Network Working Group
Internet-Draft
Intended status: Informational
Expires: June 25, 2022

Z. Du
China Mobile

December 22, 2021

Service Routing in Multi-access Edge Computing
draft-du-intarea-service-routing-in-mec-00

Abstract

This document introduces a service routing mechanism in the scenario of Multi-access Edge Computing.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

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 June 25, 2022.

Copyright Notice

Table of Contents

1. Introduction

The operators are deploying Multi-access Edge Computing (MEC) to provide services with lower latency to their users. Comparing to accessing service in the cloud, the MECs can provide service much nearer to the users.

However, in the current architecture of Internet, we need to send a DNS request to get the IP address of the service firstly, and then access the service [RFC1035]. It is not the optimal solution in the MEC scenarios which are sensitive to the latency of service accessing. In this document, we introduce a mechanism that can access the service directly without the DNS procedure.

In the 5G architecture, a UE (User Equipment) need to connect to a UPF (User Plane Function) working as a gateway, and then access service via the destination IP address.

In the scenarios of MEC, the service may be accessed within the MEC, meanwhile the MEC also provides a UPF Function for the UEs. Therefore, in fact, the service access takes place in a limited domain [RFC8799]. In this limited domain, we can use a specific IP address to directly access the service.

2. Proposed Mechanism Description

In the proposed mechanism, a UE should have a session with the UPF in the MEC. Also, the UE should be aware that it can access the service more quickly within the MEC if the service is available in the MEC. The proposed mechanism is described briefly as below.

Firstly, the UE send a normal DNS request if it wants to access a service, such as "www.local-weather.com". Meanwhile, it can make a destination IP address itself by hashing the URL, and try to establish a TCP connection directly.

Secondly, the UE may establish the connection successfully by using the specific IP address, and get access to the service bypassing the DNS procedure. If it fails, the UE can wait for the normal destination IP address received from the DNS procedure.

In this mechanism, the IP address can contain some information about the service, so we call it service routing in this document. The specific IP address is called the Service Routing IP address or the SR IP address.

3. SR IP Address

There are many options for the Service Routing IP address.

In the first option, we can assume that the UE can receive an MEC prefix for the service routing in the procedure of establishing the session between the UE and the UPF in the MEC. For example, an MEC prefix is 64 bits, and the hashed URL is also 64 bits. In the MEC, the server of the service should use the same hash algorithm to generate the SR IP address, and the 128 bits IPv6 address should be routed correctly within the MEC. Hence, the MEC works like a virtual network node containing services, with the MEC prefix as a Location, and the hashed URL as a Function.

In the second option, we can use a ULA IP address for the SR IP address [RFC8799]. The procedure is similar to the first option, but the SR IP address becomes the format of <MEC_ULA_Preifx: Hashed_URL>. The MEC_ULA_Prefix contains a specific subnet-ID.

In the last option, we can use all the 128 bits as the Hashed_URL. In this situation, the UE does not need to receive a specific prefix in advanced, and all the services in different MECs have the same IP address for the same service to support this quick access.

4. IANA Considerations

TBD.

5. Security Considerations

TBD.

6. Acknowledgements

TBD.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <https://www.rfc-editor.org/info/rfc4291>.

[RFC8799] Carpenter, B. and B. Liu, "Limited Domains and Internet Protocols", RFC 8799, DOI 10.17487/RFC8799, July 2020, <https://www.rfc-editor.org/info/rfc8799>.

7.2. Informative Referenc

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, DOI 10.17487/RFC1035, November 1987, <https://www.rfc-editor.org/info/rfc1035>.

Author’s Address

Zongpeng Du
China Mobile
No.32 XuanWuMen West Street
Beijing 100053
China

Email: duzongpeng@foxmail.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: 26 April 2023

Z. Du
China Mobile

23 October 2022

Service Routing in Multi-access Edge Computing
draft-du-intarea-service-routing-in-mec-02

Abstract

This document introduces a service routing mechanism in the scenario of Multi-access Edge Computing, in which the server's preferred address mechanism in QUIC can be used.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

The operators are deploying Multi-access Edge Computing (MEC) to provide services with lower latency to their users. Comparing to accessing service in the clouds, the MECs can provide service much nearer to the users.

However, in the current architecture of Internet, we need to send a DNS query to get the IP address of the service firstly, and then access the service [RFC1035]. It is not the optimal solution in the MEC scenarios which are sensitive to the latency of service accessing. In this document, we introduce a mechanism that can access the service directly without the DNS procedure.

In the 5G architecture, a UE (User Equipment) needs to connect to a UPF (User Plane Function) working as a gateway by using a tunnel, and then access service via the destination IP address.

In the scenarios of MEC, the service may be accessed within the MEC, meanwhile the MEC also provides a UPF Function for the UEs. Therefore, in fact, the service access takes place in a limited domain [RFC8799]. In this limited domain, we can use a specific IP address to directly access the service.

2. Proposed Mechanism Description

In the proposed mechanism, a UE should have a session with the UPF in the MEC. Also, the UE should be aware that it can access the service more quickly within the MEC if the service is available in the MEC. The proposed mechanism is described briefly as below.

Firstly, the UE sends a normal DNS query to the attached MEC if it wants to access a service, such as "www.local-weather.com". Meanwhile, it can send a connection establishment request for the service to the attached MEC, and try to establish a TCP/QUIC connection directly. In the request, the destination IP address is a specific IP made by the UE itself by hashing the domain name.

Secondly, the UE may establish the connection successfully by using the specific IP address, and get access to the service bypassing the DNS procedure. It will take place when the UE receives the response of the connection establishment request before receiving the response of the normal DNS query. If the DNS response returns firstly, the UE will do the normal service access procedure. It means that if the UE fails to establish a connection using the specific IP firstly, the UE can wait for the normal destination IP address received from the DNS procedure.

In this mechanism, the specific IP address can contain some information about the service, so we call it service routing in this document. The specific IP address is called the Service Routing IP address.

3. Service Routing IP Address

There are several options for the Service Routing IP address. The address has the same structure as the IPv6 address defined in [RFC4291].

In the first option, we can assume that the UE can receive an MEC prefix for the service routing in the procedure of establishing the session between the UE and the UPF in the MEC. For example, the length of an MEC prefix is 64 bits, and the length of the hashed domain name is also 64 bits. In the MEC, the server of the service should use the same hash algorithm to generate the Service Routing IP address, and the 128 bits IPv6 address should be routed correctly within the MEC. Hence, the MEC works like a virtual network node containing services, with the MEC prefix as a Location, and the hashed domain name as a Function.

In the second option, we can use a ULA IP address (Unique Local Address) for the Service Routing IP address [RFC8799]. The procedure is similar to the first option, but the Service Routing IP address becomes the format of <MEC_ULA_Preifx: Hashed_Domain_Name>. The MEC_ULA_Prefix contains a specific subnet-ID.

In the last option, we can use all the 128 bits as the Hashed_Domain_Name. In this situation, the UE does not need to receive a specific prefix in advanced, and all the services in different MECs have the same IP address for the same service to support this quick access.

4. Requirements of Service Routing Network Nodes

In the MEC, the network should support forwarding the Service Routing IP. In the client and server, they should support the binding of the Service Routing IP and the traditional DA IP. The value of the Service Routing IP exists mainly in the period of establishing the connection. After the connection is established, we can use the normal DA IP instead.

In the mechanism of this document, the MEC will receive a normal DNS query, and a connection establishment request for the service based on service routing. The MEC will try to establish the connection directly with the UE. Meanwhile, the MEC also does the normal DNS procedure for the UE. They take place independently, so that after the procedure of DNS, the MEC will response a target IP address to the UE no matter whether the connection establishment successes or fails by using the Service Routing IP address.

5. Server’s Preferred Address in QUIC

In QUIC [RFC9000], there is a "Server’s Preferred Address" mechanism. Perhaps it can help the DA changing process. QUIC allows servers to accept connections on one IP address and attempt to transfer these connections to a more preferred address shortly after the handshake.

We assume that the mechanism about the "Server’s Preferred Address" is supported both in the client and server, and the connection is a QUIC connection. Thus, the UE can use the hashed DA address to establish the connection, and after that, use the Server’s Preferred Address instead. In this situation, the Server’s Preferred Address should be the same as the normal DA IP address obtained in the DNS process mentioned before.

6. HASH Conflict between Services in MEC

At the beginning of the adoption of the mechanism, we do not think there would be too many essential services requiring this ultimate user experience, so that we assume that there would be no Hash conflict between the services. Besides, if there is any conflict in the MEC, the MEC can find it before deploying the service.

If the mechanism is adopted widely, and conflict exists between hashed domain names in the MEC, we can enable the mechanism only on the most essential service. Another option is to change the HASH algorithm that is running on the clients and severs to make a better Hash result.

7. Service Routing for Fixed Clients

MEC can also support accessing via fixed clients. In this situation, the BNG (Broadband Network Gateway) as the gateway of the client can work similarly to the UPF. A tunnel between the BNG and the MEC may be needed, and the MEC prefix can be obtained in the procedure of authentication. In the authentication of a fixed client, a more static session can be established because the client will not move.

8. IANA Considerations

TBD.

9. Security Considerations

TBD.

10. Acknowledgements

TBD.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <https://www.rfc-editor.org/info/rfc4291>.

[RFC8799] Carpenter, B. and B. Liu, "Limited Domains and Internet Protocols", RFC 8799, DOI 10.17487/RFC8799, July 2020, <https://www.rfc-editor.org/info/rfc8799>.

[RFC9000] Iyengar, J., Ed. and M. Thomson, Ed., "QUIC: A UDP-Based Multiplexed and Secure Transport", RFC 9000, DOI 10.17487/RFC9000, May 2021, <https://www.rfc-editor.org/info/rfc9000>.

11.2. Informative References

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, DOI 10.17487/RFC1035, November 1987, <https://www.rfc-editor.org/info/rfc1035>.

Author's Address

Zongpeng Du
China Mobile
No.32 XuanWuMen West Street
Beijing
100053
China
Email: duzongpeng@foxmail.com

intarea
Internet-Draft
Intended status: Best Current Practice
Expires: 8 September 2022

L. Giuliano

M. Aelmans

T. Li

7 March 2022

Regional Internet Blocking Considerations
draft-giuliano-blocking-considerations-00

Abstract

Geopolitical conflicts can cause policy makers to question whether or not blocking the Internet connectivity for an opposing region is a constructive tactic. This document provides an overview of the various technologies that can be used to implement regional blocking of Internet connectivity and discusses the implications of these options. This document does not advocate any policy or given blocking mechanism, but does attempt to articulate the implications of these blocking technologies for policy makers. The document also intends to help inform policy makers from countries who could be exposed to such blocking techniques on the implications of these methods.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Geopolitical conflicts can cause policy makers to question whether or not blocking the Internet connectivity for an opposing region is a constructive tactic. This document provides an overview of the various technologies that can be used to implement regional blocking of Internet connectivity and discusses the implications of these options. This document does not advocate any policy or given blocking mechanism, but does attempt to articulate the implications of these blocking technologies for policy makers. The document also intends to help inform policy makers from countries who could be exposed to such blocking techniques on the implications of these methods.

The content expressed in this document solely reflects the views of the authors and do not necessarily reflect the views or positions of any of our organizations, affiliates, friends, or enemies.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Scope

The scope of this document it limited to a description of well-known methods for disrupting core Internet services including physical cabling, Internet routing, filtering, and the Domain Name System.

The document does not intend to give any political directions or advocate for implementing the described methods, nor does it intend to be a guide for malicious attackers hence it will purely describe concepts and does not provide actual configuration or implementation methods.

3. Disconnection Methods

There are many ways of blocking a region’s Internet connectivity. In this section, we discuss some of them, their implications, capabilities, advantages, and disadvantages.

3.1. Physical layer

Disconnection at the physical layer is the most definitive method. Cutting cables and severing fibers will most definitely stop bits from passing. Unfortunately, this approach is also the most expensive to repair. In the optimistic view that disconnection is only intended to be temporary, creating downstream costs of physical repair is distinctly suboptimal. This approach may also be selected by the unscrupulous who have physical access to the media, but do not have further physical or managerial access.

A less destructive physical layer disconnection is simply disconnecting the fiber or cable, either at the terminating device, optical module, patch panel, amplifier, repeater, or transponder. This is easily repaired, but still requires physical access. Unscrupulous parties that wish to prevent easy repairs would be unlikely to select this option.

The simplest physical layer disconnection is administrative shutdown. Managerially disabling a physical circuit is a trivial configuration option that will sever communications. It is trivial to revert.

3.2. Routing layer

The Internet is a collection of many enterprises, web and cloud hosting, access-providers, etc. networks connected to each other using the Border Gateway Protocol (BGP) [RFC4271] to exchange routing information between those networks.

The simplest explanation of how BGP works is to compare it to a group of networks using the earlier described physical connections to gossip with each other on their knowledge about their own and neighboring networks. In other words, they exchange routing information describing how to reach destinations on the Internet.

Connected entities play different roles in this ecosystem, some will know how to reach all destinations on the Internet. These networks are so-called Tier 1 providers and provide connectivity to Tier 2 and Tier 3 networks. They offer services on a global scale and connect thousands of networks.

Tier 2 providers are large regional or national operators (for example the national service providers) offering services in country or regional. They have connections to many other networks, provide services to Tier 3 networks, but also purchase services from Tier 1 networks to reach destinations on the Internet they cannot reach themselves.

Tier 3 providers don’t provide routing information knowledge to other networks and are dependent from Tier 1 and Tier 2 networks to reach destinations on the Internet. In this category are small service providers or webhosting providers.

The BGP protocols offers several options to manipulate the routing information that is exchanged between these networks.

3.2.1. Autonomous System Number Filtering

Networks participating in the exchange of routing information with other networks use unique Autonomous System Numbers (ASNs) to identify themselves. These numbers are assigned to them by the Regional Internet Registries.

The ASN is used to construct a path to a destination prefix on the Internet. For example, a Tier 1 advertises its prefix to a Tier 2 originating from its own ASN. Next the Tier 2 advertises the prefix to a Tier 1 and adds its own ASN to the path. The route to reach this prefix now looks as follows: ASN2 ASN1.

As networks are highly connected there are many ASN paths through the Internet to reach destinations. The ’further away’ the destination is the longer the ASN path will be. On average most destinations on the Internet are reachable in a maximum of 5 hops. In other words, most destinations on the Internet can be reached via a maximum of 5 networks.

Networks that have a need to filter out another network with which they don’t have a direct peering session completely have, next to filtering prefixes, the option to filter on ASN. When applying an ASN filter, it will filter out all prefixes originating from that specific ASN.

Mitigating ASN filtering requires similar measures as mitigating prefix filters; networks with many upstream connections to Tier 1 and 2 networks will have a much lower chance of being completely filtered as, if one out of many upstream peers filters the ASN (and so its originating prefixes), others might still propagate them. This could still result in prefixes not being globally reachable anymore, but the chances are much lower.

3.2.2. De-peering

BGP uses a TCP session between two networks to exchange routing information. Such a session is called a peering session. Disabling such a session is referred to as de-peering.

3.2.2.1. De-peering Tier 3 networks

In many cases Tier 3 networks are using a single Tier 1 or 2 network for their connectivity to the Internet. In that case it's relatively easy to disconnect such networks from the Internet by disabling their peering sessions on the Tier 1 or 2 side.

3.2.2.2. De-peering Tier 2 networks

As described earlier these networks have multiple connections to other Tier 2 providers and typically between 2-8 Tier 1 providers to provide connectivity to the Internet. Subsequently, they could also receive routing information via Internet Exchange Points giving them even more options to reach destinations on the Internet.

De-peering such a network is much harder as one would need to disable peering sessions in many networks and at multiple (probably international) locations. Tier 2 networks will likely have international connections as well. Pursuing networks to disable these peering sessions in another jurisdiction could be very complicated.

3.2.2.3. De-peering Tier 1 networks

By their nature, Tier 1 networks have global span and have thousands of connections with other Tier 1, 2 and 3 networks. Fully disconnecting such networks is considered almost impossible without having physical and administrative access to the network itself. Pursuing other networks to de-peer a Tier 1 network is impossible because of the many countries they are present in and their jurisdictions.

3.2.3. Countering De-peering

Entities that want to protect themselves against de-peering would have a diversified connectivity strategy including multiple Tier 1 and 2 peers, actively peering on Internet Exchange Points, and preferably possessing its own physical infrastructure to connect to other networks in different countries or regions.

Tier 3 networks are most vulnerable to de-peering.

3.2.4. Prefix Filtering

Each network that is part of the Internet uses unique IPv4 and IPv6 address prefixes ranges to expose services to its directly connected (local) customers but also those connected via the Internet. These prefixes are advertised over a BGP peering session to the neighboring network so they will learn which prefixes originate from their neighbor and know how to reach them. Subsequently, they will advertise any routing knowledge they have about their neighboring networks and the neighboring network of their neighbors, etc. This way every network builds it own view of the Internet and map of how to reach destinations.

For example, Tier 1 and 2 networks will both have 'downstream' (customer) peering sessions with networks of which they have knowledge about; the prefixes they are advertising. If one of these networks wants to filter a neighbor, they could de-peer them as discussed earlier but that would basically filter all prefixes. In many cases, for example when intending to filter out social media, a subset of the prefixes is enough to accomplish this goal.

With this method a Tier 1 could also filter out prefixes from a Tier 3 that it learns via a Tier 2. De-peering the Tier 2 would result in all Tier 2 and all its customer prefixes becoming unreachable via this Tier 1. If only prefixes advertised by the single Tier 3 need to be filtered, the Tier 1 applies a prefix filter to the peering session from which it receives the advertisements.

Contrary, networks with many upstream connections to Tier 1 and 2 networks will have a much lower chance of being completely filtered as, if one out of many upstream peers filters the prefixes, others might still propagate them. This could still result in the prefix not being globally reachable anymore, but the chances are much lower.

3.3. Packet Filtering

Most network layer devices have the ability to filter traffic. The mechanism for doing this is commonly called an "Access Control List" (ACL). This is a possible mechanism for implementing a disconnection. Typically, an ACL allows filtering on a combination of source address, destination address, protocol number, and TCP/UDP source or destination port number.

3.3.1. GeoIP ACLs

The question then becomes one of ACL construction. However, this is not simple. IP address space is delegated in large sets, commonly known as 'prefixes.' Each prefix is assigned to an organization. Some organizations, such as Internet Service Providers (ISPs) will in turn delegate a portion of their address space to a customer. Customers and service providers do not necessarily fall along clean regional lines. Large multi-national corporations can receive a prefix from an ISP in one region and may use it in an entirely different region or even globally. They may also receive a prefix directly from a Regional Internet Registry (RIR). Service providers can obtain a prefix from an RIR and delegate parts of that prefix to customers from another region. This can create both false positives and false negatives when trying to map between a prefix and a region.

There are services which attempt to provide mappings from an IP address or prefix to a region, commonly called 'GeoIP' services. However, due to the above issues, these services cannot guarantee their accuracy. Constructing an ACL based on GeoIP services is likely to have unintended consequences, both filtering unintended addresses and not filtering intended addresses. Some commercial applications (notably streaming video) are willing to accept these inaccuracies, but this may not be acceptable in all circumstances.

Virtual Private Networks (VPNs) and other tunneling mechanisms can be used to create virtual topologies. If a single VPN server within a target region is not blocked, then it can provide access to innumerable other systems within the region, effectively bypassing GeoIP filtering services. When these are discovered, they are typically added to GeoIP databases, but this creates an ongoing battle between VPN service providers and GeoIP providers. As a result, this is an imperfect solution that may or may not be sufficiently accurate.

3.4. DNS

Blocking DNS capabilities can be an effective method for inhibiting end users from easily accessing Internet resources in a given region. For example, removing the delegation entries in the root servers for a given country code can prevent users from resolving the names of all domains for that country code. This approach can be circumvented to an extent with the creation of stub zones on resolving nameservers, which can provide a shortcut delegation to the country code top-level domain servers (ccTLDs) that are authoritative for that country code. But these stub zone entries would have to be manually created on any resolving nameserver that serves the resolution requests of users seeking resolution of domains for that

particular country code.

In the opposite direction, blocking resolution requests can inhibit users coming from a region from easily accessing Internet resources. Specifically, filters can be used to block resolving nameservers from a given region, or can block resolution requests from end users within a given region from making resolution requests to resolving nameservers that reside outside that region.

4. Gaps

The mechanisms discussed above cover the salient technical points for blocking a region. In this section, we discuss the various other considerations that are relevant to regional blocking.

4.1. Information Dissemination

At the very lowest level, the Internet copies bits from one location to another. Bits that are injected at one point are packetized, forwarded, and hopefully show up at their intended destination. The technology of the Internet does not care what is encoded in those bits. Whether it is state secrets or yesterday's grocery list, the Internet will happily ship it all the way around the world in milliseconds. The intrinsic value, properties, and attributes of the information conveyed in those bits is immaterial at the technological level.

4.1.1. Information Value

Policies considering blocking the transfer of information must also consider the value of the information that is being blocked. Filtering mechanisms can be extremely coarse and block all information, and this may not match the purposes of the policy. Thus, a blocking policy may need to be extremely specific about its goals and purposes.

A policy may want some information to be able to enter into a region. Sending certain messages into the region may be beneficial to the policy maker. Similarly, being able to get information out of a region may be beneficial. Further, parties within a region may be depending on global Internet connectivity to coordinate activities. A policy that blocks too much information may be counterproductive to the aims of the policy maker. A more selective policy would want some information to be communicated and not other information. Further, a selective policy is likely to be highly directional. Information that should flow into the region may not be permitted back out, and vice versa.

4.2. Information Concealment

If a policy allows any information to transit a boundary, then there is the technological possibility that other information may also transit that boundary. Information can be disguised or concealed through the use of cryptography, steganography, or other techniques. Policy makers should assume that any mechanism that allows any information to transit a boundary would eventually be used to transfer information against the purposes of the policy.

4.3. Misinformation

If a policy blocks information from flowing into a region, that may allow parties within that region to generate misinformation that is not disputed by outside information. This may be highly advantageous to the parties within the region. In the past, there have been many occurrences when parties within a region disconnected from the Internet precisely so that internal information could not be disputed.

4.4. Target Inaccuracy

The Internet infrastructure does not assign address space or ASNs according to strict regional, national, or continental boundaries. While there is some rough correlation, that is the result of administrative convenience. Thus, a prefix that is allocated from the general pool of European address space may end up covering part of Europe and Greenland. An ASN that was allocated for Singapore may be used in Australia.

This is further complicated by the fact that the parties that receive an ASN or prefix are not obligated or constrained to use it in a given region. If an organization acquires an ASN and subsequently grows outside of its original region, it may still use that ASN. If a company is assigned a prefix and the company is acquired by another firm, then that prefix could be used in a completely different hemisphere.

Consequently, if a policy elects to block traffic based on ASNs and prefixes, it may have unintended consequences, potentially blocking unintended traffic and not blocking proscribed traffic.

4.4.1. Accuracy of Registry Information

Many public resources are available to query Internet routing related information including, IPv4, IPv6 and ASN resource holders, routing intentions and actual reachability data. Unfortunately, the data doesn't always represent the actual situation, can be incomplete and in quite a few occasions outdated.

4.4.1.1. Internet Routing Registries

Internet Routing Registry (IRR) databases hold information about network operators routing intentions. For example, ASN holders can specify with whom they have peering relationships. This could give an indication which networks a specific ASN is connected to, however the data is entered (manually or automated) by network operators and isn't per se verified.

In practice IRR databases are between 40-70% accurate. However, some show an accuracy of around 95%.

4.4.1.2. RPKI

Resource Public Key Infrastructure (RPKI) is a public key infrastructure (PKI) framework to support improved security for the Internet's BGP routing infrastructure. The most important property is that in RPKI only legitimate resource holders can make statements about the IPv4, IPv6 and ASN resources they hold. This means that any information, right or wrong, found in RPKI databases represents the intention of, or at least is entered into RPKI by, the rightful holder.

RPKI is therefore considered to be 100% accurate. The downside of RPKI is that there aren't records for every resource and a large portion of the IPv4, IPv6 and ASN resources don't have records in RPKI.

4.5. Spoofing ASNs and Hijacking Prefixes

If a policy attempts to filter routing advertisements based on an ASN, then the opposition may attempt to counter that filtering attempt by using an alternate ASN. The alternate ASN may be an unused one, an ASN that has been assigned but is not actively in use elsewhere, or could be one that is actively assigned to another party. Using this ASN, the opposition could advertise its prefixes into BGP, bypass the ASN filter, and regain connectivity.

Similarly, if a policy attempts to filter routing advertisements or implement forwarding plane filters based on assigned prefixes, then the opposition may attempt to circumvent these policies by obtaining, advertising, and deploying alternate prefixes. As with ASNs, these prefixes could come from unassigned address space, address space that has been assigned but is not actively advertised, or even address space that is actively advertised by other parties.

There are security mechanisms that have been developed to help counter these possible attacks (IRR filtering [RFC7682], RPKI [RFC6811], and BGPsec [RFC8205]), but they are not ubiquitously deployed and may or may not be effective, depending on the operational procedures of ISPs that provide connectivity to the region.

4.6. Porous Borders

The Internet is, by design, a decentralized system of interconnections. Thus, it is nearly impossible to completely block Internet access for a region. Simply put, there will always be ways to circumvent any blocking attempts by sufficiently motivated parties. However, there are certain chokepoints and various methods, as described above, that can significantly inhibit connectivity and throughput for users going to/coming from a given region.

4.7. Acknowledgments

This document was inspired by the thoughtful comments of many friends and colleagues.

5. IANA Considerations

This document makes no requests of IANA.

6. Security Considerations

This document discusses technical and policy considerations of blocking Internet access for regions, and their potential impact on global security.

This document does not present new attack or defense strategies and merely discusses the implications of a variety of technical approaches. This document does not advocate or dissuade any policy about blocking Internet connectivity, it discusses various considerations that policy makers should understand prior to setting policy.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

7.2. Informative References

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[RFC6811] Mohapatra, P., Scudder, J., Ward, D., Bush, R., and R. Austein, "BGP Prefix Origin Validation", RFC 6811, DOI 10.17487/RFC6811, January 2013, <https://www.rfc-editor.org/info/rfc6811>.

[RFC7682] McPherson, D., Amante, S., Osterweil, E., Blunk, L., and D. Mitchell, "Considerations for Internet Routing Registries (IRRs) and Routing Policy Configuration", RFC 7682, DOI 10.17487/RFC7682, December 2015, <https://www.rfc-editor.org/info/rfc7682>.

[RFC8205] Lepinski, M., Ed. and K. Sriram, Ed., "BGPsec Protocol Specification", RFC 8205, DOI 10.17487/RFC8205, September 2017, <https://www.rfc-editor.org/info/rfc8205>.

Authors' Addresses

Lenny Giuliano
Email: lenny@lenny.net

Melchior Aelmans
Email: melchior@aelmans.eu

Tony Li
Email: tony.li@tony.li

Internet Area Working Group
Internet-Draft
Intended status: Informational
Expires: 7 September 2022

Y. Jia
D. Trossen
L. Iannone
Huawei
P. Mendes
Airbus
N. Shenoy
R.I.T.
L. Toutain
IMT-Atlantique
A. Y. Chen
Avinta
D. Farinacci
lispers.net
6 March 2022

Gap Analysis in Internet Addressing
draft-jia-intarea-internet-addressing-gap-analysis-02

Abstract

There exist many extensions to Internet addressing, as it is defined in [RFC0791] for IPv4 and [RFC8200] for IPv6, respectively. Those extensions have been developed to fill gaps in capabilities beyond the basic properties of Internet addressing. This document outlines those properties as a baseline against which the extensions are categorized in terms of methodology used to fill the gap together with examples of solutions doing so.

While introducing such extensions, we outline the issues we see with those extensions. This ultimately leads to consider whether or not a more consistent approach to tackling the identified gaps, beyond point-wise extensions as done so far, would be beneficial. The benefits are the ones detailed in the companion document [I-D.jia-intarea-scenarios-problems-addressing], where, leveraging on the gaps identified in this memo and scenarios provided in [I-D.jia-intarea-scenarios-problems-addressing], a clear problem statement is provided.

Status of This Memo

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

Copyright Notice

Table of Contents

1. Introduction

[I-D.jia-intarea-scenarios-problems-addressing] outlines scenarios and problems in Internet addressing through presenting a number of cases of communication that have emerged over the many years of utilizing the Internet and for which various extensions to the network interface-centric addressing of IPv6 have been developed. In order to continue the discussion on the emerging needs for addressing, initiated with [I-D.jia-intarea-scenarios-problems-addressing], this memo aims at identifying gaps between the Internet addressing model and desirable features that have been added by various extensions, in various contexts.

The approach to identifying the gaps is guided by key properties of Internet addressing, outlined in Section 2, namely (i) the fixed length of the IP addresses, (ii) the ambiguity of IP addresses semantic, while still (iii) providing limited IP address semantic support. Those properties are derived directly as a consequence of the respective standards that provide the basis for Internet addressing, most notably [RFC0791] for IPv4 and [RFC8200] for IPv6, respectively.

Those basic properties, and the potential issues that arise from those properties, give way to extensions that have been proposed over the course of deploying new Internet technologies. Section 3 discusses those extensions, summarized as gaps against the basic properties in Section 4.

Finally, this memo outlines issues that arise with the extension-driven approach to the basic Internet addressing, discussed in Section 6, arguing that any requirements for solutions that would revise the basic Internet addressing would require to address those issues.

2. Properties of Internet Addressing

As the Internet Protocol adoption has grown towards the global communication system we know today, its characteristics have evolved subtly, with [RFC6250] documenting various aspects of the IP service model and its frequent misconceptions, including Internet addressing. In this section, the three most acknowledged properties related to _Internet addressing_ are detailed. Those are (i) fixed IP address length, (ii) ambiguous IP address semantic, and (iii) limited IP address semantic support.

Section 3 elaborates on various extensions that aim to expand Internet addressing beyond those properties; those extensions are positioned as intentions to close perceived gaps against those key properties.

2.1. Property 1: Fixed Address Length

The fixed IP address length is specified as a key property of the design of Internet addressing, with 32 bits for IPv4 ([RFC0791]), and 128 bits for IPv6 ([RFC8200]), respectively. Given the capability of the hardware at the time of IPv4 design, a fixed length address was considered as a more appropriate choice for efficient packet forwarding. Although the address length was once considered to be variable during the design of Internet Protocol Next Generation ("IPng", cf., [RFC1752]) in the 1990s, it finally inherited the design of IPv4 and adopted a fixed length address towards the current IPv6. As a consequence, the 128-bit fixed address length of IPv6 is regarded as a balance between fast forwarding (i.e., fixed length) and practically boundless cyberspace (i.e., enabled by using 128-bit addresses).

2.2. Property 2: Ambiguous Address Semantic

Initially, the meaning of an IP address has been to identify an interface on a network device, although, when [RFC0791] was written, there were no explicit definitions of the IP address semantic.

With the global expansion of the Internet protocol, the semantic of the IP address is commonly believed to contain at least two notions, i.e., the explicit ’locator’, and the implicit ’identifier’. Because of the increasing use of IP addresses to both identify a node and to

indicate the physical or virtual location of the node, the intertwined address semantics of identifier and locator was then gradually observed and first documented in [RFC2101] as 'locator/identifier overload' property. With this, the IP address is used as an identification for host and server, very often directly used, e.g., for remote access or maintenance.

2.3. Property 3: Limited Address Semantic Support

Although IPv4 [RFC0791] did not add any semantic to IP addresses beyond interface identification (and location), time has proven that additional semantics are desirable (c.f., the history of 127/8 [HISTORY127] or the introduction of private addresses [RFC1918]), hence, IPv6 [RFC4291] introduced some form of additional semantics based on specific prefix values, for instance link-local addresses or a more structured multicast addressing. Nevertheless, systematic support for rich address semantics remains limited and basically prefix-based.

3. Filling Gaps through Extensions to Internet Addressing Properties

Over the years, a plethora of extensions has been proposed in order to move beyond the native properties of IP addresses, outlined in the previous section. The development of those extensions can be interpreted as filling gaps between the original properties of Internet addressing and desired new capabilities that those developing the extensions identified as being missing and yet needed and desirable.

3.1. Length Extensions

Extensions in this subsection aim at extending the property described in Section 2.1, i.e., the fixed IP address length.

When IPv6 was designed, the main objective was to create an address space that would not lead to the same situation as IPv4, namely to address exhaustion. To this end, while keeping the same addressing model like IPv4, IPv6 adopted a 128-bit address length with the aim of providing a sufficient and future-proof address space. The choice was also founded on the assumption that advances in hardware and Moore's law would still allow to make routing and forwarding faster, and the IPv6 routing table manageable.

We observe, however, that the rise of new use cases but also the number of new, e.g., industrial/home or small footprint devices, was possibly unforeseen. Sensor networks and more generally the Internet of Things (IoT) emerged after the core body of work on IPv6, thus different from IPv6 assumptions, 128-bit addresses were costly in

certain scenarios. On the other hand, given the huge investments that IPv6 deployment involved, certain solutions are expected to increase the addressing space of IPv4 in a compatible way, and thus extend the lifespan of the sunk investment on IPv4.

