INTERNET-DRAFT
Intended Status: Standard
Expires: July 2019

T. Herbert
Quantonium
Vikram Siwach
Independent consultant

January 28, 2019

Address Mapping System
draft-herbert-intarea-ams-00

Abstract

This document describes the Address Mapping System that is a generic, extensible, and scalable system for mapping network addresses to other network addresses. The Address Mapping System is intended to be used in conjunction with overlay techniques which facilitate transmission of packets across overlay networks. Information returned by the Address Mapping System can include the particular network overlay method and instructions related to the method. The Address Mapping System has a number of potential use cases networking including identifier-locator protocols, network virtualization, and promotion of privacy.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/1id-abstracts.html

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

Copyright and License Notice

Table of Contents

1 Introduction

This document describes the Address Mapping System (AMS). AMS is a system that maps network addresses to other network addresses. The canonical use case is to map "identifiers" to "locators" (applying identifier-locator split terminology). Identifiers are logical addresses that identify a node, and locators are addresses that indicate the current location of a node. Identifiers are mapped to locators at points in the data path to facilitate device mobility or or network virtualization.

The address mapping system may be queried on a per packet basis in the data path. For instance, an encapsulating tunnel ingress node for virtualization would perform a lookup on each destination virtual address to discover the address of the physical node to which a packet should be forwarded. It follows that access to the mapping system is expected to be tightly coupled with nodes that query the system to perform packet forwarding.

The mapping system contains a database or table of all the address mappings for a mapping domain. The database may be distributed across some number of nodes, sharded for scalability, and caches may be used to optimize communications. The mappings in a mapping system may be very dynamic, for instance end user devices in a mobile network may change location within the network at a high rate (e.g. a mobile device in fast moving automobile may frequently connect to different cells). Protocols are defined to synchronize the mapping information across devices that participate in the address mapping system.

1.1 Use cases

This section describes some of the use cases of the Address Mapping System.

- Network virtualization
- Identifier/locator protocols
- Address resolution
- Privacy in Internet addressing
- Mobile networks

1.2 Requirements

Requirements for the Internet Addressing Mapping system are:

- Allow use of different overlay protocols

  The mapping system should be agnostic to the protocol used to implement an underlying network overlay. An overlay could be implemented using an encapsulation protocol, such as GTP, GUE, LISP, VXLAN, etc., or using and identifier/locator address split protocol such as ILA. A network may simultaneously use different protocols per its needs. Mapping information provided by the address mapping system could include instructions that indicate the overlay protocol to be used when sending to a destination.

- Secure access to mapping system

  An address mapping system may contain sensitive information, particularly in the case that locators would reveal location or identity of specific users. Access to the mapping system must be tightly controlled. Law enforcement considerations may require maintaining a history of mappings to provide under legal order.

- Mapping caches (anchorless mobility)

  Mapping caches may be implemented at the network edge to perform overlay forwarding and avoid triangular routing through centralized anchor points. A cache may be implemented as a working set cache or could be pre-populated with mappings for common destinations. The purpose of the cache is to optimize communications for critical communications, however the use of caches should not be required for viable communications.

- Scalability

  Address mapping systems should be able to scale to at least a billion mappings in a single mapping system domain. This accounts for a large number of devices, where each device may have some number of associated mappings. It follows that a large deployment will likely need a number of sharded mapping servers each of which may be replicated for reliability.

- Resiliency against Denial of Service attack

  An address mapping system must be resistant to Denial of Service attacks. For instance, if a mapping cache is used then a resource exhaustion attack on a mapping cache must not result in loss of service to users.

- User privacy

  An address mapping system must facilitate user privacy. As

mentioned above, the mapping system must be secured to prevent intrusion for sensitive personal information. The mapping system can also foster privacy in addressing by supporting untrackable, per-flow IP addresses.

- o Seamless handover

  When a mobile device switches from one point of attachment to another (handover), existing communications should continue without packet loss or substantial delay. The mapping system must be dynamic to handle handover events with bounded latency.

- o Roaming

  Devices may roam from one administrative domain to another. The mapping systems in the home domain and remote domain may coordinate to persist existing communications using addresses that are local to the home domain.

- o Stateless mapping mode

  An address mapping system may provide a communication mode where the mapping information is carried in packets themselves. When a packet the contains such information enters a network, the information can be decoded to determine the identifier to locator mapping. This obviates the need for lookup in the mapping system for each packet.

1.3 Terminology

Address Mapping System (AMS)
: A system for mapping addresses to other addresses.

Address mapping system domain
: An administrative domain in which an address mapping system is run. The address mappings and related addresses are considered to be in a domain. An address mapping system domain implements a security policy to prevent unauthorized viewing or manipulation of mapping information.

Mapping database/mapping table
: A logical or real database that contains all of the address mappings for an address mapping system domain.

Mapping addresses
: The network addresses that are the objects in the address mapping system table. These are typical IPv4 or IPv6

addresses, but can generically be any type of fixed length network addresses.

Identifier
: A mapping address that identifies an end node in network communication. In AMS, identifier generically refers to the key in an address mapping system database.

Locator
: A mapping address that refers to the location of a node. In AMS, locator generically refers to the addresses that a key maps to in the mapping system database.

Mapping entry
: A single entry in a mapping domain. A mapping entry is composed of the key address (the identifier), one or more locators that the key maps to, and optional ancillary information.

Mapping query
: A lookup in the address mapping system database. A key address (identifier) is provided and the corresponding map entry (containing locators) is returned if the key is matched in the table.

Overlay forwarding
: The processing performed to implement a network overlay that forwards packets to the location for their destination address based on a mapping entry in the address mapping system. Overlay forwarding may be encapsulation, address transformations, etc.

Overlay termination
: The processing done at the terminal endpoint of overlay protocol used in overlay forwarding.

AMS router (AMS-R)
: A node that contains all or a shard of the addressing mapping system database. An AMS-R node serves mapping system information to AMS forwarding nodes. An AMS router node will often act at a packet router that performs overlay forwarding for addresses that it manages in the mapping system.

AMS forwarders (AMS-F)
: A node that performs overlay forwarding and/or overlay termination. The AMS forwarder contains a mapping cache to facilitate overlay forwarding. End hosts may participate in the address mapping system as a

specialized type of a forwarder.

Addressing Mapping Routing Protocol (AMRP)
: A protocol used amongst AMS routers to synchronize the mapping system database.

Addressing Mapping Forwarder Protocol (AMFP)
: A control protocol run between AMS routers and AMS forwarders that is used to manage mapping caches in AMS forwarders.

Firewall and Service Tickets (FAST)
: A protocol in which packets carry "tickets" in extension headers. Tickets provide arbitrary information about how a network processes packets.

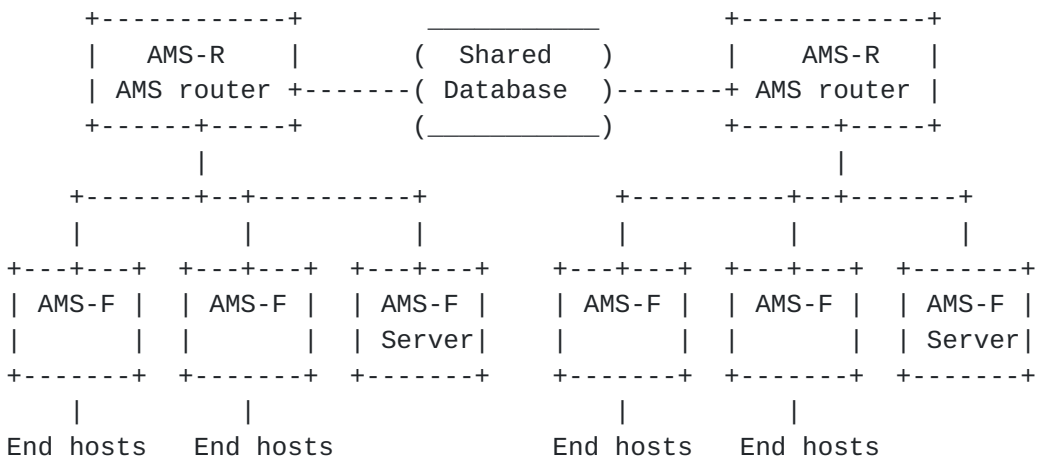
Hidden aggregation
: A method to encode aggregation in network addresses where the aggregation is visible to trusted devices within a network, but is transparent to external observers of the addresses.

2 Architecture

This section describes the architecture of the Address Mapping System.

2.1 Reference topology

This section provides a generic reference topology for AMS.



2.2 Functional components

As shown in the reference topology, there are two types of functional nodes in the AMS architecture:

AMS-R: AMS routers

AMS-F: AMS forwarders

2.3 AMS router (AMS-R)

AMS routers are deployed within the network infrastructure and collectively contain the address mapping database for an address mapping system domain. The database may be sharded across some number of routers for scalability. AMS routers that maintain the database or a shard may be replicated for scalability and availability. AMS routers share and synchronize mapping information amongst themselves using an Address Mapping Routing Protocol.

AMS routers serve mapping information to AMS forwarders via the Address Mapping Forwarder Protocol. Mapping information is provided by a request/reply protocol, a push mechanism, or mapping redirects.

An AMS router may perform overlay forwarding for the destination addresses it serves in the address mapping system database. Network

routing is configured so that packets with identifier addresses served by an AMS-R will be routed to that AMS-R.

AMS routers are considered authoritative for the portion of that mapping database that they serve. For instance, if a packet with an identifier address is routed to an AMS-R then, either a mapping is found and the packet is forwarded via overlay forwarding, or the packet is dropped. In this sense, AMS routers can be thought of as anchor point when they are forwarding packets (in 3GPP terminology).

An AMS router can send mapping redirects to AMS forwarders in order to inform them of a direct path they can take to a destination. A redirect is sent to the upstream AMS forwarder of the source which can be determined by a mapping query the source address. When an AMS forwarder receives a redirect, it can create a mapping cache entry and apply overlay forwarding on subsequent packets to directly send to the destination instead routing packets through a AMS router.

2.3.1 AMS router operation

The operation of a forwarding AMS router is:

1) Packet are routed to the AMS-R

2) For each received packet, a lookup on the destination address is done in the mapping system database

3) If a matching mapping entry is found in the address mapping database:

   - The packet is forwarded over a network overlay per the returned locator and ancillary information

   - Optionally, a mapping redirect is be sent to an AMS forwarder that is in that path from the source of the packet

4) Else, the packet is dropped

2.4 AMS forwarder (AMS-F)

As indicated in the reference topology, forwarding nodes may deployed near the point of device attachment (e.g. base station, eNodeB) of user devices (e.g. UEs).

