HAW HAMBURG

HAUPTSEMINAR
Isabell Egloff

Analyse und Charakterisierung von Scanner-Aktivitäten im IPv6-Adressraum

FAKULTÄT TECHNIK UND INFORMATIK
Department Informatik

Faculty of Engineering and Computer Science
Department Computer Science

Betreuung durch: Prof. Dr. Thomas C. Schmidt / Raphael Hiesgen
Eingereicht am: 26. Juni 2024

HOCHSCHULE FÜR ANGEWANDTE
WISSENSCHAFTEN HAMBURG
Hamburg University of Applied Sciences

Inhaltsverzeichnis

1 Einleitung

IPv6 ist der Nachfolger von IP Version 4 (IPv4) [14] und damit die neue Version des Internetprotokolls. Ein großer Vorteil der Einführung von IPv6 ist die erweiterte Adressierungsmöglichkeit. IPv6 erhöht die IP-Adressgröße von 32 Bit auf 128 Bit [3]. Diese Größe bringt aber auch neue Herausforderungen. Mit Tools wie ZMap [4] lässt sich der IPv4-Adressraum in unter 45 Minuten komplett durchscannen. Der IPv6-Adressraum ist zu groß, um ihn komplett zu durchscannen. Die Forschung, kommerzielle Dienste und böswillige Akteure verfolgen daher mehrere unterschiedliche Strategien, um Teile des IPv6-Adressraums möglichst effizient zu scannen. Diese Scan-Stratgien werden wiederum von anderen beobachtet und analysiert. In dieser Arbeit wird eine solche Analyse durchgeführt, um mit Hilfe der Beobachtungen Aussagen über das IPv6-Scan-Verhalten treffen zu können.

Im Vorfeld dieser Arbeit wurde sich intensiv mit der Analyse des IPv6-Scan-Verhaltens beschäftigt, einschließlich der Ermittlung des aktuellen Forschungsstands. Infolgedessen wurden Methoden zur Untersuchung von IPv6-Paketen und IPv6-Scannern entwickelt. Unter anderem wurde das reaktive Netzwerk-Teleskop Spoki [9] erweitert, um neben IPv4 auch auf IPv6-TCP-Pakete reagieren zu können und somit IPv6-Scanner besser analysieren zu können als bisher. Darüber hinaus wurde der Netzwerkverkehr von vier unterschiedlichen Präfixen analysiert, wobei das reaktive Netzwerk-Teleskop auf einem spezifischen Präfix eingesetzt wurde.

In dieser Arbeit wird nun das IPv6-Scan-Verhalten anhand der vier Netzwerkpräfixe untersucht. Diese Präfixe wurden über verschiedene Zeiträume (9 Monate, 7 Monate, 4 Monate und 2 Monate) hinweg gemessen und weisen unterschiedliche Eigenschaften auf. Es wird analysiert, welche Auswirkungen diese Eigenschaften auf die Paketankunft in den Präfixen haben könnten.

Die Struktur dieser Arbeit ist wie folgt gestaltet: In Abschnitt 2 wird zunächst die Problemstellung erläutert, gefolgt von der Vorstellung des theoretischen Hintergrunds in Abschnitt 3. Anschließend werden in Abschnitt 4 verwandte Arbeiten zu diesem Thema ausführlich dargelegt, wodurch der bisherige Stand der Forschung veranschaulicht wird. In Abschnitt 5 werden die Begriffe *Scanner* und *Scan* im Kontext dieser Arbeit definiert. Zusätzlich dazu werden die Eigenschaften der Präfixe vorgestellt. Die Analyse der zwei /48-Subnetze und die daraus gewonnenen Erkenntnisse werden in Abschnitt 6 erläutert.

Anschließend werden in Abschnitt 7 die Ergebnisse des /48-Präfixes präsentiert, gefolgt von den Ergebnissen des /32-Präfixes in Abschnitt 8. Die wichtigsten Erkenntnisse der gesamten Analyse werden dann in Abschnitt 9 zusammengefasst, während Abschnitt 10 eine abschließende Zusammenfassung sowie einen Ausblick auf mögliche zukünftige Forschungsarbeiten bietet.

2 Problemstellung

Aufgrund der Größe des IPv6-Adressraums sind IPv6-Scans auf Teiladressbereiche beschränkt. Genauso können auch nur Teiladressbereiche beispielsweise mit Netzwerkpräfixen untersucht werden. Ein zentraler Aspekt dieser Arbeit ist das Anziehen von Scannern durch verschiedene Eigenschaften von Präfixen. Durch Nutzung dieser Eigenschaften sollen gezielt Netzwerkpakete angezogen werden, um zu analysieren, wie Scanner auf die Präfixeigenschaften reagieren. Zusätzlich dazu ist auch die Analyse der Scans an sich ebenso wertvoll, um zu untersuchen, welche Scan-Strategien verfolgt werden.

3 Hintergrund

Zufällig eine aktive IPv6-Adresse in einem Adressraum mit 2^{128} Adressen zu finden, ohne Berücksichtigung bekannter annoncierter Präfixe oder die Anwendung von verschiedenen Algorithmen, die Millionen von Adressen innerhalb kurzer Zeit abfragen können, ist äußerst unwahrscheinlich. Eine weitere Möglichkeit ist nach manuell konfigurierten Adressen zu suchen [8]. (*i*) Bei *Low-Byte-Adressen* werden alle Bytes des Interface Identifier (IID), mit Ausnahme der niedrigstwertigen Bytes, auf Null gesetzt. Wie beispielsweise *2001:db8::1*. Und es kommt vor, dass beide 16-Bit-Wörter der niedrigsten Ordnung gesetzt sind, wie in *2001::db8::1:10*. (*ii*) Service-Port-Adressen betten den Port eines laufenden Dienstes in die IID, wie in *2001:db::443* für HTTPS ein. (*iii*) IPv4-eingebettete-Adressen betten die IPv4-Adresse der Netzwerkschnittstelle in die IID ein, zum Beispiel *2001:db8::192.0.0.1*. (*iv*) Sogenannte *Wordy-Adressen* beinhalten Wörter wie *2001:db8::cafe*.

Ullrich *et al.* [20] waren die ersten, die musterbasierte Scan-Techniken veröffentlichten. In dieser Arbeit werden die Zieladressen der empfangenen Pakete analysiert und Rückschlüsse aus den Zieladressmustern gezogen.

4 Verwandte Arbeiten

Während auf der einen Seite Scanner den IPv6-Adressraum scannen, wird auf der anderen Seite dieses Vorgehen analysiert. In verwandten Arbeiten wurden aus beiden Perspektiven bereits unterschiedliche Vorgehensweisen untersucht.

Mit der Arbeit von Ford *et al.* [5] begann 2005 die Analyse der IPv6-Scanner. Sie untersuchten die Hintergrundstrahlung eines /48-Präfixes. Zu der Zeit war das neue Internetprotokoll (IPv6) noch nicht sehr verbreitet. Aus diesem Grund erhielten sie auch nicht mehr als 12 Pakete.

Mit dem gleichen /48-Präfix wurde 2023 von Ronan *et al.* [16] ein erneutes Messexperiment gestartet. Zu der Zeit wurden 5 Tausend Pakete innerhalb von sechs Monaten empfangen. Enthalten waren 74 % ICMPv6-Pakete, 21 % TCP-Pakete und 4 % UDP-Pakete.

Czyz *et al.* [2] analysierten 2013 die IPv6-Hintergrundstrahlung von fünf angekündigten /12-Adressblöcken, die den fünf regionalen Internet Registries (RIRs) zugewiesen waren. Sie sammelten ausschließlich den Darknetverkehr, also die Pakete zu den Subpräfixen der /12, die nie zugewiesen oder geroutet wurden. Insgesamt entsprach das 5 % (209 Millionen) von allen empfangenen Paketen. Dabei beobachteten sie RIR-Präfixe (ARIN, APNIC), bei denen schon ein bis zwei Tausend Quell-IP-Adressen 90 % des Datenverkehrs ausmachten. Wobei bei den anderen (AFRINIC und LACNIC) mehr als 10 Tausend bzw. 100 Tausend Quell-IP-Adressen zusammen erst 90 % der Pakete abdeckten.

2018 schlugen Fukuda *et al.* [6] vor, DNS-Backscatter zu verwenden, um IPv6-Scan-Verhalten zu erkennen. Sie identifizierte Scanner auf der Grundlage von DNS-Reverse-Lookups, die von Routern durchgeführt wurden. Die Methodik wurde durch den Vergleich von Scannern mit MAWI-Daten und einem /37 IPv6-Darknet verifiziert. Über einen Zeitraum von sechs Monaten identifizierten sie 16 aktive IPv6-Scanner pro Woche. Aus dem analysierten Scan-Verhalten konnten sie drei Scan-Methoden ableiten: (*i*) das Scannen von *Low-Byte-Adressen*, (*ii*) Scanner die auf IPs abzielen, die über Reverse DNS auffindbar sind und (*iii*) die Verwendung des Zielgenerierungsalgorithmus 6Gen von Murdock *et al.* [13].

Strowes *et al.* [18] analysierten 2020 den Netzwerkverkehr eines neu annoncierten /12-Präfixes. Es wurden vier /32- und vier /48-Präfixe aus einem /29-Covering-Präfix separat

annonciert. 95 % der fast 5,5 Millionen TCP-Traceroute-Pakete stammten von einem einzelnen AS. 54 % der übrigen Pakete zielten auf die Antwortadressen ab, von denen eine als Testadresse auf einer Mailingliste beworben wurde. Die verbleibenden 46 % der Pakete wurden breiter innerhalb des Präfixes verteilt und bevorzugt an Low-Byte-Adressen versendet. Die Autoren hatten den Großteil ihres aufgezeichneten Datenverkehrs selbst gesendet, um eine tiefer gehende Routenfilterungsanalyse durchzuführen.

Liu *et al.* [11] analysierten 2021 ebenfalls die Hintergrundstrahlung eines vorher ungenutzten /20 Präfixes. Sie beobachteten 2,9 Millionen Pakete innerhalb von sechs Monaten. Sie sahen am meisten ICMPv6-Pakete (67 %), darauf folgte die Menge an TCP-Paketen (33 %) und schließlich sahen sie am wenigsten UDP-Pakete (< 1 %). Dabei versendeten nur 10 Quelladressen schon 95 % aller empfangenen Pakete.

Richter *et al.* [15] analysierten 2022 das IPv6-Scan-Verhalten mithilfe der Firewallprotokolle eines großen CDN. Dabei wurden ICMP-Pakete ausgeschlossen, genauso wie TCP- und UDP-Pakete mit Zielportangaben 80 oder 443. Sie stellen bei der Analyse fest, dass die beiden aktivsten Scanner beide von Rechenzentrums-ASes in China stammten. Darauf folgte ein Cybersicherheitsunternehmen aus den USA und eine Reihe von globalen Hosting- und Cloud-Anbietern. Wobei die Top-5-Quell-ASes bereits 93 % der Scan-Pakete ausmachten.

Tanveer *et al.* [19] untersuchten wie IPv6-Hostaktivitäten (Web-Crawls, NTP-Pool-Server, öffentliche NTP-Server, Tor, DNS-Anfragen, DNS-Zonen) das Verhalten von Scannern im IPv6-Adressraum beeinflussen. Dabei analysieren sie den Netzwerkverkehr eines vorher noch ungenutzten /56-Subnetzes. Jedes der sechs Experimente wurde auf vier zufällig ausgewählte /64-Subnetze bereitgestellt. Es zeigte sich, dass öffentlich sichtbare aktive Dienste (NTP, Tor, DNS-Zones) deutlich mehr Scan-Aktivitäten anzogen als Kommunikationen, die von den Teleskopen ausgehen (Web-Crawls und DNS-Anfragen). Sie stellten fest, dass die Quelladressen zufällige Adressen scannten oder sie sich auf *Low-Byte-Adressen* konzentrierten. 65 % aller Scanner verwendeten nur eine der beiden Strategien. Die *Low-Byte-Scanner* generierten nur 4 % aller empfangene Pakete, während die Scanner, die zufällige IIDs scannten nur 5 % generierten. Die verbleibenden 91 % der Pakete wurden von Scannern versendet, die eine Mischung aus beiden Strategien anwendeten.

Zhao *et al.* [21] analysierten Auswirkung der Adressoffenlegung über DNS mithilfe eines zuvor ungenutzten /56-Netzwerks. Sie veröffentlichten Adressen über vier Methoden (*i*) IPv6-Adressen, die Domänen zugeordnet sind, die über IPv4-PTR-Einträge verfügen,

(*ii*) PTR-Eintrag für zufällige Adressen, (*iii*) PTR-Datensätze für *Wordy*- und Port-eingebettete Adressen und (*iv*) IPv6-Adressen mit beliebten Domänennamen. Jede Methode wurde jeweils in einem /64-Darknet und jeweils in einem /64-Honeynet eingesetzt. Die Zuordnung von IPv6-Adressen zu Domänen mit IPv4-PTR-Einträgen machten über 99,99 % der Scans aus. Die Scans im Darknet konzentrierten sich auf bekannte Adressen (97 %), zielten aber auch auf zufällige, nicht bekannte Adressen ab. In Honeynet waren die Scanner weniger fokussiert und zielten gleichmäßiger auf bereits bekannte und unbekannte Adressen in der Nähe der Bekannten ab. Low-Byte-Scans machten nur 0,03 % aller Scans aus.

In dieser Arbeit wird ebenfalls das Scan-Verhalten von IPv6-Scannern analysiert, jedoch wird der Netzwerkverkehr von zwei Subnetzen und zwei annoncierten Präfixen miteinander verglichen, die alle jeweils andere Eigenschaften besitzen. Es wird untersucht, inwieweit die jeweiligen Präfixeigenschaften Einfluss auf den Netzwerkverkehr haben.

5 Methodik

In dieser Arbeit werden die folgenden vier Präfixe (P1, P2, P3, P4) genutzt, um IPv6-Netzwerkverkehr zu untersuchen. Jedes Präfix besitzt unterschiedliche Eigenschaften. Es wird daher ebenfalls untersucht, welchen Einfluss die Eigenschaften der jeweiligen Präfixe auf den empfangenen Netzwerkverkehr haben. Ein Überblick über die Eigenschaften wird in der Tabelle 1 dargestellt.

