IPv6 Operations
Internet-Draft
Intended status: Informational
Expires: 11 June 2026

G. Ren
W. Zhang
X. Yin
L. He
Tsinghua University
H. Yu
CNNIC

8 December 2025

Measurement and Analysis of IPv6 Interface Identifier Patterns in the Real World
draft-ren-v6ops-ipv6-iid-patterns-measurement-00

Abstract

Interface Identifiers (IIDs) are critical components of IPv6 addresses, significantly impacting user privacy and the feasibility of network reconnaissance. RFC 7707 previously provided a comprehensive analysis of IID patterns based on data from the early stages of IPv6 deployment. However, with the widespread adoption of privacy-enhancing standards such as RFC 7217, historical data no longer accurately reflects the current IPv6 ecosystem. This document provides updated measurements of IID patterns by utilizing an improved pattern recognition method and incorporating novel data sources, such as public mailing lists. The measurement data reveals that while "Low-byte" patterns have decreased significantly in server addresses, a substantial number of seemingly random addresses actually belong to non-random, specific patterns, implying that heuristic scanning remains a viable vector. Furthermore, while client devices have widely adopted randomized addresses-effectively enhancing privacy-Client Premise Equipment (CPE) routers continue to exhibit a high usage rate of IEEE EUI-64 addresses, constituting an often-overlooked privacy risk. This document aims to update the statistics and analysis regarding IID pattern distribution found in RFC 7707, providing essential insights for modern network defense strategies and standard compliance.

About This Document

This note is to be removed before publishing as an RFC.

Status information for this document may be found at https://datatracker.ietf.org/doc/draft-ren-v6ops-ipv6-iid-patterns-measurement/.

Discussion of this document takes place on the v6ops Working Group mailing list (mailto:v6ops@ietf.org), which is archived at https://mailarchive.ietf.org/arch/browse/v6ops/. Subscribe at https://www.ietf.org/mailman/listinfo/v6ops/.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 11 June 2026.

Copyright Notice

Table of Contents

1. Introduction

RFC 7707 [RFC7707] provided a pioneering analysis of network reconnaissance techniques and defense strategies in IPv6 networks. That document discussed the feasibility of address scanning in detail and provided statistical data on the distribution of IPv6 Interface Identifier (IID) patterns for various device types (servers, clients, and routers) at that time. However, the data cited in RFC 7707 was primarily based on measurements conducted around 2012-2013. In the subsequent decade, both the scale of IPv6 deployment and the relevant standards have undergone profound changes.

First, to address the privacy leakage and scanning risks associated with traditional EUI-64 addresses, the IETF has published a series of updated standards. These include RFC 8981 [RFC8981], RFC 7217 [RFC7217] (which defines a method for generating semantically opaque stable addresses), and RFC 8064 [RFC8064] (which recommends deprecating EUI-64 in all cases). Second, the default behaviors of network stacks in mainstream operating systems (such as Windows, Linux, Android, and iOS) have adjusted accordingly, widely adopting these privacy protection mechanisms. Collectively, these factors have resulted in an IPv6 address ecosystem that differs significantly from the era of RFC 7707.

This document aims to update the community's understanding of IPv6 IID allocation through the latest measurement data. By employing a broader range of data sources and a more accurate pattern recognition methodology, this document presents a latest panoramic view of IPv6 address patterns. The data reveals the evolution of server address configuration strategies, the current state of endpoint privacy protection, and potential risks existing within edge network devices, providing a factual basis for updating the relevant sections on address distribution statistics in RFC 7707.

The insights provided in this document are critical for network operators, equipment vendors, and the research community. By revealing the persistence of scannable patterns in infrastructure and

the specific privacy vulnerabilities in edge devices, this document underscores the need for updated defense strategies-moving beyond simple "security by obscurity"-and calls for stricter adherence to standards like RFC 8064 in CPE manufacturing. Furthermore, the updated data models presented here serve as a foundational reference for future network measurements and security assessments.

2. IPv6 Interface Identifiers: Mechanisms, Patterns, and Mapping

This section outlines the generation mechanisms of Interface Identifiers (IIDs) in modern IPv6 networks, the observable sequence patterns, and the mapping relationship between the two.

2.1. Allocation Mechanisms

The generation mechanism of an IID determines the underlying properties of the address. Common allocation mechanisms include:

* Stateless Address Autoconfiguration (SLAAC): Traditionally based on the IEEE EUI-64 specification, expanding a 48-bit MAC address and embedding it into the IID [RFC4862].