End hosts may act as AMS forwarders. These could be servers that provide overlay forwarding and termination on behalf of VMs or containers for virtualization. Since the source of packets is local on a host that is an AMS forwarder is, there may be some datapath

optimizations that can be applied.

AMS forwarders have two functions:

- Overlay termination which is restoring packet with original identifier addresses
- Optional overlay forwarding to destinations based on a mapping cache

2.4.1 Overlay termination

AMS forwarders perform overlay termination. In other words, they are typically the target node of a locator. Overlay termination is the process of removing or undoing the overlay processing that was previously done. If the overlay method is encapsulation, the overlay termination processing is to decapsulate the packet. If the overlay method is address transformation, such as in ILA, the overlay termination processing is to transform addresses back to their original values before overlay processing. Once the overlay processing is undone, an AMS forwarder forwards the resultant packet to its final destination.

2.4.2 Overlay forwarding

An AMS forwarder may perform overlay fowarding to send packets directly to the destination using a cache of address mappings. The mapping cache of an AMS forwarder may be managed as a working set cache. As a cache there must be methods to populate, evict, and timeout entries. A cache is considered an optimization, so the system should be functional without it being used use (e.g. if the cache has no entries)

The operation of overlay forwarding in an AMS forwarder is:

1) Receive packets from downstream nodes

2) Lookup up packet's destination address in the mapping cache

3) If a match is found in the mapping cache then forward the packet over a network overlay per the returned locator and instructions

4) Else, forward the unmodified packet in the network per normal routing

5) An AMS router may send a mapping redirect in response to a packet that had been forwarded by the AMS forwarder. In

response, the forwarder may create a mapping cache entry based on the contents of the redirect and use the entry to send directly to a destination for subsequent packets.

3 Address Mapping Router Protocol (AMRP)

AMS routers must synchronize the contents of the mapping database. When a change occurs to an address mapping, for instance a mobile device has moved to a new location, the AMS routers managing the shard that contains the identifier must be synchronized in as little convergence time as possible.

There are a number of options to use or have been used to implement an AMS mapping router protocol. This document highlights some alternatives, but does not prescribe a particular protocol.

3.1 Key/value database

A key/value database, such as a NoSQL database like Redis, can implement an address mapping routing protocol. The idea of the database is that each mapping shard is a distributed database instance with some number of replicas. When a write is done in the database, the change is propagated throughout all of the replicas for the shard using the standard database replication mechanisms. Mapping information is written to the database using common database API that can require authenticated write permissions. Each AMS router can read the database for the associated shard to perform its function.

3.2 BGP

BPG can be used to propagate mapping information amongst AMS routers as simple routes. [BGP-ILA] describes a method to distribute identifier to locator information using Multiprotocol Extensions for BGP-4.

3.3 3GPP uses GTP

4 Address Mapping Forwarder Protocol (AMFP)

The Address Mapping Forwarder Protocol (AMFP) is a control plane protocol that provides address to address mappings. Clients of the AMFP include AMS forwarders with mapping caches, so AMFP includes primitives for mapping cache management.

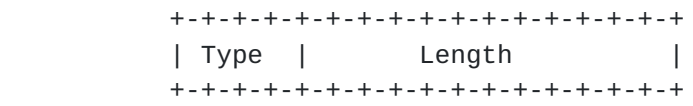
AMFP is primarily used between AMS forwarders and AMS routers. The purpose of the protocol is to populate and maintain the mapping cache in AMS forwarders.

AMFP defines mapping redirects, a request/response protocol, and a push mechanism to populate the mapping cache. AMFP runs over TCP to leverage reliability, statefulness implied by established connections, ordering, and security in the form of TLS. Secure redirects are facilitated by the use of TCP.

AMFP messages are sent over the TCP stream and must be delineated by a receiver. Different versions of AMS are allowed and the version used for communication is negotiated by Hello messages.

4.1 Common header format

All AMFP messages begin with a two octet common header:



The contents of the common header are:

- Type: Indicates the type of message. A type 0 message is a Hello message. Types greater than zero are interpreted per the negotiated version.

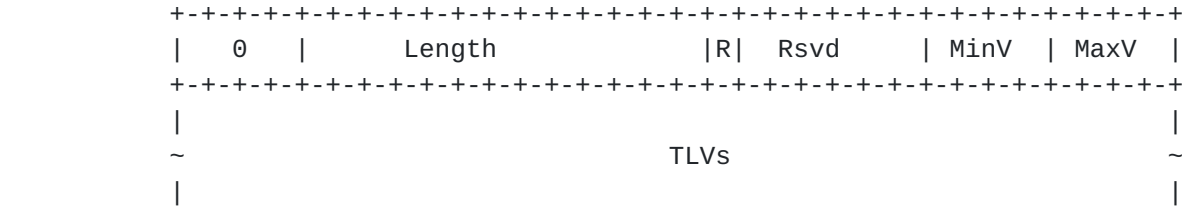
- Length: Length of the message in 32-bit words not including the first four bytes of the message. All AMFP messages are multiples of of four bytes in length and the message length includes the two bytes for the common header. The length field is computed as (message length / 4) - 1, so the minimum message size is four and the maximum size is 16,384 bytes.

Following the two octet common header is variable length data that is specific to the version and type the message.

4.2 Hello messages

Hello messages indicate the versions of AMFP that a node supports. Hello message MUST be sent by each side as the first message in the connection.

The format of an AMFP Hello message is:



The contents of the Hello message are:

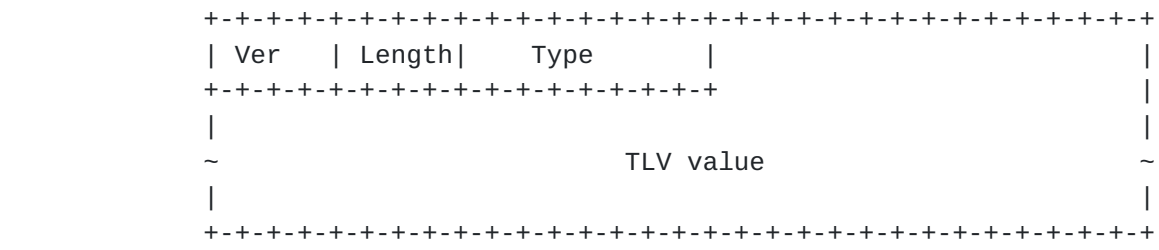
- Type = 0. This is indicates the type is a Hello message.
- Router bit: Indicates the sender is an AMS router. If the sender is an AMS forwarder this bit is cleared.
- Rsvd: Reserved bits. Must be set to zero on transmit.
- MinV: Minimum version number supported by the sending node.
- MaxV: Maximum version number supported by the sending node.
- TLVs: An optional list of TLVs that describe capabilities or requested options.

Version numbers are from 0 to 15. This document describes version 0 of AMFP.

4.3 Hello Message TLVs

TLVs (Type Length Values) MAY be used in AMFP Hello messages to convey optional information and parameters pertaining for negotiation. For example, a TLV could be used by an AMS-F to indicate the maximum number of mappings in its cache so that the AMS-R can managed redirects and pushed mappings accordingly.

The format of the Hello message TLVs is:



The fields of the Hello message TLV are:

- Ver: Which version of AMFP the TLV refers to. This be one of the version numbers in the proposed range of the Hello message.
- Length: Length of the TLV in 32-bit words not including the first four bytes. All Hello message TLVs have a size that is a multiple of four bytes. The minimum length of a TLV is four bytes and the maximum length is sixty-four bytes.

- Type: The type of the TLV. Type values are relative to the version stated in the Ver field so that each AMFP version has its own set of TLV types.

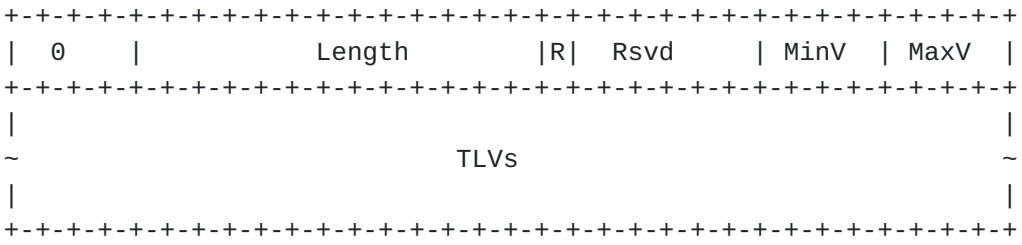
A receiver MUST only process TLVs that are for the negotiated version. Before parsing the TLV list, the r3eceiver determines the negotiated AMFP version as described above. When parsing the list of TLVs, the node should skip over any TLVs that are not for the negotiated version.

The high order bit if the TLV type indicates disposition of a type that is unrecognized. If the high bit is set for a TLV type that the receiver does not recognizes and the TLV if for the AMFP version that is being negotiated, then the receiver MUST terminate the connection. It MAY log the error.

4.2 Hello messages

Hello messages indicate the versions of AMCP that a node supports. Hello message MUST be sent by each side as the first message in the connection.

The format of an AMCP Hello message is:



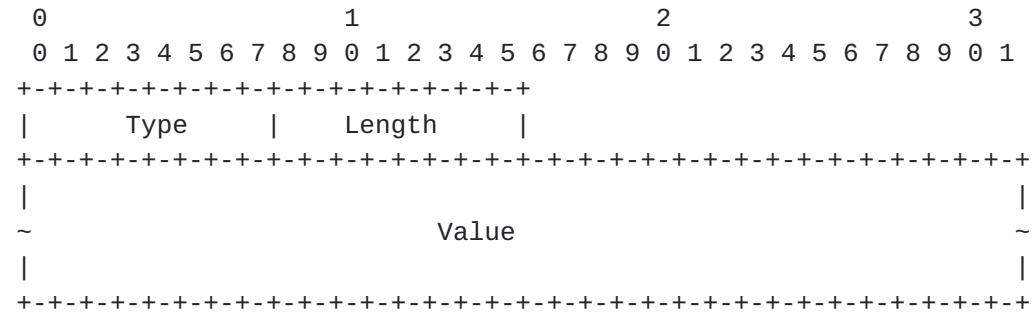
The contents of the Hello message are:

- Type = 0. This is indicates the type is a Hello message.
- Router bit: Indicates the sender is an AMS router. If the sender is an AMS forwarder this bit is cleared.
- Rsvd: Reserved bits. Must be set to zero on transmit.
- MinV: Minimum version number supported by the sending node.
- MaxV: Maximum version number supported by the sending node.
- TLVs; An optional list of Type Length Value structures (TLVs). Use of TLVs in Hello messages is described below.

