Towards an IPv6-only WLCG: more successes in reducing IPv4

Garhan Attebury[1], *Marian* Babik[2], *Martin* Bly[3], *Nick* Buraglio[4], *Tim* Chown[5], *Dimitrios* Christidis[2], *Jiři* Chudoba[6], *Phil* DeMar[7], *José* Flix Molina[8], *Costin* Grigoras[2], *Bruno* Hoeft[9], *Hiro* Ito[10], *David* Kelsey[3*], *Edoardo* Martelli[2], *Shawn* McKee[11], *Carmen* Misa Moreira[2], *Raja* Nandakumar[3], *Kars* Ohrenberg[12], *Francesco* Prelz[13], *Duncan* Rand[14], *Andrea* Sciabà[2], *Tim* Skirvin[7], and *Christopher* Walker[5]

[1]University of Nebraska-Lincoln (UNL), 1400 R St, Lincoln, NE 68588, USA
[2]European Organization for Nuclear Research (CERN), CH-1211 Geneva 23, Switzerland
[3]UKRI STFC Rutherford Appleton Laboratory (RAL), Harwell Campus, Didcot OX11 0QX, UK
[4]Energy Sciences Network (ESnet), Lawrence Berkeley National Laboratory, Berkeley CA 94720, USA
[5]Jisc, Portwall Lane, Bristol BS1 6NB, UK
[6]Institute of Physics, Academy of Sciences of the Czech Republic, Prague 8, Czech Republic
[7]Fermi National Accelerator Laboratory (FNAL), P.O. Box 500, Batavia IL 60510, USA
[8]Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain
[9]Karlsruhe Inst. of Technology, Hermann-v-Helmholtz-Pl. 1, D-76344 Egg.-Leopoldshafen, Germany
[10]Brookhaven National Laboratory (BNL), 98 Rochester St., Upton NY 11973, USA
[11]University of Michigan Physics, 450 Church St, Ann Arbor, MI 48109, USA
[12]Deutsches Elektronen-Synchrotron (DESY), Notkestraße 85, D-22607 Hamburg, Germany
[13]INFN, Sezione di Milano, via G. Celoria 16, I-20133 Milan, Italy
[14]Imperial College London, South Kensington Campus, London SW7 2AZ, UK

Abstract. The Worldwide Large Hadron Collider Computing Grid (WLCG) community's deployment of dual-stack IPv6/IPv4 on its worldwide storage infrastructure has been very successful. Dual-stack is not, however, a viable long-term solution; the HEPiX IPv6 Working Group has focused on studying where and why IPv4 is still being used, and how to flip such traffic to IPv6. The agreed end goal is to turn IPv4 off and run IPv6-only over the wide-area network to simplify both operations and security management.

This paper reports our work since the CHEP2023 conference. Firstly, we present our campaign to deploy IPv6 on CPU services and Worker Nodes, with a deadline of end of June 2024. Then, the WLCG Data Challenge (DC24) performed in February 2024 was an excellent opportunity to observe the percentage of data transfers carried by IPv6. We observed the predominance of IPv6 in data transfers during DC24 and were able to understand yet more reasons for the use of IPv4 and areas for remedial action.

The paper ends with the working group's plans for moving WLCG to "IPv6-only". One aspect of this is the possible automated use of IPv6-only clients configured with a customer-side translator, or CLAT, together with a deployment of NAT64 using what is often known as "IPv6-Mostly", enabling IPv6-only sites to connect to non-WLCG IPv4-only services.

*e-mail: david.kelsey@stfc.ac.uk

1 Introduction

The Worldwide Large Hadron Collider Computing Grid (WLCG) infrastructure has been increasing its use of the Internet Protocol version 6, known as "IPv6", alongside the old Internet Protocol version 4, called "IPv4", in a mode known as "dual-stack". The HEPiX IPv6 Working Group [1] continues to investigate issues related to enabling IPv6 for WLCG services and to enable the use of IT resources that **only** communicate over IPv6, as agreed by the WLCG Management Board and originally presented by us in 2018 [2]. The agreed endpoint of the WLCG transition is to perform wide-area data transfers over only IPv6 and to remove the complexity and security concerns of operating dual-stack as presented by us in 2019 [3].