At the same time, it may also be possible to use variable and longer address lengths to address current networking demands. For example in content delivery networks, longer addresses such as URLs are required to fetch content, an approach that Information-Centric Networking (ICN) applied for any data packet sent in the network, using information-based addressing at the network layer. Furthermore, as an approach to address the routing challenges faced in the Internet, structured addresses may be used in order to avoid the need for routing protocols. Using variable length addresses allow as well to have shorter addresses. So for requirements for smaller network layer headers, shorter addresses could be used, maybe alleviating the need to compress other fields of the header. Furthermore, transport layer port numbers can be considered short addresses, where the high order bits of the extended address is the public IP of a NAT. Hence, in IoT deployments, the addresses of the devices can be really small and based on the port number, but they all share the global address of the gateway to make each one have a globally unique address.

3.1.1. Shorter Address Length

3.1.1.1. Description:

In the context of IoT [RFC7228], where bandwidth and energy are very scarce resources, the static length of 128-bit for an IP address is more a hindrance than a benefit since 128-bit for an IP address may occupy a lot of space, even to the point of being the dominant part of a packet. In order to use bandwidth more efficiently and use less energy in end-to-end communication, solutions have been proposed that allow for very small network layer headers instead.

3.1.1.2. Methodology:

* Header Compression/Translation: One of the main approaches to reduce header size in the IoT context is by compressing it. Such technique is based on a stateful approach, utilizing what is usually called a ’context’ on the IoT sensor and the gateway for communications between an IoT device and a server placed somewhere in the Internet - from the edge to the cloud.

The role of the 'context' is to provide a way to 'compress' the original IP header into a smaller one, using shorter address information and/or dropping some field(s); the context here serves as a kind of dictionary.

* Separate device from locator identifier: Approaches that can offer customized address length that is adequate for use in such constrained domains are preferred. Using different namespaces for the 'device identifier' and the 'routing' or 'locator identifier' is one such approach.

3.1.1.3. Examples

* Header Compression/Translation: Considering one base station is supposed to serve hundreds of user devices, maximizing the effectiveness for specific spectrum directly improves user quality of experience. To achieve the optimal utilization of the spectrum resource in the wireless area, the RObust Header Compression (ROHC) [RFC5795] mechanism, which has been widely adopted in cellar network like WCDMA, LTE, and 5G, utilizes header compression to shrink existing IPv6 headers onto shorter ones.

  Similarly, header compression techniques for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPAN) have been around for several years now, constituting a main example of using the notion of a 'shared context' in order to reduce the size of the network layer header ([RFC6282], [RFC7400], [ITU9959]). More recently, other compression solutions have been proposed for Low Power Wide Area Networks (LPWAN - [RFC8376]). Among them, the Static Context Header Compression (SCHC - [RFC8724]) generalized the compression mechanism developed by 6lo. Instead of a standard compression behavior implemented in all 6lo nodes, SCHC introduces the notion of rule shared by two nodes. The SCHC compression technique is generic and can be applied to IPv6 and above layers. Regarding the nature of the traffic, IPv6 addresses (source and destination) can be elided, partially sent, or replaced by a small index. Instead of the versatile IP packet, SCHC defines new packet formats dedicated to specific applications. SCHC rules are equivalence functions mapping this format to standard IP packets.

  Also, constraints coming from either devices or carrier links would lead to mixed scenarios and compound requirements for extraordinary header compression. For native IPv6 communications on DECT ULE and MS/TP Networks [RFC6282], dedicated compression mechanisms are specified in [RFC8105] and [RFC8163], while the transmission of IPv6 packets over NFC and PLC, specifications are being developed in [I-D.ietf-6lo-nfc] and [I-D.ietf-6lo-plc].

* Separate device from locator identifier: Solutions such as proposed in [EIBP] and [I-D.ietf-lisp-rfc6830bis] can utilize a separation of device from locator, where only the latter is used for routing between the different domains using the same technology, therefore enabling the use of shorter addresses in the (possibly constrained) local environment. Device IDs used within such domains are carried as part of the payload by EIBP and hence can be of shorter size suited to the domain, while, for instance, in LISP a flexible address encoding [RFC8060] allows shorter addresses to be supported in the LISP control plane [I-D.ietf-lisp-rfc6833bis].

3.1.2. Longer Address Length

3.1.2.1. Description

Historically, obtaining adequate address space is considered as the primary and raw motivation to invent IPv6. Longer address (more than 32-bit of IPv4 address), which can accommodate almost inexhaustible devices, used to be considered as the surest direction in 1990s. Nevertheless, to protect the sunk cost of IPv4 deployment, certain efforts focus on IPv4 address space depletion question but engineer IPv4 address length in a more practical way. Such effort, i.e., NAT (Network Address Translation), unexpectedly and significantly slows IPv6 deployment because of its high cost-effectiveness in practice.

Another crucial need for longer address lengths comes from "semantic extensions" to IP addresses, where the extensions themselves do not fit within the length limitation of the IP address. Section 3.3 discusses extensions which extend address semantics that are not limited by the IP address length.

This sub-section focuses on address length extensions that aim at reducing the IPv4 addresses depletion, while Section 3.3, i.e., address sematic extensions, may still refer to extensions when longer address length are suitable to accommodate different address semantic. See Section 3.3 for details of sematic-driven address lengthening.

3.1.2.2. Methodology

* Split address zone by network realm: This methodology first split the network realm into two types: one public realm (i.e., the Internet), and innumerable private realms (i.e., local networks, which may be embedded and/or having different scope). Then, it splits the IP address space into two type of zones: global address zone (i.e., public address) and local address zone (e.g., private address, reserved address). Based on this, it is assumed that in

public realm, all devices attached to it should be assigned an address that belongs to the global address zone. While for devices attached to private realms, only addresses belonging to the local address zone will be assigned. Local realms may have different scope or even be embedded one in another, like for instance, light switches local network being part of the building local network, which in turn connects to the Internet. In the local realms address may have a pure identification purpose. For instance in last example, addresses of the light switches identify the switches themselves, while the building local network is used to locate them.

Given that the local address zone is not globally unique, certain mechanisms are designed to express the relationship between the global address zone (in public realm) and the local address zone (in any private realm). In this case, global addresses are used for forwarding when a packet is in the public realm, and local addresses are used for forwarding when a packet is in a private realms.

3.1.2.3. Examples

* Split address zone by network realm: Network Address Translation (NAT), which was first laid out in [RFC2663], using private address and a stateful address binding to translate between the realms. As outlined in [RFC2663], basic address translation is usually extended to include port number information in the translation process, supporting bidirectional or simple outbound traffic only. Because the 16-bits port number is used in the address translation, NAT theoretically increase IPv4 address length from 32-bit to 48-bit, i.e., 281 trillion address space.

  Similarly, EzIP [EzIP] expects to utilize a reserved address block, i.e., 240/4, and an IPv4 header option to include it. Based on this, it can be regarded as EzIP is carrying a hierarchical address with two parts, where each part is a partial 32-bit IPv4 address. The first part is a public address residing in the "address field" of the header from globally routable IPv4 pool [IPv4pool], i.e., ca. 3.84 billion address space. The second part is the reserved address residing in "option field" and belongs to the 240/4 prefix, i.e., ca. 2^28=268 million. Based on that, each EzIP deployment is tethered on the existing Internet via one single IPv4 address, and EzIP then have 3.84B * 268M address, ca. 1,000,000 trillion. Collectively, the 240/4 can also be used as end point identifier and form an overlay network providing services parallel to the current Internet, yet independent of the latter in other aspects.

Compared to NAT, EzIP is able to establish a communication session from either side of it, hence being completely transparent, and facilitating a full end-to-end networking configuration.

3.1.3. Summary

Table 1 summarizes methodologies and examples towards filling gaps on IP address length extensions.

	Methodology	Examples
Shorter Address Length	Header compression/ translation	6LoWPAN, ROHC, SCHC
	Separate device from locator identifier	EIBP, LISP, ILNP, HIP
Longer Address Length	Split address zone by network realm	NAT, EzIP

Table 1: Summary Length Extensions

3.2. Identity Extensions

Extensions in this subsection attempt extending the property described in Section 2.2, i.e., 'locator/identifier overload' of the ambiguous address semantic.

From the perspective of Internet users, on the one hand, the implicit identifier semantic results in a privacy issue due to network behavior tracking and association. Despite that IP address assignments may be dynamic, they are nowadays considered as 'personal data' and as such undergoes privacy protection regulations like General Data Protection Regulation ("GDPR" [GDPR]). Hence, additional mechanisms are necessary in order to protect end user privacy.

For network regulation of sensitive information, on the other hand, dynamically allocated IP addresses are not sufficient to guarantee device or user identification. As such, different address allocation systems, with stronger identification properties are necessary where security and authentication are at highest priority. Hence, in order to protect information security within a network, additional mechanism are necessary to identify the users or the devices attached to the network.

3.2.1. Anonymous Address Identity

3.2.1.1. Description

As discussed in Section 2.2, IP addresses reveal both 'network locations' as well as implicit 'identifier' information to both traversed network elements and destination nodes alike. This enables recording, correlation, and profiling of user behaviors and historical network traces, possibly down to individual real user identity. The IETF, e.g., in [RFC7258], has taken a clear stand on preventing any such pervasive monitoring means by classifying them as an attack on end users' right to be left alone (i.e., privacy). Regulations such as the EU's General Data Protection Regulation (GDPR) classifies, for instance, the 'online identifier' as personal data which must be carefully protected; this includes end users' IP addresses [GDPR].

Even before pervasive monitoring [RFC7258], IP addresses have been seen as something that some organizational owners of networked system may not want to reveal at the individual level towards any non-member of the same organization. Beyond that, if forwarding is based on semantic extensions, like other fields of the header, extension headers, or any other possible extension, if not adequately protected it may introduce privacy leakage and/or new attack vectors.

3.2.1.2. Methodology:

* Traffic Proxy: Detouring the traffic to a trusted proxy is a heuristic solution. Since nodes between trusted proxy and destination (including the destination per se) can only observe the source address of the proxy, the 'identification' of the origin source can thereby be hidden. To obfuscate the nodes between origin and the proxy, the traffic on such route would be encrypted via a key negotiated either in-band or off-band. Considering that all applications' traffic in such route can be seen as a unique flow directed to the same 'unknown' node, i.e., the trusted proxy, eavesdroppers in such route have to make more efforts to correlate user behavior through statistical analysis even if they are capable of identifying the users via their source addresses. The protection lays in the inability to isolate single application specific flows. According to the methodology, such approach is IP version independent and works for both IPv4 and IPv6.

* Source Address Rollover: Privacy issues related to address 'identifier' semantic can be mitigated through regular change (beyond the typical 24 hours lease of DHCP). Due to the semantics of 'identifier' that an IP address carries, such approach promotes

to change the source IP address at a certain frequency. Under such methodology, the refresh cycling window may reach to a balance between privacy protection and address update cost. Due to the limited space that IPv4 contains, such approach usually works for IPv6 only.

* Private Address Spaces: Their introduction in [RFC1918] foresaw private addresses (assigned to specific address spaces by the IANA) as a means to communicate purely locally, e.g., within an enterprise, by separating private from public IP addresses. Considering that private addresses are never directly reachable from the Internet, hosts adopting private addresses are invisible and thus 'anonymous' for the Internet. Besides, hosts for purely local communication used the latter while hosts requiring public Internet service access would still use public IP addresses.

* Address Translation: The aforementioned original intention for using private IP addresses, namely for purely local communication, resulted in a lack of flexibility in changing from local to public Internet access on the basis of what application would require which type of service.

  If eventually every end-system in an organization would require some form of public Internet access in addition to local one, an adequate number of public Internet addresses would be required for providing to all end systems. Instead, address translation enables to utilize many private IP addresses within an organization, while only relying on one (or few) public IP addresses for the overall organization.

  In principle, address translation can be applied recursively. This can be seen in modern broadband access where Internet providers may rely on carrier-grade address translation for all their broadband customers, who in turn employ address translation of their internal home or office addresses to those (private again) IP addresses assigned to them by their network provider.

  Two benefits arise from the use of (private to public IP) address translation, namely (i) the hiding of local end systems at the level of the (address) assigned organization, and (ii) the reduction of public IP addresses necessary for communication across the Internet. While the latter has been seen for long as a driver for address translation, we focus on the first issue in this section, also since we see such privacy benefit as well as objective as still being valid in addressing systems like IPv6 where address scarcity is all but gone [GNATCATCHER].

* Separate device from locator identifier: Solutions that make a clear separation between the routing locator and the identifier, can allow for a device ID of any size, which in turn can be encrypted by a network element deployed at the border of routing domain (e.g., access/edge router). Both source and end-domain addresses can be encrypted and transported, as in the routing domain, only the routing locator is used.

3.2.1.3. Examples:

* Traffic Proxy: Although not initially designed as a traffic proxy approach, a Virtual Private Network (VPN [VPN]) is widely utilized for packets origin hiding as a traffic detouring methodology. As it evolved, VPN derivatives like WireGuard [WireGuard] have become a mainstream instance for user privacy and security enhancement.

  With such methodology in mind, onion routing [ONION], instantiated in the TOR Project [TOR], achieves high anonymity through traffic hand over via intermediates, before reaching the destination. Since the architecture of TOR requires at least three proxies, none of them is aware of the entire route. Given that the proxies themselves can be deployed all over cyberspace, trust is not the prerequisite if proxies are randomly selected.

  In addition, dedicated protocols are also expected to be customized for privacy improvement via traffic proxy. For example, Oblivious DNS over HTTPS (ODoH [ODoH]) use a third-party proxy to obscure identifications of user source addresses during DNS over HTTPS (DoH [RFC8484]) resolution. Similarly, Oblivious HTTP [OHTTP] involve proxy alike in the HTTP environment.

* Source Address Rollover: As for source address rollover, it has been standardized that IP addresses for Internet users should be dynamic and temporary every time they are being generated [RFC8981]. This benefits from the available address space in the case of IPv6, through which address generation or assignment should be unpredictable and stochastic for outside observers.

  More radically, [EPHEMERALv6] advocates an ’ephemeral address’, changing over time, for each process. Through this, correlating user behaviors conducted by different identifiers (i.e., source address) becomes much harder, if not impossible, if based on the IP packet header alone.

* Private Addresses: The use and assignment of private addresses for IPv4 is laid out in [RFC1918], while unique local addresses (ULAs) in IPv6 [RFC4193] take over the role of private address spaces in IPv4.

* Network Address Translation: Given address translation can be performed several times in cascade, NATs may exist as part of existing customer premise equipment (CPE), such as a cable or an Ethernet router, with private wired/wireless connectivity, or may be provided in a carrier environment to further translate ISP-internal private addresses to a pool of (assigned) public IP addresses. The latter is often dynamically assigned to CPEs during its bootstrapping.

* Separate device from locator identifier: EIBP [EIBP] utilizes a structured approach to addressing. It separates the routing ID from the device ID, where only the former is used for routing. As such, the device IDs can be encrypted, protecting the end device identity. Similarly, LISP uses separate namespaces for routing and identification allowing to 'hide' identifiers in encrypted LISP packets that expose only known routing information [RFC8061].

3.2.2. Authenticated Address Identity

3.2.2.1. Description

In some scenarios (e.g., corporate networks) it is desirable to being able authenticate IP addresses in order to prevent malicious attackers spoofing IP addresses. This is usually achieved by using a mechanism that allows to prove ownership of the IP address.

3.2.2.2. Methodology

* Self-certified addresses: This method is usually based on the use of nodes' public/private keys. A node creates its own interface ID (IID) by using a cryptographic hash of its public key (with some additional parameters). Messages are then signed using the nodes' private key. The destination of the message will verify the signature through the information in the IP address. Self-certification has the advantage that no third party or additional security infrastructure is needed. Any node can generate its own address locally and then only the address and the public key are needed to verify the binding between the public key and the address.

* Third party granted addresses: DHCP (Dynamic Host Configuration Protocol) is widely used to provide IP addresses, however, in its basic form, it does not perform any check and even an unauthorized user without the right to use the network can obtain an IP address. To solve this problem, a trusted third party has to grant access to the network before generating an address (via DHCP or other) that identifies the user. User authentication done securely either based on physical parameters like MAC addresses or based on an explicit login/password mechanism.

3.2.2.3. Examples

* Self-certified Addresses: As an example of this methodology serves [RFC3972], defining IPv6 cryptographically Generated Addresses (CGA). A Cryptographically Generated Address is formed by replacing the least-significant 64 bits of an IPv6 address with the cryptographic hash of the public key of the address owner. Packets are then signed with the private key of the sender. Packets can be authenticate by the receiver by using the public key of the sender and the address of the sender. The original specifications have been already amended (cf., [RFC4581] and [RFC4982]) in order to support multiple (stronger) cryptographic algorithms.

* Third party granted addresses: [RFC3118] defines a DHCP option through which authorization tickets can be generated and newly attached hosts with proper authorization can be automatically configured from an authenticated DHCP server. Solutions exist where separate servers are used for user authentication like [UA-DHCP] and [RFC4014]. The former proposing to enhance the DHCP system using registered user login and password before actually providing an IP address lease and recording the MAC address of the device the user used to sign-in. The latter, couples the RADIUS authentication protocol ([RFC2865]) with DHCP, basically piggybacking RADIUS attributes in a DHCP sub-option, with the DHCP server contacting the RADIUS server to authenticate the user.

3.2.3. Summary

Table 2, summarize the methodologies and the examples towards filling the gaps on identity extensions.

	Methodology	Examples
Anonymous Address Identity	Traffic Proxy	VPN, TOR, ODoH
	Source Address Rollover	SLAAC
	Private Address Spaces	ULA
	Address Translation	NAT
	Separate device from locator identifier	EIBP, LISP
Authenticated Address Identity	Self-certified Addresses	CGA
	Third party granted addresses	DHCP-Option

Table 2: Summary Identity Extensions

3.3. Semantic Extensions

Extensions in this subsection try extending the property described in Section 2.3, i.e., limited address semantic support.

As explained in Section 2.2, IP addresses carry both locator and identification semantic. Some efforts exist that try to separate these semantics either in different address spaces or through different address formats. Beyond just identification, location, and the fixed address size, other efforts extended the semantic through existing or additional header fields (or header options) outside the Internet address.

How much unique and globally routable an address should be? With the effect of centralization, edges communicate with (rather) local DCs, hence a unique address globally routable is not a requirement anymore. There is no need to use globally unique addresses all the time for communication, however, there is the need of having a unique address as a general way to communicate to any connected entity without caring what transmission networks the packets traverse.

3.3.1. Utilizing Extended Address Semantics

3.3.1.1. Description

Several extensions have been developed to extend beyond the limited IPv6 semantics. Those approaches may include to apply structure to the address, utilize specific prefixes, or entirely utilize the IPv6 address for different semantics, while re-encapsulating the original packet to restore the semantics in another part of the network. For instance, structured addresses have the capability to introduce delimiters to identify semantic information in the header, therefore not constraining any semantic by size limitations of the address fields.

We note here that extensions often start out as being proposed as an extended header semantic, while standardization may drive the solution to adopt an approach to accommodate their semantic within the limitations of an IP address. This section does include examples of this kind.

3.3.1.2. Methodology

*Semantic prefixes: Semantic prefixes are used to separate the IPv6 address space. Through this, new address families, such as for information-centric networking [HICN], service routing or other semantically rich addressing, can be defined, albeit limited by the prefix length and structure as well as the overall length limitation of the IPv6 address.

* Separate device/resource from locator identifier: The option to use separate namespaces for the device address would offer more freedom for the use of different semantics. For instance, the static binding of IP addresses to servers creates a strong binding between IP addresses and service/resources, which may be a limitation for large Content Distribution networks (CDNs) [FAYED21].

  As an extreme form of separating resource from locator identifier, recent engineering approaches, described in [CLOUDFLARE_SIGCOMM], decouple web service (semantics) from the routing address assignments by using virtual hosting capabilities, thereby effectively mapping possibly millions of services onto a single IP address.

* Structured addressing: One approach to address the routing challenges faced in the Internet is the use of structured addresses, e.g., to void the need for routing protocols. Benefits of this approach can be significant, with the structured addresses

capturing the relative physical or virtual position of routers in the network as well as being variable in length. Key to the approach, however, is that the structured addresses capturing the relative physical or virtual position of routers in the network, or networks in an internetwork may not fit within the fixed and limited IP address length (cf., Section 3.1.2). Other structured approaches may be the use of application-specific structured binary components for identification, generalizing URL schema used for HTTP-level communication but utilized at the network level for traffic steering decisions.

* Localized forwarding semantics: Layer 2 hardware, such as SDN switches, are limited to the use of specific header fields for forwarding decisions. Hence, devising new localized forwarding mechanisms may be based on re-using differently existing header fields, such as the IPv6 source/destination fields, to achieve the desired forwarding behavior, while encapsulating the original packets in order to be restored at the local forwarding network boundary. Networks in those solutions are limited by the size of the utilized address field, e.g., 256 bits for IPv6, thereby limiting the way such techniques could be used.

3.3.1.3. Examples

* Semantic prefixes: Newer approaches to IP anycast suggest the use of service identification in combination with a binding IP address model [SFCANYCAST] as a way to allow for metric-based traffic steering decisions; approaches for Service Function Chaining (SFC) [RFC7665] utilize the Network Service Header (NSH) information and packet classification to determine the destination of the next service.

  Another example of the usage of different packet header extensions based on IP addressing is Segment Routing. In this case, the source chooses a path and encodes it in the packet header as an ordered list of segments. Segments are encoded using new Routing Extensions Header type, the Segment Routing Header (SRH), which contains the Segment List, similar to what is already specified in [RFC8200], i.e., a list of segment ID (SID) that dictate the path to follow in the network. Such segment IDs are coded as 128 bit IPv6 addresses [RFC8986].

  Approaches such as [HICN] utilize semantic prefixing to allow for ICN forwarding behavior within an IPv6 network. In this case, an HICN name is the hierarchical concatenation of a name prefix and a name suffix, in which the name prefix is encoded as an IPv6 128 bits word and carried in IPv6 header fields, while the name suffix is encoded in transport headers fields such as TCP. However, it

is a challenge to determine which IPv6 prefixes should be used as name prefixes. In order to know which IPv6 packets should be interpreted based on an ICN semantic, it is desirable to be able to recognize that an IPv6 prefix is a name prefix, e.g. to define a specific address family (AF_HICN, b0001::/16). This establishment of a specific address family allows the management and control plane to locally configure HICN prefixes and announce them to neighbors for interconnection.

* Separate device from locator identifier: EIBP [EIBP] separates the routing locator from the device identifier, relaxing therefore any semantic constraints on the device identifier. Similarly, LISP uses a flexible encoding named LISP Canonical Address Format (LCAF [RFC8061]), which allows to associate to routing locators any possible form (and length) of identifier. ILNP [RFC6740] introduces as well a different semantic of IP addresses, while aligning to the IPv6 address format (128 bits). Basically, ILNP introduces a sharper logical separation between the 64 most significant bits and the 64 least significant bits of an IPv6 address. The former being a global locator, while the latter being an identifier that can have different semantics (rather than just being an interface identifier).

* Structured addressing: Network topology captures the physical connectivity among devices in the network. There is a structure associated with the topology. Examples are the core-distribution-access router structure commonly used in enterprise networks and clos topologies that are used to provide multiple connections between Top of Rack (ToR) devices and multiple layers of spine devices. Internet service providers use a tier structure that defines their business relationships. A clear structure of connected networks can be noticed in the Internet. EIBP [EIBP] proposes to leverage the physical structure (or a virtual structure overlaid on the physical structure) to auto assign addresses to routers in a network or networks in an internetwork to capture their relative position in the physical/virtual topology. EIBP proposes to administratively identify routers/networks with a tier value based on the structure.

* Localized forwarding semantics: Approaches such as those outlined in [REED] suggest using a novel forwarding semantic based on path information carried in the packet itself, said path information consists in a fixed size bit-field (see [REED] for more information on how to represent the path information in said bit-field). In order to utilize existing, e.g., SDN-based, forwarding switches, the direct use of the IPv6 source/destination address is suggested for building appropriate match-action rules (over the suitable binary information representing the local output ports),

while preserving the original IPv6 information in the encapsulated packet. As mentioned above, such use of the existing IPv6 address fields limits the size of the network to a maximum of 256 bits (therefore paths in the network over which such packets can be forwarded). [ICNIP], however, goes a step further by suggesting to use the local forwarding as direct network layer mechanism, removing the IP packet and only leaving the transport/application layer, with the path identifier constituting the network-level identifier albeit limited by using the existing IP header for backward compatibility reasons (the next section outlines the removal of this limitation).

3.3.2. Utilizing Existing or Extended Header Semantics

3.3.2.1. Description:

While the former sub-section explored extended address semantic, thereby limiting any such extended semantic with that of the existing IPv6 semantic and length, additional semantics may also be placed into the header of the packet or the packet itself, utilized for the forwarding decision to the appropriate endpoint according to the extended semantic.

Reasons for embedding such new semantics may be related to traffic engineering since it has long been shown that the IP address itself is not enough to steer traffic properly since the IP address itself is not semantically rich enough to adequately describe the forwarding decision to be taken in the network, not only impacting WHERE the packet will need to go but also HOW it will need to be sent.

3.3.2.2. Methodology:

* In-Header extensions: One way to add additional semantics besides the address fields is to use other fields already present in the header.

* Headers option extensions: Another mechanism to add additional semantics is to actually add additional fields, e.g., through Header Options in IPv4 or through Extension Headers in IPv6.

* Re-encapsulation extension: A more radical approach for additional semantics is the use of a completely new header that is designed so to carry the desired semantics in an efficient manner (often as a shim header).

* Structured addressing: Similar to the methodology that structures addresses within the limitations of the IPv6 address length, outlined in the previous sub-sections, structured addressing can

also be applied within existing or extended header semantics, e.g., utilizing a dedicated (extension) header to carry the structured address information.

* Localized forwarding semantics: This set of solutions applies capabilities of newer (programmable) forwarding technology, such as [P4], to utilize any header information for a localized forwarding decision. This removes any limitation to use existing header or address information for embedding a new address semantic into the transferred packet.

3.3.2.3. Examples:

* In-Header extensions: In order to allow additional semantic with respect to the pure Internet addressing, the original design of IPv4 included the field 'Type of Service' [RFC2474], while IPv6 introduced the 'Flow label' and the 'Traffic Class' [RFC8200]. In a certain way, those fields can be considered 'semantic extensions' of IP addresses, and they are 'in-header' because natively present in the IP header (differently from options and extension headers). However, they proved not to be sufficient. Very often a variety of network operation are performed on the well-known 5-tuple (source and destination addresses; source and destination port number; and protocol number). In some contexts all of the above mentioned fields are used in order to have a very fine grained solution ([RFC8939]).

* Headers option extensions: Header options have been largely under-exploited in IPv4. However, the introduction of the more efficient extension header model in IPv6 along with technology progress made the use of header extensions more widespread in IPv6. Segment Routing re-introduced the possibility to add path semantic to the packet by encoding a loosely defined source routing ([RFC8402]). Similarly, in the aim to overcome the inherent shortcoming of the multi-homing in the IP context, SHIM6 ([RFC5533]) also proposed the use of an extension header able to carry multi-homing information which cannot be accommodated natively in the IPv6 header.

  To serve a moving endpoint, mechanisms like Mobile IPv6 [RFC6275] are used for maintaining connection continuity by a dedicated IPv6 extension header. In such case, the IP address of the home agent in Mobile IPv6 is basically an identification of the on-going communication. In order to go beyond the interface identification model of IP, the Host Identity Protocol (HIP) tries to introduce an identification layer to provide (as the name says) host identification. The architecture here relies on the use of another type of extension header [RFC7401].

* Re-encapsulation extension: Differently from the previous approach, re-encapsulation prepends complete new IP headers to the original packet introducing a completely custom shim header between the outer and inner header. This is the case for LISP, adding a LISP specific header right after an IP+UDP header ([I-D.ietf-lisp-rfc6830bis]). A similar design is used by VxLAN ([RFC7348]) and GENEVE ([RFC8926]), even if they are designed for a data center context. IP packets can also be wrapped with headers using more generic and semantically rich names, for instance with ICN [ICNIP].

* Structured addressing: Solutions such as those described in the previous sub-section, e.g., EIBP [EIBP], can provide structured addresses that are not limited to the IPv6 address length but instead carry the information in an extension header to remove such limitation.

  Also Information-Centric Networking (ICN) naming approaches usually introduce structures in the (information) names without limiting themselves to the IP address length; more so, ICN proposes its own header format and therefore radically breaks with not only IP addressing semantic but the format of the packet header overall. For this, approaches such as those described in [RFC8609] define a TLV-based binary application component structure that is carried as a 'name' part of the CCN messages. Such a name is a hierarchical structure for identifying and locating a data object, which contains a sequence of name components. Names are coded based on 2-level nested Type-Length-Value (TLV) encodings, where the name-type field in the outer TLV indicates this is a name, while the inner TLVs are name components including a generic name component, an implicit SHA-256 digest component and a SHA-256 digest of Interest parameters. For textual representation, URIs are normally used to represent names, as defined in [RFC3986].

  In geographic addressing, position based routing protocols use the geographic location of nodes as their addresses, and packets are forwarded when possible in a greedy manner towards the destination. For this purpose, the packet header includes a field coding the geographic coordinates (x, y, z) of the destination node, as defined in [RFC2009]. Some proposals also rely on extra fields in the packet header to code the distance towards the destination, in which case only the geographic coordinates of neighbors are exchanged. This way the location of the destination is protected even if routing packets are eavesdropped.

* Localized forwarding semantics: Unlike the original suggestion in [REED] to use existing SDN switches, the proliferation of P4 [P4] opens up the possibility to utilize a locally limited address semantic, e.g., expressed through the path identifier, as an entirely new header (including its new address) with an encapsulation of the IP packet for E2E delivery (including further delivery outside the localized forwarding network or positioning the limited address semantic directly as the network address semantic for the packet, i.e., removing any IP packet encapsulation from the forwarded packet, as done in [ICNIP]. Removing the IPv6 address size limitation by not utilizing the existing IP header for the forwarding decision also allows for extensible length approaches for building the path identifier with the potential for increasing the supported network size. On the downside, this approach requires to encapsulate the original IP packet header for communication beyond the local domain in which the new header is being used, such as discussed in the previous point above on 're-encapsulation extension'.

3.3.3. Summary

Table 3, summarize the methodologies and the examples towards filling the gaps on semantic extensions.

	Methodology	Examples
Utilizing Extended Address Semantics	Semantic prefixes	HICN
	Separate device from locator identifier	EIBP, ILNP, LISP, HIP
	Structured addressing	EIBP, ILNP
	Localized forwarding semantics	REED
Utilizing Existing or Extended Header Semantics	In-Header extensions	DetNet
	Headers option extensions	SHIM6, SRv6, HIP
	Re-encapsulation extension	VxLAN, ICNIP
	Structured addressing	EIBP
	Localized forwarding semantics	REED

Table 3: Summary Semantic Extensions

4. Overview of Approaches to Extend Internet Addressing

The following Table 4 describes the objectives of the extensions discussed in this memo with respect to the properties of Internet addressing (Section 2}. As summarized, extensions may aim to extend one property of the Internet addressing, or extend other properties at the same time.

	Length Extension	Identity Extension	Semantic Extension
6LoWPAN	x		
ROHC	x		

EzIP	x		
TOR		x	
ODoH		x	
SLAAC		x	
CGA		x	x
NAT	x	x	
HICN		x	x
ICNIP	x	x	x
CCNx names	x	x	x
EIBP	x	x	x
Geo addressing	x		x
REED	x (with P4)		x
DetNet		x	
Mobile IP			x
SHIM6			x
SRv6			x
HIP		x	x
VxLAN		x	x
LISP		x	x
SFC		x	x

Table 4: Relationship between Extensions and Internet Addressing

5. A System View on Address

In the following, we investigate in which parts of the overall Internet system extensions have been proposed and developed. For this, we divide the possible innovation across two dimensions:

* Horizontal: Internet edge vs core. The criticality, scale, investment on the core of the Internet makes it more difficult to introduce innovation, while at the edges there is more flexibility. As general purpose processors have drastically improved in performance, data-plane features can be implemented in software. At the edge of the Internet, it easier to introduce innovation for several reasons: Economics, faster ROI because of faster deployment; No need of large scale deployment (and hence less standardization effort); less stakeholders involved (sometimes just one, see following point). Furthermore, the fact that the edge is a place where there is less coordination and cooperation from the core, is another factor that ease the innovation.

* Vertical: at which layer of the protocol stack. The difficulty to innovate varies as well depending at which layer the innovation takes place. One thing is to innovate at application layer where the app developer has large degree of freedom, another is to innovate at network layer, which is more constrained because of its central point in the architecture. Innovation at higher layer sometimes leads to walled gardens (aka limited domains [RFC8799]). Indeed because of the centralization phenomena, an actor offering a certain service may very well develop and deploy a custom technology that does not need to be actually standardized because it is done for its own internal usage.

* Horizontal vs Vertical Innovation:

  - In the public Internet, core innovation at lower layer is harder, often reduced to app-level innovation or building an overlay limited domain (aka a walled garden).

  - At the edges it is easier to innovate at lower layers (more vertical flexibility) but some form of adaptation is needed if global reachability is wanted.

Despite these two orthogonal dimensions, innovation does not happen either horizontally or vertically, rather in both dimensions simultaneously at various degree.

6. Issues in Extensions to Internet Addressing

While the extensions to the original Internet properties, discussed in Section 3, demonstrate the benefits of more flexibility in addressing, they also bring with them a number of issues, which are discussed in the following section. To this end, the problems hereafter outlined link to the approaches to extensions summarized in Section 4. These issues may not be present all the time and everywhere, since as explained in Section 5, extensions are developed and deployed in different part of the Internet, which may worsen things.