Version numbers are from 0 to 15. This document describes version 0 of AMCP.

4.1.1 TLV format

Hello message TLVs have the following format:



Fields:

- Type: Type for TLV. Defined types are described below
- Length: Length in bytes of a TLV Value. Note that this length does not include the two bytes for Type and Length.
- Value: Data for the TLV

4.1.2 TLV types

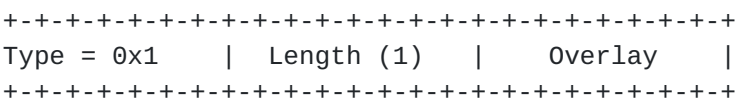
The table below lists the TLVs defined in this document. The "Length" column indicates any required limits on TLVs, and the "Typical Length" column indicates the most useful lengths for the TLV.

Type	Length	Sender	Meaning
0			RESERVED
1	1	Router	Default overlay method
2	4	Router	Default timeout
3	1	Router	Default priority
4	1	Router	Default weight
5	variable	Router	Default instructions
6	variable	Either side	Supported overlay methods
5-127			UNASSIGNED (assignable by IANA)
128-255			User defined

The TLVs defined in this document apply to version 0.

4.1.3 Default overlay method

This TLV reports the default overlay method in report mapping information when the method is not explicitly provided in a mapping information message. The format of the TLV is:



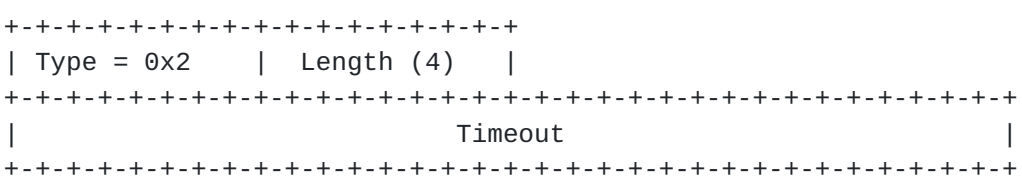
Fields are:

- Overlay: Indicates the overlay method to be used when sending to the locator in the entry (e.g. ILA, LIST, SRv6, IPIP, GRE, etc.). A value of zero indicates that the default overlay method for the network or that negotiated by Hello messages.

On AMS routers SHOULD send this TLV. If the TLV is received by a router it is considered an error.

4.1.4 Default timeout

This TLV reports the default timeout for report mapping information when the timeout is not explicitly provided in a mapping information.



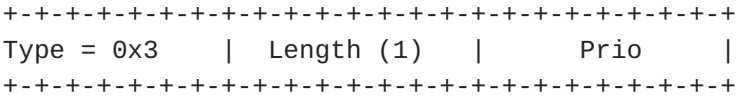
Fields are:

- Timeout: The time to live for the identifier information in seconds.

Only AMS routers send this TLV. If the TLV is received by a router it is considered an error.

4.1.5 Default priority

This TLV reports the default overlay priority in reported mapping information when the priority is not explicitly provided in a mapping information message. The format of the TLV is:



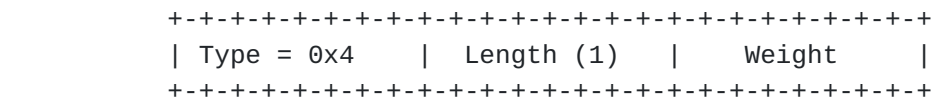
Fields are:

- Priority: Relative priority of a locator. Locators with higher priority values have preference to be used. Locators that have the same priority may be used for load balancing.

Only AMS routers send this TLV. If the TLV is received by a router it is considered an error.

4.1.6 Default weight

This TLV reports the default weight in reported mapping information when the priority is not explicitly provided in a mapping information message. The format of the TLV is:



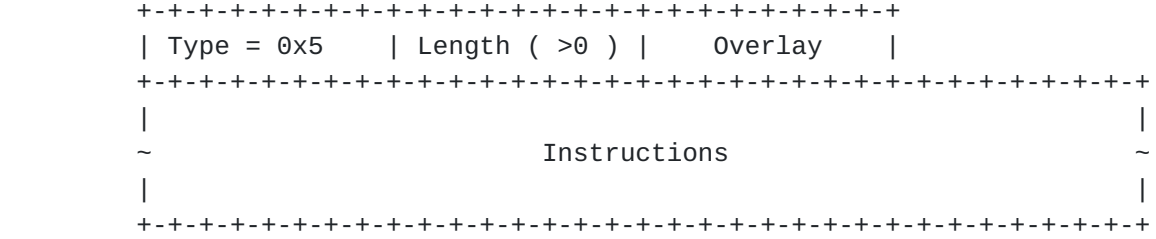
Fields are:

- Weight: Relative weights assigned to each locator. In the case that locators have the same priority the weights are used to control how traffic is distributed. A weight of zero indicates no weight and the mapping is not used unless all locators for the same priority have a weight of zero.

Only AMS routers send this TLV. If the TLV is received by a router it is considered an error.

4.1.6 Default instructions

This TLV reports the default overlay specific instructions in reported mapping information when instructions are not explicitly provided in a mapping information message. The format of the TLV is:



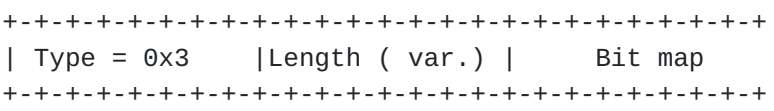
Fields are:

- Instructions: Optional data with format and semantics that are specific to an overlay method and can describe options for the method how the overlay method is used.

- Overlay: Indicates the overlay method to be used when sending to the locator in the entry (e.g. ILA, LIST, SRv6, IPIP, GRE, etc.). A value of zero indicates that the default overlay method for the network or that negotiated by Hello messages. Only AMS routers send this TLV. If the TLV is received by a router it is considered an error.

4.1.5 Supported overlay methods

This TLV reports the support overlay methods that a node supports. The TLV can be sent by either and MAS-R or an AMS-F and is a hint to the peer about what methods are supported. The format of the message is:



Fields are:

- Bit map: A variable length bit map that indicates overlay methods. The position in the bip map corresponds to the defined values for the various overlay methods.

Relative priority of a locator. Locators with higher priority values have preference to be used. Locators that have the same priority may be used for load balancing.

Only AMS routers send this TLV. If the TLV is received by a router it is considered an error.

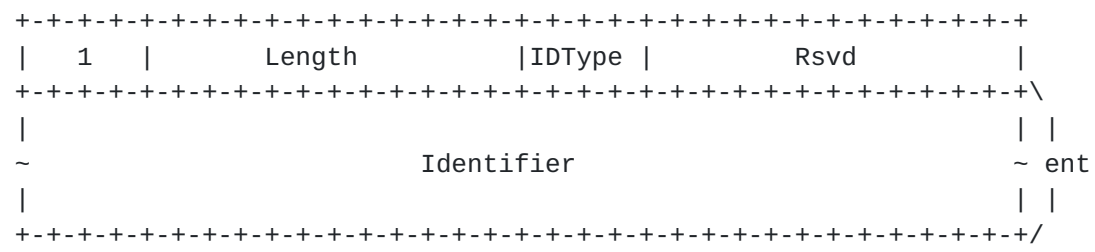
4.4 AMFP Version 0

The message types in version 0 of AMFP are:

- Map request (Type = 1)
- Map information (Type = 2)
- Compressed map information (Type = 3)
- Locator unreachable (Type = 4)

4.4.1 Map request

A map request is sent by an AMS forwarder to an AMS router to request mapping information for a list of identifiers. The format of a map request is:



The contents of the map request message are:

- Type = 1. This is indicates the type is map request.
- Length: Message length is set to size of an identifier times the number of identifiers in the list. The Length field is computed as identifier_size * number_of identifiers.
- Rsvd: Reserved bits. Must be set to zero when sending.
- IDType: Identifier type. Specifies the identifier type. This also implies the length of each identifier in the request list. Identifier types are defined below.
- Identifier: An identifier of type indicated by IDType. The size of an identifier is specified by the type.

The Identifier field is repeated for each identifier in the list. The number of identifiers being requested is (message length - 4) / (identifier size).

4.4.2 Map information

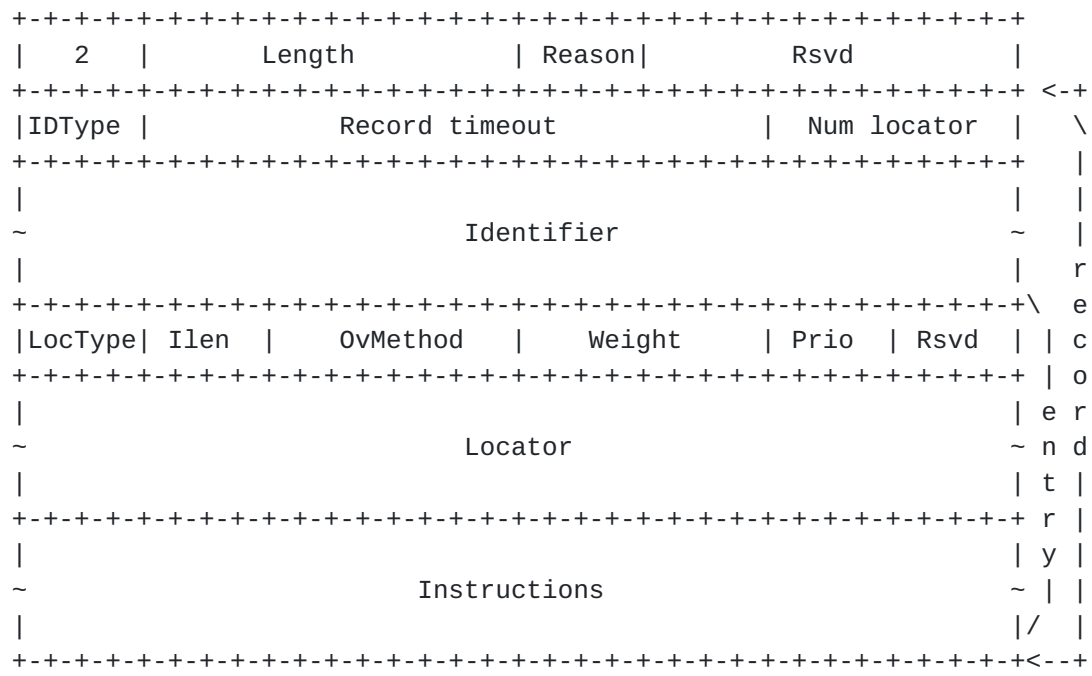
A map information message is sent by an AMS router to provide identifier to locator mapping information. In addition to providing locators for an identifier, the message also contains the overlay method to use and related instructions for sending to an identifier.