After the successful campaign to have WLCG storage IPv6-enabled through dual-stack in 2023, the Working Group requested the implementation of IPv6, again initially dual-stack, on worker nodes and computing elements, in order to make sure that IPv6-only resources could be reached. Having IPv4-only clients is a block to the implementation of IPv6-only services. A GGUS ticket campaign was launched with the deadline of June 2024. As of today, only ~65% of the sites have completed the request. Figure 1 shows how the WLCG's adoption of IPv6 has grown over the last 10 years, reaching an encouraging ~80% of all WLCG services. Another indicator, namely that ~90% of traffic on the LHCOPN links is IPv6, also confirms good progress.

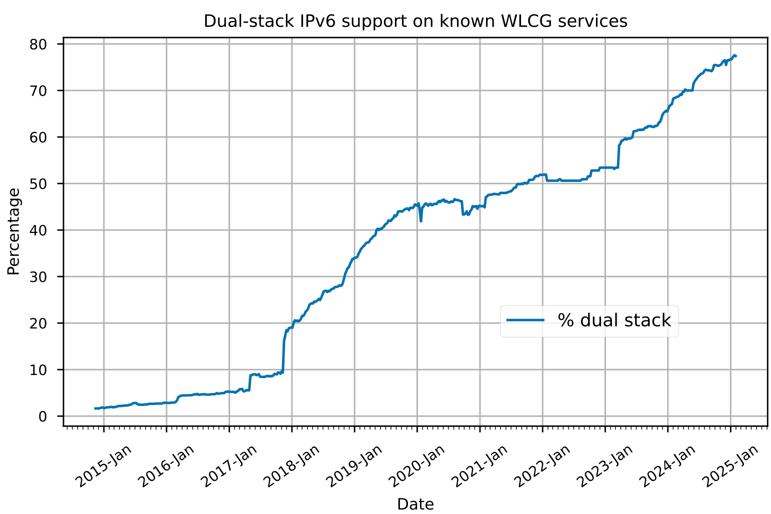

Figure 1. Percentage of WLCG services running on dual-stack IPv6/IPv4 servers

This paper is organised as follows. Firstly, we present the status of enabling IPv6 on CPU services and worker-nodes in WLCG. Then we describe our observations during the WLCG Data Challenge (DC24) performed in February 2024. Next, we present our ongoing work to understand and fix continued use of the IPv4 protocol [4]. The paper ends with the working group's current plans for moving WLCG to the use of "IPv6-only" in wide-area data transfers, together with the possible use of the IETF "IPv6-Mostly" approach by WLCG sites.

2 CPU ticket campaign

The motivations for mandating WLCG sites to deploy IPv6 also on their computing resources are several, were extensively discussed in HEPiX and WLCG, and are as follows:

- Address the scarcity of globally unique IPv4 addresses, which will likely affect new sites, avoiding the performance and management complexity of NAT
- Comply with the US federal government's mandate [5] to provide IPv6-only network access on systems on a relatively short timescale
- Add the IPv6-specific SciTag [6] packet marking functionality for traffic to/from worker nodes
- Remove a large reason for residual IPv4 traffic on WLCG
- The increasing likelihood of IPv6-only sites appearing, and ensuring access to them

The fact that most experiment software has been fully IPv6-compatible for several years, and that all WLCG sites have IPv6 networking, thanks to the IPv6 storage deployment campaign, seemed to make this goal readily attainable.

Moreover, all LHC experiments, as well as the DUNE collaboration, requested all their sites to deploy IPv6 on their entire infrastructure (at the very least, on all their Grid-facing services and resources); CMS even proposed, and it was agreed by everyone, to set a deadline for 30 June 2024.

The WLCG Management Board approved the new deployment campaign, which started on 28 November 2023, and sites were requested to enable IPv6 connectivity on all their computing elements and worker nodes, or explain why they might not be able to meet the deadline. The campaign did not request removal of IPv4, meaning sites would likely use a dual-stack approach. The campaign was conducted using GGUS support tickets, whose state is used to track progress. The current status of the campaign is shown in Figure 2, while progress over time is shown in Figure 3.