6.1. Limiting Address Semantics

Many approaches changing the semantics of communication, e.g., through separating host identification from network node identification [RFC7401], separating the device identifier from the routing locator ([EIBP], [I-D.ietf-lisp-introduction]), or through identifying content and services directly [HICN], are limited by the existing packet size and semantic constraints of IPv6, e.g., in the form of its source and destination network addresses.

While approaches such as [ICNIP] may override the addressing semantics, e.g., by replacing IPv6 source and destination information with path identification, a possible unawareness of endpoints still requires the carrying of other address information as part of the payload.

Also, the expressible service or content semantic may be limited, as in [HICN] or the size of supported networks [REED] due to relying on the limited bit positions usable in IPv6 addresses.

6.2. Complexity and Efficiency

A crucial issue is the additional complexity introduced for realizing the additional addressing semantics. This is particularly an issue since we see those additional semantics particularly at the edge of the Internet, utilizing the existing addressing semantic of the Internet to interconnect the domains that require those additional semantics.

Furthermore, any additional complexity often comes with an efficiency and cost penalty, particularly at the edge of the network, where resource constraints may play a significant role. Compression processes, taking [ROHC] as an example, require additional resources both for the sender generating the compressed header but also the gateway linking to the general Internet by re-establishing the full IP header.

Conversely, the performance requirements of core networks, in terms of packet processing speed, makes the accommodation of extensions to addressing often prohibitive. This is not only due to the necessary extra processing that is specific to the extension, but also due to the complexity that will need to be managed in doing so at significantly higher speeds than at the edge of the network. The observations on the dropping of packets with IPv6 extension headers in the real world is (partially) due to such a implementation complexity [RFC7872].

Another example for lowering the efficiency of packet forwarding is the routing in systems like TOR [TOR]. As detailed before, traffic in TOR, for anonymity purposes, should be handed over by at least three intermediates before reaching the destination. Frequent relaying enhances the privacy, however, because such kind of solutions are implemented at application level, they come at the cost of lower communication efficiency. May be a different privacy enhanced address semantic would enable efficient implementation of TOR-like solutions at network layer.

6.2.1. Repetitive encapsulation

Repetitive encapsulation is an issue since it bloats the packets size due to additional encapsulation headers. Addressing proposals such as those in [ICNIP] utilize path identification within an alternative forwarding architecture that acts upon the provided path identification. However, due to the limitation of existing flow-based architectures with respect to the supported header structures (in the form of IPv4 or IPv6 headers), the new routing semantics are being inserted into the existing header structure, while repeating the original, sender-generated header structure, in the payload of the packet as it traverses the local domain, effectively doubling the per-packet header overhead.

The problem is also present in a number of solutions tackling different issues, e.g., mobility [I-D.ietf-lisp-mn], DC networking ([RFC8926], [RFC7348], [I-D.ietf-intarea-gue]), traffic engineering [RFC8986], and privacy ([TOR], [SPHINX]). Certainly these solutions are able to avoid other issues, like path lengthening or privacy issues, as described before, but they come at the price of multiple encapsulations that reduce the effective payload. This, not only hampers efficiency in terms of header-to-payload ratio, but also introduces 'encapsulation points', which in turn add complexity to the (often edge) network as well as fragility due to the addition of possible failure points; this aspect is discussed in further details in Section 6.4.

6.2.2. Compounding issues with header compression

IP header overhead requires header compression in constrained environments, such as wireless sensor networks and IoT in general. Together with fragmentation, both tasks constitute significant energy consumption, as shown in [HEADER_COMP_ISSUES1], negatively impacting resource limited devices that often rely on battery for operation. Further, the reliance on the compression/decompression points creates a dependence on such gateways, which may be a problem for intermittent scenarios.

According to the implementation of _contiki-ng_ [CONTIKI], an example of operating system for IoT devices, the source codes for 6LowPan requires at least 600Kb to include a header compression process. In certain use cases, such requirement can be an obstacle for extremely constrained devices, especially for the RAM and energy consumption.

6.2.3. Introducing Path Stretch

Mobile IP [RFC6275], which was designed for connection continuity in the face of moving endpoints, is a typical case for path stretch. Since traffic must follow a triangular route before arriving at the destination, such detour routing inevitably impacts transmission efficiency as well as latency.

6.2.4. Complicating Traffic Engineering

While many extensions to the original IP address semantic target to enrich the decisions that can be taken to steer traffic, according to requirements like QoS, mobility, chaining, compute/network metrics, flow treatment, path usage, etc., the realization of the mechanisms as individual solutions likely complicates the original goal of traffic engineering when individual solutions are being used in combination. Ultimately, this may even prevent the combined use of more than one mechanism and/or policy with a need to identify and prevent incompatibilities of mechanisms. Key here is not the issue arising from using conflicting traffic engineering policies, rather conflicting realizations of policies that may well generally work well alongside ([ROBUSTSDN], [TRANSACTIONSDN]).

This not only increases fragility, as discussed separately in Section 6.4, but also requires careful planning of which mechanisms to use and in which combination, likely needing human-in-the-loop approaches alongside possible automation approaches for the individual solutions.

6.3. Security

The properties described in Section 2 have, obviously, also consequences in terms of security and privacy related issues, as already mentioned in other parts of this document.

For instance, in the effort of being somehow backward compatible, HIP [RFC7401] uses a 128-bit Host Identity, which may be not sufficiently cryptographically strong in the future because of the limited size (future computational power may erode 128-bit security). Similarly, CGA [RFC3972] also aligns to the 128-bit limit, but may use only 59 bits of them, hence, the packet signature may not be sufficiently robust to attacks [I-D.rafiee-6man-cga-attack].

IP addresses, even temporary ones meant to protect privacy, have been long recognized as a 'Personal Identification Information' that allows even to geolocate the communicating endpoints [RFC8280]. The use of temporary addresses provides sufficient privacy protection only if the renewal rate is high [EPHEMERALv6]. However, this causes additional issues, like the large overhead due to the Duplicate Address Detection, the impact on the Neighbor Discovery mechanism, in particular the cache, which can even lead to communication disruption. With such drawbacks, the extensions may even lead to defeat the target, actually lowering security rather than increasing it.

The introduction of alternative addressing semantics has also been used to help in (D)DoS attacks mitigation. This leverages on changing the service identification model so to avoid topological information exposure, making the potential disruptions likely remain limited [ADDRLESS]. However, this increased robustness to DDoS comes at the price of important communication setup latency and fragility, as discussed next.

6.4. Fragility

From the extensions discussed in Section 3, it is evident that having alternative or additional address semantic and formats available for making routing as well as forwarding decisions dependent on these, is common place in the Internet. This, however, adds many extension-specific translation/adaptation points, mapping the semantic and format in one context into what is meaningful in another context, but also, more importantly, creating a dependency towards an additional component, often without explicit exposure to the endpoints that originally intended to communicate.

For instance, the re-writing of IP addresses to facilitate the use of private address spaces throughout the public Internet, realized through network address translators (NATs), conflicts with the end-to-end nature of communication between two endpoints. Additional (flow) state is required at the NAT middle-box to smoothly allow communication, which in turn creates a dependency between the NAT and the end-to-end communication between those endpoints, thus increasing the fragility of the communication relation.

A similar situation arises when supporting constrained environments through a header compression mechanism, adding the need for, e.g., a ROHC [RFC5795] element in the communication path, with communication-related compression state being held outside the communicating endpoints. Failure will introduce some inefficiencies due to context regeneration, which may affects the communicating endpoints, increasing fragility of the system overall.

Such translation/adaptation between semantic extensions to the original 'semantic' of an IP address is generally not avoidable when accommodating more than a single universal semantic. However, the solution-specific nature of every single extension is likely to noticeably increase the fragility of the overall system, since individual extensions will need to interact with other extensions that may be deployed in parallel, but were not designed taking into account such deployment scenario (cf., [I-D.ietf-intarea-tunnels]). Considering that extensions to traditional per-hop-behavior (based on IP addresses) can essentially be realized over almost 'any' packet field, the possible number of conflicting behaviors or diverging interpretation of the semantic and/or content of such fields, among different extensions, may soon become an issue, requiring careful testing and delineation at the boundaries of the network within which the specific extension has been realized.

7. Summary of issues

Table 5, derived from Section 6, summarizes the issues related to each extension. While each extension involves at least one issue, some others, like ICNIP, may create several issues at the same time.

	Limiting Address Semantics	Complexity and Efficiency	Security	Fragility
6LoWPAN		x		x
ROHC		x		x

EzIP		x		
TOR		x		x
ODoH		x		
SLAAC		x		
CGA	x		x	
NAT		x		x
HICN	x			
ICNIP	x	x		
CCNx name	x			
EIBP				x
Geo addressing	x			x
REED	x			
DetNet		x		
Mobile IP		x		x
SHIM6				x
SRv6				x
HIP			x	x
VxLAN		x		
LISP		x		x
SFC		x		x

Table 5: Issues in Extensions to Internet Addressing

8. Conclusions

The examples of extensions discussed in Section 3 to the original Internet addressing scheme show that extensibility beyond the original model (and its underlying per-hop behavior) is a desired capability for networking technologies and has been so for a long time. Generally, we can observe that those extensions are driven by the requirements of stakeholders, expecting a desirable extended functionality from the introduction of the specific extension. If interoperability is required, those extensions require standardization of possibly new fields, new semantics as well as (network and/or end system) operations alike.

The issues we identified in this document with the extension-specific solution approach, point to the need for a discussion on Internet addressing, as formulated in the companion document [I-D.jia-intarea-scenarios-problems-addressing] that formalizes the problem statement through scenarios that highlight the shortcomings of the Internet addressing model.

It is our conclusion that the existence of the many extensions to the original Internet addressing is clear evidence for gaps that have been identified over time by the wider Internet community, each of which come with a raft of issues that we need to deal with daily: We believe that it is time to develop an architectural but more importantly a sustainable approach to make Internet addressing extensible in order to capture the many new use cases that will still be identified for the Internet to come.

To jumpstart any such effort from an addressing perspective, it will be key to suitable define what an address is at which layer of the overall system, let alone the network layer. We argue that any answer to this question must be derived from what features we may want from the network instead of being guided by the answers that the Internet can give us today, e.g., being a mere ephemeral token for accessing PoP-based services (as indicated in related arch-d mailing list discussions).

This is not to ’second guess’ the market and its possible evolution, but to outline clear features from which to derive clear principles for a design. Any such design must not skew the technical capabilities of addressing to the current economic situation of the Internet since this bears the danger of locking down innovation capabilities as an outcome of those technical limitations introduced. Instead, addressing must be aligned with enabling the model of permissionless innovation that the IETF has been promoting, ultimately enabling the serendipity of new applications that has led to many of those applications we can see in the Internet today. Most

importantly, any inaction on our side in that regard will only compound the issues identified, eventually hampering the future Internet's readiness for those new uses.

9. Security Considerations

The present memo does not introduce any new technology and/or mechanism and as such does not introduce any security threat to the TCP/IP protocol suite.

As an additional note, and as discussed in this document, security and privacy aspects were not considered as part of the key properties for Internet addressing, which led to the introduction of a number of extensions intending to fix those gaps. The analysis presented in this memo (non-exhaustively) shows those issues are either solved in an ad-hoc manner at application level, or at transport layer, while at network level only few extensions tackling specific aspects exist, albeit often with limitations due to the adherence to the Internet addressing model and its properties.

10. IANA Considerations

This document does not include any IANA request.

11. Informative References

[ADDRLESS] Hao, S., Liu, R., Weng, Z., Chang, D., Bao, C., and X. Li, "Addressless: A new internet server model to prevent network scanning", PLOS ONE Vol. 16, pp. e0246293, DOI 10.1371/journal.pone.0246293, February 2021, <https://doi.org/10.1371/journal.pone.0246293>.

[CLOUDFLARE_SIGCOMM] Fayed, M., Bauer, L., Giotsas, V., Kerola, S., Majkowski, M., Odintsov, P., Sitnicki, J., Chung, T., Levin, D., Mislove, A., Wood, C., and N. Sullivan, "The ties that un-bind: decoupling IP from web services and sockets for robust addressing agility at CDN-scale", Proceedings of the 2021 ACM SIGCOMM 2021 Conference, DOI 10.1145/3452296.3472922, August 2021, <https://doi.org/10.1145/3452296.3472922>.

[CONTIKI] "Contiki-NG: The OS for Next Generation IoT Devices", n.d., <https://github.com/contiki-ng/contiki-ng>.

[EIBP] Shenoy, S Chandraiah, P Willis, N., "A Structured Approach to Routing in the Internet", June 2021, <First Intl Workshop on Semantic Addressing and Routing for Future Networks>.

[EPHEMERALv6] Gont, F. and G. Gont, "IPv6 Addressing Considerations", Work in Progress, Internet-Draft, draft-gont-v6ops-ipv6-addressing-considerations-01, 21 February 2021, <https://www.ietf.org/archive/id/draft-gont-v6ops-ipv6-addressing-considerations-01.txt>.

[EzIP] Chen, A. Y., Ati, R. R., Karandikar, A., and D. R. Crowe, "Adaptive IPv4 Address Space", Work in Progress, Internet-Draft, draft-chen-ati-adaptive-ipv4-address-space-10, 8 December 2021, <https://www.ietf.org/archive/id/draft-chen-ati-adaptive-ipv4-address-space-10.txt>.

[FAYED21] Fayed, M., Bauer, L., Giotsas, V., Kerola, S., Majkowski, M., Odintsov, P., Sitnicki, J., Chung, T., Levin, D., Mislove, A., Wood, C., and N. Sullivan, "The ties that unbind: decoupling IP from web services and sockets for robust addressing agility at CDN-scale", Proceedings of the 2021 ACM SIGCOMM 2021 Conference, DOI 10.1145/3452296.3472922, August 2021, <https://doi.org/10.1145/3452296.3472922>.

[GDPR] Voigt, P. and A. von dem Bussche, "The EU General Data Protection Regulation (GDPR)", Springer International Publishing book, DOI 10.1007/978-3-319-57959-7, 2017, <https://doi.org/10.1007/978-3-319-57959-7>.

[GNATCATCHER] "Global Network Address Translation Combined with Audited and Trusted CDN or HTTP-Proxy Eliminating Reidentification", n.d., <https://github.com/bslassey/ip-blindness>.

[HEADER_COMP_ISSUES1] Mesrinejad, F., Hashim, F., Noordin, N., Rasid, M., and R. Abdullah, "The effect of fragmentation and header compression on IP-based sensor networks (6LoWPAN)", The 17th Asia Pacific Conference on Communications, DOI 10.1109/apcc.2011.6152926, October 2011, <https://doi.org/10.1109/apcc.2011.6152926>.

[HICN] Muscariello, L., "Hybrid Information-Centric Networking: ICN inside the Internet Protocol", March 2018, <https://datatracker.ietf.org/meeting/interim-2018-icnrg-01/materials/slides-interim-2018-icnrg-01-sessa-hybrid-icn-hicn-luca-muscariello>.

[HISTORY127] "History of 127/8 as localhost/loopback addresses", n.d., <https://elists.isoc.org/pipermail/internet-history/2021-January/006920.html>.

[I-D.ietf-6lo-nfc] Choi, Y., Hong, Y., Youn, J., Kim, D., and J. Choi, "Transmission of IPv6 Packets over Near Field Communication", Work in Progress, Internet-Draft, draft-ietf-6lo-nfc-17, 23 August 2020, <https://www.ietf.org/archive/id/draft-ietf-6lo-nfc-17.txt>.

[I-D.ietf-6lo-plc] Hou, J., Liu, B., Hong, Y., Tang, X., and C. E. Perkins, "Transmission of IPv6 Packets over PLC Networks", Work in Progress, Internet-Draft, draft-ietf-6lo-plc-10, 17 February 2022, <https://www.ietf.org/archive/id/draft-ietf-6lo-plc-10.txt>.

[I-D.ietf-intarea-gue] Herbert, T., Yong, L., and O. Zia, "Generic UDP Encapsulation", Work in Progress, Internet-Draft, draft-ietf-intarea-gue-09, 26 October 2019, <https://www.ietf.org/archive/id/draft-ietf-intarea-gue-09.txt>.

[I-D.ietf-intarea-tunnels] Touch, J. and M. Townsley, "IP Tunnels in the Internet Architecture", Work in Progress, Internet-Draft, draft-ietf-intarea-tunnels-10, 12 September 2019, <https://www.ietf.org/archive/id/draft-ietf-intarea-tunnels-10.txt>.

[I-D.ietf-lisp-introduction] Cabellos, A. and D. S. (Ed.), "An Architectural Introduction to the Locator/ID Separation Protocol (LISP)", Work in Progress, Internet-Draft, draft-ietf-lisp-introduction-15, 20 September 2021, <https://www.ietf.org/archive/id/draft-ietf-lisp-introduction-15.txt>.

[I-D.ietf-lisp-mn]
Farinacci, D., Lewis, D., Meyer, D., and C. White, "LISP Mobile Node", Work in Progress, Internet-Draft, draft-ietf-lisp-mn-11, 30 January 2022, <https://www.ietf.org/archive/id/draft-ietf-lisp-mn-11.txt>.

[I-D.ietf-lisp-rfc6830bis]
Farinacci, D., Fuller, V., Meyer, D., Lewis, D., and A. Cabellos, "The Locator/ID Separation Protocol (LISP)", Work in Progress, Internet-Draft, draft-ietf-lisp-rfc6830bis-36, 18 November 2020, <https://www.ietf.org/archive/id/draft-ietf-lisp-rfc6830bis-36.txt>.

[I-D.ietf-lisp-rfc6833bis]
Farinacci, D., Maino, F., Fuller, V., and A. Cabellos, "Locator/ID Separation Protocol (LISP) Control-Plane", Work in Progress, Internet-Draft, draft-ietf-lisp-rfc6833bis-30, 18 November 2020, <https://www.ietf.org/archive/id/draft-ietf-lisp-rfc6833bis-30.txt>.

[I-D.jia-intarea-scenarios-problems-addressing]
Jia, Y., Trossen, D., Iannone, L., Shenoy, N., Mendes, P., 3rd, D. E. E., and P. Liu, "Challenging Scenarios and Problems in Internet Addressing", Work in Progress, Internet-Draft, draft-jia-intarea-scenarios-problems-addressing-02, 23 October 2021, <https://www.ietf.org/archive/id/draft-jia-intarea-scenarios-problems-addressing-02.txt>.

[I-D.rafiee-6man-cga-attack]
Rafiee, H. and C. Meinel, "Possible Attack on Cryptographically Generated Addresses (CGA)", Work in Progress, Internet-Draft, draft-rafiee-6man-cga-attack-03, 8 May 2015, <https://www.ietf.org/archive/id/draft-rafiee-6man-cga-attack-03.txt>.

[ICNIP]
Trossen, D., Robitzsch, S., Reed, M., Al-Naday, M., and J. Riihijarvi, "Internet Services over ICN in 5G LAN Environments", Work in Progress, Internet-Draft, draft-trossen-icnrg-internet-icn-5glan-04, 1 October 2020, <https://www.ietf.org/archive/id/draft-trossen-icnrg-internet-icn-5glan-04.txt>.

[IPv4pool] "IANA IPv4 Address Space Registry", n.d., <https://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xhtml>.

[ITU9959] Badenhop, C., Fuller, J., Hall, J., Ramsey, B., and M. Rice, "Evaluating ITU-T G.9959 Based Wireless Systems Used in Critical Infrastructure Assets", IFIP Advances in Information and Communication Technology pp. 209-227, DOI 10.1007/978-3-319-26567-4_13, 2015, <https://doi.org/10.1007/978-3-319-26567-4_13>.

[ODoH] Kinnear, E., McManus, P., Pauly, T., Verma, T., and C. A. Wood, "Oblivious DNS Over HTTPS", Work in Progress, Internet-Draft, draft-pauly-dprive-oblivious-doh-11, 17 February 2022, <https://www.ietf.org/archive/id/draft-pauly-dprive-oblivious-doh-11.txt>.

[OHTTP] Thomson, M. and C. A. Wood, "Oblivious HTTP", Work in Progress, Internet-Draft, draft-thomson-http-oblivious-02, 24 August 2021, <https://www.ietf.org/archive/id/draft-thomson-http-oblivious-02.txt>.

[ONION] Goldschlag, D., Reed, M., and P. Syverson, "Onion routing", Communications of the ACM Vol. 42, pp. 39-41, DOI 10.1145/293411.293443, February 1999, <https://doi.org/10.1145/293411.293443>.

[P4] Bosshart, P., Daly, D., Gibb, G., Izzard, M., McKeown, N., Rexford, J., Schlesinger, C., Talayco, D., Vahdat, A., Varghese, G., and D. Walker, "P4: programming protocol-independent packet processors", ACM SIGCOMM Computer Communication Review Vol. 44, pp. 87-95, DOI 10.1145/2656877.2656890, July 2014, <https://doi.org/10.1145/2656877.2656890>.

[REED] Reed, M., Al-Naday, M., Thomos, N., Trossen, D., Petropoulos, G., and S. Spirou, "Stateless multicast switching in software defined networks", 2016 IEEE International Conference on Communications (ICC), DOI 10.1109/icc.2016.7511036, May 2016, <https://doi.org/10.1109/icc.2016.7511036>.

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <https://www.rfc-editor.org/info/rfc791>.

[RFC1752] Bradner, S. and A. Mankin, "The Recommendation for the IP Next Generation Protocol", RFC 1752, DOI 10.17487/RFC1752, January 1995, <https://www.rfc-editor.org/info/rfc1752>.

[RFC1918] Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G. J., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, DOI 10.17487/RFC1918, February 1996, <https://www.rfc-editor.org/info/rfc1918>.

[RFC2009] Imielinski, T. and J. Navas, "GPS-Based Addressing and Routing", RFC 2009, DOI 10.17487/RFC2009, November 1996, <https://www.rfc-editor.org/info/rfc2009>.

[RFC2101] Carpenter, B., Crowcroft, J., and Y. Rekhter, "IPv4 Address Behaviour Today", RFC 2101, DOI 10.17487/RFC2101, February 1997, <https://www.rfc-editor.org/info/rfc2101>.

[RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, DOI 10.17487/RFC2474, December 1998, <https://www.rfc-editor.org/info/rfc2474>.

[RFC2663] Srisuresh, P. and M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations", RFC 2663, DOI 10.17487/RFC2663, August 1999, <https://www.rfc-editor.org/info/rfc2663>.

[RFC2865] Rigney, C., Willens, S., Rubens, A., and W. Simpson, "Remote Authentication Dial In User Service (RADIUS)", RFC 2865, DOI 10.17487/RFC2865, June 2000, <https://www.rfc-editor.org/info/rfc2865>.

[RFC3118] Droms, R., Ed. and W. Arbaugh, Ed., "Authentication for DHCP Messages", RFC 3118, DOI 10.17487/RFC3118, June 2001, <https://www.rfc-editor.org/info/rfc3118>.

[RFC3972] Aura, T., "Cryptographically Generated Addresses (CGA)", RFC 3972, DOI 10.17487/RFC3972, March 2005, <https://www.rfc-editor.org/info/rfc3972>.

[RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, DOI 10.17487/RFC3986, January 2005, <https://www.rfc-editor.org/info/rfc3986>.

[RFC4014] Droms, R. and J. Schnizlein, "Remote Authentication Dial-In User Service (RADIUS) Attributes Suboption for the Dynamic Host Configuration Protocol (DHCP) Relay Agent Information Option", RFC 4014, DOI 10.17487/RFC4014, February 2005, <https://www.rfc-editor.org/info/rfc4014>.

[RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, DOI 10.17487/RFC4193, October 2005, <https://www.rfc-editor.org/info/rfc4193>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <https://www.rfc-editor.org/info/rfc4291>.

[RFC4581] Bagnulo, M. and J. Arkko, "Cryptographically Generated Addresses (CGA) Extension Field Format", RFC 4581, DOI 10.17487/RFC4581, October 2006, <https://www.rfc-editor.org/info/rfc4581>.

[RFC4982] Bagnulo, M. and J. Arkko, "Support for Multiple Hash Algorithms in Cryptographically Generated Addresses (CGAs)", RFC 4982, DOI 10.17487/RFC4982, July 2007, <https://www.rfc-editor.org/info/rfc4982>.

[RFC5533] Nordmark, E. and M. Bagnulo, "Shim6: Level 3 Multihoming Shim Protocol for IPv6", RFC 5533, DOI 10.17487/RFC5533, June 2009, <https://www.rfc-editor.org/info/rfc5533>.

[RFC5795] Sandlund, K., Pelletier, G., and L-E. Jonsson, "The RObust Header Compression (ROHC) Framework", RFC 5795, DOI 10.17487/RFC5795, March 2010, <https://www.rfc-editor.org/info/rfc5795>.

[RFC6250] Thaler, D., "Evolution of the IP Model", RFC 6250, DOI 10.17487/RFC6250, May 2011, <https://www.rfc-editor.org/info/rfc6250>.

[RFC6275] Perkins, C., Ed., Johnson, D., and J. Arkko, "Mobility Support in IPv6", RFC 6275, DOI 10.17487/RFC6275, July 2011, <https://www.rfc-editor.org/info/rfc6275>.

[RFC6282] Hui, J., Ed. and P. Thubert, "Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks", RFC 6282, DOI 10.17487/RFC6282, September 2011, <https://www.rfc-editor.org/info/rfc6282>.

[RFC6740] Atkinson, RJ. and SN. Bhatti, "Identifier-Locator Network Protocol (ILNP) Architectural Description", RFC 6740, DOI 10.17487/RFC6740, November 2012, <https://www.rfc-editor.org/info/rfc6740>.

[RFC7228] Bormann, C., Ersue, M., and A. Keranen, "Terminology for Constrained-Node Networks", RFC 7228, DOI 10.17487/RFC7228, May 2014, <https://www.rfc-editor.org/info/rfc7228>.

[RFC7258] Farrell, S. and H. Tschofenig, "Pervasive Monitoring Is an Attack", BCP 188, RFC 7258, DOI 10.17487/RFC7258, May 2014, <https://www.rfc-editor.org/info/rfc7258>.

[RFC7348] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", RFC 7348, DOI 10.17487/RFC7348, August 2014, <https://www.rfc-editor.org/info/rfc7348>.

[RFC7400] Bormann, C., "6LoWPAN-GHC: Generic Header Compression for IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs)", RFC 7400, DOI 10.17487/RFC7400, November 2014, <https://www.rfc-editor.org/info/rfc7400>.

[RFC7401] Moskowitz, R., Ed., Heer, T., Jokela, P., and T. Henderson, "Host Identity Protocol Version 2 (HIPv2)", RFC 7401, DOI 10.17487/RFC7401, April 2015, <https://www.rfc-editor.org/info/rfc7401>.

[RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function Chaining (SFC) Architecture", RFC 7665, DOI 10.17487/RFC7665, October 2015, <https://www.rfc-editor.org/info/rfc7665>.

[RFC7872] Gont, F., Linkova, J., Chown, T., and W. Liu, "Observations on the Dropping of Packets with IPv6 Extension Headers in the Real World", RFC 7872, DOI 10.17487/RFC7872, June 2016, <https://www.rfc-editor.org/info/rfc7872>.

[RFC8060] Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical Address Format (LCAF)", RFC 8060, DOI 10.17487/RFC8060, February 2017, <https://www.rfc-editor.org/info/rfc8060>.

[RFC8061] Farinacci, D. and B. Weis, "Locator/ID Separation Protocol (LISP) Data-Plane Confidentiality", RFC 8061, DOI 10.17487/RFC8061, February 2017, <https://www.rfc-editor.org/info/rfc8061>.

[RFC8105] Mariager, P., Petersen, J., Ed., Shelby, Z., Van de Logt, M., and D. Barthel, "Transmission of IPv6 Packets over Digital Enhanced Cordless Telecommunications (DECT) Ultra Low Energy (ULE)", RFC 8105, DOI 10.17487/RFC8105, May 2017, <https://www.rfc-editor.org/info/rfc8105>.

[RFC8163] Lynn, K., Ed., Martocci, J., Neilson, C., and S. Donaldson, "Transmission of IPv6 over Master-Slave/Token-Passing (MS/TP) Networks", RFC 8163, DOI 10.17487/RFC8163, May 2017, <https://www.rfc-editor.org/info/rfc8163>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[RFC8280] ten Oever, N. and C. Cath, "Research into Human Rights Protocol Considerations", RFC 8280, DOI 10.17487/RFC8280, October 2017, <https://www.rfc-editor.org/info/rfc8280>.

[RFC8376] Farrell, S., Ed., "Low-Power Wide Area Network (LPWAN) Overview", RFC 8376, DOI 10.17487/RFC8376, May 2018, <https://www.rfc-editor.org/info/rfc8376>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[RFC8484] Hoffman, P. and P. McManus, "DNS Queries over HTTPS (DoH)", RFC 8484, DOI 10.17487/RFC8484, October 2018, <https://www.rfc-editor.org/info/rfc8484>.

[RFC8609] Mosko, M., Solis, I., and C. Wood, "Content-Centric Networking (CCNx) Messages in TLV Format", RFC 8609, DOI 10.17487/RFC8609, July 2019, <https://www.rfc-editor.org/info/rfc8609>.

[RFC8724] Minaburo, A., Toutain, L., Gomez, C., Barthel, D., and JC. Zuniga, "SCHC: Generic Framework for Static Context Header Compression and Fragmentation", RFC 8724, DOI 10.17487/RFC8724, April 2020, <https://www.rfc-editor.org/info/rfc8724>.

[RFC8799] Carpenter, B. and B. Liu, "Limited Domains and Internet Protocols", RFC 8799, DOI 10.17487/RFC8799, July 2020, <https://www.rfc-editor.org/info/rfc8799>.

[RFC8926] Gross, J., Ed., Ganga, I., Ed., and T. Sridhar, Ed., "Geneve: Generic Network Virtualization Encapsulation", RFC 8926, DOI 10.17487/RFC8926, November 2020, <https://www.rfc-editor.org/info/rfc8926>.

[RFC8939] Varga, B., Ed., Farkas, J., Berger, L., Fedyk, D., and S. Bryant, "Deterministic Networking (DetNet) Data Plane: IP", RFC 8939, DOI 10.17487/RFC8939, November 2020, <https://www.rfc-editor.org/info/rfc8939>.

[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/info/rfc8981>.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., and Z. Li, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

[ROBUSTSDN] Canini, M., Kuznetsov, P., Levin, D., and S. Schmid, "A distributed and robust SDN control plane for transactional network updates", 2015 IEEE Conference on Computer Communications (INFOCOM), DOI 10.1109/infocom.2015.7218382, April 2015, <https://doi.org/10.1109/infocom.2015.7218382>.

[ROHC] Fitzek, F., Rein, S., Seeling, P., and M. Reisslein, "RObust Header Compression (ROHC) Performance for Multimedia Transmission over 3G/4G Wireless Networks", Wireless Personal Communications Vol. 32, pp. 23-41, DOI 10.1007/s11277-005-7733-2, January 2005, <https://doi.org/10.1007/s11277-005-7733-2>.

[SFCANYCAST] Wion, A., Bouet, M., Iannone, L., and V. Conan, "Distributed Function Chaining with Anycast Routing", Proceedings of the 2019 ACM Symposium on SDN Research, DOI 10.1145/3314148.3314355, April 2019, <https://doi.org/10.1145/3314148.3314355>.

[SPHINX] Danezis, G. and I. Goldberg, "Sphinx: A Compact and Provably Secure Mix Format", 2009 30th IEEE Symposium on Security and Privacy, DOI 10.1109/sp.2009.15, May 2009, <https://doi.org/10.1109/sp.2009.15>.

[TOR] "The Tor Project", n.d., <https://www.torproject.org/>.

[TRANSACTIONSDN] Curic, M., Despotovic, Z., Hecker, A., and G. Carle, "Transactional Network Updates in SDN", 2018 European Conference on Networks and Communications (EuCNC), DOI 10.1109/eucnc.2018.8442793, June 2018, <https://doi.org/10.1109/eucnc.2018.8442793>.

[UA-DHCP] Komori, T. and T. Saito, "The secure DHCP system with user authentication", 27th Annual IEEE Conference on Local Computer Networks, 2002. Proceedings. LCN 2002., DOI 10.1109/lcn.2002.1181774, n.d., <https://doi.org/10.1109/lcn.2002.1181774>.

[VPN] Khanvilkar, S. and A. Khokhar, "Virtual private networks: an overview with performance evaluation", IEEE Communications Magazine Vol. 42, pp. 146-154, DOI 10.1109/mcom.2004.1341273, October 2004, <https://doi.org/10.1109/mcom.2004.1341273>.

[WireGuard] Donenfeld, J., "WireGuard: Next Generation Kernel Network Tunnel", Proceedings 2017 Network and Distributed System Security Symposium, DOI 10.14722/ndss.2017.23160, 2017, <https://doi.org/10.14722/ndss.2017.23160>.

Acknowledgments

Thanks to all the people that shared insightful comments both privately to the authors as well as on various mailing list, especially on the INTArea Mailing List. Also thanks for the interesting discussions to Carsten Borman, Brian E. Carpenter.