5.1 Scans und Scanner

Sämtliche Netzwerkpakete, die empfangen werden, werden als PCAPs gesammelt. Bei der Analyse wird jede Quelladresse für sich betrachtet. Daher werden in dieser Arbeit keine Aggregationen von Quelladressen in die Analyse mit einbezogen. Es werden keine Scans identifiziert. Stattdessen werden bei genaueren Untersuchungen der Zieladressen die Pakete eines Tages oder eine Stichprobe eines Tages visualisiert, an denen es ein relativ hohes Verkehrsaufkommen des jeweiligen Scanners gab. Die Scanner werden in dieser Arbeit mit der dazugehörigen AS-Organisation benannt, die mit Hilfe des Tools pyasn[1] identifiziert wird. Das Tool pyasn ist ein Python-Erweiterungsmodul, das eine sehr schnelle Suche nach IP-Adressen und Autonome Systemnummern (ASN) ermöglicht.

[1] https://catalog.caida.org/software/pyasn

5.2 Präfixeigenschaften

Die Eigenschaften der Präfixe umfassen die Größe des Präfixes, den Messbeginn, das Datum des Announcements, mögliche Aktivität innerhalb des Präfixes und das Auftauchen auf der Hitliste der TU München [7, 22]. Dies ist eine sehr beliebte IPv6-Adressensammlung von aktiven Hosts. Es gibt insgesamt drei Listen: *Responsive-Adressen*, *Aliased-Präfixe* und *Non-aliased-Präfixe*. P2, P3 und P4 befanden sich bereits vor Beginn der Messungen auf der Hitliste, während P1 fünf Tage nachdem es annonciert wurde dort auftaucht. Mit der Veröffentlichung in der Hitliste ist die *Non-aliased-Präfix-Liste* gemeint. Der Nutzer der Hitliste weiß damit, dass die dortigen Einträge jeweils exklusiv für ein bestimmtes Netzwerk genutzt werden.

Tabelle 1: Eigenschaften der Präfixe

Name	**Größe**	**Mess-beginn**	**Announcement**	**Aktivität**	**Eintrag in der TUM-Hitliste**
P1	/32	24.08.2023	Zeitgleich mit Messbeginn	Nein	Seit dem 29.08.2023
P2	/48	30.06.2023	Vor Messbeginn	Aktives /56-Subnetz	Vor Messbeginn
P3	/48	31.01.2023	Das /29-Präfix vor Messbeginn	Reagiert auf TCP-Anfragen	Das /29-Präfix vor Messbeginn
P4	/48	30.03.2023		Nein	

P1: Neu annonciertes /32-Präfix. Am 24. August 2023 ist das /32-Präfix erstmals annonciert worden. Alle Adressen in diesem Präfix sind passiv und keine ist einem Endpunkt zugewiesen. Seit dem 29. August 2023 befindet es sich auf der Hitliste.

P2: /48-Präfix mit aktivem /56-Subnetz. Die Messung des ersten annoncierten /48-Präfix startet am 30. Juni 2023. Es ist vor 13 Jahren annonciert worden. Eines der /56-Subnetze wird produktiv genutzt und hostet Dienste wie Webserver und IoT-Geräte. Dieses /56-Subnetz wird nicht gesondert bekannt gegeben. Die empfangenen Pakete des /56-Subnetzes werden aus der Analyse ausgeschlossen und nicht mit betrachtet.

P3: Reaktives /48-Subnetz. Die Messung des ersten /48-Subnetzes startet am 31. Januar 2023. Es ist Teil eines /29-Covering-Präfixes. Im Gegensatz zu P4 nimmt es mit Einsatz des reaktiven Netzwerk-Teleskops Spoki seit dem 16. Mai 2023 aktiv TCP-Verbindungen entgegen und reagiert auf Scan-Anfragen.

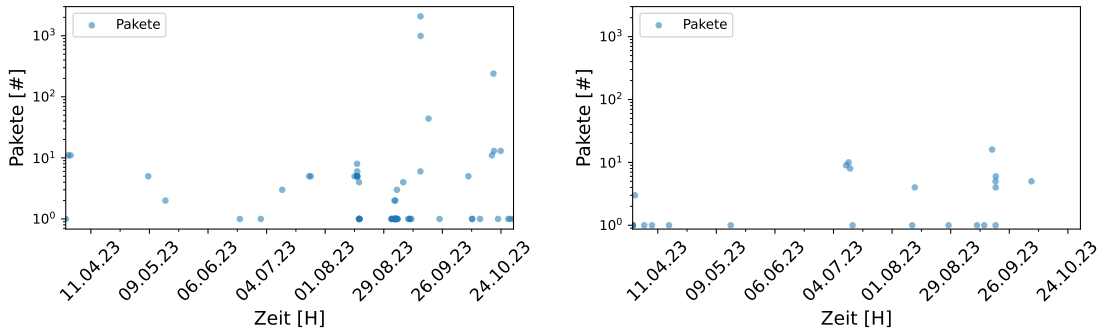
P4: Passives /48-Subnetz. Die Messung des anderen /48-Subnetzes wird am 30. März 2023 gestartet. Es ist ebenfalls ein Subnetz des /29-Covering-Präfixes und es besitzt nur passive Adressen. Das /29-Covering-Präfix befand sich vor Messbeginn bereits in der öffentlichen Hitliste der TU München.

6 Das passive und das reaktive Subnetz

Am 31. Januar 2023 beginnt die erste Messung mit P3. Auf P3 läuft seit dem 16. Mai 2023 das reaktive Netzwerk-Teleskop Spoki, weshalb es seitdem auf TCP-Anfragen reagiert. Die zweite Messung startet am 30. März 2023 mit P4. Beides sind Subnetze eines /29-Covering-Präfixes. Die /48-Subnetze könnten über die Veröffentlichung des /29-Covering-Präfix in der Hitliste gefunden werden. Für den Vergleich der Subnetze wird der Zeitraum ab dem Messbeginn von P4 am 30. März 2023 bis 30. Oktober 2023 festgelegt. In dieser Zeit empfängt P3 über 3.500 Pakete, während P4 innerhalb des Messzeitraums 83 Pakete empfängt.

6.1 Beobachtung des Netzwerkverkehrs

In der Abbildung 1 wird die Paketankunft pro Stunde dargestellt. Die Paketmenge wird auf der y-Achse logarithmisch dargestellt.



(a) P3 (b) P4

Abbildung 1: Paketankunft pro Stunde

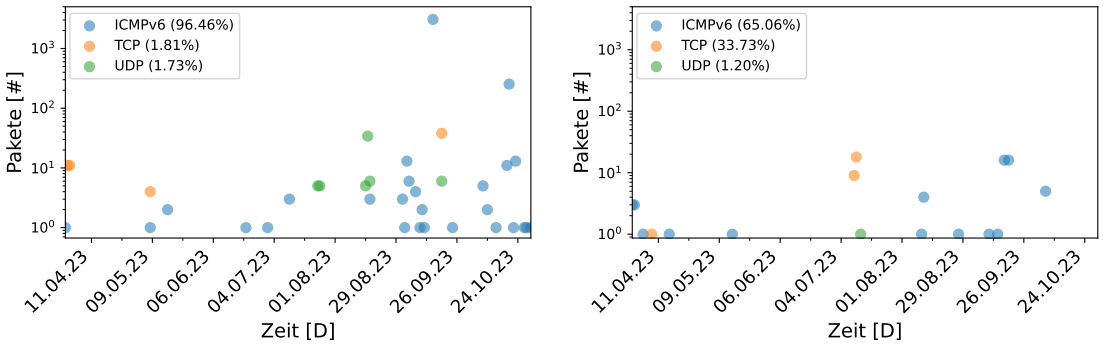
Im Diagramm 1a ist ein signifikanter Anstieg auf über 3 Tausend Pakete am 15. September 2023 in P3 zu erkennen. An den übrigen Tagen ist die Anzahl der Pakete pro

Stunde deutlich geringer. Jedoch ist kein merklicher Anstieg der Paketzahlen im Mai festzustellen, obwohl P3 in diesem Zeitraum beginnt, auf TCP-Anfragen zu antworten. Im Vergleich dazu werden in P4 weniger Pakete empfangen.

Es wurden 258 Quelladressen identifiziert, die Pakete an P3 senden, wobei die Mehrzahl dieser Adressen Hoster sind. Darüber hinaus gehören die Quelladressen zu Networkserviceprovidern, Internetserviceprovidern, Business Networks, Bildungs- und Forschungseinrichtungen. Die Quelladresse mit der höchsten Paketzahl sendet über 3 Tausend Pakete, während jede der übrigen Quelladressen insgesamt weniger als 60 Pakete versendet hat. Es werden zwei Pakete am 15. September 2023 in P3 entdeckt die beide an eine Zieladresse gesendet werden die mit *::0* endet. Gesendet wurden die Pakete von der Quelladresse der AS-Organisation *tencent-net-ap-cn tencent building* kurz *Tencent*.

Im Gegensatz dazu wurden nur acht Quelladressen festgestellt, die Pakete an P4 senden, wobei die maximale Anzahl der empfangenen Pakete pro Quelladresse 32 beträgt. Auch hier handelt es sich hauptsächlich um Hoster, ergänzt durch Pakete von Networkserviceprovidern, Bildungs- und Forschungseinrichtungen.

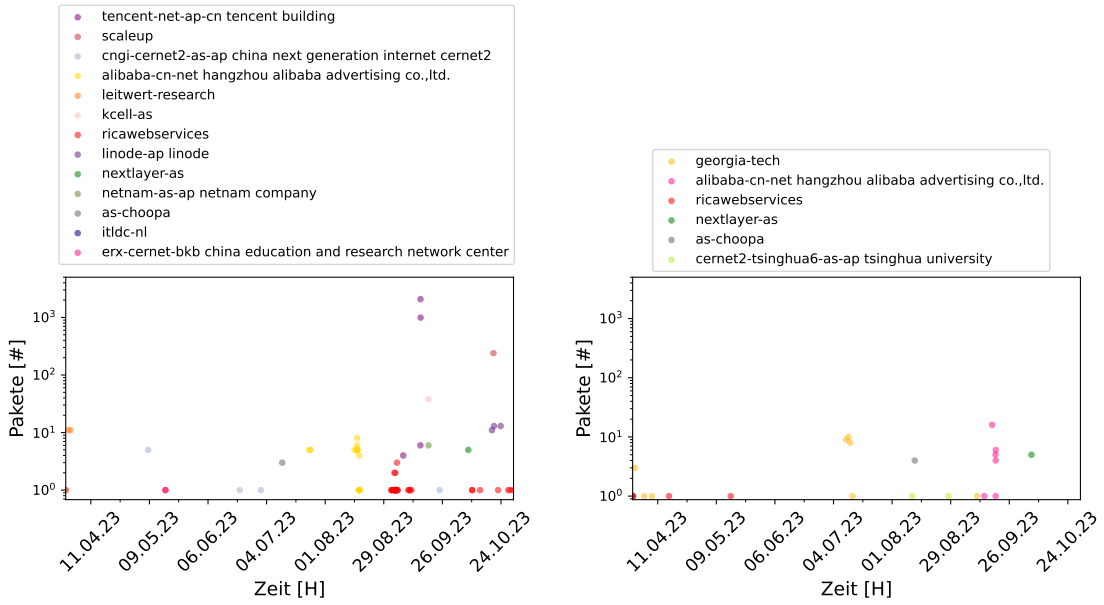
In Abbildung 2 werden die verwendeten Protokolle und in Abbildung 3 die AS-Organisationen visualisiert. Aus diesen Ergebnissen lässt sich schließen, welche Protokolle wie häufig verwendet wurden und welche AS-Organisationen mehr oder weniger Pakete in das jeweilige Präfix versendet haben. Die Protokolle werden pro Tag zusammengefasst, da diese sich sonst sehr stark überlappen. Die AS-Organisationen werden pro Stunde visualisiert, da diese häufig in kürzeren Abständen mehrmals scannen.



(a) P3 (b) P4

Abbildung 2: Protokollverwendung pro Tag

Aus den Abbildungen 2a und 2b geht hervor, dass in beiden Subnetzen ein Großteil der empfangenen Pakete ICMPv6-Pakete sind. In P3 sind es mit einer prozentualen Häufigkeit von über 96 % fast ausschließlich ICMPv6-Pakete. Das sind insgesamt 3.409 Pakete. Es werden dort außerdem 64 TCP-Pakete und 61 UDP-Pakete empfangen. Es gab insgesamt sechs SYN-Packete, die nach dem 16. Mai 2023 empfangen wurden. Auf diese wurde mit entsprechendem SYN-ACK reagiert und es folgten sechs ACK- und sechs FIN-ACK-Pakete. All diese Pakete wurden innerhalb von 30 Minuten empfangen. Fünf der insgesamt sechs Quelladressen der Pakete stammen aus einem AS. Daher könnte es sich hierbei um eine Quelle handeln, die mit mehreren Quelladressen die Pakete versendet hat. Dazu werden noch 23 PUSH-ACK-Pakete beobachtet, die ebenfalls von diesen fünf Quellen stammen und zur selben Zeit versendet wurden. Die TCP-Pakete sind in Abbildung 2a am 19. September erkennbar. Danach werden keine weiteren TCP-Pakete mehr empfangen. In P4 werden 54 ICMPv6-Pakete, 28 TCP-Pakete und ein UDP-Paket empfangen. Bei den TCP-Paketen handelt es sich ausschließlich um SYN-Pakete. Die ICMPv6-Pakete, die in P3 und P4 empfangen werden sind ausschließlich *Echo Requests*. Die empfangenen Pakete in P3 und P4 weisen keine Eigenschaften von Backscatter auf.



(a) P3 (b) P4

Abbildung 3: Paketankunft der AS-Organisationen pro Stunde

In Abbildung 3a ist zu erkennen, dass der hohe Paketanstieg am 15. September 2023 der AS-Organisation *Tencent* zuzuordnen ist. Von der AS-Organisation *Scale Up* wird ebenfalls eine relativ hohe Menge an Paketen festgestellt. Hierbei handelt es sich um zwei Hoster. *Tencent* hat mit 3.080 Paketen die meisten Pakete in P3 versendet. Von *Scale Up* werden 240 Pakete gezählt und die verbleibenden AS-Organisationen haben zusammen 214 Pakete an P3 versendet. In P4 hat das *Georgia Institute of Technology* (Bildungs- und Forschungseinrichtung) mit 34 Paketen die meisten Pakete in das Präfix versendet, während von der AS-Organisation *Hangzhou Alibaba Advertising* (Business) 33 Pakete gezählt werden. Vom Rest der AS-Organisationen werden zusammen 16 Pakete empfangen.