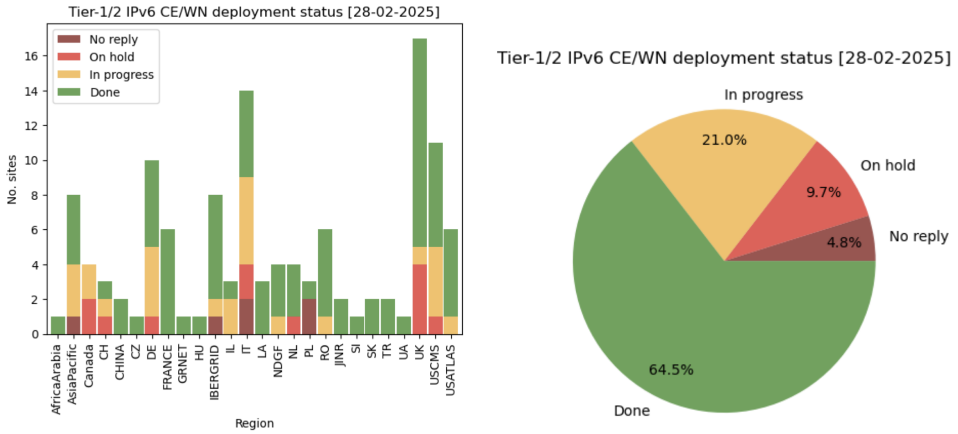

Figure 2. WLCG IPv6 CPU. Site status by region (left). Overall site status (right)

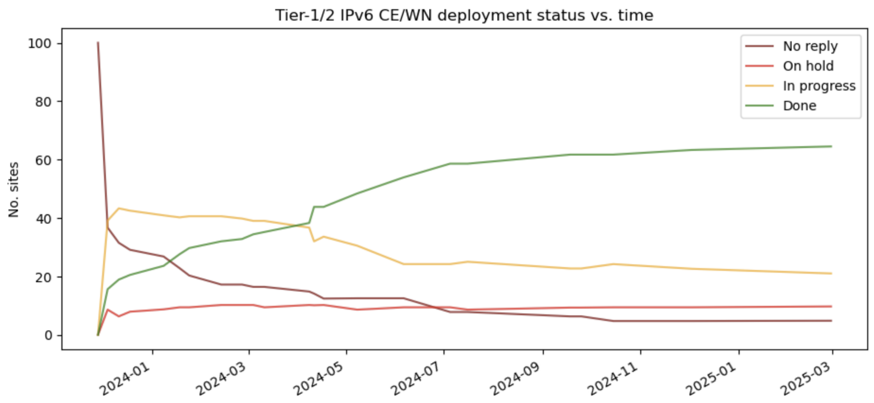

Figure 3. WLCG IPv6 CPU status vs time

As can be seen, the campaign is only two-thirds through, even six months after the deadline has passed. This is not completely unexpected, as it closely resembles the experience of the storage deployment campaign. Some recurring patterns could be identified, notably:

- Sites which took charge of the task, added it to their planning and eventually completed it, either before or after the deadline
- Sites which were, or are, constrained by external events (e.g. moving to a new data centre, or new hardware deployment, or improvements to the site networking infrastructure), and cannot yet proceed with the IPv6 deployment
- Sites which have higher priority tasks to deal with, and/or suffer from lack of manpower
- Sites which never acknowledged the request, a few of those finally announcing that the deployment had been completed

The conclusion is that little can be done to speed up the deployment, in addition to regularly inquiring about progress, while there is no real incentive for completing the deployment, or no negative consequences for not doing it. However, steady progress is still observed, even if slow, so we will continue driving the campaign until completion.

3 Lessons learned from DC24

The WLCG regularly conducts Data Challenges (DCs) to assess whether the required upgrades to the capabilities of the WLCG infrastructure have been made and are performant enough to meet future demands. The 2024 Data Challenge [7] covered the entire WLCG process chain, including data transfers from CERN over the wide-area network to the Tier-1/2 sites, storage systems at Tier-1/2, middleware, and tape systems at the Tier-1 sites. In 2021, a Data Challenge was held running at 10% of the expected capacity for the High Luminosity LHC (HL-LHC) run, which is scheduled to begin in 2030. The 2024 Data Challenge, held from February 12 to 23, targeted running at 25% of the HL-LHC data rates. The flexible target wide-area network data rate of 2.5 Tbps was successfully achieved for around 6 hours in the second week. The data transfer over the LHCOPN network reached 87% IPv6, as shown in Table 1.

The LHC experiments (ALICE, ATLAS, CMS, LHCb) also monitor their transferred data. However, the assignment to the underlying protocol (IPv6/IPv4) was not accurate. This discrepancy became evident when comparing their monitoring data with that of interfaces as-

 https://doi.org/10.1051/epjconf/202533701260

Table 1. Data rate during Data Challenge 2024 for incoming and outgoing traffic (in Gbps), including the percentage for both IPv4 and IPv6 protocols.

