Task Force.

[4] A. Dhamdhere, M. Luckie, B. Huffaker, kc claffy, A. Elmokashfi, and E. Aben. 2012. Measuring the deployment of IPv6: topology, routing and performance. In *Proc. ACM Internet Measurement Conference (IMC).*

[5] DomCop. 2023. Top 10 million Websites Based on Open PageRank data. https://www.domcop.com/top-10-million-websites [Last Accessed: September 9th, 2024].

[6] Google. 2024. IPv6 deployment statistics. (2024). https://www.google.com/intl/en/ipv6/statistics.html [Last Accessed: September 7th, 2024].

[7] G. Huston. 2013–2024. *IPv4 Address Report.* https://ipv4.potaroo.net [Last Accessed: September, 8th 2024].

[8] G. Huston. 2015. Examining IPv6 Performance. https://labs.ripe.net/author/gih/examining-ipv6-performance/ [Last Accessed: Semptember, 8th 2024].

[9] Juniper. 2023. *Understanding Dual Stacking of IPv4 and IPv6 Unicast Addresses.* https://www.juniper.net/documentation/us/en/software/junos/is-is/topics/concept/ipv6-dual-stack-understanding.html [Last Accessed: September, 5th 2024].

[10] M. Nikkhah and R. Guérin. 2015. Migrating the Internet to IPv6: an exploration of the when and why. *IEEE/ACM Transactions on Networking (ToN)* 24, 4 (August 2015), 2291 – 2304.

[11] Selenium. 2023. Selenium automates browsers. That's it! https://www.selenium.dev [Last Accessed: August 19th, 2024].