6.2 Analyse des Netzwerkverkehrs

Es fällt auf, dass beide Subnetze nicht wirklich viele Pakete empfangen. Die Veröffentlichung des /29-Covering-Präfixes macht es den Scannern nicht einfacher, die Subnetze zu finden. Genauso zeigt sich keine signifikante Veränderung bei der Menge der empfangenen Paketen, sobald P3 auf TCP-SYN-Pakete reagiert. Während dieses Zeitraums werden lediglich sechs TCP-SYNs empfangen, auf die P3 reagiert. Die insgesamt geringe Paketmenge erschwert eine aussagekräftige Analyse über das IPv6-Scan-Verhalten, um unter anderem Muster in der Zieladressgenerierung der Scanner zu erkennen und eine Veränderung des Scan-Verhaltens über einen längeren Zeitraum zu untersuchen. Da nur wenige Pakete über den kompletten Analysezeitraum empfangen wurden, ist eine Untersuchung von Scannern statistisch nicht aussagekräftig und ein Vergleich von unterschiedlichen Scannern ist aus diesem Grund nicht repräsentativ. Daher wird sich auf umfangreichere Analysen bei den letzteren beiden Präfixen konzentriert.

7 Das /48-Präfix mit aktivem Subnetz

Am 30. Juni 2023 startet die Messung des annoncierten /48-Präfix (P2). Es befindet sich ebenso vor Messbeginn auf der TUM-Hitliste. Außerdem wird festgestellt, dass das Präfix in kürzerer Zeit viel mehr Pakete empfängt als P3 und P4. Das Präfix besitzt darüber hinaus ein aktives /56-Subnetz, das auch bereits vor Messbeginn aktiv war. Vier Monate nach dem Start der Messung werden über 4 Millionen Pakete im Präfix aufgezeichnet.

7.1 Überblick des Netzwerkverkehrs

In der Abbildung 4 ist die Paketankunft pro Stunde im Messzeitraum vom 30. Juni bis zum 30. Oktober 2023 von P2 zu sehen. Im Diagramm ist im gesamten Messzeitraum von Juli bis Ende Oktober 2023 häufig ein Anstieg von Paketankünften pro Stunde im Bereich von 100 bis 1.000 festzustellen. Es sind außerdem in sehr konstanten Abständen hohe Paketraten pro Stunde zu erkennen, die ungefähr 100 Tausend Pakete pro Stunde beinhalten. Im gesamten Messzeitraum werden über 4 Millionen Pakete von über 8 Tausend Quelladressen empfangen. Etwas mehr als 2 Tausend TCP-Pakete haben nur das ACK-Flag gesetzt. Sie stammen von 11 Quelladressen. Dabei scheint es sich wahrscheinlich um ACK-Scans zu handeln. Mit einem ACK-Scan kann festgestellt werden, ob Firewalls zustandsorientiert sind oder nicht und welche Ports gefiltert werden[2]. Bei den ICMPv6-Paketen handelt es sich ausschließlich um *Echo Requests*. Es gibt einen kleinen Anteil der Pakete, der auf Backscatter schließen lässt. Dazu gehören fünf ICMPv6-Pakete mit dem Typ *Destination unreachable* und sieben RST-Pakete. Es werden außerdem acht UDP-Pakete mit einem reservierten Quellport (<1024) gesehen, die von drei Quelladressen stammen. Es werden unter anderem Internetserviceprovider, Networkserviceprovider, Hoster, Bildungs- und Forschungseinrichtungen beobachtet.
In P2 werden an 119 der 123 Tage des Messzeitraums insgesamt 16.072 Pakete entdeckt, die an genau 1.500 unterschiedliche Zieladressen gesendet wurden die alle mit *::0* enden. Gesendet wurden die Pakete von 571 unterschiedlichen Quelladressen.

[2] https://nmap.org/book/scan-methods-ack-scan.html

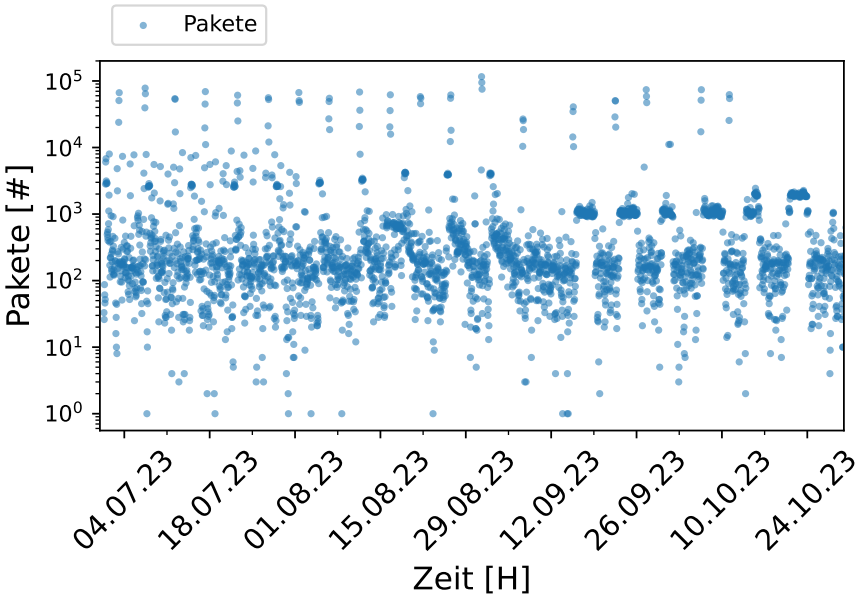


Abbildung 4: Paketankunft pro Stunde

Die Paketankünfte der vier AS-Organisationen, von denen die meisten Pakete im Messzeitraum empfangen wurden, sind in Abbildung 5 dargestellt. Die Pakete der vier AS-Organisationen decken 75 % aller Pakete ab. Das entspricht etwa 3,2 Millionen Pakete. In der Abbildung 5 sind die Anteile der Paketmengen in Prozent dargestellt, basierend auf den Paketen der vier AS-Organisationen. Die Quelladresse mit den meisten empfangenen Paketen macht 65 % aller Pakete aus und 74 %, wenn alleine die Top vier AS-Organisationen berücksichtigt werden. Laut pyasn gehört diese Quelladresse zur AS-Organisation *Tencent*. Auf dem Diagramm fällt die AS-Organisation *Hurricane* ebenso deutlich auf. Ab dem Beginn der Messung ist die Paketankunft pro Stunde, die von *Hurricane* beobachtet wird, knapp über zweieinhalb Monate lang sehr konstant. Daraufhin folgen ab dem 17. September Pausen in gleichmäßigen Abständen. Bei den Paketen der AS-Organisation *Ovh* sind ebenfalls Scans in regelmäßigen Abständen wie bei *Tencent* zu erkennen. Die Paketmenge pro Stunde, die von der AS-Organisation *Cdn77* empfangen wird, schwankt im ersten Monat stark und sinkt danach sehr schnell. Die Organisation *Cdn77* scannt insgesamt deutlich weniger als die Quelladressen der anderen drei AS-Organisationen. Es bleiben noch über 525 Tausend Pakete, die von den anderen nicht genannten 624 AS-Organisationen empfangen wurden.

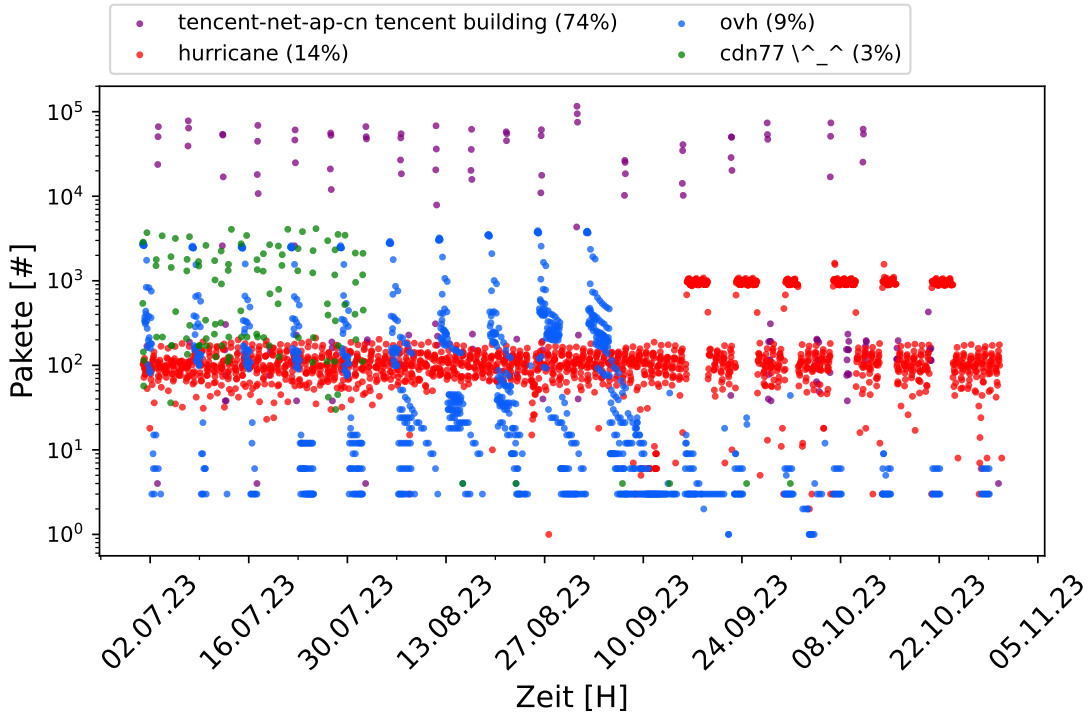


Abbildung 5: Paketankünfte der Top 4 AS-Organisationen

7.2 Protokolle

In Abbildung 6 wird eine Übersicht der verwendeten Protokolle gegeben. Da es eine so hohe Menge an Paketankünften gibt, werden die Pakete im Diagramm pro Tag dargestellt. Es wird unter anderem die Häufigkeit der Verwendung und der Zeitraum der verwendeten Protokolle sichtbar.
Bei einem Vergleich mit Abbildung 5 wird ersichtlich, dass der aktivste Scanner (*Tencent*) ICMPv6-Pakete versendet. Kleine Scans werden in diesem Präfix vor allem mit UDP durchgeführt. In der Mitte der logarithmischen Achse ist sowohl ICMPv6 als auch TCP zu sehen.

Tabelle 2: Pakete und Quelladressen pro Netzwerkprotokoll

Protokoll	Pakete		Quelladressen	
	Count	%	Count	%
ICMPv6	3 573 408	82	5216	64
TCP	756 245	17	6383	78
UDP	19 293	1	1957	24

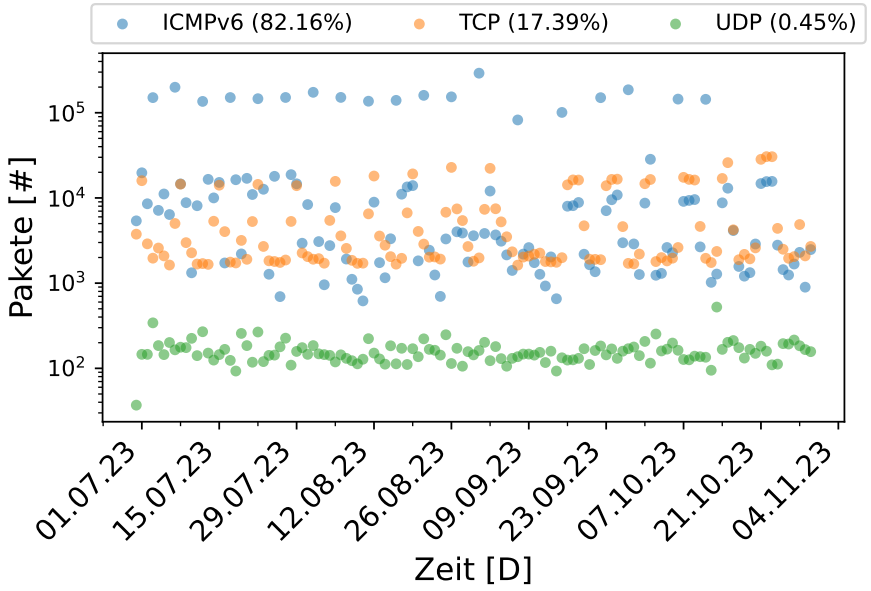


Abbildung 6: Protokollverwendung pro Tag

Es erfolgt eine zusätzliche Analyse, bei der berechnet wird wie viele Quelladressen welches Netzwerkprotokoll nutzen. Dabei wird pro Quelladresse jedes verwendete Protokoll jeweils nur einmal gezählt. Die Menge der Pakete pro Protokoll und die Menge der Quelladressen pro Protokoll werden in Tabelle 2 jeweils dargestellt. Man sollte berücksichtigen, dass einige Quelladressen während des Messzeitraums nicht ausschließlich ein Protokoll verwendet haben, sondern möglicherweise mehrere. 67 % der Quelladressen haben mehr als ein Protokoll verwendet. Aus diesem Grund ergibt die Summe der prozentualen Häufigkeiten bei den Quelladressen in Tabelle 2 nicht 100 %, sondern mehr. Aus der Tabelle wird ersichtlich, dass das TCP-Protokoll pro Quelladresse häufiger verwendet wird als die anderen Protokolle. Insgesamt gibt es mehr ICMPv6-, als TCP- und UDP-Pakete.

UDP wird sowohl pro Quelladresse, als auch insgesamt im Vergleich zu den anderen Protokollen seltener verwendet.