	Data Rate Outgoing (from CERN) [Gbps]	Data Rate Incoming (to CERN) [Gbps]	Percentage [%]
IPv4 traffic	87.3	32.6	13
IPv6 traffic	516	282	87

signed to a single protocol, such as IPv6, as revealed during the analysis of the entire monitoring data from DC24. Current efforts are focused on identifying the necessary corrections to the monitoring data, and ensuring these will be in place for the next DC.

Further analysis was conducted on the LHCOPN Tier-0 to Tier-1 VPN connection (see Figure 4) between CERN and DE-KIT, the German Tier-1, specifically analyzing the remaining IPv4 traffic. While IPv6 is enabled through dual-stack, there is still the potential for applications to, for some reason, use IPv4.

Although IPv4 and IPv6 traffic are routed over the LHCOPN on different overlay interfaces, the IPv4 traffic load can still be identified through interface monitoring. A summary of the findings is presented in the following paragraphs.

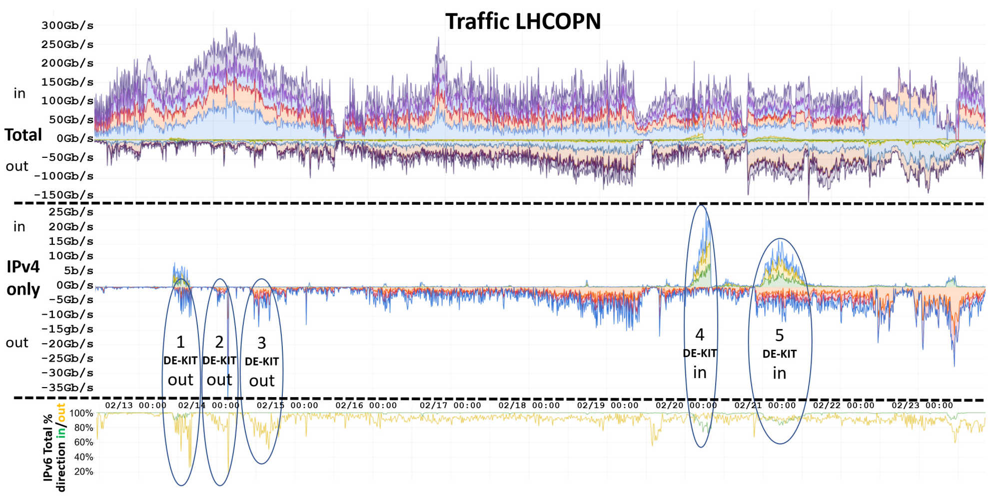

Figure 4. DE-KIT LHCOPN traffic overview during DC24. The first plot shows the total in/out traffic (in Gbps) for the LHCOPN network (Tier-0 to Tier-1 VPN). The plot in the middle, shows the IPv4-only traffic (in Gbps), with five highlighted regions where IPv4 traffic is significant, which are further discussed in the following paragraphs. Each VLAN is represented by a different color. The third plot displays the percentage in/out of IPv6 traffic.

The first significant IPv4 traffic load from DE-KIT to CERN was observed on the second day of the Data Challenge 2024, between 10:30 and 12:35 (CET). Using NetFlow data, more detailed insights were obtained. Over 98% of the 3 Terabytes of data exchanged between DE-KIT and CERN is presumed to be related to the ALICE experiment, as the servers involved at CERN are primarily associated with ALICE. Further investigation revealed that all the servers are dual-stack enabled. One possible reason could be that the initial transfer was not able to start with the IPv6 protocol, causing it to fall back to IPv4. Additionally, there was a small data transfer to a Tier-1 and some traffic between the BDII servers between DE-KIT and other Tier-1 sites, as well as IPv4 test traffic related to perfSONAR network monitoring servers. During DC24, the BDII server at DE-KIT was still IPv4-only, but since then IPv6 has been enabled.

The second block of data involved an additional 3 Terabytes of outgoing transfers from DE-KIT through CERN to another Tier-1 site. During the investigation of file transfer errors, it was found that the routing device was unable to handle the large number of IPv6 Neighbor Discovery entries. This resulted in a memory overflow due to the excessive number of concurrent IPv6 connections, which caused the file transfers to fail. To prevent further transfer failures, this Tier-1 site has been forced for now to initiate communication using only the IPv4 protocol, while still allowing external IPv6 communication.

