ETSI GR IPE 001 V1.1.1 (2021-08)

IPv6 Enhanced Innovation (IPE); Gap Analysis

Disclaimer

The present document has been produced and approved by the IPv6 Enhanced Innovation ETSI Industry Specification Group (ISG) and represents the views of those members who participated in this ISG.
It does not necessarily represent the views of the entire ETSI membership.

Reference
DGR/IPE-001

Keywords
automation, gap analysis, IP, IPv6, network management, security

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM®** and the GSM logo are trademarks registered and owned by the GSM Association.

BLUETOOTH® is a trademark registered and owned by Bluetooth SIG, Inc.

Foreword

This Group Report (GR) has been produced by ETSI Industry Specification Group (ISG) IPv6 Enhanced Innovation (IPE).

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Executive summary

IPv6 is about addressing the future. The world is in a serious trajectory on the transition to IPv6. IPv6 is growing much faster than IPv4. The IPv6 penetration has reached a remarkable rate of 44,7 % worldwide as of June 2021 with China, India, USA, Europe, Brazil, and Japan driving the IPv6 adoption with their large IPv6 user base. However, the adoption is still unevenly creating a new digital divide between those taking leadership in deploying IPv6 and those still staying on the fence on IPv4. There are many reasons for such wide-ranging progress and gaps that need addressing.

With the global spread and rapidly increasing adoption of Internet Protocol version six (IPv6), the present document is reviewing the state of the protocol/technology and some additional work that needs to be completed for smoother and wider adoption of the protocol. IPv4 has been going through this process for many years and then IETF releases the: end work on IPv4, draft-ietf-sunset4-ipv6-ietf-01 [i.149] which states that "The IETF will stop working on IPv4, except where needed to mitigate documented security issues, to facilitate the transition to IPv6, or to enable IPv4 decommissioning".

IPv6 has been deployed on several physical mediums with adaptation layers such as Ethernet, ATM, point-to-point links, fibre optic, and wireless technologies such as Wi-Fi™, 4G, and 5G cellular as well as low bandwidth and low power protocols. For example, IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) is one of the evolving standards that integrates high speed wired IPv6 networks with a low power and bandwidth wireless network where there are independent access points interconnected by radio technologies that provide dynamic links with varying bandwidths, distances in which the state of the link can change. A customized Neighbour Discovery, routing, and other protocols had to be developed to handle these use cases and this edge case is still being tested and modified with operational experience.

In early 2000 there was preliminary testing and deployment of IPv6 usually in research and test networks. By 2009 most of the core protocols had stabilized and the industry was beginning to receive operational feedback on the deployment of the protocol. In 2011, the IANA central address pool was completely depleted. By 2012 network equipment vendors had transitioned the protocol processing from software/firmware to hardware and microcode on routers, switches, firewalls, and other networks, computing, and cybersecurity devices which provided higher speed, full line rate interfaces. In 2020, most of the address space at the registries has been depleted and the cost of buying IPv4 addresses on the resale market has increased above 20 dollars an IPv4 address which is another factor driving the adoption of IPv6.

The industry has now had ten to twelve years of operational experience with IPv6 with vastly larger networks than IPv4 has typically supported. Early uses cases such as cellular systems required tens to hundreds of millions of IP addresses that are not available in IPv4 even with duplicate or multiples of IETF RFC 1918 [i.77] space.

One aspect that hampered the initial adoption and deployment of IPv4 TCP/IP was the multiple protocols that it was replacing such as SNA, DECnet, IPX, and others, until the time when the technology refresh cycles supported the replacement of these network devices and interfaces. IPv4-only protocol stacks were then baked into OSs on computers, networking devices as well as Network Interface Cards that would plug into a larger variety of equipment and chassis. By 2010, IPv4 became the older protocol that needed to be deprecated and replaced by IPv6.

The size of networks and the internets, that connect them, have grown rapidly including the worldwide Internet which now has an estimated 12 billion or more devices connected, with a majority of these devices coming through NAT with IETF RFC 1918 [i.77] addresses. IPv4 only has 3,5 billion usable addresses out of the total 4,2 billion.

Because of the requirements for IPng (the next generation IP protocol) the new protocol, that became IPv6, because IPv5 was already a protocol, could not be backward compatible with IPv4 and the protocol designers were free to develop a scalable protocol to support hundreds of billions of devices or more. A full discussion of the IPv6 protocol is beyond the scope of this introduction but some of the changes include: 128-bit addresses, fixed-size packet header to allow packet processing in hardware, an enhanced link-layer protocol (ICMPv6) to scale with the number of devices attached on links and sub-network.

Technical clauses of the document start with clause 4 which has new industry challenges which discusses cloud and IP convergence, distributed edge computing, next generation IPv6-based transport infrastructure, research and education networks, and 3GPP narrowband IoT & IPv6. Clause 5 discusses existing solutions and gaps which starts with a brief overview of the differences between IPv4 and IPv6 packet headers as a foundation for the other clause which include: IP link and the implications for IPv6 link-layer support for subnets, IPAM, DHCP, DNS, and Orchestration systems which are needed for the vast number of IPv6 and IPv4 addresses that need to be managed, privacy, wired IPv6, wireless IPv6, Multicast Gaps, transition to IPv6-only for Carriers Broadband Services, Operations, Administrations, and Maintenance.

Clause 6 covers non-technical gaps including knowledge and experience, support for the current hardware and software, design and migration, as well as maintenance and support challenges. Clause 7 has general industry trends for technology transition; proactive ND; flexible Extension Headers; SRv6-based forwarding; service-based slicing, telemetry-based OAM, advanced multicast, service oriented networking, computing-based networking with clause 8 conclusions.

Introduction

IPv6 is growing very fast in every dimension:

- websites;
- users; and

- traffic.

Figure 1 is based on W3Tech [i.111] statistics. It shows the percentage of websites supporting IPv6 between January 2015 and January 2021. The Compound Annual Growth Rate (CAGR) has stable growth at the 23 %.

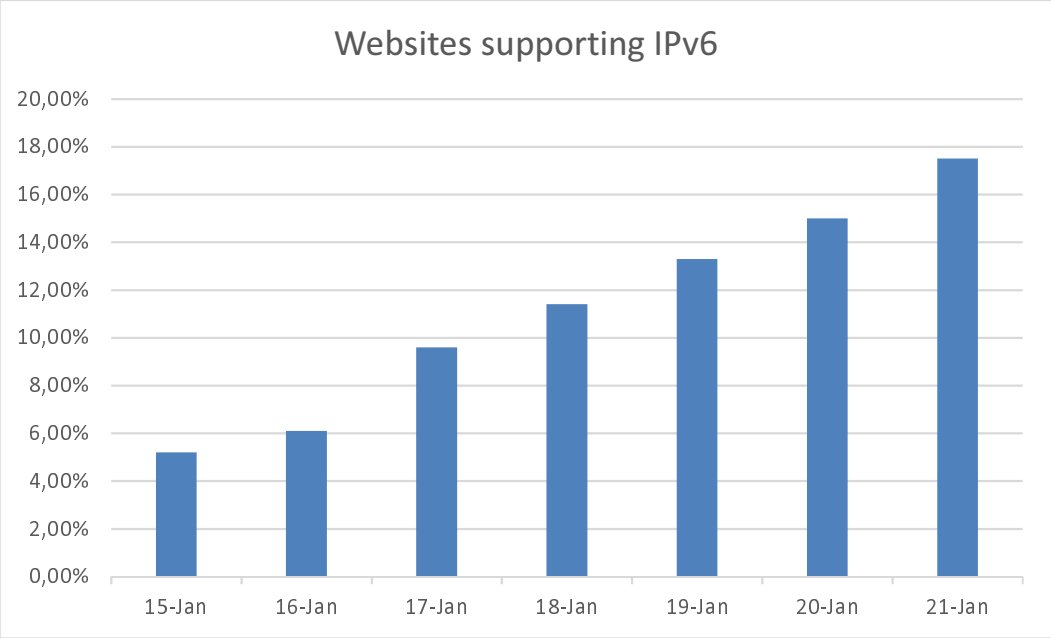


Figure 1: Websites support for IPv6 by W3tech

Websites statistics by Eric Vyncke [i.112] is based on Alexa's top-50 websites in every country. It shows even better IPv6 support (+10 %) and similar CAGR. The better IPv6 support is explained by the measurement of only the 50 biggest websites per country.

Although the percentage of websites supporting IPv6 is moderate (about 1/4), it allows for more than 50 % of traffic for some countries and more than 90 % for some carriers because traffic is not spread equally among sites. Most countries have a situation where a small list of websites generates the most traffic, and all of these sites typically support IPv6. Therefore, it may be concluded that most of the traffic is able to be serviced over IPv6.

World IPv6 user's number is consistent among many sources. Google™ user statistics [i.28] show a little better IPv6 proportion because all Google™ services are IPv6-enabled - see Figure 2.

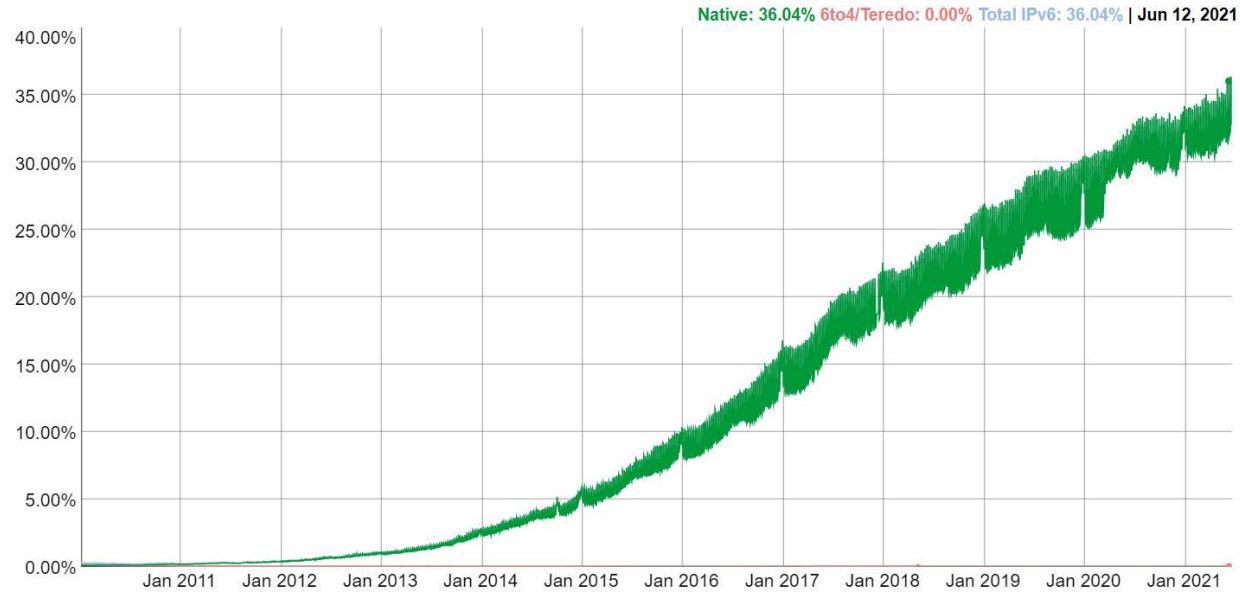


Figure 2: World IPv6 user's proportion by Google™

Google™ statistics does not cover China where Google™ services are filtered. China's statistics [i.142] claims 528 million IPv6 users that represent 8,7 % of Internet users (4 136 million from APNIC statistics). Hence, the overall IPv6 user proportion could be as high as 44,7 % every Saturday when the world has the spike of IPv6 usage.

APNIC [i.113] has a uniform collection of statistics from all users, except some countries filtering the Internet on borders. The IPv6 proportion of traffic or user's CAGR has the same healthy 20 %.

Akamai statistics [i.115] has from 1 % to 2 % per country difference because Akamai is more reflecting of the enterprise market.

Session statistics presented above is a good reflection of user's adoption but it may be not accurate for traffic estimation. From one point of view traffic per user is very similar worldwide, but from another point of view, the IPv6 proportion may be distorted by popular local websites that do not support IPv6. Carriers typically do not publish traffic statistics (except Internet exchanges). Rear published reports permit aggregate neither for global nor for regional scale.

India (58 %), Belgium (53 %), China (53 %), Germany (49 %), and US (47 %) are the most developed for IPv6 that is easy to check by all 3 sources mentioned above. Reliance™ India (94 %) and T-Mobile® US (93 %) are the leaders among big carriers.

Critical gaps are not possible, with such a high IPv6 adoption rate. Therefore, IPv6 has already been well-proven in production.

1 Scope

The present document discusses gaps that still exist in IPv6-related use cases.

Gaps investigation is split into four major blocks:

- New industry challenges - clause 4.
- Existing solutions and gaps - clause 5.
- Non-technical gaps - clause 6.
- Innovative solutions for the gaps - clause 7.

Every solution or technology is discussed briefly, to put gaps in the proper context.
It is assumed that the reader has good knowledge of IPv6. Recommended readings are [i.67], [i.69], [i.146], [i.147], [i.148].

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI does not guarantee their long-term validity.

The following referenced documents are not necessary for the application of the present document, but they assist the user with regard to a particular subject area.

[i.1] IETF RFC 4861: "Neighbor Discovery for IP version 6".

[i.2] IETF RFC 4862: "IPv6 Stateless Address Autoconfiguration".

[i.3] IETF RFC 8499: "DNS Terminology".

[i.4] IETF RFC 8415: "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)".

[i.5] IETF RFC 8028: "First-Hop Router Selection by Hosts in a Multi-Prefix Network".

[i.6] IETF RFC 8200: "Internet Protocol, Version 6 (IPv6) Specification".

[i.7] IETF RFC 8981: "Temporary Address Extensions for Stateless Address Autoconfiguration in IPv6".

[i.8] IETF RFC 6724: "Default Address Selection for Internet Protocol Version 6 (IPv6)".

[i.9] IETF RFC 7610: "DHCPv6-Shield: Protecting against Rogue DHCPv6 Servers".

[i.10] IETF RFC 6105: "IPv6 Router Advertisement Guard".

[i.11] IETF RFC 6877: "Combination of Stateful and Stateless Translation".

[i.12] IETF RFC 7597: "Mapping of Address and Port with Encapsulation (MAP-E)".

[i.13] IETF RFC 7599: "Mapping of Address and Port using Translation".

[i.14] IETF RFC 6333: "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion".

[i.15] IETF RFC 7596: "Lightweight 4over6: An Extension to the Dual-Stack Lite Architecture".

[i.16] IETF RFC 7252: "The Constrained Application Protocol (CoAP)".

[i.17] IETF RFC 7799: "Active and Passive Metrics and Methods (with Hybrid Types In-Between)".

[i.18] IETF draft-jdurand-ipv6-multicast-ra: "Route Advertisement Option for IPv6 Multicast Prefixes".

[i.19] IETF draft-pfister-moving-net-autoconf: "Routers auto-configuration using Route Information Option from ICMPv6 Router Advertisements".

[i.20] IDC FutureScape Highlights.

NOTE: Available at https://www.idc.com/getdoc.jsp?containerId=prUS46963620.

[i.21] IETF draft-ietf-v6ops-IPv6-deployment-status: "IPv6 Deployment Status".

[i.22] IETF RFC 7368: "IPv6 Home Networking Architecture Principles".

[i.23] IETF RFC 7157: "IPv6 Multihoming without Network Address Translation".

[i.24] IETF draft-ietf-6man-grand: "Gratuitous Neighbor Discovery: Creating Neighbor Cache Entries on First-Hop Routers".

[i.25] draft-pref64folks-6man-ra-pref64: "Discovering PREF64 in Router Advertisements".

[i.26] IETF RFC 8106: "IPv6 Router Advertisement Options for DNS Configuration".

[i.27] IETF RFC 5175: "IPv6 Router Advertisement Flags Option".

[i.28] Google™ IPv6 statistics

NOTE: Available at https://www.google.com/intl/en/ipv6/statistics.html.

[i.29] draft-ietf-idr-bgpls-srv6-ext: "BGP Link State Extensions for SRv6".

[i.30] IETF draft-ioametal-ippm-6man-ioam-ipv6-deployment: "Deployment Considerations for In-situ OAM with IPv6 Options".

[i.31] IETF RFC 8979: "Reaction of IPv6 Stateless Address Autoconfiguration (SLAAC) to Flash-Renumbering Events".

[i.32] IETF RFC 8660: "Segment Routing with the MPLS Data Plane".

[i.33] IETF RFC 8754: "IPv6 Segment Routing Header (SRH)".

[i.34] IETF RFC 8986: "Segment Routing over IPv6 (SRv6) Network Programming".

[i.35] IETF RFC 6146: "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers".

[i.36] IETF RFC 6052: "IPv6 Addressing of IPv4/IPv6 Translators".

[i.37] IETF RFC 7915: "IP/ICMP Translation Algorithm".

[i.38] IETF RFC 8402: "Segment Routing Architecture".

[i.39] IETF RFC 3031: "Multiprotocol Label Switching Architecture".

[i.40] IETF RFC 7348: "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks".

[i.41] IETF draft-ietf-lsr-isis-sr-vtn-mt: "Using IS-IS Multi-Topology (MT) for Segment Routing based Virtual Transport Network".

[i.42] IETF RFC 7011: "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information".

NOTE: Available at https://datatracker.ietf.org/doc/html/rfc7011.

[i.43] IETF RFC 8321: "Alternate-Marking Method for Passive and Hybrid Performance Monitoring".

[i.44] IETF draft-ietf-ippm-ioam-ipv6-options: "In-situ OAM IPv6 Options".

NOTE: Available at https://tools.ietf.org/html/draft-ietf-ippm-ioam-ipv6-options-04.

[i.45] IETF RFC 7039: "Source Address Validation Improvement (SAVI) Framework".

[i.46] OIF FLEXE-02.1: "Flexible Ethernet 2.1 Implementation Agreement".

[i.47] IETF draft-kumar-rtgwg-grpc-protocol: "gRPC Protocol".

[i.48] IETF draft-openconfig-rtgwg-gnmi-spec: "gRPC Network Management Interface (gNMI)".

[i.49] IETF draft-ietf-opsawg-ntf: "Network Telemetry Framework".

[i.50] IETF draft-ietf-bier-use-cases: "BIER Use Cases".

[i.51] IETF draft-ietf-idr-segment-routing-te-policy: "Advertising Segment Routing Policies in BGP".

[i.52] IETF RFC 7217: "A Method for Generating Semantically Opaque Interface Identifiers with IPv6 Stateless Address Autoconfiguration (SLAAC)".

[i.53] IETF draft-ietf-opsec-v6-27: "Operational Security Considerations for IPv6 Networks".

[i.54] IETF RFC 2236: "Internet Group Management Protocol, Version 2".

[i.55] IETF RFC 3376: "Internet Group Management Protocol, Version 3".

[i.56] IETF RFC 2475: "An Architecture for Differentiated Services".

[i.57] IETF RFC 1633: "Integrated Services in the Internet Architecture: an Overview".

[i.58] IEEE 802.15.4™: "Low-rate wireless personal area networks".

[i.59] IETF draft-ietf-pce-segment-routing-ipv6: "PCEP Extensions for Segment Routing leveraging the IPv6 data plane".

[i.60] IETF RFC 8955: "Dissemination of Flow Specification Rules".

[i.61] IETF RFC 7209: "Requirements for Ethernet VPN (EVPN)".

[i.62] IETF RFC 7432: "BGP MPLS-Based Ethernet VPN".

[i.63] IETF RFC 8214: "Virtual Private Wire Service Support in Ethernet VPN".

[i.64] IETF draft-ietf-teas-enhanced-vpn: "A Framework for Enhanced Virtual Private Network (VPN+) Services".

[i.65] IETF draft-ietf-spring-resource-aware-segments: "Introducing Resource Awareness to SR Segments".

[i.66] IETF RFC 8724: "SCHC: Generic Framework for Static Context Header Compression and Fragmentation".

[i.67] ETSI White Paper #35: "IPv6 Best Practices, Benefits, Transition Challenges and the Way Forward".

[i.68] Compendium on IPv6 Based Solutions/Architecture/Case Studies for Different Industry Verticals "National IPv6 Deployment Roadmap version II" in India.

NOTE: Available at https://dot.gov.in/sites/default/files/Compendium_on_IPv6_Based_Solutions.pdf.

[i.69] ETSI GR IP6 001 (V1.1.1): "IPv6 Deployment in the enterprise".

[i.70] IETF RFC 7381: "Enterprise IPv6 Deployment Guidelines".

[i.71] IETF RFC 8466: "A YANG Data Model for Layer 2 Virtual Private Network (L2VPN) Service Delivery".

[i.72] IETF RFC 8299: "YANG Data Model for L3VPN Service Delivery".

[i.73] IETF draft-ietf-lsr-flex-algo: "IGP Flexible Algorithm".

[i.74] IETF RFC 2710: "Multicast Listener Discovery (MLD) for IPv6".

[i.75] IETF RFC 7761: "Protocol Independent Multicast - Sparse Mode (PIM-SM)".

[i.76] IETF RFC 6513: "Multicast in MPLS/BGP IP VPNs".

[i.77] IETF RFC 1918: "Address Allocation for Private Internets".

[i.78] IETF RFC 6037: "Cisco Systems' Solution for Multicast in BGP/MPLS IP VPNs".

[i.79] IETF RFC 7113: "RFC 7113 - Implementation Advice for IPv6 Router Advertisement Guard".

[i.80] IETF RFC 7872: "Observations on the Dropping of Packets with IPv6 Extension Headers in the Real World".

[i.81] IETF RFC 5578: "PPP over Ethernet (PPPoE) Extensions for Credit Flow and Link Metrics".

[i.82] ETSI TS 129 281: "Universal Mobile Telecommunications System (UMTS); LTE; 5G; General Packet Radio System (GPRS) Tunnelling Protocol User Plane (GTPv1-U) (3GPP TS 29.281)".

[i.83] BBF TR-177: "IPv6 in the context of TR-101".

[i.84] IETF RFC 7084: "Basic Requirements for IPv6 Customer Edge Router".

[i.85] IETF RFC 8585: "Requirements for IPv6 Customer Edge Routers to Support IPv4-as-a-Service".

[i.86] IETF RFC 8305: "Happy Eyeballs Version 2: Better Connectivity Using Concurrency".

[i.87] BBF TR-187: "IPv6 for PPP Broadband Access".

[i.88] IEEE 802.1Q™: "IEEE Standard for Local and Metropolitan Area Networks-Bridges and Bridged Networks".

[i.89] IETF RFC 6775: "Neighbor Discovery Optimization for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)".

[i.90] IETF RFC 8505: "Registration Extensions for IPv6 over Low-Power Wireless Personal Area Network (6LoWPAN) Neighbor Discovery".

[i.91] IETF RFC 8928: "Address-Protected Neighbor Discovery for Low-Power and Lossy Networks".

[i.92] IETF RFC 8929: "IPv6 Backbone Router".

[i.93] IETF RFC 4291: "IP Version 6 Addressing Architecture".

[i.94] IETF RFC 6830: "The Locator/ID Separation Protocol".

[i.95] IETF draft-ietf-rift-rift: "RIFT: Routing in Fat Trees".

[i.96] ETF IETF RFC 7668: "IPv6 over BLUETOOTH® Low Energy".

[i.97] IETF RFC 6550: "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks".

[i.98] IETF RFC 9010: "Routing for RPL (Routing Protocol for Low-Power and Lossy Networks) Leaves".

[i.99] IETF RFC 6762: "Multicast DNS".

[i.100] IETF RFC 5942: "IPv6 Subnet Model: The Relationship between Links and Subnet Prefixes".

[i.101] IETF draft-ietf-6man-spring-srv6-oam: "Operations, Administration, and Maintenance (OAM) in Segment Routing Networks with IPv6 Data plane (SRv6)".

[i.102] IETF draft-yourtchenko-6man-dad-issues: "A survey of issues related to IPv6 Duplicate Address Detection".

[i.103] IETF RFC 3971: "SEcure Neighbor Discovery (SEND)".

[i.104] IETF RFC 3972: "Cryptographically Generated Addresses (CGA)".

[i.105] IETF BCP 38: "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing".

[i.106] IETF RFC 6959: "Source Address Validation Improvement (SAVI) Threat Scope".

[i.107] IETF RFC 8074: "Source Address Validation Improvement (SAVI) for Mixed Address Assignment Methods Scenario".

[i.108] IEEE 802.11™: "IEEE Standard for Information technology -- Telecommunications and information exchange between systems Local and metropolitan area networks -- Specific requirements -- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification".

[i.109] IEEE 802.11s™: "IEEE Standard for Information Technology -- Telecommunications and information exchange between systems -- Local and metropolitan area networks -- Specific requirements -- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 10: Mesh Networking".

[i.110] IEEE 802.15.10™: "IEEE Recommended Practice for Routing Packets in IEEE 802.15.4 Dynamically Changing Wireless Networks".

[i.111] W3Tech.

NOTE: Available at https://w3techs.com/technologies/overview/site_element.

[i.112] IPv6 Deployment Aggregated Status.

NOTE: Available at https://www.vyncke.org/ipv6status/index.php?small-samples=on.

[i.113] APNIC.

NOTE: Available at https://stats.labs.apnic.net/ipv6.

[i.114] ARCEP: "Annual Barometer of the transition to IPv6 in France" .

NOTE: Available at https://www.arcep.fr/fileadmin/reprise/observatoire/ipv6/Arcep_2020_Barometer_of_the_Transition_to_IPv6_dec2020.pdf.

[i.115] Akamai: "IPv6 Adoption By Country".

NOTE: Available at https://www.akamai.com/us/en/resources/our-thinking/state-of-the-internet-report/state-of-the-internet-ipv6-adoption-visualization.jsp.

[i.116] ETSI GS MEC 001: "Multi-access Edge Computing (MEC); Terminology".

[i.117] ETSI GS MEC 002: "Multi-access Edge Computing (MEC); Phase 2: Use Cases and Requirements".

[i.118] ETSI GS MEC 003: "Multi-access Edge Computing (MEC); Framework and Reference Architecture".

[i.119] IETF draft-ietf-bier-mld: "BIER Ingress Multicast Flow Overlay using Multicast Listener Discovery Protocols".

[i.120] IETF RFC 6514: "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs".

[i.121] IETF RFC 8279: "Multicast Using Bit Index Explicit Replication (BIER)".

[i.122] IETF RFC 8296: "Encapsulation for Bit Index Explicit Replication (BIER) in MPLS and Non-MPLS Networks".

[i.123] IETF draft-ietf-bier-ipv6-requirements: "BIER IPv6 Requirements".

[i.124] IETF draft-ietf-bier-problem-statement: "Bit Indexed Explicit Replication (BIER) Problem Statement".

[i.125] IETF RFC 8556: "Multicast VPN Using Bit Index Explicit Replication (BIER)".