A map information message is composed four byte header followed by a set of identifier records. Each identifier record describes mapping information for one identifier. An identifier record is composed of a four byte header, an identifier, and a set of locator entries. Each locator entry provides the information about one locator used to reach an identifier. Each locator entry is composed of a four byte header that includes the overlay method to use, the locator, and instructions specific to the overlay method for the locator.

The identifier record is repeated for each mapping being reported and the locator entry is repeated for each locator being reported for an identifier. The total number of identifiers being reported is

determined by parsing the message.

The format of a map information message is:



The contents of the map information message header are:

- Type = 2. This indicates an extended map information message

- Length: The message length is four bytes plus the sum of lengths of all the identifier records in the message. The length of a record is four bytes plus the sum of all the lengths of locator entries in the record. The length of a locator entry is four plus the size of a locator plus the length of instruction field.

- Reason: Specifies the reason that the message was sent. Reasons are:

  - 0: Map reply to a map request

  - 1: Redirect

  - 2: Push map information

- Rsvd: Reserved bits. Must be set to zero when sending.

The contents of an identifier record are:

- IDType: Identifier type. Specifies the identifier type. This also

implies the length of each identifier in the list. Identifier types are defined below.

- Record timeout: The time to live for the identifier information in seconds. A value of zero indicates the default is used.

- Num locator: Number of locators (entries) being reported for an identifier.

- Identifier: An identifier of type specified in IdType.

The contents of a locator entry are:

- LocType: Locator type. Specifies the locator type. This also implies the length of each locator in the list. Locator types are defined below.

- Ilen: Length in 32-bit words of optional instructions in the entry (length of the instructions field). Instructions are overlay method specific and can describe options or or how the overlay is used. The instructions length is from zero to sixty bytes.

- OvMethod: The overlay method to use for sending to the identifier using the given locator. This is an indication of the encapsulation method (e.g. GUE, GTP, LISP, etc.) or address transformation method (e.g. ILA). Specific values are listed in IANA section.

- Weight: Relative weights assigned to each locator. In the case that locators have the same priority the weights are used to control how traffic is distributed. A weight of zero indicates no weight and the mapping is not used unless all locators for the same priority have a weight of zero.

- Priority: Relative priority of a locator. Locators with higher priority values have preference to be used. Locators that have the same priority may be used for load balancing.

- Rsvd: Reserved bits. Must be set to zero when sending

- Locator: A locator of type specified in LocType.

- Instructions: Optional data with format and semantics that are specific to an overlay method and can describe options for the method how the overlay method is used. Ilen indicates the length of the field.

4.4.3 Compressed map information

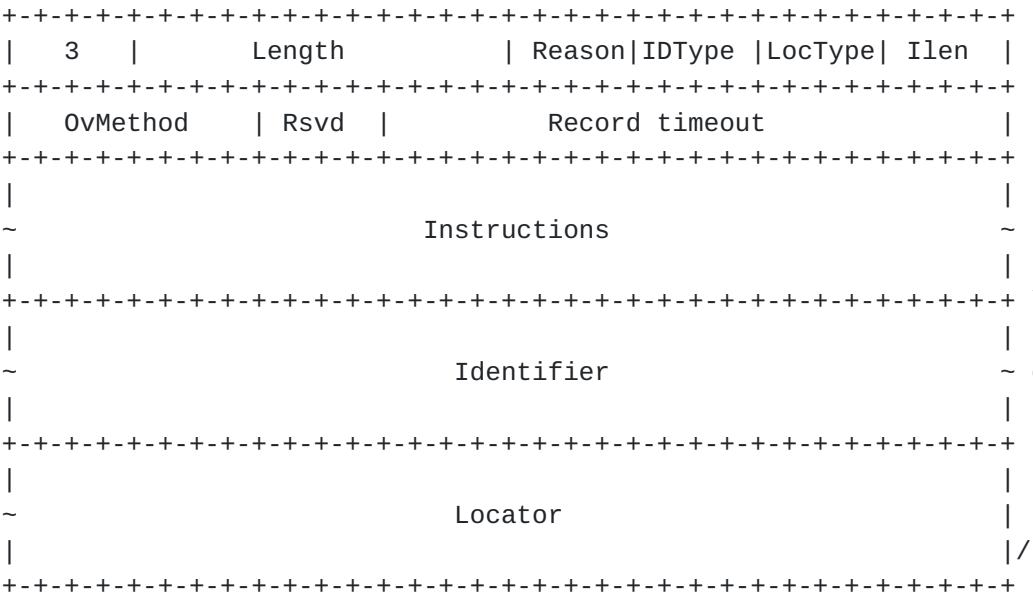
The compressed map information message may be used as a mode efficient alternative to the map information message. The compressed map information can be used when:

- There is only locator provided for each identifier, and
- The overlay method, identifier type, locator type, and overlay instructions are common for all the mappings reported in the message, and
- The identifier record timeout is common amongst the provided identifiers

A compressed map information message is composed of an eight byte header followed by an optional instruction field, followed by a set of identifier/locator pairs. Each pair on identifier to locator mapping, and stated the overlay method and instructions are common to all the mappings in the message.

The identifier/locator pairs are repeated for each mapping being reported. The total number of identifiers being reported can determined by parsing the message or computing it based on the message length and the lengths of the identifiers and locators.

The format of the compressed map information message header is:

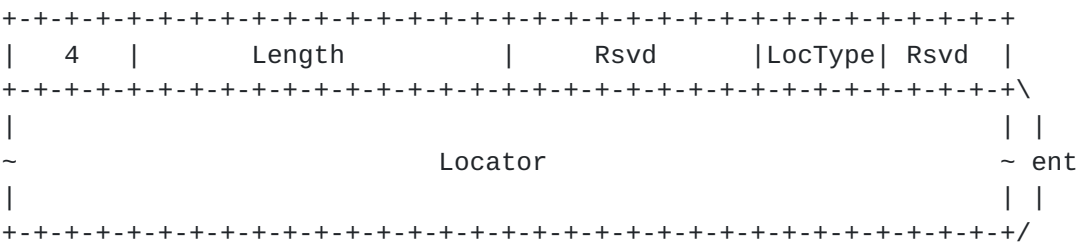


The contents of the compressed map information message header are:

- Type = 3. This indicates an compressed map information message
- Length: Set to four plus the length of the instructions field plus the sum of lengths of the identifier/locator pairs in the message.
- Reason: Specifies the reason that the message was sent. Reasons are:
  - 0: Map reply to a map request
  - 1: Redirect
  - 2: Push map information
- IDType: Identifier type. Specifies the identifier type. This also implies the length of each identifier in the list. Identifier types are defined below.
- LocType: Locator type. Specifies the locator type. This also implies the length of each locator in the list. Locator types are defined below.
- Ilen: Length in 32 bit words of optional instructions in the entry (length of the instructions field). Instructions are overlay method specific and can describe options or or how the overlay is used. The instructions length is from zero to sixty bytes.
- OvMethod: The overlay method to use for sending to the identifier using the given locator. This is an indication of the encapsulation method (e.g. GUE, GTP, LISP, etc.) or address transformation method (e.g. ILA). Specific values are listed in IAMA section.
- Rsvd: Reserved bits. Must be set to zero when sending
- Record timeout: The time to live for the identifier information in seconds. A value of zero indicates the default is used.
- Instructions: Optional data with format and semantics that are specific to an overlay method and can describe options for the method or how the overlay methos is used. Ilen indicates the length of the field.
- Identifier: An identifier of type specified in IdType.
- Locator: A locator of type specified in LocType.

4.4.4 Locator unreachable

A locator unreachable message is sent by AMS routers to AMS forwarders in the event that a locator or locators are known to no longer be reachable. The format of a locator unreachable message is:



The fields of the locator unreachable message are:

- Type = 4. This is indicates the type is a locator unreachable message.
- Length: Set to the size of the locator times the number of locators in the list.
- Rsvd: Reserved bits. Must be set to zero when sending.
- LocType: Locator type. Specifies the locator type. This also implies the length of each locator in the list. Locator types are defined below.
- Locator: A locator of type indicated by LocType. The size of a locator is specified by the type.

The Locator field is repeated for each locator in the list. The number of locators being reported is (message length - 4) / (locator size).

4.4.5 Identifier and locator types

Identifier and locator values used in IDType and LocType fields of AMCP messages are:

- 0: Null value, 0 bit vlaue. This indicates that absence of locator or identifier information.
- 1: IPv6 address, 128 bit value
- 2: IPv4 address, 32 bit value

- o 3: 32 bit index
- o 4: 64 bit index
- o 5: ILA value. A 64 bit value that represent a canonical ILA identifier when used in an IDType field and a canonical ILA locator when used in a LocType field.

Note that the types for index values are use to index into tables for locators or identifiers.

4.5 Operation

This section describes the operation of AMFP.

4.5.1 Version negotiation

The first message sent by each side of an AMFP connection is a Hello message. Hello messages contain the minimum and maximum versions of AMCP supported. The minimum and maximum values form an inclusive range.

When a host receives a AMFP Hello message, it determines which version is negotiated. The negotiated version is the maximum version number supported by both sides. For instance, if a node advertises a minimum version of 0 and maximum of 1 and receives a peer Hello message with a minimum version of 0 and maximum of 2; then the negotiated version is 1 since that is the greatest version supported by both sides. The peer host will also determine that 1 is the negotiated version.

If there is no common version supported between the peers, that is their supported version ranges are disjoint, then version negotiation fails. The connection MUST be terminated and error message SHOULD be logged.