Authors' Addresses

Yihao Jia
Huawei Technologies Co., Ltd
156 Beiqing Rd.
Beijing
100095
P.R. China
Email: jiayihao@huawei.com

Dirk Trossen
Huawei Technologies Duesseldorf GmbH
Riesstr. 25C
80992 Munich
Germany
Email: dirk.trossen@huawei.com

Luigi Iannone
Huawei Technologies France S.A.S.U.
18, Quai du Point du Jour
92100 Boulogne-Billancourt
France
Email: luigi.iannone@huawei.com

Paulo Mendes
Airbus
Willy-Messerschmitt Strasse 1
81663 Munich
Germany
Email: paulo.mendes@airbus.com

Nirmala Shenoy
Rochester Institute of Technology
New-York, 14623
United States of America
Email: nxsvks@rit.edu

Laurent Toutain
IMT-Atlantique
2 rue de la Chataigneraie
CS 17607
35576 Cesson-Sevigne Cedex
France
Email: laurent.toutain@imt-atlantique.fr

Abraham Y. Chen
Avinta Communications, Inc.
142 N. Milpitas Blvd.
Milpitas, CA, 95035-4401
United States of America
Email: AYChen@Avinta.com

Dino Farinacci
lispers.net
United States of America
Email: farinacci@gmail.com

Internet Area Working Group
Internet-Draft
Intended status: Informational
Expires: 7 September 2022

Y. Jia
D. Trossen
L. Iannone
Huawei
N. Shenoy
R.I.T.
P. Mendes
Airbus
D. Eastlake 3rd
Futurewei
P. Liu
China Mobile
D. Farinacci
lispers.net
6 March 2022

Challenging Scenarios and Problems in Internet Addressing
draft-jia-intarea-scenarios-problems-addressing-03

Abstract

The Internet Protocol (IP) has been the major technological success in information technology of the last half century. As the Internet becomes pervasive, IP has been replacing communication technology for many domain-specific solutions. However, domains with specific requirements as well as communication behaviors and semantics still exist and represent what [RFC8799] recognizes as "limited domains".

This document describes well-recognized scenarios that showcase possibly different addressing requirements, which are challenging to be accommodated in the IP addressing model. These scenarios highlight issues related to the Internet addressing model and call for starting a discussion on a way to re-think/evolve the addressing model so to better accommodate different domain-specific requirements.

The issues identified in this document are complemented and deepened by a detailed gap analysis in a separate companion document [I-D.jia-intarea-internet-addressing-gap-analysis].

Status of This Memo

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

Copyright Notice

Table of Contents

1. Introduction

The Internet Protocol (IP), positioned as the unified protocol at the (Internet) network layer, is seen by many as key to the innovation stemming from Internet-based applications and services. Even more so, with the success of TCP/IP protocol stack, IP has been gradually replacing existing domain-specific protocols, evolving into the core protocol of the entire communication eco-system. At its inception, roughly 40 years ago [RFC0791], the Internet addressing system, represented in the form of the IP address and its locator-based (topological) semantics, has brought the notion of a 'common namespace for all communication'. Compared to proprietary technology-specific solutions, such 'common namespace for all communication' advance ensures end-to-end communication from any device connected to the Internet to another.

However, use cases, associated services, node behaviors, and requirements on packet delivery have since been significantly extended, with the Internet technology being developed to accommodate them in the framework of addressing that stood at the beginning of the Internet's development. This evolution is reflected in the concept of "Limited Domains", first introduced in [RFC8799]. It refers to a single physical network, attached to or running in parallel with the Internet, or is defined by a set of users and nodes distributed over a much wider area, but drawn together by a single virtual network over the Internet. Key to a limited domain is that requirements, behaviors, and semantics could be noticeable local and, more importantly, specific to the limited domain. Very often, the realization of a limited domain is defined by specific communication scenario(s) and/or use case(s) that exhibit the domain-specific behaviors and pose the requirements that lead to the establishment of the limited domain. Identifying limited domains may sometime be not obvious because of blurry boundaries depending on the point of view. For instance, from an end user perspective there is no vision at all on limited domains, hence for end users the dichotomy Internet vs limited domains more transparent. In such cases, it is harder to ensure (and detect) that no limited domain specific semantics leak in the Internet or other limited domains.

One key architectural aspect, when communicating within limited domains, is that of addressing and, therefore, the address structure, as well as the semantic that is being used for packet forwarding (e.g., service identification, content location, device type). The topological location centrality of IP is fundamental when reconciling the often differing semantics for 'addressing' that can be found in those limited domains. The result of this fundamental role of the single IP addressing is that limited domains have to adopt specific solutions, e.g., translating/mapping/converting concepts, semantics, and ultimately, domain-specific addressing, into the common IP addressing used across limited domains.

This document advocates flexibility in addressing in order to accommodate limited domain specific semantics, while, if possible, ensuring a single holistic addressing scheme able to reduce, or even entirely remove, the need for aligning the address semantics of different limited domains, such as the current topological location semantic of the Internet. Ultimately, such holistic addressing could be beneficial to those communication scenarios realized within limited domains by improving efficiency, removing of constraints imposed by needing to utilize the limited semantics of IP addressing, and/or in other ways.

In other words, this document revolves around the following question:

> "Should interconnected limited domains purely rely on IP addresses and therefore deal with the complexity of translating any semantic mismatch themselves, or should flexibility for supporting those limited domains be a key focus for an evolved Internet addressing?"

To that end, this document describes well-recognized scenarios in limited domains that could benefit from greater flexibility in addressing and overviews the problems encountered throughout these scenarios due to the lack of that flexibility. A detailed gap analysis can be found in {I-D.jia-intarea-internet-addressing-gap-analysis}}, which elaborates on the issues identified in this memo in reference to extensions to Internet addressing that have attempted to address those issues. The purpose of this memo is rather to stimulate discussion on the emerging needs for addressing at large with the possibility to fundamentally re-think the addressing in the Internet beyond the current objectives of IPv6 [RFC8200].

It is important to remark that any change in the addressing, hence at the data plane level, leads to changes and challenges at the control plane level, i.e., routing. The latter is an even harder problem than just addressing and might need more research efforts that are beyond the objective of this document, which focuses solely on the data plane.

2. Communication Scenarios in Limited Domains

The following sub-sections outline a number of scenarios, all of which belong to the concept of "limited domains" [RFC8799]. While the list of scenarios may look long, this document focuses on scenarios with a number of aspects that can be observed in those limited domains, captured in the sub-section titles. For each scenario, possible challenges are highlighted, which are then picked upon in Section 4, when describing more formally the existing shortcomings in current Internet addressing.

2.1. Communication in Constrained Environments

In a number of communication scenarios, such as those encountered in the Internet of Things (IoT), a simple, communication network demanding minimal resources is required, allowing for a group of IoT network devices to form a network of constrained nodes, with the participating network and end nodes requiring as little computational power as possible and having small memory requirements in order to reduce the total cost of ownership of the network. Furthermore, in the context of industrial IoT, real-time requirements and scalability make IP technology not naturally suitable as communication technology ([OCADO]).

In addition to IEEE 802.15.4, i.e., Low-Rate Wireless Personal Area Network [LR-WPAN], several limited domains exist through utilizing link layer technologies such as Bluetooth Low Energy (BLE) [BLE], Digital European Cordless Telecommunications (DECT) - Ultra Low Energy (ULE) [DECT-ULE], Master-Slave/Token-Passing (MS/TP) [BACnet], Near-Field-Communication (NFC) [ECMA-340], and Power Line Communication (PLC) [IEEE_1901.1].

The end-to-end principle (detailed in [RFC2775]) requires IP addresses (e.g., IPv6 [RFC8200]) to be used on such constrained nodes networks, allowing IoT devices using multiple communication technologies to talk on the Internet. Often, devices located at the edge of constrained networks act as gateway devices, usually performing header compression ([RFC4919]). To ensure security and reliability, multiple gateways must be deployed. IoT devices on the network must select one of those gateways for traffic passthrough by the devices on the (limited domain) network.

Given the constraints imposed on the computational and possibly also communication technology, the usage of a single addressing semantic in the form of a 128-bit endpoint identifier, i.e., IPv6 address, may pose a challenge when operating such networks.

Another type of (differently) constrained environment is an aircraft, which encompasses not only passenger communication but also the integration of real-time data exchange to ensure that processes and functions in the cabin are automatically monitored or actuated. The goal for any aircraft network is to be able to send and receive information reliably and seamlessly. From this perspective, the medium with which these packets of information are sent is of little consequence so long as there is a level of determinism to it. However, there is currently no effective method in implementing wireless inter- and intra-communications between all subsystems. The emerging wireless sensor network technology in commercial applications such as smart thermostat systems, and smart washer/dryer units could be transposed onto aircraft and fleet operations. The proposal for having an Wireless Avionics Intra-Communications (WAIC) system promises reduction in the complexity of electrical wiring harness design and fabrication, reduction in wiring weight, increased configuration, and potential monitoring of otherwise inaccessible moving or rotating aircraft parts. Similar to the IoT concept, WAIC systems consist of short-range communications and are a potential candidate for passenger entertainment systems, smoke detectors, engine health monitors, tire pressure monitoring systems, and other kinds of aircraft maintenance systems.

While there are still many obstacles in terms of network security, traffic control, and technical challenges, future WAIC can enable real-time seamless communications between aircraft and between ground teams and aircraft as opposed to the discrete points of data leveraged today in aircraft communications. For that, WAIC infrastructure should also be connected to outside IP based networks in order to access edge/cloud facilities for data storage and mining. However, the restricted capacity (energy, communication) of most aircraft devices (e.g. sensors) and the nature of the transmitted data - periodic transmission of small packets - may pose some challenges for the usage of a single addressing semantic in the form of a 128-bit endpoint identifier, i.e., an IPv6 address. Moreover, most of the aircraft applications and services are focused on the data (e.g. temperature of gas tank on left wing) and not on the topological location of the data source. This means that the current topological location semantic of IP addresses is not beneficial for aircraft applications and services.

Greater flexibility in Internet addressing may avoid complex and energy hungry operations, like header compression and fragmentation, necessary to translate protocol headers from one limited domain to another, while enabling semantics different from locator-based addressing may better support the communication that occurs in those environments.

2.2. Communication within Dynamically Changing Topologies

Communication may occur over networks that exhibit dynamically changing topologies. One such example is that of satellite networks, providing global Internet connections through a combination of inter-satellite and ground station communication. With the convergence of space-based and terrestrial networks, users can experience seamless broadband access, e.g., on cruise ships, flights, and within cars, often complemented by and seamlessly switching between Wi-Fi, cellular, or satellite based networks at any time [WANG19].

The satellite network service provider will plan the transmission path of user traffic based on the network coverage, satellite orbit, route, and link load, providing potentially high-quality Internet connections for users in areas that are not, or hard to be, covered by terrestrial networks. With large scale LEO (Low Earth Orbit) satellites, the involved topologies of the satellite network will be changing constantly while observing a regular flight pattern in relation to other satellites and predictable overflight patterns to ground users [CHEN21].

Although satellite bearer services are capable of transporting IPv4 and IPv6, as well as associated protocols such as IP Multicast, DNS services and routing information, no IP functionality is implemented on-board the spacecraft limiting the capability of leveraging for instance large scale satellite constellations.

One of the major constraints of deploying routing capability on board of a satellite is power consumption. Due to this, space routers may end up being intermittently powered up during a daytime sunlit pass. Another limitation of the first generation of IP routers in space was the lack of capability to remotely manage and upgrade software while in operation.

The limitations faced in early development of IP based satellite communication payloads, showed the need to develop a flexible networking solution that would enable delay tolerant communications in the presence of intermittent connectivity. Further, in order to reduce latency, which is the major impairment of satellite networks, there was a need of a networking solution able to perform in a scenario encompassing mobile devices with the capability of storing data, leading to a significant reduction of latency, which is the major impairment of satellite networks.

Moreover, due to the current IP addressing scheme and its focus on IP unicast addressing with extended deployment of IP multicast and some IP anycast, current deployments do not take advantage of the broadcast nature of satellite networks.

Moreover networking platforms based on a name (data or service) based addressing scheme would bring several potential benefits to satellite networks aiming to tackle their major challenges, including high propagation delay and changing network topology in the case of LEO constellations.

Another example is that of vehicular communication, where services may be accessed across vehicles, such as self-driving cars, for the purpose of collaborative objection recognition (e.g., for collision avoidance), road status conveyance (e.g., for pre-warning of road-ahead conditions), and other purposes. Communication may include Road Side Units (RSU) with the possibility to create ephemeral connections to those RSUs for the purpose of workload offloading, joint computation over multiple (vehicular) inputs, and other purposes [I-D.ietf-lisp-nexagon]. Communication here may exhibit a multi-hop nature, not just involving the vehicle and the RSU over a direct link. Those topologies are naturally changing constantly due to the dynamic nature of the involved communication nodes.

The advent of Flying Ad-hoc NETworks (FANETs) has opened up an opportunity to create new added-value services [CHRIKI19]. Although these networks share common features with vehicular ad hoc networks, they present several unique characteristics such as energy efficiency, mobility degree, the capability of swarming, and the potential large scale of swarm networks. Due to high mobility of FANET nodes, the network topology changes more frequently than in a typical vehicular ad hoc network. From a routing point of view, although ad-hoc reactive and proactive routing approaches can be used, there are other type of routing protocols that have been developed for FANETS, such as hybrid routing protocols and position based routing protocols, aiming to increase efficiency in large scale networks with dynamic topologies.

Both type of protocols challenge the current Internet addressing semantic: in the case of hybrid protocols, two different routing strategies are used inside and outside a network zone. While inside a zone packets are routed to a specific destination IP address, between zones, query packets are routed to a subset of neighbors as determined by a broadcast algorithm. In the case of position based routing protocol, the IP addressing scheme is not used at all, since packets are routed to a different identifier, corresponding to the geographic location of the destination and not its topological location. Hence, what is needed is to consolidate the geo-spatial addressing with that of a locator-based addressing in order to optimize routing policies across the zones.

Moreover most of the application/services deployed in FANETs tend to be agnostic of the topological location of nodes, rather focusing on the location of data or services. This distinction is even more important because is dynamic network such as FANET robust networking solutions may rely on the redundancy of data and services, meaning that they may be found in more than one device in the network. This in turn may bring into play a possible service-centric semantic for addressing the packets that need routing in the dynamic network towards a node providing said service (or content).

In the aforementioned network technologies, there is a significant difference between the high dynamics of the underlying network topologies, compared to the relative static nature of terrestrial network topology, as reported in [HANDLEY]. As a consequence, the notion of a topological network location becomes restrictive in the sense that not only the relation between network nodes and user endpoint may change, but also the relation between the nodes that form the network itself. This may lead to the challenge of maintaining and updating the topological addresses in this constantly changing network topology.

In attempts to utilize entirely different semantics for the addressing itself, geographic-based routing, such as in [CARTISEAN], has been proposed for MANETs (Mobile Ad-hoc NETworks) through providing geographic coordinates based addresses to achieve better routing performance, lower overhead, and lower latency [MANET1].

Flexibility in Internet addressing here would allow for accommodating such geographic address semantics into the overall Internet addressing, while also enabling name/content-based addressing, utilizing the redundancy of many network locations providing the possible data.

2.3. Communication among Moving Endpoints

When packet switching was first introduced, back in the 60s/70s, it was intended to replace the rigid circuit switching with a communication infrastructure that was more resilient to failures. As such, the design never really considered communication endpoints as mobile. Even in the pioneering ALOHA [ALOHA] system, despite considering wireless and satellite links, the network was considered static (with the exception of failures and satellites, which fall in what is discussed in Section 2.2). Ever since, a lot of efforts have been devoted to overcome such limitations once it became clear that endpoint mobility will become a main (if not THE main) characteristic of ubiquitous communication systems.

The IETF has for a long time worked on solutions that would allow extending the IP layer with mobility support. Because of the topological semantic of IP addresses, endpoints need to change addresses each time they visit a different network. However, because routing and endpoint identification is also IP address based, this leads to a communication disruption.

To cope with such a situation, sometimes, the transport layer gets involved in mobility solutions, either by introducing explicit in-band signaling to allow for communicating IP address changes (e.g., in SCTP [RFC5061] and MPTCP [RFC6182]), or by introducing some form of connection ID that allows for identifying a communication independently from IP addresses (e.g., the connection ID used in QUIC [RFC9000]).

Concerning network layer only solutions, anchor-based Mobile IP mechanisms have been introduced ([RFC5177], [RFC6626] [RFC5944], [RFC5275]). Mobile IP is based on a relatively complex and heavy mechanism that makes it hard to deploy and it is not very efficient. Furthermore, it is even less suitable than native IP in constrained environments like the ones discussed in Section 2.1.

Alternative approaches to Mobile IP often leverage the introduction of some form of overlay. LISP [I-D.ietf-lisp-introduction], by separating the topological semantic from the identification semantic of IP addresses, is able to cope with endpoint mobility by dynamically mapping endpoint identifiers with routing locators [I-D.ietf-lisp-mn]. This comes at the price of an overlay that needs its own additional control plane [I-D.ietf-lisp-rfc6833bis].

Similarly, the NVO3 (Network Virtualization Overlays) Working Group, while focusing on Data Center environments, also explored an overlay-based solution for multi-tenancy purposes, but also resilient to mobility since relocating Virtual Machines (VMs) is common practice.

NVO3 considered for a long time several data planes that implement slightly different flavors of overlays ([RFC8926], [RFC7348], [I-D.ietf-intarea-gue]), but lacks an efficient control plane specifically tailored for DCs.

Alternative mobility architectures have also been proposed in order to cope with endpoint mobility outside the IP layer itself. The Host Identity Protocol (HIP) [RFC7401] introduced a new namespace in order to identify endpoints, namely the Host Identity (HI), while leveraging the IP layer for topological location. On the one hand, such an approach needs to revise the way applications interact with the network layer, by modifying the DNS (now returning an HI instead of an IP address) and applications to use the HIP socket extension. On the other hand, early adopters do not necessarily gain any benefit unless all communicating endpoints upgrade to use HIP. In spite of this, such a solution may work in the context of a limited domain.

Another alternative approach is adopted by Information-Centric Networking (ICN) [RFC7476]. By making content a first class citizen of the communication architecture, the "what" rather than the "where" becomes the real focus of the communication. However, as explained in the next sub-section, ICN can run either over the IP layer or completely replace it, which in turn can be seen as running the Internet and ICN as logically completely separated limited domains.

Unmanned Aircraft Systems (UAS) are examples of moving devices that require a stable mobility management scheme since they consist of a number of Unmanned Aerial Vehicles (UAV) subordinated to a Ground Control Station (GCS) [MAROJEVIC20]. The information produced by the different sensors and electronic devices available at each UAV is collected and processed by a software or hardware data acquisition unit, being transmitted towards the GCS, where it is inspected and/or analyzed. Analogously, control information transmitted from the GCS to the UAV enables the execution of control operations over the aircraft, such as changing the route planning or the direction pointed by a camera.

Although UAVs may have redundant links to maintain communications in long-range missions (e.g., satellite), most of the communications between the GCS and the UAVs take place over wireless data links, e.g., based on a radio line-of-sight technology, Wi-Fi or 3G/4G/5G. While in some scenarios, UAVs will operate always under the range of the same cellular base station, in missions with large range, UAVs will move between different cellular or wireless ground infrastructure, meaning that the UAV needs to upload its topological locator and re-start the ongoing communication sessions. In such cases, most of existing Mobile IP approaches may play a role, as well as approaches to split the UAV identifier and the topological locator, such as HIP.

However, while the industry is given the first steps towards evolved UAS architectures and communication models, the data-centric communication plays an increasing role, where information is named and decoupled from its location, and applications/services operate over these named data rather than on host-to-host communications.

In this context, the Data Distribution Service ([DDS]) has emerged as an industry-oriented open standard that follows this approach. The space and time decoupling allowed by DDS is very relevant in any dynamic and distributed system, since interacting entities are not forced to know each other and are not forced to be simultaneously present to exchange data. Time decoupling can significantly simplify the management of intermittent data-links, in particular for wireless connectivity between UAS, as well as facilitate seamless UAV mobility between GCSs. This model of communication, in turn, questions the locator-based addressing used in IP and instead utilizes a data-centric naming.

In the case of using TCP/IP, mobility of UAVs introduces a significant challenge. Consider the case where a GCS is receiving telemetry information from a specific UAV. Assuming that the UAV moves and changes its point of attachment to the network, it will have to configure a new IP address on its wireless interface. However, this is problematic, as the telemetry information is still being sent by to the previous IP address of the UAV. This simple example illustrates the necessity to deploy mobility management solutions to handle this type of situations.

However, mobility management solutions increase the complexity of the deployment and may impact the performance of data distribution, both in terms of signaling/data overhead and communication path delay. Considering the specific case of multicast data streams, mobility of content producers and consumers is inherently handled by multicast routing protocols, which are able to react to changes of location of mobile nodes by reconstructing the corresponding multicast delivery

trees. Nevertheless, this comes with a cost in terms of signaling and data overhead (data may still flow through branches of a multicast delivery tree where there are no receivers while the routing protocol is still converging).

Another alternative is to perform the mobility management of producers and consumers not at the application layer based on IP multicast trees, but on the network layer based on an Information Centric Network approach, which was already mentioned in this section.

Greater flexibility in addressing may help in dealing with mobility more efficiently, e.g., through an augmented semantic that may fulfil the mobility requirements [RFC7429] in a more efficient way or through moving from a locator- to a content or service-centric semantic for addressing.

2.4. Communication Across Services

As a communication infrastructure spanning many facets of life, the Internet integrates services and resources from various aspects such as remote collaboration, shopping, content production as well as delivery, education, and many more. Accessing those services and resources directly through URIs has been proposed by methods such as those defined in ICN [RFC7476], where providers of services and resources can advertise those through unified identifiers without additional planning of identifiers and locations for underlying data and their replicas. Users can access required services and resources by virtue of using the URI-based identification, with an ephemeral relationship built between user and provider, while the building of such relationship may be constrained with user- as well as service-specific requirements, such as proximity (finding nearest provider), load (finding fastest provider), and others.

While systems like ICN [CCN] provide an alternative to the topological addressing of IP, its deployment requires an overlay (over IP) or native deployment (alongside IP), the latter with dedicated gateways needed for translation. Underlay deployments are also envisioned in [RFC8763], where ICN solutions are being used to facilitate communication between IP addressed network endpoints or URI-based service endpoints, still requiring gateway solutions for interconnection with ICN-based networks as well as IP routing based networks (cf., [ICN5G][ICNIP]).

Although various approaches combining service and location-based addressing have been devised, the key challenge here is to facilitate a "natural", i.e., direct communication, without the need for gateways above the network layer.

Another aspect of communication across services is that of chaining individual services to a larger service. Here, an identifier would be used that serves as a link to next hop destination within the chain of single services, as done in the work on Service Function Chaining (SFC). With this, services are identified at the level of Layer 2/3 ([RFC7665], [RFC8754], [RFC8595]) or at the level of name-based service identifiers like URLs [RFC8677] although the service chain identification is carried as a Network Service header (NSH) [RFC7665], separate to the packet identifiers. The forwarding with the chain of services utilizes individual locator-based IP addressing (for L3 chaining) to communicate the chained operations from one Service Function Forwarder [RFC7665] to another, leading to concerns regarding overhead incurred through the stacking of those chained identifiers in terms of packet overhead and therefore efficiency in handling in the intermediary nodes.

Greater flexibility in addressing may allow for incorporating different information, e.g., service as well as chaining semantics, into the overall Internet addressing.

2.5. Communication Traffic Steering

Steering traffic within a communication scenario may involve at least two aspects, namely (i) limiting certain traffic towards a certain set of communication nodes and (ii) restraining the sending of packets towards a given destination (or a chain of destinations) with metrics that would allow the selection among one or more possible destinations.

One possibility for limiting traffic inside limited domains, towards specific objects, e.g., devices, users, or group of them, is subnet partition with techniques such as VLAN [RFC5517], VxLAN [RFC7348], or more evolved solution like TeraStream [TERASTREAM] realizing such partitioning. Such mechanisms usually involve significant configuration, and even small changes in network and user nodes could result in a repartition and possibly additional configuration efforts. Another key aspect is the complete lack of correlation of the topological address and any likely more semantic-rich identification that could be used to make policy decisions regarding traffic steering. Suitably enriching the semantics of the packet address, either that of the sender or receiver, so that such decision could be made while minimizing the involvement of higher layer mechanisms, is a crucial challenge for improving on network operations and speed of such limited domain traffic.

When making decisions to select one out of a set of possible destinations for a packet, IP anycast semantics can be applied albeit being limited to the locator semantic of the IP address itself.

Recent work in [SFCANYCAST] suggests utilizing the notion of IP anycast address to encode a "service identifier", which is dynamically mapped onto network locations where service instances fulfilling the service request may be located. Scenarios where this capability may be utilized are provided in [SFCANYCAST] and include, but are not limited to, scenarios such as edge-assisted VR/AR, transportation, smart cities, smart homes, smart wearables, and digital twins.

The challenge here lies in the possible encoding of not only the service information itself but the constraint information that helps the selection of the "best" service instance and which is likely a service-specific constraint in relation to the particular scenario. The notion of an address here is a conditional (on those constraints) one where this conditional part is an essential aspect of the forwarding action to be taken. It needs therefore consideration in the definition of what an address is, what is its semantic, and how the address structure ought to look like.

As outlined in the previous sub-section, chaining services are another aspect of steering traffic along a chain of constituent services, where the chain is identified through either a stack of individual identifiers, such as in Segment Routing [RFC8402], or as an identifier that serves as a link to next hop destination within the chain, such as in Service Function Chaining (SFC). The latter can be applied to services identified at the level of Layer 2/3 ([RFC7665], [RFC8754], [RFC8595]) or at the level of name-based service identifiers like URLs [RFC8677]. However, the overhead incurred through the stacking of those chained identifiers is a concern in terms of packet overhead and therefore efficiency in handling in the intermediary nodes.

Flexibility in addressing may enable more semantic rich encoding schemes that may help in steering traffic at hardware level and speed, without complex mechanisms usually resulting in handling packets in the slow path of routers.

2.6. Communication with built-in security

Today, strong security in the Internet is usually implemented as a general network service ([PILA], [RFC6158]). Among the various reasons for such approach is the limited semantic of current IP addresses, which do not allow to natively express security features or trust relationships. Efforts like Cryptographically Generated Addresses (CGA) [RFC3972], provide some security features by embedding a truncated public key in the last 57-bit of IPv6 address, thereby greatly enhancing authentication and security within an IP network via asymmetric cryptography and IPsec [RFC4301]. The

development of the Host Identity Protocol (HIP) [RFC7401] saw the introduction of cryptographic identifiers for the newly introduced Host Identity (HI) to allow for enhanced accountability, and therefore trust. The use of those HIs, however, is limited by the size of IPv6 128bit addresses.

Through a greater flexibility in addressing, any security-related key, certificate, or identifier could instead be included in a suitable address structure without any information loss (i.e., as-is, without any truncation or operation as such), avoiding therefore compromises such as those in HIP. Instead, CGAs could be created using full length certificates, or being able to support larger HIP addresses in a limited domain that uses it. This could significantly help in constructing a trusted and secure communication at the network layer, leading to connections that could be considered as absolute secure (assuming the cryptography involved is secure). Even more, anti-abuse mechanisms and/or DDoS protection mechanisms like the one under discussion in PEARG ([PEARG]) Research Group may leverage a greater flexibility of the overall Internet addressing, if provided, in order to be more effective.

2.7. Communication protecting user privacy

See Comments in Section "Issues".

2.8. Communication in Alternative Forwarding Architectures

The performance of communication networks has long been a focus for optimization due to the immediate impact on cost of ownership for communication service providers. Technologies like MPLS [RFC3031] have been introduced to optimize lower layer communication, e.g., by mapping L3 traffic into aggregated labels of forwarding traffic for the purposes of, e.g., traffic engineering.

Even further, other works have emerged in recent years that have replaced the notion of packets with other concepts for the same purpose of improved traffic engineering and therefore efficiency gains. One such area is that of Software Defined Networks (SDN) [RFC7426], which has highlighted how a majority of Internet traffic is better identified by flows, rather than packets. Based on such observation, alternate forwarding architectures have been devised that are flow-based or path-based. With this approach, all data belonging to the same traffic stream is delivered over the same path, and traffic flows are identified by some connection or path identifier rather than by complete routing information, possibly enabling fast hardware based switching (e.g. [DETNET], [PANRG]).

On the one hand, such a communication model may be more suitable for real-time traffic like in the context of Deterministic Networks ([DETNET]), where indeed a lot of work has focused on how to "identify" packets belonging to the same DETNET flow in order to jointly manage the forwarding within the desired deterministic boundaries.

On the other hand, it may improve the communication efficiency in constrained wireless environments (cf., Section 2.1), by reducing the overhead, hence increasing the number of useful bits per second per Hz.

Also, the delivery of information across similar flows may be combined into a multipoint delivery of a single return flow, e.g., for scenarios of requests for a video chunk from many clients being responded to with a single (multi-destination) flow, as outlined in [BIER-MC] as an example. Another opportunity to improve communication efficiency is being pursued in ongoing IETF/IRTF work to deliver IP- or HTTP-level packets directly over path-based or flow-based transport network solutions, such as in [TROSSEN][BIER-MC][ICNIP][ICN5G] with the capability to bundle unicast forward communication streams flexibly together in return path multipoint relations. Such capability is particularly opportune in scenarios such as chunk-based video retrieval or distributed data storage. However, those solutions currently require gateways to "translate" the flow communication into the packet-level addressing semantic in the peering IP networks. Furthermore, the use of those alternative forwarding mechanisms often require the encapsulation of Internet addressing information, leading to wastage of bandwidth as well as processing resources.

Providing an alternative way of forwarding data has also been the motivation for the efforts created in the European Telecommunication Standards Institute (ETSI), which formed an Industry Specification Group (ISG) named Non-IP Networking (NIN) [ETSI-NIN]. This group sets out to develop and standardize a set of protocols leveraging an alternative forwarding architecture, such as provided by a flow-based switching paradigm. The deployment of such protocols may be seen to form limited domains, still leaving the need to interoperate with the (packet-based forwarding) Internet; a situation possibly enabled through a greater flexibility of the addressing used across Internet-based and alternative limited domains alike.

As an alternative to IP routing, EIBP (Extended Internet Bypass Protocol) [EIBP] offers a communications model that can work with IP in parallel and entirely transparent and independent to any operation at network layer. For this, EIBP proposes the use of physical and/or virtual structures in networks and among networks to auto assign

routable addresses that capture the relative position of routers in a network or networks in a connected set of networks, which can be used to route the packets between end domains. EIBP operates at Layer 2.5 and provides encapsulation (at source domain), routing, and de-encapsulation (at destination domain) for packets. EIBP can forward any type of packets between domains. A resolver to map the domain ID to EIBP's edge router addresses is required. When queried for a specific domain, the resolver will return the corresponding edge router structured addresses.

EIBP decouples routing operations from end domain operations, offering to serve any domain, without point solutions to specific domains. EIBP also decouples routing IDs or addresses from end device/domain addresses. This allows for accommodation of new and upcoming domains. A domain can extend EIBP's structured addresses into the domain, by joining as a nested domain under one or more edge routers, or by extending the edge router's structure addresses to its devices.

A greater flexibility in addressing semantics may reduce the aforementioned wastage by accommodating Internet addressing in the light of such alternative forwarding architectures, instead enabling the direct use of the alternative forwarding information.

3. Desired Network Features

From the previous subsection, we recognize that Internet technologies are used across a number of scenarios, each of which brings their own (vertical) view on needed capabilities in order to work in a satisfactory manner to those involved.

In the following, we complement those vertical-specific insights with answers to the question of network features that end users (in the form of individuals or organizations alike) desire from the networked system at large. Answers to this question look at the network more from a horizontal perspective, i.e. not with a specific usage in mind beyond communication within and across networks. The text here summarizes the discussion that took place on the INT Area mailing list after IETF112 on this issue. For some of those identified features, we can already identify how innovations on addressing may impact the realization of a particular feature.

We then combine the insights from both scenario-specific and wider horizontal views for the identification of issues when realizing the specific capability of addressing, presented in Section 4.

1. Always-On: The world is getting more and more connected, leading to being connected to the Internet, anywhere, by any technology (e.g., cable, fiber, or radio), even simultaneously, "all the time", and, most importantly, automatically (without any switch turning). However, when defining "all the time" there is a clear and important difference to be made between availability and reliability vs "desired usage". In other words, "always on" can be seen as a desirable perception at the end user level or as a characteristic of the underlying system. From an end user perspective, clearly the former is of importance, not necessarily leading to an "always on" system notion but instead "always-app-available", merely requiring the needed availability and reliability to realize the perception of being "always on" (e.g., for earthquake alerts), possibly complemented by app-specific methods to realize the "always on" perception (e.g., using local caching rather than communication over the network).

2. Transparency: Being agnostic with respect to local domains network protocols (Bluetooth, ZigBee, Thread, Airdrop, Airplay, or any others) is key to provide an easy and straightforward method for contacting people and devices without any knowledge of network issues, particularly those specific to network-specific solutions. While having a flexible addressing model that accommodates a wide range of use cases is important, the centrality of the IP protocol remains key as a mean to provide global connectivity.

3. Multi-homing: Seamless multi-homing capability for the host is key to best use the connectivity options that may be available to an end user, e.g., for increasing resilience in cases of failures of one available option. Protocols like LISP, SHIM6, QUIC, MPTCP, SCTP (to cite a few) have been successful at providing this capability in an incremental way, but too much of that capability is realized within the application, making it hard to leverage across all applications. While today each transport protocol has its own way to perform multi-address discovery, the network layer should provide the multi-homing feature (e.g., SHIM6 can be used to discover all addresses on both ends), and then leave the address selection to the transport. With that, multi-address discovery remains a network feature exposed to the upper layers. This may also mean to update the Socket API (which may be actually the first thing to do), which does not necessarily mean to expose more network details to the applications but instead be more address agnostic yet more expressive.