[i.126] draft-ietf-bier-bierin6: "Supporting BIER in IPv6 Networks (BIERin6)".

[i.127] IETF RFC 8534: "Explicit Tracking with Wildcard Routes in Multicast VPN".

[i.128] IETF RFC 8401: "Bit Index Explicit Replication (BIER) Support via IS-IS".

[i.129] IETF RFC 8444: "OSPFv2 Extensions for Bit Index Explicit Replication (BIER)".

[i.130] IETF draft-ietf-bess-srv6-services: "SRv6 BGP based Overlay Services".

[i.131] IETF draft-ietf-bier-oam-requirements: "Operations, Administration and Maintenance (OAM) Requirements for Bit Index Explicit Replication (BIER) Layer".

[i.132] IETF draft-ietf-bier-pmmm-oam: "Performance Measurement (PM) with Marking Method in Bit Index Explicit Replication (BIER) Layer".

[i.133] IETF RFC 4176: "Framework for Layer 3 Virtual Private Networks (L3VPN) Operations and Management".

[i.134] IETF RFC 4364: "BGP/MPLS IP Virtual Private Networks (VPNs)".

[i.135] IETF RFC 4664: "Framework for Layer 2 Virtual Private Networks (L2VPNs)".

[i.136] IETF RFC 6624: "Layer 2 Virtual Private Networks Using BGP for Auto-Discovery and Signaling".

[i.137] IETF RFC 8077: "Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)".

[i.138] IETF RFC 4448: "Encapsulation Methods for Transport of Ethernet over MPLS Networks".

[i.139] IETF RFC 4761: "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling".

[i.140] IETF RFC 4762: "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling".

[i.141] IETF draft-ietf-spring-sr-for-enhanced-vpn: "Segment Routing based Virtual Transport Network (VTN) for Enhanced VPN".

[i.142] China IPv6 statistics.

NOTE: Available at https://www.china-ipv6.cn/#/.

[i.143] IETF draft-ietf-idr-sr-policy-ifit: "BGP SR Policy Extensions to Enable IFIT".

[i.144] IETF draft-ietf-netconf-distributed-notif: "Subscription to Distributed Notifications".

[i.145] IETF draft-ietf-ippm-ioam-yang: "A YANG Data Model for In-Situ OAM".

[i.146] ETSI GR IP6 031 (V1.1.1): "IPv6 Security, Cybersecurity, Blockchain".

[i.147] ETSI GR IP6 006 (V1.1.1): "Generic migration steps from IPv4 to IPv6".

[i.148] ETSI GR IP6 010 (V1.1.1): "IPv6-based SDN and NFV; Deployment of IPv6-based SDN and NFV".

[i.149] IETF draft-ietf-sunset4-ipv6-ietf-01: "IETF: End Work on IPv4".

[i.150] CPNI (US).

NOTE: Available at https://www.fcc.gov/general/customer-privacy.

[i.151] HIPPA (US).

NOTE: Available at https://www.hhs.gov/hipaa/index.html.

[i.152] FCRA (US).

NOTE: Available at https://www.ftc.gov/enforcement/statutes/fair-credit-reporting-act.

[i.153] ECPA (US).

NOTE: Available at https://bja.ojp.gov/program/it/privacy-civil-liberties/authorities/statutes/1285.

[i.154] GDPR (Europe).

NOTE: Available at https://gdpr-info.eu/.

[i.155] CSL (China).

NOTE: Available at http://www.informatica-juridica.com/ley/chinas-cyber-security-law-csl/.

[i.156] DPB (India).

NOTE: Available at https://prsindia.org/billtrack/the-personal-data-protection-bill-2019.

[i.157] Data Protection Act 2018 (UK).

NOTE: Available at https://www.gov.uk/data-protection#:~:text=The%20Data%20Protection%20Act%202018%20is%20the%20UK%27s%20implementation%20of,used%20fairly%2C%20lawfully%20and%20transparently.

3 Definition of terms, symbols and abbreviations

3.1 Terms

For the purposes of the present document, the following terms apply:

464XLAT: combination of Stateful and Stateless Translation

NOTE: See IETF RFC 6877 [i.11].

6bone: test bed for Internet Protocol version 6

A: DNS record for translation of a domain name into a 32-bit IPv4 address

AAAA: DNS record for translation of a domain name into a 128-bit IPv6 address

application-aware networking: application characteristic information such as application-aware identification and network performance requirements is carried in the packet encapsulation in order to facilitate service provisioning, perform fine-granularity traffic steering and network resource adjustment

ARCEP: France's Electronic Communications, Postal and Print media distribution Regulatory Authority

binding Segment IDentifier (SID): type of label introduced in Segment Routing

control plane: part of the router architecture that is concerned with drawing the network topology, or the information in a routing table that defines what to do with incoming packets

data plane: part of the networking node that processes the data requests

DECnet: suite of network protocols

DHCPv6- Shield: mechanism for protecting hosts connected to a switched network against rogue DHCPv6 servers

NOTE: See IETF RFC 7610 [i.9].

Evolved Node B (eNodeB): element in E-UTRA of LTE, radio base station that supports LTE technology

flow label: maintain the sequential flow of the packets belonging to a communication

FlowSpec: propagation of traffic filtering rules by BGP

NOTE: See IETF RFC 8955 [i.60].

GÉANT: pan-European data network for the research and education community

homenet: home network, comprising host and router equipment, with one or more CE routers providing connectivity to a service provider network(s)

in-situ Operations, Administration, and Maintenance (iOAM): used for recording and collecting operational and telemetry information

Internet2: not-for-profit United States computer networking consortium led by members from the research and education communities, industry, and government

lw4o6: extension to the Dual-Stack Lite Architecture

NOTE: see IETF RFC 7596 [i.15].

mesh-under: Link Layer technology that emulates a broadcast domain by flooding the broadcast packets at the Link Layer

provider-independent: block of IP addresses assigned by a Regional Internet Registry (RIR) directly to an end-user organization

provider-aggregatable: block of IP addresses assigned by a regional Internet registry to an Internet service provider

Quality of Experience (QoE): measure of the delight or annoyance of a customer's experiences with a service

Quality of Service (QoS): description or measurement of the overall performance of a service

RA-guard: filtering in the layer-2 network against rogue Router

NOTE: See IETF RFC 6105 [i.10].

reverse-path forwarding: technique used in modern routers for the purposes of ensuring loop-free forwarding of multicast packets in multicast routing and to help prevent IP address spoofing in unicast routing

route-over: Multi-Link Subnet (MLSN) in LLN

sampled flow: industry standard for packet export at Layer 2 of the OSI model

NOTE: See https://sflow.org.

Segment Routing (SR): steering packets through an ordered list of instructions to realizes end-to-end policy without creating any per-flow state in the network

Service Function Chaining (SFD): "steering" of traffic through the set of services

TCPDUMP: data-network packet analyser

telemetry: in-situ collection of measurements or other data at remote points and their automatic transmission to receiving equipment (telecommunication) for monitoring

Wi-Fi™: family of wireless network protocols

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

4G	The fourth generation technology standard for broadband cellular networks
5G	The fifth generation technology standard for broadband cellular networks
6LoWPAN	IPv6 over Low-Power Wireless Personal Area Networks
AI	Artificial Intelligence
AP	Access Point
APAN	Asia Pacific Advanced Network
API	Application Programming Interface
APNIC	Asia Pacific Network Information Centre
AR	Address Resolution
ARP	Address Resolution Protocol
AS	Autonomous System
ASIC	Application-Specific Integrated Circuit
ASN	Autonomous System Number
ATM	Asynchronous Transfer Mode
BBF	Broadband Forum
BCP	Best Current Practice

NOTE: IETF type for the document.

BE	Best Effort

NOTE: IETF working group.

BESS	BGP Enabled ServiceS
BFIR	Bit Forwarding Ingress Routers
BFR	Bit Forwarding Routers
BGP	Border Gateway Protocol
BGP-LS	BGP Link-State protocol
BIER	Bit Index Explicit Replication
BLE	Bluetooth® Low Energy
BNG	Broadband Network Gateway
BSS	Basic Service Set
BSS	Business Support System
CAGR	Compound Annual Growth Rate
CDN	Content Delivery Network
CERNET	China Education and Research Network
CERNET2	China Education and Research Network v2
CGN	Carrier-Grade NAN
CLI	Command Line Interface
CPNI	Centre for the Protection of National Infrastructure
CPU	Central Processing Unit
CSL	China's Cyber Security Law
DAD	Duplicate Address Detection
DDI	DNS, DHCP and IPAM
DEX	Direct Export
DHCP	Dynamic Host Configuration Protocol

NOTE: See IETF RFC 8415 [i.4].

DHCPv4	Dynamic Host Configuration Protocol for IPv4

DHCPv6 Dynamic Host Configuration Protocol for IPv6

NOTE: See IETF RFC 8415 [i.4].

DNA Detecting Network Attachment
DNS Domain Name System

NOTE: See IETF RFC 8499 [i.3].

DNS64 DNS with translation from IPv4 to IPv6
DOS Denial-of-Service attack
DPB Personal Data Protection Bill
DPI Deep packet inspection
DSL Digital subscriber line
DS-Lite Dual-Stack Lite Broadband Deployments

NOTE: See IETF RFC 6333 [i.14].

E2E End to End
EARO Address Registration Option

NOTE: See IETF RFC 8505 [i.90].

ECMP Equal-cost Multi-path routing
ECPA Electronic Communications Privacy Act
EDAC Extended Duplicate Address Confirmation

NOTE: See IETF RFC 8505 [i.90].

EDAR Extended Duplicate Address Request

NOTE: See IETF RFC 8505 [i.90].

EH Extension headers
eMBB Enhanced Mobile Broadband
eNodeB Evolved Node B
EPS Evolved Packet System
ESP Encapsulating Security Payload
ESS Extended Service Set
EVPN Ethernet VPN
FBB Fixed Broadband
FCRA Fair Credit Reporting Act
FQDN Fully Qualified Domain Name
FRR Fast Reroute
Gbps Billions of bits per second
GDPR General Data Protection Regulation
GPON Gigabit Passive Optical Networks
GRE Generic Routing Encapsulation
gRPC Open Source Remote Procedure Call
GRT Global Routing Table
GS Group Specification
GTP GPRS Tunnelling Protocol
GUA Global Unicast Addresses
HD High-Definition
HIPAA Health Insurance Portability and Accountability Act
HTTP Hyper-Text Transfer Protocol
IAM Identity and Access Management
IANA Internet Assigned Numbers Authority
ICT Information and Communications Technology
ID Identification
IDC International Data Corporation
IDR Inter-Domain Routing

NOTE: IETF working group

IGMP Internet Group Management Protocol
IGP Interior gateway protocol
iOAM In-situ Operations, Administration, and Maintenance
IoT Internet of Things
IoV Internet of Vehicles
IP Internet Protocol
IPALM IP Address Lifecycle Management
IPAM IP Address Management
IPE IPv6 Enhanced Innovations

NOTE: ETSI working group.

IPFIX Internet Protocol Flow Information Export

NOTE: See IETF RFC 7011 [i.42].

IPng The next generation IP protocol
IPoE Internet Protocol over Ethernet
IPS Intrusion Prevention System
IPSec Internet Protocol Security
IPTV Internet Protocol Television
IPv4 Internet Protocol version 4
IPv4aaS IPv4 as a service
IPv5 Internet Protocol version 5
IPv6 Internet Protocol version 6
IPX Internetwork Packet Exchange
IS-IS Intermediate System to Intermediate System protocol
ISP Internet Service Provider
IT Information Technology
IXP Internet Exchange Point
L2 Layer 2 of OSI model
L2R Layer 2 Routing
LAN Local Area Network
LISP Locator/ID Separation Protocol

NOTE: See IETF RFC 6830 [i.94].

LLA Link-local address
LLN Low-Power and Lossy Network
LPWAN Low-power wide-area network
LSR Label Switching Router
lw4o6 Lightweight 4over6
MAC Media Access Control address
MAP-E/T Mapping of Address and Port using Encapsulation/Translation

NOTE: See IETF RFC 7597 [i.12] and IETF RFC 7599 [i.13].

MBB Mobile Broadband
mDNS Multicast DNS

NOTE: See IETF RFC 6762 [i.99].

MEC Multi-access Edge Computing
MLD Multicast Listener Discovery
MLSN Multi-Link Subnet

NOTE: MLSN is as a collection of Hub-and-Spoke link layer domains where the Hubs form a connected dominating set of the member nodes, and IPv6 routing takes place between the Hubs within the subnet.

mMTC Massive Machine Type Communication
MPLS Multiprotocol Label Switching
MTU Maximum Transmission Unit
MVPN Multicast VPN
NA Neighbor Advertisement

NAT Network address translation
NAT46 Network address translation from IPv4 to IPv6
NAT64 Network address translation from IPv6 to IPv4
NB-IoT Narrowband IoT
NBMA Non-broadcast Multiple Access Network
NCE Neighbour Cache Entry
ND Neighbour Discovery protocol

NOTE: See IETF RFC 4861 [i.1].

NETCONF Network Configuration Protocol
NFV Network Function Virtualisation
NG-MVPN Next Generation Multicast VPN
NMS Network Management System
NREN National Research and Education Network
NS Neighbour Solicitation

NOTE: See IETF RFC 4861 [i.1].

OAM Operations, Administration, and Maintenance
OIF Optical Internetworking Forum
OPEX Operating Expenses
OS Operating System
OSPF Open Shortest Path First
OSPFv3 Open Shortest Path First version 3
OSS Operational Support Systems
P2MP Point-to-Multipoint
P2P Point-to-Point
PA Provider-Aggregatable
PCEP Path computation element
PCRF Policy and Charging Rules Function
PHY Physical Layer
PI Provider-Independent
PIM Protocol Independent Multicast
PIM-SM Protocol Independent Multicast - Sparse-Mode
PIO Prefix Information Options

NOTE: See IETF RFC 4861 [i.1].

PPPoE Point-to-Point Protocol over Ethernet

NOTE: See IETF RFC 5578 [i.81].

QoE Quality of Experience
QoS Quality of Service
R&E Research and Education
RA Router Advertisement of ND protocol

NOTE: See IETF RFC 4861 [i.1].

RAN Radio Access Network
RFC Request for Comments
RG Residential Gateway
RIFT Routing In Fat Trees

NOTE: See IETF draft-ietf-rift-rift [i.95].

RIPE Réseaux IP Européens
RIR Regional Internet Registry
RISC Reduced Instruction Set Computer
ROHC Robust Header Compression
RPF Reverse-Path Forwarding
RPL IPv6 Routing Protocol for Low-Power and Lossy Networks
RRs Resource Records

RS Router Solicitation
RSVP Resource Reservation Protocol
SAFI (BGP) Subsequent Address Family Identifier
SAVI Source Address Validation Improvement

NOTE: See IETF RFC 7039 [i.45].

SCHC Static Context Header Compression
SDN Software Defined Network
SD-WAN Software-Defined WAN
SID Segment IDentifier
SIEM Security Information and Event Management
SLA Service-Level Agreement
SLAAC Stateless Address Autoconfiguration

NOTE: See IETF RFC 4861 [i.1] and IETF RFC 4862 [i.2].

SLO Service Level Objectives
SMB Small and Medium Business
SMS Short Message Service
SNA Systems Network Architecture
SNMA Solicited-Node Multicast Address

NOTE: See IETF RFC 4291 [i.93].

SPRING Source Packet Routing in NetworkinG

NOTE: IETF working group.

SR Segment Routing
SRH Segment Routing Header
SRv6 Segment Routing for IPv6
SSID Service Set IDentifier
STA (Wi-Fi) Station
TCP Transmission Control Protocol
TE Traffic Engineering
TEAS Traffic Engineering Architecture and Signaling
TV Television

NOTE: IETF working group.

UDP User Datagram Protocol
UE User Equipment
UK United Kingdom
ULA Unique Local Address
UPF User Plane Function
uRLLC Ultra-Reliable Low-Latency Communication
US United States
VLAN Virtual Local Area Network
VPLS Virtual Private LAN Service
VPN Virtual private network
VR Virtual Reality
VTN Virtual Transport Network
VxLAN Virtual Extensible LAN
WAN Wide Area Network
WG Working Group
WLAN Wireless LAN
WPAN Wireless Personal Area Network
YANG Yet Another Next Generation data modelling language

4 New industry challenges

4.1 Industry trends

The industry has many new technology domains like 5G, Internet of Things (IoT), cloud computing, cloud-network convergence, and distributed edge for services. Wide-spread digitalization and automation are accelerating. As more and more devices, vehicles, sensors, and home appliances go online, enabling IP on everything becomes a vision and industry trend.

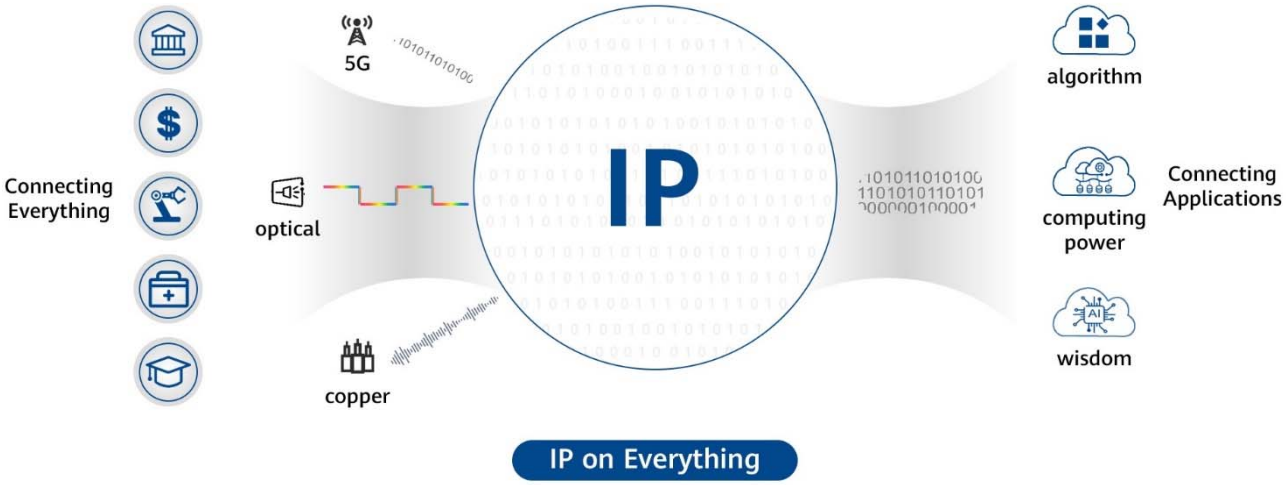


Figure 3: IP on everything

IPv6 is the suitable choice for this expansion, because of the IPv4 address shortage and lack of IPv4 innovations. The functionality and ecosystem of IPv6 is mature and has a more and more important role for the telecommunication industry. New scenarios like cloud & IP network convergence, distributed edge computing, 5G transport, and IoT are introducing new challenges and triggering IPv6 enhanced innovation.

4.2 Cloud & IP network convergence

4.2.1 Background

More and more computing and storage resources are moving to public and private cloud environments. Typical 3-tiered application architectures (based on user interface, application and database) are becoming split between different physical locations. It creates additional requirements for networking: increased performance, greater flexibility, increasingly virtualized and segmented networks, and stronger QoS and security requirements. Ubiquitous connectivity, deterministic quality, low latency, automation, and security are becoming the strong requirements.

Enterprises prefer competitive proposals from different public cloud providers. They may also need different cloud functionalities in different cloud providers. Enterprises would like to keep some critical parts of applications under their control for security and resiliency reasons and host in their own private cloud. All of these applications need to interact with each other, no matter where they are located. All the above creates the requirement for connections to many cloud providers at the same time with flexibility to adapt at any time, hence, the multi-cloud requirement.

During its FutureScapes 2020 event, the International Data Corporation (IDC) predicted [i.20] that by the end of 2021, 80 % of enterprises will put a mechanism in place to shift to cloud-centric infrastructure and applications twice as fast as before the pandemic.

New technology innovations like SD-WAN, NFV, and containerization provide considerably improved ICT value for the business: through agility, flexibility, and cost, which is an additional driver for architectural changes.

Enterprises would like to enhance their competitive advantages with the help of multi-cloud deployment, high-performance edge-cloud collaboration, and integrated service provisioning. Network-to-cloud integration needs to adapt to many constantly evolving scenarios. The network requirements are summarized below:

- Adaptive and optimized workload placement in customer premises, carrier's edge cloud or multiple public cloud environments, and application-aware networking, are important for hybrid-multi-cloud deployment. Think of this as 'network native' applications that are built leveraging network intelligence and/or without as much reliance on network resiliency.
- In the hybrid-multi-cloud paradigm, the possibility to connect the workload with SLAs is an essential requirement.
- The workload in hybrid-multi-cloud has to be harmonized within the whole lifecycle of the service, especially in service creation and service scaling phases.
- The cloud application needs a flexible configuration (by service provider's console or API) with rules to allow them to quickly address the service workload variation over time.
- The connectivity setup needs to be able to select the path matching the target SLA, have the ability to scale, and reserve resources according to service need.
- Network resources reservation needs flexibility to meet the requirements of most demanding applications (path, slice, and service chain).
- Network connectivity needs fine-grain performance measurement & visualization.

4.2.2 Current enterprise access connections limitations

Carriers are currently selling mostly dumb access. It is defined by static characteristics, e.g. the maximum bandwidth of the access medium. For private WAN, the current connectivity services are based on MPLS Virtual Private Network (VPN) that has a long list of deficiencies:

- high cost (10x);
- fixed SLA parameters;
- limited assurance: the SLA violation is common. That pushes many enterprises to have separate monitoring systems to check SLA violations;
- long provisioning time (weeks);
- inadequate automation to connect dynamic virtual cloud instances;
- lack of availability for cross-carrier MPLS VPN.

Enterprises need the possibility to request new connectivity services with specific SLAs. Moreover, there is the need for the flexibility to define scaling criteria in the case of overload to match the current public cloud capabilities.

There is a trend for evolution from MPLS to IPv6-based technologies to overcome those functional gaps.

4.2.3 IPv6 benefits

Cloud services are based on software and hardware virtualization. Virtualization may be applied to a wide range of concepts, such as overlay networks, storage, and servers where operating systems are run in virtual machines or containers, while applications are run in an isolated environment.

IPv6 has several benefits for such an environment:

- Large address space: many virtual instances, especially with micro-services, create a strong demand for many IP addresses. IPv6 has a much bigger address space. It is easy to create a scalable and public address scheme for the cloud of any size with address summarization, room for growth, and the possibility to have connectivity from any service to any other service.

- Scalability: IPv6 provides the ability to create a structured address plan with summarization on all levels; this decreases the size of the forwarding table.
- Simplified header: the processing burden may be reduced for the network equipment because of the simplified IPv6 basic header.
- Enhanced security: IPv6 enhanced functionality gives more possibilities for security improvements.
- Privacy protection: IPv6 allows many IP addresses for the host on the link, which provides additional privacy protection.
- Extension headers: IPv6 has unlimited flexibility because it has EH (Extension Headers). SRv6 is an example of technology that benefits from IPv6 extensibility. It helps to create an on-demand path with high flexibility. Technology like telemetry provides near real-time performance measurement up to a single flow, and enables end-to-end SLA guaranties.

NAT avoidance: IPv6 end-to-end connectivity is valuable for many services. Those functionalities provide agility in service creation and predictive SLA.

4.2.4 Conclusion

IPv6 is a mature and flexible protocol. Cloud and IP network convergence would benefit from the improvements brought by IPv6. It has flexibility to add new functionality to satisfy cloud requirements. It is especially important for service-rich cloud deployments.

The trend of cloud-network convergence and IPv6 transition is happening simultaneously. It is possible to leverage the latest IPv6 technologies to make cloud services more secure, agile, flexible, and scalable.

4.3 Distributed Edge Computing

4.3.1 Background

With the rise of cloud computing, more and more application services are migrating from enterprise servers to cloud data centres. Applications benefit from maintenance-free hardware, elastic scaling, and flexible deployment of network and security solutions. A data centre may serve users within a radius of hundreds to thousands of kilometres. This distance introduces a significant transmission delay, which degrades the Quality of Experience (QoE) for some delay-sensitive applications. Streaming applications may introduce continuous usage of bandwidth from the edge to the backbone and then the data centres. This may put significant pressure on the network. Deterministic quality, low latency, automation, and security are becoming the strong requirements.

Some scenarios have high requirements for low latency performance in the 5G and cloud era:

- For the cloud gaming scenario, the rendering of game graphics/videos and the transmission of rendered graphics and instructions need to be completed quickly to ensure real-time performance and a good gaming experience. This requires the game server, especially the video rendering service, to be deployed closer to the users.
- In smart home and smart city scenarios, IoT devices like cameras do not have powerful computing capability. Therefore, some AI tasks need to be executed on servers, such as recognitions of voice, image, and video. A significant transmission delay will be introduced if tasks are dispatched to servers in a cloud data centre. As a result, the total processing time of AI tasks will be high, and as a consequence, the AI identification speed will not be fast enough. In addition, the streaming media from many devices to the data centre will put significant bandwidth pressure on the network. Therefore, the AI service needs to be deployed closer to users and devices.
- For the Internet of Vehicles (IoV) scenario, vehicles need to interact with service platforms or roadside service endpoints to improve intelligent driving, participate in road-network monitoring, and do traffic scheduling. The service function also needs to be deployed closer to the vehicle to quickly and accurately control the driving status of the vehicle.

ETSI GS MEC 002 [i.117] describes more scenarios.

4.3.2 Requirement

Future services have an increasingly strong demand for edge computing. Edge computing is implemented by multiple distributed edge sites located closer to the users. ETSI GS MEC 001 [i.116] defines edge computing and extends it from mobile edge computing to Multi-access Edge Computing (MEC). ETSI also has a series of specifications for edge computing requirements, architecture, best practices, and so on.

Based on the requirements of various new scenarios, edge computing needs the following features:

- Low latency. Tasks of applications that require a faster response or processing speed are dispatched to edge sites instead of cloud data centres. Edge computing is mandatory to achieve real-time performance for some scenarios.
- Low transmission cost. Flows of applications that stream videos or upload voice samples and images are frequently steered, terminated, and processed by edge sites. That will avoid resource competition on the higher network levels.
- Coordinated computing resources. The computing resources and capabilities of edge sites are much less than those of data centres due to the limited space of equipment rooms. Therefore, some edge sites may be saturated easily during peak hours, while other sites may be idle or with a light load. As a result, the response time for users in hotspot areas becomes high and service quality decreases. There is a need for a mechanism to balance workload between neighbouring edge computing sites to guarantee service performance during peak hours. The sessions of the requests need to stick to the same edge site once the user's requests are dispatched to the certain edge site.
- Mobility. The user equipment is constantly moving in some scenarios, for example, in IoV. Therefore, mobility is an important feature of edge computing. According to ETSI GS MEC 003 [i.118], there are three mobility scenarios for applications: application state relocation, application instance relocation, and no relocation. The network needs to consider how to maintain users' sessions during the user movement.
- Privacy protection. Campus, enterprise, and industrial internet scenarios require local storage and processing of enterprise data, so edge computing resources need to be deployed within the geographic scope of their campus. The data communication between the edge site and the external network also needs to be strictly controlled to prevent production data and trade secret leakage.