7.3 Ports

In der Tabelle 3 werden die fünf UDP-Ziel-Ports dargestellt, die am häufigsten verwendet wurden. In Tabelle 4 werden dagegen die fünf häufigsten TCP-Ziel-Ports dargestellt. Dabei muss beachtet werden, dass insgesamt 244 verschiedene UDP-Ziel-Ports und 1.196 verschiedene TCP-Ziel-Ports beobachtet werden. Die Tabellen geben nur einen Teil dieser Ergebnisse wieder.
Bei etwa 34 % der UDP-Pakete liegen die UDP-Ziel-Ports zwischen den Ports 33434 bis 33523, die auf Traceroute hinweisen[3]. Traceroute ermöglicht die Verfolgung des Datenpakets von einem Ursprungsort zu einem spezifischen Ziel im Internet, wobei die Hops und die benötigte Zeit für jeden Hop angezeigt werden [12]. Traceroute wird häufig für Netzwerkanalysen eingesetzt.

Tabelle 3: Die Top 5 der UDP-Ziel-Ports

Port	Dienste	Anzahl
33434 bis 33523	Traceroute	6530
53	DNS	2871
161	SNMP	2246
500	ISAKMP	2234
123	NTP	2221

Tabelle 4: Die Top 5 der TCP-Ziel-Ports

Port	Dienste	Anzahl
80	HTTP	316 530
443	HTTPS	267 044
8080	HTTP	4436
8090	HTTP	4419
8001	SHOUTcast	4418

DNS wird bei über 2.800 Paketen und damit neben Traceroute am häufigsten verwendet. Danach folgen SNMP, das ist ein Netzwerkverwaltungsprotokoll, ISAKMP, das Internet-Security-Association- und Schlüsselverwaltungsprotokoll und schließlich NTP. Die mit Abstand am öftesten verwendeten TCP-Ziel-Ports sind HTTP (80), das bei über 316 Tausend Paketen als Ziel-Port verwendet wird, und HTTPS (443), das bei über 267 Tausend Paketen als Ziel-Port verwendet wird.

7.4 Zieladressen-Generierung

Ein Ziel ist es, Muster oder Strukturen in der Generierung von Zieladressen zu identifizieren und zu untersuchen, ob diese Muster zur Erkennung von Scannern verwendet

[3] https://elinux.org/Traceroute_-_Tracing_Route

werden können. Dazu werden detaillierte Analysen der Zieladressgenerierung durchgeführt. Die Ergebnisse mehrere Scanner werden im Anschluss verglichen, um Änderungen und Eigenschaften festzustellen.

7.4.1 Zieladressen-Permutationen

In diesem Abschnitt wird eine Untersuchung der Permutationsmenge innerhalb der Zieladressen eines Scanners visualisiert. Dabei werden die Zieladressen in Zwei-Byte-Segmenten getrennt betrachtet. Die Menge der Permutation der ersten drei Segmente liegt immer bei eins, da sich dort das feste Präfix befindet (48 Bits). Die y-Achse verwendet eine logarithmische Skala und reicht bis zu einem Maximum von 65.536 möglichen Permutationen pro Segment. 65.536 ist die Anzahl der maximal möglichen Nibble-Abfolgen, die in einem Segment stehen können. Ein Byte sind somit zwei Nibble und ein Segment sind in diesem Fall vier Nibble. Eine Nibble-Abfolge eines Segmentes ist beispielsweise *1234* oder *ffff*. Die Abbildungen 7 und 8 visualisieren die Anzahl der Permutationen pro Segment für die Zieladressen der drei aktivsten Scanner. Unter den aktivsten Scannern sind Quelladressen gemeint, die im Messzeitraum Pakete an die größte Anzahl unterschiedlicher Zieladressen gesendet haben. Diese drei Quelladressen werden von *pyasn* den AS-Organisationen *Tencent*, der *Tsinghua University* und der *Chaos Computer Club Veranstaltungsgesellschaft mbH* zugeordnet. Abbildung 7 zeigt das Ergebnis zwei Monate, nachdem P2 annonciert wurde. Im Gegensatz dazu visualisiert Abbildung 8 die Permutationen der Segmente innerhalb der Zieladressen nach einem Zeitraum von insgesamt vier Monaten, also am Ende des Messzeitraums. Es ist wichtig zu beachten, dass insgesamt 2^{80} mögliche Zieladressen in P2 (/48-Präfix) gescannt werden können. Doch in dieser Analyse wird jedes Segment (65.536 Möglichkeiten) einzeln betrachtet. Dennoch lassen sich Aussagen über das Scan-Verhalten treffen.

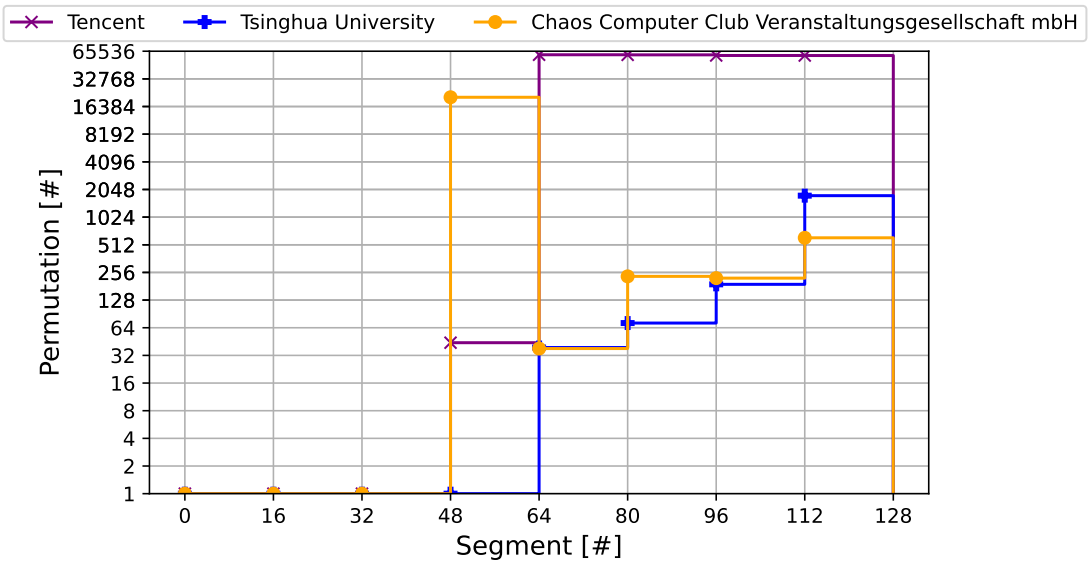


Abbildung 7: Segmentanalyse der Top Scanner nach zwei Monaten

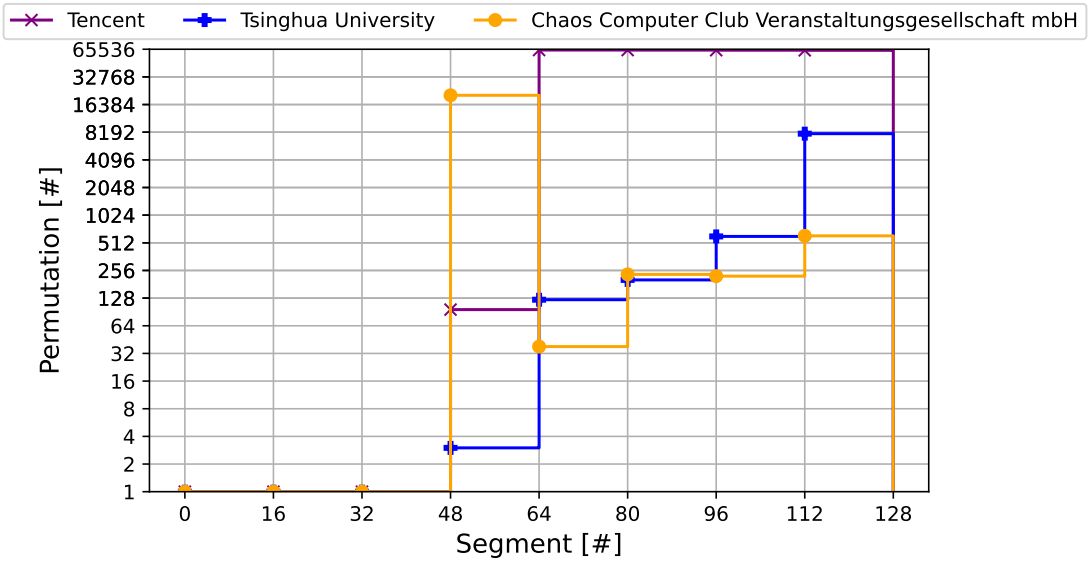


Abbildung 8: Segmentanalyse der Top Scanner nach vier Monaten

Beim Vergleich beider Abbildungen fällt auf, dass der *Tencent-Scanner* nur geringe Veränderungen in den Permutationen zeigt, obwohl es auch noch Scans zwischen dem zweiten und vierten Monat des Messzeitraums gab. Es gab aber nur eine geringe Erhöhung

der Permutationsmenge pro Segment. Auffällig ist ebenfalls, dass die Segmente im IID-Bereich des *Tencent-Scanners* eine identische Permutationsmenge aufweisen. Neben den letzten Segmenten gibt es im vierten Segment der Zieladressen nur wenig Permutationen. Daraus lässt sich ableiten, dass sich der Scanner eher auf weniger Subnetze konzentriert beim Scannen. Insgesamt versendet die Quelladresse Pakete an über 1 Millionen Zieladressen.

Bei dem Scanner der AS-Organisation *Tsinghua University* gibt es in den ersten zwei Monaten keine Veränderungen bei den Zieladressen zwischen dem 48. und 64. Bit. Zwei Monate danach steigt die Menge der Permutationen im vierten Segment auf drei und innerhalb der IIDs werden ebenso neue Permutationen festgestellt. Anders als bei *Tencent* gibt es keine identische Menge an Permutationen in den Segmenten der IID. Vom 64. bis zu 128. Bit steigt die Menge der Permutationen pro Segment stufenartig an. Dennoch konzentriert sich der Scanner ebenso auf bestimmte Subnetze beim Scannen. Nach vier Monaten versendete die Quelladresse Pakete an über 33 Tausend Zieladressen.

Der Scanner der AS-Organisation *Chaos Computer Club* scannt im Gegensatz zu den anderen beiden mehrere /64-Subnetze. Es sind daher mehr Permutationen im vierten Segment zu beobachten. In den letzteren Segmenten gibt es im Vergleich dazu nicht viele Permutationen. Dies deutet darauf hin, dass der Scanner sich im Vergleich zu den anderen beiden mehr auf das Scannen der Subnetze konzentriert. Es werden Pakete an über 24 Tausend Zieladressen versendet. Ebenso gibt es über 16 Tausend Permutationen zwischen dem 48. und 64. Bit. Von der Quelladresse werden außerdem nur in den ersten zwei Monaten innerhalb von vier Tagen Pakete empfangen.

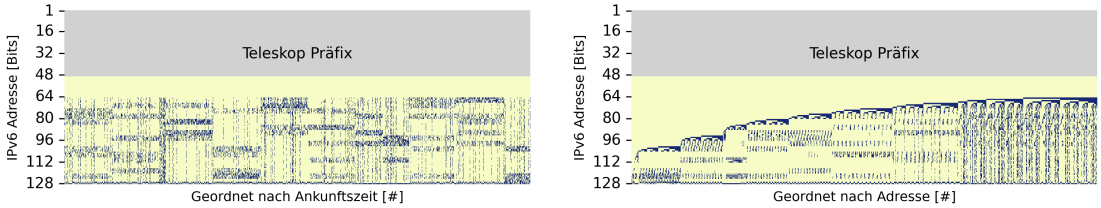
7.4.2 Analyse des aktivsten Scanners

In der folgenden Analyse werden Zieladressen untersucht, die innerhalb eines bestimmten Tages von einem Scanner der AS-Organisation *Tencent* gescannt wurden. *Tencent* ist eines der größten und meistgenutzten Internet-Service-Portale in China und betreibt führende Internet-Plattformen. *Tencent* wird dem Netzwerktyp *Hosting* zugeordnet.

Die Abbildungen 9 und 10 zeigen die Zieladressen, die am 07. Juli 2023 von der Quelladresse von *Tencent* gescannt wurden. Die einzelnen Zieladressen werden in Abbildung 9 und 10 vertikal angeordnet. Abbildung 9 zeigt dabei die Zieladressen in binärer Darstel-

lung und 10 in hexadezimaler Darstellung. Es werden über 181 Tausend Pakete innerhalb dieses Tages ermittelt. Da das zu überwachende Präfix ein /48-Präfix ist, befindet sich ein grauer Balken über dem /48-Präfix-Bereich. Dieser ist bei jeder Zieladresse identisch und es wird somit der Fokus auf den Teil der Adressen gelegt, der analysiert wird.
Bei Abbildung 9a und 10a werden die Zieladressen entlang der x-Achse von links nach rechts basierend auf ihrer Ankunftszeit im Präfix angeordnet. In der Abbildung 9b und 10b wird nach IPv6-Adresse sortiert und dementsprechend von links nach rechts angeordnet. Beide Heatmaps in Abbildung 10 werden durch Farbskalen ergänzt, die Aufschluss über die Zuordnung der Farben für das jeweilige Hexadezimalzeichen geben. In Abbildung 9 wird jede binäre *1* blau markiert und jede binäre *0* gelb markiert.

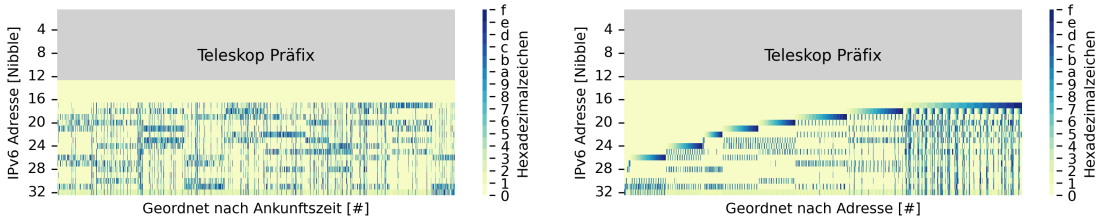
Die binäre Repräsentation der Zieladressen wird verwendet, um Muster auf der niedrigsten Repräsentationsebene zu erkennen. Auf diese Weise können Veränderungen in der Generierung von Zieladressen auf der binären Ebene sichtbar gemacht werden. In der hexadezimalen Ansicht lassen sich die einzelnen Nibble erkennen. Ein Vorteil der hexadezimalen Darstellung ist beispielsweise, dass in Abbildung 10b ersichtlich wird, wie der Scanner das Präfix durchläuft. Es ist ein stufenartiger Anstieg zu erkennen, der darauf hindeutet, dass alle Hexadezimalzeichen vom 26. Nibble bis zum 17. Nibble (bis auf das 23. und das 25. Nibble) mindestens einmal an diesen Positionen in den Zieladressen verwendet wurden.