4. Mobility: A lot of work has been put in MobileIP ([RFC5944],[RFC6275]) to provide seamless and lossless communications for moving nodes (vehicle, satellites). However, it has never been widely deployed for several reasons, like complexity of the protocol and the fact that the problem has often been tackled at higher layers, with applications resilient to address changes. However, similar to multi-homing, solving the problem at higher layers means that each and every transport protocol and application have their own way to deal with mobility, leading to similar observations as those for the previous multi-homing aspect.

5. Security and Privacy: The COVID-19 pandemic has boosted end users' desire to be protected and protect their privacy. The balance among privacy, security, and accountability is not simple to achieve. There exist different views on what those properties should be, however the network should provide the means to provide what is felt as the best trade-off for the specific use case.

6. Performance: While certainly desirable, "performance" is a complex issue that depends on the objectives of those building for but also paying for performance. Examples are (i) speed (shorter paths/direct communications), (ii) bandwidth (10petabit/s for a link), (iii) efficiency (less overlays/encapsulations), (iv) high efficacy or sustainability (avoid waste). From an addressing perspective, length/format/semantics that may adapt to the specific use case (e.g. use short addresses for low power IoT, or, where needed, longer for addresses embedding certificates for strong authentication, authorization and accountability) may contribute to the performance aspects that end users desire, such as reducing waste through not needed encapsulation or needed conversion at network boundaries.

7. Availability, Reliability, Predictability: These three properties are important to enable wide-range of services and applications according to the desired usage (cf. point 1).

8. Do not do harm: Access to the Internet is considered a human right [RFC8280]. Access to and expression through it should align with this core principle. This issue transcends through a variety of previously discussed 'features' that are desired, such as privacy, security but also availability and reliability. However, lifting the feature of network access onto a basic rights level also brings in the aspect of "do not do harm" through the use of the Internet with respect to wider societal objectives. Similar to other industries, such as electricity or cars, preventing harm usually requires an interplay of

commercial, technological, and regulatory efforts, such as the enforcement of seat belt wearing to reduce accident death. As a first step, the potential harmfulness of a novel method must be recognized and weighted against the benefits of its introduction and use. One increasingly important consideration in the technology domain is "sustainability" of resource usage for an end user's consumption of and participation in Internet services. As an example, Distributed Ledger Technologies (DLT) are seen as an important tool for a variety of applications, including Internet decentralization ([DINRG]). However, the non-linear increase in energy consumption means that extending proof-of-work systems to the entire population of the planet would not only be impractical but also possibly highly wasteful, not just at the level of computational but also communication resource usage [DLT-draft]. This poses the question on how novel methods for addressing may improve on sustainability of such technologies, particularly if adopted more widely.

9. Maximum Transmission Unit (MTU): One long standing issue in the Internet is related to the MTU and how to discover the path MTU in order to avoid fragmentation ([I-D.ietf-6man-mtu-option], [I-D.templin-6man-aero]). While it makes sense to always leverage as much performance from local systems as possible, this should come without sacrificing the ability to communicate with all systems. Having a solid solution to solve the issue would make the overall interconnection of systems more robust.

4. Issues in Addressing

The desired properties outlined in the previous section have implications that go beyond addressing and need to be tackled from a larger architectural point of view. Such a discussion is left as future action, limiting the present document at discussing only the addressing viewpoint and identifying shortcomings perceived from this perspective.
There are a number of issues that we can identify from the communication scenarios in Section 2 and the network features generally desire from the network, presented in Section 3. We do not claim to be exhaustive in our list:

1. Limiting Alternative Address Semantics: Several communication scenarios pursue the use of alternative semantics of what constitute an 'address' of a packet traversing the Internet, which may fall foul of the defined network interface semantic of IP addresses.

2. Hampering Security: Aligning with the semantic and length limitations of IP addressing may hamper the security objectives of any new semantic, possibly leading to detrimental effects and possible other workarounds (at the risk of introducing fragility rather than security).

1. Hampering Privacy:

   * Easy individual identification

   * Flow linkability

   * App/Activity profiling

2. Complicating Traffic Engineering: Utilizing a plethora of non-address inputs into the traffic steering decision in real networks complicates traffic engineering in that it makes the development of suitable policies more complex, while also leading to possible contention between methods being used.

3. Hampering Efficiency: Extending IP addressing through point-wise solutions also hampers efficiency, e.g., through needed re-encapsulation (therefore increasing the header processing overhead as well as header-to-payload ratio), through introducing path stretch, or through requiring compression techniques to reduce the header proportion of large addresses when operating in constrained environments.

4. Fragility: The introduction of point solutions, each of which comes with possibly own usages of address or packet fields, together with extension-specific operations, increases the overall fragility of the resulting system, caused, for instance, through contention between feature extensions that were neither foreseen in the design nor tested during the implementation phase.

5. Extensibility: Accommodating new requirements through ever new extensions as an extensibility approach to addressing compounds aspects discussed before, i.e., fragility, efficiency etc. It complicates engineering due to the clearly missing boundaries against which contentions with other extensions could be managed. It complicates standardization since extension-based extensibility requires independent, and often lengthy, standardization processes. And ultimately, deployments are complicated due to backward compatibility testing required for any new extension being integrated into the deployed system.

The table below shows how the above identified issues do arise somehow in our outlined communication scenarios in Section 2. This overview will be deepened in more details in the gap analysis document [I-D.jia-intarea-internet-addressing-gap-analysis].

	Issue 1	Issue 2	Issue 3	Issue 4	Issue 5	Issue 6
Constrained Environments				x	x	x
Dynamically Changing Topologies	x		x	x	x	x
Moving Endpoints	x		x	x	x	x
Across Services	x		x	x	x	x
Traffic Steering	x		x	x	x	x
Built-in Security	x	x		x	x	x
Alternative Forwarding Architectures	x			x		x

Table 1: Issues Involved in Challenging Scenarios

5. Problem Statement

This document identifies a number of scenarios as well as general features end users would want from the network, positioning the existing Internet addressing structure itself as a potential hindrance in solving key problems for Internet service provisioning. Such problems include supporting new, e.g., service-oriented, scenarios more efficiently, with improved security and efficient traffic engineering, as well as large scale mobility. We can observe that those new forms of communication are particularly driven by the conceptual framework of limited domains, realizing the requirements of stakeholders for an optimized communication in those limited domains, while still utilizing the Internet for interconnection as

well as for access to the wealth of existing Internet services.

This co-existence of optimized LD-level as well as Internet communication creates a tussle between those requirements on addressing stemming from those limited domains and those coming from the Internet in the form of agreed IPv6 semantics. This tussle directly refers back to our introductory question on flexibility in addressing (or leaving the problem to limited domain solutions to deal with). It is also captured in the discussion on where new features are being introduced, i.e. at the edge or core of the Internet.

But more importantly, the question on 'what is an address anyway' (derived from what features we may want from the network) should not be guided by the answers that the Internet can give us today, e.g., being a mere ephemeral token for accessing PoP-based services (as indicated in related arch-d mailing list discussions), but instead what features could be enabled by a particular view of what an address is. However, that is not to 'second guess' the market and its possible evolution, but to outline clear features from which to derive clear principles for a design.

For this, it is important to recognize that skewing the technical capabilities of any feature, let alone addressing, to the current economic situation of the Internet bears the danger of locking down innovation capabilities as an outcome of those technical limitations being introduced. Instead, addressing must align with enabling the model of permissionless but compatible innovation that the IETF has been promoting, ultimately enabling the serendipity of new applications that has led to many of those applications we can see in the Internet today.

At this stage, this document does not provide a definite answer nor does it propose or promote specific solutions to the problems here portrayed. Instead, this document aims at stimulating discussion on the emerging needs for addressing, with the possibility to fundamentally re-think the addressing in the Internet beyond the current objectives of IPv6, in order to provide the flexibility to suitably support the many new forms of communication that will emerge. Addressing can be rather flexible and can be of any form that applications may need. There is no limitation on the address to preclude any future applications.

To complement the problem statement in this document, the companion gap analysis document [I-D.jia-intarea-internet-addressing-gap-analysis] deepens the issues identified in Section 4 along key properties of today's Internet addressing.

6. Security Considerations

The present memo does not introduce any new technology and/or mechanism and as such does not introduce any security threat to the TCP/IP protocol suite.

Nevertheless, it is worth to observe whether or not greater flexibility of addressing (as suggested in previous sections) would allow to introduce fully featured security in endpoint identification, potentially able to eradicate the spoofing problem, as one example. Furthermore, it may be used to include application gateways' certificates in order to provide more efficiency, e.g., using web certificates also in the addressing of web services. While increasing security, privacy protection may also be improved.

7. IANA Considerations

This document does not include an IANA request.

8. References

8.1. Normative References

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <https://www.rfc-editor.org/info/rfc791>.

8.2. Informative References

[ALOHA] Kuo, F., "The ALOHA System", ACM SIGCOMM Computer Communication Review Vol. 25, pp. 41-44, DOI 10.1145/205447.205451, January 1995, <https://doi.org/10.1145/205447.205451>.

[BACnet] "BACnet-A Data Communication Protocol for Building Automation and Control Networks", ANSI/ASHRAE Standard 135-2016, January 2016, <https://www.techstreet.com/ashrae/standards/ashrae-135-2016?product_id=1918140>.

[BIER-MC] Trossen, D., Rahman, A., Wang, C., and T. Eckert, "Applicability of BIER Multicast Overlay for Adaptive Streaming Services", Work in Progress, Internet-Draft, draft-ietf-bier-multicast-http-response-06, 10 July 2021, <https://www.ietf.org/archive/id/draft-ietf-bier-multicast-http-response-06.txt>.

[BLE] "Bluetooth Specification", Bluetooth SIG Working Groups, n.d., <https://www.bluetooth.com/specifications>.

[CARTISEAN] Hughes, L., Shumon, K., and Y. Zhang, "Cartesian Ad Hoc Routing Protocols", Ad-Hoc, Mobile, and Wireless Networks pp. 287-292, DOI 10.1007/978-3-540-39611-6_27, 2003, <https://doi.org/10.1007/978-3-540-39611-6_27>.

[CCN] Jacobson, V., Smetters, D., Thornton, J., Plass, M., Briggs, N., and R. Braynard, "Networking named content", Proceedings of the 5th international conference on Emerging networking experiments and technologies - CoNEXT '09, DOI 10.1145/1658939.1658941, 2009, <https://doi.org/10.1145/1658939.1658941>.

[CHEN21] Chen, Y., Li, H., Liu, J., Wu, Q., and Z. Lai, "GAMS: An IP Address Management Mechanism in Satellite Mega-constellation Networks", 2021 International Wireless Communications and Mobile Computing (IWCMC), DOI 10.1109/iwcmc51323.2021.9498722, June 2021, <https://doi.org/10.1109/iwcmc51323.2021.9498722>.

[CHRIKI19] Chriki, A., Touati, H., Snoussi, H., and F. Kamoun, "FANET: Communication, mobility models and security issues", Computer Networks Vol. 163, pp. 106877, DOI 10.1016/j.comnet.2019.106877, November 2019, <https://doi.org/10.1016/j.comnet.2019.106877>.

[DDS] AL-Madani, B., Elkhider, S., and S. El-Ferik, "DDS-Based Containment Control of Multiple UAV Systems", Applied Sciences Vol. 10, pp. 4572, DOI 10.3390/app10134572, July 2020, <https://doi.org/10.3390/app10134572>.

[DECT-ULE] "Digital Enhanced Cordless Telecommunications (DECT); Common Interface (CI); Part 1: Overview", ETSI European Standard, EN 300 175-1, V2.6.1, May 2009, <https://www.etsi.org/deliver/etsi_en/300100_300199/30017501/02.06.01_60/en_30017501v020601p.pdf>.

[DETNET] "Deterministic Networking (DetNet)", n.d., <https://datatracker.ietf.org/wg/detnet/about/>.

[DINRG] "Decentralized Internet Infrastructure - DINRG", n.d., <https://datatracker.ietf.org/rg/dinrg/about/>.

[DLT-draft]
Trossen, D., Guzman, D., Bride, M. M., and X. Fan, "Impact of DLTs on Provider Networks", Work in Progress, Internet-Draft, draft-trossen-rtgwg-impact-of-dlts-01, 2 March 2022, <https://www.ietf.org/archive/id/draft-trossen-rtgwg-impact-of-dlts-01.txt>.

[ECMA-340] EECMA-340, "Near Field Communication - Interface and Protocol (NFCIP-1) 3rd Ed.", June 2013.

[EIBP] Shenoy, S Chandraiah, P Willis, N., "A Structured Approach to Routing in the Internet", June 2021, <First Intl Workshop on Semantic Addressing and Routing for Future Networks>.

[ETSI-NIN] ETSI - European Telecommunication Standards Institute, "Non-IP Networking - NIN", n.d., <https://www.etsi.org/technologies/non-ip-networking>.

[HANDLEY] Handley, M., "Delay is Not an Option: Low Latency Routing in Space", Proceedings of the 17th ACM Workshop on Hot Topics in Networks, DOI 10.1145/3286062.3286075, November 2018, <https://doi.org/10.1145/3286062.3286075>.

[I-D.ietf-6man-mtu-option]
Hinden, R. M. and G. Fairhurst, "IPv6 Minimum Path MTU Hop-by-Hop Option", Work in Progress, Internet-Draft, draft-ietf-6man-mtu-option-13, 28 February 2022, <https://www.ietf.org/archive/id/draft-ietf-6man-mtu-option-13.txt>.

[I-D.ietf-intarea-gue]
Herbert, T., Yong, L., and O. Zia, "Generic UDP Encapsulation", Work in Progress, Internet-Draft, draft-ietf-intarea-gue-09, 26 October 2019, <https://www.ietf.org/archive/id/draft-ietf-intarea-gue-09.txt>.

[I-D.ietf-lisp-introduction]
Cabellos, A. and D. S. (Ed.), "An Architectural Introduction to the Locator/ID Separation Protocol (LISP)", Work in Progress, Internet-Draft, draft-ietf-lisp-introduction-15, 20 September 2021, <https://www.ietf.org/archive/id/draft-ietf-lisp-introduction-15.txt>.

[I-D.ietf-lisp-mn]
Farinacci, D., Lewis, D., Meyer, D., and C. White, "LISP Mobile Node", Work in Progress, Internet-Draft, draft-ietf-lisp-mn-11, 30 January 2022, <https://www.ietf.org/archive/id/draft-ietf-lisp-mn-11.txt>.

[I-D.ietf-lisp-nexagon]
Barkai, S., Fernandez-Ruiz, B., Tamir, R., Rodriguez-Natal, A., Maino, F., Cabellos-Aparicio, A., and D. Farinacci, "Network-Hexagons: H3-LISP GeoState & Mobility Network", Work in Progress, Internet-Draft, draft-ietf-lisp-nexagon-19, 14 September 2021, <https://www.ietf.org/archive/id/draft-ietf-lisp-nexagon-19.txt>.

[I-D.ietf-lisp-rfc6833bis]
Farinacci, D., Maino, F., Fuller, V., and A. Cabellos, "Locator/ID Separation Protocol (LISP) Control-Plane", Work in Progress, Internet-Draft, draft-ietf-lisp-rfc6833bis-30, 18 November 2020, <https://www.ietf.org/archive/id/draft-ietf-lisp-rfc6833bis-30.txt>.

[I-D.jia-intarea-internet-addressing-gap-analysis]
Jia, Y., Trossen, D., Iannone, L., Shenoy, N., and P. Mendes, "Gap Analysis in Internet Addressing", Work in Progress, Internet-Draft, draft-jia-intarea-internet-addressing-gap-analysis-01, 23 October 2021, <https://www.ietf.org/archive/id/draft-jia-intarea-internet-addressing-gap-analysis-01.txt>.

[I-D.templin-6man-aero]
Templin, F. L., "Automatic Extended Route Optimization (AERO)", Work in Progress, Internet-Draft, draft-templin-6man-aero-39, 22 February 2022, <https://www.ietf.org/archive/id/draft-templin-6man-aero-39.txt>.

[ICN5G]
Ravindran, R., Suthar, P., Trossen, D., Wang, C., and G. White, "Enabling ICN in 3GPP’s 5G NextGen Core Architecture", Work in Progress, Internet-Draft, draft-irtf-icnrg-5gc-icn-04, 10 January 2021, <https://www.ietf.org/archive/id/draft-irtf-icnrg-5gc-icn-04.txt>.

[ICNIP] Trossen, D., Robitzsch, S., Reed, M., Al-Naday, M., and J. Riihijarvi, "Internet Services over ICN in 5G LAN Environments", Work in Progress, Internet-Draft, draft-trossen-icnrg-internet-icn-5glan-04, 1 October 2020, <https://www.ietf.org/archive/id/draft-trossen-icnrg-internet-icn-5glan-04.txt>.

[IEEE_1901.1] "Standard for Medium Frequency (less than 15 MHz) Power Line Communications for Smart Grid Applications", IEEE 1901.1 IEEE-SA Standards Board, May 2018, <https://ieeexplore.ieee.org/document/8360785>.

[LR-WPAN] "IEEE 802.15.4 - IEEE Standard for Low-Rate Wireless Networks", IEEE 802.15 WPAN Task Group 4, May 2020, <https://standards.ieee.org/standard/802_15_4-2020.html>.

[MANET1] Abdallah, A., Abdallah, E., Bsoul, M., and A. Otoom, "Randomized geographic-based routing with nearly guaranteed delivery for three-dimensional ad hoc network", International Journal of Distributed Sensor Networks Vol. 12, pp. 155014771667125, DOI 10.1177/1550147716671255, October 2016, <https://doi.org/10.1177/1550147716671255>.

[MAROJEVIC20] Marojevic, V., Guvenc, I., Dutta, R., Sichitiu, M., and B. Floyd, "Advanced Wireless for Unmanned Aerial Systems: 5G Standardization, Research Challenges, and AERPAW Architecture", IEEE Vehicular Technology Magazine Vol. 15, pp. 22-30, DOI 10.1109/mvt.2020.2979494, June 2020, <https://doi.org/10.1109/mvt.2020.2979494>.

[OCADO] "Ocado Technologys robot warehouse a Hive of IoT innovation", n.d., <https://techmonitor.ai/tech-leaders/ocado-technology-robot-hive-innovation>.

[PANRG] "Path Aware Networking Research Group - PANRG", n.d., <https://datatracker.ietf.org/rg/panrg/about/>.

[PEARG] "Privacy Enhancements and Assessments Research Group - PEARG", n.d., <https://irtf.org/pearg>.

[PILA] Krahenbuhl, C., Legner, M., Bitterli, S., and A. Perrig, "Pervasive Internet-Wide Low-Latency Authentication", 2021 International Conference on Computer Communications and Networks (ICCCN), DOI 10.1109/icccn52240.2021.9522235, July 2021, <https://doi.org/10.1109/icccn52240.2021.9522235>.

[RFC2775] Carpenter, B., "Internet Transparency", RFC 2775, DOI 10.17487/RFC2775, February 2000, <https://www.rfc-editor.org/info/rfc2775>.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <https://www.rfc-editor.org/info/rfc3031>.

[RFC3972] Aura, T., "Cryptographically Generated Addresses (CGA)", RFC 3972, DOI 10.17487/RFC3972, March 2005, <https://www.rfc-editor.org/info/rfc3972>.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI 10.17487/RFC4301, December 2005, <https://www.rfc-editor.org/info/rfc4301>.

[RFC4919] Kushalnagar, N., Montenegro, G., and C. Schumacher, "IPv6 over Low-Power Wireless Personal Area Networks (6LoWPANs): Overview, Assumptions, Problem Statement, and Goals", RFC 4919, DOI 10.17487/RFC4919, August 2007, <https://www.rfc-editor.org/info/rfc4919>.

[RFC5061] Stewart, R., Xie, Q., Tuexen, M., Maruyama, S., and M. Kozuka, "Stream Control Transmission Protocol (SCTP) Dynamic Address Reconfiguration", RFC 5061, DOI 10.17487/RFC5061, September 2007, <https://www.rfc-editor.org/info/rfc5061>.

[RFC5177] Leung, K., Dommety, G., Narayanan, V., and A. Petrescu, "Network Mobility (NEMO) Extensions for Mobile IPv4", RFC 5177, DOI 10.17487/RFC5177, April 2008, <https://www.rfc-editor.org/info/rfc5177>.

[RFC5275] Turner, S., "CMS Symmetric Key Management and Distribution", RFC 5275, DOI 10.17487/RFC5275, June 2008, <https://www.rfc-editor.org/info/rfc5275>.

[RFC5517] HomChaudhuri, S. and M. Foschiano, "Cisco Systems' Private VLANs: Scalable Security in a Multi-Client Environment", RFC 5517, DOI 10.17487/RFC5517, February 2010, <https://www.rfc-editor.org/info/rfc5517>.

[RFC5944] Perkins, C., Ed., "IP Mobility Support for IPv4, Revised", RFC 5944, DOI 10.17487/RFC5944, November 2010, <https://www.rfc-editor.org/info/rfc5944>.

[RFC6158] DeKok, A., Ed. and G. Weber, "RADIUS Design Guidelines", BCP 158, RFC 6158, DOI 10.17487/RFC6158, March 2011, <https://www.rfc-editor.org/info/rfc6158>.

[RFC6182] Ford, A., Raiciu, C., Handley, M., Barre, S., and J. Iyengar, "Architectural Guidelines for Multipath TCP Development", RFC 6182, DOI 10.17487/RFC6182, March 2011, <https://www.rfc-editor.org/info/rfc6182>.

[RFC6275] Perkins, C., Ed., Johnson, D., and J. Arkko, "Mobility Support in IPv6", RFC 6275, DOI 10.17487/RFC6275, July 2011, <https://www.rfc-editor.org/info/rfc6275>.

[RFC6626] Tsirtsis, G., Park, V., Narayanan, V., and K. Leung, "Dynamic Prefix Allocation for Network Mobility for Mobile IPv4 (NEMOv4)", RFC 6626, DOI 10.17487/RFC6626, May 2012, <https://www.rfc-editor.org/info/rfc6626>.

[RFC7348] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", RFC 7348, DOI 10.17487/RFC7348, August 2014, <https://www.rfc-editor.org/info/rfc7348>.

[RFC7401] Moskowitz, R., Ed., Heer, T., Jokela, P., and T. Henderson, "Host Identity Protocol Version 2 (HIPv2)", RFC 7401, DOI 10.17487/RFC7401, April 2015, <https://www.rfc-editor.org/info/rfc7401>.

[RFC7426] Haleplidis, E., Ed., Pentikousis, K., Ed., Denazis, S., Hadi Salim, J., Meyer, D., and O. Koufopavlou, "Software-Defined Networking (SDN): Layers and Architecture Terminology", RFC 7426, DOI 10.17487/RFC7426, January 2015, <https://www.rfc-editor.org/info/rfc7426>.

[RFC7429] Liu, D., Ed., Zuniga, JC., Ed., Seite, P., Chan, H., and CJ. Bernardos, "Distributed Mobility Management: Current Practices and Gap Analysis", RFC 7429, DOI 10.17487/RFC7429, January 2015, <https://www.rfc-editor.org/info/rfc7429>.

[RFC7476] Pentikousis, K., Ed., Ohlman, B., Corujo, D., Boggia, G., Tyson, G., Davies, E., Molinaro, A., and S. Eum, "Information-Centric Networking: Baseline Scenarios", RFC 7476, DOI 10.17487/RFC7476, March 2015, <https://www.rfc-editor.org/info/rfc7476>.

[RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function Chaining (SFC) Architecture", RFC 7665, DOI 10.17487/RFC7665, October 2015, <https://www.rfc-editor.org/info/rfc7665>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[RFC8280] ten Oever, N. and C. Cath, "Research into Human Rights Protocol Considerations", RFC 8280, DOI 10.17487/RFC8280, October 2017, <https://www.rfc-editor.org/info/rfc8280>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[RFC8595] Farrel, A., Bryant, S., and J. Drake, "An MPLS-Based Forwarding Plane for Service Function Chaining", RFC 8595, DOI 10.17487/RFC8595, June 2019, <https://www.rfc-editor.org/info/rfc8595>.

[RFC8677] Trossen, D., Purkayastha, D., and A. Rahman, "Name-Based Service Function Forwarder (nSFF) Component within a Service Function Chaining (SFC) Framework", RFC 8677, DOI 10.17487/RFC8677, November 2019, <https://www.rfc-editor.org/info/rfc8677>.

[RFC8754] Filsfils, C., Ed., Dukes, D., Ed., Previdi, S., Leddy, J., Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header (SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020, <https://www.rfc-editor.org/info/rfc8754>.

[RFC8763] Rahman, A., Trossen, D., Kutscher, D., and R. Ravindran, "Deployment Considerations for Information-Centric Networking (ICN)", RFC 8763, DOI 10.17487/RFC8763, April 2020, <https://www.rfc-editor.org/info/rfc8763>.

[RFC8799] Carpenter, B. and B. Liu, "Limited Domains and Internet Protocols", RFC 8799, DOI 10.17487/RFC8799, July 2020, <https://www.rfc-editor.org/info/rfc8799>.

[RFC8926] Gross, J., Ed., Ganga, I., Ed., and T. Sridhar, Ed., "Geneve: Generic Network Virtualization Encapsulation", RFC 8926, DOI 10.17487/RFC8926, November 2020, <https://www.rfc-editor.org/info/rfc8926>.

[RFC9000] Iyengar, J., Ed. and M. Thomson, Ed., "QUIC: A UDP-Based Multiplexed and Secure Transport", RFC 9000, DOI 10.17487/RFC9000, May 2021, <https://www.rfc-editor.org/info/rfc9000>.

[SFCANYCAST] Wion, A., Bouet, M., Iannone, L., and V. Conan, "Distributed Function Chaining with Anycast Routing", Proceedings of the 2019 ACM Symposium on SDN Research, DOI 10.1145/3314148.3314355, April 2019, <https://doi.org/10.1145/3314148.3314355>.

[TERASTREAM] "Deutsche Telekom tests TeraStream, the network of the future, in Croatia", n.d., <https://www.telekom.com/en/media/media-information/archive/deutsche-telekom-tests-terastream-the-network-of-the-future-in-croatia-358444>.

[TROSSEN] Trossen, D., Sarela, M., and K. Sollins, "Arguments for an information-centric internetworking architecture", ACM SIGCOMM Computer Communication Review Vol. 40, pp. 26-33, DOI 10.1145/1764873.1764878, April 2010, <https://doi.org/10.1145/1764873.1764878>.

[WANG19] Wang, P., Zhang, J., Zhang, X., Yan, Z., Evans, B., and W. Wang, "Convergence of Satellite and Terrestrial Networks: A Comprehensive Survey", IEEE Access Vol. 8, pp. 5550-5588, DOI 10.1109/access.2019.2963223, 2020, <https://doi.org/10.1109/access.2019.2963223>.

Acknowledgments

Thanks to all the people that shared insightful comments both privately to the authors as well as on various mailing list, especially on the INTArea Mailing List. Also thanks for the interesting discussions to Stewart Bryant, Ron Bonica, Toerless Eckert, Brian E. Carpenter, Kiran Makhijani, Fred Templin.

Authors' Addresses

Yihao Jia
Huawei Technologies Co., Ltd
156 Beiqing Rd.
Beijing
100095
P.R. China
Email: jiayihao@huawei.com

Dirk Trossen
Huawei Technologies Duesseldorf GmbH
Riesstr. 25C
80992 Munich
Germany
Email: dirk.trossen@huawei.com

Luigi Iannone
Huawei Technologies France S.A.S.U.
18, Quai du Point du Jour
92100 Boulogne-Billancourt
France
Email: luigi.iannone@huawei.com

Nirmala Shenoy
Rochester Institute of Technology
New-York, 14623
United States of America
Email: nxsvks@rit.edu

Paulo Mendes
Airbus
Willy-Messerschmitt Strasse 1
81663 Munich
Germany
Email: paulo.mendes@airbus.com

Donald E. Eastlake 3rd
Futurewei Technologies
2386 Panoramic Circle
Apopka, FL, 32703
United States of America
Email: d3e3e3@gmail.com

Peng Liu
China Mobile
32 Xuanwumen West Ave
Xicheng, Beijing
100053
P.R. China
Email: liupengyjy@chinamobile.com

Dino Farinacci
lispers.net
United States of America
Email: farinacci@gmail.com

Network Working Group
Internet-Draft
Intended status: Informational
Expires: 7 July 2024

L. Han, Ed.
R. Li
A. Retana
Futurewei Technologies, Inc.
M. Chen
L. Su
Tianji. Jiang
China Mobile

4 January 2024

Problems and Requirements of Satellite Constellation for Internet
draft-lhan-problems-requirements-satellite-net-06

Abstract

This document presents the detailed analysis about the problems and requirements of satellite constellation used for Internet. It starts from the satellite orbit basics, coverage calculation, then it estimates the time constraints for the communications between satellite and ground-station, also between satellites. How to use satellite constellation for Internet is discussed in detail including the satellite relay and satellite networking. The problems and requirements of using traditional network technology for satellite network integrating with Internet are finally outlined.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Satellite constellation for Internet is emerging. Even there is no constellation network established completely yet at the time of the publishing of the draft (June 2021), some basic internet service has been provided and has demonstrated competitive quality to traditional broadband service.

This memo will analyze the challenges for satellite network used in Internet by traditional routing and switching technologies. It is based on the analysis of the dynamic characters of both ground-station-to-satellite and inter-satellite communications and its impact to satellite constellation networking.

The memo also provides visions for the future solution, such as in routing and forwarding.

The memo focuses on the topics about how the satellite network can work with Internet. It does not focus on physical layer technologies (wireless, spectrum, laser, mobility, etc.) for satellite communication.

2. Terminology

LEO	Low Earth Orbit with the altitude from 180 km to 2000 km.
VLEO	Very Low Earth Orbit with the altitude below 450 km
MEO	Medium Earth Orbit with the altitude from 2000 km to 35786 km
GEO	Geosynchronous orbit with the altitude 35786 km
GSO	Geosynchronous satellite on GEO
ISL	Inter Satellite Link
ISLL	Inter Satellite Laser Link

3GPP	3rd Generation Partner-ship Project
NTN	Non-Terrestrial Network, it includes satellite networks (satellite could be on GEO, MEO, LEO or VLEO), high altitude platform systems (HAPS) and other types of air-to-ground networks
EIRP	Effective isotropic radiated power
P2MP	Point to Multiple Points
GS	Ground Station, a device on ground connecting the satellite. In the document, GS will hypothetically provide L2 and/or L3 functionality in addition to process/send/receive radio wave. It might be different as the reality that the device to process/send/receive radio wave and the device to provide L2 and/or L3 functionality could be separated.
SGS	Source ground station. For a specified flow, a ground station that will send data to a satellite through its uplink.
DGS	Destination ground station. For a specified flow, a ground station that is connected to a local network or Internet, it will receive data from a satellite through its downlink and then forward to a local network or Internet.
PGW	Packet Gateway
UPF	User Packet Function
NodeB	The base station in 3G
eNodeB	The base station in 4G
gNB	gNodeB, the base station in 5G
PE router	Provider Edge router
CE router	Customer Edge router
P router	Provider router
LSA	Link-state advertisement

LSP	Link-State PDUs
L1	Layer 1, or Physical Layer in OSI model [OSI-Model]
L2	Layer 2, or Data Link Layer in OSI model [OSI-Model]
L3	Layer 3, or Network Layer in OSI model [OSI-Model], it is also called IP layer in TCP/IP model
BGP	Border Gateway Protocol [RFC4271]
eBGP	External Border Gateway Protocol, two BGP peers have different Autonomous Number
iBGP	Internal Border Gateway Protocol, two BGP peers have same Autonomous Number
IGP	Interior gateway protocol, examples of IGPs include Open Shortest Path First (OSPF [RFC2328]), Routing Information Protocol (RIP [RFC2453]), Intermediate System to Intermediate System (IS-IS [RFC7142]) and Enhanced Interior Gateway Routing Protocol (EIGRP [RFC7868]).

3. Overview

The traditional satellite communication system is composed of few GSO and ground stations. For this system, each GSO can cover 42% Earth's surface [GEO-Coverage], so as few as three GSO can provide the global coverage theoretically. With so huge coverage, GSO only needs to amplify signals received from uplink of one ground station and relay to the downlink of another ground station. There is no inter-satellite communications needed. Also, since the GSO is stationary to the ground station, there is no mobility issue involved.

Recently, more and more LEO and VLEO satellites have been launched, they attract attentions due to their advantages over GSO and MEO in terms of higher bandwidth, lower cost in satellite, launching, ground station, etc. Some organizations [ITU-6G][Surrey-6G][Nttdocomo-6G] have proposed the non-terrestrial network using LEO, VLEO as important parts for 6G to extend the coverage of Internet. 3GPP has been working on the NTN integration with 5G and beyond. SpaceX has started to build the satellite constellation called StarLink that will deploy over 10 thousand LEO and VLEO satellites finally [StarLink]. China also started to request the spectrum from ITU to establish a constellation that has 12992 satellites [China-constellation]. European Space Agency (ESA) has proposed "Fiber in the sky" initiative to connect satellites with fiber network on Earth [ESA-HydRON].