If both sides set the router bit or both clear the router bit in a Hello message, then this is an error and the connection MUST be terminated and error message SHOULD be logged. Both sides cannot have the same role in an AMFP session.

4.5.2 Populating an mapping cache

AMS forwarders can maintain a cache of identifier to locator mappings. There are three means for populating this cache:

- o Redirects

- Mapping request/reply

- Pushed mappings

Redirects are RECOMMNDED as the primary means of dynamically obtaining mapping information. Request/reply and push mappings may be used in limited circumstances, however generally these techniques don't scale and are susceptible to DOS attack.

AMS forwarders (and AMS routers as well) are work conserving, they do not hold packets that are pending mapping resolution. If a node does not have a mapping for a destination in its cache then the packet is forwarded into the network; the packet should be processed by an AMS router and sent to the proper destination node.

4.5.3 Redirects

An AMS router can send redirects in conjunction with forwarding packets. Redirects are sent to AMS forwarders in order to inform them of a direct AMS path. A redirect is sent to the upstream AMS forwarder of the source which is determined by a lookup in the mapping system on the source address of the packet being forwarded. The found locator is used to infer an address of the AMS forwarder. Note that this technique assumes a symmetric path towards the source.

4.5.3.1 Proactive push with redirect

In addition to sending an AMFP redirect to the AMS forwarder, an AMS router MAY send an AMFP push to the AMS forwarder associated with the destination to inform it of the identifier to locator mapping for the source address in a packet. This is an optimization to push the mapping entry that can be used in the reverse direction of the communications. In order to do this, the AMS router performs a mapping lookup on the source address (which should already be done to perform the redirect). An AMFP push message is then sent to the forwarding node or host based on its locator.

4.5.3.2 Redirect rate limiting

An AMS router SHOULD rate limit the number of redirects it sends to a forwarder for each redirected address. The rate limit SHOULD be configurable. The default rate limit SHOULD be to send no more than one redirect per second per redirected identifier. If a mapping change is detected the rate limiting SHOULD be reset so that redirects for a new mapping can be sent immediately.

4.5.4 Map request/reply

An AMS forwarder may send a map request message to obtain mapping information for a locator. If the receiving AMS router has the mapping information it responds with a map information message. If the router does not have a mapping entry for the requested identifier, it MAY reply with in a locator type of Null.

Map requests are NOT RECOMMENDED as the primary means to dynamically populate entries in a mapping cache. The problem with this technique is that an AMS forwarder may generate a map request for each new destination that it gets from a downstream end host. A downstream end host could launch a Denial of Service (DOS) attack whereby it sends packets with random destination addresses that requires a mapping lookup. In the worst case scenario, the forwarder would send a map request for every packet received. Rate limiting the sending of map requests does not mitigate the problem since that would prevent the cache from getting mappings for legitimate destinations.

4.5.5 Push mappings

An AMS router may push mappings to an AMS forwarder without being requested to do so. This mechanism could be used to pre-populate a mapping cache. Pre-populating the cache might be done if the network has a very small number of identifiers or there are a set of identifiers that are likely to be used for forwarding in most AMS forwarders (identifiers for common services in the network for instance). When a mapping router detects a changed mapping, the locator changes for instance, a new mapping can be pushed to the AMS forwarders.

The push model is NOT RECOMMENDED as a primary means to populate an mapping cache since it does not scale. Conceivably, one could implement a pub/sub model and track of all AMS mappings and to which nodes the mapping information was provided. When a mapping changes, mapping information could be sent to those nodes that expressed interest. Such a scheme will not scale in deployments that have many mappings.

4.5.6 Cache maintenance

This section describes maintenance of a mapping cache.

4.5.6.1 Timeouts

A node SHOULD apply a timeout for a mapping entry that was indicated in a map information message. If the timeout fires then the mapping entry is removed. Subsequent packets may cause an AMS router to send a redirect so that the mapping entry gets repopulated in the cache.

The RECOMMENDED default timeout for identifiers is five minutes. If a node sends a map request to refresh a mapping, the RECOMMENDED default is to send the request ten seconds before the the mapping expires.

4.5.6.2 Cache refresh

In order to avoid cycling a mapping entry with a redirect for a mapping that times out, a node MAY try to refresh the mapping before timeout. This should only be done if the cache entry has been used to forward a packet during the timeout interval.

A cache refresh is performed by sending a map request for an identifier before its cache entry expires. If a map information message is received for the identifier, then the timeout can be reset and there are no other side effects.

4.5.7 AMS forwarder processing

If an AMS forwarder receives to its local address (i.e. a locator address) a packet that has undergone overlay forwarding, it will perform overlay termination. It will check its local mapping database to determine if the identifier revealed in the packet after overlay termination is local. If the identifier is local, the forwarder will forward the packet on to its destination which is either a downstream node that the forwarder has a route to, or a local VM or container in the case that the forwarder is an end host.

If the identifier is not local then the AMS forwarder forwards the packet back into the network after overlay termination. This may happen if an end node has moved to be attached to a different AMS forwarder and the new locator has not yet been propagated to all AMS nodes. The packet should traverse an AMS router which can send a mapping redirect back the source's AMS forwarder as described above. To avoid infinite loop in this process, the forwarder must decrement TTL in the packet being forwarded.

When a node migrates its point of attachment from one forwarder to another, the local mapping on the old node is removed so that any packets that are received and destined to the migrated identifier are re-injected without the overlay. A "negative" mapping with timeout may also be set ensure that the node is able to infer the destination address is a proper identifier for the mapping domain (e.g. would be needed with foreign identifiers).

4.5.8 Locator unreachable handling

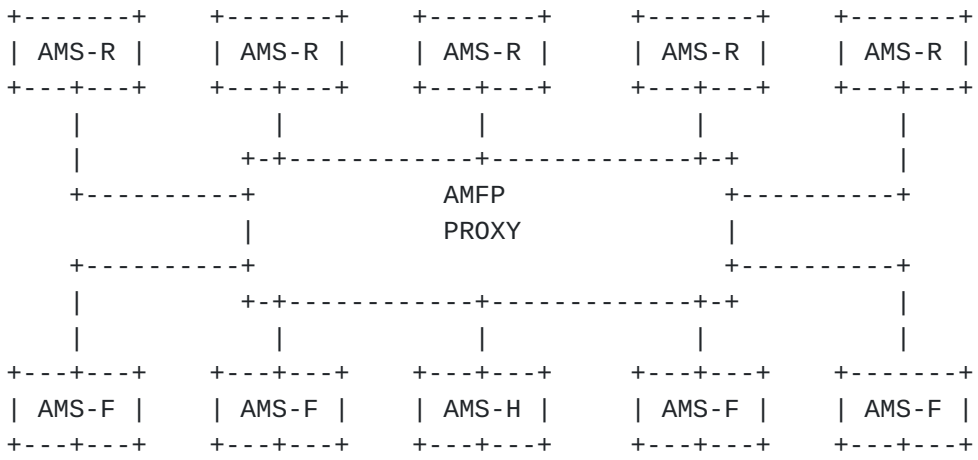
When connectivity to a locator is loss, the mapping system should

detect this. A locator unreachable message MAY be sent by AMS routers to AMS forwarders to inform them that a locator is no longer reachable. Each forwarder SHOULD remove any cache entries using that locator and MAY send a map request for the affected identifiers.

4.5.9 Control Connections

AMS forwarders must create AMFP connections to all the AMS routers that might provide routing information. In a simple network there may be just one router to connect to. In a more complex network with AMS routers for a sharded and replicated mapping system database there may be many. A list of AMS routers to connect to is provided to each AMS forwarder. This list could be provided by configuration, a shared database, or an external protocol to AMCP.

Conceivably, the number of AMS routers in a network that might report mapping information could be quite large (into the thousands). If managing a large number of connections at the AMS forwarders is problematic, AMS router proxies could be used that consolidate connections as illustrated below:



In the above diagram a single AMS router proxy serves five AMS routers and five AMS forwarders. The proxy creates one connection to each AMS router and each AMS forwarder creates one connection to the proxy.

4.5.10 Protocol errors

If a protocol error is encountered in processing AMFP messages then a node MUST terminate the connection. It SHOULD log the error and MAY attempt to restart the connection. There are no error messages defined in AMFP.

Protocol errors include mismatch of length for given data, reserved

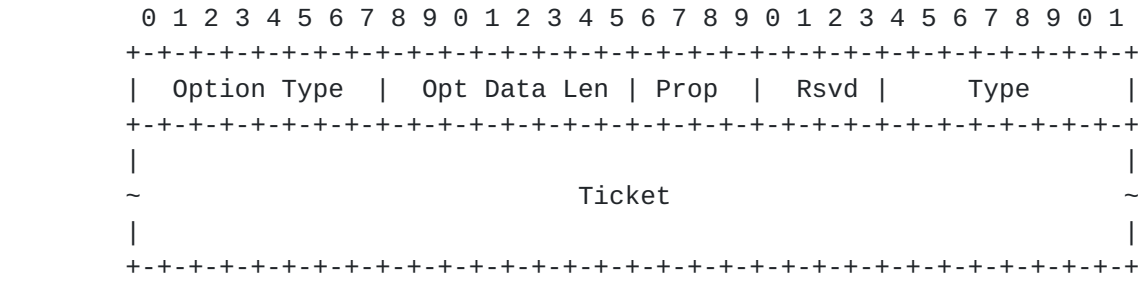
bits not set to zero, unknown identifier type or locator types, unknown reason, unknown overlay type or instructions, and loss of message synchronization in a TCP stream. Note that if the end of a message does not end on field or record boundary this also considered a protocol error.

5 Stateless mapping optimization

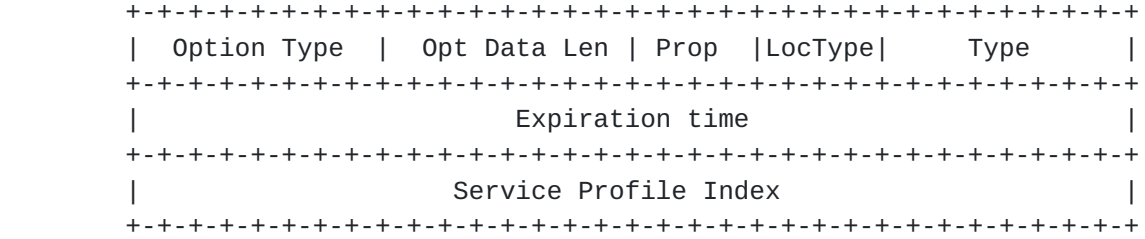
An alternative to requiring a mapping lookup on each packet is to encode the mapping information in packets themselves. This can be achieved by encoding mapping information in Firewall and Service Tickets. The basic concept is that mapping information is encoded in FAST tickets which are attached in packets at the end hosts and interpreted by the network. Tickets are associated with flows and are set in all the packets for the flow. Ticket reflection ensures that packets sent in the return path of a flow include a ticket.