(a) Scan nach Ankunftszeit sortiert

(b) Scan nach IPv6-Adresse sortiert

Abbildung 9: Tencent-Scan-Tag am 07.07.2023 in binärer Darstellung



(a) Scan nach Ankunftszeit sortiert (b) Scan nach IPv6-Adresse sortiert

Abbildung 10: Tencent-Scan-Tag am 07.07.2023 in hexadezimaler Darstellung

In den Darstellungen 9a und 10a sind Bit-Wechsel erkennbar. Diese Bit-Wechsel sorgen für blockartige Muster, während der Rest der Zieladressabschnitte sich nur geringfügig ändert. Die hexadezimale Darstellung bekräftigt die Tatsache noch, indem die Zeichenwechsel noch einmal mit den dort farblich sichtbaren Hexadezimalzeichen dargestellt sind.
Beim letzten Nibble in den Zieladressen ist in den Abbildungen ebenfalls ein Bit-Wechsel feststellbar, der sich jedoch durch den ganzen Tag zieht. Das ist in der binären Darstellung ersichtlich, indem zwischen den Werten *0b01* und *0b10* gewechselt wird. In der hexadezimalen Darstellung werden diese mit den Farben für *1* und *2* erkennbar.

Zusammengefasst kann den Abbildungen 9 und 10 entnommen werden, dass es sich hierbei sehr wahrscheinlich nicht um eine zufällige Generierung handelt, da eine gewisse Struktur festgestellt wird. Es lässt sich außerdem sagen, dass die Zieladressen so generiert wurden, dass einzelne Zeichen an bestimmten Nibble mehrmals durch getauscht werden und sich der Rest der Zieladressen nur geringfügig verändert. Außerdem werden am letzten Nibble größtenteils die Hexadezimalzeichen eins und zwei verwendet. Ein ähnliches Ergebnis wird vom selben Scanner ebenfalls an anderen Tagen innerhalb des Präfixes und auch in P1 beobachtet.

Zusätzlich wird noch das Tool addr6 aus dem *IPv6Toolkit* [17] verwendet. Mit Hilfe des Tools kann die IID in Adresstypen klassifiziert werden. Die Adresstypen sind nach den Spezifikationen von RFC 7707 [8] definiert. Die Einteilung ist in Abbildung 11 dargestellt.

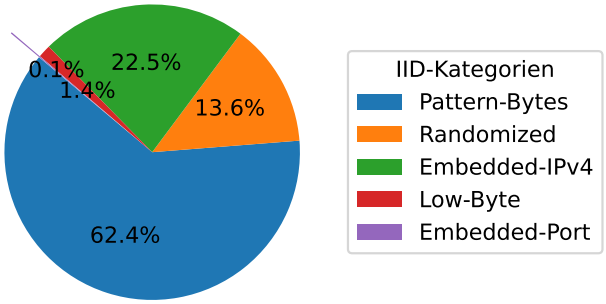


Abbildung 11: Klassifizierung der Interface-IDs der Zieladressen

Im Kreisdiagramm dominiert die Kategorie *Pattern-Bytes* mit 62,4 %. Eine Zieladresse wird dieser Kategorie zugeordnet, wenn kein anderer Adresstyp gefunden werden konnte und sich mindestens drei aufeinanderfolgende Null-Bytes in ihrer IPv6-IID befinden. 22,5 % der Zieladressen werden der Kategorie *Embedded-IPv4* zugeordnet. Sie repräsentiert IPv6-Adressen, die eine eingebettete IPv4-Adresse in der IID enthalten. Diese Adressen bestehen aus einem variablen Präfix, der eingebetteten IPv4-Adresse und einem variablen Suffix gemäß RFC 6052 [1]. Zusätzlich dazu werden 1.4 % der Zieladressen zu der Kategorie *Low-Byte* zugeordnet. Hierbei handelt es sich um Adressen, bei denen alle Bytes der IID außer dem niederwertigsten Byte auf *0* gesetzt sind. Es gibt auch andere Muster wie Adressen, bei denen die IID bis auf die letzten beiden 16-Bit-Wörter auf *0* gesetzt sind. Solche Adressmuster gehören ebenfalls zu der Kategorie *Low-Byte*. Schließlich ist die *Embedded-Port-Kategorie* am geringsten vertreten und umfasst Adressen, bei denen die letzten Bytes der IID den Service-Port einbetten. Einige Zieladressen fallen auch in die *Randomized-Kategorie*. Sie dient eher als Ausschlusskriterium für andere Kategorien, da sie keine Zufälligkeit testet. Stattdessen fallen in diese Kategorie alle IIDs rein, die den anderen Kategorien nicht zugeordnet werden können.
Es lässt sich feststellen, dass *Tencent* mit diesem Scan unterschiedliche Zieladresskategorien abdeckt. Durch dieses Ergebnis wird die Vermutung einer nicht zufälligen Generierung bekräftigt. Das Ergebnis lässt darauf schließen, dass mindestens ein Teil der Zieladressen bewusst gescannt wurde.

7.5 Erkenntnisse aus den Analysen

P2 empfing bis zum 30. Oktober im Vergleich zu P3 und P4 sehr viele Netzwerkpakete. Die meisten Pakete werden von der AS-Organisation *Tencent* empfangen. Die Analyse der Protokolle zeigt, dass die meisten Pakete, die empfangen wurden, ICMPv6-Pakete sind. Die Visualisierungen der Zieladressen vom aktivsten Scanner ermöglichen Einblicke in Muster und Strukturen. Dabei werden unter anderem Bit-Wechsel erkannt. Die IIDs der Zieladressen dieses Scanners können verschiedenen Kategorien wie *Pattern-Bytes*, *Low-Byte* und *Embedded-IPv4* zugeordnet werden, wodurch sich eine bestimmte Vorgehensweise in der Zieladressgenerierung des Scanners ableiten lässt und eine zufällige Generierung somit eher ausgeschlossen wird. Mit dieser Vorgehensweise werden nur wenig Subnetze abgedeckt, die durch die Bit-Wechsel an unterschiedlichen Positionen in der IID, teilweise gescannt werden.

8 Das neu annoncierte /32-Präfix

P1 wird am 24. August 2023 annonciert und es wird ebenfalls mit der Messung begonnen. Somit ist eine Analyse der ersten Pakete in P1 ebenfalls möglich. Am 29. August wird P1 in der TUM-Hitliste veröffentlicht. Wie bei den anderen Präfixen endet der Messzeitraum dieser Arbeit am 30. Oktober. Innerhalb des Messzeitraums werden über 2 Millionen Pakete empfangen. Alle Adressen in diesem Präfix sind passiv.

8.1 Überblick des Netzwerkverkehrs

In Abbildung 12 wird die Paketmenge pro Stunde innerhalb des Messzeitraums abgebildet.

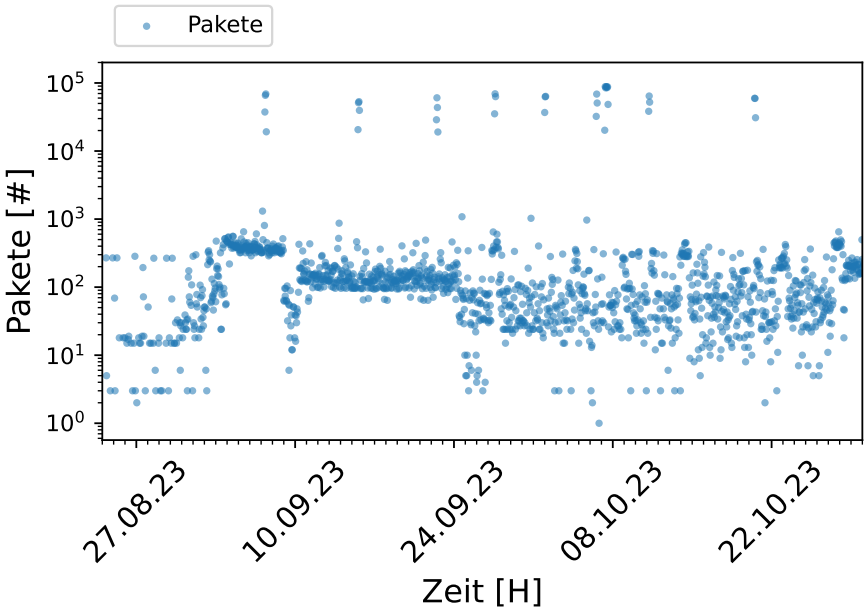


Abbildung 12: Paketankunft pro Stunde

Die Paketmengen werden auf der y-Achse logarithmisch dargestellt. Auffällig sind die deutlich hohen Paketraten, die in ähnlichen Abständen auftauchen. Neben diesen Peaks ist die Anzahl der Paketankünfte pro Stunde anfangs sehr unterschiedlich, bis sie schließlich immer weiter steigt. Ab dem 04. September ist die Menge der ankommenden Pakete sehr konstant auf einer Höhe von 100 bis 1.000 Pakete pro Stunde. Nach dem 24. September 2023 wird die Anzahl der Pakete insgesamt geringer und die Menge variiert stark von unter 10 bis 1.000 Paketen pro Stunde.

Im Vergleich zu P3 und P4 empfängt P1 eine deutlich höhere Anzahl an Paketen. P2 weist dagegen mehr als die doppelte Anzahl der von P1 empfangenen Pakete auf. Es ist jedoch zu beachten, dass der Messzeitraum von P2 um 55 Tage länger ist. Zudem wurde P2 knapp 13 Jahre früher annonciert und besitzt im Gegensatz zu P1 ein aktives Subnetz. Diese Faktoren könnten die Anzahl der empfangenen Pakete beeinflusst haben. In beiden Diagrammen, die die Paketankünfte für P2 und P1 darstellen, sind erhebliche Anstiege in der Anzahl der empfangenen Pakete zu beobachten, die ungefähr 100 Tausend Pakete pro Stunde erreichen. Abgesehen davon sind keine weiteren Ähnlichkeiten zwischen den beiden Diagrammen erkennbar.

Bei der Suche nach Backscatter werden drei ICMPv6-Pakete mit dem Typ *Destination unreachable*, ein ICMPv6-Paket mit dem Typ *Time Exceeded*, ein ACK-Paket und ein RST-Paket beobachtet. Insgesamt werden Pakete von über 1.200 Quelladressen empfangen. Es werden unter anderem Internetserviceprovider, Networkserviceprovider, Hoster, Bildungs- und Forschungseinrichtungen beobachtet.
In P1 wurde an jedem Tag der Messung, außer einem, mindestens ein Paket an Zieladressen gesendet, die mit *::0* enden. Eine Stichprobe eines Tages, an dem über 1.000 Pakete an solche Zieladressen gesendet wurden, zeigt, dass pro Quelladresse stets genau vier Pakete verschickt wurden. An diesem spezifischen Tag, dem 7. September 2023, haben insgesamt 283 Quelladressen solche Pakete gesendet. Dabei wurden drei verschiedene Zieladressen mit der Endung *::0* gescannt. Im gesamten Messzeitraum wurden insgesamt 8.357 Pakete an Zieladressen mit der Endung *::0* von insgesamt 452 Quelladressen gesendet. Insgesamt wurden 64 unterschiedliche Zieladressen mit der Endung *::0* gescannt.

In Abbildung 13 sind die drei AS-Organisationen dargestellt, von denen im Messzeitraum die meisten Pakete empfangen werden. Die höchste Anzahl empfangener Pakete stammt von *Tencent*. Zwei Quelladressen innerhalb dieser AS-Organisation führten regelmäßig Scan-Aktivitäten in ähnlichen zeitlichen Abständen durch. Im Gegensatz dazu wird von einer Quelladresse, die der AS-Organisation *Next-Layer* zugeordnet ist, nur an einem einzigen Tag eine signifikante Anzahl von Paketen empfangen. An keinem weiteren Tag wurde eine vergleichbare Menge an Paketen pro Quelladresse verzeichnet. Es werden drei Quelladressen gesehen, die der AS-Organisation *Rica Web Services* zugeordnet sind. Diese drei Quelladressen sind über einen größeren Zeitraum hinweg aktiv und decken somit mehr Tage ab, an denen Pakete in das Präfix gesendet werden, als die Quelladressen der beiden anderen Organisationen.

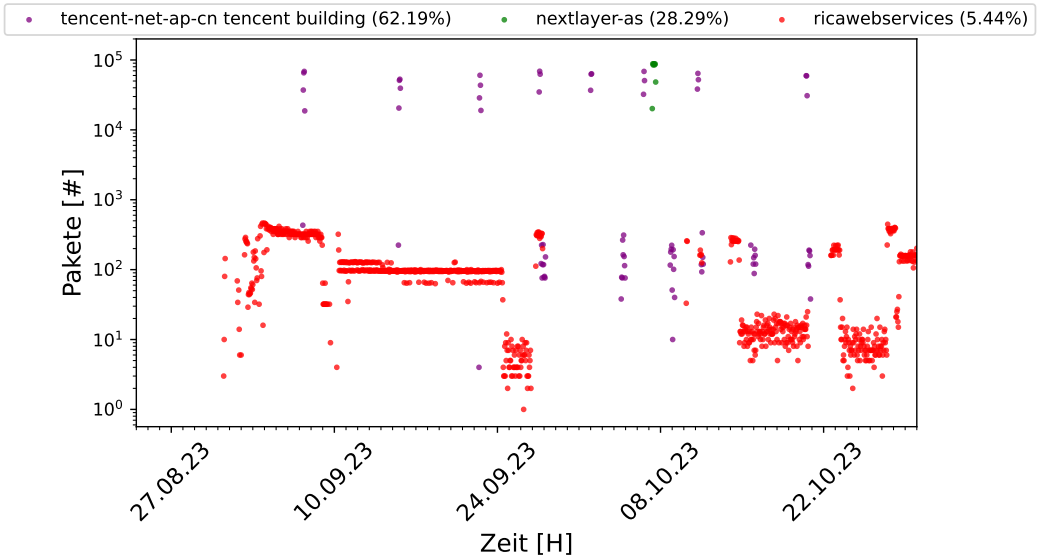


Abbildung 13: Paketankünfte der Top 3 AS-Organisationen

Tencent zeigt eine nahezu periodische Scan-Aktivität, da die Scans in sehr ähnlichen Zeitabständen auftauchen. Während *Next-Layer* zwar nur einmal, jedoch die größte Anzahl an Paketen pro Tag sendet. *Rica Web Services* weist anfänglich eine sehr konstante, wenn auch geringere Paketmenge pro Stunde auf als bei den anderen beiden AS-Organisationen. Von *Rica Web Services* werden gerade in der Anfangsphase täglich Pakete empfangen. Ab dem 25. September 2023 variiert die Menge als auch die Zeitintervalle zwischen den Ankünften von *Rica Web Services*. Die Analyse des Diagramms zeigt, dass es viele Ähnlichkeiten zwischen der Abbildung 13 und Abbildung 12 gibt. Ein Grund dafür ist, dass nur noch etwa 4 % der gesamten Paketmenge von anderen AS-Organisationen stammt.