* Temporary Addresses: To protect privacy, operating systems periodically generate random IIDs that change over time [RFC8981].

* Stable Opaque Addresses: Generates a stable IID per prefix that exhibits random characteristics, intended to replace EUI-64 as the default configuration [RFC7217].

* DHCPv6: Stateful address assignment managed by a server. Server policies can be sequential, random, or based on specific algorithms [RFC8415].

* Manual Configuration: Administrators manually specify static addresses, commonly used for server and router interfaces, often employing patterns that are easy to remember.

* Transition Technologies: Mechanisms such as Teredo [RFC4380] or ISATAP [RFC5214], which generate IIDs containing IPv4 address or port information via specific algorithms.

2.2. Interface Identifier Sequence Patterns

"Sequence Pattern" refers to the byte structure characteristics of an IID as perceived by an external observer. RFC 7707 established a well-known taxonomy for these patterns. This document adopts and extends this taxonomy. The primary patterns are listed below (roughly in order of identification priority):

1. Transition Technology: IIDs conforming to specific transition protocol specifications (e.g., ISATAP addresses beginning with 0000:5efe [RFC5214]).

2. IEEE-based: IIDs conforming to the EUI-64 format, containing ff:fe in the middle, with the first three bytes corresponding to a valid vendor OUI [RFC4862].

3. Embedded-IPv4: IIDs containing a complete IPv4 address.

4. Embedded-Port: IIDs where the low-order bits contain common service port numbers (e.g., 80, 443), with the remainder typically being zero. This is essentially a special case of the Low-byte pattern.

5. Low-byte: IIDs where the high-order bits are all zero, and only the lowest bytes (typically a small portion of the final 64 bits) are non-zero.

6. Byte-pattern: IIDs containing a large number of zero bytes (e.g., more than 3 bytes are 00), exhibiting sparse characteristics but not fitting the Low-byte definition.

7. Seed-Similar (New): A new classification introduced in this document. This refers to IIDs that would traditionally be classified as "Randomized" but are identified as having non-random characteristics through a specific algorithm (detailed in subsequent sections).

8. Randomized: IIDs remaining after filtering out all the above rules. They exhibit no obvious observable structure.

The original methodology of RFC 7707 (implemented in tools like addr6 [IPv6-Toolkit]) classified all addresses remaining after the first six filtering steps as "Randomized". This approach resulted in the misclassification of many non-random, manually configured addresses (such as ffff:ffff:ffff:ffff or specific wordy addresses). By introducing "Seed-Similar" Patterns, this document aims to further strip away non-random components from these "remaining addresses", thereby more accurately assessing the entropy of the address space.

2.3. Mapping Patterns to Mechanisms

While the sequence pattern of an IID is publicly visible, the specific mechanism generating it is often opaque. A single pattern may be produced by multiple different mechanisms. The table below summarizes the most typical correspondences between sequence patterns and allocation mechanisms in the current network environment:

Sequence Pattern	Primary Mechanisms	Notes
Transition	Teredo [RFC4380], ISATAP [RFC5214], etc.	Depends on specific transition protocol specs.
IEEE-based	SLAAC (Traditional EUI-64)	Decreased in modern clients, but still common in CPEs.
Embedded-IPv4	Manual, Some Transition Techs	Common in dual-stack network planning.
Embedded-Port	Manual Configuration	Configured by admins for service mnemonics.
Low-byte	Manual, DHCPv6 (Sequential)	Mainstream for infrastructure/servers; easy to scan.
Byte-pattern	Manual Configuration	Contains many zeros but not strictly Low-byte.
Randomized	Temporary [RFC8981], Stable Opaque [RFC7217], DHCPv6 (Random)	Pattern cannot distinguish between "Temporary" and "Stable" devices.
Seed-Similar	Manual, DHCPv6 (Specific Algo)	Previously misclassified as random; likely follows specific organizational norms.

Table 1: Mapping between Sequence Patterns and Allocation Mechanisms

3. Measurement Methodology

This section details the methods used to collect IPv6 address data and analyze IID patterns. To update the statistics in RFC 7707, we have not only expanded the scope of data sources but also improved the IID pattern recognition algorithm to more accurately assess the randomness of the address space.

3.1. Data Sources

To comprehensively cover different types of IPv6 devices (servers, clients, and routers), we utilized multiple data collection channels. In particular, obtaining address data with temporal attributes via public mailing lists provides a new perspective for analyzing the evolutionary trends of IID patterns.