4.3.3 Conclusion

Clause 4.2 already highlighted the benefit of IPv6 in supporting the service scenarios that require computing tasks to be moved from the user equipment to servers and offloaded from cloud data centres to edge sites. Distributing both cache and computing functions at the network edge will greatly decrease the response time and bandwidth consumption of these services and help to improve the battery efficiency of user equipment, thus improving the user experience.

4.4 Next generation IPv6-based transport infrastructure

4.4.1 Background

IP infrastructure had some well-known requirements for many years that are still important: fixed and mobile convergence, all types of end-to-end services over a single network, high security, cost optimization, and others. New services (like 5G, cloud, IoT) are demanding new quality and functionality from IP transport:

- Deterministic quality through resource reservations in addition to isolation for VPN services. It is typically achieved by "slicing". SLA compliance is becoming mandatory, especially for low latency.
- The high level of automation for service and assurance that is typically resolved by SDN technology. It may be augmented by artificial intelligence, intend-based networking, and model-driven architecture.
- The requirement to change the architecture to move subscriber's termination and applications close to a subscriber, which is typically achieved by "distributed edge".
- Virtualization for complex networking services. It is typically associated with NFV technology.

- Strict time synchronization (with phase) for 5G services.
- Ultra-higher bandwidth driven primarily by residential and mobile broadband.
- Security requirements are growing together with network importance.
- Automation is becoming more challenging for distributed environment.
- All of the above at the same or even lower cost. It is especially important to control operational expenses because it is 3 times bigger than capital expenses.

Additionally, it is important to mention a general trend to move some complexity from the control plane to the data plane that gives a positive synergy effect. It has been discovered that the modern data plane is much more capable and would not be affected by a few additional headers, while the control plane may be greatly simplified as a result. The market has several new technologies that greatly improve functionality, like SRv6, BIER, EVPN to specifically satisfy new transport requirements.

4.4.2 New approaches

Big networks are typically split into core, aggregation, access, and platforms domains for scalability, functional, and administrative reasons. From an architectural perspective, the challenge is to integrate all these domains and provide end-to-end services.

IPv6 comes to the infrastructure data plane in synergy with new transport technologies such as SRv6, BIER over IPv6. SRv6 and BIER are new concepts of traffic forwarding when information about the next hop is populated in the packet, not in the forwarding table of every device on the traffic path. It is a more scalable and feature-rich approach.

In addition, modern networks' operations require a solid control plane and network management based on Software Defined Network (SDN).IPv6 is the prerequisite to have a flexible, manageable, and optimizable data plane:

- Dynamic tunnels based on SRv6 may use the Flexible Algorithm for path optimization through the network. The IGP Flexible Algorithm is currently an IETF draft-ietf-lsr-flex-algo [i.73].
- Predefined SRv6 tunnels may be deployed, managed, and optimized by the SDN platform, accounting for resource constraints such as IGP cost, latency, jitter, or link quality.
- Data plane readiness depends on the particular vendor implementation, but, in general, all vendors have made good progress in the last 4 years to support SRv6 functionality. This may need a bigger parsing buffer in ASICs and network processors. Older platforms may have gaps - they need a check on a project-by-project basis.

A new concept of resources reservation (i.e. slicing) may be needed for many use cases:

- The main concept of slicing is to abstract physical network resources into logical network resources that reserved or assigned, based on the business needs, to a service or a customer.
- There are two principally different types of slicing: hardware-based ("hard") and software-based ("soft") slicing.
- Hardware slicing is based on physical segregation using entire physical interfaces or sub-interface. FlexE (Flexible Ethernet) is an example of a new technology that provides the ability to bundle physical interfaces or split these bundles into more granular sub-interfaces. The standard granularity is in the range of Gbps.
- Software slicing is based on some QoS mechanisms. The principle is similar to hardware slicing but in this case, it is logical separation. The granularity is not restricted to particular bandwidth.
- Slicing is a current hot topic in the industry. Physical Layer mechanisms are clear and specified, as in OIF Flexible Ethernet 2.1 Implementation Agreement [i.46]. Many topics are still in discussion in several IETF working groups:
  - TEAS WG (framework);
  - SPRING WG;

  - 6MAN WG;
  - MPLS WG (data plane);
  - LSR WG; and
  - IDR WG (control plane).

- IETF is also working on the terminology related to slicing. The emerging definition of "IETF Network Slice" addresses the instantiation of slicing in a transport network.
- Among the different solutions, it is worth mentioning draft-ietf-teas-enhanced-vpn [i.64], which describes the approach to realize network slicing as enhanced VPN services. Several specifications related to enhanced VPN and Virtual Transport Network (VTN, which is the virtual underlay of enhanced VPN services) are currently in progress. But it may still take another 2 years to settle the standardization process for slicing in the IETF.

IPv6 fully integrates the QoS mechanisms currently in use (e.g. differentiated services, integrated services, and hierarchical scheduling). Innovative QoS-related ideas based on IPv6 extension headers, like carrying packet's time requirements, are still under discussion.

A solid control plane is needed for routing and the control information exchange in the network. This goal is achieved by:

- A robust IGP. IS-IS is recommended for service providers as it is IP agnostic and capable of supporting dual-stack in a transition phase. IS-IS offers good scalability, protocol extensibility, and flexibility. Some IS-IS implementations may need additional features to support SR or SRv6.
- A solid BGP. BGP has been considerably extended recently with the primary target for automation and rich inter-domain functionality.
- IGP and BGP have been refreshed recently (primarily by the efforts of IETF) and are ready even for the latest functionalities like SRv6, BGP policy management, and BGP-LS topology collection, unless for the gaps highlighted in the following.

SDN is the basis for advanced automation. Only SDN provides an automatic closed-loop reaction to network events. SDN platform may have some functionalities still in development:

- The SDN platform needs to be aware of the network topology and status. It is possible to collect it by BGP-LS, PCEP, and YANG. BGP-LS is the most developed currently. BGP-LS still has minor updates for the latest extensions of IS-IS. PCEP is following closely. YANG needs many models that are still in long-term development.
- The SDN platform needs to provide data path configuration to routers, it is typically done by BGP policy and NETCONF. BGP has been extended with new address families (SAFI) and new communities - see IETF draft-ietf-idr-segment-routing-te-policy [i.51] for deploying the SRv6 Traffic Engineering policies.
- SDN is developed in association with comprehensive data information models, which are nowadays expressed in YANG semantic. YANG is considered the main standard for describing data models in the network domain. The key YANG models (like IETF RFC 8466 [i.71] for "L2 service model" or IETF RFC 8299 [i.72] for "L3 service model") have been ready for a couple of years, but hundreds of other YANG models are in development now by many standard bodies.
- SDN's multi-vendor interoperability aspects are still to be addressed because the list of YANG models needed for rich functionality is very long and private extensions are used for vendor-based configuration.

BGP EVPN is a proven technology, even if more features are still under development to offer a more complete functionality. It already offers the opportunity to build emulated L1 wire, L2 Ethernet LAN services, or L3 services. It is the obvious replacement for VPLS and pseudo-wire and maybe the replacement for L3 VPN in some scenarios. It is primarily developed in IETF BESS WG. The current BESS agenda is still very packed with EVPN-related activities but the majority of activities are for some very advanced EVPN functions that are usually optional.

4.4.3 Conclusion

IP transport infrastructure is at a big transformation point. New service and business requirements come to the market with many innovations at the same time. The transformation process has been active for many years. The majority of functions are standardized and implemented by most vendors. Some new and advanced functionality related to SRv6, Slicing, and SDN is still in active development. It does not create a problem for ongoing migration towards IPv6 and on top technologies since it will be possible to smoothly add new functionalities, readily getting the related benefits.

4.5 Research and Education networks

4.5.1 Background

A National Research and Education Network (NREN) is a specialized internet service provider dedicated to supporting the needs of the research and education communities within a country. The regional research and education networks are connected to NREN and provide services to campus networks. The NRENs are connected to form global research and education networks through bilateral or multilateral agreements. NREN business processes are becoming more dependent on networking that creates the bigger pressure for the network quality. Multimedia services are demanding low latency and deterministic quality. Recent shift to spend more time on remote or home location pressure ubiquitous connectivity and high bandwidth. Constantly growing infrastructure makes automation more important.

The network design differs from case to case. However, they are all similar by the support of broadband infrastructure that provides dedicated links to the member institutes and offers over-the-top services for academia. The concept of a specialized network is not exclusive to academia, VPNs are well-known services for service providers. Nevertheless, among other carrier's networks, NRENs converge between VPN service and the normal commercial Internet because researchers are not isolated from the other world. This dual nature of NRENs necessitates special treatment in network design and routing setup. Therefore, to be more flexible and efficient NRENs mostly use public IP addresses.

IPv4 does not satisfy the need of NRENs for end-to-end connectivity because the shortage in IPv4 addresses imposes multiple levels of Network Address Translation (NAT) that is typically called Carrier-Grade NAT (CGN) solution. Hence, NRENs always had strong motivation to use IPv6. Many NRENs have finished the first trial on IPv6 and have deployed IPv6 networks afterward, contributing to technologies and best current practices. The 6Bone is the first global IPv6 backbone based on IPv6 over IPv4 tunnels back in the 1990s, its major players are research and education institutions. Since then, Internet2 in 1997, GEANT2 in 2004, and APAN in 1997 have developed dual-stack networks, while CERNET has been developing IPv6-only backbone since 2004.

The backbones of the research and education network are providing native IPv6 services. BGP is used for the IPv6 route distribution at upstream and downstream peering. However, some regional networks are connected via static routes for better aggregation.

Most of the regional networks are dual-stack and have the following features:

- They typically have /48 IPv6 address blocks, some big ones have /32 IPv6 address block.
- IPv6 addresses assigned to end systems are either via SLAAC or via DHCPv6.
- The end systems, especially those connecting via WLAN, typically use a separate IPv6-only SSID.
- The DNS system is dual-stack and provides AAAA and A records for authoritative and recursive services via IPv4, IPv6, or both.
- While it would be preferable to have homogeneous access control policy for both IPv6 and IPv4, some differences may exist. For example, DHCPv4 is used for centralized IPv4 address configuration for the end systems. In the case of IPv6, the end system may determine its address dynamically by SLAAC, bringing differences to protocol security and traceability.

4.5.2 IPv6 for Research and Education Networks

Key benefits of using IPv6 in Research and Education Networks are:

- Research and Education Networks are natural testbeds for new cutting-edge technology, including IPv6 itself and its future enhanced innovations.

- IPv6 has huge address space and provides better end-to-end address transparency.
- IPv6 provides more scalable, secure, and controllable multicast for one-to-many applications. IPv6-enabled information resources (like IPTV and online courses) will attract students to use IPv6 services and be familiar with IPv6 technologies. This will promote the migration to a full IPv6 Internet.
- IPv6 is the very natural and possibly the only choice for some IoT devices because of IPv4 address shortage.
- IPv6 has technology advantages compared to IPv4. For example:
    1) the "flow label" in the IPv6 header permits to simplify headers parsing on the transit nodes that improves throughput; or
    2) the source route functionality provided by SRv6 used to improve forwarding paths management.
- Last, but not least, the research and education networks are used to train the future native IPv6 engineers, scientists, and enterprisers that create a positive outcome for decades.

4.5.3 Adopting IPv6 in Research and Education Networks

4.5.3.1 Strength and Opportunity

NREN has its autonomy to accelerate the migration to IPv6 to satisfy its member's needs. Moreover, the not-for-profit nature of NREN's model alleviates the business case needed to justify the migration expenses. It is possible to minimize the migration budget to IPv6 by planning it with other hardware and software upgrades in the normal depreciation cycle. Hence, inventory control is to be conducted to evaluate the IPv6 enabled resources and plan for gradual replacement when it comes to its useful lifecycle. However, if a budget is secure then the migration process to IPv6 may be accelerated that would be quickly reflected in the IPv6 traffic transited.

Video services, including online courses, educational video clips, and IPTV streams are the major contributors to the IPv6 traffic on NREN. Collaborating with broadcasting companies, the CERNET2 IPv6 video system provides TV channels as an innovative use case. IPv6-only streaming and stateless IPv4/IPv6 translation systems have been installed and become valuable for regional network users.

NRENs are not in competition with local ISP. Instead, it gets support from all ICT stakeholders as it considers one of the empowering elements of quality education. Such support from commercial Internet providers makes a ground for a good price of Internet transit and national connectivity. The appearance of Internet Exchange Points (IXPs) contributes further to reduce the local transit costs. Furthermore, the typical support received from the local governments in terms of facilitating license, intergovernmental permits, and the allocated funds usually dedicated to Research and Education (R&E) institutes allow the NREN to sustain in a cost-effective operation.

The regional and international firms support NRENs too. For example, RIRs provide huge price cuts and other alleviates in requirements of obtaining ASN and IP resources.

In history, a lot of start-ups in the IP domain came up from research and educational networks. The research and educational network provides an important playground for new start-ups in the IPv6 age.

4.5.3.2 Challenges

There are some challenges of IPv6 adoption in research and education networks despite that NRENs are early players of IPv6 and the IPv6 penetrations are relatively high compare to other networks.

Technical challenges include:

- Academic regional networks are usually multihomed for connecting to NRENs and carriers networks. In IPv4, it is reasonable to use private IPv4 addresses in the internal network and use NAT to mapping private addresses to the public address of different upper streams. It is especially the practice if the regional network does not have PI (provider-independent) address blocks. However, there is no NAT in IPv6, so this approach does not work and regional networks need new IPv6 multi-homing solutions that may be very complex.

- Regional networks provide different services with different requirements, for example, for their users to access the Internet, to provide connectivity to internal and public servers, to support high-performance applications, and IoT devices. Since IPv6 has a much larger address space and new features, the network-layer slicing in IPv6 may be different compared to IPv4. Hence, the opportunities and challenges will coexist when using a new scheme of network-layer slicing by IPv6 (like SRv6).

- Due to the COVID-19 Pandemic worldwide from 2020, the way of traditional education has dramatically changed. The students and professors at home have to use public networks with video conferencing to access library resources and labs inside regional networks and NRENs. In this scenario, VPN services over IPv6 may be a potential problem because of legacy video conferencing systems readiness to IPv6.

- Multicast is a fundamental feature of IPv6. A lot of academic applications are based on one-to-many and many-to-many communications, so, IPv6 multicast is supposed to be a mandatory infrastructure requirement for NREN. However, deploying a scalable, manageable, and secure IPv6 multicast service in the NREN backbone is a challenge that needs careful configuration and upgrade of many components.

Organizational challenges include:

- Due to budget issue, the high-quality IPv6-enabled equipment and IPv6 engineers are rare resources in developing countries. Therefore, the IPv6 adopting rate in research and education networks in these countries is still low.

- Making a balance between operational security and innovation is not easy for IPv6.

4.5.4 Conclusion

IPv6 is the only technology to support the growth of the global Internet in the predictable future. The research and education networks have been playing a very important role in IPv6 development for a long time. More and more new technologies are expected to grow from the academic field. It will benefit the whole Industry.
The majority of challenges and gaps are non-technological. It is budget and other types of resources.

4.6 NB-IoT & IPv6

4.6.1 Background

Narrow-band IoT (NB-IoT) is a technological standard developed and supported by 3GPP for devices with low bandwidth and power consumption requirements used in cellular networks. It has widely-known benefits (improved coverage, energy efficiency, massive connections) in comparison to other LPWAN technologies that stimulated NB-IoT adoption over the last few years. It has many use-cases ranging from automotive to manufacturing and from agriculture to consumer electronics that create expectations of NB-IoT penetration to grow several folds in this decade. The latest 3GPP releases aim to optimize NB-IoT architecture for wider adoption and address the challenges faced during deployments.

As per 3GPP specifications, there are two major options for User Equipment (UE) connectivity:

- With Packet Data Network - Traditional way for all UEs to attach to the Evolved Packet System (EPS).

- Without Packet Data Network - UEs remain unconnected for long periods, and data transmission is infrequent. Only SMS service is available for data transmission.

In the first option, With Packet Data Network, data connectivity options include:

- IP over Control Plane - Both UDP and TCP are supported.

- IP over User Plane - Both UDP and TCP are supported.

- Non-IP over Control Plane.

- Non-IP over User Plane.

This clause of NB-IoT & IPv6 primarily aims to address data connectivity mechanisms using IP over User Plane or Control Plane when UEs are connected to the EPS with Packet Data Network.

4.6.2 IPv6 Benefits for NB-IoT Networks

Most NB-IoT deployments consider that using an IP-based device communication protocol, i.e. standard TCP/IP, introduces significant overheads, which may lead to increased costs for the end-user. These overheads may also lead to increased bandwidth requirements and power consumption of the end devices, which is quite contrary to what NB-IoT was originally designed for. However, lightweight IPv6-based protocol stacks may significantly reduce these overheads while leveraging the inherent benefits of IPv6 and provide secure end-to-end connectivity.

Some key benefits of using IPv6 in NB-IoT include:

- Massive IoT scale support. IPv6 has a large address space to accommodate the ever-growing number of end-devices without the need to use any address translation techniques.
- Autoconfiguration mechanism to support plug-n-play of NB-IoT sensors.
- Security features of IPv6 answer the growing end-to-end security concerns for NB-IoT devices. For example, additional privacy of end-nodes specified in IETF RFC 8981 [i.7] and IETF RFC 7217 [i.52].
- IPv6 Mobility support is important to address the roaming challenges in NB-IoT.
- IPv6 features like multicast, anycast, and address scope bring new services to the NB-IoT domain.
- Address architecture simplification is important for large-scale deployments like smart cities.

As most eNodeB devices already support IPv6, the only major requirement would be IPv6 support (with a limited stack) on the NB-IoT sensors. The end-to-end connectivity provided by IPv6 between the sensors and the application servers creates several opportunities ranging from new services to automated troubleshooting.

4.6.3 Challenges of Adopting IPv6 in NB-IoT

NB-IoT is designed to address situations where bandwidth requirements and power consumption of the end devices are constrained. The introduction of the basic IPv6 into NB-IoT introduces significant overhead that requests more bandwidth and increases power consumption. 40-byte IPv6 header (excluding the extension headers) does not seem an ideal solution for typical payload sizes in NB-IoT deployments that are in the range of 1-byte to 10-bytes. Additionally, to leverage the other benefits of IPv6, like security and mobility, more overheads would need to be incorporated as part of the corresponding extension headers involved.

While header compression techniques such as Robust Header Compression (ROHC) and 6LoWPAN header compression have tried to resolve this issue for LPWANs in general, they do not specifically cater to the needs of NB-IoT networks due to other limitations. ROHC sends uncompressed packets for initialization and resynchronization and does not support CoAP (IETF RFC 7252 [i.16]). Whereas a 6LoWPAN header compression may reduce a typical 48-byte IPv6-UDP header to 7-bytes, it is better suited for IEEE 802.15.4 [i.58] networks in which the frame payload is much higher than in NB-IoT deployments (~100-bytes). Moreover, the 6LoWPAN fragmentation header produces additional overheads and is primarily designed for IEEE 802.15.4 [i.58] mesh networks, whereas NB-IoT uses a star topology.

Static Context Header Compression (SCHC), see IETF RFC 8724 [i.66] looks promising for NB-IoT, as it not just defines a better scheme for header compression but also provides an efficient fragmentation mechanism. Typical NB-IoT deployment consists of static applications on the end devices that episodically transmit consistent data. SCHC uses this knowledge for efficiently compressing a 48-byte IPv6/UDP Header to 3-bytes. The IETF's LPWAN WG is currently working to develop an SCHC standard specified in the context of NB-IoT. The actual deployment challenges faced by adopting SCHC for IPv6 in NB-IoT will only be known later because the standard is currently in a draft state. Additionally, the generic framework for SCHC does not define any specific rules for compressing IPv6 Extension Headers. This feature may be required to leverage all the benefits of IPv6, including security and mobility.

Thus, there is a need for relevant standards to be evolved and matured to address NB-IoT challenges. An additional mechanism may also be required as not all features of IPv6 (such as Routing scheme, Neighbour discovery, etc.) are needed in the NB-IoT networks.

There are some non-technical challenges in NB-IoT adoption:

- Knowledge and training on IPv6.
- Legacy systems upgrade.

- Regulatory requirements -considering NB-IoT operates in licensed bands.
- New IPv6 tools are required during the development and testing of NB-IoT solutions.

However, addressing these challenges will be easy over time of IPv6 adoption in NB-IoT networks.

4.6.4 Conclusion

As a valuable part of 5G technology, NB-IoT will continue to evolve in 3GPP's specifications and co-exist with other 5G use-cases. IPv6 end-to-end secure connectivity, along with other features such as autoconfiguration and multicast, enables new services and use-cases for NB-IoT. The challenges discussed above may be addressed by evolving standards to meet the current and future deployment scenarios' requirements and by sharing IPv6 best practices.

4.7 New industry challenges conclusion

New scenarios keep emerging in the 5G and Cloud era, such as convergence of cloud and network, edge computing, 5G transport, etc. These scenarios impose highly demanding requirements on the IPv6 networks. Some IPv6 enhanced innovations are enabling these new scenarios, including:

- Ubiquitous connectivity: Networks will be well integrated with multiple cloud data centres because of cloud and network convergence. Enterprises will be able to access the services deployed among multiple clouds via a single access line and dynamically select which clouds to access and the related SLA. 5G transport infrastructure will help devices distributed in various places to have access everywhere, which is especially important for IoT scenarios. The research and education networks also request a large public address space for global connectivity.
- Ultra-high bandwidth: Increasing network traffic poses a high network bandwidth requirement, especially for access to the cloud and distributed edge computing. It will consume more bandwidth in the data centres, metro, and bearer networks, the networks may need support 400GE to cover the increasing traffic. The general rise in computing and transmission performance of commercial technologies creates more pressure for infrastructure.
- Deterministic quality: Some services of vertical industries need networks to provide strict SLA assurance, e.g. 1ms bounded jitter and very low packet loss (10^{-6}). The distributed edge computing deployed in the industry campus and IPv6 transport network for 5G are promising to achieve deterministic quality.
- Low latency: Many new services demand networks to provide low latency quality. IoT, virtual and augmented reality services may need latency as low as 1ms. Some new technologies like distributed edge computing and 5G low latency communication are developed to satisfy the latency requirements of these new services.
- Automation: New services request networks to support fast service provisioning and provide faster and more reliable self-healing capabilities to cope with possible network faults in several minutes. The network needs to be able to recover from failures in a short time to reduce the impact on services. Constantly increasing network scale and complexity may not be sustainable without a new level of automation. IPv6 transport infrastructure for 5G pays much more attention to operation and management automation in all the phases of the lifecycle.
- Security: Network security is becoming more and more important. All scenarios discussed in this clause, including research and education network, NB-IoT, 5G transport, cloud, pay great attention to security. IPv6 has a special role here because of its extensibility and flexibility. AI technologies are being applied to networking along with automation which enable the identification and prevention of complex network threats in several minutes.

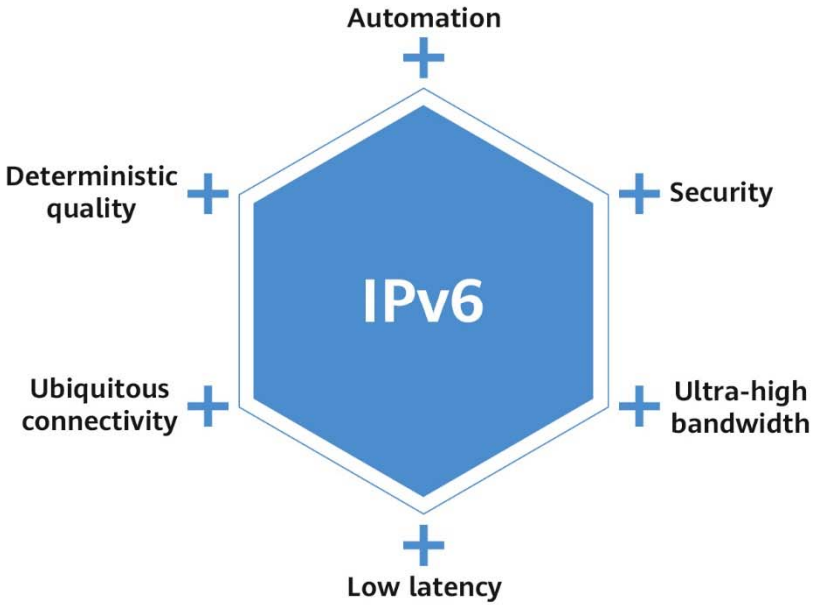


Figure 4: Six key features of IPv6 Enhanced Innovations

5 Existing solutions and gaps

5.1 IPv6 differences from IPv4

The primary driver for IPv6 deployments is the scalability of ever-expanding networks because the number of attached devices has grown exponentially. The IPv4 protocol was developed when the costs of end-user equipment and computers ranged from tens of thousands to millions of dollars and therefore the approximately 4 billion address was sufficient for the Internet to expand for many years. To extend the life of IPv4 four sets of IPv4 address blocks were reserved for use as private space within an organization. This allows the reuse of these address blocks for an unlimited number of organizations. IPv4 space size was improved by the wide adoption of Network Address Translation (NAT) but not enough. The number of devices attached to the Internet grew exponentially as the cost of computers dropped to under 40 dollars and intelligent cellular phones became very popular. Since then bigger addressing of IPv6 has become mandatory.

Major advantages of the IPv6 protocol are:

- IPv6 Addresses are 128 bits in length with 64 bits for the network and subnet addresses and 64 bits of a fixed size device address on a default subnet. This amounts to 16 billions of billions networks and the same amount of devices per network.
- IPv6 has a fixed packet header for easy hardware-based processing in routers or switches.
- Extension headers support functionality that was not retained from the IPv4 packet header, as well as the capability to expand protocol functionality transparently to old non-compatible network devices.
- Multiple address assignment methods (DHCP ad SLAAC) efficiently support large numbers of devices with a diverse set of use cases.

To illustrate how IPv6 fulfils its design purpose and as a foundation for the following more complex protocols discussed later. The IPv4 packet header is variable length and the IPv6 packet header is fixed length with potentially several extension headers that are illustrated below. The IPv6 protocol was designed to be scalable and the IPv6 header was made a fixed length so that its processing occurs in hardware to speed up switching and routing operations. The header checksum has been removed as the need for the field has diminished because of digital and optical links which are very robust.