The third block of outgoing IPv4 traffic from DE-KIT was towards the same Tier-1 site as the previous block. It was determined that the IPv4-only servers at a different Tier-1 belong to the Belle-II experiment, and that Tier-1 agreed to upgrade these servers to dual-stack during the next available downtime.

The fourth and fifth analyzed IPv4 blocks, with sizes of 33 and 43 Terabytes, respectively, were inbound IPv4 traffic at DE-KIT from CERN. A more detailed analysis revealed that most of the involved servers are configured with dual-stack. At CERN, a significant number of servers are involved compared to DE-KIT, with the server count approximately 100 times greater, as indicated by the data. For instance, in the fifth analyzed block, 2426 servers at CERN participated in the data transfer, whereas only 25 servers at DE-KIT were involved. A closer look at DE-KIT revealed that 17 of these servers were storage servers, and 8 were Squid servers. The reason why at both sides dual-stack deployed servers are still exchanging data over the protocol IPv4 may be explained by a suggestion of one of the experiment representatives, that if the transfer is not starting over IPv6, for whatever reason, the transfer is initiated over the protocol IPv4. Another explanation is that only the latest Squid server versions fully support the configuration to prioritize IPv6 as the preferred protocol.

4 Identification of on-going use of IPv4

At DE-KIT, a detailed monitoring system is deployed to analyze the worker node traffic. The data clearly indicates that DNS-related traffic has transitioned from being predominantly IPv4 to now operating at 99% over the IPv6 protocol. Similarly, CERN-VM File System (CVMFS) and Squid data exchanges have also reached 99% IPv6 traffic.

The Local Resource Management System (LRMS) has successfully migrated to 99% IPv6 usage. However, XRootD currently operates slightly above 90% on IPv6, indicating room for further improvement. Full migration to IPv6-only traffic has not yet been achieved, as only 40% of the total traffic is currently processed over IPv6.

A recently conducted IPv6-only test on worker nodes revealed that NFS at DE-KIT was still operating over IPv4. However, this issue was promptly addressed by enabling IPv6. Additionally, some worker nodes were deployed with IPv6-only configuration on both internal and external interfaces, and all jobs executed successfully. Nevertheless, job failures were observed on worker nodes where an IPv6-only loopback interface was also deployed. Further investigation, in collaboration with the experiments, revealed that the experiment software contained hard-coded references to the IPv4 loopback address. This issue has since been resolved, ensuring that jobs now run without failures.

Sporadic surges in IPv4 utilization are still observed within the WLCG networks. These occurrences are clearly visible in the network utilization statistics of the LHCOPN links, where IPv4 and IPv6 traffic are routed on different VLANs.

Thanks to the sFlow data generated and collected at CERN, it is possible to identify the specific applications and servers responsible for these significant IPv4 traffic flows. Although this analysis is highly time-consuming, it has proven valuable in detecting clients or servers that have not been IPv6-enabled, as well as identifying applications that may tend to use as default IPv4 instead of IPv6.

A recent case of this issue was observed between two Tier-1 sites exchanging data via GridFTP over IPv4, despite both source and destination being dual-stack. GridFTP requires explicit configuration to use IPv6, suggesting that some sites may not yet be setting it up correctly. However, since GridFTP is already deprecated for WLCG transfers, this IPv4 traffic is expected to diminish in the near future.

Another case identified through sFlow found Squid flows that preferred IPv4 even for communication between dual-stack client-server pairs. This behavior was confirmed by the developers, and efforts are currently in progress to modify this behavior and ensure proper IPv6 utilization.

5 Plans for IPv6-only

The Working Group's (WG) main goal is enabling the use of IPv6 in all WLCG wide-area data transfers. Once IPv6 is fully deployed on all services and clients we will concentrate on discouraging the use of IPv4 in data transfers and in other wide-area connections. A dual-stack IPv6/IPv4 network is the most complex to operate. The easiest way of ensuring IPv4 is not used is to move all WLCG clients and services to IPv6-only. The use of IPv6-only networking also simplifies operations, eases troubleshooting and minimizes the security attack surface.

The plan for moving to an IPv6-only WLCG includes the following components:

1. By the end of LHC Run 3 ensure all WLCG services and clients are fully IPv6-enabled. Once all services are dual-stack we will concentrate on removing the use of IPv4.
2. Ensure appropriate monitoring is in place to continue to identify and remove the use of legacy IPv4 on the LHCOPN network to a very small residual fraction thereby allowing the removal of IPv4-peerings on LHCOPN, making the core WLCG network IPv6-only.
3. Similarly, continue to identify, investigate and remove any ongoing use of IPv4 on the LHCONE network, as preparation towards making the LHCONE overlay IPv6-only.
4. Aim to complete this work during the first half of the LHC Long Shutdown 3 well before the start of HL-LHC Run 4.