8.2 Protokolle

In Abbildung 14 werden die Pakete pro Tag für jedes Protokoll dargestellt. Da es in P1 zu so vielen Paketankünften kommt, wird hier ebenfalls die Einheit pro Tag gewählt, um einen besseren Überblick darzustellen. Es sind signifikante Unterschiede in der Häufigkeit zu sehen. ICMPv6-Pakete wurden in bedeutend höherer Anzahl empfangen, während UDP-Pakete vergleichsweise selten erfasst wurden. TCP wurde etwas häufiger als UDP

Tabelle 5: Pakete und Quelladressen pro Netzwerkprotokoll

Protokoll	Pakete		Quelladressen	
	Count	%	Count	%
ICMPv6	2 064 522	98,94	1029	83
TCP	18 633	0,89	25	17
UDP	3586	0,17	205	2

empfangen, jedoch bleibt die Anzahl im Vergleich zu ICMPv6-Paketen immer noch vergleichsweise gering.

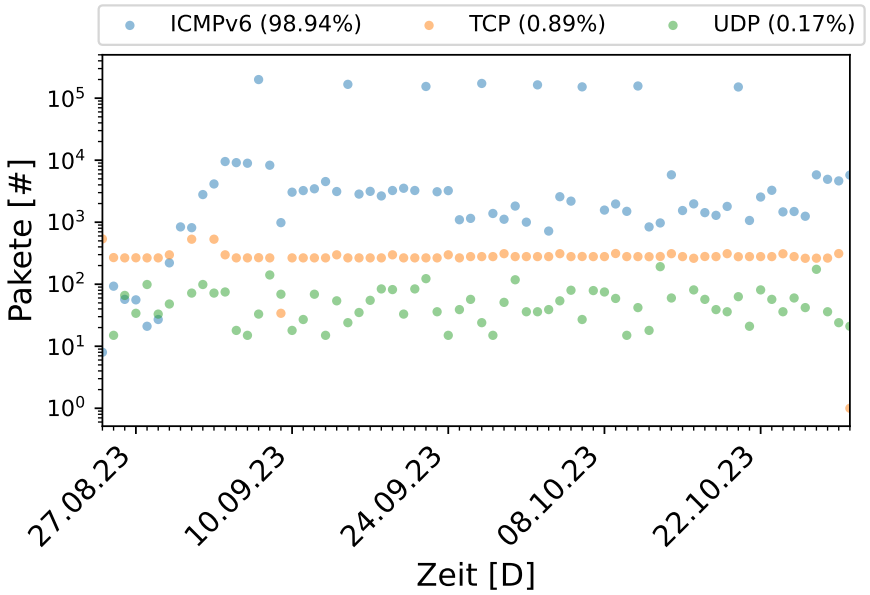


Abbildung 14: Protokollverwendung pro Tag

Die hohe Anzahl an ICMPv6-Paketen wird stark von den aktivsten Scannern beeinflusst, da diese überwiegend ICMPv6-Pakete versenden. Des Weiteren fällt auf, dass zu Beginn der Messung die Anzahl der ICMPv6-Pakete zunimmt und sich nach einer gewissen Zeit erst zu einer so großen Mehrheit entwickelt.

Der Anteil der Pakete pro Netzwerkprotokoll und der Anteil der Quelladressen werden in Tabelle 5 dargestellt. In der Tabelle ist zu sehen, dass ICMPv6 auch bei den meisten Quelladressen mindestens einmal verwendet wird. Von den über 1.200 Quelladressen haben 1.029 mindestens ein ICMPv6-Paket in das Präfix versendet. Es haben nur 205

verschiedene Quelladressen UDP-Pakete verschickt und nur 25 verschiedene Quelladressen haben TCP-Pakete in das Präfix versendet. 1,45 % der Quelladressen verwendeten im gesamten Messzeitraum mehr als nur ein Protokoll.

In Abbildung 14 ist eine nahezu konstante Anzahl an empfangenen TCP-Paketen pro Tag zu erkennen. Um herauszufinden, was dafür der Grund sein könnte, werden in Tabelle 6 die Top 10 AS-Organisationen aufgelistet, von denen am meisten TCP-Pakete in P1 empfangen wurden.

Tabelle 6: TCP-Verwendung

AS-Organisation	**AS-Nummer**	**Pakete**
leitwert-research	29108	14 220
reliablesite	23470	1072
asseflow	49367	1001
leaseweb-apac-sin-11 leaseweb asia pacific pte. ltd.	59253	603
mwn-as	12816	416
-reserved as-	206271	2
msf	2027	2
nassist-as	29632	2
reliancejio-in.	55836	2
fastweb	12874	1

Diese Analyse verdeutlicht, dass die TCP-Pakete nicht ausschließlich von einer einzigen Quelle stammen. Insbesondere ragen die drei führenden Einträge in der Tabelle heraus, von denen eine beträchtliche Menge an TCP-Paketen versendet wurde. Bei den Paketmengen dieser führenden AS-Organisationen wird festgestellt, dass diese in gleichbleibenden Abständen immer in etwa die gleiche Menge an Paketen versenden. Diese sorgen für die linienartige Darstellung der TCP-Paketankünfte in Abbildung 14.

8.3 Ports

Um einen Überblick der am häufigsten verwendeten UDP- und TCP-Ziel-Ports und Dienste zu erhalten, werden diese aus den empfangenen Paketen ermittelt und in der Tabelle 7 und 8 dargestellt.

Tabelle 7: Die Top 5 der UDP-Ziel-Ports

Port	Dienste	Anzahl
33434 bis 33523	Traceroute	3.586

Tabelle 8: Die Top 5 der TCP-Ziel-Ports

Port	Dienste	Anzahl
80	HTTP	18.624
443	HTTPS	6
4843	opcua-tls	1
24915	Unbekannt	1
10171	Unbekannt	1

Es haben über 18 Tausend TCP-Pakete HTTP als Ziel-Port-Angabe und sind damit unverschlüsselt. Port 443, der für verschlüsselte HTTPS-Verbindungen steht, wurde sechs Mal beobachtet. Zu den Ports 10171 und 24915 werden keine Dienste gefunden. Port 4843 (OPC UA Protocol over TLS/SSL) wird für den Informationsaustausch zwischen verschiedenen Geräten und Systemen in industriellen Umgebungen verwendet[4]. Daneben werden noch drei weitere Ports (2207, 43101, 34767) jeweils einmal beobachtet. Abgesehen vom Dienst hpssd (HP Status and Services) von Port 2207 sind die anderen beiden Dienste unbekannt.

Bei der Analyse der UDP-Ziel-Ports sind 70 unterschiedliche Ports bis zum 30. Oktober protokolliert worden. Sämtliche Zielports der UDP-Pakete liegen im Bereich von 33434 bis 33534 (Traceroute).

8.4 Eintrag in der Hitliste

P1 wurde fünf Tage nach seiner Ankündigung, am 29. August 2023, in der TUM-Hitliste verzeichnet. Innerhalb der ersten 15 Tage sind immer mehr Pakete empfangen worden. Diese sind in der Tabelle 9 dargestellt.

Die Tabelle zeigt einen deutlichen Anstieg der Paketanzahl nach dem 29. August. Es wird daher angenommen, dass ein Eintrag in der Hitliste Scanner dazu veranlasst, ein Präfix zu scannen, was zu einem erhöhten Netzwerkverkehr führt. Jedoch lässt sich dies nicht sicher bestätigen.

[4] https://www.iana.org/assignments/service-names-port-numbers/service-names-port-numbers.xhtml?search=ssl

Tabelle 9: Pakete pro Tag

Tag	Paketmenge
24.08.2023	546
25.08.2023	377
26.08.2023	389
27.08.2023	357
28.08.2023	386
29.08.2023	326
30.08.2023	567
31.08.2023	840
01.09.2023	1.420
02.09.2023	2.890
03.09.2023	4.742
04.09.2023	9.887
05.09.2023	9.415
06.09.2023	9.218
07.09.2023	199.863

8.5 Zieladressen-Generierung

In diesem Abschnitt wird eine Untersuchung der Zieladressen von den aktivsten Scannern durchgeführt, mit dem Ziel, spezifische Muster und Strukturen zu identifizieren. Diese Analyse zielt darauf ab, Einblicke in die Methoden und Charakteristika der Scanner zu gewinnen, insbesondere hinsichtlich der von ihnen generierten Zieladressen.

8.5.1 Zieladressen-Permutationen

Die Abbildungen 15 und 16 visualisieren die Anzahl der Permutationen pro Segment (zwei Bytes) für alle Zieladressen der drei aktivsten Scanner des /32-Präfixes. Diese drei Quelladressen können den AS-Organisationen *Tencent*, *Rica Web Services* und *Next-Layer* zugeordnet werden.

Abbildung 15 zeigt das Ergebnis einen Monat nach dem Messbeginn von P1, während Abbildung 16 zwei Monate nach dem Messbeginn zeigt. Die Quelladresse, der AS-Organisation *Tencent* versendet im Messzeitraum Pakete an über 561 Tausend Zieladressen. Gefolgt von einer Quelladresse der AS-Organisation *Rica Web Services*, die Pakete

an über 86 Tausend Zieladressen verschickte. Schließlich versendete die Quelladresse der AS-Organisation *Next-Layer* Pakete an über 65 Tausend Zieladressen.

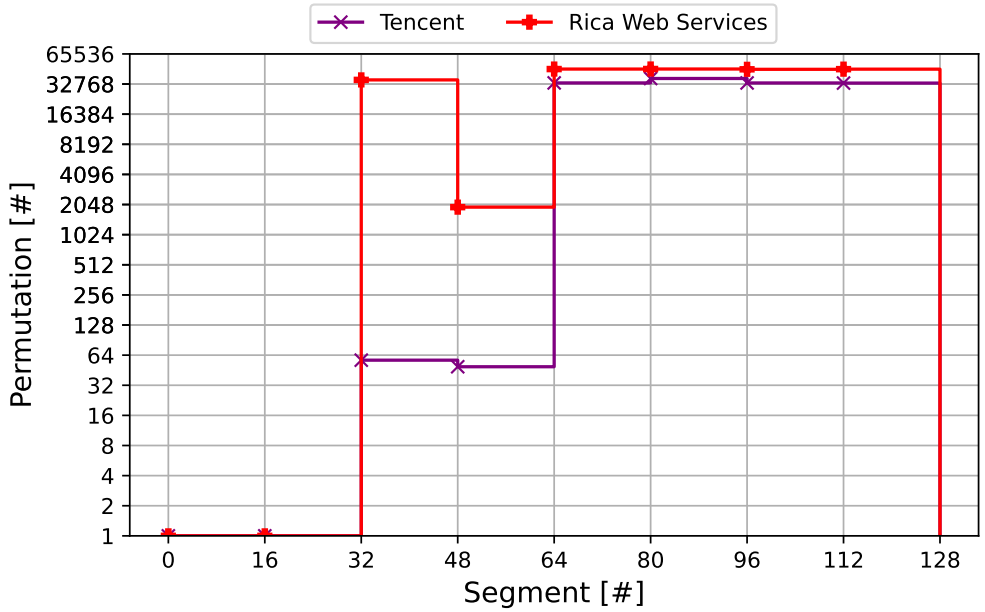


Abbildung 15: Segmentanalyse der Top Scanner nach einem Monat

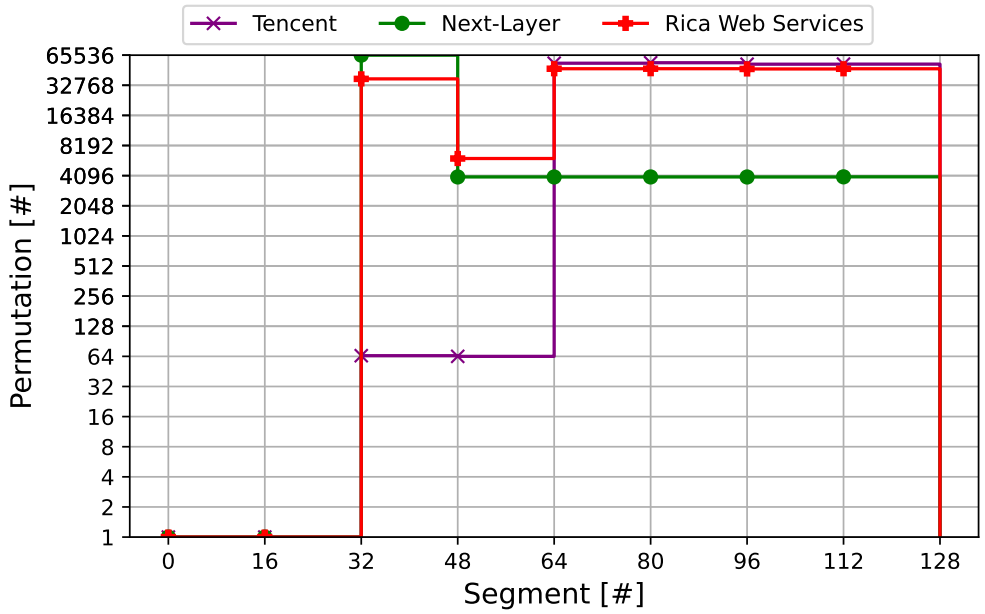


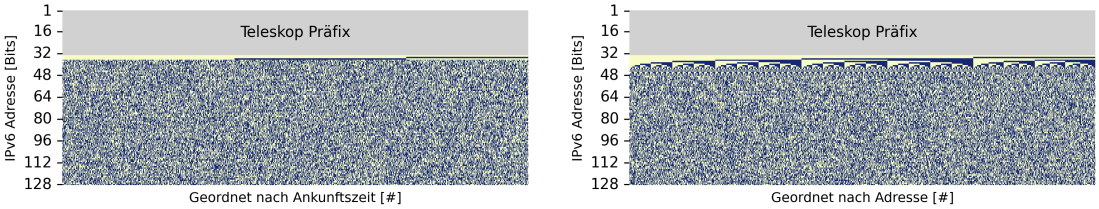
Abbildung 16: Segmentanalyse der Top Scanner nach zwei Monaten