When satellites on MEO, LEO and VLEO are deployed, the communication problem becomes more complicated than for GSO satellites. This is because the altitude of MEO/LEO/VLEO satellites are much lower. As a result, the coverage of each satellite is much smaller than for GSO, and the satellite is moving very fast on the ground reference and not relatively stationary to the ground. This will lead to:

1. More satellites than GSO are needed to provide the global coverage. Appendix A will brief the satellite orbit parameters; analyze the coverage area, and the minimum number of satellites required to cover the earth surface; discuss the real deployment for LEO satellite network.

2. The point-to-point communication between satellite and ground station can only last a few minutes. Mobility issue has to be considered. Detailed analysis about the lifetime of communication is done in Appendix B.1.

3. The inter-satellite communication is needed, and all satellites need to form a network. details are described in Appendix B.2 that includes the communication between satellites on different orbit and different geographic areas.

In Section 4, we will discuss couple of topics of satellite network integration with Internet, such as using satellite network for broadband access and wireless access, the current 3GPP works for satellite network in 5G and beyond.

Finally, the problems and requirements for satellite network integration with Internet will be discussed and analyzed in Section 5.

As the 1st satellite constellation company in history, the SpaceX/ StarLink will be inevitably mentioned in the draft. But it must be noted that all information about SpaceX/StarLink in the draft are from the public. Authors of the draft have no relationship or relevant inside knowledge of SpaceX/Starlink.

4. Satellite Network Integrated with Internet

Since there is no complete satellite network established yet, all following analysis is based on the predictions from the traditional GEO communication. The analysis also learnt how other type of network has been used in Internet, such as Broadband access network, Mobile access network, Enterprise network and Service Provider network.

To integrate the satellite network with Internet, many other technologies are needed to provide the functions on different layeres. Currently, there are four major international SDO (standard organizations) involved in the development of different technologies: IETF, 3GPP, CCSDS (Consultative Committee for Space Data Systems [CCSDS]), DVB (Digital Video Broadcast [DVB]). Section 4.1 will discuss the different protocol stacks based on different combinations of technologies from different SDOs.

As a criteria to be part of Internet, any device connected to any satellite should be able to communicate with any public IP4 or IPv6 address in Internet. There could be three types of methods to deliver IP packet from source to destination by satellite:

1. Data packet is relayed between ground station and satellite. For this method, there is no inter-satellite communication and networking. Data packet is bounced once or couple times between ground stations and satellites until the packet arrives at the destination in Internet.

2. Data packet is delivered by inter-satellite networking. For this method, the data packet traverses with multiple satellites connected by ISL and inter-satellite networking is used to deliver the packet to the destination in Internet.

3. Both satellite relay and inter-satellite networking are used. For this method, the data packet is relayed in some segments and traverse with multiple satellites in other segments. It is a combination of the method 1 and method 2.

Using the above methods for IP packet delivery via satellite network, we will have two typical use cases for satellite network. One is for the general broadband access (see Section 4.2), another is for the integration with 3GPP wireless network including 4G and 5G (see Section 4.3 and Section 4.4).

It must be noted that we use 3GPP as an use case example does not mean other technologies cannot be used, e.g., using DVB instead of 3GPP for the satellite network integration. We use 3GPP is because 3GPP has done the most thorough research and produced lots of solutions for satellite networking, such as using 5G NR for satellite up link and down link, use transparent pay load and regenerative pay load for different scenarios, etc. (see Section 4.3).

4.1. Protocol Stack for Satellite Networking with Different Technologies

Figure 1 illustrates three typical protocol statcks that use different technology combinations. This does not include the use of private technologies for wireless and Link Layer such as Starlink.

The stacks show obviously that TCP/IP is the common technologies, IETF has to (at least) provide the L3 and L4 technologies for satellite networking integrated with Internet.

```
  +-------------------+ +--------------------+ +-------------------+
  |5G upper layers(SA)| | CCSDS upper layers | |  DVB upper layers |
  +-------------------+ +--------------------+ +-------------------+
  |      TCP/UDP      | |      TCP/UDP       | |      TCP/UDP      |
  |         IP        | |         IP         | |        IP         |
  +-------------------+ +--------------------+ +-------------------+
  |       SDAP        | |        IPoC        | |      IPoGSE       |
  |       PDCP        | |                    | |                   |
  |        RLC        | |        USLP        | |    DVB MPE/GSE    |
  |        MAC        | |                    | |                   |
  +-------------------+ +--------------------+ +-------------------+
  |      3GPP RF;     | |      CCSDS RF;     | |   DVB S2/S2X/T2;  |
  | 3GPP Spectrum;FSO | | CCSDS Spectrum;FSO | | DVB Spectrum; FSO |
  +-------------------+ +--------------------+ +-------------------+
        3GPP Stacks           CCSDS Statks            DVB Stacks
```

Figure 1: Protocol Statcks for Different Standard Technologies

Some meaning of symbols in Figure 1 are as follows:

SA Service Architecture

SDAP SService Data Adaption Protocol

PDCP	Packet Data Convergence Protocol
RLC	Radio Link Control
MAC	Medium Access Control
RF	Radio Frequency
FSO	Free Space Optics, provided by ITU
IPoC	IP over CCSDS
USLP	Unified Space Link Protocol
IPoGSE	IP over Generic Stream Encapsulation
MPE	Multiprotocol Encapsulation
GSE	Generic Stream Encapsulation
S2	Digital Video Broadcasting Satellite Second Generation
S2X	Digital Video Broadcasting Satellite Second Generation Extension
T2	Digital Video Broadcasting Satellite Second Generation Terrestrial

4.2. Use Satellite Network for Broadband Access

For this use case, the end user terminal or local network is connected to a ground station, and another ground station is connected to Internet. Two ground stations will have IP connectivity via a satellite network. The satellite network could be by satellite relays or by inter-satellite network.

Follows are typical deployment scenarios that a Satellite network is used for broadband access of Internet.

1. The end user terminal access Internet through satellite relay (Figure 2 for one satellite relay, Figure 3 for multiple satellite relay).

2. The end user terminal access Internet through inter-satellite-networking (Figure 4).

3. The local network access Internet through satellite relay (Figure 5 for one satellite relay, Figure 6 for multiple satellite relay).

4. The local network access Internet through inter-satellite-networking (Figure 7).

```
                    S1----\            /-----------\
                   /       \          /             \
            T---GW--GS1--S2--GS2-------PE   Internet   +
                   \       /          \             /
                    \---S3/            \-----------/
```

Figure 2: End user terminal access Internet through one satellite relay

```
                  S1----\    S4----\       /-----------\
                 /       \  /       \     /             \
          T---GW--GS1--S2--GS2---S5--GS3---PE   Internet   +
                 \       /  \       /     \             /
                  \---S3/    \---S6/       \-----------/
```

Figure 3: End user terminal access Internet through multiple satellite relay

```
             S1-----S2-----S3--\            /----------\
            /                  \          /            \
      T---GW--GS1--S4----S5---S6---GS2-------PE  Internet   +
            \                  /          \            /
             \---S7----S8----S9/            \----------/
```

Figure 4: End user terminal access Internet through inter-satellite-networking

```
       /-----------\           S1----\           /-------\
      /             \         /       \         /         \
     + Local network CE------GS1--S4--GS2-------PE Internet +
      \             /         \       /         \         /
       \-----------/           \---S7/           \-------/
```

Figure 5: Local network access Internet through one satellite relay

```
       /-----------\         S1----\   S4----\       /-------\
      /             \       /       \  /      \     /         \
     + Local network CE----GS1--S2--GS2--S5--GS3---PE Internet +
      \             /       \       / \       /     \         /
       \-----------/         \---S3/   \---S6/       \-------/
```

Figure 6: Local network access Internet through multiple satellite relay

```
   /-----------\          S1-----S2-----S3---\            /------\
  /             \         /                   \          /        \
 + Local network CE------GS1--S4----S5---S6---GS2-------PE Internet+
  \             /         \                   /          \        /
   \-----------/           \---S7----S8----S9/            \------/
```

Figure 7: Local network access Internet through inter-satellite-networking

In above Figure 2 to Figure 7, the meaning of symbols are as follows:

T	The end user terminal
GW	Gateway router
GS1, GS2, GS3	Ground station with L2/L3 routing/switch functionality.
S1 to S9	Satellites
PE	Provider Edge Router
CE	Customer Edge Router

The above configuration may have different variations, e.g., the GW and GS functions can be merged into one ohysical devices.

4.3. Use Satellite Network with 3GPP Wireless Access Network

For this use case, the wireless access network (4G, 5G) defined in 3GPP is used with satellite network. By such integration, a user terminal or local network can access Internet via 3GPP wireless network and satellite network. The End user terminal or local network access Internet through satellite network and Mobile Access Network. There are two cases: 1) From mobile access network to satellite network or 2) From satellite network to mobile access network, Satellite network includes inter satellite network and relay network. See Figure 8 for mobile access network to satellite network, and Figure 9 for satellite network to mobile access network.

```
+--------------+    +-------------+    +---------+    +--------+
|    T or      |    |Mobile Access|    |Satellite|    |Internet|
| Local network+----+  Network    +----+ Network +----+        |
+--------------+    +-------------+    +---------+    +--------+
```

Figure 8: End user terminal or local network access Internet through Mobile Access Network and Satellite Network

```
+--------------+   +---------+   +-------------+   +--------+
|    T or      |   |Satellite|   |Mobile Access|   |Internet|
| Local network+---+ Network +---+  Network    +---+        |
+--------------+   +---------+   +-------------+   +--------+
```

Figure 9: End user terminal or local network access Internet through Satellite Network and Mobile Access Network

4.4. Recent Development and Study in 3GPP for Satellite Network

3GPP SA Working Groups (WG) feature a couple of satellite-related projects (or SIDs).

For Release 18, 3GPP has finished the project 'Study on 5G System with Satellite Backhaul' [TR-23.700-27] and 'Study on 5GC enhancement for satellite access Phase 2' [TR-23.700-28].

For Release 19, 3GPP will study more topics for satellite network used for 5G system, such as Regenerative payload generic architecture study, Store and Forward Satellite operation, etc.

One key aspect is to investigate the potential architecture requirements and enhancements to deploy UPFs on satellites (LEO/MEO/GEO) with gNBs on the ground. Specifically, it targets at enhancing the local-switching capability for UE-to-UE data communication when UEs are served by UPFs on-board satellite(s). Similarly, the SA1 WG proposed a new satellite-based SID in which the service end points (could also be called UEs in a broader sense) may continuously move in a fast way. The UEs can be ships, boats, and cars, etc., which are located in remote regions that need the connection to LEO's for achieving communication.

In all the SIDs, satellite based backhaul is important for mission critical scenarios in remote areas. Here, we want to clarify that while 3GPP documents TS 23.501 [TS-23.501] and 23.502 [TS-23.502] specify that a ground base station, i.e., gNB, may have multiple types of satellite backhauls (BH), e.g., GEO BH, LEO BH and LEO-BH with ISL, this use case focuses specifically on the LEO-BH with ISL. ISL stands for inter-satellite link.

Clearly, when a satellite backhaul involves multi-hop ISL path connected via different satellites, the capabilities provided by the satellite path would be changed and adjusted dynamically. For example, in the LEO case, the peering relationship between two neighboring satellites changes roughly every 5 minutes thanks to the

orbital movement (see Table 2). This will definitely impair the networking performance and stability, and, in worst case, may cause the loss of connectivity. Even if some overlay tunneling mechanisms could be used to address the multi-hop ISL issue, the extra delay and potentially less bandwidth as introduced naturally by the ever-changing backhaul path would still impact the traffic engineering over the links.

The following diagram Figure 10 demonstrate the dynamic characteristics of satellite backhaul between two UEs. In the figure, UEs are connected, via gNBs, to UPFs on-board satellites. Both UPFs are connected via multi-hop ISLs to the 5G core (5GC) on the ground. There are two different multi-hop ISL paths: o A UE has to rely on a multi-hop ISL path to connect to 5GC on the ground. o When two UEs intend to communicate via the local data switching on satellite(s), some new ISL-based peering has to be established which would bring in the multi-hop ISL scenario. For example, the ISL between the Sat#1 and Sat#2 helps form a multi-hop path (marked N19 in the diagram) between the two UEs. Note that if the UPF-based local data switching involves only one UPF, then it is designated as intra-UPF local switching and relatively simpler. This is compared to the case of inter-UPF local switching as shown in the diagram.

```
     Sat#: Satellite                    GS:  Ground Station
     UPF:  User Plane Function (5G)     gNB: Next Generation NodeB

                         Sat#1
                  +----------+       +--------+
                  |  UPF#1   | (ISL) | Sat#.. |
    UE#1--gNB#1--|(on-board)|- - - -|        |-----+
                  +----------+       +--------+     |
                       :                            |      --------
                       :(N19)                       v     /        \
                       :(ISL)                      GS ---+ 5G Core +
                       :                            ^     \        /
                  +----------+       +--------+     |      --------
                  |  UPF#2   | (ISL) | Sat#.. |     |
    UE#2--gNB#2--|(on-board)|- - - -|        |-----+
                  +----------+       +--------+
                     Sat#2
```

Figure 10: Use Satellite network as back haul for 5G

In this diagram, both UEs are served by different satellite backhauls. If the local data switching via LEO UPFs on-board could be established (via the N19 ISL forwarding), then the system efficiency and QoE improvement would be achieved. Here, since UEs are served by different satellites, a multi-hop ISL scenario must be

supported. But, this scenario posts challenges due to the dynamic satellite network topology and distinguished transmission capabilities from different satellites.

For example, if the UE-to-UE session has to maintain a service over longer time (> 5 minutes) such that the Sat#1 and Sat#2 move apart, then a new ISL path with potentially a new N19-ISL might be established. In worst case, if newly-involved satellites in the path happen to be polar-orbit ones and they do not support cross-seam ISLs, the communication latency may change dramatically when cross-seam transits or leaves. In another example, if both UEs belong to the same entity and need to form a 5G-VN group, then the 5G LAN-type service with PSA UPF-based local-switching must be applied among them.

Regardless, more efficient satellite communication mechanisms must be adopted, e.g., running efficient satellite-based routing protocols, establishing tunnels between LEO UPFs on-board, etc., for better local-data switching.

Further, 5GS may collaborate with satellite networks to improve QoS. One 5GC NF (i.e., SMF) can initiate UP path monitoring, and accordingly receive UP path monitoring results indicating observed delay. After that, the SMF takes corresponding actions like further verifying network statistics, updating sessions, etc. The coordination with the satellite networks would improve the process, which suggests satellites networks respond better to the (monitor-based) polling from 5GS.

One more thing we want to point out is that, while the propagation delay of satellite backhaul paths may change dramatically with the movement of satellite, this kind of change normally be periodic and can be well predicated based on the operation information of satellite constellation. Thus, making use of these information would also help for better services.

5. Problems and Requirements for Satellite Constellation for Internet

As described in Section 4, satellites in a satellite constellation can either relay internet traffic or multiple satellites can form a network to deliver internet traffic. More detailed analysis are in following sub sections. There might have multiple solutions for each method described in Section 4, following contexts only discuss the most plausible solution from networking perspectives.

Section 5.1 will list the common problems and requirements for both satellite relay and satellite networking.

Section 5.2 and Section 5.3 will describe key problems, requirement and potential solution from the networking perspective for these two cases respectively.

5.1. Common Problems and Requirements

For both satellite relay and satellite networking, satellite-ground-station communication must be used, so, the problems and requirements for the satellite-ground-station communication is common and will apply for both methods.

When one satellite is communicating with ground station, the satellite only needs to receive data from uplink of one ground station, process it and then send to the downlink of another ground station. Figure 2 illustrates this case. Normally microwave is used for both links.

Additionally, from the coverage analysis in Appendix A.2 and real deployment in Appendix A.3, we can see one ground station may communicate with multiple satellites. Similarly, one satellite may communicate with multiple ground stations. The characters for satellite-ground-station communication are:

1. Satellite-ground-station communication is P2MP.
   Since microwave physically is the carrier of broadcast communication, one satellite can send data while multiple ground stations can receive it. Similarly, one ground station can send data and multiple satellites can receive it.

2. Satellite-ground-station communication is in open space and not secure.
   Since electromagnetic fields for microwave physically are propagating in open space. The satellite-ground-station communication is also in open space. It is not secure naturally.

3. Satellite-ground-station communication is not steady.
   Since the satellite is moving with high speed, from Appendix B.1, the satellite-ground-station communication can only last a certain period of time. The communication peers will keep changing.

4. Satellite-to-Satellite communication is not steady.
   For some satellites, even they are in the same altitude and move in the same speed, but they move in the opposite direction, from Appendix B.2.2, the satellite-to-satellite communication can only last a certain period of time. The communication peers will keep changing.

5. Satellite-to-Satellite distance is not steady.
   For satellites with the same altitude and same moving direction, even their relative position is steady, but the distance between satellites are not steady. This will lead to the inter-satellite-communication's bandwidth and latency keep changing.

6. Satellite physical resource is limited.
   Due to the weight, complexity and cost constraint, the physical resource on a satellite, such as power supply, memory, link speed, are limited. It cannot be compared with the similar device on ground. The design and technology used should consider these factors and take the appropriate approach if possible.

The requirements of satellite-ground-station communication are:

R1. The bi-directional communication capability
   Both satellites and ground stations have the bi-directional communication capability

R2. The identifier for satellites and ground stations
   Satellites and ground stations should have Ethernet and/or IP address configured for the device and each link. More detailed address configuration can be seen in each solution.

R3. The capability to decide where the IP packet is forwarded to.
   In order to send Internet traffic or IP date to destination correctly, satellites and ground station must have Ethernet hub or switching or IP routing capability. More detailed capability can be seen in each solution.

R4. The protocol to establish the satellite-ground-station communication.
   For security and management purpose, the satellite-ground-station communication is only allowed after both sides agree through a protocol. The protocol should be able to establish a secured channel for the communication when a new communication peer comes up. Each ground station should be able to establish multiple channels to communicate with multiple satellites. Similarly, each satellite should be able to establish multiple channels to communicate to multiple ground stations.

R5. The protocol to discover the state of communication peer.
   The discover protocol is needed to detect the state of communication peer such as peer's identity, the state of the peer and other info of the peer. The protocol must be running securely without leaking the discovered info.

R6. The internet data packet is forwarded securely.
When satellite or ground station is sending the IP packet to its peer, the packet must be relayed securely without leaking the user data.

R7. The internet data packet is processed efficiently on satellite
Due to the resource constraint on a satellite, the packet may need more efficient mechanism to be processed on satellite. The process on satellite should be very minimal and offloaded to ground as much as possible.

5.2. Satellite Relay

One of the reasons to use satellite constellation for internet access is it can provide shorter latency than using the fiber underground. But using ISL for inter-satellite communication is the premise for such benefit in latency. Since the ISL is still not mature and adopted commercially, satellite relay is a only choice currently for satellite constellation used for internet access. In [UCL-Mark-Handley], detailed simulations have demonstrated better latency than fiber network by satellite relay even the ISL is not present.

5.2.1. One Satellite Relay

One satellite relay is the simplest method for satellite constellation to provide Internet service. By this method, IP traffic will be relayed by one satellite to reach the DGS and go to Internet.

The solution option and associated requirements are:

S1. The satellite only does L1 relay or the physical signal process.

For this solution, a satellite only receives physical signal, amplify it and broadcast to ground stations. It has no further process for packet, such as L2 packet compositing and processing, etc. All packet level work is done only at ground station. The requirements for the solution are:

R1-1. SGS and BGS are configured as IP routing node. Routing protocol is running in SGS and BGS
SGS and BGS is a IP peer for a routing protocol (IGP or BGP). SGS will send internet traffic to DGS as next hop through satellite uplink and downlink.

R1-2. DGS must be connected with Internet.
DGS can process received packet from satellite and forward the packet to the destination in Internet.

In addition to the above requirements, following problem should be solved:

P1-1. IP continuity between two ground stations
This problem is that two ground stations are connected by one satellite relay. Since the satellite is moving, the IP continuity between ground stations is interrupted by satellite changing periodically. Even though this is not killing problem from the view point that IP service traditionally is only a best effort service, it will benefit the service if the problem can be solved. Different approaches may exist, such as using hands off protocols, multipath solutions, etc.

S2. The satellite does the L2 relay or L2 packet process.

For this solution, IP packet is passing through individual satellite as an L2 capable device. Unlike in the solution S1, satellite knows which ground station it should send based on packet's destination MAC address after L2 processing. The advantage of this solution over S1 is it can use narrower beam to communicate with DGS and get higher bandwidth and better security. The requirements for the solution are:

R2-1. Satellite must have L2 bridge or switch capability
In order to forward packet to properly, satellite should run some L2 process such as MAC learning, MAC switching. The protocol running on satellite must consider the fast movement of satellite and its impact to protocol convergence, timer configuration, table refreshment, etc.

R2-2. same as R1-1 in S1

R2-3. same as R1-2 in S1

In addition to the above requirements, the problem P1-1 for S1 should also apply.

5.2.2. Multiple Satellite Relay

For this method, packet from SGS will be relayed through multiple intermediate satellites and ground station until reaching a DGS.

This is more complicated than one satellite relay described in Section 5.2.1.

One general solution is to configure both satellites and ground-stations as IP routing nodes, proper routing protocols are running in this network. The routing protocol will dynamically determine forwarding path. The obvious challenge for this solution is that all links between satellite and ground station are not static, according to the analysis in Appendix B.1, the lifetime of each link may last only couple of minutes. This will result in very quick and constant topology changes in both link state and IP adjacency, it will cause the distributed routing algorithms may never converge. So this solution is not feasible.

Another plausible solution is to specify path statically. The path is composed of a serials of intermediate ground stations plus SGS and DGS. This idea will make ground stations static and leave the satellites dynamic. It will reduce the fluctuation of network path, thus provide more steady service. One variant for the solution is whether the intermediate ground stations are connected to Internet. Separated discussion is as below:

S1. Manual configuring routing path and table

For this solution, the intermediate ground stations and DGS are specified and configured manually during the stage of network planning and provisioning. Following requirements apply:

R1-1. Specify a path from SGS to DGS via a list of intermediate ground stations.
The specified DGS must be connected with internet. Other specified intermediate ground stations does not have to

R1-2. All Ground stations are configured as IP routing node.
Static routing table on all ground stations must be pre-configured, the next hop of routes to Internet destination in any ground station is configured to going through uplink of satellite to the next ground station until reaching the DGS.

R1-3. All Satellites are configured as either L1 relay or L2 relay.
The Satellite can be configured as L1 relay or L2 relay described in S1 and S2 respectively in Section 5.2.1

In addition to the above requirements, the problem P1-1 in Section 5.2.1 should also apply.

S2. Automatic decision by routing protocol.

This solution is only feasible after the IP continuity problem (P1-1 in Section 5.2.1) is solved. Following requirements apply:

R2-1. All Ground stations are configured as IP routing node. Proper routing protocols are configured as well.
The satellite link cost is configured to be lower than the ground link. In such a way, the next hop of routes for the IP forwarding to Internet destination in any ground station will be always going through the uplink of satellite to the next ground station until reaching the DGS.

R2-2. All Satellites are configured as either L1 relay or L2 relay.
The Satellite can be configured as L1 relay or L2 relay described in S1 and S2 respectively in Section 5.2.1

In addition to the above requirements, the problem P1-1 in Section 5.2.1 should also apply.

5.3. Satellite Networking by ISL

In the draft, satellite Network is defined as a network that satellites are inter-connected by inter-satellite links (ISL). One of the major difference of satellite network with the other type of network on ground (telephone, fiber, etc.) is its topology and links are not stationary, some new issues have to be considered and solved. Follows are the factors that impact the satellite networking.

5.3.1. L2 or L3 network

The 1st question to answer is should the satellite network be configured as L2 or L3 network? As analyzed in Appendix A.2 and Appendix A.3, since there are couple of hundred or over ten thousand satellites in a network, L2 network is not a good choice, instead, L3 or IP network is more appropriate for such scale of network.

5.3.2. Inter-satellite-Link Lifetime

If we assume the orbit is circular and ignore other trivial factors, the satellite speed is approximately determined by the orbit altitude as described in the Appendix B.1. The satellite orbit can determine if the dynamic position of two satellites is within the range of the inter-satellite communication. That is 2000km for laser communication [Laser-communication-range] by Inter Satellite Laser Link (ISLL).

When two satellites' orbit planes belong to the same group, or two orbit planes share the same altitude and inclination, and when the satellites move in the same direction, the relative positions of two satellites are relatively stationary, and the inter-satellite communication is steady. But when the satellites move in the

opposite direction, the relative positions of two satellites are not stationary, the communication lifetime is couple of minutes. The Appendix B.2.2 has analyzed the scenario.

When two satellites' orbit planes belong to the different group, or two orbit planes have different altitude, the relative position of two satellite are unstable, and the inter-satellite communication is not steady. As described in Appendix B.2, The life of communication for two satellites depends on the following parameters of two satellites:

1. The speed vectors.

2. The altitude difference

3. The intersection angle

From the examples shown in Table 4 to Table 7, we can see that the lifetime of inter-satellite communication for the different group of orbit planes are from couple of hundred seconds to about 18 hours. This fact will impact the routing technologies used for satellite network and will be discussed in Section 5.3.3.

5.3.3. Problems for Traditional Routing Technologies

When the satellite network is integrated with Internet by traditional routing technologies, following provisioning and configuration (see Figure 11) will apply:

1. The ground stations connected to local network and internet are treated as PE router for satellite network (called PE_GS1 and PE_GS2 in the following context), and all satellites are treated as P router.

2. All satellites in the network and ground stations are configured to run IGP.

3. The eBGP is configured between PE_GS and its peered network's PE or CE.

The work on PE_GS1 are:

* The local network routes are received at PE_GS1 from CE by eBGP. The routes are redistributed to IGP and then IGP flood them to all satellites. (Other more efficient methods, such as iBGP or BGP reflectors are hard to be used, since the satellite is moving and there is no easy way to configure a full meshed iBGP session for all satellites, or configure one satellite as BGP reflector in satellite network.)

* The internet routes are redistributed from IGP to eBGP running on PE_GS1, and eBGP will advertise them to CE.

The work on PE_GS2 are:

* The Internet routes are received at PE_GS2 from PE by eBGP. The routes are redistributed to IGP and then IGP flood them to all satellites. (Similar as in PE_GS1, Other more efficient methods, such as iBGP or BGP reflector cannot be used.)

* The local network routes are redistributed from IGP to eBGP running on PE_GS2, and eBGP will advertise them to Internet.

```
  /--------\             S1---S2----S3----\               /------\
 /          \            /    IGP domain   \             /        \
+ Local net CE--eBGP--PE_GS1---S4---S5---PE_GS2--eBGP--PE Internet +
 \          /            \                 /             \        /
  \--------/              \---S6---S7---S8/               \------/
```

Figure 11: Local access Internet through inter-satellite-networking

Local access Internet through inter-satellite-networking

On PE-GS1, due to the fact that IGP link between PE_GS1 and satellite is not steady; this will lead to following routing activity:

1. When one satellite is connecting with PE_GS1, the satellite and PE_GS1 form a IGP adjacency. IGP starts to exchange the link state update.

2. The local network routes received by eBGP in PE_GS1 from CE are redistributed to IGP, and IGP starts to flood link state update to all satellites.

3. Meanwhile, the Internet routes learnt from IGP in PE_GS1 will be redistributed to eBGP. eBGP starts to advertise to CE.

4. Every satellite will update its routing table (RIB) and forwarding table (FIB) after IGP finishes the SPF algorithm.

5. When the satellite is disconnecting with PE-GS1, the IGP adjacency between satellite and PE_GS1 is gone. IGP starts to exchange the link state update.

6. The routes of local network and satellite network that were redistributed to IGP in step 2 will be withdrawn, and IGP starts to flood link state update to all satellites.

7. Meanwhile, the Internet routes previously redistributed to eBGP in step 3 will also be withdrawn. eBGP starts to advertise route withdraw to CE.

8. Every satellite will update its routing table (RIB) and forwarding table (FIB) after the SPF algorithm.

Similarly on PE_GS2, due to the fact that IGP link between PE_GS2 and satellite is not steady; this will lead to following routing activity:

1. When one satellite is connecting with PE_GS2, the satellite and PE_GS2 form a IGP adjacency. IGP starts to exchange the link state update.

2. The Internet routes previously received by eBGP in PE_GS2 from PE are redistributed to IGP, IGP starts to flood the new link state update to all satellites.

3. Meanwhile, the routes of local network and satellite network learnt from IGP in PE_GS2 will be redistributed to eBGP. eBGP starts to advertise to Internet peer PE.

4. Every satellite will update its routing table (RIB) and forwarding table (FIB) after IGP finishes the SPF algorithm.

5. When the satellite is disconnecting with PE-GS2, the IGP adjacency between satellite and PE_GS2 is gone. IGP starts to exchange the link state update.

6. The internet routes previously redistributed to IGP in step 2 will be withdrawn, and IGP starts to flood link state update to all satellites

7. Meanwhile, the routes of local network and satellite network previously redistributed to eBGP in step 3 will also be withdrawn. eBGP starts to advertise route withdraw to PE.

8. Every satellite will update its routing table (RIB) and forwarding table (FIB) after the SPF algorithm.

For the analysis of detailed events above, the estimated time interval between event 1 and 5 for PE_GS1 and PE_GS2 can use the analysis in Appendix B.1. For example, it is about 398s for LEO and 103s for VLEO. Within this time interval, the satellite network including all satellites and two ground stations must finish the works from 1 to 4 for PE_GS1 and PE_GS2. The normal internet IPv6 and IPv4 BGP routes size are about 850k v4 routes + 100K v6 routes [BGP-Table-Size]. There are couple critical problems associated with the events:

P1. Frequent IGP update for its link cost
Even for satellites in different orbit with the steady relative positions, the distance between satellites is keep changing. If the distance is used as the link cost, it means the IGP has to update the link cost frequently. This will make IGP keep running and update its routing table.

P2. Frequent IGP flooding for the internet routes
Whenever the IGP adjacency changes (step 1 and 5 for PE_GS2), it will trigger the massive IGP flooding for the link state update for massive internet routes learnt from eBGP. This will result in the IGP re-convergency, RIB and FIP update.

P3. Frequent BGP advertisement for the internet routes
Whenever the IGP adjacency changes (step 3 and 7 for PE_GS1), it will trigger the massive BGP advertisement for the internet routes learnt from IGP. This will result in the BGP re-convergency, RIB and FIB update. BGP convergency time is longer than IGP. The document [BGP-Converge-Time1] has shown that the BGP convergence time varies from 50sec to couple of hundred seconds. The analysis [BGP-Converge-Time2] indicated that per entry update takes about 150us, and it takes o(75s) for 500k routes, or o(150s) for 1M routes.

P4. More frequent IGP flooding and BGP update in whole satellite network

To provide the global coverage, a satellite constellation will have many ground stations deployed. For example, StarLink has applied for the license for up to one million ground stations [StarLink-Ground-Station-Fcc], in which, more than 50 gateway ground stations (equivalent to the PE_GS2) have been registered [SpaceX-Ground-Station-Fcc] and deployed in U.S. [StarLink-GW-GS-map]. It is expected that the gateway ground station will grow quickly to couple of thousands [Tech-Comparison-LEOs]. This means almost each satellite in the satellite network would have a ground station connected. , Due to the fact that all satellites are moving, many IGP adjacency changes may occur in a shorter period of time described in Appendix B.1 and result in the problem P1 and P2 constantly occur.

P5. Service is not steady

Due to the problems P1 to P3, the service provider of satellite constellation is hard to provide a steady service for broadband service by using inter-satellite network and traditional routing technologies.

As a summary, the traditional routing technology is problematic for large scale inter-satellite networking for Internet. Enhancements on traditional technologies, or new technologies are expected to solve the specific issues associated with satellite networking.

6. IANA Considerations

This memo includes no request to IANA.

7. Security Considerations

Security considerations for communication between satellite and ground station, or between satellites are described in corresponding sections. There is no extra security issue introduced by this memo.

8. Contributors

9. Acknowledgements

10. References

10.1. Normative References

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[RFC2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, DOI 10.17487/RFC2328, April 1998, <https://www.rfc-editor.org/info/rfc2328>.

[RFC7142] Shand, M. and L. Ginsberg, "Reclassification of RFC 1142 to Historic", RFC 7142, DOI 10.17487/RFC7142, February 2014, <https://www.rfc-editor.org/info/rfc7142>.

[RFC2453] Malkin, G., "RIP Version 2", STD 56, RFC 2453, DOI 10.17487/RFC2453, November 1998, <https://www.rfc-editor.org/info/rfc2453>.

[RFC7868] Savage, D., Ng, J., Moore, S., Slice, D., Paluch, P., and R. White, "Cisco’s Enhanced Interior Gateway Routing Protocol (EIGRP)", RFC 7868, DOI 10.17487/RFC7868, May 2016, <https://www.rfc-editor.org/info/rfc7868>.

10.2. Informative References

[KeplerianElement] "Keplerian elements", <https://en.wikipedia.org/wiki/Orbital_elements>.

[GEO-Coverage] "Coverage of a geostationary satellite at Earth", <https://www.planetary.org/space-images/coverage-of-a-geostationary>.

[Nttdocomo-6G] "NTTDPCOM 6G White Paper", <https://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/whitepaper_6g/DOCOMO_6G_White_PaperEN_20200124.pdf>.