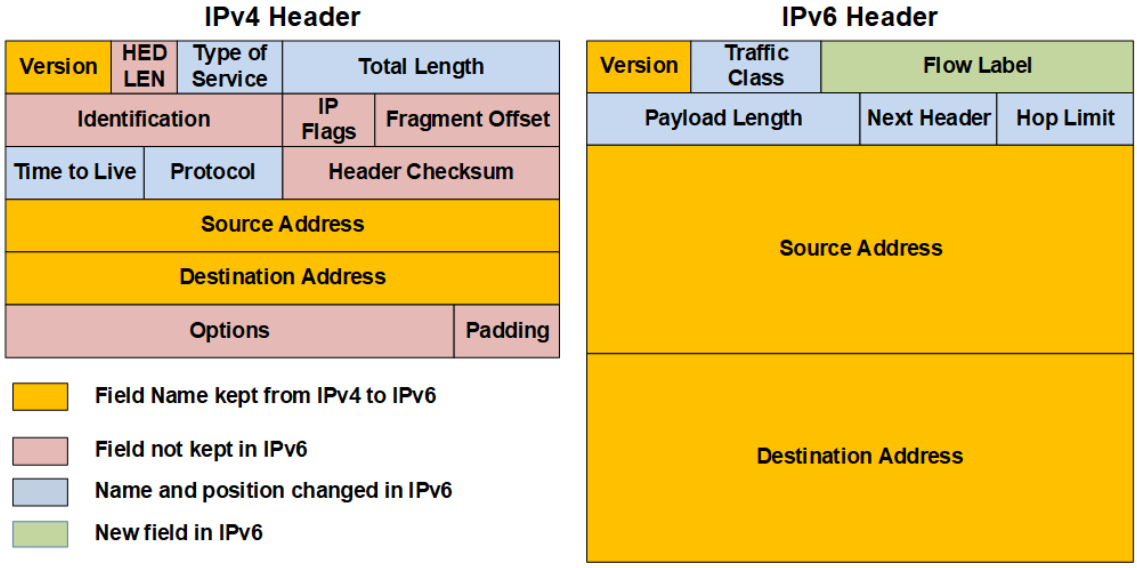


Figure 5: IPv4 and IPv6 Packet Header

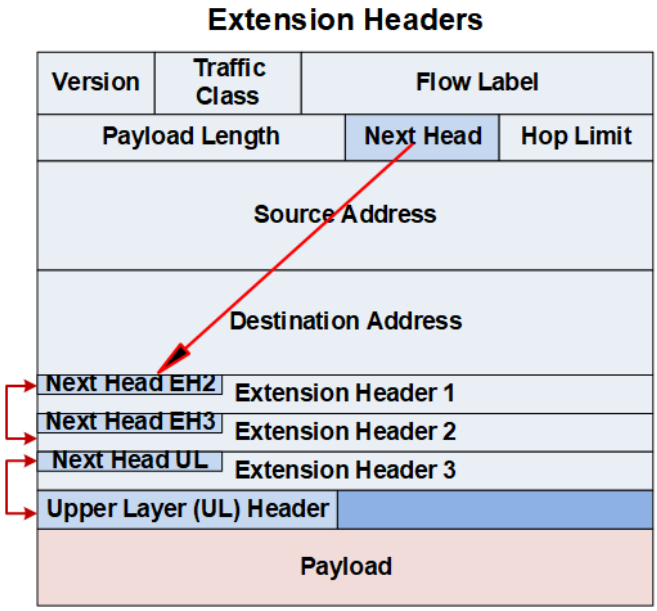
The options field in the IPv4 header is variable length and used to convey additional information on the packet or on the way to process it. Routers, unless instructed otherwise, process the options in the IPv4 header. The processing of most IPv4 header options pushes the packet into the slow firmware processing path leading to a forwarding performance hit.

IPv4 Options perform a very important role in the overall IP protocol operation therefore the capabilities had to be preserved in IPv6. On the other hand, the impact of IPv4 options on the protocol performance was taken into consideration in the development of IPv6 protocol and packet headers. The functionality of options is removed from the main header and implemented through a set of additional headers called extension headers. The main header remains fixed in size (40 bytes) while customized EHs are added as needed. Figure 6 shows how the headers are linked together in an IPv6 packet as well as the types and sequence of EHs. Extension Headers provide a standard way of introducing new functionality in the protocol over time and older nodes that do not support the new EHs will ignore them.

Extension headers are an intrinsic part of the IPv6 protocol and they support some fundamental functions and services. The following is a list of several circumstances where EHs are commonly used:

- Hop-by-Hop EH when used with the Router Alert option, is an integral part of the operation of IPv6 Multicast through Multicast Listener Discovery (MLD) and RSVP for IPv6.
- Destination EH is used in IPv6 Mobility and supports other applications.
- Routing EH is used in IPv6 Mobility and in Source Routing, which will be discussed later in the present document.
- Fragmentation EH is the method used to fragment packets in the IPv6 protocol. Traffic sources control fragmentation as the routers and interim hop devices are not allowed to do fragmentation, as a scalability factor.
- Mobility EH is used to support Mobile IPv6 services.

The detailed discussion of EHs used in IPv6 protocols is beyond the scope of the present document.



Order	Header Type	Next Header Code
1	Basic IPv6 Header	-
2	Hop-by-Hop Options	0
3	Destination Options with Route	60
4	Routing Header	43
5	Fragment Header	44
6	Authentication Header	51
7	Encapsulation Security Payload	50
8	Destination Options	60
9	Mobility Header	135
	No next header	59
Upper Layer	TCP	6
Upper Layer	UDP	17
Upper Layer	ICMPv6	58

Figure 6: Packet and Extension Headers

To support almost unlimited scalability, the IPv6 Address is 128 bits of which the first 64 bits are the network and subnetwork identifier/address and the lower half 64 bits are the host address.

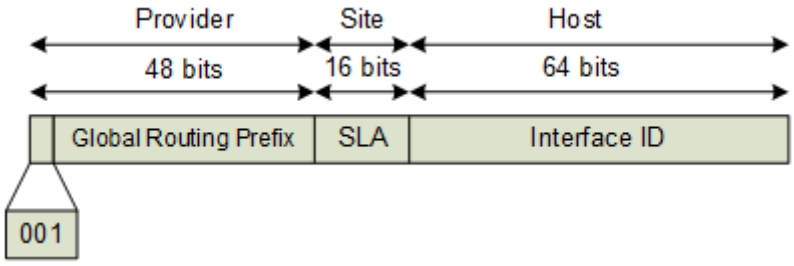


Figure 7: Typical IPv6 Address Format

The typical Global Unicast Address format is shown in figure 7. It has a 48-bit Global Routing Prefix, a 16-bit subnet identifier, and a 64-bit Interface ID.

The first and evident difference is IP address size. It has a few consequences:

- All official registries (RIR) and many RFCs are calling not to save IPv6 addresses in the same way as it was for IPv4. It makes sense to have an address plan with sufficient capacity for any enhancements in the future. For example, give a subscriber subnet with /56 size, even for a mobile subscriber. It is easy to obtain proper address space from RIR and get as much address space as needed from a long-term perspective.
- It is mandatory to reserve 64 bits for any link where stateless address assignment (SLAAC) would be used, while DHCP link is possible to be of any size.
- It is recommended for privacy reasons to use random bits in the subnet and the interface ID part. It creates addresses difficult to memorize and use manually. It is becoming more important to have DNS and automated IP Address Management (IPAM) even for infrastructure addresses to simplify troubleshooting and maintenance.

The network administrator may not receive the address space from RIR (Provider-Independent) but uses carrier address space (Provider-Aggregatable). Multiple addresses from different carriers are probable on every host if the site has redundancy connections to different carriers. Multi-prefix multihome is a very tricky configuration that is not yet fully supported in the industry because the host needs:

- To choose the proper source IP that would permit to reach the destination resource (clients from other networks would not be served by the walled garden zone of carrier).

- To send the packet in the direction of the respective carrier (carriers expect packets only from their clients, packets with other source addresses would be filtered out by RPF security check).

The solution for the multi-prefix multihome environment is proposed in IETF RFC 8028 [i.5], but it is not yet widely implemented. The primary problems are:

1) to choose proper source address (from many available on the host);

2) to choose proper router on the link that is capable to send the packet to a particular carrier delegated address space for this source address.

   IPv4 did solve a similar problem with NAT at the carrier border that capable to translate packets to proper address space at the last hop just before the packet would be sent to the carrier. But NAT is discouraged in IPv6.

IPv6 Link Layer is significantly enhanced and consolidated from the IPv4 implementation, which is specified in IETF RFC 4861 [i.1] and IETF RFC 4862 [i.2]:

- IPv6 assumes and promotes the usage of many GUA or ULA IP addresses for the link (from different prefixes, temporarily addresses, ephemeral addresses). Link Local Address (LLA) is mandatory. Global Unicast Address (GUA) is very often. Unique Local Address (ULA) is popular for communication inside the administrative domain. Moreover, it is very common in the IPv6 world for a host to have many GUA temporary addresses for privacy extension (IETF RFC 8981 [i.7]).

- Such a liberal IPv6 address usage makes it impractical to filter one host on the subnet because the host may get many IPv6 addresses in a very dynamic way. Filtering has to be dynamic or based on the subnet.

- IPv6 link may have many prefixes on the link (announced by the router) and the host may assign many addresses to each prefix at the same time. Many prefixes may be needed because of different address types (GUA, ULA), network topology transition, multi-homing (addresses from different carriers), or security concerns (some prefixes are filtered inside or outside). IETF RFC 6724 [i.8] regulates the choice of source and destination IP addresses from many candidates on the host.

- It is not possible to suggest local link connectivity looking just at the prefix and the mask. A router does announce a special flag for every prefix (L=Local) to inform hosts that they are capable to communicate directly on this prefix. The host sends traffic to the router for prefixes that are not explicitly announced as local. Only LLA is considered always local.

- The router may redirect a host inside the link, even for addresses that do not belong to any prefix on the link. The host may cache such information and treat such address as "on the link" from now on.

- IPv6 has Stateless Address Autoconfiguration (SLAAC) in addition to DHCP that permits the host to choose an IP address from the router announced prefixes then check that this address is not a duplicate. The Duplicate Address Detection (DAD) check is typically negative because random numbers collision has a low probability in the 2^64 address space.

- Address resolution protocol (ND) relies on multicast instead of broadcast for many protocol messages. It is very important for layer 2 to properly support or optimally emulate multicast.

Security issues are similar for the first hop in IPv4 and IPv6. DHCPv6-Shield [i.9] and RA-Guard [i.10] are needed to filter out rogue DHCP server or router advertisement packets.

There is a separate problem caused by the huge address space of every subnet (2^64) that it is no longer possible to scan for active hosts. It is a disadvantage for inventory systems but an advantage for security because reconnaissance is more difficult for the attacker.

IPv6 values any-to-any connectivity. Hence, Network Address Translation (NAT) from IPv6 to IPv6 is discouraged. A firewall is recommended for security filtering, because it is more secure than NAT.

IPv6 has a long list of well-developed NAT implementations to translate IPv4 into IPv6 and back (464XLAT, MAP-E/T, DS-Lite, lw4o6) that allows migration to an IPv6-only network but still has access to all IPv4 resources on the Internet. Stateful NAT64 (IETF RFC 6146 [i.35]) or stateless NAT64 (IETF RFC 6052 [i.36], IETF RFC 7915 [i.37]) are employed in order to reach the global IPv4 Internet.

All modern host's operating systems have IPv6 preference. A host would prefer IPv6 for communication if both IPv4 and IPv6 are available. It is defined in "Happy Eyeballs" (IETF RFC 6724 [i.8]) and implemented in the major operation systems for a long time.

IPv6-only servers are becoming more popular, especially for video content distribution. They use stateless translation technology to provide services for users from IPv4 Internet.

The last, but not the least, principal difference of IPv6 is the fragmentation only at the source node. Transit nodes are not permitted to fragment packets by IETF RFC 8200 [i.6]. It may create a problem for the case of tunnelling or the usage of IPv6 extension headers due to MTU oversized issues.

5.2 IP Link

5.2.1 IPv6 Link vs. IP Subnet vs. Layer-2 broadcast domain

Like IPv4, IPv6 has its concepts of IP Link and IP subnets, to denote a direct (Link-Layer) connectivity between Nodes and the smallest (atomic) aggregation of Nodes that may be injected in the routing system and aggregated again.

In contrast with IPv4, IPv6 defines link access methods that are all based on a Layer-3 abstraction called the link-local addresses. Services such as DHCP and address resolution are obtained in a media-independent fashion using unicast and multicast link-scoped IPv6 addresses. One may consider a link-local address as a Layer-3 abstraction of the MAC address. In practice, though, link-scoped IPv6 multicast transmissions are mapped into Link-Layer broadcasts for some L2 technologies, and the multicast property is lost in translation.

On wired media, the scope of IP links and subnets are may be confused with the physical broadcast domains because both are determined by the cabling, e.g. a serial cable or an Ethernet shared wire. Ethernet Bridging reinforces that illusion with a Link Layer broadcast domain that emulates a physical broadcast domain over the mesh of wires. But the difference shows on legacy Non-Broadcast Multi-Access (NBMA) networks such as ATM and Frame-Relay, on shared links, and newer types of NBMA networks such as radio and composite radio-wires networks. It also shows when private VLANs or Link Layer cryptography restrict the capability to read a frame to a subset of the inter-connected Nodes.

In the case of an IEEE 802.1Q [i.88] bridged domain, a switched fabric, a Mesh-Under LLN, and a Wi-Fi™ Infrastructure BSS, the IP Link extends beyond the physical broadcast domain to the emulated Link Layer broadcast domain. Relying on Multicast for the ND operation remains feasible but becomes highly detrimental to the unicast traffic and becomes less and less cost/energy-efficient and reliable as the network grows.

On Ethernet, an IP subnet is often congruent with an IP link because both are determined by the physical attachment to a shared wire or an IEEE 802.1Q [i.88] bridged domain. In that case, the connectivity over the link is both symmetric and any-to-any, the subnet may appear as on-link, and any Node may resolve a destination MAC address of any other Node directly using Classic IPv6 ND.

An IP link and an IP subnet are not always congruent. In the case of a shared link, individual subnets may each encompass only a subset of the Nodes connected to the link. But another model is possible, whereby routers federate the links between Nodes that belong to the subnet and the subnet extends beyond any of the federated links.

5.2.2 Decoupling the IPv6 Subnet from the Layer-2 domain to simplify hand-over

IPv6 defines the concept of a subnet for Global and Unique Local Addresses (GUA and ULA). All the addresses in a subnet share the same prefix, and by extension, a Node belongs to a subnet if it has an address in that subnet. IPv6 enables a Node to form multiple addresses, some temporary or ephemeral, and with particular attention paid to privacy. As opposed to IPv4, the IPv6 address allocation is very dynamic, and addresses may be formed and deprecated asynchronously to the Layer-2 state going up or down for mobile or sleeping nodes.

The IPv6 aggregation model relies on the property that a packet from the outside of a subnet may be routed to any router that belongs to the subnet and that this router will be able to either resolve the destination Link Layer address and deliver the packet or route the packet to the destination within the subnet using a host route. The subnet prefix is unique within a routing system; for ULAs, the routing system is typically a limited domain; whereas, for GUAs, it is the whole Internet. For that reason, it is sufficient to determine that an address that is formed from a subnet prefix is unique within the scope of that subnet to guarantee that it is globally unique within the whole routing system. Note that a subnet may become partitioned due to the loss of a wired or wireless link, so even that subnet-wide determination is not necessarily obvious.

Though it is possible to assign an IPv6 subnet when and where an IPv4 subnet is assigned, a /64 IPv6 subnet may in theory grow enormously larger, to the point that the protocols that rely on IPv6 multicast cause so-called broadcast storms. The excessive use of layer-3 multicast thus becomes the gating factor to the scalability of a subnet that is congruent with the layer-2 broadcast domain (more in clause 7.2).

To limit that effect, IPv6 ND advertises whether a subnet prefix is on-link versus not-on-link; if the prefix is on-link, then any Node may also resolve the Link Layer address of any destination in the same subnet and deliver the packet; but if the prefix is not on-link, then a host in the subnet that does not have a Neighbour Cache Entry (NCE) for the destination will also need to pass the packet to a router, more in IETF RFC 5942 [i.100].

This provides a way to reduce the propagation of Layer-2 broadcast while preserving a large single subnet, effectively forming a so-called Multi-Link Subnet (see figure 8). MLSN definition is in IETF RFC 8929 [i.92]. Each link in the MLSN, including a federating backbone if any, is its broadcast domain. A key property of MLSNs is that link-local unicast traffic, link-scope multicast, and traffic with a hop limit of 1 may not reach all the Nodes in the same subnet. This enables to scale the IPv6 subnet beyond the limitations of the local broadcast domain and avoid large broadcast storms, relying on routers to forward to the not on-link destination, but may produce unexpected behaviours in software that expects any member of a subnet to be reachable over an IP Link.

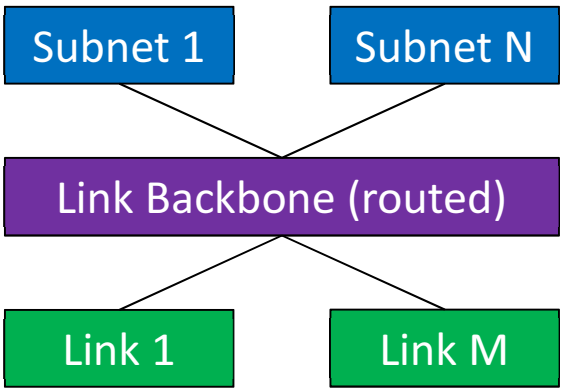


Figure 8: Multi-Link Subnet

5.2.3 ND and SLAAC unresolved problems

5.2.3.1 DAD, Address Lookup and Broadcast Storms

Classic IPv6 ND was defined at the time a LAN was a single Ethernet (yellow) wire, memory was so expensive that a node may only retain information of a few of its peers, processors were early RISC architecture that required word alignment to operate efficiently, and there was plenty of bandwidth versus CPU. With this in mind, it is natural that the IPv6 ND PDUs are aligned to 8 bytes, that the peer information is a cache as opposed to a stateful table, and that peers are discovered reactively, that is, on-demand, with a least recently used strategy to clean up the cache when it gets full. The issue in that model is that the peer discovery, when the subnet is on-link, is a broadcast; arguably this is a Layer-3 multicast but at Layer-2, it is mapped to a broadcast anyway.

In addition to the DHCP protocol that is evolved from IPv4, classic IPv6 ND protocol proposes a Stateless Address Auto-configuration IETF RFC 4862 [i.2] model that enables a host to select its addresses and claim them over the network. With SLAAC, a single Node may form many addresses but does not have a way to tell the network when its temporary ones are flushed. It results that a large amount of stale NCE state may remain in the network.

When all nodes in a network may form any address, any time, and using any method, duplicates may appear. Duplicate Address Detection (DAD) is thus a critical operation for SLAAC addresses. The IPv6 ND Neighbour Solicitation (NS) IETF RFC 4861 [i.1] message is used for DAD and address lookup when a Node moves or wakes up and reconnects to the wireless network. The NS message is targeted to a Solicited-Node Multicast Address (SNMA) IETF RFC 4291 [i.93] and, in theory, only reach a very small group of Nodes. But since IPv6 multicast messages are typically broadcast for Link Layer technologies, they consume enough bandwidth to cause a substantial cost and/or degradation to the unicast traffic service. And if the targeted node is a mobile node or a sleeping IoT node and is not reachable at this time, then the ND protocol will fail to obtain the address mapping or detect the duplication.

On radio access networks where the Link-layer emulates the wired broadcast operation, IP links between peers come and go as the individual physical broadcast domains of the transmitters meet and overlap not in the full mesh. The DAD operation may not work for the radio transmitter if that transmitter keeps meeting new peers on the go. Since broadcast may be unreliable over wireless media, DAD often fails to discover duplications - see IETF draft-yourtchenko-6man-dad-issues [i.102]. In practice, the fact that IPv6 addresses collision very rarely conflict is mostly attributable to the entropy of the 64-bit Interface IDs as opposed to the successful operation of the DAD mechanism.

5.2.3.2 Remote Denial of Service Attacks

A well-known weakness in the reactive approach used in classic ND (IETF RFC 4861 [i.1]) for on-link prefixes is how easily the ND cache may be attacked to:

1) flood the local network with broadcasts; and

2) deny the normal lookup operation by saturating the router cache or forcing it to throttle ND lookups to protect itself and the network.

Indeed, a packet coming from the inside or even outside the local network for an address that is not present in the ND cache of the access router will result in an address lookup by the router that is broadcasted over the local network. Also, the router is supposed to create a state and keep one message per address while the resolution is performed, the neighbour cache entry being typically maintained for several seconds until the router finally decides that the address is not reachable if such is the case.

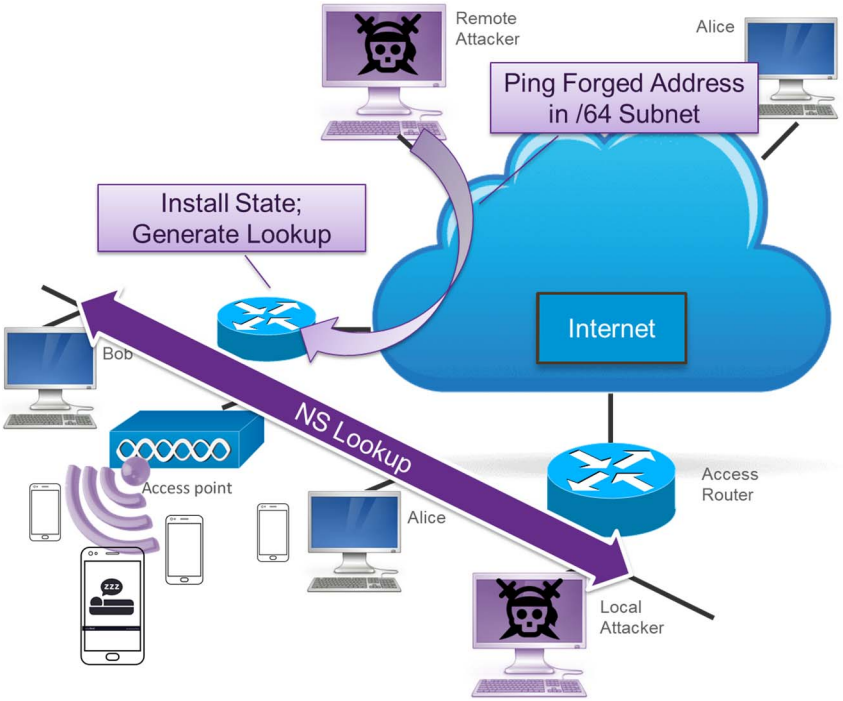


Figure 9: Remote Denial of Service Attack

The attack vector already exists with IPv4 but is a lot more severe with IPv6, because of the sheer number of possible addresses in a subnet. With an IPv4 /24, the router needs to build at most 256 entries and may implement protection to throttle the lookups per individual address, allowing legitimate addresses to operate normally. With IPv6, a remote attacker may generate messages (e.g. ping) to any of the possible 2^64 interface IDs in the attacked /64 subnet in any random order and at any rate. The router will typically implement protection such as a basic throttling mechanism, but the protection is global to the subnet, which means throttling legitimate lookups as well as rogue ones.

5.2.3.3 Captive Portals and Temporary Hiccups at connection time

When a node joins a network, it will typically assert its Internet connectivity to discover captive portals by sending a series of HTTP requests to various internet services in a rapid sequence. This is used as a trigger by captive portals in a hotel to prompt a login page that a mobile device will present to the user to obtain access.

This behaviour and the reactive lookup causes side effects that are observable in any network, regardless of whether there is a captive portal or not. Routers typically store at most one message per address being resolved during the lookup phase to limit the DOS attack surface against the reactive lookup procedure.

When the fastest service responds, the router will store the first packet and generate a broadcast lookup for the destination address. The broadcast lookup may be lost in the local network, e.g. due to the way a wireless network handles multicasts at the MAC Layer. During that window of time, the response from the other services may reach the router, in which case they are dropped pending resolution of the destination address. The result is that for several seconds - till the host tries again to reach the external services- the host reports to the user that the network has limited Internet connectivity, and the user applications are stalled for tens of seconds.

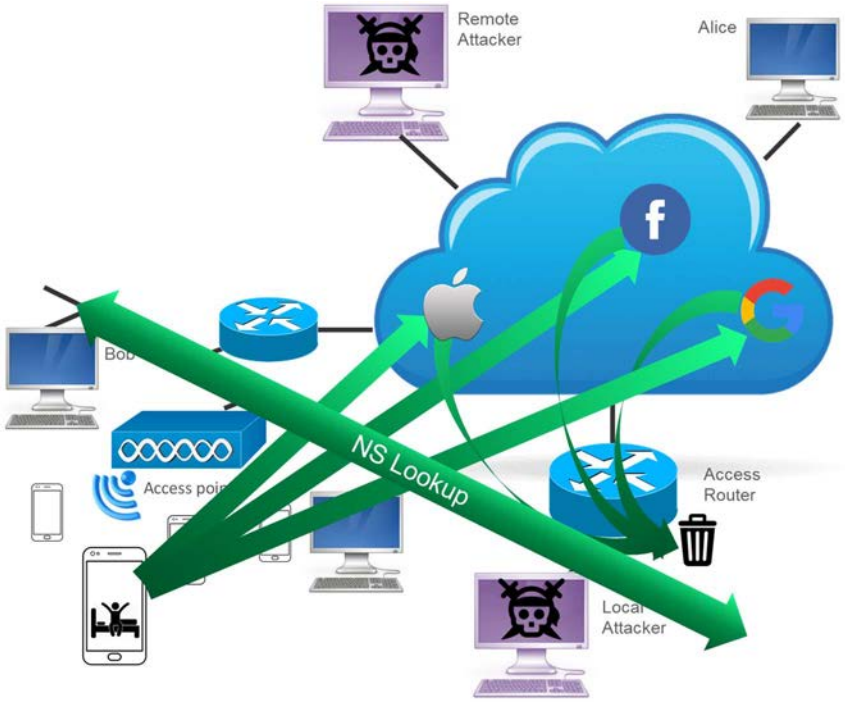


Figure 10: Temporary Hiccups at connection time

5.2.3.4 Impersonation attacks

Another form of attack that classic IPv6 ND is vulnerable to is impersonation attacks whereby an attacker that is connected to the local network may:

1) attracts traffic that is destined to a victim; or

2) generates traffic that appears to be coming from the victim.

Any node at any time may claim an address using a NA message with the override bit set. The message is broadcast and any node with a neighbour cache for the target address will update its neighbour cache with the Layer-2 address of the new node, including any router that serves the subnet. This is the IPv6 version of a gratuitous ARP in IPv4, and the problem is pretty much the same for both versions of the IP protocol. There was an attempt to solve that problem with Secure ND (IETF RFC 3971 [i.103], IETF RFC 3972 [i.104]) but the approach failed, due to the complexity and the need to maintain infrastructure for the key management.