5.1 Firewall and Service Tickets encoding

FAST tickets are encoded in Hop-by-Hop options. The format of a FAST ticket in a Hop-by-Hop option is:



[FAST] suggests a simple and efficient encoding of a Service Profile Index:

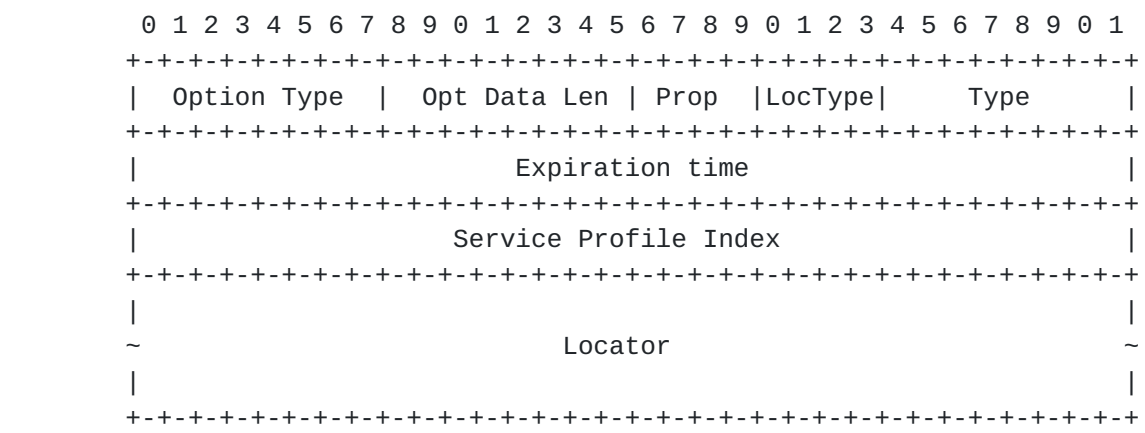


This format can be amended to include address mapping encoding.

5.2 Address mapping encoding

A locator address can directly encode in a ticket. Different address types can be used. A ticket with expiration time, service profile and

locator address may have format:



Pertinent fields are:

- LocType: Locator type. Specifies the locator type. This also implies the length the locator in the list. Locator types are defined above.

- Service Profile Index: Can encode the overlay method and a limited set of instructions for overlay forwarding.

- Locator: A locator of type indicated by LocType. The size of a locator is specified by the type.

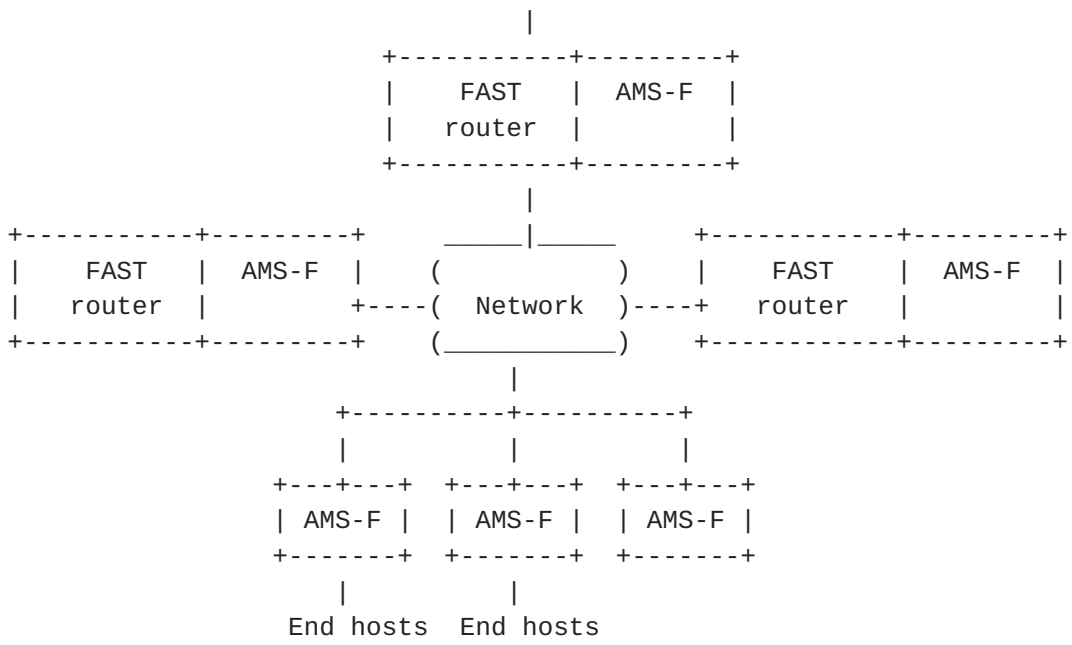
A network may have a comparatively small number of locators. For instance, a mobile provider might associate each eNodeB with a locator and there may only be a few million of these. In this case, the border routers might maintain a static table of locator addresses that can simply be indexed by number in a small range. Similarly, the backend server in the layer 4 load balancing case might also be indicated by an index into a table of backend servers.

5.3 Reference topology

As show in the reference topology below, FAST routers and AMS forwarders are involved in the stateless mapping datapath. AMS routers are not directly involved in the data path, however they serve the mapping information to be encoded into FAST tickets.

FAST routers interpret tickets and perform overlay forwarding. AMS forwarders terminate overlay forwarding. Note that an AMS forwarder and FAST router would be co-located so that a node processes FAST tickets and does AMS forwarding base on that.

Internet



5.4 Operation

This section describes the operation of encoding mapping entries in FAST tickets.

5.4.1 Ticket requests

Applications request FAST tickets from a ticket agent in the network local to the application. The ticket agent can return a ticket for the application to use in its data packets. The ticket includes information that is parsed by elements in the issuing network. The ticket information may include routing information. For example, if the application is on a mobile device, the network may provide a ticket that has a locator indicating the current location of the device.

[FAST] describes the process of an application requesting tickets and setting them in packets. An application will not normally need to make any special requests for routing information and the use of routing information is expected to be transparent to the application.

5.4.2 Qualified locators

There are two possibilities for locator information in an issued ticket:

- o The locator is fully qualified.

- 0 The locator is not qualified.

5.4.2.1 Fully qualified locators

If a locator is qualified then the issued ticket contains the locator for the end node. If the locator changes, that is the node moves, then a new ticket will need to be issued to the application.

5.4.2.2 Unqualified locators

If the locator is not qualified, then the locator information in the issued ticket contains a "not set" value. For instance, in the case the locator is expressed by a Locator Index then the "not set" value may be -1 (all ones). The AMS forwarder in the upstream path of an end node may write a locator value into the locator information to make it qualified; most often this would just be its own locator value in cases where it is the first upstream hop of an end devices that coincides with an AMS forwarder that provides location in the network. The implication is that this will be the locator used in the network overlay on the return path to reach the end node. Note that to write a locator into to a ticket requires that the ticket is in a modifiable Hop-by-Hop option.

5.4.3 AMS forwarder processing

Once an application has been issued a ticket with mapping information it will set the ticket in all packets sent to the peer node. The first hop upstream router, which might also be an AMS forwarder, in the FAST domain parses the ticket-- this may typically be the first hop router in a provider network closest to end user nodes.

If the ticket contains a qualified locator, the first hop node may validate it (as part of FAST ticket validation). If the ticket has unqualified locator information, the first hop node may set it to a qualified locator value in the packet. As described above, the locator information written is likely to be that corresponding to the locator of the first hop device which is an AMS forwarder.

5.4.4 Transit to the peer

Beyond the first hop router to the ultimate peer destination, no processing of mapping information in a ticket should be needed. Intervening networks and routers should deliver the ticket to the destination host unchanged.

At the peer host, the procedures described in [FAST] are followed to save the received ticket in a flow context and to reflect it in subsequent packets. As with other reflected tickets, one containing mapping information is treated as an opaque value that is not parsed or modified by the peer or any network outside of the origin network.

Packets sent by a peer will include reflected tickets for a flow. No processing of reflected mapping information in a ticket should be needed until the packet reaches the origin network of the ticket. Intervening networks and routers should deliver the ticket to the destination origin network unchanged.

5.4.5 Ingress into the origin network

At a border FAST router for the origin network, tickets are parsed and the encoded services are applied. If a ticket contains mapping information then the FAST router uses the information to perform overlay forwarding to the destination (the function of an AMS-F). Note that the FAST router performs no map query and does not need to maintain a mapping cache.

The service parameters contained in the ticket may provide additional instructions about how the packet is to be sent over the network overlay. For instance, the service parameters might indicate the packet is encrypted or to use some extensions of an encapsulation protocol.

5.4.6 Overlay termination

When a forwarded packet is received at the targeted AMS-F, normal procedures for overlay termination and forwarding the packet on to its destination are done.

At the end host, received reflected tickets are validated for acceptance as described in [FAST]. This is done by comparing the received ticket to that which was sent on the corresponding flow.

5.4.7 Fallback

The proposal described here is considered an optimization. Routing information in FAST tickets is not intended to completely replace a routing infrastructure. In particular, this solution relies on several parties to implement protocols correctly. For instance, the use of extension headers requires that they can be successfully sent through a network. As reported in [RFC7872], Internet support for forwarding packets with extension headers is not yet ubiquitous.

Therefore, a fallback is required when encoding mapping information in FAST is not viable for a flow. The fallback in AMS is to route packets through AMS routers.

5.4.8 Mobile events

When a mobile node moves and its locator changes, it is desirable to

converge to using the new locator as a quickly as possible. With tickets that contain locator information, a modified ticket needs to be sent to a peer host.

If an application was issued a ticket with qualified locator information then a new ticket needs to be issued. This can be done by the application receiving a signal that a mobile event has occurred causing it to make new ticket requests for established flows.

If an application has a ticket with an unqualified locator then the network should start writing the new locator information into packets that are sent by the application after the mobile event. This should be transparent to the application.