* Public Domain Names: To measure server addresses, we utilized multiple public top-level domain lists (such as Alexa Top 1M, OpenIntel, Tranco, etc.). By performing DNS queries for AAAA records (Web servers), MX records (Mail servers), and NS records (Name servers) on these domains, we collected a large-scale set of server IPv6 addresses. This continues the traditional measurement approach of RFC 7707, ensuring data consistency and comparability.

* BitTorrent DHT Network: To obtain client addresses of end-users, we participated in the BitTorrent network. By deploying passive nodes, we collected active client IPv6 addresses. This method does not rely on server logs and can more directly reflect the address configuration of end-users. This methodology was inspired by [Draft-P2P].

* Traceroute Probes: To measure network infrastructure (router) addresses, we performed Traceroute probes on all advertised BGP prefixes, continuing the traditional approach of RFC 7707. Additionally, we performed traceroutes to the collected server and client addresses. Specifically, we distinguished the edge router (the last hop), which is crucial for analyzing the configuration habits of Customer Premises Equipment (CPE).

* Public Mailing Lists: This is a novel data source introduced in this document. Many open-source communities and organizations maintain public mailing list archives. Email header information (Headers) typically contains the IP address of the sending client as well as the addresses of Mail Transfer Agent (MTA) servers along the path.

  - Advantage: This data not only distinguishes between clients and servers but, more importantly, carries explicit timestamps (Date header). This allows us to construct a longitudinal dataset spanning over a decade, thereby tracking the evolutionary trends of IID patterns over time (e.g., the adoption process of RFC 7217).

3.2. Interface ID Pattern Recognition Methodology

Regarding pattern recognition, we largely followed the methodology established in RFC 7707 (specifically the logic implemented in the addr6 tool [IPv6-Toolkit]). However, as noted previously, the traditional method lumps all addresses not matching specific rules into "Randomized", leading to a high false-positive rate. To address this, we added a recognition step for "Seed-Similar Patterns" at the end of the original identification flow-specifically, before classifying an address as "Randomized".

Recognition Principle:

The core idea of this method is based on statistical probability: if an IID to be tested is generated via a cryptographic algorithm or random generator (i.e., true random), the probability of it colliding with or exhibiting high similarity to any other known IID (whether random or manually configured) in a 64-bit space is negligible. Conversely, if an IID exhibits significant similarity to an IID in a known address list, we can conclude that the IID is highly likely non-randomly generated (e.g., it may be a variation of manual configuration, specific organizational norms, etc.).

Implementation:

1. Seed List Construction: We first construct a large-scale "Seed Address List" based on all addresses collected in Section 3.1.

2. Similarity Detection: For any IID remaining after filtering through the preceding rules, we compare it against the IIDs in the seed list.

    * Criterion: Theoretically, calculating the Hamming Distance between two IIDs is an accurate measure of similarity. However, calculating pairwise Hamming Distances on datasets of hundreds of millions scales poorly. Therefore, we adopted a more efficient heuristic rule: if the first 4 bytes or the last 4 bytes of two IIDs are identical, they are determined to have similarity.

3. Classification Decision: If the IID under test is determined to be similar to an IID (from a different prefix) in the seed list, it is classified as "Seed-Similar"; otherwise, it is finally classified as "Randomized".

Validation:

By introducing this improvement step, we successfully identified a large number of manually configured addresses that were previously misclassified as random (e.g., significantly non-random addresses like ffff:ffff:ffff:abcd). In our tests on the server dataset, this method reduced the proportion of addresses originally flagged as "Randomized" by approximately 69%. This indicates that RFC 7707 indeed significantly overestimated the randomness of server addresses, and the measurement results of this method are closer to the true state of network configuration.

4. Measurement Results and Analysis

This section presents the measurement results of IPv6 IID patterns based on data collected in 2024, compared with historical data cited in RFC 7707 (circa 2012). By analyzing this data, we evaluate the current state of IPv6 address scanning feasibility and privacy risks in the real world.

4.1. Servers

The distribution of IID patterns for server addresses shows significant evolution, particularly in the decline of easily predictable patterns. The table below displays the distribution for Web servers, Mail servers, and Name servers (NS):

Type	Randomized	Seed-Similar	Embedded-IPv4	Byte-pattern	IEEE-based	Port-Embed	Low-byte
Web	21.52%	47.93%	12.75%	8.76%	0.27%	0.40%	8.36%
NS	1.86%	4.62%	20.62%	4.38%	1.07%	6.86%	59.52%
Mail	3.22%	13.06%	27.45%	3.52%	1.53%	3.50%	46.11%

Table 2: IID Pattern Distribution in Server Addresses

The most significant change observed is the marked decline in Low-byte patterns within Web and Mail servers (compared to ~90% in the RFC 7707 era). In the past, attackers could discover the vast majority of servers by simply scanning a small range (e.g., ::1 through ::ff). The current data suggests that the hit rate for such simple linear scans has dropped drastically.