[ITU-6G] "ITU 6G vision", <https://www.itu.int/dms_pub/itu-s/opb/itujnl/S-ITUJNL-JFETF.V1I1-2020-P09-PDF-E.pdf>.

[Surrey-6G] "Surrey 6G vision", <https://www.surrey.ac.uk/sites/default/files/2020-11/6g-wireless-a-new-strategic-vision-paper.pdf>.

[OSI-Model]
"OSI Model", <https://en.wikipedia.org/wiki/OSI_model>.

[StarLink] "Star Link", <https://en.wikipedia.org/wiki/Starlink>.

[China-constellation]
"China Constellation", <https://www.itu.int/ITU-R/space/asreceived/Publication/DisplayPublication/23706>.

[ESA-HydRON]
"HydRON: Fiber in the sky", <https://www.esa.int/ESA_Multimedia/Videos/2021/04/HydRON_Fibre_in_the_sky>.

[SpaceX-Non-GEO]
"FCC report: SPACEX V-BAND NON-GEOSTATIONARY SATELLITE SYSTEM", <https://fcc.report/IBFS/SAT-LOA-20170301-00027/1190019.pdf>.

[Satellite-coverage]
Alan R.Washburn, Department of Operations Research, Naval Postgraduate School, "Earth Coverage by Satellites in Circular Orbit", <https://faculty.nps.edu/awashburn/Files/Notes/EARTHCOV.pdf>.

[Microwave-vs-Laser-communication]
International Journal for Research in Applied Science and Engineering Technology (IJRASET), "Comparison of Microwave and Optical Wireless Inter-Satellite Links", <https://www.ijraset.com/fileserve.php?FID=7815>.

[Laser-communication-range]
"Interferometric optical communications can potentially lead to robust, secure, and naturally encrypted long-distance laser communications in space by taking advantage of the underlying physics of quantum entanglement.", <https://www.laserfocusworld.com/optics/article/16551652/interferometry-quantum-entanglement-physics-secures-spacetospace-interferometric-communications>.

[BGP-Table-Size]
"BGP in 2020 - BGP table", <https://blog.apnic.net/2021/01/05/bgp-in-2020-the-bgp-table/>.

[BGP-Converge-Time1]
"BGP in 2020 - BGP Update Churn",
<https://labs.apnic.net/?p=1397>.

[BGP-Converge-Time2]
"Bringing SDN to the Internet, one exchange point at the time",
<https://www.cs.princeton.edu/courses/archive/fall14/cos561/docs/SDX.pdf>.

[StarLink-Ground-Station-Fcc]
"APPLICATION FOR BLANKET LICENSED EARTH STATIONS",
<https://fcc.report/IBFS/SES-LIC-INTR2019-00217/1616678>.

[SpaceX-Ground-Station-Fcc]
"List of SpaceX applications for ground stations",
<https://fcc.report/IBFS/Company/Space-Exploration-Technologies-Corp-SpaceX>.

[Tech-Comparison-LEOs]
"A Technical Comparison of Three Low Earth Orbit Satellite Constellation Systems to Provide Global Broadband",
<http://www.mit.edu/~portillo/files/Comparison-LEO-IAC-2018-slides.pdf>.

[StarLink-GW-GS-map]
"StarLink gateway ground station map",
<https://www.google.com/maps/d/u/0/viewer?mid=1H1x8jZs8vfjy60TvKgpbYs_grargieVw>.

[UCL-Mark-Handley]
"Using ground relays for low-latency wide-area routing in megaconstellations",
<https://discovery.ucl.ac.uk/id/eprint/10090242/1/hotnets-ucl.pdf>.

[TR-23.700-27]
3GPP, "Study on support of satellite backhauling in 5GS", December 2022.

[TR-23.700-28]
3GPP, "Study on 5GC enhancement for satellite access Phase 2", December 2022.

[TS-23.501]
3GPP, "System architecture for the 5G System (5GS)", December 2021.

[TS-23.502]
3GPP, "Procedures for the 5G System (5GS)", December 2021.

[CCSDS] CCSDS, "Consultative Committee for Space Data Systems", <https://public.ccsds.org/default.aspx>.

[DVB] DVB, "Digital Video Broadcasting", <https://en.wikipedia.org/wiki/DVB>.

Appendix A. Basics of Satellite Constellation

This section will introduce some basics for satellite such as orbit parameters, coverage estimation, minimum number of satellite and orbit plane required, real deployments.

A.1. Satellite Orbit

The orbit of a satellite can be either circular or ecliptic, it can be described by following Keplerian elements [KeplerianElement]:

1. Inclination (i)

2. Longitude of the ascending node (Omega)

3. Eccentricity (e)

4. Semimajor axis (a)

5. Argument of periapsis (omega)

6. True anomaly (nu)

For a circular orbit, two parameters, Inclination and Longitude of the ascending node, will be enough to describe the orbit.

A.2. Coverage of LEO and VLEO Satellites and Minimum Number Required

The coverage of a satellite is determined by many physical factors, such as spectrum, transmitter power, the antenna size, the altitude of satellite, the air condition, the sensitivity of receiver, etc. EIRP could be used to measure the real power distribution for coverage. It is not deterministic due to too many variants in a real environment. The alternative method is to use the minimum elevation angle from user terminals or gateways to a satellite. This is easier and more deterministic. [SpaceX-Non-GEO] has suggested originally the minimum elevation angle of 35 degrees and deduced the radius of the coverage area is about 435km and 1230km for VLEO (altitude 335.9km) and LEO (altitude 1150km) respectively. The details about

how the coverage is calculated from the satellite elevation angle can be found in [Satellite-coverage].

Using this method to estimate the coverage, we can also estimate the minimum number of satellites required to cover the earth surface.

It must be noted, SpaceX has recently reduced the required minimum elevation angle from 35 degrees to 25 degrees. The following analysis still use 35 degrees.

Assume there is multiple orbit planes with the equal angular interval across the earth surface (The Longitude of the ascending node for sequential orbit plane is increasing with a same angular interval). Each orbit plane will have:

1. The same altitude.

2. The same inclination of 90 degree.

3. The same number of satellites.

With such deployment, all orbit planes will meet at north and south pole. The density of satellite is not equal. Satellite is more dense in the space above the polar area than in the space above the equator area. Below estimations are made in the worst covered area, or the area of equator where the satellite density is the minimum.

Figure 12 illustrates the coverage area on equator area, and each satellite will cover one hexagon area. The figure is based on plane geometry instead of spherical geometry for simplification, so, the orbit is parallel approximately.

Figure 13 shows how to calculate the radius (Rc) of coverage area from the satellite altitude (As) and the elevation angle (b).

```
                        x
                 |      |
                 |      |
                 x   *********
                 |   *    |    *
                 |  *     |     *
            *********   x      *
           *     |  *    |     *
          *      |   *   |    *
        *        x   *********
         *       |   *   |
          *      |  *    |
           *********   x
                 |      |
                 |   orbit 2          ^ north
                 x                    |
                 |                    |
                 |                    |
             orbit 1                  +-------> east
```

Figure 12: Satellite coverage on ground

```
                |<---  2*Rc --->|

                        + Satellite
                       /|
                      / |
                     /  |
                    / b |
                   /-\  +
                  /   * |     __Earth surface
                 /  *   |    /
                / *_----+----__
               +                +
                *              *
                 *            *
                  *   2*a   *
                   *  ___  *
                    *_   _*
                     *   *
                      * *
                       * Earth center
```

Figure 13: Satellite coverage estimation

x The vertical projection of satellite to Earth

Re The radius of the Earth, Re=6378(km)

As The altitude of a satellite

Rc The radius (arc length) of the coverage, or, the arc length of hexagon center to its 6 vertices. Rc=Re*(a*pi)/180

a The cap angle for the coverage area (the RC arc). a = arccos((Re/(Re+As))*cos(b))-b.

b The least elevation angle that a ground station or a terminal can communicate with a satellite, b = 35 degree.

Ns The minimum number of satellites on one orbit plane, it is equal to the number of the satellite's vertical projection on Earth, so, Ns = 180/(a*cos(30))

No The minimum number of orbit (with same inclination), it is equal to the number of the satellite orbit's vertical projection, so, No = 360/(a*(1+sin(30)))

For a example of two type of satellite LEO and VEO, the coverages are calculated as in Table 1:

Parameters	VLEO1	VLEO2	LEO1	LEO2
As(km)	335.9	450	1100	1150
a(degree)	3.907	5.078	10.681	11.051
Rc(km)	435	565	1189	1230
Ns	54	41	20	19
No	62	48	23	22

Table 1: Satellite coverage estimation for LEO and VLEO examples

A.3. Real Deployment of LEO and VLEO for Satellite Network

Obviously, the above orbit parameter setup is not optimal since the sky in the polar areas will have the highest density of satellite.

In the real deployment, to provide better coverage for the areas with denser population, to get redundance and better signal quality, and to make the satellite distance within the range of inter-satellite communication (2000km [Laser-communication-range]), more than the minimum number of satellites are launched. For example, different orbit planes with different inclination/altitude are used.

Normally, all satellites are grouped by orbit planes, each group has a number of orbit planes and each orbit plane has the same orbit parameters, so, each orbit in the same group will have:

1. The same altitude

2. The same inclination, but the inclination is less than 90 degrees. This will result in the empty coverage for polar areas and better coverage in other areas. See the orbit picture for phrase 1 for [StarLink].

3. The same number of satellites

4. The same moving direction for all satellites

The proposed deployment of SpaceX can be seen in [SpaceX-Non-GEO] for StarLink.

The China constellation deployment and orbit parameters can be seen in [China-constellation].

Appendix B. Communications for Satellite Constellation

Unlike the communication on ground, the communication for satellite constellation is much more complicated. There are two mobility aspects, one is between ground-station and satellite, another is between satellites.

In the traditional mobility communication system, only terminal is moving, the mobile core network including base station, front haul and back haul are static, thus an anchor point, i.e., PGW in 4G or UPF in 5G, can be selected for the control of mobility session. Unfortunately, when satellite constellation joins the static network system of Internet on ground, there is no such anchor point can be selected since the whole satellite constellation network is moving.

Another special aspect that can impact the communication is that the fast moving speed of satellite will cause frequent changes of communication peers and link states, this will make big challenges to the network side for the packet routing and delivery, session control and management, etc.

B.1. Dynamic Ground-station-Satellite Communication

All satellites are moving and will lead to the communication between ground station and satellite can only last a certain period of time. This will greatly impact the technologies for the satellite networking. Below illustrates the approximate speed and the time for a satellite to pass through its covered area.

In Table 2, VLEO1 and LEO3 have the lowest and highest altitude respectively, VLEO2 is for the highest altitude for VLEO. We can see that longest communication time of ground-station-satellite is less than 400 seconds, the longest communication time for VLEO ground-station-satellite is less than 140 seconds.

The "longest communication time" is for the scenario that the satellite will fly over the receiver ground station exactly above the head, or the ground station will be on the diameter line of satellite coverage circular area, see Figure 12.

Re The radius of the Earth, Re=6378(km)

As The altitude of a satellite

AL The arc length(in km) of two neighbor satellite on the same orbit plane, AL=2*cos(30)*(Re+As)*(a*pi)/180

SD The space distance(in km) of two neighbor satellite on the same orbir plane, SD=2*(Re+As)*sin(AL/(2*(Re+As))).

V the velocity (in m/s) of satellite, V=sqrt(G*M/(Re+As))

G Gravitational constant, G=6.674*10^(-11)(m^3/(kg*s^2))

M Mass of Earth, M=5.965*10^24 (kg)

T The time (in second) for a satellite to pass through its cover area, or, the time for the station-satellite communication. T= ALs/V

Parameters	VLEO1	VLEO2	LEO1	LEO2	LEO3
As(km)	335.9	450	1100	1150	1325
a(degree)	3.907	5.078	10.681	11.051	12.293
AL(km)	793	1048	2415	2515	2863
SD(km)	792.5	1047.2	2404	2503.2	2846.1
V(km/s)	7.7	7.636	7.296	7.272	7.189
T(s)	103	137	331	346	398

Table 2: The time for the ground-station-satellite communication

B.2. Dynamic Inter-satellite Communication

B.2.1. Inter-satellite Communication Overview

In order to form a network by satellites, there must be an inter-satellite communication. Traditionally, inter-satellite communication uses the microwave technology, but it has following disadvantages:

1. Bandwidth is limited and only up to 600M bps [Microwave-vs-Laser-communication].

2. Security is a concern since the microwave beam is relatively wide and it is easy for 3rd party to sniff or attack.

3. Big antenna size.

4. Power consumption is high.

5. High cost per bps.

Recently, laser is used for the inter-satellite communication, it has following advantages, and will be the future for inter-satellite communication.

1. Higher bandwidth and can be up to 10G bps [Microwave-vs-Laser-communication].

2. Better security since the laser beam size is much narrower than microwave, it is harder for sniffing.

3. The size of optical lens for laser is much smaller than microwave's antenna size.

4. Power saving compared with microwave.

5. Lower cost per bps.

The range for satellite-to-satellite communications has been estimated to be approximately 2,000 km currently [Laser-communication-range].

From Table 2, we can see the Space Distance (SD) for some LEO (altitude over 1100km) are exceeding the celling of the range of laser communication, so, the satellite and orbit density for LEO need to be higher than the estimation values in the Table 1.

Assume the laser communication is used for inter-satellite communication, then we can analyze the lifetime of inter-satellite communication when satellites are moving. The Figure 14 illustrates the movement and relative position of satellites on three orbits. The inclination of orbit planes is 90 degrees.

```
                    + North pole
                   /|\
                  | s |
                 s  |  s
                /   s   \
                s   |   s
                |   s1  |
                s4  |   s6
                |   s2  |   -------- Equator
                s5  |   s7
                |   s3  |
                s   |   s
                \   s   /
                 s  |  s
                  | s |
                   \|/
                    +  South pole
```

Figure 14: Satellite movement

There are four scenarios:

1. For satellites within the same orbit
   The satellites in the same orbit will move to the same direction with the same speed, thus the interval between satellites is relatively steady. Each satellite can communicate with its front and back neighbor satellite as long as satellite's orbit is maintained in its life cycle. For example, in Figure 14, s2 can communication with s1 and s3.

2. For satellites between neighbor orbits in the same group at non-polar areas
   The orbits for the same group will share the same orbit altitude and inclination. So, the satellite speed in different orbit are also same, but the moving direction may be same or different. Figure 15 illustrates this scenario. When the moving direction is the same, it is similar to the scenario 1, the relative position of satellites in different orbit are relatively steady as long as satellite's orbit is maintained in its life cycle. When the moving direction is different, the relative position of satellites in different orbit are un-steady, this scenario will be analyzed in more details in Appendix B.2.2.

3. For satellites between neighbor orbits in the same group at polar areas
   For satellites between neighbor orbits with the same speed and moving direction, the relative position is steady as described in #2 above, but the steady position is only valid at areas other than polar area. When satellites meet in the polar area, the relative position will change dramatically. Figure 16 shows two satellites meet in polar area and their ISL facing will be swapped. So, if the range of laser pointing angle is 360 degrees and tracking technology supports, the ISL will not be flipping after passing polar area; Otherwise, the link will be flipping and inter-satellite communication will be interrupted.

4. For satellites between different orbits in the different group
   The orbits for the different group will have different orbit altitude, inclination and speed. So, the relative position of satellite is not static. The inter-satellite communication can only last for a while when the distance between two satellite is within the limit of inter-satellite communication, that is 2000km for laser [Laser-communication-range], this scenario will be analyzed in more details in Appendix B.2.3

```
                 i+N/2         i+1+N/2       i+2+N/2
               / \           / \           / \
              /   \         /   \         /   \
             S1    \       S2    \       S3    \
            /       S4    /      S5     /      S6
           /         \   /         \   /         \
          /           \ /           \ /           \
          i-1           i             i+1
```

* The total number of orbit planes are N
* The number (i-1, i, i+1,...) represents the Orbit index
* The bottom numbers (i-1, i, i+1) are for orbit planes on which satellites (S1, S2, S3) are moving from bottom to up.
* The top numbers (i+N/2, i+1+N/2, i+2+N/2) are for orbit planes on which satellites (S4, S5, S6) are moving from up to bottom.

Figure 15: Two satellites with same altitude and inclination (i) move in the same or opposite direction

```
\      /
 P3   P4
  \  /
   \/
   /\
  /  \
 P1   P2
/      \
```

* Two satellites S1 and S2 are at position P1 and P2 at time T1
* S1's right facing ISL connected to S2's left facing ISL
* S1 and S2 move to the position P4 and P3 at time T2
* S1's left facing ISL connected to S2's right facing ISL

Figure 16: Two satellites meeting in the polar area will change its facing of ISL

B.2.2. Satellites on Adjacent Orbit Planes with Same Altitude

For satellites on different orbit planes with same altitude, the estimation of the lifetime when two satellite can communicate are as follows.

Figure 17 illustrates a general case that two satellites move and intersect with an angle A.

```
                                        ^ V2
                                       /
                                      /
                                    +-
                                   /  \ A
                     -------------+----+----> V1
                                 /
                                /
```

Figure 17: Two satellites (speed vector V1 and V2) intersect with angle A

More specifically, for orbit planes with the inclination angle i, Figure 18 illustrates two satellites move in the opposite direction and intersect with an angle 2*i.

```
                                 ^ move from south to north
                            \      /
                             \    /
                              \  /\
                               \/  | A = 2*i
                               /\  |
                              /  \/
                             /    \
                            /      V move from north to south
```

Figure 18: Two satellites with same altitude and inclination (i) intersect with angle A=2*i

Follows are the math to calculate the lifetime of communication. Table 3 are the results using the math for two satellites with different altitudes and different inclination angles.

Dl The laser communication limit, Dl=2000km [Laser-communication-range]

A The angle between two orbit’s vertical projection on Earth. A=2*i

V1 The speed vector of satellite on orbit1

V2 The speed vector of satellite on orbit2

|V| the magnitude of the difference of two speed vector V1 and V2, |V|=|V1-V2|=sqrt((V1-V2*cos(A))^2+(V2*sin(A))^2). For satellites with the same altitude and inclination angle i, V1=V2, so, |V|=V1*sqrt(2-2*cos(2*i))=2V1*sin(i)

T The lifetime two satellites can communicate, or the time of two satellites' distance is within the range of communication, T = 2*Dl/|V|.

i (degree)	80	80	65	65	50	50
Alt (km)	500	800	500	800	500	800
|V| (km/s)	14.98	14.67	13.79	13.5	11.66	11.41
T(s)	267	273	290	296	343	350

Table 3: The lifetime of communication for two LEOs (with two altitudes and three inclination angles)

B.2.3. Satellites on Adjacent Orbit Planes with Different Altitude

For satellites on different orbit planes with different altitude, the estimation of the lifetime when two satellite can communicate are as follows.

Figure 19 illustrates two satellites (with the altitude difference Da) move and intersect with an angle A.

```
                                   ^ V2
                                  /
                                 /
                        -------+  /
                        Da    /| +-
                             / |/  \ A
                    ----------/--+----+----> V1
                              /  /
                                /
                               /
                              /
```

Figure 19: Satellite (speed vector V1 and V2, Altitude difference Da) intersects with Angle A

Follows are the math to calculate the lifetime of communication

Dl The laser communication limit, Dl=2000km [Laser-communication-range]

Da Altitude difference (in km) for two orbit planes

A The angle between two orbit's vertical projection on Earth

Vl The speed vector of satellite on orbit 1

V2 The speed vector of satellite on orbit 2

|V| the magnitude of the difference of two speed vector V1 and V2, |v|=|V1-V2|=sqrt((V1-V2*cos(A))^2+(V2*sin(A))^2)

T The lifetime two satellites can communicate, or the time of two satellites' distance is within the range of communication, T = 2*sqrt(Dl^2-Da^2)/|V|

Using formulas above, below is the estimation for the life of communication of two satellites when they intersect. Table 4 and Table 5 are for two VLEOs with the difference of 114.1km for altitude. (VLEO1 and VLEO2 on Table 2). Table 6 and Table 7 are for two LEOs with the difference of 175km for altitude (LEO2 and LEO3 on Table 2).

Parameters	VLEO1	VLEO2
As(km)	335.9	450
V (km/s)	7.7	7.636

Table 4: Two VLEO with different altitude and speed

A (degree)	0	10	45	90	135	180
|V| (km/s)	0.065	1.338	5.869	10.844	14.169	15.336
T(s)	61810	2984	680	368	282	260

Table 5: Two VLEO intersects with different angle and the life of communication

Parameters	LEO1	LEO2
As(km)	1150	1325
V (km/s)	7.272	7.189

Table 6: Two LEO with different altitude and speed

A (degree)	0	10	45	90	135	180
|V| (km/s)	0.083	1.263	5.535	10.226	13.360	14.461
T(s)	47961	3155	720	390	298	276

Table 7: Two LEO intersects with different angle and the life of communication

Appendix C. Change Log

* Initial version, 07/03/2021

* 01 version, 10/20/2021

* 02 version, 2/13/2022

* 03 version, 7/5/2022

* 04 version, 1/4/2023

* 05 version, 7/5/2023

* 06 version, 1/3/2024

Authors' Addresses

Lin Han (editor)
Futurewei Technologies, Inc.
2330 Central Expy
Santa Clara, CA 95050,
United States of America
Email: lhan@futurewei.com

Richard Li
Futurewei Technologies, Inc.
2330 Central Expy
Santa Clara, CA 95050,
United States of America
Email: rli@futurewei.com

Alvaro Retana
Futurewei Technologies, Inc.
2330 Central Expy
Santa Clara, CA 95050,
United States of America
Email: alvaro.retana@futurewei.com

Meiling Chen
China Mobile
32, Xuanwumen West
BeiJing 100053
China
Email: chenmeiling@chinamobile.com

Li Su
China Mobile
32, Xuanwumen West
BeiJing 100053
China
Email: suli@chinamobile.com

Tianji Jiang
China Mobile
1525 McCathy Blvd.
Milpitas, CA 95035,
United States of America
Email: tianjijiang@chinamobile.com

INTAREA Working Group
Internet-Draft
Intended status: Best Current Practice
Expires: 4 August 2022

Tony. Li
Juniper Networks

31 January 2022

On Higher Levels of Address Aggregation
draft-li-int-aggregation-00

Abstract

Routing and addressing are inexorably tied, and the scalability of the routing system is wholly dependent on the abstraction and allocation of the address space. The addressing architecture for the Internet was set forth in [RFC1518], [RFC4632], and [RFC4291]. These describe how address aggregation can be performed at the ISP and local level.

Address allocation and assignment procedures by the Regional Internet Registries (RIRs) have created large address blocks. This creates an opportunity for further aggregation above the ISP level without any change to existing allocations.

This document discusses issues regarding address aggregation above the ISP level, for continents or regions, thereby providing additional address space aggregation and efficiency in the routing system. Small changes to address allocation policies can help to ensure futher aggregations and improvements in routing efficiency. Some of these concepts were discussed as part of the Routing and Addressing meetings [RFC1380] and extended further here.

This document is not advocating geographical assignment below the continental level. That has been thoroughly discussed previously.

Status of This Memo

This Internet-Draft will expire on 4 August 2022.

Table of Contents

1. Introduction

The Internet depends upon the efficiency and scalability of the routing system. Without effective routing, no traffic can be delivered. Ensuring that routing scales is key to the Internet architecture. History has shown that the architectural changes made as part of [RFC1518] and [RFC4632] have been extremely effective.

However further improvements are possible. While prefix aggregation at the ISP level has helped provide good routing efficiency, more aggregation is possible. This document discusses how aggregation could be performed above the provider level, forming aggregates at the regional or continental level based on the address allocations that have already been performed.

This document also suggests ways that ICANN and the RIRs can change address allocation procedures to enable better future regional and continental aggregation. These changes would have no perceptible effect on ISP or end-site address allocations, and would simply cause the allocations to come from different address blocks from the same RIR.

IPv4 and IPv6 addresses and prefixes are used throughout this document as examples. The concepts presented here apply equally to both address spaces.

2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Routing and Addressing

The routing subsystem of the Internet is responsible for discovering paths from any point on the Internet to any other point. The results of the routing system are paths instantiated in the forwarding plane of routers throughout the network, resulting in end-to-end connectivity.

For routing to function, there must be specific names for hosts. The architecture of the namespace for these names is a critical decision, as these names will have global scope. When we also use these names as a binding to a location in the network, we call these names ’addresses’ and the namespace that they are taken from an ’address space’.

Scalability is a central concern for routing. Each item of information that routing must propagate around the network requires processing power and memory for storage throughout the network. This scales with the size of the network. If routing also scaled linearly with the number of hosts, then the cost of running the routing system would grow as the size of the network times the number of hosts, which is clearly problematic. For this reason, we cannot have a routing subsystem that just carries individual host routes.

4. Abstraction and Aggregation

Instead, we seek to define groups of hosts and treat them together as a single abstraction, commonly known as a 'prefix'. We call the process of combining addresses together into a prefix 'aggregation'. Under some circumstances, prefixes themselves may also be aggregated to form another prefix, resulting in a recursive structure. If prefix A is a proper subset of prefix B, we say that A is 'more specific' than B and that B is 'less specific' than A.

We can then define the routing efficiency of a specific prefix as the cost of carrying that prefix, plus all of its more specifics, integrated across the entire network, and divided by the number of host addresses subsumed by the prefix.

It is well known that abstraction obscures important detail and that abstraction in routing can cause sub-optimal paths, resulting in extra hops, wasted bandwidth, and managerial difficulties. As a result, there will always be a trade-off between scalability and optimality when introducing abstractions.

When optimality is paramount and simple reachability is insufficient, the routing subsystem has additional mechanisms that allow network operators to make different path selection choices, sometimes intentionally ignoring or explicitly working against abstraction. We call this broad set of mechanisms 'traffic engineering'.

5. Abstraction Boundaries

Abstractions have three different boundaries that we will be concerned with:

Abstraction Administrative Boundary
: An abstraction's administrative boundary occurs at the topological interface between the abstraction's administratively controlled network and other administrations.

Abstraction Naming Boundary
An abstraction's naming boundary is the topological container of all of the host addresses within that abstraction.

Abstraction Action Boundary
An abstraction's action boundary is the topological container where the abstraction has an effect on routing's path computation.

In simple cases of abstraction, these boundaries are aligned. For example, consider a university that has been assigned the prefix 128.125/16. It has a pair of routers that interface to its two ISP's and it advertises that prefix to both of it's ISPs and no more specifics. The entire university's infrastructe utilizes this one prefix. In this case, the administrative, naming, and action boundary all occur between the university's router's and the ISPs'.

As a more interesting example, consider an ISP X that has been assigned the prefix 2001:1234::/32. For its own internal purposes, the ISP chooses to partition this prefix and assigns 2001:1234:5600::/40 to city C, which is an IGP area in the ISP's network. For traffic engineering purposes, ISP X also advertises more specifics of the city C prefix to ISP Y, but not to ISP Z. The more specifics are constrained so that ISP Y cannot propagate them to its neighbors (e.g., using the BGP NO_EXPORT community).

```
                                ISP X

                              XXXXXXXXXXXXXX           ISP Y
                           XXX             XX
      ISP Z               XX                 XX
                        XX                    X       YYYYYYYYYY
                      XX                      X     YYY         YY
    ZZZZZZ          XX                        X---YYY            YY
   ZZ     ZZZZ     XX                         X   Y               YY
  ZZ         ZZ----XX              CCCCCC    XX   YY               Y
 ZZ           Z     XX            CC    CC   X     YY              Y
ZZ            Z       X          CC      CC  X------YY            YY
ZZZ          ZZ        X         C        C XX        YYY        YY
   ZZZZ    ZZZ          XX       C        CXX            YYYYYYYYY
       ZZZZZ             X       C      CCCX
                          XX      CCCCCCCX
                            XXXXXXXXXX
```

The administrative boundary for the city C prefix is now ISP X's entire network. The naming boundary is city C itself, but the action boundary includes ISP X and Y, but not ISP Z.

Placing the abstraction action boundary at an acceptable location is key. If the action boundary is not located correctly, then it may allow prefixes to propagate too far, unnecessarily damaging routing efficiency, or it may not allow prefixes to propagate far enough, causing traffic to take suboptimal paths.

```
R1-------R2
 |         \
 |          \
 |           \
 |            R3-----R6
 |           /
 |          /
 |         /
R4-------R5
```

In the above figure, suppose that prefix A is advertised by router R1 and prefix B is advertised by R4. If aggregation is performed at router R6, then that is inefficient. Router R3 is the next hop for both prefix A and B, so if R3 had aggregated A and B, R6 would have less state to carry. Conversely, if aggregation happened at routers R2 and R5, then R3 would likely make a suboptimal forwarding decision, possibly sending traffic for prefix B via R2 instead of the optimal R5.

From this, we can see that the perfect location for the action boundary is the first point where all paths would merge. For pragmatic reasons, it’s easier to put action boundaries at administrative boundaries. Thus, the optimal action boundary would be the network boundary after the path merge point.

6. Regional Aggregation

Abstraction can be used to help aggregate routing information above the ISP level, as well as below. While this is currently not commonly done, it should be considered as a means of further reducing the size of the global routing table and improving routing efficiency.

Address allocation today is performed at the behest of the Internet Assigned Numbers Authority (IANA), currently delegated to five Regional Internet Registries (RIRs): AFRINIC, APNIC, ARIN, LACNIC, RIPE NCC. Those registries are assigned large blocks of address space that they in turn assign to ISPs and other networks. We can take advantage of the topological connectivity within a region and consider aggregation across ISPs.

Current address allocation policies have given each RIR a number of address blocks. While aggregation of these blocks is not a requirement, they are a convenient starting point. Any ISP can look at their routing table and determine whether or not they are within the optimal abstraction boundary of such a prefix. If not, then that prefix is good candidate for aggregation by the ISP.

While the ISP could aggregate the prefix on routers where it would be advertised to other ASes, it might be more beneficial if the ISP performed the aggregation on the routers where the more specifics enter the network. This saves the ISP from carrying the more specific prefixes internally, which is an economic incentive to aggregate.

There are several important considerations when contemplating this. Since more specifics are not carried internally, traffic for the prefix will find the closest exit point. This is sometimes known as 'hot potato routing'. This may or may not be compatible with the ISP's routing policy. Aggregation may also cause global traffic patterns for the prefix to change. Since more specific prefixes are preferred, if an ISP starts advertising an aggregate to its neighbors and those neighbors are still receiving more specifics from other sources, the traffic for the prefix may be diverted away from the ISP through the more specifics. This may also be in conflict with the ISP's routing policy. If not, then it may actually be beneficial to the ISP as they are still offering transit for that prefix, but not actually carrying any traffic for it. This is another economic incentive to aggregate.

ISPs that wish to perform this aggregation are able to do so unilaterally. No coordination with other entities is required, although prior discussion beforhand might avoid operational issues.

A common traffic engineering practice today is to advertise more specifics of a prefix today at different entry points with different path attributes in an attempt to influence inbound traffic patterns. This has proven very effective and become very common. This is very harmful to overall routing efficiency because the more specifics propagate very widely, far beyond their point of actual impact. Regional aggregation could help to recapture much of this inefficiency.

7. Continental Aggregation

Continental aggregation was previously discussed in [RFC1518]. Today, some RIRs are closely aligned with a continent, so continental and regional aggregation are aligned. However, for RIRs that serve more than one continent, there is a natural opportunity for additional aggregation. Continental boundaries are commonly aligned with a few very expensive links. In graph theoretic terms, these links form a natural cut-set, making a continent a possible valuable abstraction. Since routing across the cut-set is likely to be expensive, providers will want to optimize it, however, it is also likely that the optimal abstraction action boundary for a continental abstraction is just beyond the links in the cut-set.

To maximize the ability to create continental abstractions, RIRs that serve more than one continent should consider allocating blocks by continent and delegating from within those blocks to entities within those continents.

8. ICANN Considerations

The current IPv6 Global Unicast Address Assignments are found in [v6guaa]. While some hierarchical allocation is being practiced, there have been numerous blocks allocated that are not aggregatable.

This document recommends that ICANN review their policy on global address allocation and consider reserving shorter prefixes for RIRs, such as /4's or /8's, and then making further allocations to the RIRs from these shorter prefixes. This would, in the distant future, enable more opportunities for regional aggregation.

9. RIR Considerations

This document recommends that RIRs optimize their address allocation policies to maximize the opportunities for regional and continental aggregation.

RIRs that serve more than one continent should consider allocating address blocks per continent and delegating from those blocks to providers on that continent.

Allocations to multi-regional providers should be done from separate blocks than regional providers. This will maximize the opportunity to aggregate regional providers.

10. ISP Considerations

This document encourages ISPs to aggregate more, both to help the overall routing efficiency of the overall Internet also to help contain local costs, as well as churn from oscillating more specific prefixes and accidental deaggregation.

11. IANA Considerations

The author would like to take this opportunity to thank IANA for years of selfless service. This document makes no further requests of IANA.

12. Security Considerations

This document creates no new security issues.

13. Acknowledgments

The author would like to acknowledge the contributions of J. Noel Chiappa to routing architecture in general and for his specific insights in defining the abstraction naming boundary and the abstraction action boundary.

14. References

14.1. Normative References

[RFC1518] Rekhter, Y. and T. Li, "An Architecture for IP Address Allocation with CIDR", RFC 1518, DOI 10.17487/RFC1518, September 1993, <https://www.rfc-editor.org/info/rfc1518>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <https://www.rfc-editor.org/info/rfc4291>.