Any node may also source a packet with any IPv6 address, whether the address is it own or forged, and use a forged address for flood and reflection attacks. If the source address of a packet is not topologically correct (the subnet is elsewhere), then IETF BCP 38 [i.105] applies to both IP versions to block the rogue traffic. Sadly, it is very difficult to perform that operation far from the origination network, e.g. based on a routing table, because the return path to a prefix may not be congruent with the forward path from which packets are received.

It results that the most efficient protection is to block at the access router of a network going outwards any packet that originated in the domain and for which the source address is not topologically present in the domain. This protects against attackers using impersonation to flood the impersonated victim but does not have effect if the attacker uses a topologically correct address, e.g. picks a source address from the local Subnet.

Those and more specific attacks are discussed in detail in the Source Address Validation Improvement (SAVI) Threat Scope IETF RFC 6959 [i.106].

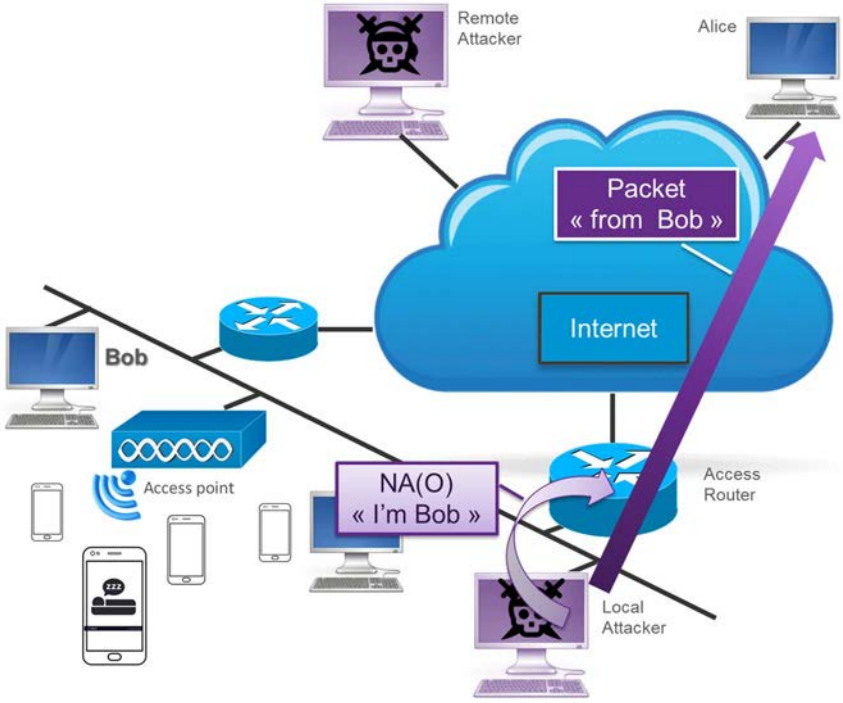


Figure 11: Impersonation attacks

5.2.3.5 Incomplete knowledge for SAVI, Routing and Proxy ND

As the security requirements on the network increase, it becomes necessary to ensure that only the owner of an address may use it to source or receive packets, a function generally referred to as SAVI (IETF RFC 7039 [i.45], IETF RFC 8074 [i.107]) for Source Address Validation Improvements. SAVI requires that the network that enforces a protection policy has the full knowledge of which node owns which address, and where it is located, which is not what the reactive IPv6 ND operation was designed to produce.

The same requirement arises when the routing fabric operates on host routes (e.g. in a data centre) and requires a complete state on which addresses are present in the network to either inject the routes in the IGP or populate a Mapping Server/Resolver (LISP).

The IEEE 802.11 [i.108] standard requires that the Access Point (AP) performs a proxy-ND operation to avoid the propagation of ND multicast messages when none of the intended destinations are connected to this AP. Then again, a full knowledge of which STA owns which address is required.

Even if the knowledge of IPv6 addresses used by a Node may be obtained by snooping protocols such as IPv6 ND and DHCPv6, or by observing data traffic sourced at the Node, such methods provide only an imperfect knowledge of the state of the Node at the AP. Silent Hosts such as printers may be forgotten, Sleeping Nodes such as IoT devices may appear to have moved away, and the router may not rely on its own ND cache as it only contains partial and possibly stale information. The link may have 2 routers for redundancy and only one router would have a particular address in the cache.

This may result in a loss of connectivity for some IPv6 addresses, e.g. for addresses rarely used (silent nodes), and in a situation of mobility when the mobility is not tracked efficiently.

For those reasons, snooping the IPv6 ND and DHCPv6 protocols is not a recommended technique, and it only be used as a last resort. The recommended alternative is to use the IPv6 Registration method specified in IETF RFC 8505 [i.90]. By that method, the AP exposes its capability to proxy ND to the Node specified in Router Advertisement messages. In turn, the Node may request proxy ND services from the AP for one or more IPv6 addresses, using an Address Registration Option. The Registration state has a lifetime that limits unwanted state remanence in the network. The registration is optionally secured using IETF RFC 8928 [i.91] to prevent address theft and impersonation. The registration carries a sequence number, which enables fast mobility without a loss of connectivity.

5.3 IPAM, DHCP, DNS, and Orchestration

5.3.1 IPAM motivation

The primary focus of the development of IPv6 was to allow the Internet to scale to hundreds of billions of interconnected devices. This design goal meant that IPv6 would not be backward compatible with IPv4 and several mechanisms used in IPv4 would have to be replaced by scalable ones for IPv6. The first was the size of IP addresses expanded from 32 bits to 128 bits to allow for substantial growth in the number of attached devices. Multiple protocols had to be enhanced or replaced to support IPv6 and these include ICMPv6, DHCPv6, OSPFv3, and BGP. In addition to these protocol enhancements, network management systems had to be reengineered to support the scale-up systems.

Because of the massive increase in IPv6 address ranges and individual IPv6 addresses in addition to the IPv4 addresses consumed during the transition phase to IPv6, an automated IP Address Management (IPAM) systems needs to support both IPv6 and IPv4 addresses, hierarchical IP address blocks, large numbers of subnets and a major increase in the total number of individual IP addresses. As the number and complexity of network-attached equipment are growing, associating both physical and logical interfaces and ports should also be supported. The IPv6 protocol defines multiple IP address assignment methods such as SLAAC, and different modes of DHCPv6 that are documented in Table 1. Most if not all current IPAM systems do not allow the configuration of the different address assignment methods as both the routers and end devices have to be configured depending on the required assignment method. An industry-standard configuration method needs to be developed to accommodate the multiple assignment methods.

Existing DNS servers will need to be enhanced to account for larger and more complex zone files as well as helper apps on switches and routers for DNS relays and forwarders. DNS has added additional Resource Records such as AAAA records for IPv6 address support as well as a new zone file ip6.arpa formatted to support reverse address lookups. Dynamic DNS updates are also supported from DHCPv6 assignments.

With the exponential increase in the number of all types of attached devices, more efficient and effective methods of configuration, turn-ups, and operations needed to be adopted and implemented to manage, secure, and operate the networks. Cloud with many virtual instances exacerbates the problem.

5.3.2 Background

Most IPv4 networks have evolved over time and did not usually have an overall architecture and a coherent IP addressing scheme, which leads to routing inefficiencies. IPv6 gives the opportunity to architect a hierarchical address plan and addressing scheme that may overlay the existing IPv4 network to promote more efficient routing of packets which reduces packet latency, simplifies management and improves overall network performance. There is a DDI acronym that stands for DNS, DHCP, and IPAM. The order of the addressing solution is assumed to be IP Address Plan, IP Address Management (IPAM), DHCPv6 or SLAAC, and then DNS.

The method to maintain, discover, and distribute IP addresses (IPAM, DHCP) and DNS records are needed to operate an IP network. DNS records are needed to map the FQDN to IP addresses and the reverse. The most popular accounting method for IPv4 addresses is a spreadsheet or small custom database developed by the network engineer whose job is to manage the addresses. From those early beginnings, a robust DDI marketplace has developed linking the management of IP address blocks with configuring DHCP scopes to advertise addresses and configuring zone files to allow the DNS servers to perform their role in DNS lookups.

Historically, in a typical building, a sufficient number of IP addresses were allocated per floor or per LAN based on the work being performed in that space and taking into account the density of required addresses. Printers, servers, and other shared devices would get a FQDN in DNS to allow common access to the resources. Frequent Moves, Adds, and Changes (MAC) were not documented timely. Therefore leading to more motivation for automated DDI systems and policies allowing devices to be added, moved, or changed only if the DDI system was updated before the change took place.

With the widespread deployment of IETF RFC 1918 [i.77] space after 1996, overlapping address block (typically /24) may be assigned behind the NAT. Various methods and workarounds were developed that obscured the purpose, design, and architecture of networks because of the increasing scarcity of globally routable IPv4 addresses.

5.3.3 IPv6 Engineering influence on Addressing

While IPv6 addressing is simpler than IPv4 addressing, an understand IPv6 engineering and associated addressing is still required. It is needed to understand design, operational goals, and how to reengineer subnets while deprecating IPv4 engineering practices that do not apply to IPv6. Some examples of these practices include:

- Multiple IPv4 subnets are required when there are not a sufficient number of IP addresses available in a single subnet. A single IPv6 subnet gives 2^64 numbers of end-user addresses/devices so multiple subnets are no longer required. It is suggested that the subnet structure be transitioned to a similar IPv6 subnet structure for testing and turn up of the IPv6 subnets. Once the network is stable and operating, redo the multiple subnets into one.

- Remove NAT devices and convert IETF RFC 1918 [i.77] space to routable IPv6 blocks and addresses with the addition of router ACLs and firewall rules for security.

- In IPv4 networks, packet fragmentation can occur at each hop, but in IPv6 networks, packet fragmentation is initiated by the source node if one of the links does not support the required MTU.

5.3.4 Challenges for Address Planning and IPAM systems

In the era of large to very large (IPv6) networks with multiple ASNs, IP Address Management systems need to be enhanced to manage the entire IP address lifecycle. It is needed not just the assignment of individual addresses and subnets but a system to manage the entire IP address lifecycle as IPALM (IP Address Lifecycle Management) system. Spreadsheets and small custom databases are not scalable solutions for IPv6 and hybrid IPv6 and IPv4 address management.

The short list of IP address planning challenges:

3) Design, develop and deploy either a hybrid IPv6/IPv4 address plan or a separate IPv6 address plan.

4) Choose a method to deploy the IPv6 addresses manually, DHCPv6, SLAAC, or by an orchestration system.

5) The hardware, virtual machines, and containers have to work on dual-stacked or IPv6-only networks depending on the overall IPv6 transition plan.

6) While subnet scanning was used to identify "active" IP attached devices on IPv4 networks, a single IPv6 subnet a /64 that is 18 446 744 073 709 551 616 addresses. It will no longer work to have one address per row in a spreadsheet or one line in a table. DDI system is needed to manage blocks of addresses and then the list of individual addresses assigned from those blocks.

7) IPAM systems need to manage the release of IPv4 addresses and blocks, aggregate the free addresses to the largest aggregable block, maintain the address block in "reclaim" status for some time (to allow the addresses to expire in DNS and route tables), and then "free" the blocks for reassignment or removal.

8) An equipment template object may be defined within an IPAM system to specify the number of physical and logical ports that have IP addresses assigned. This increases the accuracy of address assignments by not allow non-existing ports to be assigned addresses.

9) IPAM systems need to organize assigning IPv6 addresses to physical and virtual interfaces that already have IPv4 addresses assigned to support dual-stack. Dual stacking may require multiple virtual interfaces with each virtual interface being assigned a single IP address from either IPv6 or IPv4.

5.3.5 Challenges for Address Assignment Methods

Once the top-level master address block has been instantiated in the IPALM/IPAM system then:

1) Allocate subordinate address blocks from the master block based on the overall address plan taking into account organization or network structure and facilities locations.

2) These subordinate blocks may be a separate IP block aggregation level, or a national, or a regional level reporting to high levels.

3) If this is an aggregation point that aggregates multiple buildings or actual campuses then the allocation formula would be:

   a) Adding up the total number of buildings or campuses being aggregated and ascertain how close the total is to a power of two.

   b) Put in a buffer factor (at each level) so if a design has seven buildings that instead of using eight as the number of address blocks, select 16 to support additional aggregable address blocks for future growth.

   c) Add another buffer factor; for example, is to allocate a /63 for a subnet which gives two /64s in case expansion would be needed then the subnet would preserve aggregation.

4) As part of the design process of selecting block sizes for the number of subnets that will be supported per building, it is needed to layout a subnet structure supporting business goals and objectives using IPv6 engineering practices and principles supporting cybersecurity and privacy goals. The limit on the number of devices on a subnet is very high (/64). Hence, the standard layout of subnets for most buildings and campuses is suggested to simplify management, security, and privacy. An IPv6 /56 will allow 256 subnets, 8 bits of address between a /56 and a /64. While a /48 is 16 bits of address space giving a total of 65 535 subnets.

5) Address assignment mechanisms are an important choice. It may be a manual assignment, equipment, or interfaces configurations in an orchestration language, SLAAC, Stateful DHCPv6, Stateless DHCPv6, and a combination of Stateless and stateful DHCPv6. See Table 1 for the important characteristics of all these methods.

6) It is imperative to integrate address assignment methods with actual equipment, equipment configuration management, orchestration, and network monitoring and management systems so that there is only one repository of address information that feeds all of the other systems including cybersecurity.

Table 1: IPv6 configuration options

Address Autoconfiguration Method	ICMPv6 RA (Type 134) Flags		ICMPv6 RA (Type 134) ICMPv6 Option Prefix Info		Prefix Derived from	Interface ID Derived from	Other Configuration Options	# of IPv6 Addr
	M Flag	O Flag	A Flag	L Flag				
Link-Local (always configured)	N/A	N/A	N/A	N/A	Internal (fe80::)	M-EUI-64 or Privacy	Manual	1
Manual	Off	Off	Off	On	Manual	Manual	Manual	2 (LL, Manual)
SLAAC	Off	Off	On	On	RA	M-EUI-64 or Privacy	Manual	3 (LL, IPv6, IPv6 temp)
Stateful (DHCPv6)	On	N/R	Off	On	DHCPv6	DHCPv6	DHCPv6	2 (LL, Manual)
Stateless DHCPv6	Off	On	On	On	RA	M-EUI-64 or Privacy	DHCPv6	3 (LL, IPv6, IPv6 temp)
Combination Stateless & DHCPv6	On	N/R	On	On	RA and DHCPv6	M-EUI-64 or Privacy and DHCPv6	DHCPv6	4 (LL, IPv6, IPv6 temp, DHCPv6)

5.3.6 Challenges for DNS Systems

The DNS system developed for IPv4 has been extended to include IPv6 RRs (Resource Records) as well as different formats and names in-addr file "ip6.arpa".

5.3.7 Network Transition Strategies

There are two major scenarios during an IPv6 transition depending on whether there is an existing IPv4 network or if the deployment is a greenfield with no existing network.

Many current IPv4 networks are mission-critical. Hence, it is needed to use the existing network until a fully operational replacement would be ready. Dual stacking or IPv4aaS may be the solution for that scenario.

A dual-stack configuration has two independent protocols, native IPv4 and native IPv6 running on the same links, interconnected to the same interfaces into the routers, switches, or servers. Other criteria of a dual-stacked network that neither protocol is aware of nor interacts with the other protocol, in any way. Equipment in the network has to be tested to be compliant with dual-stacked environments before it would be configured in the network with the second protocol. The IPv6 address assigned to interfaces on dual-stacked equipment allows integration testing to occur. Users would not find services till IPv6 addresses would be populated in DNS servers. It permits to test systems and activate the IPv6 when it would be ready for production operation.

IPv4aaS assumes that only IPv6 is used for services inside the company. It makes sense for the scenario when the big proportion of services are delivered as native IPv6 traffic. Some hosts may still generate IPv4 requests (for example literal "192.168.1.1") or, more probably, many hosts would still need access to the IPv4-only part of the Internet. Hence, IPv4aaS additionally needs translation technology on the client-side and a gateway to the Internet to carry such exceptional IPv4 traffic over IPv6. The primary choices for translation technology are 464XLAT and DS-Lite but many other technologies exist (MAP-E/T, lw4o6). Some IPv4aaS transition technologies (like MAP-E/T) have the requirement for very strict coordination between IPv4 and IPv6 address space.

It may be a good choice to bypass Dual-Stack and go directly to IPv4aaS for new network deployments. Some translation technology would be probably needed anyway because of the requirement to access the IPv4-only part of the Internet.

For the foreseeable future, it is envisioned that there will be islands of both IPv6 and IPv4 only networks interconnected by dual-stack environments, tunnels, or NAT64/DNS64, or similar mechanisms. DDI management for both IPv6 as well as IPv4 would be needed for a long time.

One of the first steps in a transition process, once management and governance issues have been addressed, is to develop an accurate inventory of network devices, communication links, network topology, and all IP address blocks including those used both in the network and any interconnected networks. Ideally, device configurations would be under management by a configuration control system and all addresses and blocks would be accurately maintained in an IPAM or IPALM system for all network devices in the data and control plane, network monitoring and management equipment as well as any monitoring and management applications, GUIs, databases, analytics, and reporting.

It is recommended for enterprises to begin the IPv6 deployment process at the Internet perimeter with the upstream carrier. Dual-stacking Internet facing websites allows customers and partners to interact with the business utilizing either protocol.

Phased deployments of IPv6 network segments, backbones, or larger but manageable sections of the network allow the technical staff to gain experience with dual-stack or IPv6-only while still keeping the network operational.

5.3.8 Network Modelling and Orchestration

Historically, network equipment had element managers and Command Line Interfaces (CLI) to configure the device and activate services on it. CLI was good for humans but far from optimal for automation tools. Hence, Netconf and YANG have become very popular in the last decade. Netconf and YANG have been standardized as the primary Application Programmer Interfaces (API) for all OSS tools. Network Management System (NMS) and Software Defined Network (SDN) controllers are merged in the last years to one universal platform that does not have the standard name yet. This platform is the primary OSS tool to provision IP addresses on the network. Integration with IPAM is important for operational success.

5.4 Privacy considerations

To implement privacy, security is needed first to have a reliable and secure network as a foundation for privacy controls. Many countries have regulations related to privacy:

- CPNI [i.150], HIPAA [i.151], FCRA [i.152], ECPA [i.153] in US;
- GDPR [i.154] in Europe;
- CSL [i.155] in China;
- DPB [i.156] in India;
- Data Protection Act 2018 [i.157] in the UK.

Many regulations have been refreshed or fully redeveloped in recent years.

IPv6 was designed initially with mandatory encryption in mind. It has been understood later that the cost of encryption and key management is a too big burden for many types of business. Mandatory requirement for network encryption has been released. The second wave of encryption is now but this time it is applied at the application level. Modern hosts are powerful enough for encryption, challenge still exists only for key management.

Encryption at the application level does not hide the pattern of communication or the communication itself. IPv6 has additional capabilities to help here. IETF RFC 8981 [i.7] is widely accepted by the market. It permits to generate temporary addresses for new sessions that block the possibility to correlate sessions in transit. Additionally, the big address space and sparse population of it does permit to have random not just IP addresses but subnet numbers too - see IETF draft-ietf-opsec-v6-27 [i.53]. "Security by obscurity" is not the replacement for other mechanisms but it helps to complicate attacker's task. It is not possible to easily guess for an attacker the system type by just seeing the IP address.

It may be reasonable for some scenarios to use specialized security solutions such as firewalls, network access controls, identity and access management, security information, and event management systems (SIEM), IDS/IPS, event logs, and physical security systems as well as application firewalls.

Equally critical are role-based Identity and Access Management (IAM) systems tied in to allow user access only to the systems they are authorized to use for business purposes with the minimum amount of account privileges to complete the task.

Privacy and security, in general, are based on the systematic approach that is well discussed in IETF draft-ietf-opsec-v6-27 [i.53]. Many things improve IPv6 security. Well-designed and implemented privacy and data security systems mitigate insider threats, maintain data integrity, security of the data in transit and at rest, and allow only authorized access to the minimum amount of data to perform the tasks.

5.5 Wired & IPv6

Wired links are mostly Ethernet in our times, but carriers are still using some specialized link layer technologies (DOCSIS, DSL, and GPON). All other wired link-layer technologies (Serial, ATM, and Frame Relay) are history.

Infrastructure connections are typically P2P links even if Ethernet framing is used. It eliminated all problems discussed below because ND protocol is disabled or operates in a special restricted functionality mode.

IPv6 leverages Ethernet's multicast and broadcast efficiency for ND protocol. IPv6 ND is better suited to the nature of Ethernet multi-access links than IPv4 ARP, because only interested nodes process ND packets at the IP layer.

Many IPv6 problems are common for all wired scenarios:

- The problem of "Remote Denial of Service Attacks" discussed in clause 5.2.3.2 applies to a wired environment too. External ping sweep to /64 subnet would DoS router serving this subnet. The common solution for the problem is to just rate limit the number of new address resolutions per second that saves the router but affects legal traffic.
  The same technical problem (huge size of the typical subnet) is considered as the security advantage because it creates a big additional problem for an attacker in network reconnaissance.

- The problem "Captive Portals and Temporary Hiccups at connection time" discussed in clause 5.2.3.3 applies to a wired environment too. The router would start address resolution only after real traffic would be received from the Internet and probably heavily dropped (except for the first packet). IETF draft-ietf-6man-grand [i.24] proposes the solution for this problem by proactive unsolicited host announcement.

- A host may need many bootstrap parameters. ND has been patched already many times, but it is still behind DHCP for options richness: IETF draft-pref64folks-6man-ra-pref64 [i.25], IETF RFC 8106 [i.26], IETF RFC 5175 [i.27], IETF draft-jdurand-ipv6-multicast-ra [i.18], IETF draft-pfister-moving-net-autoconf [i.19]. It is an additional incentive for businesses to choose DHCP instead of SLAAC because a business may need additional flexibility.

- Brach offices of enterprises and SMB may have more than one device on-site and, at the same time, this site would probably have PA address space dynamically received from the carrier. It is a big challenge to distribute address space between many routers and links inside the site, even if the carrier would give a big enough prefix (/56 or better).

  NOTE: The homenet architecture (IETF RFC 7368 [i.22]) proposes a solution that was not adopted by the market.

- A multi-homing multi-prefix environment (when an enterprise or home subscriber is connected to many carriers for redundancy reason) is still an unresolved issue discussed in IETF RFC 7157 [i.23]. Hosts needs to use proper source address and forward the traffic in the direction of the respective carrier or else traffic would be dropped. IETF RFC 8028 [i.5] has a general recommendation on how to resolve this issue but it has not been incorporated into respective protocol standards (for example ND) and is not implemented by products. The same problem in the IPv4 home environment is typically resolved by separate NAT in the direction of every carrier but NAT is discouraged in IPv6.

- The problem of so-called "flash renumbering" is discussed in IETF RFC 8979 [i.31]. If connectivity to a carrier would be lost and then re-establish again, 37 % of carriers would give the subscriber the new IP prefix. If a subscriber has a complex network in-house that has many link layer hops (additional Wi-Fi™ access point, switch, or router) then hosts (smart TV, computers, smartphones, etc.) would not be informed that the previous prefix is invalid. A host would continue attempts to use the invalid IP address for a week. 6man WG is working now on the resolution of this problem.

DHCP by itself requires more resources for address configuration, it needs at least a DHCP server. But theoretically, DHCP has an advantage against SLAAC because DHCP may be snooped on the router and then this information may be fed into the MAC-IP binding table. It may help against problems from the above but such functionality is not widely implemented.

Distributed SLAAC procedure for address auto-configuration makes it difficult to trace IP address usage history for legal or troubleshooting purposes. It is an additional challenge to map the IP address to DNS name if IPv6 address is randomly chosen by the server itself. Hence, all types of businesses (enterprises and carriers) typically choose to stay on DHCP after IPv4 transition to IPv6.

For DOCSIS, DSL, and GPON access, carriers may use the service model when every subscriber is separated by its link-layer media (typically VLAN) up to the service point (BNG). It effectively emulates a separate link-layer P2P connection between subscriber and BNG that simplifies requirements to ND protocol.
Some carriers do use a shared link layer between the service point and many subscribers that potentially may create additional challenges in the SLAAC environment. But in practice, all fixed broadband carriers use DHCP for address assignment that eliminates the need for discussion of additional SLAAC problems.

NAT absence is an advantage because it permits two-way E2E connectivity that is important for new cloud services. But NAT was some sort of weak security protection that was considered good enough by many subscribers. A firewall is more important now for every subscriber or small business that leads to the associated complexity of firewall maintenance.

IPv6 is primarily used without encryption or authentication at the link layer that creates impersonation security problems discussed in clause 5.2.3. Any other address resolution protocol (like ARP in IPv4) would have the same security vulnerability under the absence of digital certificates. IPv6 has a SAVI solution (snooping by switches) to mitigate this problem, see clause 5.2.3. The industry has not found yet enough motivation to maintain digital certification authority in every legal entity.

Rogue DHCP server or Router is the attack vector of the same severity for IPv6 as well as for IPv4. Switches need to have RA-Guard filtering (IETF RFC 6105 [i.10]) and DHCP-Shield filtering (IETF RFC 7610 [i.9]).

In general, IPv6 was designed for Ethernet. Ethernet is the primary wired technology. Hence, IPv6 has good compliance to it.

5.6 Wireless & IPv6

5.6.1 Wireless Physical domain

At the physical (PHY) Layer, a radio broadcast domain is the set of nodes that may receive a transmission that one sends over an interface, in other words the set of nodes in range of the radio transmission. This set may comprise a single peer on a serial cable used as Point-to-Point (P2P) link. It may also comprise multiple peer nodes on a broadcast radio or a shared physical resource such as the Ethernet wires and hubs for which Classic IPv6 ND was initially designed.

On WLAN and WPAN radios, the physical broadcast domain is defined relative to a particular transmitter, as the set of nodes that are susceptible to receive the transmission. Literally every frame defines its own broadcast domain since the chances of reception of a given frame are statistical. In average and in stable conditions, the broadcast domain of a particular node remains consistent and is used, e.g. in Wi-Fi, to define a closure of nodes on which an upper Layer abstraction is built.

In short, a bidirectional PHY Layer communication is possible between two nodes when their physical broadcast domains of their unicast transmissions overlap enough to encompass both nodes consistently. On WLAN and LoWPAN radios, that relation is often symmetric, meaning that if B receives a frame from A, then A receives a frame from B. But asymmetries may be observed, e.g. due to power levels, interferers near one of the receivers, or differences in the quality of the hardware (e.g. crystals, PAs and antennas) that may affect the balance to the point that the connectivity becomes mostly unidirectional, e.g. A to B but practically not B to A.