Auffällig ist, dass es bei dem *Tencent-Scanner* in den Segmenten drei und vier relativ wenig Permutation gibt. Dies deutet darauf hin, dass dieser Scanner sich beim Auskundschaften auf bestimmte Subnetze fokussiert. Bei der Analyse fällt auf, dass der Scanner häufig die gleichen Zieladressen an unterschiedlichen Tagen scannt. An allen acht Tagen, an denen die Quelladresse Pakete sendet, werden insgesamt über eine Million Pakete empfangen. 47 Zieladressen erhalten an jedem dieser Tage mindestens ein Paket vom Scanner. Darunter auch die *Low-Byte-Adressen*, die ausschließlich Nullen in der IID besitzen und mit *::1* oder *::2* enden. Daraus lässt sich ableiten, das der Scanner sich eher auf bestimmte Zieladressen konzentriert, als das komplette Präfix zufällig durch zu scannen.
Bei *Rica Web Services* fällt dagegen auf, dass es mehr Permutation im dritten und vierten Segment gibt als bei *Tencent*. Darüber hinaus versendet die Quelladresse Pakete jedes Mal an eine andere Zieladresse. Daraus lässt sich schließen, dass der Scanner darauf abzielt das Präfix auf neue unbekannt aktive Hosts zu durchsuchen. Das Ziel scheint darin zu liegen das Präfix zu erkunden.
Next-Layer scannt nur am 07. Oktober 2023 und ist aus diesem Grund, nur in der Abbildung 16 zu sehen. An diesem Tag werden gleich alle Permutationen im dritten Segment abgedeckt, während die Anzahl der Permutationen in den hinteren Segmenten konstant, aber etwas geringer ausfällt. Dieser Scanner scheint sich auf das Scannen in die Breite und damit auf das Auskundschaften mehrerer Subnetze zu konzentrieren. Genauso wie der Scanner von *Rica Web Services* scheint der Scanner das Ziel zu verfolgen das Präfix weitestgehend zu erkunden.
Es zeigt sich, dass allgemein mehr Permutationen im dritten Segment (32. bis 48. Bit) auftreten als im vierten Segment (48. bis 64. Bit). Außerdem wird im vierten Segment oft das hexadezimale Zeichen *0* gesetzt. Dieses Muster wird bei weiteren Scannern beobachtet, jedoch konnte keine genaue Ursache dafür festgestellt werden. Es wird vermutet, dass die Scanner eine bestimmte Struktur beim Auskundschaften des Präfixes verfolgen. Indem sie die Zieladressen zwischen dem 32. und 48. Bit permutieren lassen und sich somit eher auf bestimmte Subnetze fokussieren.
Außerdem fällt auf, dass mehrere Scanner eine identische Permutationsmenge bei den Segmenten innerhalb der IID haben. Dies ist in Abbildung 16 ebenfalls bei allen Scannern zu sehen. Ein Grund ist dafür nicht bekannt.

Damit sich ein besseres Bild von den Zieladressen gemacht werden kann, werden in den folgenden Abschnitten die beiden Scanner von *Next-Layer* und *Rica Web Services* weiter untersucht. Der Scanner von *Tencent* zeigt ähnliche Ergebnisse wie aus Abschnitt 7.4.2.

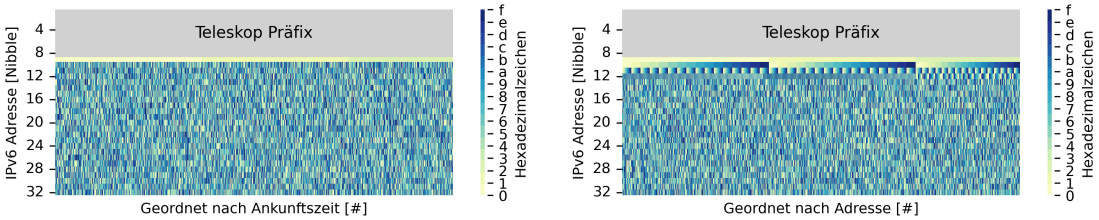
8.5.2 Analyse des größten Scan-Tags

Die folgenden Abbildungen 17 und 18 zeigen die ersten 100 Tausend Zieladressen, die von der Quelladresse der AS-Organisation *Next-Layer* empfangen wurden. Es handelt sich bei der Organisation um eine Telekommunikationsdienstleistungs- und Beratungs GmbH. Dieser Scanner sendet ausschließlich am 07. Oktober 2023 Pakete. An diesem Tag werden insgesamt 590 Tausend Pakete von dieser Quelladresse empfangen. Ein Einblick in die Zieladressen des Scan-Tags zeigt das Vorgehen des Scanners.



(a) Sortiert nach Ankunftszeit (b) Sortiert nach IPv6-Adresse

Abbildung 17: Die ersten 100.000 Pakete vom Scan-Tag am 07.10.2023 in binärer Darstellung



(a) Sortiert nach Ankunftszeit (b) Sortiert nach IPv6-Adresse

Abbildung 18: Die ersten 100.000 Pakete vom Scan-Tag am 07.10.2023 in hexadezimaler Darstellung

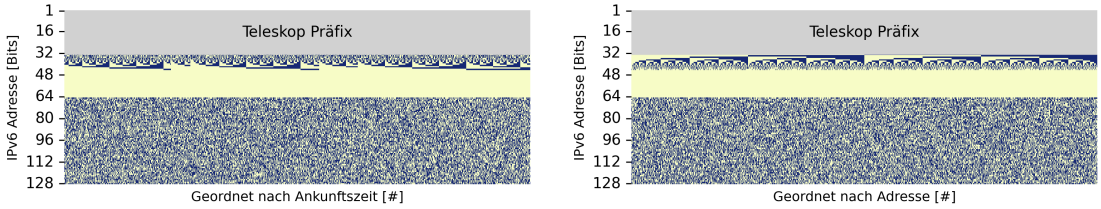
In Abbildung 17 lässt sich auf dem ersten Blick eine Struktur in den ersten Bits nach dem Präfix erkennen. Abbildung 18a zeigt, dass die Hexadezimalzeichen *0* und *1* im obersten Nibble (9. Nibble) der Zieladressen in sequentieller Reihenfolge gesetzt wurden. Bei einer genaueren Analyse zeigt sich zudem, dass sich diese sequentielle Reihenfolge aller Hexadezimalzeichen bis *f* bis zum letzten Paket fortsetzt. Abbildung 18b verdeutlicht, dass an der nachfolgenden Position (10. Nibble) alle Hexadezimalzeichen in sequentieller Reihenfolge pro Zeichen am 9. Nibble gesetzt wurden. Analog dazu wird pro Zeichen am

10. Nibble jedes Hexadezimalzeichen im 11. Nibble in sequentieller Reihenfolge gesetzt. Abbildung 17b zeigt ebenfalls eine erkennbare Struktur bis zum 47. Bit. Ab dem 48. Bit ist jedoch keine Struktur mehr feststellbar. Wie in Abbildung 16 gezeigt, werden die /48-Subnetze vollständig gescannt, wobei die Hexadezimalzeichen am 10. und 11. Nibble der Zieladressen offenbar durch einen Algorithmus vor dem Versenden der Pakete vermischt wurden. Ab dem 48. Bit wurde sehr wahrscheinlich eine andere Vorgehensweise angewandt, da dort kein Muster identifiziert wird.

Bei der Analyse des IID-Typs mit dem IPv6ToolKit werden alle IIDs der Zieladressen dieses Tages vom Scanner klassifiziert. Das Tool klassifiziert alle IIDs als *Randomized*, da diese Adressen keiner anderen Kategorie zugeordnet werden können. Beim Vergleich der Kategorisierung und der Scan-Analysen aus Abbildung 16, 17 und 18 wird festgestellt, dass beim dritten Segment alle Permutationen verwendet wurden, um viele Subnetze zu scannen. Der restliche Bereich der Zieladressen kann kein bestimmtes Muster entnommen werden. Jedoch wird anhand der Segmentanalyse in Abbildung 16 erkannt, dass die Permutationen in den Segmenten der IID alle auf einer konstante Menge von 4.096 liegen. Dies könnte auf eine Gleichverteilung hindeuten, die durch einen bestimmten Algorithmus erzeugt wird.

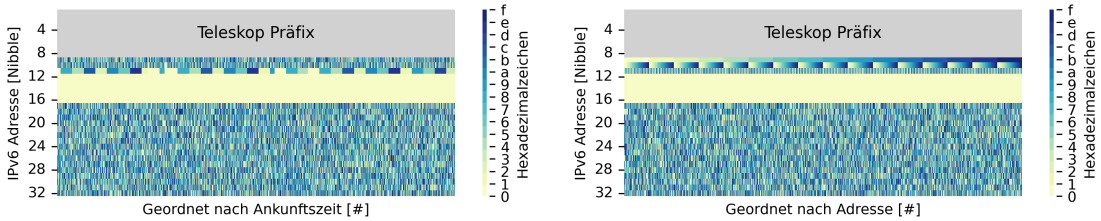
8.5.3 Analyse des meistgesehenen Scanners

Bei der Untersuchung der Pakete, die in P1 gesendet wurden, ist ein weiterer Scanner auffällig, der der AS-Organisation *Rica Web Services* zugeordnet ist. Es handelt sich dabei um einen Hoster. Im Rahmen der Untersuchungen wird festgestellt, dass dieser Scanner keine signifikante Aktivität innerhalb eines kurzen Zeitraums aufwies. Stattdessen erstreckte sich seine Aktivität über mehrere Tage, wobei immer wieder eine relativ gleichmäßige Anzahl von Paketen erfasst wurde.
Um dieses Verhalten genauer zu analysieren, werden die vom Scanner empfangenen Pakete eines ganzen Tages für die Analyse erfasst. Am 04. September 2023 wurden insgesamt 8.816 Pakete von diesem Scanner empfangen. In Abbildung 19 und 20 werden die Zieladressen des Scanners von diesem Tag dargestellt.



(a) Sortiert nach Ankunftszeit (b) Sortiert nach IPv6-Adresse

Abbildung 19: Scan-Tag am 04.09.2023 in binärer Darstellung



(a) Sortiert nach Ankunftszeit (b) Sortiert nach IPv6-Adresse

Abbildung 20: Scan-Tag am 04.09.2023 in hexadezimaler Darstellung

Bei der Betrachtung sowohl auf binärer als auch auf hexadezimaler Ebene zeigen sich klare Wiederholungen in der Verwendung bestimmter Hexadezimalzeichen innerhalb der Adressen. Es ist auffällig, dass alle Nibble vom 12. bis zum 16. auf *0* gesetzt sind. Nach dem 16. Nibble sind keine erkennbaren Muster oder Wiederholungen ersichtlich.

In Abbildung 20b sind, ähnlich wie beim Scanner der vorherigen Analyse, alle Hexadezimalzeichen am 9. Nibble sichtbar. An der darauf folgenden Position (10. Nibble) wurden ebenfalls alle Hexadezimalzeichen pro Zeichen am 9. Nibble gesetzt.

Das IPv6ToolKit klassifiziert die IIDs der Zieladressen von diesem Scanner wieder ausschließlich als zufällig generierte IIDs. Ein Muster bei der Generierung von Zieladressen ist daher nicht ersichtlich. Dennoch sind in der Abbildung 20a am 11. Nibble Wiederholungen der gesetzten Hexadezimalzeichen zu erkennen. Somit lässt sich dort eine Struktur feststellen. Vom 12. bis zum 16. Nibble wird die *0* gesetzt und innerhalb der IID kann in der Darstellung keine Struktur oder ein Muster festgestellt werden.

Zusammenfassend kann festgestellt werden, dass es Ähnlichkeiten bei der Generierung

der Zieladressen bei *Next-Layer* und *Rica Web Services* gibt. Vom 9. bis zum 11. Nibble scheint es so, als ob die Scanner versuchen, möglichst viele Subnetze zu scannen. Zusätzlich dazu ist innerhalb der IIDs beider Scan-Tage kein Muster ableitbar.

8.6 Erkenntnisse aus den Analysen

Die Analyse von P1 zeigt eine hohe Anzahl von ICMPv6-Paketen, während TCP-Pakete seltener, aber in einer sehr konstanten Menge empfangen wurden. UDP-Pakete wurden noch seltener empfangen. Eine wichtige Erkenntnis aus den UDP-Paketen war aber dafür, dass es sich bei den UDP-Paketen im gesamten Messzeitraum immer um Traceroute gehandelt hat. Die Analyse der TCP-Ziel-Ports zeigt eine dominante Nutzung von Port 80 für unverschlüsselten HTTP-Verkehr. In Bezug auf die Generierung von Zieladressen wird zusätzlich aus den vorherigen Analysen festgestellt, dass es Ähnlichkeiten bei der Zieladressgenerierung bei *Rica Web Services* und *Next-Layer* gibt. Beide Scanner verfolgen das Ziel, eine möglichst große Anzahl an Subnetzen zu erfassen, um einen maximalen Adressbereich abzudecken. Das primäre Ziel dieser Scanner scheint darin zu bestehen, das Präfix umfassend zu erkunden und dabei neue aktive Hosts zu identifizieren.

9 Ergebnisse aus den Scan-Analysen

In diesem Abschnitt werden die zentralen Ergebnisse der Analysen erläutert. Dazu gehören Erkenntnisse über den Einfluss der spezifischen Präfixeigenschaften auf den Netzwerkverkehr. Zudem werden die Resultate der Untersuchung des Scan-Verhaltens der aktivsten Scanner miteinander verglichen. Darüber hinaus wird eine Scan-Strategie identifiziert, die über die bisher bekannten Adresstypen hinausgeht.