[RFC4632] Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan", BCP 122, RFC 4632, DOI 10.17487/RFC4632, August 2006, <https://www.rfc-editor.org/info/rfc4632>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

14.2. Informative References

[RFC1380] Gross, P. and P. Almquist, "IESG Deliberations on Routing and Addressing", RFC 1380, DOI 10.17487/RFC1380, November 1992, <https://www.rfc-editor.org/info/rfc1380>.

[v6guaa] IANA, "Address Family Numbers", November 2019, <https://www.iana.org/assignments/ipv6-unicast-address-assignments/ipv6-unicast-address-assignments.xhtml>.

Author's Address

Tony Li,
Juniper Networks
1133 Innovation Way
Sunnyvale, California 94089
United States of America

Email: tony.li@tony.li

Internet Engineering Task Force
Internet-Draft
Intended status: Best Current Practice
Expires: 23 February 2023

S.D. Schoen
J. Gilmore
D. Täht
IPv4 Unicast Extensions Project

22 August 2022

The IETF Will Continue Maintaining IPv4
draft-schoen-intarea-ietf-maintaining-ipv4-01

Abstract

This document confirms the consensus of the IETF that IETF and its affiliated working groups will continue to maintain the IPv4 protocol family.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

It might seem surprising to imagine the IETF ceasing to maintain the IP version 4 protocol suite which it has led to worldwide success. However, just such a change has been advanced in the past. [ipv6-ietf], and the issue continues to produce confusion and uncertainty during discussions of unrelated technical questions.

This document explicitly confirms the IETF's prior practice of maintaining IPv4 in the interest of its user and implementer community, and affirms that doing so is the considered and continued consensus of the IETF.

IETF actions or inactions whose motivation or effect is to fail to maintain IPv4 disrupt the ordinary practice of IETF working groups and functions, and bear a burden of justification.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. The Evolution of the Internet

Since version 4 of the Internet Protocol (IPv4) was created as an experiment in 1981 in [RFC0791], the Internet has grown enormously to become a vital resource for humanity.

IPv4 is easily the most popular network-layer protocol in the world, carrying the majority of the world's commercial data traffic, as well as the majority of traffic on private intranets. For more than 40 years, the IPv4 protocol has formed the central common agreement that has enabled technologists, entrepreneurs, and policymakers to build a worldwide network-of-networks containing billions of nodes, and serving billions of users. Use of the IPv4 protocol remains a necessity for the vast majority of Internet nodes today.

The Internet has grown by many orders of magnitude in physical size, bandwidth, and traffic volume. It has increased dramatically in organizational, administrative, and operational complexity. With that growth, the original specifications and understandings underlying IPv4 and its related protocol suite have required adaptation and adjustment. Congestion control, security, address allocation, routing, and many other areas were adjusted gradually over time as the world gained experience in managing and growing a single worldwide network that puts every user only a few hundred milliseconds away from every other user. Most such adjustments have been done with gradual, compatible changes. On a few occasions, this adjustment required protocol changes in every node on the Internet, such as the transition to CIDR [RFC1519] in the 1990s, or in every node that supported a particular protocol, such as the removal for security reasons of SSL v3 in [RFC7568].

3. Internet Evolution and the IETF

To promote the reliability and stability of the Internet, the user and implementer base for IPv4 have gathered together for coordination, guidance in its use, and both technical and policy development and evolution. Since 1986, the Internet Engineering Task Force (IETF) has been the home of that community gathering.

Discussions in the IETF community have exposed a variety of attitudes toward the continued existence of IPv4 and toward occasions and circumstances in which the IETF is called upon to maintain, support, and coordinate the use of IPv4 and IPv4-based protocols. These occasions will likely continue to arise as the Internet continues to evolve.

This document confirms the consensus among IETF members and IETF leadership that the IETF will continue to maintain the IPv4 protocol suite.

4. Challenges to the IETF's Role as Maintainer of IPv4

Some IETF participants would prefer that the IETF act to hasten the day when IPv6 would completely replace IPv4. Others see an ongoing role for both protocols. These differing points of view play out in the IETF in various ways.

The most radical position the authors have encountered views some limitations and problems with IPv4 as actively beneficial, because its proponents view increased pain or cost of using IPv4 as encouraging people to adopt IPv6 as a substitute.

Holders of this position have suggested that the IETF should sometimes deliberately allow breakage or degradation in the IPv4 protocol. [nat-undocumented] Or that IETF should declare that IPv4 has "historic" status and should no longer be implemented.[v4historic] They may wish that IETF or some other body could take on the power to actively put an end to use or deployment of IPv4, much as the ARPA funding agency could compel all ARPANET hosts to switch from NCP to IPv4 protocols between 1981 and 1983 [RFC0801]; or how the Defense Communications Agency, as the source of funding for the ARPANET, physically took apart the ARPANET by 1990-02-28 in order to force its users (who were all using IPv4) to switch to connecting via the NSFnet or other more modern networks. [decommission-arpanet]

Other positions do not necessarily view problems with IPv4 as a good thing in themselves. But they promote the view that the total resources available for standardization, coordination, and/or implementation efforts, are inherently limited in such a way that doing any work related to the IPv4 protocol would cause less work to be done on the IPv6 protocol. Multiple people who seem to hold this position respond to requests to improve IPv4 with an objection that "we should not be improving IPv4; we should be deploying IPv6 instead".

The objection takes the form of a false dilemma; that is, it assumes that there are only two possible actions, and that taking one prevents the other from being taken. But in actual fact, it is possible to do neither, either, or both of those actions; they are unrelated. It is exceedingly unlikely that if this objection prevents someone from improving IPv4, as it did in 2008 during discussions of the unicast use of the 240/4 address block in the intarea Working Group [FLM], for example, that they would immediately turn their efforts to deploying IPv6. And most of the work required for increased deployment of IPv6 involves neither coordinating new standards nor implementing IPv6; IPv6 is already widely standardized and implemented.

Those expressing either of these views may worry that ongoing IPv4 work provides an "excuse" for decision makers, such as network operators, to delay IPv6 adoption, because they will seemingly perceive the IETF's blessing for doing so, because they will perceive IPv4 as not obsolete, or because specific technical problems they would otherwise encounter with IPv4 will be mitigated.

5. Neglecting IPv4 is Not Our Transition Strategy

A strong consensus exists to continue the IETF's work in support of IPv6 and to promote its adoption [RFC6540]. However, no consensus has been found to actively discourage IPv4. Instead, one serious attempt to form such a consensus was definitively rejected.

The IETF chartered a working group, "sunset4", which existed between 2012 and 2018. Its original remit included:

> The IETF is committed to the deployment of IPv6 to ensure the evolution of the Internet. However, the IPv4-only components of the Internet must continue to operate as much as possible during the transition from IPv4 to IPv6.
>
> The Working Group will standardize technologies that facilitate the graceful sunsetting of the IPv4 Internet in the context of the exhaustion of IPv4 address space while IPv6 is deployed.

A year later, the charter was revised to say:

> In order to fully transition the Internet to IPv6, individual applications, hosts, and networks that have enabled IPv6 must also be able to operate fully in the absence of IPv4. The Working Group will point out specific areas of concern, provide recommendations, and standardize protocols that facilitate the graceful "sunsetting" of the IPv4 Internet in areas where IPv6 has been deployed. This includes the act of shutting down IPv4 itself, as well as the ability of IPv6-only portions of the Internet to continue to connect with portions of the Internet that remain IPv4-only.

A participant in this working group produced an Internet-Draft [v4historic] entitled "IPv4 Declared Historic", whose abstract was the single sentence, "IPv4 has been superseded by IPv6, and is therefore Historic." It stated:

> The use of IPv4 is deprecated. The term "deprecated" is used to indicate a feature, characteristic, or practice that should be avoided, in this case because it is being superseded by a newer protocol. The term does not indicate that the practice is harmful, but that there will be no further development in IPv4...

This draft was discussed by the working group (with some of the results documented in its author's blog entry, [IPv4-NOT-Declared-Historic]. The working group's co-chair remarked on the mailing list, "That's part of the reason this discussion is happening - it's looking for rough consensus from the IETF that we are done making changes to IPv4." [wes-george-sunset4-2016-03-22] A later draft [ipv6-support] explicitly stated, "new functionality should be developed in IPv6, and IETF effort SHOULD NOT be spent retrofitting features into the legacy protocol."

Eventually an evolved draft [ipv6-ietf] went through a Working Group last call. The document was entitled "IETF: End Work on IPv4" and its abstract was:

> The IETF will stop working on IPv4, except where needed to mitigate documented security issues, to facilitate the transition to IPv6, or to enable IPv4 decommissioning.

It specifically declared: "The IETF will not initiate new IPv4 extension technology development." The WG chair officially summarized it as a request for "IETF to stop working on IPv4 except for security issues." He noted that

> The working group last call got strong support but only very few people participated in the last call. Given the relative inactivity of the working group for quite a while, it is possible that the mailing list is not watched. Given that this document has widespread implications to any work within IETF, the real wide review should be done IETF wide.

(Only three people had responded to the WG Last Call.) A Routing Directorate reviewer, Ron Bonica, reviewed it for the IESG, saying it "Needs Work", and noted, "Given that the majority of Internet traffic still runs over IPv4, is that a good idea?" as well as asking, "Does this mean that RFC 791 cannot be updated? Or does it mean more than this?"

The draft went through an IETF Last Call, and generated a lot of controversy. While some participants were supportive, other participants expressed concerns that the IETF was proposing to abdicate a vital responsibility for maintaining one of the most important and widely-used technologies in the world. If the IETF

would not do this work, some suggested, other standards-development organizations would be compelled to take it up instead -- but they would likely be less qualified to do so because they would have far less historic expertise and experience with the technology, and would also not necessarily share other IETF values such as openness. The result was that the draft expired without attaining IETF-wide consensus. The IESG counted this as a rejection.

This history demonstrates that there was no IETF-wide consensus to neglect the maintenance of IPv4. However, it did not demonstrate the opposite consensus either, but left that question for a later day.

This document is in some sense that opposite proposal, demonstrating that there IS an IETF-wide consensus to continue to maintain IPv4.

6. IPv4 Requires Ongoing Maintenance

As a protocol in use in billions of nodes, IPv4 continues to evolve. New situations and new realizations have resulted in numerous proposals for protocol modifications that have reached consensus in the recent past. Below are several examples of recent IETF work that contributed to this evolution. Each one identifies the IETF working group in which it was approved.

* In 2022, [RFC9293] restated the definition of the TCP protocol, giving detailed advice to implementers on the interactions between the IP and TCP layers, including IPv4-specific considerations. (WG tcpm)

* In 2020, [RFC8899] defined a new way to detect path MTUs, both for datagram protocols and stream protocols, without the use of ICMP error messages. (WG tsvwg)

* In 2020, [RFC8815] deprecated any-source multicast packets for interdomain uses, and recommended application support of source-specific multicast. (WG mboned)

* In 2017, [RFC8029] defined a way to "ping" the data plane of an MPLS network that carries IPv4 or IPv6 traffic. The details changed the behavior of IPv4 routers which receive certain UDP packets that use destination addresses in the 127/8 range. (WG mpls)

* In 2016, [RFC7766] updated the host requirements for DNS resolvers and servers, requiring them to implement TCP as well as UDP. It also changed various other requirements to improve DNS implementations' compatibility with larger resource records used for DNS Security. It superseded a similar update in [RFC5966] in 2010. (WG dnsop)

* In 2014, [RFC7413] created the TCP Fast Open option to allow reduced latency in some applications of TCP (both v4 and v6) such as http GET requests or DNS lookups. (WG tcpm)

* The IPv4 header's "ID" field is used in fragmentation and reassembly of IP packets. Under the original specifications for IPv4, this 16-bit field had to have a unique value in every packet, that would remain unique for the lifetime of the packets in transit between a particular pair of source and destination addresses. With a potential packet lifetime of about 2 minutes (120 seconds) during routing flaps, and typical packet sizes, this limited transmission speeds to only about 6 megabits per second. In 2013, [RFC6864] updated the specifications for this field in the header of every IPv4 packet, to allow for implementations that meet the standards and can operate at gigabit and greater speeds. (WG intarea)

* Also in 2013, [RFC6918] formally deprecated some ICMP message types that had become obsolete in practice, such as the Information Request, Information Reply, Address Mask Request, and Address Mask Reply messages that have been replaced by DHCP. (This was an individual submission to the IESG and did not go through any working group.)

* Also in 2013, [RFC6762] defined Multicast DNS and required that DNS queries for names of the form "foo.local" must be sent to a link-local multicast address. This protocol is part of the IETF's Zeroconf effort to reduce manual configuration of IPv4 and IPv6 networks. (WG dnsext)

* In 2012, [RFC6528] standardized a revised algorithm for generating Initial Sequence Numbers in the TCP protocol, which reduces the chance of an off-path attacker guessing those sequence numbers. This makes some kinds of automated attacks on network connections harder to accomplish. (WG tcpm)

* Also in 2012, [RFC6633] deprecated the ICMP Source Quench mechanism for congestion control, which has not been reliably used in the Internet since the 1990s. (WG tsvwg)

* In 2011, [RFC6093] clarified the specifications and limitations of the TCP "Urgent Data" mechanism. (WG tcpm)

* Also in 2011, [RFC6298] changed how the TCP retransmission timer is calculated, for recovering from a failure by the receiver to acknowledge sent data. (WG tcpm)

In recent years, various other draft proposals did not reach consensus, partly due to confusion about the proper role of the IETF in working on IPv4-related protocols. Had this IPv4-maintenance issue been resolved independently, as proposed in this document, those proposals would have had a better chance of reaching consensus on their technical merits, rather than being pulled into unrelated issues about IPv4 versus IPv6.

In addition, various errata have been noted in the IPv4 standards, including significant technical errors in [RFC0791] noted in 2016 and 2021.

7. The IETF is Uniquely Positioned to Maintain IPv4

The Internet Engineering Task Force (IETF) was initially an informal work group of government-funded grant recipients involved with building the Internet technologies. It has grown into a major standards development organization, while retaining its traditional values such as transparency, consensus, and informality. As changes in protocol specifications and operational practices have been needed, the IETF has provided a forum where these can be discussed, agreed upon, and publicized.

Implementers and operators care a great deal about the IETF's recommendation for the technologies that were developed here, and questions affecting interoperability in IPv4 continue to arise. There is no other organization that would be as clearly empowered to do this work as the IETF. If the IETF actively neglected to coordinate IPv4 work, it would squander some of the trust that the community places in it.

While predicting is hard, especially about the future, decades of experience suggest that the IPv4 and IPv6 protocols will continue to co-exist for the foreseeable future. Increased IPv6 adoption by individual sites does not typically eliminate those sites' need for continued use of IPv4 services in reaching parts of the Internet that do not use IPv6. In addition, even if IPv6 soon becomes the predominant network-layer protocol on the global Internet, IPv4 is likely to remain important on LANs and private networks, with corresponding needs for suppliers and operators to continue to coordinate interoperability.

Implementers and operators continue to look to the IETF as the authority for IPv4 standardization efforts. The IETF is better-positioned than any other organization to play this role both because of its conspicuous position in evolving IPv4 and IPv6, and because of its deep institutional knowledge and broadly representative participation model.

Since IPv4 is still the world's most-used networking protocol, many parties will look for a standards-development organization to coordinate its ongoing standardization and to maximize interoperability among systems using it. Though the IETF could attempt to make IPv4 less attractive by deprecating it or refusing to maintain it, it's not clear that this course of action would lead to faster IPv6 adoption. Instead, it might encourage non-IETF organizations to take up responsibility for IPv4's maintenance, which could lead to IPv4 being a stronger competitor against IPv6, or greater fragmentation in Internet standards development, as the location of the authority to define and coordinate IPv4 would no longer be clear.

8. The IETF Continues to Support IPv4 as Well as IPv6

There are many reasons to encourage IPv6 adoption and support everywhere on the Internet. This document does not change the IETF's policy in favor of IPv6, but merely makes it clear that the IETF intends to continue fully maintaining and supporting IPv4, in addition to continuing the promotion and evolution of IPv6.

9. IANA Considerations

This document makes no change to IANA's existing role in providing and maintaining IPv4-related registries.

10. Security Considerations

The IETF's ongoing responsibility for IPv4 includes remaining apprised of emerging security threats to IPv4 users and applications, and developing or publicizing guidance for how to mitigate these threats. In some cases, the IETF may modify existing and deployed protocols as required or useful in adjusting to security concerns. [RFC2644]

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC6540] George, W., Donley, C., Liljenstolpe, C., and L. Howard, "IPv6 Support Required for All IP-Capable Nodes", BCP 177, RFC 6540, DOI 10.17487/RFC6540, April 2012, <https://www.rfc-editor.org/info/rfc6540>.

11.2. Informative References

[decommission-arpanet] Living Internet, "NSFNET -- National Science Foundation Network", <https://livinginternet.com/i/ii_nsfnet.htm>.

[FLM] Fuller, V., Lear, E., and D. Meyer, "Reclassifying 240/4 as usable unicast address space", Work in Progress, Internet-Draft, draft-fuller-240space-02, 25 March 2008, <https://datatracker.ietf.org/doc/html/draft-fuller-240space-02>.

[IPv4-NOT-Declared-Historic] Howard, L., "IPv4 NOT Declared Historic", 29 April 2016, <https://web.archive.org/web/20200708045029/http://www.wleecoyote.com/blog/ietf95-sunset4.htm>.

[ipv6-ietf] Howard, L., "IETF: End Work on IPv4", Work in Progress, Internet-Draft, draft-ietf-sunset4-ipv6-ietf-01, 18 September 2017, <https://datatracker.ietf.org/doc/html/draft-ietf-sunset4-ipv6-ietf-01>.

[ipv6-support] George, W. and L. Howard, "IPv6 Support Within IETF work", Work in Progress, Internet-Draft, draft-george-ipv6-support-03, 30 September 2014, <https://datatracker.ietf.org/doc/html/draft-george-ipv6-support-03>.

[nat-undocumented] Perlman, R., "Keynote: Do the Wrong Thing! (starting from 41:50 within the presentation)", 25 February 2022, <https://www.youtube.com/watch?v=5D1v42nw25E>.

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <https://www.rfc-editor.org/info/rfc791>.

[RFC0801] Postel, J., "NCP/TCP transition plan", RFC 801, DOI 10.17487/RFC0801, November 1981, <https://www.rfc-editor.org/info/rfc801>.

[RFC1519] Fuller, V., Li, T., Yu, J., and K. Varadhan, "Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy", RFC 1519, DOI 10.17487/RFC1519, September 1993, <https://www.rfc-editor.org/info/rfc1519>.

[RFC2644] Senie, D., "Changing the Default for Directed Broadcasts in Routers", BCP 34, RFC 2644, DOI 10.17487/RFC2644, August 1999, <https://www.rfc-editor.org/info/rfc2644>.

[RFC5966] Bellis, R., "DNS Transport over TCP - Implementation Requirements", RFC 5966, DOI 10.17487/RFC5966, August 2010, <https://www.rfc-editor.org/info/rfc5966>.

[RFC6093] Gont, F. and A. Yourtchenko, "On the Implementation of the TCP Urgent Mechanism", RFC 6093, DOI 10.17487/RFC6093, January 2011, <https://www.rfc-editor.org/info/rfc6093>.

[RFC6298] Paxson, V., Allman, M., Chu, J., and M. Sargent, "Computing TCP's Retransmission Timer", RFC 6298, DOI 10.17487/RFC6298, June 2011, <https://www.rfc-editor.org/info/rfc6298>.

[RFC6528] Gont, F. and S. Bellovin, "Defending against Sequence Number Attacks", RFC 6528, DOI 10.17487/RFC6528, February 2012, <https://www.rfc-editor.org/info/rfc6528>.

[RFC6633] Gont, F., "Deprecation of ICMP Source Quench Messages", RFC 6633, DOI 10.17487/RFC6633, May 2012, <https://www.rfc-editor.org/info/rfc6633>.

[RFC6762] Cheshire, S. and M. Krochmal, "Multicast DNS", RFC 6762, DOI 10.17487/RFC6762, February 2013, <https://www.rfc-editor.org/info/rfc6762>.

[RFC6864] Touch, J., "Updated Specification of the IPv4 ID Field", RFC 6864, DOI 10.17487/RFC6864, February 2013, <https://www.rfc-editor.org/info/rfc6864>.

[RFC6918] Gont, F. and C. Pignataro, "Formally Deprecating Some ICMPv4 Message Types", RFC 6918, DOI 10.17487/RFC6918, April 2013, <https://www.rfc-editor.org/info/rfc6918>.

[RFC7413] Cheng, Y., Chu, J., Radhakrishnan, S., and A. Jain, "TCP Fast Open", RFC 7413, DOI 10.17487/RFC7413, December 2014, <https://www.rfc-editor.org/info/rfc7413>.

[RFC7568] Barnes, R., Thomson, M., Pironti, A., and A. Langley, "Deprecating Secure Sockets Layer Version 3.0", RFC 7568, DOI 10.17487/RFC7568, June 2015, <https://www.rfc-editor.org/info/rfc7568>.

[RFC7766] Dickinson, J., Dickinson, S., Bellis, R., Mankin, A., and D. Wessels, "DNS Transport over TCP - Implementation Requirements", RFC 7766, DOI 10.17487/RFC7766, March 2016, <https://www.rfc-editor.org/info/rfc7766>.

[RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <https://www.rfc-editor.org/info/rfc8029>.

[RFC8815] Abrahamsson, M., Chown, T., Giuliano, L., and T. Eckert, "Deprecating Any-Source Multicast (ASM) for Interdomain Multicast", BCP 229, RFC 8815, DOI 10.17487/RFC8815, August 2020, <https://www.rfc-editor.org/info/rfc8815>.

[RFC8899] Fairhurst, G., Jones, T., Tüxen, M., Rüngeler, I., and T. Völker, "Packetization Layer Path MTU Discovery for Datagram Transports", RFC 8899, DOI 10.17487/RFC8899, September 2020, <https://www.rfc-editor.org/info/rfc8899>.

[RFC9293] Eddy, W., Ed., "Transmission Control Protocol (TCP)", STD 7, RFC 9293, DOI 10.17487/RFC9293, August 2022, <https://www.rfc-editor.org/info/rfc9293>.

[v4historic] Howard, L., "IPv4 Declared Historic", Work in Progress, Internet-Draft, draft-howard-sunset4-v4historic-00, 14 March 2016, <https://datatracker.ietf.org/doc/html/draft-howard-sunset4-v4historic-00>.

[wes-george-sunset4-2016-03-22] George, W., "Re: [sunset4] Declaring IPv4 Historic", 22 March 2016, <https://mailarchive.ietf.org/arch/msg/sunset4/bvuqbcPPazA9WF1I3hye7envYRU/>.

Authors' Addresses

Seth David Schoen
IPv4 Unicast Extensions Project
San Francisco, CA
United States of America
Email: schoen@loyalty.org

John Gilmore
IPv4 Unicast Extensions Project
PO Box 170640-rfc
San Francisco, CA 94117-0640
United States of America
Email: gnu@rfc.toad.com

David M. Täht
IPv4 Unicast Extensions Project
Half Moon Bay, CA
United States of America
Email: dave@taht.net

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: 15 August 2024

F. L. Templin, Ed.
Boeing Research & Technology

12 February 2024

IPv4 Parcels and Advanced Jumbos (AJs)
draft-templin-intarea-parcels-99

Abstract

IPv6 Parcels and Advanced Jumbos (AJs) present new data packaging constructs and a new link model for Internetworking. As is often the case, technologies developed in the IPv6 space can also be applied in IPv4 and vice-versa. This document presents the adaptations necessary to support Parcels and AJs in IPv4.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

IPv6 Parcels and Advanced Jumbos (AJs) [I-D.templin-6man-parcels] present new data packaging constructs and a new link model for Internetworking. As is often the case, technologies developed in the IPv6 space [RFC8200] can also be applied in IPv4 [RFC0791] and vice-versa. This document presents the differences that need to be addressed to adapt IPv6 Parcels and AJs to IPv4.

All aspects of IPv6 Parcels and AJs, including the use of IPv6 extension headers and control messaging, apply also to IPv4. Only differences in the IP header format and some control option encapsulations need to be accounted for as discussed below. This document therefore specifies IPv4 parcels and AJs.

2. Requirements

IPv4 parcels and AJs observe all requirements established for IPv6 [I-D.templin-6man-parcels] including the use of IPv6 Hop-by-Hop Options headers. This means that nodes that recognize IPv4 parcels/AJs MUST recognize and correctly process IP protocol 0 (Hop-by-Hop) extension headers the same as for IPv6 when they occur in an extension header chain following the IPv4 header but before the upper

layer payload.

When an IPv4 router or destination end system processes a parcel/AJ probe for which the IPv4 Protocol field encodes an unrecognized value (such as 0 for Hop-by-Hop Options), it drops the probe and returns an ICMPv4 "Destination Unreachable - Protocol Unreachable" message [RFC0792]. The source should regard any such messages as an advisory indication that OMNI protocol UDP encapsulation may be necessary in future probes.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119][RFC8174] when, and only when, they appear in all capitals, as shown here.

3. IPv4 Parcels and Advanced Jumbos (AJs)

All aspects of [I-D.templin-6man-parcels] are imported as normative specifications for IPv4 parcels and AJs, with the exception of the following differences:

3.1. IPv4 Total Length

The IPv6 header includes a "Payload Length" field defined as the: "Length of the IPv6 payload, i.e., the rest of the packet following this IPv6 header, in octets". The IPv4 header instead includes a "Total Length" field defined as: "the length of the datagram, measured in octets, including internet header and data".

IPv6 parcels/AJs always set the Payload Length field to 0; thus, a node that does not understand the parcel/AJ format may truncate the message and send only the IPv6 header forward.

IPv4 parcels/AJs instead always set the Total Length to the length of the IPv4 header to ensure the same truncation behavior.

3.2. IPv4 Time To Live (TTL)

The IPv4 "Time To Live (TTL)" and IPv6 "Hop Limit" values are treated in exactly the same way in both protocol versions. In particular, the source sets the TTL/Hop Limit to an initial value and each router in the path to the destination decrements the TTL/Hop Limit by 1 when it forwards a parcel/AJ/probe. (Note that this represents a parcel/AJ-specific requirement for IPv4 routers, since [RFC1812] permits routers to decrement TTL by values other than 1.)

3.3. IPv4 Parcel/Jumbo Payload Length

The same as for IPv6, the Parcel Payload Length field in the Parcel Payload Option and the Jumbo Payload Length field in the Jumbo Payload Option of IPv4 parcels/AJs encode the length of the IPv6 extension headers plus the length of the {TCP,UDP} header plus the combined length of all concatenated segments with their per-segment headers/trailers.

Therefore, the length of the IPv4 header itself is not included in the Parcel/Jumbo Payload Length field the same as for IPv6. The IPv4 header length for IPv4 parcels and AJs is instead available in both the IPv4 Total Length and Internet Header Length (IHL) fields.

3.4. IPv4-Compatible IPv6 Addresses

Whenever an IPv4 address needs to be coded in an IPv6 address field, the address is coded as an IPv4-compatible IPv6 address as specified in [RFC4291].

3.5. IPv4 Parcel Packetization/Restoration

When a node performs packetization on a {TCP,UDP}/IPv4 parcel, it inserts a Parcel Parameters {TCP,UDP} option the same as for IPv6 [I-D.templin-6man-parcels].

The IPv4 destination then performs restoration by gathering up IPv4 packets that arrive with the same upper layer 5-tuple and with Parcel Parameters information including the same Identification. The Parcel Parameters Index then determines the ordinal position of each packet segment to be concatenated into the restoration buffer, i.e., the same as for IPv6. (Note: if the IPv4 destination does not recognize the {TCP,UDP} Parcel Parameters option, it simply processes the packet as a singleton IPv4 packet. This would result in correct behavior, but with Generic Receive Offload (GRO) disabled.)

3.6. Parcel Probing

When an IPv4 router or destination receives an intact parcel probe, it processes the probe the same as specified for IPv6 [I-D.templin-6man-parcels].

When an IPv4 router forwards the parcel probe, it MUST decrement the TTL by exactly 1 the same as specified for the IPv6 Hop Limit.

When an IPv4 destination receives a parcel probe, it should return a Parcel Parameters option in any {TCP,UDP}/IPv4 packet to be returned to the source the same as for IPv6.

When the IPv4 source receives a {TCP,UDP} packet that includes a Parcel Parameters option it matches the Identification value with its recently-transmitted probes. If there is a match, the source accepts the MTU and Parcel Limit values found in the Parcel Parameters option.

The same as for IPv6, if the source or destination is located outside of a controlled environment / limited domain [RFC8799] the source should send probes including the IPv4 header followed by an OMNI UDP encapsulation header followed by the Hop-by-Hop Options header and finally followed by the {TCP,UDP} header plus protocol data.

3.7. Parcel/Jumbo Replys

When an IPv4 router returns a Parcel/Jumbo Reply, it prepares the message in exactly the same way as for IPv6 (i.e., including the ICMPv6 message body and optional Encapsulating Security Payload (ESP) encapsulation) then wraps the message in UDP/IPv4 headers for transmission to the source according to the OMNI protocol specification.

3.8. Advanced Jumbos (AJs)

All aspects of IPv4 Advanced Jumbos (AJ) are processed the same as for IPv6 AJs.

3.9. Jumbo-in-Jumbo Encapsulation

Original IPv4 parcels/AJs can follow the "e(X)treme" forwarding paths across successive OMNI links in the path using "jumbo-in-jumbo" encapsulation the same as for IPv6. The OMNI link ingress encapsulates each IPv4 parcel/AJ in an OMNI IPv6 header plus any outer L2 encapsulations which may include an IPv4 header with an Advanced Jumbo option Hop-By-Hop extension header. All aspects of this "jumbo-in-jumbo" encapsulation are the same as for IPv6.

3.10. Integrity

To support the IPv4 parcel/AJ header checksum calculation, the network layer uses modified versions of the {TCP,UDP}/IPv4 pseudo-header found in [RFC9293]. Note that while the contents of the two IP protocol version-specific pseudo-headers beyond the address fields are the same, the order in which the contents are arranged differs and must be honored according to the specific IP protocol version.

The IPv6 pseudo-header is found in [I-D.templin-6man-parcels], while the IPv4 pseudo-header is shown in Figure 1. The similarities between the two pseudo-headers allows for maximal reuse of widely deployed code while ensuring interoperability.

```
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      IPv4 Source Address                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    IPv4 Destination Address                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      zero     |  Next Header  |        Segment Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Index   |C|S|D|X|          Parcel Payload Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: {TCP,UDP}/IPv4 Parcel/AJ Pseudo-Header Format

Note: The same as for IPv6, the "Index/C/S/D/X" and Parcel Payload Length fields in the IPv4 parcel pseudo-header are replaced by the single 4-octet Jumbo Payload Length field in the IPv4 AJ pseudo-header.

4. Implementation Status

An early prototype of UDP/IPv4 parcels (draft version -15) has been implemented relative to the linux-5.10.67 kernel and ION-DTN ion-open-source-4.1.0 source distributions. Patch distribution found at: "https://github.com/fltemplin/ip-parcels.git".

5. IANA Considerations

This document does not include any IANA instructions.

6. Security Considerations

Security Considerations are the same as for IPv6 as found in [I-D.templin-6man-parcels].

7. Acknowledgements

This work was inspired by ongoing AERO/OMNI/DTN investigations. The concepts were further motivated through discussions with colleagues.

Honoring life, liberty and the pursuit of happiness.

8. References

8.1. Normative References

[I-D.templin-6man-parcels]
Templin, F., "IPv6 Parcels and Advanced Jumbos (AJs)", Work in Progress, Internet-Draft, draft-templin-6man-parcels-18, 26 January 2024, <https://datatracker.ietf.org/doc/html/draft-templin-6man-parcels-18>.

[I-D.templin-intarea-omni]
Templin, F., "Transmission of IP Packets over Overlay Multilink Network (OMNI) Interfaces", Work in Progress, Internet-Draft, draft-templin-intarea-omni-64, 26 January 2024, <https://datatracker.ietf.org/doc/html/draft-templin-intarea-omni-64>.

[RFC0768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, DOI 10.17487/RFC0768, August 1980, <https://www.rfc-editor.org/info/rfc768>.

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <https://www.rfc-editor.org/info/rfc791>.

[RFC0792] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, DOI 10.17487/RFC0792, September 1981, <https://www.rfc-editor.org/info/rfc792>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, DOI 10.17487/RFC4291, February 2006, <https://www.rfc-editor.org/info/rfc4291>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[RFC9293] Eddy, W., Ed., "Transmission Control Protocol (TCP)", STD 7, RFC 9293, DOI 10.17487/RFC9293, August 2022, <https://www.rfc-editor.org/info/rfc9293>.

8.2. Informative References

[RFC1812] Baker, F., Ed., "Requirements for IP Version 4 Routers", RFC 1812, DOI 10.17487/RFC1812, June 1995, <https://www.rfc-editor.org/info/rfc1812>.

[RFC8799] Carpenter, B. and B. Liu, "Limited Domains and Internet Protocols", RFC 8799, DOI 10.17487/RFC8799, July 2020, <https://www.rfc-editor.org/info/rfc8799>.

Appendix A. Change Log

<< RFC Editor - remove prior to publication >>

Changes from earlier versions:

* Submit for review.

Author's Address

Fred L. Templin (editor)
Boeing Research & Technology
P.O. Box 3707
Seattle, WA 98124
United States of America
Email: fltemplin@acm.org