It may be difficult to maintain every node in a multi-node Subnet in reach of every other, in a fashion that all their physical broadcast domains fully overlap. In other words, the relation of radio connectivity is generally not transitive, meaning that A in range with B and B in range with C does not necessarily imply that A is in range with C. As a result, a large wireless Subnet typically forms a Non-broadcast Multi-Access (NBMA) topology unless the MAC-Layer provides a Broadcast Domain Emulation.

5.6.2 MAC-Layer Broadcast Domain Emulation

An IEEE 802.11 [i.108] Infrastructure Basic Service Set (BSS) provides any-to-any connectivity to nodes as defined by the broadcast domain of the Access Point (AP). The AP relays both unicast and broadcast packets and provides the symmetric and any-to-any emulation of a shared wire between the associated nodes, with the capability to signal link-up/link-down to the upper layer. Within a BSS, the physical broadcast domain of the AP serves as emulated broadcast domain for all the nodes that are associated to the AP. Broadcast packets are relayed by the AP and are not acknowledged.

Figure 12: Wireless connectivity for Infrastructure Basic Service Set

To optimize the chances that a broadcast is received by all nodes in the closure, the AP transmits the broadcasted frames at the slowest PHY speed that reaches all nodes in the BSS. This translates into maximum co-channel interferences for others and the longest occupancy of the medium, possibly hundred times longer than that of the unicast transmission of a frame of the same size.

For that reason, upper layer protocols tend to avoid the use of broadcast when operating over Wi-Fi. To cope with this problem, APs may implement strategies such as turn a broadcast into a series of unicast transmissions, or drop the message altogether, which may impact the upper layer protocols. For instance, some APs may not copy Router Solicitation (RS) messages under the assumption that there is no router across the wireless interface. This assumption may be correct at some point in time and may become incorrect in the future. Another strategy used in Wi-Fi™ APs is to proxy protocols that heavily rely on broadcast, such as the Address Resolution in ARP and Classic IPv6 ND, and either respond on behalf or preferably forward the broadcast frame as a unicast to the intended target.

In an IEEE 802.11 [i.108] Infrastructure Extended Service Set (ESS), infrastructure BSSes are interconnected by a bridged network, typically running Transparent Bridging and the Spanning tree Protocol or a more advanced Layer 2 Routing (L2R) scheme. In the original model of learning bridges, the forwarding state is set by observing the source MAC address of the frames. When a state is missing for a destination MAC address, the frame is broadcasted with the expectation that the response will populate the state on the reverse path. This is a reactive operation, meaning that the state is populated reactively to the need to reach a destination. It is also possible in the original model to broadcast a gratuitous frame to advertise self throughout the bridged network, and that is also a broadcast.

The process of the Wi-Fi™ association prepares a bridging state proactively at the AP, which avoids the need for a reactive broadcast lookup over the wireless access. In an ESS, the AP may also generate a gratuitous broadcast sourced at the MAC address of the STA to prepare or update the state in the learning bridges so they point towards the AP for the MAC address of the STA. IETF RFC 8505 [i.90] emulates that proactive method at the Network Layer for the operations of AR, DAD and ND proxy.

In some instances of WLANs and LoWPANs, a Mesh-Under technology (e.g. IEEE 802.11s [i.109] or IEEE 802.15.10 [i.110]) provides meshing services that are similar to bridging, and the broadcast domain is well-defined by the membership of the mesh. Mesh-Under emulates a broadcast domain by flooding the broadcast packets at the Link Layer. When operating on a single frequency, this operation is known to interfere with itself, and requires inter-frame gaps to dampen the collisions, which reduces further the amount of available bandwidth. As the cost of broadcast transmissions becomes increasingly expensive, there is a push to rethink the upper Layer protocols to reduce the dependency on broadcast operations.

5.6.3 Unmet any-to-any expectation at the link layer

The traditional IT networking topology inherited from IPv4 Ethernet is that of an IP Link that is either Point-to-Point (P2P) or Transit (any to any). In that legacy model, the Layer-2 network is typically a P2P serial cable or a switched Ethernet fabric where switches autonomously form a L2 broadcast domain (e.g. using the Spanning Tree Protocol [i.88]).

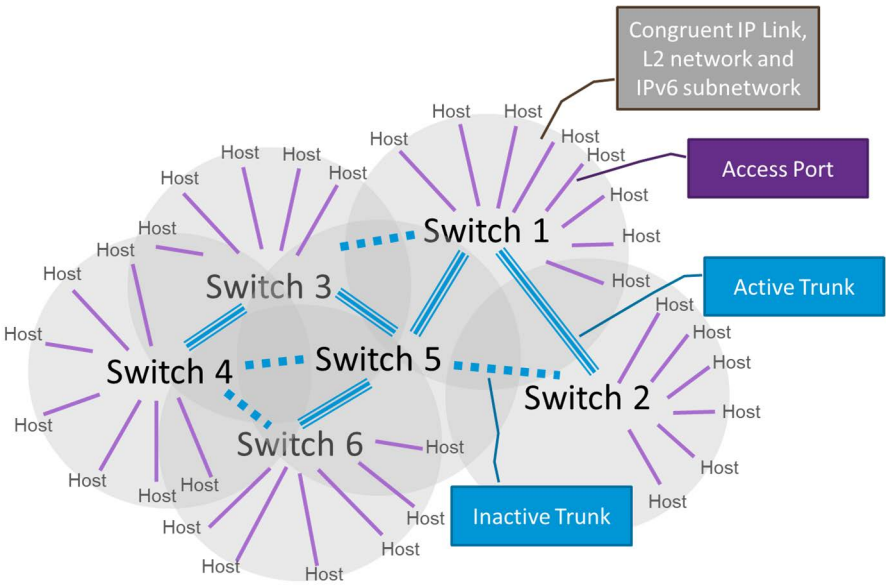


Figure 13: Typical switched network

This model may be applied to an IPv6 Subnet deployed over the broadcast domain, using IPv6 ND on-link subnet operation as a replacement for IPv4 ARP. Doing so, the issues related with broadcast storms in IPv4 only get worse with IPv6 as each host may form a collection of temporary addresses for privacy.

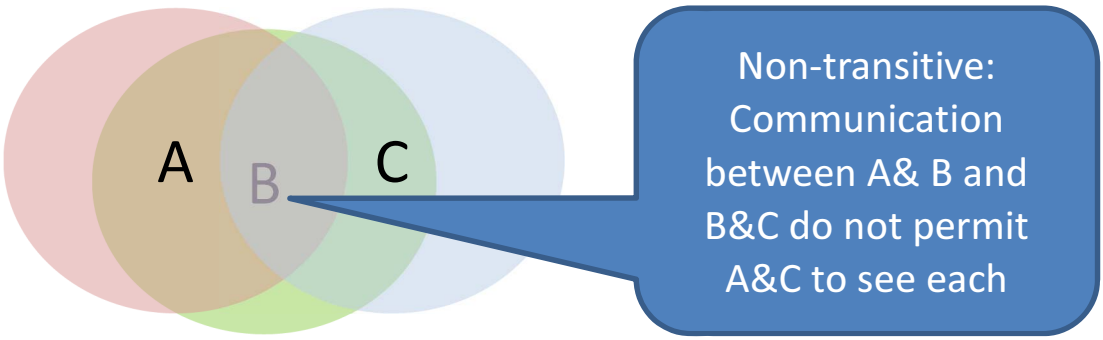


Figure 14: No any-to-any connectivity

Without a MAC-Layer Broadcast Domain Emulation, a multi-node radio subnet typically forms an NBMA structure, and the Classic IPv6 ND standard is very explicit that it requires extensions for operations over NBMA networks. And though it mostly works, the Classic IPv6 ND on-link Subnet operation over MAC-layer transit emulation such as L2R, learning bridges, and Wi-Fi™ Infrastructure Mode is not wireless-friendly due to the excessive use of MAC-Layer broadcast.

5.6.4 Roaming problem

Wireless improves the ease of connection, and mostly, it enables mobile scenarios where devices roam from an access to another. As a device moves and change its point of attachment, it may change its access router as well and it is possible that the addresses that the node formed from the router advertisements of a previous router are not topologically correct with the new router.

Classic IPv6 ND supports a case where a node is connected to multiple routers and the routers expose different prefixes; in that case IPv6 ND allows the node to form an address from any of the routers and use any other the router to forward the packets, even if the other router does not expose the same prefix.

To support mobility, a node needs to match the source address and the router, in a fashion like the first-hop router selection in IETF RFC 8028 [i.5] - see Figure 15.

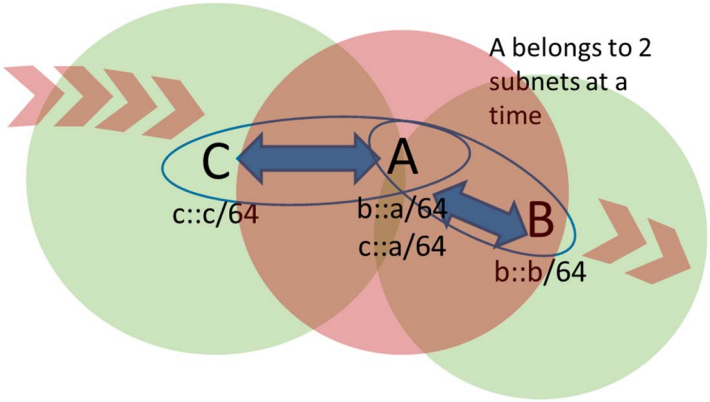


Figure 15: Transition at roaming

Linked to this, the discovery of a new router that is now reachable over the radio, and the deprecation of an old router that is no more reachable, are part of the overall fast roaming flow. For seamless mobility scenarios, a very responsive DNA (Detecting Network Attachment) method is required.

5.7 Multicast gaps

IP multicast enables point-to-multipoint data transmission in IP networks that supports a range of applications. IPTV is the primary application dependent on multicast in carrier networks. Software distribution and real-time financial updates are good examples for enterprise networks.

IP multicast protocols are divided into multicast membership management protocols and multicast routing protocols. Multicast membership management protocols run between hosts and routing devices. It is primarily Internet Group Management Protocol (IGMP) specified in IETF RFC 2236 [i.54], IETF RFC 3376 [i.55] and Multicast Listener Discovery (MLD) IETF RFC 2710 [i.74]. Multicast routing protocols are protocols that run between routing devices and mainly include Protocol Independent Multicast (PIM) IETF RFC 7761 [i.75] and Multicast VPN (MVPN) IETF RFC 6513 [i.76].

PIM is the most widely deployed multicast routing protocol. PIM is used to record the join information of a multicast receiver and build a multicast forwarding tree by sending a PIM join message to the multicast source.

IETF has proposed MVPN to run multicast applications as a VPN service where multiple services run simultaneously with VPN isolation from each other. The original MVPN was built based on the PIM protocol and IP GRE encapsulation which is the so-called Rosen MVPN defined in IETF RFC 6037 [i.78]. With the deployment of MPLS, carriers wish to use the MPLS data plane to carry all services including multicast. Hence, NG-MVPN has been proposed where BGP was extended to the new address family, see IETF RFC 6514 [i.120]. RSVP P2MP and mLDP P2MP may be used to build up a P2MP tunnel according to that standard.

Previous multicast technologies limitations are discussed in IETF draft-ietf-bier-problem-statement [i.124]:

- The existing multicast solutions request explicit tree-building protocols which bring the state to the transit nodes. Multicast convergence times are negatively impacted by tree state, which causes that multicast transition converges slowly than unicast after a failure in the network.
- There is a trade-off between flow state scalability and optimal delivery. In current multicast methods, if optimal delivery of multicast packets is provided then a different explicitly built tree is needed for each multicast flow which brings much more state in the network. Flow aggregation brings less optimal delivery.

- Maintaining and troubleshooting multicast networks may be very difficult. The available solutions are different from unicast, which often creates unique corner cases. Thus specialized skills and dedicated staff are required just to operate multicast services on a network.

BIER is the answer to the challenges above. It is discussed in clause 7.6.

5.8 Transition to IPv6-only for Carriers Broadband services

Broadband services have primarily an overlay structure, especially in the Mobile environment. It means that subscribers' traffic is isolated inside the tunnel, the tunnel itself has been built on a completely independent header and technology. Mobile traffic was always using GTP (ETSI TS 129 281 [i.82]) for tunnelling. Fixed broadband was primarily using PPPoE (IETF RFC 5578 [i.81]) for L2 tunnelling but a considerable part of carriers (especially in Asia) did use just the chain of VLANs for tunnelling at L2. A small proportion of fixed broadband carriers did not have separate virtual L2 media for every subscriber, anyway, they do use L2 isolation technologies to permit communications only between subscriber and carrier, they block direct communications between subscribers for security, compliance, and business reasons.

All these use cases greatly simplify the transition to IPv6 because IPv6 communication is needed only between devices on the tunnel ends. These devices are RG (by the terminology of BBF) or UE (by the terminology of 3GPP) on the subscriber side and respectively BNG or Packet Core UPF on the carriers' side. A high-level view of the Fixed and Mobile Broadband Network architecture may be seen in Figure 16.

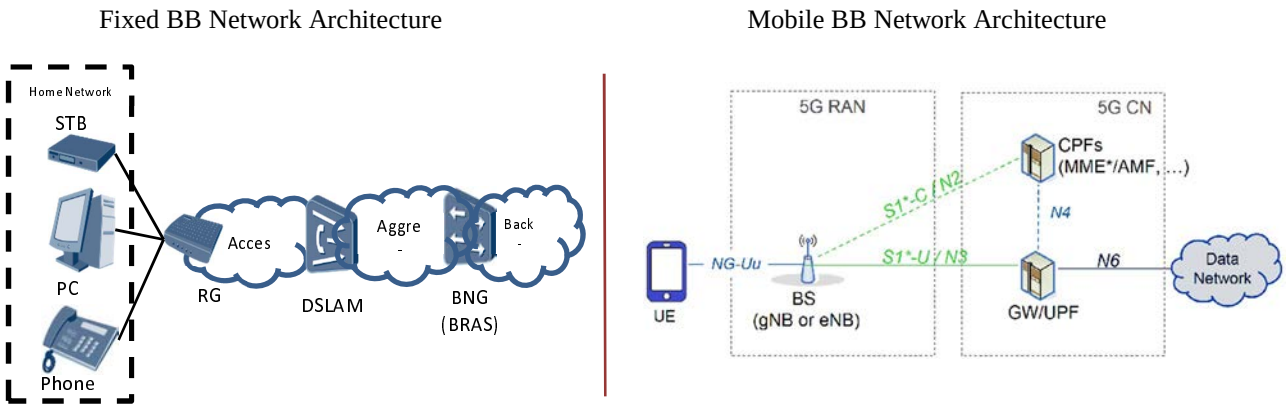


Figure 16: Fixed and Mobile broadband Network Architectures

The IPv6 support on carriers' broadband termination devices is very good, for more than a decade - no gap is possible. Mobile broadband has good coverage for IPv6 from the year 1999 in 3GPP specifications. Fixed broadband discussion is summarized in BBF TR-177 [i.83]. Albeit, it may have some performance implications because the dual-stack transition has twice as many instances (IP addresses, security, and QoS filters) on carrier's termination devices to manage. An IPv6-only environment does not have this additional challenge.

The IPv6 support on mobile UE has a shorter time in the market: a big push for IPv6 on smartphones and modems did happen around 2015. Six years is more than enough time for UE refreshment for markets where UEs are sold together with service contracts of carriers. UE refreshment may not be finished for other countries where subscribers do buy their smartphones by themselves. Hence, it is still possible to have a considerable proportion of UEs not supporting IPv6. It prevents carriers in such countries from direct migration to IPv6-only, they need to support conservative subscribers with old smartphones. In general, the refreshment cycle in mobile is very fast (around three years). It gives hope that the number of users that need IPv4 support would sharply decrease in the next few years. 5G deployment would speed up UE refreshment because it is not possible to buy 5G UE without IPv6 support. There is a non-zero probability that some UEs would not support IPv6 in all cases. Hence, Carriers have to manage IPv4-Only and IPv6-Only UEs at the same Packet Core concurrently.

Regulations in particular countries may impose some additional restrictions that influence the IPv6 transition strategy. For example, to facilitate threat tracking the number of subscribers sharing one IPv4 address may be restricted, or 5-tuple logging (IP addresses and TCP/UDP ports) may be mandatory.

IPv6 support on fixed broadband RGs is not so good for 2 reasons:

1) RG's refreshment cycle is much longer (up to 7 years for some countries);

2) a proper set of IPv6 requirements were introduced as late as the year 2019.

It is still very probable that the RGs of subscribers do not properly support IPv6, despite that carriers may have requested such support from the year 2013. Unfortunately, the initial set of requirements (IETF RFC 7084 [i.84]) had a restricted list of translation technologies (only DS-lite and Dual-Stack). A requirements update from IETF RFC 8585 [i.85] is very needed for an effective IPv6-only strategy. All in all, it means that IPv6 transition in fixed broadband needs a few years of preparation because many RGs may get IPv6 support only through refreshment.

PPPoE (FBB) and GTP (MBB) tunnels are fully transparent. They would not create any additional requirements to transit infrastructure for any IP version used inside tunnels.
It is important to consider the case where the separation of a fixed broadband subscriber from the carrier infrastructure is not so strict. It may be just VLAN isolation, where the FBB service is typically called IPoE. It is not good isolation because the transit L2 devices need to look into DHCP packets to insert user authentication information. Authentication is typically done based on physical ports that are visible only on the first hop switch. DHCPv4 is different from DHCPv6, it would demand from transit switches to parse DHCPv6 to insert option 18. It is not a big functional requirement for transit switches, but the gap may exist in the real deployment, hence, switches may need an upgrade.

It is possible to implement Dual Stack for the initial stages of IPv6 transition. The subscriber's device (UE or RG) and carrier termination device (BNG or UPF) would need to support both IP stacks. It is additional scalability requirements to the carrier's termination that may influence the cost and additional operating expenses for double-stack support. Hence, there is a big temptation to migrate to IPv6-only as soon as possible, at least for the carrier's part of the network. It is possible for so-called IPv4aaS strategy when IPv4 is carried over IPv6 carrier services - see Figure 17.

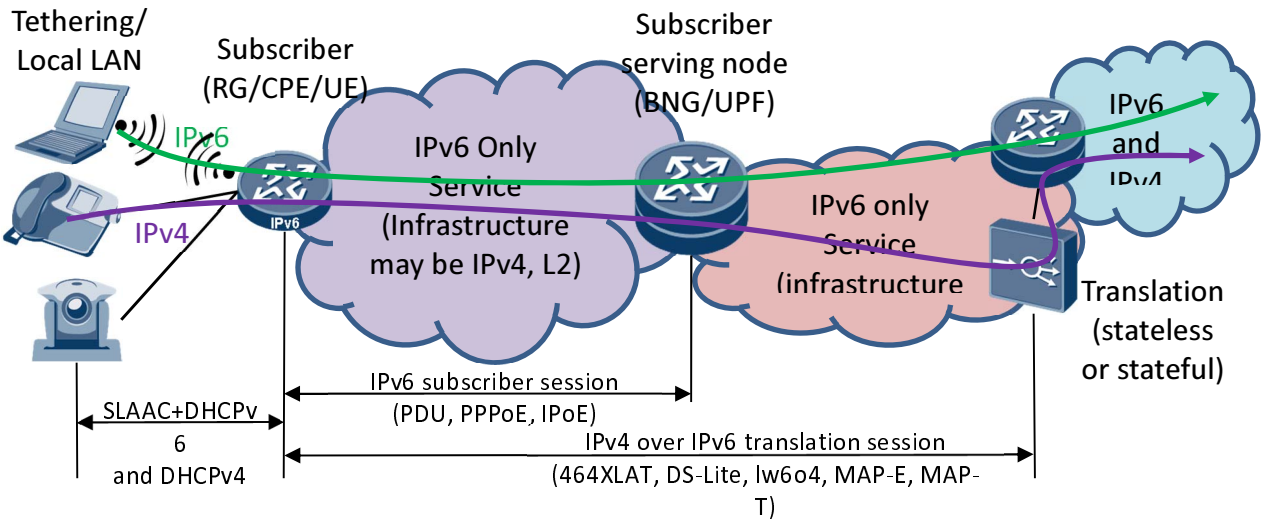


Figure 17: IPv4aaS in broadband architecture

Many external resources on the Internet may remain IPv4-only for a long time, hence, translation of IPv6 into IPv4 is needed. Such type of translation is the best in the form of NAT64+DNS64: DNS would synthesize IPv6 address from Internet IPv4 address and fixed prefix, then the host would believe that this resource is available on IPv6 and prefer IPv6 by default, then NAT64 on the carrier's border would translate it to IPv4.
NAT64+DNS64 is a very powerful combination because the majority of hosts do support the "Happy Eyeballs" library (IETF RFC 8305 [i.86]). It does help to support smooth transition because the host prefers IPv6 if available. Real deployments show that it is possible to achieve up to 99 % of IPv6 traffic inside the carrier in a mobile environment where almost all UEs support IETF RFC 8305 [i.86].

Unfortunately, there are old devices that support neither "Happy Eyeballs" nor plain IPv6. It is at least 1 % on mobile broadband (primarily connected by tethering) and up to 15 % on fixed broadband (printer, camera, and some other non-upgradable things). There is a need to translate IPv4 into IPv6 on the subscriber's side, use the IPv6-only service on the carrier side, and then translate it back to IPv4. 464XLAN is the best translation technology for this because it has a natural synergy with NAT64. It is an extension of NAT64. With NAT46 translation on the subscriber's side, the carrier's side (NAT64) does not need any change. All mobile UEs have very good 464XLAT support for the last 5 years, restrictions in the mobile environment are effectively the same as discussed above for plain IPv6 support on UEs, the gap is possible but the probability is decreasing very fast by every year. Fixed broadband RGs have low compliance to 464XLAT requirements because this requirement has become mandatory for RGs only from the year 2019 after IETF RFC 8585 [i.85] was specified. RGs gaps are very probable and are considered as the reason to do RG replacements in a refreshment cycle.

Some production networks have different types of middleboxes used for special subscriber services: Firewall, Deep Packet Inspection, Load Balancer, Intrusion Prevention Systems, and others. During the transition to IPv6, Carriers also manage their subscriber services controlled by Radius, PCRF, and DPI. Hence, they may need an upgrade and configuration changes to support IPv6 properly. See Figure 18.

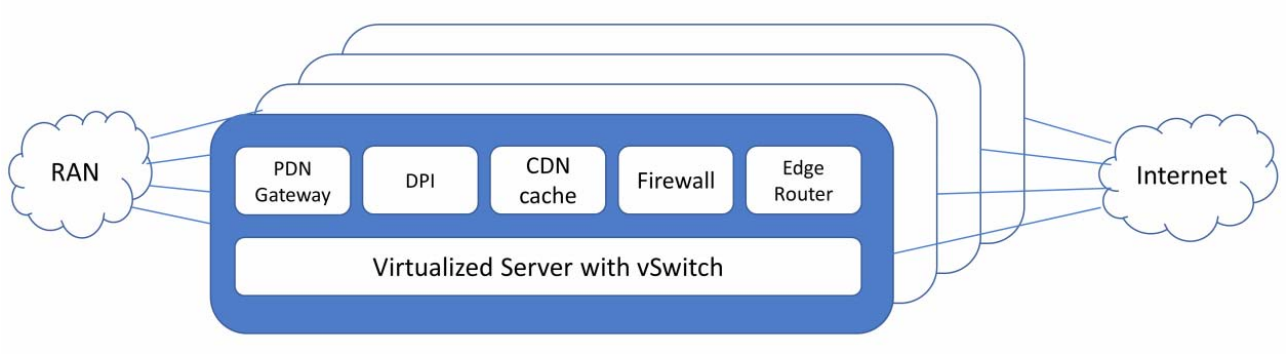


Figure 18: Mobile broadband architecture with special services

Any IPv6 transition would probably need an update of the OSS system because many carriers already have some level of automation, and some configurations would have different parameters for IPv6. BBF TR-187 [i.87] discusses new configuration parameters that BNGs need. Authentication, Authorization, and Accounting would also need the upgrade to properly handle the new IPv6 service. Lawful intercept would need changes too.
Pure IPv6-only installation (including IPv6 for underlay) would need a bigger upgrade because all OSS tolls (for fault, configuration, accounting, performance, and security) currently use IPv4 transport to manage devices.

IPv6 has one principal advantage for broadband services, because there is no NAT on the subscriber side. All hosts inside the subscriber's home network are possible to address with public addresses. It creates the possibility for new services such as surveillance, security, IoT, etc. From another point of view, this functionality makes it more important to have a firewall for home environment protection. It is not mandatory to have a firewall on the subscriber side because, as discussed above, all subscriber's connections are separate overlay connections. Hence, a Firewall is possible to put on the Carrier's side of the connection. It may be integrated or coincide with BNG or Packet Core.

The gaps discussed above, concerning broadband services, are all related to legacy hardware and software. The most recent standard that is vital for IPv6 broadband services is 464XLAT dated April 2013, everything else is older. It is an expected situation, because fixed and mobile broadband services had the address shortage first. It was evident that IPv6 migration is mandatory.

5.9 Operations, administration, and maintenance

The value of the network is strictly dependent on the quality of support and maintenance. It needs good OAM tools. OAM is primarily for:

1) SLA control;
2) troubleshooting;
3) compliance control;
4) security monitoring;

5) network optimization; and

6) event tracking and prediction.

Historically, OAM tools were mostly "active" by definition of IETF RFC 7799 [i.17]. It means that synthetic traffic was created and sent over tested paths. The purpose is:

1) continuity checks;

2) connectivity verification;

3) path discovery;

4) performance monitoring.

There are additional "passive" tools by definition of IETF RFC 7799 [i.17]. It means that there is no additional information sent over the network. Only logs and statistics are collected observing user traffic. Popular examples are TCPDUMP, IPFIX (IETF RFC 7011 [i.42]), sampled flow (see https://sflow.org/), and traffic mirroring.

Classic OAM tools have some challenges:

- Developed mostly for manual OAM. Automation by a controller is challenging.
- Limited flexibility in the definition of collected statistics.
- Networks have become extremely multi-path for the last decade. There are real production examples for up to 16k potential paths between pairs of locations for just 5 networking hops (links with respective multi-path 4*8*16*8*4). Active tools (synthetic traffic) would test a very limited number of paths, user traffic may still have an undetected problem.
- Massive statistics collected by some protocols (especially passive) may have an extremely big size that is expensive to centralized processing; it is very important to do aggregation and pre-processing on the Data Plane of network devices.
- Problems may be short spikes of the traffic (milliseconds level) that is not possible to investigate by polling detailed statistics all the time. It is needed to program the logic for statistics collection in a more programmable way (for example very detailed collection after a particular event).