Note that in either case, in order to update the tickets that a peer is reflecting, the application needs to send packets to the peer that includes an updated ticket. There is no guarantee when an application may send packets, so there is the possibility of a window where the peer node is sending reflected tickets with outdated locator information. The window should be limited by the expiration time of a ticket (see below), however it is recommended to implement mechanisms to avoid communication blackholes. For instance, a "care of address" mapping entry could be installed at the old locator node to forward to the new one. Such solutions are also used to mitigate database convergence time or cache synchronization time.

5.4.10 Interaction with expired tickets

FAST typically expects ticket to have an expiration time. If a ticket is received before the expiration time and it is otherwise valid, then the packet is forwarded per the services indicated by the ticket. If a packet is received with an expired ticket, it might still be accepted subject to rate limiting. Accepting expired tickets is useful in the case that a connection goes idle and after some time the remote peer starts to send.

For tickets that are expired and contain mapping information, a FAST router should ignore the mapping information and take the fallback path. When an application sends new packets, it can include a fresh ticket so that the fast path is taken on subsequent packets. Ignoring the mapping information in expired tickets puts an upper bound on the window that outdated information can be used.

6 Privacy in Internet addresses

This section discusses the interaction between the address mapping system and privacy in Internet addressing. An address mapping system can facilitate strong privacy in Internet addressing. [ADDR-PRIV]

discusses privacy in addressing.

6.1 Criteria for privacy in addressing

Per [ADDR-PRIV], the ideal criteria for IPv6 addresses that provide strong privacy are:

- Addresses are composed of a global routing prefix and a suffix that is internal to an organization or provider. This is the same property for IP addresses [RFC4291].
- The registry and organization of an address can be determined by the network prefix. This is true for any global address. The organizational bits in the address should have minimal hierarchy to prevent inference. It might be reasonable to have an internal prefix that divides identifiers based on broad geographic regions, but detailed information such as location, department in an enterprise, or device type should not be encoded in a globally visible address.
- Given two addresses and no other information, the desired properties of correlating them are:
  - It can be inferred if they belong to the same organization and registry. This is true for any two global IP addresses.
  - It may be inferred that they belong to the same broad grouping, such as a geographic region, if the information is encoded in the organizational bits of the address.
  - No other correlation can be established. It cannot be inferred that the IP addresses address the same node, the addressed nodes reside in the same subnet, rack, or department, or that the nodes for the two addresses have any geographic proximity to one another.
- Geographic location of a node cannot be deduced from an address with accuracy.
- Given two observed addresses, no strong correlations can be drawn. In particular it must not be possible to correlate that two different flows originate from the same user.

6.2 Achieving strong privacy

Strong privacy in addressing can be achieved by using a different randomly generated identifier source address for each flow. Conceptually, this would entail that the network creates

and assigns a unique and untrackable address to a host for every flow created by the host.

In this scheme, each host would be assigned many addresses which are non-topological in the local network to both promote privacy and mobility. An identifier-locator protocol with an address mapping system can provide reachability. This would entail that the addressing mapping system contains a mapping entry for each ephemeral address.

In large networks this solution presents an obvious scaling problem. Assigning an address per connection is a potential scaling problem on two accounts:

- The amount of state needed in the address mapping system is significant.

- Bulk host address assignment is inefficient.

6.3 Scaling network state

The amount of state necessary to assign each flow its own unique source IP address is equivalent, or at least proportional, to the amount of state needed for CGNAT-- basically this is one state element for every connection in the network. So in one sense this solution should scale as well as NAT has.

6.3.1 Hidden aggregation

A possible solution to reduce state is to make addresses aggregable, but use an aggregation method that is known only by the network provider and hidden to the rest of the world. The network could use a reversible hash or encryption function to create addresses. This method is called "hidden aggregation".

The input to an address generation function includes a group identifier, a secret key, and a generation index.

The function may have the form:

```
Address = Func(key, group_ident, gen)
```

Where "key" is secret to network, "group_ident" is a network internal identifier for an aggregated set of addresses (roughly equivalent to "identity" in IDEAS), and "gen" is generation number 0,1,2,... N. The generation value is changed for each invocation to create different addresses for assignment to a node.

When a network ingress node is forwarded a packet it performs the inverse function on an address.

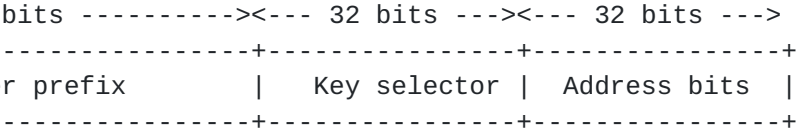
The inverse function has the form:

(group_ident, gen) = FuncInv(key, Address)

The returned group_ident value is used as the identifier in the mapping lookup for a locator address. In this manner, the network can generate many addresses to assign to a node where they all share a single entry in the mapping system.

6.3.2 Address format

A possible address format for hidden aggregation is shown below.



Note the that provider prefix is not hidden, so the address does identify the network provider of a user. Key selector is an index into a table of keys. A key table should have at least 2^16 entries that are randomly generated and securely shared amongst AMS routers. Hosts can be assigned addresses in blocks based on a key, however the same key should be used for different hosts assignments and end hosts should be assigned blocks from different keys.

The address bits are used to create unique addresses per key. A decoded address may contain a magic value to verify the hash function.

Keys should be rotated periodically. Addresses assigned using a particular key will therefore have an expiration, the default expiration time should be one week (assuming one of 2^16 keys in table are rotated each minute).

6.3.3 Practicality of hidden aggregation methods

The premise of hidden aggregation is that only trusted devices in the network are able decode the aggregation hidden within IPv6 addresses. This implies that the network must keep secrets about the process. In the above examples, the secrets are keys used in the hash or encryption. The security of the key is then paramount, so techniques for key management, rotation, and using different key sets for obfuscation are pertinent.

To perform a mapping lookup a node must apply the inverse address generation function to map addresses to their group identifiers. This lookup would occur in the critical data path so performance is important. Encryption and hashing are notoriously time consuming and computationally complex functions.

Some possible mitigating factors for performance impact are:

- The input to address generation functions is a small amount of data and has fixed size. The input is a key (presumably 128 or 256 bits), part of all of an IPv6 address (128 bits), and a generation number (sixteen to twenty-four bits should work).

- Given that the input is fixed size, specialized hardware might be used to optimize performance of the inverse address generation function. For instance, modern CPUs include instructions to perform crypto [AES-NI]. Since the keys used in these functions are secret to the network and there are relatively few of them, they might be preloaded into a crypto engine to reduce setup costs.

- The output of an inverse address generation function is cacheable. A cache on a device could contain address to locator mappings. When the inverse function and lookup on group_ident are performed, a mapping of address to the discovered locator could be created in the cache. The node could then map addresses in subsequent packets sent on the same flow to the proper locator by looking up the address in the cache.

6.4 Scaling bulk address assignment

Assigning multiple addresses without aggregation is difficult to scale. Each address would need to be individually specified in an assignment sent to a host.

DHCPv6 might allow bulk singleton address assignment. As stated in [RFC7934]:

> Most DHCPv6 clients only ask for one non-temporary address, but the protocol allows requesting multiple temporary and even multiple non- temporary addresses, and the server could choose to provide multiple addresses. It is also technically possible for a client to request additional addresses using a different DHCP Unique Identifier (DUID), though the DHCPv6 specification implies that this is not expected behavior ([RFC3315], Section 9). The DHCPv6 server will decide whether to grant or reject the request based on information about the client, including its DUID, MAC

address, and more. The maximum number of IPv6 addresses that can be provided in a single DHCPv6 packet, given a typical MTU of 1500 bytes or smaller, is approximately 30.

7 Address Mapping System in 5G networks

The section describes applying AMS for use in 5G networks. AMS is instantiated as a function in the 5G services architecture described in [3GPPTS].

7.1 Architecture

The figure below depicts the use of AMS in a 5G reference point architecture. AMS is logically a network function and AMS interfaces to the 5G control plane via service based interfaces.

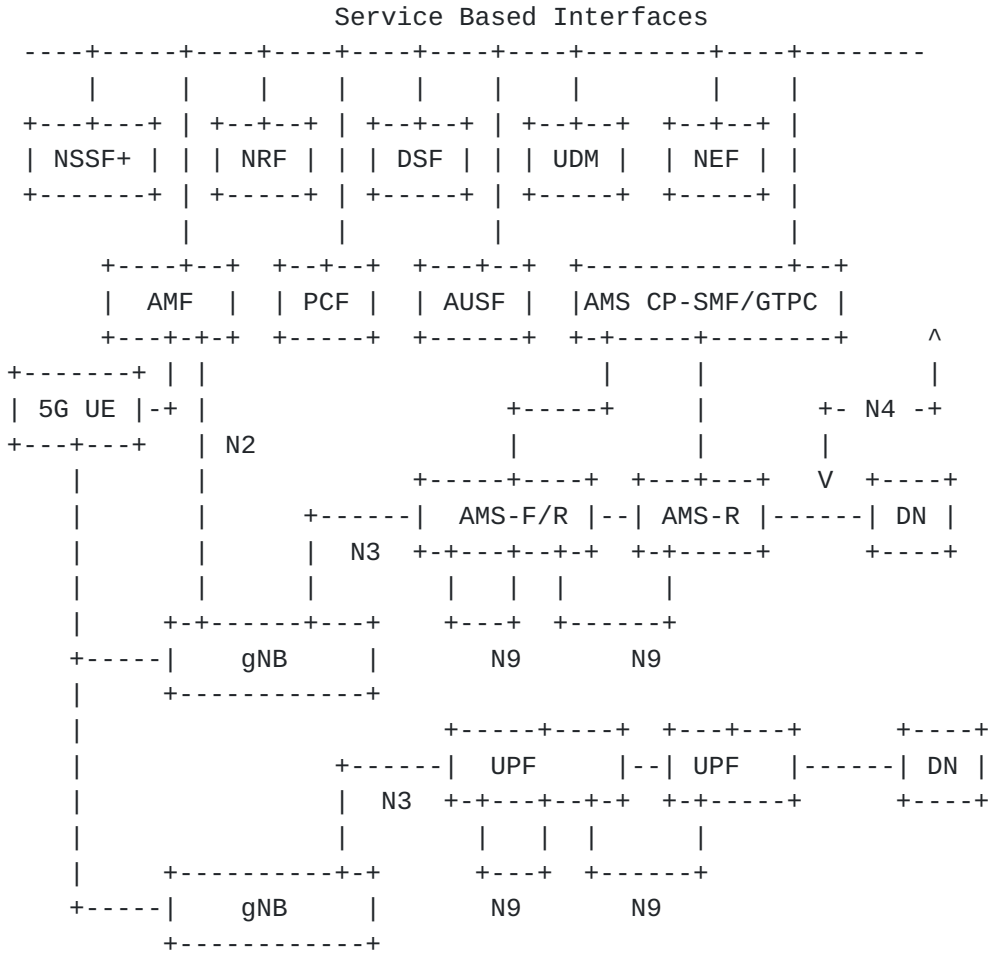


Figure 2: AMS in 5G reference point architecture

AMS is used over the N3 and N9 interface. Address mappings in the downlink from the data network are done by an AMS-R. Transformations

for edge traffic can be done by an AMS-F close to the gNB or by an AMS-R in the case of a cache miss.