9.1 Einfluss der Präfixeigenschaften

Größe des Präfixes. Das /48-Präfix P2 empfing innerhalb der ersten 68 Tage genau 2.883.644 Pakete während das /32-Präfix P1 insgesamt in den 68 Tagen des Messzeitraums 2.086.742 Pakete empfing. Beim Vergleich empfing P2 daher etwas mehr. Obwohl im /32-Präfix mehr Adressen vorhanden sind, scheint dies in diesem Fall keinen signifikanten Einfluss auf die Paketmenge zu haben. Daher lässt sich feststellen, dass ein größeres Präfix nicht zwangsläufig zu einer höheren Paketmenge führt.

Announcement. Ein Vergleich der empfangenen Paketmengen zwischen den vier Präfixen zeigt einen signifikanten Unterschied. P1 und P2 weisen eine weitaus größere Menge an empfangenen Paketen auf als P3 und P4. Der Unterschied besteht darin, dass P1 und P2 in BGP annonciert wurden, während P3 und P4 beides Subnetze eines annoncierten /29-Covering-Präfixes sind. Dies deutet darauf hin, dass BGP-Annoncements die Aufmerksamkeit von Scannern auf das Präfix lenken und somit mehr Netzwerkverkehr empfangen wird. Ein Subnetz eines solchen Präfixes ist für Scanner schwieriger zu finden und erhält daher wahrscheinlich weniger Pakete. Diese Vermutung soll in weiteren Arbeiten genauer untersucht werden.

Aktivität. Im Präfix P1 sind alle Adressen passiv. Im Gegensatz dazu enthält P2 ein aktives /56-Subnetz, das potenziell Scanner dazu anregen könnte, P2 weiter zu durchscannen und dadurch den Netzwerkverkehr in P2 zu erhöhen. Trotz dieser Möglichkeit zeigen die Analysen, dass beide Präfixe eine vergleichbare Anzahl an Paketen erhalten. Eine signifikant höhere Paketmenge in P2 ist nicht feststellbar. Dennoch empfängt P1 nur 72 % der Paketmenge im Vergleich zu P2. Dies lässt vermuten, dass das aktive Subnetz in P2 einen geringfügigen Einfluss auf die Paketmenge haben könnte.
Die Adressen in P4 sind passiv und reagieren nicht auf Anfragen. Im Gegensatz dazu nimmt P3 seit dem 16. Mai 2023 aktiv TCP-Verbindungen entgegen und reagiert auf Scan-Anfragen. Innerhalb eines Zeitraums von 30 Minuten werden sechs SYN-Pakete empfangen, auf die P3 reagiert. Anschließend werden ebenfalls ACK-Pakete empfangen. Somit hat diese Eigenschaft Auswirkungen auf den Netzwerkverkehr in P3. In einer weiteren Analyse sollten solche Ereignisse zu einem späteren Zeitpunkt ebenfalls untersucht werden.

Eintrag in der TUM-Hitliste. Scanner könnten P3 und P4 mit Hilfe der Hitliste nur über das /29-Covering-Präfix finden. Die geringe Paketmenge in den Subnetzen lässt darauf schließen, dass das alleinige Auftauchen des Covering-Präfixes auf der Hitliste nicht ausreicht, um viele Scanner anzuziehen. Da P1 und P2 beide auf der Hitliste stehen, könnte diese Eigenschaft ebenso für mehr Netzwerkverkehr in den Präfixen gesorgt haben. Während der ersten 15 Tage der Messung von P1 zeigt sich eine signifikante Zunahme der empfangenen Paketanzahl. Daher wird vermutet, dass der Hitlisteneintrag durchaus Scanner dazu veranlasst das Präfix zu scannen und somit Einfluss auf die Paketmenge nimmt.

9.2 Scan-Verhalten der aktivsten Scanner

Hammas Bin Tanveer *et al.* [19] identifizieren in ihrer Arbeit zwei vorherrschende Scan-Strategien: das zufällige Scannen und das Scannen von *Low-Byte-Adressen*. Das zufällige Scannen zielt darauf ab, eine möglichst gleichmäßige Verteilung über den gesamten Adressraum zu erreichen, während das Low-Byte-Scannen darauf abzielt, tiefer in den Adressraum vorzudringen.

Ein Vergleich der Auswertungen der Scans von *Tencent*, *Rica Web Services* und *Next-Layer* zeigt, dass die Verteilung in den Scans von *Rica Web Services* und *Next-Layer* eher auf das Auskundschaften des Präfixes hindeutet, um somit neue aktive Hosts zu entdecken. Dabei werden möglichst viele Subnetze beim Scannen abgedeckt. Aus der IID beider Scanner lassen sich keine Muster feststellen. Darum wird vermutet, dass die IIDs zufällig generiert werden. Im Gegensatz dazu zeigt der Scan von *Tencent* weniger Permutationen in den ersten Adresssegmenten. Mit der Verwendung unterschiedlicher Adresstypen in der IID wird vermutet, dass sich der Scanner auf bestimmte Subnetze fokussiert und dabei bestimmte Zieladressen scannt. Dabei werden unter anderem auch *Low-Byte-Scans* durchgeführt. Aus den Ergebnissen lassen sich Parallelen zu den Ergebnissen von Hammas Bin Tanveer *et al.* ableiten.

9.3 Das Scannen der Subnet-Router-Anycast-Adresse

In den empfangenen Paketen fallen neben den Zieladressen mit *::1* und *::2* am Ende auch Zieladressen mit *::0* auf. Bis zum 30. Oktober 2023 wurden in P1 insgesamt 8.357 solcher Pakete gezählt, die zu einer Zieladresse mit ::0 am Ende versendet wurden. In P2 werden sogar 16.072 solcher Pakete beobachtet. Dies deutet nicht unbedingt auf Zufall oder einen Konfigurationsfehler hin.
Ein Hinweis dazu steht in den Folien eines Vortrages der Black Hat aus dem Jahr 2021[5]. Bei diesem Vortrag wurden neue Methoden zum Scannen des IPv6-Adressraums vorgestellt, darunter das Scannen der *::0-Adressen*. Diese Adressen, die mit *::0* enden, werden als Subnet-Router-Anycast-Adressen bezeichnet [10]. Die Pakete werden an einen Router im Subnetz weitergeleitet. Der entsprechende Router antwortet auf die Anfrage der Subnet-Router-Anycast-Adresse mit seiner eigenen Adresse. Dieses Verhalten des Routers wurde selbst getestet und konnte bestätigt werden.

[5] https://i.blackhat.com/EU-21/Wednesday/EU-21-Shupeng-New-Ways-of-IPV6-Scanning.pdf

Es fällt auf, dass mehrere Quelladressen in regelmäßigen Abständen Pakete an diese Zieladressen versenden. Eine Annahme wäre, dass Scanner versuchen, von verschiedenen Peers Pakete zu versenden, um den *Border Router* raus zu finden. Ein möglicher Grund dafür wäre, dass versucht wird, die Topologie zu erlernen.

10 Zusammenfassung und Ausblick

Aus den Erkenntnissen der Analysen kann festgestellt werden, dass Präfixe im IPv6-Adressraum unterschiedliche Eigenschaften haben können. Diese Eigenschaften können Einfluss auf den empfangenen Netzwerkverkehr haben. Eine wichtige Erkenntnis ist, dass Präfixe, die in BGP annonciert werden mehr Scanner anziehen und somit mehr Netzwerkverkehr in diesen Präfixen zu beobachten ist. Darüber hinaus wird unterschiedliches Scan-Verhalten beobachtet. Es gibt sowohl Scanner, die sich auf das Auskundschaften der Präfixe konzentrieren, während andere sich eher auf bestimmte Subnetze und Zieladressen fokussieren.

Es soll in zukünftigen Arbeiten untersucht werden, wie die Scanner ein Präfix erkunden und welche Auswirkungen das Annoncieren in BGP auf die Scanner hat. Damit wird angestrebt, noch mehr Erkenntnisse über das IPv6-Scan-Verhalten zu identifizieren. Außerdem soll das bereits erweiterte reaktive Netzwerk-Teleskop Spoki [9] in strategisch ausgewählten Präfixen implementiert werden, um die Interaktion mit IPv6-Scannern zu ermöglichen, um somit den Scan-Verkehr noch detaillierter zu analysieren.

Literatur

[1] C. Bao, C. Huitema, M. Bagnulo, M. Boucadair, and X. Li. 2010. *IPv6 Addressing of IPv4/IPv6 Translators*. RFC 6052. IETF. https://doi.org/10.17487/RFC6052

[2] Jakub Czyz, Kyle Lady, Sam G. Miller, Michael Bailey, Michael Kallitsis, and Manish Karir. 2013. Understanding IPv6 internet background radiation. In *Proc. of the ACM IMC* (Barcelona, Spain). ACM, New York, NY, USA, 105–118. https://doi.org/10.1145/2504730.2504732

[3] S. Deering and R. Hinden. 1998. *Internet Protocol, Version 6 (IPv6) Specification*. RFC 2460. IETF. https://doi.org/10.17487/RFC2460

[4] Zakir Durumeric, Eric Wustrow, and J. Halderman. 2013. ZMap: Fast Internet-Wide Scanning and its Security Applications. In *In Proceedings of the 22nd USENIX Security Symposium*. 605–620.

[5] Mat Ford, J. Stevens, and John Ronan. 2006. Initial Results from an IPv6 Darknet. In *Proc. of the ICISP*. IEEE, Piscataway, NJ, USA, 13–13. https://doi.org/10.1109/ICISP.2006.14

[6] Kensuke Fukuda and John Heidemann. 2018. Who Knocks at the IPv6 Door? Detecting IPv6 Scanning. In *Proc. of IMC* (Boston, MA, USA) *(IMC '18)*. ACM, New York, NY, USA, 231–237. https://doi.org/10.1145/3278532.3278553

[7] Oliver Gasser, Quirin Scheitle, Sebastian Gebhard, and Georg Carle. 2016. Scanning the IPv6 Internet: Towards a Comprehensive Hitlist. In *Proc. of TMA* (Louvain La Neuve, Belgium). IFIP, Laxenburg, MD, Austria, 1–8.

[8] F. Gont and T. Chown. 2016. *Network Reconnaissance in IPv6 Networks*. RFC 7707. IETF. https://doi.org/10.17487/RFC7707

[9] Raphael Hiesgen, Marcin Nawrocki, Alistair King, Alberto Dainotti, Thomas C. Schmidt, and Matthias Wählisch. 2022. Spoki: Unveiling a New Wave of Scanners through a Reactive Network Telescope. In *Proc. of 31st USENIX Security Symposium*. USENIX Association, Berkeley, CA, USA, 431–448. https://www.usenix.org/system/files/sec22-hiesgen.pdf

[10] R. Hinden and S. Deering. 2006. *IP Version 6 Addressing Architecture*. RFC 4291. IETF. https://doi.org/10.17487/RFC4291

[11] ChenHuan Liu, ShanShan Hao, QianKun Liu, CongXiao Bao, and Xing Li. 2021. IPv6-Network Telescope Network Traffic Overview. In *2021 IEEE 11th ICEIEC*. IEEE, Piscataway, NJ, USA, 1–4. https://doi.org/10.1109/ICEIEC51955.2021.9463724

[12] G. Malkin. 1993. *Traceroute Using an IP Option*. RFC 1393. IETF. https://doi.org/10.17487/RFC1393

[13] Austin Murdock, Frank Li, Paul Bramsen, Zakir Durumeric, and Vern Paxson. 2017. Target Generation for Internet-Wide IPv6 Scanning. In *Proc. of IMC*. 242–253.

[14] J. Postel. 1981. *Internet Protocol*. RFC 791. IETF. https://doi.org/10.17487/RFC0791

[15] Philipp Richter, Oliver Gasser, and Arthur Berger. 2022. Illuminating Large-Scale IPv6 Scanning in the Internet. In *Proc. of the ACM IMC*. ACM, New York, NY, USA, 410–418. https://doi.org/10.1145/3517745.3561452

[16] John Ronan and David Malone. 2023. Revisiting and Revamping an IPv6 Network Telescope. In *2023 34th ISSC*. IEEE, Piscataway, NJ, USA, 1–6. https://doi.org/10.1109/ISSC59246.2023.10162033

[17] SI6 Networks. 2016. IPv6 Toolkit. https://www.si6networks.com/research/tools/ipv6toolkit/.

[18] Stephen D Strowes, René Wilhelm, Florian Obser, Riccardo Stagni, Agustín Formoso, and Emile Aben. 2020. Debogonising 2a10::/12 Analysis of one week's visibility of a new /12. In *Proc. of TMA*. IFIP, Laxenburg, MD, Austria, 1–9.

[19] Hammas Bin Tanveer, Rachee Singh, Paul Pearce, and Rishab Nithyanand. 2023. Glowing in the Dark: Uncovering IPv6 Address Discovery and Scanning Strategies in the Wild. In *Proc. of the USENIX Security Symposium*. USENIX Association, Anaheim, CA, 6221–6237. https://www.usenix.org/conference/usenixsecurity23/presentation/bin-tanveer

[20] Johanna Ullrich, Peter Kieseberg, Katharina Krombholz, and Edgar Weippl. 2015. On Reconnaissance with IPv6: A Pattern-Based Scanning Approach. In *10th International Conf on Availability, Reliability and Security*. IEEE, Piscataway, NJ, USA, 186–192.

[21] Liang Zhao, Satoru Kobayashi, and Kensuke Fukuda. 2024. Exploring the Discovery Process of Fresh IPv6 Prefixes: An Analysis of Scanning Behavior in Darknet and Honeynet. In *Proc. of PAM (LNCS, Vol. 14537)*. Springer, Berlin Heidelberg, 95–111. https://doi.org/10.1007/978-3-031-56249-5_4

[22] Johannes Zirngibl, Lion Steger, Patrick Sattler, Oliver Gasser, and Georg Carle. 2022. Rusty clusters? Dusting an IPv6 research foundation. In *Proc. of the ACM IMC*. ACM, New York, NY, USA, 395–409. https://doi.org/10.1145/3517745.3561440