IPv6 is an extensible protocol, it is possible to define a much bigger set of functionalities to resolve the challenges mentioned above. Clause 7.6 discusses the long list of new IPv6 OAM tools that would greatly improve OAM.

6 Non-technical gaps

6.1 Knowledge and experience

Clause 5.1 has shown a non-comprehensive list of IPv6 principal differences. It is primarily related to the wider functionality included in IPv6. Engineers with IPv4 background may not have proper knowledge and experience, so gaps here are very probable. It creates mistakes in transition and it may lead to sub-optimal IPv6 deployments.

It is especially harmful if security would be overlooked, because the consequences may be very harsh. Moreover, security for IPv6 is very different from IPv4, and so a lot of additional knowledge is needed.

Some publications on the Internet assumes that lack of knowledge on IPv6 is one of the top reasons for slow IPv6 adoption draft-ietf-v6ops-IPv6-deployment-status [i.21].

The potential sources of knowledge are:

- Training from companies specializing in IPv6.
- Vendor's training that is specific to IPv6.
- Self-study, the Internet has an overwhelming volume of information, and many formal books have been published.

- On the job experience in design, implementation, and support of IPv6 related functionality.

It is possible to get IPv6 certifications to test or validate that enough knowledge has been obtained. Some certifications are free of charge.
Certification may be mandatory for integrators or vendors that are contracted to help in the project fulfilment. The number of certified engineers is dependent on the project size, RIPE recommends 3 as the minimum, see https://www.ripe.net/publications/docs/ripe-554#skill_requirements.

It is best to have a plan and incentives for all stakeholders, to improve IPv6 knowledge, for any organization that has not finished the transition to IPv6. It is critically important for engineers related to networking, but it is also important for engineers responsible for applications. There is an example training plan in clause 6.1.2 of "National IPv6 Deployment Roadmap version II" in India [i.68], it has an estimation of resources that need to be spent.

6.2 Support for the current hardware and software

The general level of IPv6 support is very high for new systems and products less than 5 to 10 years old (depending on product). Unfortunately, the refreshment cycle for hardware and software has very big diversity from 1 year to 7 years (depends on the market segment). It is possible that, even after two major refreshments, a small percentage of products still lack IPv6 support. Hence, it is important to pay attention to IPv6 requirements in every refreshment cycle.

It is especially true for end-user devices. ARCEP analysis of IPv6 progress in France [i.114] is very typical in that way. It does have a list of end-user devices that support IPv6. The authors of that document have seen many investigations of other carriers with a long list of devices that do not support IPv6. It is more the case for fixed broadband CPEs that have been greatly improved just 2 years ago by IETF RFC 8585 [i.85] with IPv6 requirements.

The big problem may be with middleboxes, especially in the enterprise (like IPS and DPI). They have a lot of specialized functionalities that are still being developed by vendors. There is no functional parity yet between IPv6 and IPv4 for middleboxes because enterprises, the primary consumer of middleboxes, have made little progress on IPv6 transition.

The complexity in IPv6 dual-stack network operation increases the OPEX. IPv6-only may be a good strategy now with translation when needed for IPv4 legacy devices.

The network management, troubleshooting, performance monitoring, user management and accounting, security management, etc. for IPv6-only networks are not as mature as that of IPv4.

An IPv6 address is four times bigger than an IPv4 address. The same memory would have four times fewer IP addresses that may be used for routing, security, or QoS. Vendors typically do not install excessive memory, especially to small and cheap devices. The same memory may be enough for IPv4, but not enough for IPv6 despite smaller IPv6 tables that are better aggregated. Products that are typically affected: switches, low-end routers, IoT devices. Any product that has a graphical user interface is typically not affected, because it needs much bigger memory anyway.

In theory, IPv6 has better security features than IPv4 but in practice, there are still a lot of unknown factors concerning network and application security, as well as user privacy. Both hardware and software need to be upgraded for better IPv6 security and protection from attacks. The large-scale deployment of IPv6 and the new IPv6 security strategy may be a dilemma.

New telco software is typically IPv6-ready. But the probability is still very high that many sub-systems of OSS and BSS are not supporting IPv6. The real challenge happens if a particular product has been completely discontinued by the vendor or the vendor does not exist anymore, which is probably a good reason to decommission such a system anyway.

Software support may be still a problem for an enterprise environment. The enterprise market has many millions of applications in production, it is very expensive to upgrade everything, despite that an IPv6 upgrade is typically small and simple. There are some cases when even IP is not supported (50 years old mainframe application), not to mention IPv6. The enterprise market is conservative in principle and does not have a budget for replacement, only for expansion and new functionality. Hence, enterprise migration may be slow, and thus it requires a business level justification to ensure progress is gained and maintained during the transition.

Software IPv6 support problems maybe not visible if dual-stack is deployed, but would pop up as soon as IPv4 is disabled. This is why failover testing from IPv4 to IPv6 is critical for dual-stack environments.

The majority of the gaps above are related to the transition period. There is no principal deficiency, missing standard, or missing product on the market. There is only one critical product requirement: pay attention to IPv6 requirements for any case of replacement or upgrade of any hardware or software product. Make sure that it will support IPv6 use cases in the future.

6.3 Design and migration

ETSI has published white paper #35 [i.67] where it is possible to find wide discussion about topics of this clause.

There is an assumption here that for the majority of carrier cases the ecosystem is already mature enough to have IPv6-only as the transition goal. It is easy to have 85 % of traffic IPv6 for FBB and 99 % for MBB after all subscriber equipment is refreshed.
An enterprise is probably more reasonable to target Dual-Stack because of the lack of application readiness.

IPv6 is superior in functionality to IPv4. It is not possible to have the situation where some IPv4 functionality is not possible to implement on IPv6. The opposite is possible: there are more and more use cases that are possible to implement only on IPv6.

ETSI GR IP6 001 [i.69] and IETF RFC 7381 [i.70] have good discussions around IPv6 deployment approaches. Migration, in principle, is not much different from any expansion or upgrade project:

- Justify project inside the company with business stakeholders.
- Get budget, people, and other resources.
- Audit current services, design, and sub-systems capabilities.
  It is especially important to evaluate end-user equipment and applications.
- Improve internal knowledge on new (IPv6) capabilities.
- Document target design.
- Use a security by design approach for IPv6 & deeply integrate it into your existing IPv4 security controls, approaches, and operations.
- Prepare transition plan.
- Go through the procurement process for new products and services, it may be needed to test products.
- Supervise external partners tasks, manage project inside, pay attention that people may be reluctant to changes.
- Test and accept the result.

Some projects may omit knowledge improvement that is needed for IPv6-related projects. This causes a fundamental gap, putting the initiative at risk. This knowledge improvement should be done with management as well as technical staff, to ensure business support, operational efficiency and adequate security.

Dual-Stack implementations have a small additional cost because some devices need more scalability, or they have twice the amount of configuration (interfaces, IP addresses, filters, counters, etc.).

Security is a big challenge because it is a complex problem in general and security is different for IPv6. In addition, during the transition to IPv6 the attack surface may increase significantly as workloads, services, and devices are available over both IPv4 and IPv6.

The best option from the financial point of view, would be for projects to not be specifically targeted for IPv6, but instead IPv6 be included in all projects (upgrade, expansion, replacement of not supported or outdated equipment).

6.4 Maintenance and support challenges

It has been discussed already in many clauses of the present document that IPv6 is very different from IPv4. It is critical to prepare a support team and tools for IPv6.

All IP vendors have training and certification tracks that are mostly targeted for maintenance and support. It is typically a 3 level system: associate, professional and expert. Hence, it is easy to get proper knowledge for OAM.

The tools (NMS, Controllers, CLI) are typically multi-protocol and easily support IPv6 as well as IPv4.
It is discussed in clause 7.5 that IPv6 has many new and very advanced OAM tools. It is welcome to implement any of it to improve OAM quality. It is the positive differentiator for IPv6.

There is one negative thing related to IPv6 OAM. A dual-stack environment will need more maintenance during the transition period, and that may be long for the enterprise. This is a big incentive to migrate directly to IPv6-only.

7 IPv6 enhanced innovations for the gaps

7.1 General industry trends for technology transition

The development of the network has gone through phases for the requirements of different eras. In the 1980s, IPv4 became the basic protocol of the Internet and promoted the development of IP networks. In the 2000s, the MPLS technology enhanced the comprehensive bearing capability of voice and video services. After 2010, IPv6 deployment was accelerated. Cloud-based everything, Internet of Everything, and industry digitalization posed new requirements including ubiquitous and flexible connectivity, ultra-high bandwidth, deterministic quality, low latency, automation, and security. The IPv6 enhanced innovation represented by SRv6 emerges to meet these new gaps.

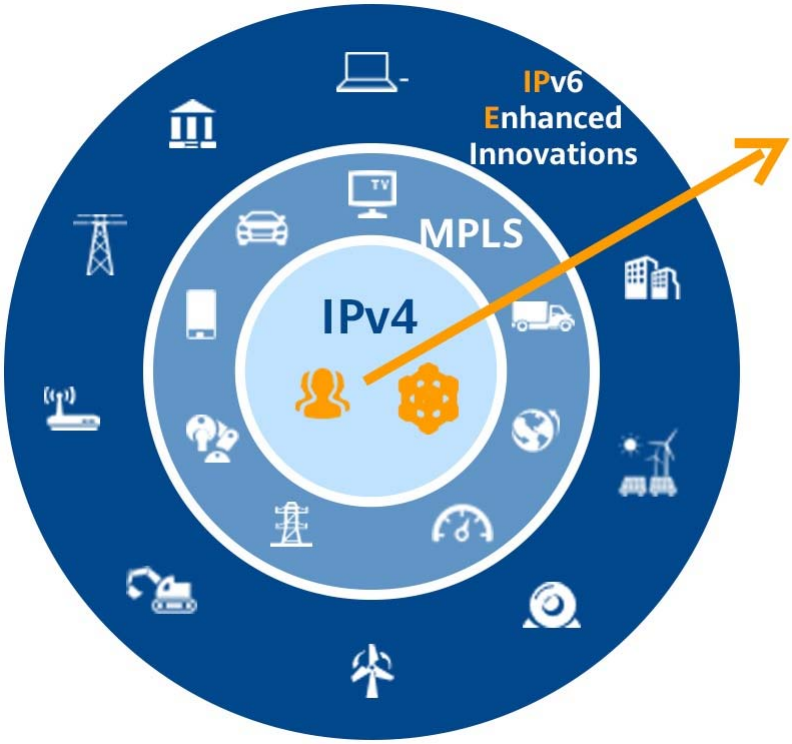


Figure 19: IPv6 Enhanced Innovations promoting the development of networking

The majority of this clause is the discussion of technologies that are still in development.

7.2 Proactive ND

7.2.1 Proactive ND justification

Ethernet transparent bridging per IEEE 802.1Q [i.88] provides an efficient and reliable Layer-2 broadcast service for wired networks. Applications and protocols have been built that heavily depend on that feature for their core operation, e.g. to provide zero-configuration discovery operations. This model enables simple operations and was a perfect match with the undelaying medium when Ethernet was a single shared wire. At that time, unicast and broadcast transmissions consumed the same resource with the same efficiency.

The IPv6 (IETF RFC 8200 [i.6]) Neighbour Discovery (IPv6 ND) protocol (IETF RFC 4861 [i.1], IETF RFC 4862 [i.2]) was introduced in the same period (the 1990s) in the same way as transparent bridging. It is a reactive protocol based on multicast transmissions to locate an on-link correspondent and ensure the uniqueness of an IPv6 address. In practice, the Layer-3 link-scoped multicast operation is operated as a Layer-2 broadcast for many Layer-2 technologies, mostly wireless, but that was not an issue when a local network was a bunch of computers on the same yellow wire.

Unfortunately, the evolution of the transmission media has made the basic assumption of broadcast efficiency less and less correct. When Ethernet Bridging moved to Switching, the phenomenon of broadcast storms appeared and became the gating factor constraining the size of the network. With wireless, the rate adaptation mechanisms that speed up unicast transmission by orders of magnitude does not apply to broadcast, adding a new dimension to the inefficiency of broadcast versus unicast.

Things only got worse with pseudo-wire and other overlay technologies that interconnect Layer-2 networks over the Internet, adding the cost of WAN transmissions to the bandwidth waste associated to broadcast operations. Networks started a slow migration towards pure Layer-3 operations and host routing in the overlay to avoid the broadcast issue completely.

Low-Power and Lossy Networks (LLNs) are typically wireless and battery-operated. Hence, they are extremely constrained in bandwidth and energy. It is not a surprise that, in the context of 6LoWPAN, the migration did happen to unicast operation on the access link under the new ND design that named Proactive ND (IETF RFC 6775 [i.89], IETF RFC 8505 [i.90]).

7.2.2 Introduction to Proactive ND

The Duplicate Address Detection (DAD) and Address Resolution (AR) procedures as defined in classic IPv6 ND expect that a node in a subnet is reachable within the broadcast domain of any other node in the subnet when that other node attempts to form an address. It is mandatory to check the duplicate status. This is why IPv6 ND is applicable for P2P and transit links, but requires extensions for more complex topologies.

Proactive ND (IETF RFC 6775 [i.89], IETF RFC 8505 [i.90], IETF RFC 8928 [i.91], IETF RFC 8929 [i.92]) defines a new operation for ND that is based on two major paradigm changes, proactive address registration by hosts to their attachment routers and routing to host routes (/128) within the subnet. This allows Proactive ND to avoid the expectations of multi-access links and subnet-wide broadcast domains. Proactive ND provides an interface between a node and a router that allows the node to register all its addresses before the router needs them resolved. The node provides attributes that characterize the use of the address, such as a lifetime, a sequence counter to maintain a sense of order when moving, and optionally a proof of ownership (IETF RFC 8928 [i.91]).

Proactive ND specifications apply to both wired and wireless technologies and were expressly designed for modern deployments such as hybrid-wireless and virtual topologies.

Proactive ND is agnostic to the method used for Address Assignment, e.g. Address Auto-configuration, DHCPv6 (IETF RFC 8415 [i.4]), or any form of IP address management (IPAM) solution such as found in modern enterprise IT and hybrid cloud environments. It does not change the IPv6 addressing architecture (IETF RFC 4291 [i.93] or the current practices of assigning prefixes, typically a /64 to a subnet. But the Duplicate Address Detection (DAD) is performed as a unicast exchange with a central registrar, using new ND Extended Duplicate Address messages (EDAR and EDAC). This unicast technique modernizes IPv6 ND and fits natively with overlay map resolvers such as LISP (IETF RFC 6830 [i.94]) enabling unicast lookups for addresses registered to the resolver.

Proactive ND enables to span a subnet over many links that are collectively called a Multi-Link Subnet (MLSN). It is for instance used to federate edge wireless links with a high-speed, typically Ethernet, backbone. This way, nodes may move freely from a wireless edge link to another without renumbering. IETF RFC 8505 [i.90] is already integrated with the RIFT (IETF draft-ietf-rift-rift [i.95]) routing protocol defined for use in datacentre fat tree Clos topologies. It is also the preferred method to set up IPv6 ND proxy operations, as described in IETF RFC 8929 [i.92].

7.2.3 A Node to a Router interface

In a model where IP Links are P2P and the subnet spans a mesh of routers, ND needs an abstract interface from the host to the router that provides the generic information that the router needs, regardless of which protocol runs in the routing fabric. Proactive ND provides an abstract way for a host to announce its addresses to the router and request reachability for that address, with the necessary information for the router to inject the address in a timely, orderly, and secure fashion.

Proactive ND registers an IPv6 address for a specified lifetime, so any stale state is removed from the routing fabric automatically and very fast. It provides ordering to enable movement where an advertisement from different places replace each other which allows a clear differentiation from anycast operations where both advertisements are usable. Finally, IETF RFC 8928 [i.91] adds proof of ownership to trust that the advertised address is not stolen or impersonated.

The sequencing operation defined for RIFT combines the sense of order from Proactive ND, which is a counter that eventually wraps, with a low precision sense of time (typically ~100 ms with NTPv4) to obtain a hybrid sense of time, using the Proactive ND sequence counter only when the advertisements are too close to determine the order from the resolution of the timestamps.

The proactive address registration is performed with a new option in NS and NA messages, which is the Extended Address Registration Option (EARO) defined in IETF RFC 8505 [i.90]. This method allows to prepare and maintain the host routes in the routers and avoids the reactive address resolution in classic IPv6 ND and the associated Layer-2 broadcasts transmissions. The router is free to obtain that reachability whichever way is used in the network, for instance by injecting the address in an IGP, or by performing ND-proxy operations towards a legacy network.

7.2.4 Links and Link-Local Addresses operation in Proactive ND

For Link-Local Addresses, DAD is typically performed between communicating pairs of nodes and ND cache is populated with a single unicast exchange. In the case of bridging proxies, though, the Link-Local traffic is bridged over the backbone and the DAD is proxied there as well.

For instance, in the case of Bluetooth® Low Energy (BLE) (IETF RFC 7668 [i.96]), the uniqueness of Link-Local Addresses needs only to be verified between the pair of communicating nodes, the central router, and the peripheral host. In that example, 2 peripheral hosts connected to the same central router are not permitted to have the same Link-Local Address because the addresses would have a collision at the central router which talks to both hosts over the same interface. The DAD operation of Proactive ND is appropriate for that use case, but the classic IPv6 ND operation is not because the peripheral hosts are not on the same broadcast domain.

On the other hand, the uniqueness of Global and Unique-Local Addresses is validated at the subnet level, using a logical registrar that is authoritative for the whole subnet. The registrar may be a centralized resolver or implemented as a distributed database in an IGP or BGP/EVPN.

7.2.5 Subnets and Global Addresses

Proactive ND extends Classic IPv6 ND for Hub-and-Spoke (example BLE) and Routed MLSNs (example RPL - see IETF RFC 6550 [i.97]). In the Hub-and-Spoke case, each Hub-and-Spoke pair is a distinct IP Link, and a subnet is mapped on a collection of links that are connected to the Hub. The subnet prefix is associated with the Hub.

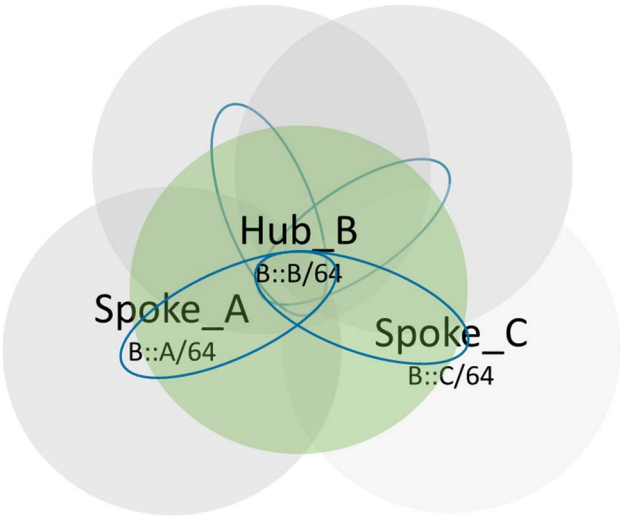


Figure 20: Subnetting in a Hub-and-Spoke topology

Acting as routers, the Hub advertises the prefix as not-on-link to spokes in Prefix Information Options (PIO) of RA messages. Acting as hosts, the spoke auto-configure addresses from that prefix and register them to the Hub with a corresponding lifetime. Acting as a registrar, the Hub maintains a binding table of all the registered IP addresses and rejects duplicate registrations, thus ensuring DAD protection for a registered address even if the registering node is sleeping. The Hub also maintains a cache for the registered addresses and delivers their packets to any of those addresses during their respective registration lifetimes. This forms the Network Layer equivalent of a Wi-Fi™ Infrastructure BSS.

A Routed MLSN is considered as a collection of Hub-and-Spoke where the Hubs form a connected dominating set of the member nodes of the subnet, and IPv6 routing takes place between the Hubs within the subnet. A single logical registrar is deployed to serve the whole mesh.

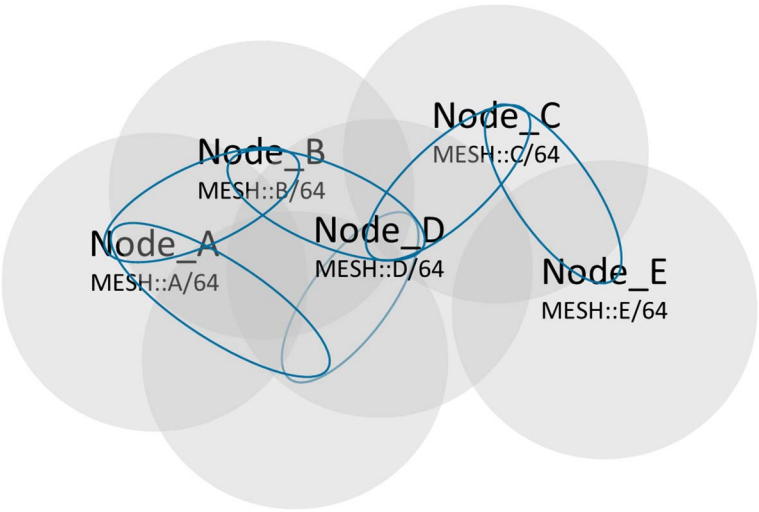


Figure 21: Subnetting in environment with routing inside the multi-link

IETF RFC 9010 [i.98] details how the Proactive ND registration is redistributed in RPL IETF RFC 6550 [i.97] for multilink subnet routed over the mesh of Low-Power Lossy Networks (LLN). In the simpler hub-and-spoke configuration, all the hubs are connected to the same high-speed backbone such as an Ethernet bridging domain where classic IPv6 ND operates. In that case, it is possible to federate the hub, spokes, and backbone nodes as a single subnet, operating ND proxy operations IETF RFC 8929 [i.92] at the Hubs, acting as backbone routers. This forms the Network Layer equivalent of a Wi-Fi™ Infrastructure ESS.

7.2.6 Proactive ND Applicability

Proactive ND applies equally to Point-to-Point (P2P) links, Point-to-Multipoint (P2MP) Hub-and-Spoke, Link-Layer Broadcast Domain Emulation such as Mesh-Under LLN and Wi-Fi™ BSS, and Route-over LLNs, Layer-3 inter-cloud connections, and pseudo-wire overlays including EVPN.

There are scenarios where link and subnet are congruent and where both classical IPv6 ND and Proactive ND may be used. These include P2P, the MAC emulation of a PHY broadcast domain, and the particular case of always-on, fully overlapping physical radio broadcast domain. But even in those cases where both are possible, Proactive ND is preferable versus Classic ND, because it reduces the need for broadcast and radio broadcast emulation are far from perfect.

There are also several practical scenarios in the real world where links and subnets are not congruent:

- The IEEE 802.11 [i.108] infrastructure BSS enables a subnet per AP and emulates a broadcast domain at the Link Layer. The IEEE 802.11 [i.108] infrastructure ESS extends that model over a backbone and recommends the use of an ND proxy to interoperate with Ethernet-connected nodes. Proactive ND incorporates an ND proxy to serve that need.

- Bluetooth® is Hub-and-Spoke at the Link Layer. It would make little sense to configure a different subnet between the central and each peripheral node (e.g. sensor). Rather, IETF RFC 7668 [i.96] allocates a prefix to the central node acting as a router, and each peripheral host (acting as a host) forms one or more addresses from that same prefix and registers it.

- Inter-cloud overlays tend to operate at Layer-3 using host routes to the cluster, meaning that the Layer-2 domain if any is constrained within the cluster. It is also possible to extend the Layer-2 domain using overlay technologies such as VxLAN to build larger layer-2 domains, but it is inefficient to extend a Layer-2 overlay beyond the local cloud because of broadcast issues.

- Large corporate buildings are typically split into many physical sectors (one or a few adjacent rooms) with local services such as printers that need to be found with the help of mDNS (IETF RFC 6762 [i.99]). Whereas the Layer-3 Subnet would typically span the building and needs to maintain the connectivity without IP renumbering as a mobile unit (phone or personal computers) moves inside the building.

In those cases, it makes sense to model the IP Link as the direct Point-to-Point (P2P) relation between a Host and a Router, regardless of the span of the Layer-2 connectivity and span the subnet using Layer-3 routing, independently of the Layer-2 limitations.

7.3 Flexible extension headers (EH)

IPv6 is famous to be flexible and agile. It is attributed primarily to the possibility to extend functionality by extension headers (EH). EHs follow the basic IPv6 header and may have many options inside. EH is specified in clause 4 of IETF RFC 8200 [i.6]. IANA registry (https://www.iana.org/assignments/ipv6-parameters/ipv6-parameters.xhtml) keeps up to date list of all extension headers (currently is 11) and all options (currently is 45).

Routers process headers in hardware for 2 decades. Some hardware was not flexible enough to process new headers. Sometimes it requires developing new microcode. Some hardware has a performance penalty for a particular EH or a long chain of headers processing. Some EHs have security implications (especially routing instructions). Many transit nodes (middleboxes) need to process all headers up to the transport layer for ECMP load balancing and filtering that exacerbate all before mentioned problems. IETF has many publications related to these problems, the most recent one is IETF draft-gont-v6ops-ipv6-ehs-packet-drops [i.78].

Security is the big problem of EH because it has been proved historically that combination of EHs may trespass the lookup capability of network devices, and then security filtering is bypassed. It is especially the problem for multi-layer switches because switches are more restricted on headers processing capabilities. Recommendations have been developed that if a switch is not capable to process some headers then it is better to drop the packet. Good explanation is IETF RFC 7113 [i.79] "Implementation Advice for IPv6 Router Advertisement Guard (RA-Guard)".

It was historically the problem to support big header lengths on network devices. It is less the problem now due to the high performance of ASICs used for packet forwarding on the latest generation of routers. It may be still the problem for multi-layer switches, even new ones. Anyway, it is still the habit for network administrators to disable EH by default then enable only the functionality that is needed. It is true even if all routers in a particular domain are capable to process a much bigger list of EHs. It is attributed to the next reasons:

1) more security and predictability for routers behaviour;

2) business typically does test all functionality that would be activated;

3) part of routers in the existing network may not support particular functionality;

4) some routers have performance penalty from EH processing.

This typical admin strategy (to disable EH by default) does create challenges to introduce new functionality in the IPv6 world. It needs many years before functionality would be standardized, developed by vendors, bought during some project or refreshment lifecycle, and finally activated by the network administrator. It results in an extremely high probability (50 % or even more) for the packet with just one EH to be dropped on the Internet and almost assurance (90+%) of packet drop for more than one EHs. The big problem that fragmentation in the IPv6 world is done with the mandatory attachment of EH that may lead to an impressive 40 % drop on the Internet, see IETF RFC 7872 [i.80].