However, the difficulty of scanning is not as high as the raw "Randomized" numbers might suggest. Our improved algorithm reveals that a large number of server addresses (approx. 46% in Web servers)

actually fall into "Seed-Similar". These are addresses that, while not strictly Low-byte, follow specific organizational templates or non-random sequences. Consequently, while simple brute-force scanning is becoming less effective, address scanning remains feasible through Target Generation Algorithms (TGA) [_6Gen-TGA] which can leverage these patterns to discover targets.

4.2. Clients

Privacy protection for client devices has been a primary focus of IETF standardization efforts. We measured client IIDs using the BitTorrent dataset (BT-Client) and the Public Mailing List dataset (Mail-Client).

Dataset	Randomized	Seed-Similar	Embedded-IPv4	Byte-pattern	IEEE-based	Port-Embed	Low-byte
Mail-Client (2024)	86.93%	0.65%	2.27%	0.97%	1.51%	0.32%	7.34%
BT-Client	77.96%	1.96%	2.44%	2.20%	8.10%	0.11%	7.15%

Table 3: IID Pattern Distribution in Client Addresses

Notably, the proportion of IEEE-based patterns in the BT-Client dataset (~8.10%) is significantly higher than in the Mail-Client(2024) dataset (~1.51%). In-depth analysis suggests this discrepancy arises because the BitTorrent network contains not only typical user endpoints but also a large number of NAS devices and home routers running embedded BT clients. These embedded devices often lag in firmware updates and still utilize traditional SLAAC EUI-64 configurations. Therefore, we consider the Mail-Client dataset to be a more representative reference for the general population of end-user client devices.

Longitudinal data based on Mail-Client shows that the usage of IEEE-based (EUI-64) addresses has dropped significantly from approximately 8.87% a decade ago to 1.51% currently. This indicates that RFC 8981 [RFC8981] (Temporary Addresses) and RFC 7217 [RFC7217] (Stable Opaque Addresses) have been widely and effectively deployed in modern operating systems (Windows, Android, iOS, Linux). This shift significantly mitigates the risk of attackers tracking specific users or identifying device manufacturers directly via endpoint IIDs. The table below shows the evolutionary trend of client IID patterns, clearly reflecting the success of privacy technologies:

Year	Randomized	IEEE-based
2013	~79.14%	~8.87%
2016	~82.50%	~5.20%
2020	~85.10%	~2.30%
2024	~86.93%	~1.51%

Table 4: Evolutionary Trend of Client IID Patterns (Randomized and IEEE-based) from Mailing Lists

From a scanning perspective, the dominance of Randomized patterns (over 85%) makes discovering specific client endpoints via wide-range scanning extremely difficult. However, it is important to note that approximately 7% of client addresses still follow Low-byte patterns (e.g., ::1, ::2). This suggests that a non-negligible fraction of client devices-potentially manually configured workstations or servers within client networks-remain vulnerable to simple brute-force scanning techniques. Attackers may specifically target this subset of addresses to gain an initial foothold in client networks.

4.3. Routers

Router address measurements reveal a massive discrepancy in configuration strategies between the general network infrastructure and the client edge network.

Dataset	Randomized	Seed-Similar	Embedded-IPv4	Byte-pattern	IEEE-based	Port-Embed	Low-byte
Router	2.65%	3.19%	12.29%	12.14%	1.87%	3.02%	64.83%
Client-Edge-Router	36.07%	2.68%	5.91%	6.21%	17.66%	0.45%	31.02%

Table 5: IID Pattern Distribution in Router Addresses

In general routers (derived from traceroutes to BGP prefixes), Low-byte patterns remain the absolute mainstream (~65%). Additionally, Embedded-IPv4 patterns have accounted for ~12.29%, likely due to dual-stack deployment strategies. This implies that brute-force scanning against network infrastructure (e.g., targeting ::1 or ::router) remains largely effective and is a viable reconnaissance vector.

For Client Edge Routers (CPEs), scanning is relatively more difficult due to a higher proportion of Randomized and IEEE-based patterns. However, Low-byte patterns still account for approximately 31% (nearly one-third) of edge devices. This indicates that while less vulnerable than the general infrastructure, a significant portion of home gateways can still be discovered using traditional scanning methods targeting small ranges.