6 The IETF: IPv6-Mostly

The recent approach proposed in the Internet Engineering Task Force (IETF) to remove unnecessary use of IPv4 is termed "IPv6-Mostly" [8]. It provides an automated means for hosts with CLAT support [9] to turn off IPv4 when the network signals the availability of a NAT64 translator [10]. The host uses DHCPv4 Option 108 [11] to confirm with its DHCP server that the network it is connecting to supports NAT64; if so it then disables IPv4. The resulting IPv6-only traffic is then either sent natively for IPv6-enabled destinations or via the CLAT for IPv4-only destinations.

This approach has been deployed in production on the campus eduroam WiFi at Imperial College London (UK), as reported at the UK IPv6 Council 2024 Annual Meeting [12]. Over

75% of a typical week's traffic was seen to be IPv6 (71,000 of 87,000 devices) thanks to the DHCPv4 Option 108 and CLAT support included in iOS, Android and macOS. As that support grows, with Windows expected soon, the percentage of IPv6 traffic will grow, leaving just legacy devices running either dual-stack or in some cases IPv4-only.

This promising approach is aimed at user-centric networks but could potentially be deployed in server networks operated by the WLCG, given that Linux CLAT implementations are available. Further investigation and WG discussion is required and will be reported on in the months ahead.

7 Conclusions

In this paper we have shown that the campaign to enable IPv6 for CPU services and worker nodes throughout WLCG is progressing well. Currently that means dual-stack, but the focus is shifting, as it has in the IETF, to identifying and removing unnecessary use of IPv4, such that IPv4 can be turned off. The working group (WG) has continued to identify and resolve such ongoing use of IPv4. The WG will continue to push the campaign to enable IPv6 on all services, all CPU and worker nodes as a stepping stone to an IPv6-only future. The LHCOPN network remains a good candidate to be the first to turn off its IPv4 peering. The exact timetable for the removal of IPv4 from wide-area network data transfers is still to be agreed by the WLCG Management Board, but the WG is planning that the date should be well before the start of the HL-LHC Run 4.

References

[1] S. Campana et al, J. Phys. Conf. Ser. **513**, 062026 (2014)
[2] M. Babik et al, EPJ Web Conf. **214**, 08010 (2019)
[3] M. Babik et al, EPJ Web Conf. **245**, 07045 (2020)
[4] M. Babik et al, EPJ Web Conf. **295**, 07036 (2024)
[5] Completing the Transition to Internet Protocol Version 6 (IPv6). USA OMB Memorandum M21-07, https://trumpwhitehouse.archives.gov/wp-content/uploads/2020/11/M-21-07.pdf (November 2020), accessed: 2025-02-26
[6] G. Attebury et al, EPJ Web Conf. **295**, 01036 (2024)
[7] The WLCG Data Challenge. K. Ellis. Plenary talk at this conference (CHEP2024) Krakow, Poland 23 Oct 2024. https://indico.cern.ch/event/1338689/contributions/6081539/ accessed: 2025-02-26
[8] IETF V6OPS Working Group, Multiple stateless dhcpv6 options for configuration of ipv4 over ipv6 tunnel, https://datatracker.ietf.org/doc/draft-ietf-v6ops-6mops/, accessed: 2025-02-17
[9] M. Mawatari, M. Kawashima, C. Byrne, 464XLAT: Combination of Stateful and Stateless Translation, RFC 6877, https://www.rfc-editor.org/rfc/rfc6877 (2013), accessed: 2025-02-17
[10] M. Bagnulo, P. Matthews, I. van Beijnum, Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers, RFC 6146, https://www.rfc-editor.org/rfc/rfc6146 (2011), accessed: 2025-02-17
[11] L. Colitti, J. Linkova, M. Richardson, T. Mrugalski, IPv6-Only Preferred Option for DHCPv4, RFC 8925, https://www.rfc-editor.org/rfc/rfc8925 (2020), accessed: 2025-02-17
[12] D. Stockdale, IPv6-Mostly at Imperial, https://www.ipv6.org.uk/wp-content/uploads/2024/10/08_IPv6-Mostly_at_Imperial.pdf (2024), accessed: 2025-02-17