The problem of EHs support in a particular network is more organizational than technological, more functionality needs more efforts from vendors and network administrators that need an additional budget that needs separate justification in business. There are already many examples when particular functionality (for example, SRH) was activated very fast through the big network because business reason has been found. Hence, despite the alarming statistics for EH support on the Internet, it is not a critical problem for one domain because it is easy to activate any EH or EH options if the business has the motivation.
The global availability of new IPv6 functionality is more difficult because there are many legacy routers worldwide with a refreshment cycle of close to 7 years. Moreover, some line cards based on old ASICs (2+ years) are still in sales and have restrictions on the length of the EHs chain. This problem would disappear by itself by hardware replacement and improvement.

A similar problem did not exist in IPv4 because IPv4 is not capable of such extensibility of functionality.

7.4 SRv6-based forwarding

Basic IPv4 was missing functionality that is very much needed for many network types, like VPN, TE (traffic engineering), and FRR (Fast Reroute). MPLS was the new technology in 1996 that addressed this gap (IETF RFC 3031 [i.39]). It is especially important for the carrier's backbone, metro, and mobile transport. Diverse functionality (hundreds of standards) has been developed for 2 decades in MPLS to support multi-service transport and promote the Internet's all IP transformation. MPLS has great success in the carrier's environment.

However, MPLS has also brought some problems and challenges, though it has achieved great success:

- MPLS islands are isolated. Although MPLS unified the technology for transport networks, IP backbone, metro, and mobile transport networks are separated and need to be interconnected using complex technologies such as inter-AS VPN. It makes E2E service deployment difficult and expensive.

- Limited programming space in IPv4 and MPLS encapsulation. Many new services require more forwarding information to be added to packets. However, the IETF announced that it has stopped developing further standards for IPv4. Also, the format of the MPLS Label field is fixed and lacks extensibility. These reasons make it difficult for IPv4 and MPLS to meet the requirements of new services for network programming.

- Decoupling of applications from transport networks is harmful for carriers. It makes it difficult to optimize networks for specific applications and improve the network's value. Many carriers found themselves stuck as a pipes provider without benefits from value-added applications. Moreover, the lack of application information means that carriers only implement network adjustment and optimization in a coarse-grained way that results in not optimal resource usage. Attempts have been made over the years to extend network technologies to hosts but all have failed. MPLS was not accepted at the application level nor in the data centres.

SR (IETF RFC 8402 [i.38]) was proposed in 2013 to address the above challenges, including network programmability. SR is a source routing paradigm that program the forwarding path of packets at the ingress of the path by inserting the set of forwarding instructions into packets. An instruction is called a segment that is executed on nodes of packet forwarding or processing. It may refer to a specific interface or the shortest path through which a packet is forwarded to a node. Such a segment is identified by a Segment Identifier (SID).

Currently, there are two data planes for SR:

1) MPLS; and

2) IPv6.

When SR is applied to the MPLS data plane, it is called SR-MPLS (IETF RFC 8660 [i.32]), and the SID is encoded as an MPLS label. When SR is applied to the IPv6 data plane, it is called SRv6, and the SID is encoded as an IPv6 address.

SRv6 is an SR network paradigm based on the IPv6 data plane by making use of a new IPv6 Routing Header, called Segment Routing Header (SRH) (IETF RFC 8754 [i.33]), allowing the ingress to insert forwarding instructions to guide data packet forwarding and processing. IETF RFC 8986 [i.34] defines extensive programmability capabilities. As of now, SRv6 has been commercially deployed by multiple carriers around the world.

SRv6 technology has the following advantages against MPLS:

- SRv6 is compatible with IPv6 forwarding and implements the interconnection of different network domains easily through basic IPv6 reachability.
- Unlike plain MPLS, SRv6 does not require additional signalling or network-wide upgrades.
- SRv6, based on SRHs, supports encapsulation of rich information into packets to greatly improve programmability, meeting diversified requirements of new services.
- SRv6's affinity to IPv6 enables it to seamlessly integrate IP transport networks with IPv6-capable applications and provide more potential value-added services for carriers through application-aware networks.

Data plane and control plane extensions are required to support SRv6. Some RFCs have been published while the others are approaching. Some key IETF standards or drafts are listed in table 2.

Table 2: SRv6 progress

Technical Subject	Status	Reference
SRv6 Architecture	RFC	IETF RFC 8402 [i.38]
SRv6 Data plane Encapsulation	RFC	IETF RFC 8754 [i.33]
SRv6 Network Programming	RFC	IETF RFC 8986 [i.34]
IS-IS for SRv6	RFC	IETF RFC 8401 [i.128]
OSPFv3 for SRv6	RFC	IETF RFC 8444 [i.129]
SRv6 OAM/Ping and Trace	WG Draft	IETF draft-ietf-6man-spring-srv6-oam [i.101]
PCEP for SRv6	WG Draft	IETF draft-ietf-pce-segment-routing-ipv6 [i.59]
BGP for SRv6	WG Draft	IETF draft-ietf-bess-srv6-services [i.130]
BGP SR Policy for SRv6	WG Draft	IETF draft-ietf-idr-segment-routing-te-policy [i.51]
BGP-LS for SRv6	WG Draft	IETF draft-ietf-idr-bgpls-srv6-ext [i.29]

However, SRv6 has its disadvantages as well. For example, in use cases like strict path TE, the overhead of SRv6 encapsulation is too large if many SRv6 SIDs are used in the SRH. The larger size of the headers will reduce the efficiency of transmission and forwarding. SRv6 compression design team has been established in the IETF SPRING working group to discuss the requirements of SRv6 compression.

For supporting better interop with existing networks. There are some requirements when designing SRv6 compression solution:

- The IPv6-based solution is preferred over MPLS-based, which benefits from IPv6 rich functionality and capability to build a unified network.

- The SRv6 based solution is preferred for compressed SRv6. SRv6 architecture, control plane, and data plane would be reused. The compressed solution may consider different data plane and control plane, as long as it derives sufficient benefit. Compatibility with SRv6 will help the solution easier deployment in the existing networks, so a smooth upgrade from SRv6 to compressed SRv6 is required.

- The segment summarization is mandatory for the solution. Summarization of segments is a key benefit of SRv6 versus SR-MPLS. It is important from a scalability point of view that any node would reach any other node via a single prefix segment in inter-domain deployments. Without summarization, border router SIDs need to be leaked and an additional global prefix segment is required for each domain border to be traversed.

- A compression solution requires a reasonable address size for not to waste too big address space.

- Carriers require flexible addressing planning because the planning of the network may vary based on the addressing scheme of the carrier. Flexible IPv6 address planning is mandatory for the compression mechanism. It is a valuable capability to use GUA from different address blocks.

A few solutions are in discussion in the SPRING working group. They aim to improve the encapsulation efficiency while will decrease the difficulty of processing the SID and increase the transmission efficiency.

7.5 Service-based Slicing

The Fifth Generation of Mobile Technologies (5G) is designed to connect everything, enabling emerging industries of an unprecedented scale and giving a new life to mobile communications. Services in the 5G era are classified into three main types:

- enhanced Mobile Broadband (eMBB) focuses on bandwidth-hungry services, such as HD video, Virtual Reality (VR), and Augmented Reality (AR);

- ultra-Reliable Low-Latency Communication (uRLLC) focuses on latency and reliability sensitive services, such as autonomous driving, industrial control, telemedicine, and drone control;

- massive Machine Type Communication (mMTC) focuses on scenarios with high connection density, such as smart city or smart farming.

These types of services require very different kinds of network characteristics and performance that are usually expressed in SLAs as a set of Service Level Objectives (SLO). There are several types of performance characteristics, including guaranteed maximum packet loss, guaranteed maximum delay, and guaranteed delay variation. Note that these guarantees apply to traffic in conformance to the contract, out-of-profile traffic is typically handled according to a separate agreement with the customer.

An additional important element of the required SLA is a guarantee that the service offered will not be affected by any other traffic in the network. This is termed "isolation" and a customer may express the requirement for isolation as an SLO parameter. Network slicing is the best technology to answer isolation challenges. It helps a single physical network to meet diverse requirements and provide a performance guarantee at the same time.

Network slicing is a method for creating multiple virtual networks over a shared physical network. Each virtual network possesses a customized network topology and provides specific network functions and resources to meet functional requirements and provide SLA guarantees for different tenants.

5G E2E network slicing covers the RAN, mobile core network, and transport network slicing. The network slicing architectures and technical specifications of the RAN and mobile core network are defined by 3GPP, whereas the transport network slicing is mainly defined by the IETF, BBF, IEEE, and ITU-T. IETF has its scope of "IETF Network Slices" that is restricted only to the IP domain.

Slicing assumes resource isolation. It is the key difference from an ordinary VPN. Several network slices need to be created on the same physical network. Transport network slices provide three levels of isolation: service, resource, and O&M:

- Service isolation: Service packets in one network slice will not be delivered to the endpoints in another network slice on the same network. In other words, service isolation ensures services of different network slices to be invisible to each other on the same network. It is important to note that service isolation does not provide guaranteed SLAs. When the network resources are shared between network slices, one network slice may still be affected by another, even if service isolation is implemented.

- Resource isolation: Network resources may be dedicated for a slice or shared among multiple slices. Resource isolation is paramount for 5G uRLLC services, which usually have strict SLAs and do not tolerate interference from other services. Resource isolation may be provided with different degrees, ranging from hard isolation to soft isolation. Carriers may select the optimal approach of resource isolation to meet their service requirements.

- O&M isolation: In addition to service isolation and resource isolation, network slice tenants may require independent operation and maintenance of network slices allocated by carriers, similar to a private network. This may be provided by the management plane of the network slices through an open management interface to the network slice tenants.

With the increasing requirement on network slicing from various vertical industrial customers, the number of network slices needed would increase accordingly. Thus the scalability of network slices becomes an important factor to consider. This includes the scalability in the forwarding plane, the data plane, the control plane and the management plane.

A network slice may span multiple network domains, and in some domains it may consists of a multi-layered network. The network slicing solution needs to support multi-domain and multi-layer coordination and integration, so as to provide an end-to-end network slice with consistent characteristics.

There are several VPN technologies defined in IETF that provide service isolation:

- L3VPN: IETF RFC 4176 [i.133], IETF RFC 4364 [i.134];

- L2VPN: IETF RFC 4664 [i.135], IETF RFC 6624 [i.136];

- PWE3: IETF RFC 8077 [i.137], IETF RFC 4448 [i.138];

- VPLS: IETF RFC 4761 [i.139], IETF RFC 4762 [i.140];

- EVPN: IETF RFC 7209 [i.61], IETF RFC 7432 [i.62], IETF RFC 8214 [i.63].

An enhanced VPN framework was proposed in IETF draft-ietf-teas-enhanced-vpn "A Framework for Enhanced VPN Services" [i.64] to provide VPN resource isolation. It describes a layered network architecture and the candidate technologies in each layer to deliver enhanced VPN service. It introduces the Virtual Transport Network (VTN) layer to provide customized virtual underlay networks with dedicated or shared resources for different enhanced VPN services. Thus enhanced VPN is a realization of network slices in the transport network.

The realization of network slice requires enhancement in the network data plane, control plane, and management plane.

In the data plane, identifiers of network slices need to be carried in data packets. These identifiers instruct packets of different network slices to be forwarded and processed according to constraints such as the topologies and resources of specific network slices. The identifiers of network slices may be based on existing fields in the data packet, or new fields may be introduced to carry the identifiers related to network slices.

In the control plane, information about logical topologies, resource attributes, and states of different network slices is defined and collected to provide essential information for generating independent slice views. The control plane also needs to provide functions such as independent route computation and service path provisioning for different network slices based on the service requirements of the slice tenants and data plane resources allocated to the slices.

In the management plane, the interfaces and related service models need to be defined for the network slice lifecycle management functions, including planning, creation, monitoring, adjustment, and deletion of network slices. This may be based on augmentation to the existing network service models, and some dedicated service models for network slice management may be introduced.

The enhanced VPN framework and the related protocol extensions are specified in relevant IETF documents:

- Enhanced VPN Framework: IETF draft-ietf-teas-enhanced-vpn [i.64].

- Resource-aware SR segments: IETF draft-ietf-spring-resource-aware-segments [i.65].

- Segment Routing for enhanced VPN: IETF draft-ietf-spring-sr-for-enhanced-vpn [i.141].

- IGP Multi-Topology for SR VTN: IETF draft-ietf-lsr-isis-sr-vtn-mt [i.41].

The interface for network slice management is under the definition in IETF as a northbound YANG data model. It will permit network slice tenants to describe the network slice requirement and collect the network slice-related performance statistics.

Research for resource reservation technologies is not finished yet. It is not fully clear yet which technology is better for a particular case.

7.6 Telemetry-based OAM

Clause 5.9 discusses new OAM market challenges that may be addressed only by new tools. The primary answer to these challenges is the "Telemetry".

A new wave of tools is in development. They are primarily classified as "hybrid" by IETF RFC 7799 [i.17].
iOAM is the most common name for these tools. "i" stands for "in-situ" which is equivalent to "on-path". It means that instructions or metadata are embedded into user packets for statistics to be collected on transit hops. The primary iOAM tolls specified for now are:

1) Tracing Options (per hop).
2) Proof of Transit Option.
3) Edge-to-Edge Option (E2E).
4) Direct Export (DEX) Option or Postcard-Based Telemetry - purely passive.
5) Passport-Based Telemetry - change metadata on transit nodes.
6) Echo Request/Reply Option.
7) IPFIX raw export of headers.

"Alternate Marking" method is another special hybrid technique of coloring users' traffic - see IETF RFC 8321 [i.43]. It is considered as one additional iOAM method to the mentioned above. Accordingly, related work have been proposed in IETF, including IPv6 Application of the Alternate Marking [draft-ietf-6man-ipv6-alt-mark] for "Alternate Marking" method implementation in IPv6 network, BGP/PCEP SR Policy Extensions to Enable On-path Telemetry [i.143], Subscription to Multiple Stream Originators [i.144], a YANG Data Model for iOAM [i.145]. IGP and BGP Extensions for Capability Advertisement are being considered.

IPv6 is the most suitable protocol for such header extensions, see IETF draft-ietf-ippm-ioam-ipv6-options [i.44] for how it is leveraged for iOAM.

There is a recent programmable approach to collect statistics directly from nodes' Data Plane. It is primarily the protocol (gRPC) to "subscribe" for interesting statistics with the encoding to "export" statistics directly from the Data Plane. See IETF draft-kumar-rtgwg-grpc-protocol [i.47] and IETF draft-openconfig-rtgwg-gnmi-spec [i.48]. It is effectively the predecessor of "Telemetry" discussed below.

Network Telemetry describes how information from various data sources may be collected using a set of automated communication processes and transmitted to one or more processing nodes for analysis. Analysis tasks may include event correlation, anomaly detection, performance monitoring, metric calculation, trend analysis, and other related processes.
Network Telemetry architecture is specified in IETF draft-ietf-opsawg-ntf [i.49]. It has the next comment "Network Telemetry refers to both, the data itself (i.e. "Network Telemetry Data"), and the techniques and processes used to generate, export, collect, and consume that data for use by potentially automated management applications. However, the term of network telemetry lacks a solid and unambiguous definition."
Network Telemetry has a special emphasis on YANG data models to improve automation capabilities. The list of primary Telemetry features is:

- Push and stream statistics instead of pull.
- Smart flow/packet selection, high volume and velocity tolerance.
- Normalization and unification.
- Smart data export and model-based data representation.

- Cross-domain, cross-device, cross-layer.
- Dynamic and Interactive network probe.

Network Telemetry may generate rich and complex statistics that benefit from AI techniques. Hence, a lot of activities to develop AI processing for network telemetry.

Network telemetry is assumed only inside one domain (trusted mode) because the market is not ready to deal with cryptography protection for intensive data plane traffic and key management with untrusted parties.

The most relevant draft for iOAM gap analysis is IETF draft-ioametal-ippm-6man-ioam-ipv6-deployment [i.30]. It does assume that no gap exists on the protocol level.

The biggest gap now is the support of Telemetry by vendor products, especially by controllers and orchestrators.
It would be improved while demand from carriers and enterprises would increase in the future.

7.7 Advanced multicast

Multicast has some challenges discussed in clause 5.6. This clause discusses BIER, the multicast technology solution for these challenges, defined in IETF RFC 8279 [i.121]. BIER uses a bit string for registering the multicast replication information, one bit for every edge router that may potentially receive the stream. BIER forwarding is stateless and decomposes the bit string node by node to replicate multicast traffic to the right interfaces. The node with BIER capabilities is named Bit Forwarding Router (BFR).
BIER uses the protection mechanisms of IGP which automatically brings extra reliability to BIER.

BIER applicability is discussed in IETF draft-ietf-bier-use-cases-12 [i.50]:

- Data Centre Virtualization scenario. Virtual eXtensible Local Area Network (VXLAN) defined in IETF RFC 7348 [i.40] emulates a layer 2 broadcast domain across the layer 3 underlay. Multicast solution for overlay would require the multicast capability in the underlay network. However, considering the complexity of the traditional multicast technologies, many data centre providers do not want to deploy multicast because of the high network operation and maintenance cost. BIER does not have multicast states in the underlay, hence, operation and maintenance are significantly simplified.
- Financial Services scenario. Financial services rely on IP multicast to deliver stock market data and its derivatives, which typically require a low latency path, deterministic convergence, and secure delivery. BIER enables the choice of the most optimal path from publisher to subscribers by leveraging unicast routing. BIER has the multicast convergence as fast as unicast, uniform and deterministic regardless of the number of multicast flows. This makes BIER more suitable multicast technology in the scenario of financial services.
- Broadcast Video Services scenario. In a video broadcast environment, the media content is sourced from various content providers across different locations. For typical broadcast services, sources are the satellite terminal nodes that may receive and feed the content with the satellite. A multicast group address is assigned for each broadcast source. With BIER in a broadcast services environment, there is no need to run multicast protocols in the core and no need to maintain any multicast state, which brings faster convergence and greatly simplifies operation and maintenance.

The basic BIER defined in IETF RFC 8296 [i.122] only covers the scenarios of BIER MPLS and BIER Ethernet, BIER over IPv6 is not included. BIER over IPv6 multicast solution is based on the existing BIER architecture. Requirements for BIER IPv6 is defined in IETF draft-ietf-bier-ipv6-requirements [i.123]:

- Possibility for transit through IPv6 nodes that do not support BIER. The BIER header is used to direct the packet from one BFR to the next BFRs. The basic IPv6 header is used to provide reachability between BFRs. Hence, it is possible to properly set the IPv6 destination address by BFR for a packet to pass the non-BIER IPv6 node.
- Inter-domain deployment enablement. IPv6 unicast address reachability provides the capability to replicate a BIER packet between routers in different domains.
- Good compatibility with other IPv6 functionality based on extension headers. For example, in the IPsec ESP case, fragmentation and reassembly are needed together with BIER.

To satisfy the above requirements, the BIER over IPv6 solution needs to cover:

- IPv6 data plane.
- BIER over IPv6 inter AS domain deployment.
- IGP support. IS-IS and OSPF extensions are needed for BIER over IPv6.
- Control plane support for MVPN. BGP procedures and messages need to be specified.
- Faster BIER over IPv6 protection may be needed for some business-critical services.
- New OAM tools.

The current discussion for BIER over IPv6 is draft-ietf-bier-bierin6 [i.126]. There are two scenarios considered, direct BIER encapsulation into link layer, and BIER encapsulation into IPv6.
In the GRT context, MLD requires extension for BIER IPv6 as described in IETF draft-ietf-bier-mld [i.119].
In the NG-MVPN context, the multicast registration information is exchanged via BGP as described in IETF RFC 8556 [i.125]. Independently of these two scenarios, the mechanism stays the same. The multicast sender (BFIR) receives all the multicast requests and creates a bit string composed of all edge routers identifiers. Then the BFIR creates packets that include the bit string information and the multicast data.

BIER standardization status is summarized in Table 3.

Table 3: BIER progress

Technical Subject	Status	Reference	Standard Published
BIER Architecture	RFC	IETF RFC 8279 [i.121]	Nov 2017
BIER Encapsulation	RFC	IETF RFC 8296 [i.122]	Jan 2018
BIER IPv6 Requirements	Draft	IETF draft-ietf-bier-ipv6-requirements [i.123]	
BIER MVPN	RFC	IETF RFC 8556 [i.125]	Apr 2018
MVPN Explicit Tracking	RFC	IETF RFC 8534 [i.127]	Feb 2019
IS-IS for BIER	RFC	IETF RFC 8401 [i.128]	Jun 2018
OSPFv3 for BIER	RFC	IETF RFC 8444 [i.129]	Nov 2018
BIER OAM/Ping and Trace	Draft	IETF draft-ietf-bier-oam-requirements [i.131]	
BIER Performance OAM	Draft	IETF draft-ietf-bier-pmmm-oam [i.132]	

BIER with MPLS encapsulation underlay is already a standard and has many vendors' implementations.
BIER over IPv6 is the draft that probably needs 1 to 2 years to reach wide acceptance on the market.

7.8 Service-oriented Networking

Quality of Service (QoS) and Quality of Experience (QoE) were always important topics for enterprises and carriers delivering services to enterprises. Networking had historically two ways to satisfy these requirements: Differentiated Services (IETF RFC 2475 [i.56]) and Integrated Services (IETF RFC 1633 [i.57]). Both ways did assume minimal complexity of the Data Plane, all complexity has been moved into the Control Plane. Integrated Services create serious challenges for Control Plane that constrained Integrated Services deployments. Differentiated Services have limited services granularity.

Driven by new services, fine-granular differentiated service provisioning becomes more and more desirable. Filters are widely used in the carriers' networks to enforce policies within their networks, typically at the network edge. However, as services and corresponding policies become more complicated, carriers will have to manage the large number of filters that impose heavy maintenance overhead. It would be very valuable to find a way to reduce the number of filters. Service tagging may be a way out. In this case, there is a need for a method to tag key services and user groups.

The Data Plane has become much more powerful in the last 20 years, especially comparing it to the Control Plane at the same time. It has become evident that it is possible to move a little more complexity to the Data Plane without any significant penalty. It may greatly improve the scalability and decrease the complexity of the Control Plane and at the same time considerably improve service quality. Data Plane needs just an additional header to carry personalized SLA parameters (identifier). IPv6 is perfectly designed for such a task: it has the concept of extension headers that carry such type of information.

Service-oriented networking needs the capability of both, Integrated Services and Differentiated Services at the same time. It will control all possible quality parameters (delay, jitter, packet loss, cost, and much more) without any burden on the control plane of all network nodes.
This new approach permits to have very personalized treatment for services or users in an extremely scalable way.
The solution is friendly to widespread encryption that prevents packet classification at layers above transport. At the same time, it fully protects privacy, because reverse mapping is not possible from an SLA identifier to a user or service. It depends on the trust model who is permitted to insert such identifier to every packet of traffic flow, in the trusted mode is assumed to be the initial host, in the untrusted mode is assumed to be the edge router in the domain delivering services.

7.9 Computing-based Networking

CDN and public DNS services usually dispatch users' requests to the server which is geographically closest to the user. These services may use the anycast IP address to achieve this goal. For the distributed edge computing scenario, there are multiple MEC sites around one user. The geographical distances between these sites and the user have no significant difference to impact network quality. So it is suitable to dispatch requests of the users to these near sites because they can provide similar low-latency quality. However, there are only a small number of servers in a MEC site. Traffic steering according to geographical distance only, like anycast IP address, will introduce two problems:

1) Some MEC sites may be overloaded that downgrade user's experience. The difference between peak and average load in mobile access could reach 5 times. It may cause heavy loads on a particular site, while others will be with light loads. That may cause a longer response time or a higher failure rate for users' requests.

2) Abnormal termination of users' sessions is probable. Considering any cast IP address technology, there is a risk that the packets belonging to an established data flow would be steered to different MEC sites.

To resolve these problems, a new dispatching technology is required. It needs to dispatch users' requests among multiple MEC sites with load balancing support based on the status of computing resources in the sites. That would avoid an unbalanced load among these sites and improve the resource utilization of the sites at the same time. It may require that the network learns the status of computing resources in the MEC sites and propagates the status to enable computing-based dispatching in the ingress router. That needs extensions for routing protocols, including OSPF, IS-IS, and BGP. The dispatching technology may be used together with some tunnel technologies on the data plane, for example, SRv6.

In addition, a technology of session affinity is also required to make sure the traffic of an established data flow will not be wrongly steered to other MEC sites in the middle of communication. That may involve extensions to data plane protocols, including functions added for IPv6 extension headers, and some extensions to IGP and BGP.

IPv6 is more suitable for computer-based networking because of its extensibility on the data plane layer.

These new technologies are under discussion in IETF. They may improve the resource utilization of the sites and provide low-latency quality for distributed edge computing services to improve user experience.

8 Conclusion

IPv6 is the future of any organization. It is possible to postpone the IPv6 transition for many organizations, but it is not possible to cancel it. The world already made excellent progress in the IPv6 transition (1/3 of the global traffic), proving IPv6 to be a mature technology in all the scenarios.

There are many new networking scenarios, some of which are discussed in clause 4 where IPv6 gaps exist because solutions in development may have gaps in general, including gaps not related to IPv6.

The gaps for new solutions are primarily evolving around new requirements for IP on everything:

- Ubiquitous and flexible connectivity.
- Ultra-high bandwidth.
- Deterministic quality.
- Low latency.

- Automation.
- Security.

Clause 5 has the analysis of typical technology challenges for IPv6 transition. The conclusion of this clause that IPv6 is principally different from IPv4. Hence, it is important to account for differences during the design, implementation, and support phases of the networking lifecycle:

- Different concept for the link, subnet, broadcast domain.
- The common practice to use many IPv6 addresses and many prefixes for one host.
- Different address acquisition technology (SLAAC).
- Greatly extended functionality based on enhanced headers.
- And many other differences.

Clause 6 has some insight into the non-technical challenges. It is important to prepare the organization, people, business processes, and tools for the transition to the new technology.

Clause 7 has the overview of new technologies that are in development specifically for the requirements of new scenarios discussed in clause 4:

- New link paradigm for L2 domains with broadcast difficulties (Proactive ND).
- New routing paradigm (SRv6).
- A new way for resources reservation (Slicing).
- New OAM tools (Telemetry).
- Next-generation multicast (BIER).
- New Services tools (signalling in the packet).

Being in the development phase, they have some gaps with a clear roadmap for resolution. It is important to point that IPv6 enhanced innovations discussed in clause 7 do not have equivalent functionality in IPv4. Hence, these gaps do not create the challenge for the transition per ser.

Following the proper guideline and best practices, the IPv6 transition may be almost for free. There are no general gaps that prevent this transition in any scenario. It is time to finalize the IPv6 transition plan to avoid ineffective investments. It is key to grant hardware and software IPv6 readiness at any refreshment cycle.

The current level of IPv6 support by the ecosystem permits to go directly to IPv6-only for the majority of cases. IPv6-only simplifies operations and security aspects, making the solution future-proof and effective.

History

Document history		
V1.1.1	August 2021	Publication