A critical finding is that approximately 17.66% of CPE devices still default to using IEEE-based patterns. This behavior constitutes a significant privacy risk. The EUI-64 address directly exposes the device manufacturer (via OUI) and provides a stable identifier that allows external observers to track the entire home network over time, effectively functioning as a "Super Cookie". This highlights the urgency of enforcing RFC 8064 [RFC8064] on edge devices to eliminate this residual privacy vulnerability.

5. Security Considerations

The implications of the observed IID patterns on network reconnaissance and user privacy (specifically regarding address scanning feasibility and CPE privacy risks) are discussed in detail in Section 4.

Regarding the measurement methodology itself, this study adhered to ethical research principles to minimize impact on the network. Active measurements (such as traceroutes) were rate-limited to avoid

congestion. For passive data collection from public mailing lists, only IP address information and timestamps were extracted; no personally identifiable information (PII), such as email addresses or message content, was stored or analyzed.

6. Conclusion

The data in this document indicates that IPv6 address Interface Identifier allocation patterns have undergone tremendous changes. While the general decrease in Low-byte patterns has increased the difficulty of traditional brute-force scanning, it remains feasible to discover the vast majority of servers and routers using heuristic methods. Furthermore, the configuration lag in edge routers remains a shortcoming in privacy protection. Future network measurements and security assessments should be based on these updated data models.

7. IANA Considerations

This document has no IANA actions.

8. References

8.1. Normative References

[RFC4380] Huitema, C., "Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)", RFC 4380, DOI 10.17487/RFC4380, February 2006, <https://www.rfc-editor.org/rfc/rfc4380>.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, DOI 10.17487/RFC4862, September 2007, <https://www.rfc-editor.org/rfc/rfc4862>.

[RFC5214] Templin, F., Gleeson, T., and D. Thaler, "Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)", RFC 5214, DOI 10.17487/RFC5214, March 2008, <https://www.rfc-editor.org/rfc/rfc5214>.

[RFC7217] Gont, F., "A Method for Generating Semantically Opaque Interface Identifiers with IPv6 Stateless Address Autoconfiguration (SLAAC)", RFC 7217, DOI 10.17487/RFC7217, April 2014, <https://www.rfc-editor.org/rfc/rfc7217>.

[RFC7707] Gont, F. and T. Chown, "Network Reconnaissance in IPv6 Networks", RFC 7707, DOI 10.17487/RFC7707, March 2016, <https://www.rfc-editor.org/rfc/rfc7707>.

[RFC8064] Gont, F., Cooper, A., Thaler, D., and W. Liu, "Recommendation on Stable IPv6 Interface Identifiers", RFC 8064, DOI 10.17487/RFC8064, February 2017, <https://www.rfc-editor.org/rfc/rfc8064>.

[RFC8415] Mrugalski, T., Siodelski, M., Volz, B., Yourtchenko, A., Richardson, M., Jiang, S., Lemon, T., and T. Winters, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 8415, DOI 10.17487/RFC8415, November 2018, <https://www.rfc-editor.org/rfc/rfc8415>.

[RFC8981] Gont, F., Krishnan, S., Narten, T., and R. Draves, "Temporary Address Extensions for Stateless Address Autoconfiguration in IPv6", RFC 8981, DOI 10.17487/RFC8981, February 2021, <https://www.rfc-editor.org/rfc/rfc8981>.

8.2. Informative References

[Draft-P2P] Defeche, M. and E. Vyncke, "Measuring IPv6 Traffic in BitTorrent Networks", Work in Progress, Internet-Draft, draft-vyncke-ipv6-traffic-in-p2p-networks-01, 2009, <https://datatracker.ietf.org/doc/html/draft-vyncke-ipv6-traffic-in-p2p-networks-01>.

[IPv6-Toolkit] Gont, F., "SI6 Networks' IPv6 Toolkit", 2013, <https://github.com/fgont/ipv6toolkit>.

[_6Gen-TGA] Murdock, A., Li, F., Bramsen, P., Durumeric, Z., and V. Paxson, "Target Generation for Internet-wide IPv6 Scanning", ACM IMC 2017, 2017.

Authors' Addresses

Gang Ren
Tsinghua University
Email: rengang@cernet.edu.cn

Wei Zhang
Tsinghua University
Email: zhang-w22@mails.tsinghua.edu.cn

Xia Yin
Tsinghua University
Email: yxia@tsinghua.edu.cn

Lin He
Tsinghua University
Email: helin1170@gmail.com

Haisheng Yu
CNNIC
Email: yuhaisheng@cnnic.cn