The control interface into AMS is via N4 interface that interacts with 5G network services. AMS Control Plane node (AMS-CP) uses RESTful APIs to make requests to network services (see section 7.3). An AMS-CP receives notifications when devices enter the network, leave it, or move within the network. The AMS-CP writes the address mapping entries accordingly.

AMS-CP communicate with other AMS-CPs, AMS-Fs, and AMS-Rs in the same routing domain via control protocols that are independent of the 5G control plane. The mapping database is shared amongst AMS-CP and AMS-Rs utilizing underlying distributed database technology deployed.

7.2 Protocol layering

Figure 3 illustrates the protocol layers of packets packets sent over various data plane interfaces in the downlink direction of data network to a mobile node. Note that this assumes the topology shown in Figure 2 where GTP-U is used over N3 and IP routing is used on N9.

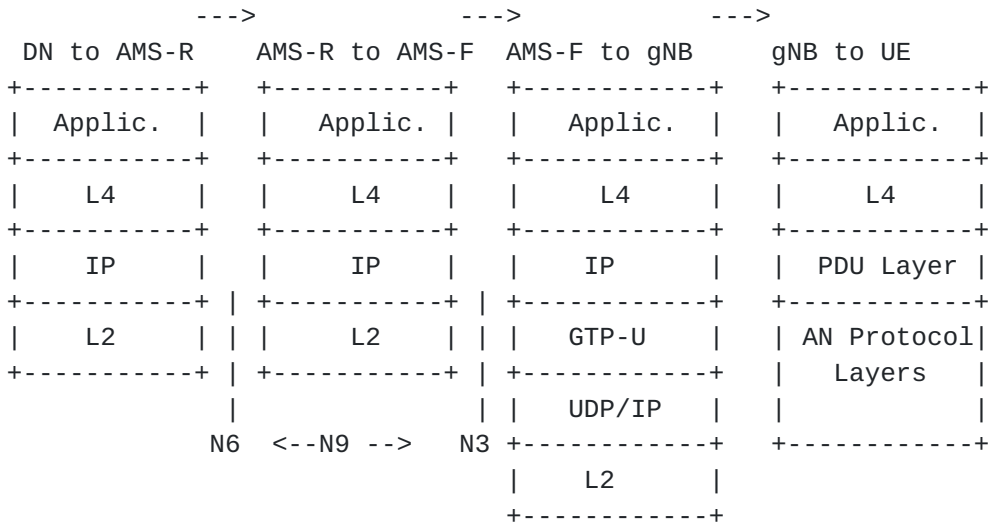
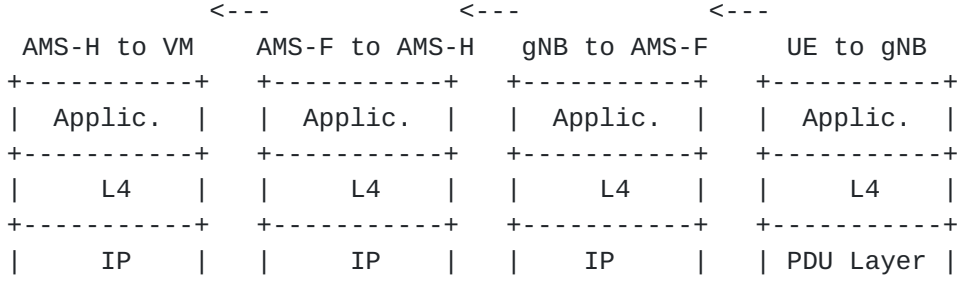


Figure 3: AMS and protocol layer in Downlink core



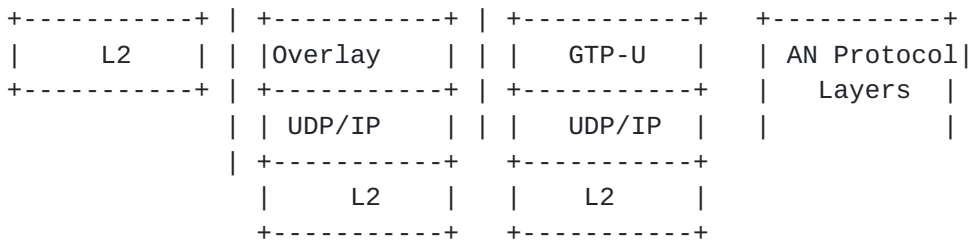


Figure 3: AMS and protocol layer in uplink MEC

7.3 Control plane between AMS and network

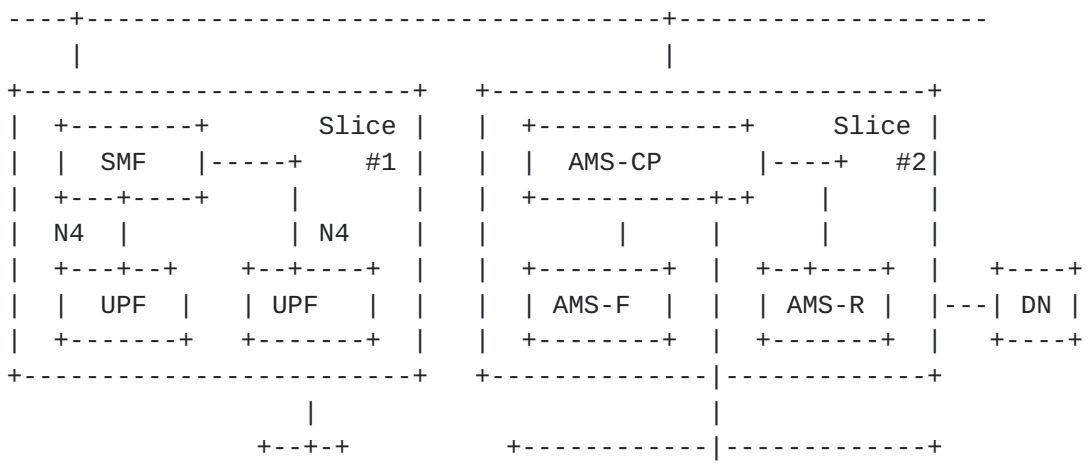
AMS is a consumer of several 5G network services. The service operations of interest to AMS are:

- Nudm (Unified Data Management): Provides subscriber information.
- Nsmf (Service Managment Function): Provides information about PDU sessions.
- Namf (Core Access and Mobility Function): Provides notifications of mobility events.

AMS-CP subscribes to notifications from network services. These notifications drive changes in the address mapping table. The service interfaces reference a UE by UE ID (SUPI or IMSI-Group Identifier), this is used as the key in the AMS identifier database to map UEs to addresses and identifier groups. Point of attachment is given by gNB ID, this is used as the key in the AMS locator database to map a gNB to an AMS-F and its locator.

7.4 AMS and network slices

Figure 4 illustrates the use of network slices with AMS.



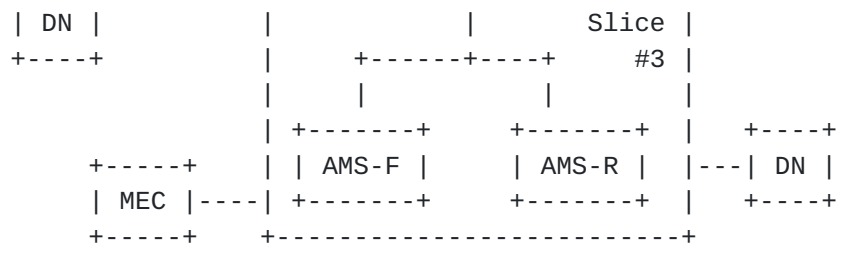


Figure 4: AMS and network slices in 5G

In this figure, slice #1 illustrates legacy use of UPFs without AMS in a slice. AMS can be deployed incrementally or in parts of the network. As demonstrated, the use of network slices can provide domain isolation for this.

Slice #2 supports AMS. Some number ofAMS-Fs and AMS-Rs are deployed. Address transformations are performed over the N9 interface. AMS-Rs would be deployed at the N6 interface to perform address transformations on packets received from a data network. AMS-Fs will be deployed deeper in the network at one side of the N3 interface. AMS-Fs may be supplemented by AMS-Rs that are deployed in the network. AMS-CP manages the mapping database within the slice.

Slice #3 shows another slice that supports AMS. In this scenario, the slice is for Mobile Edge Computing. The slice contains AMS-Rs and AMS-Fs, and as illustrated, it may also contain End hosts that run directly on edge computing servers. Note in this example, one AMS-CP, and hence one routing domain, is shared between slice #2 and slice #3. Alternatively, the two slices could each have their own AMS-CP and define separate routing domains.

7.4 AMS in 4G networks

The 4G architecture in 3GPP implements an address mapping system that is consistent with the architecture described in this document. Serving gateways have the role of AMS routers and GTP-U is the AMS routing protocol in 3GPP. 3GPP is based on an anchored routing model, the protocol can be augmented with AMS forwarders to achieve anchorless routing. Note that this can be done as an incremental addition to the 3GPP model, and in particular the core model and protocols of 3GPP, including GTP-C and GTP-C, require no change. The addition of AMS forwarders and mapping caches is done as an optimization for handling critical, low latency applications.

7.5 Overlay forwarding

As described in section X, AMS forwarders may be implemented on servers. For instance, a mobile network may have server farms that

provide VMs for running services close to users. For both performance and feasibility, it may be preferable for such servers to use an alternative overlay method than GTP. This document highlights that Generic UDP Encapsulation (UE) or Identifier Locator Addressing (ILA) may be good alternatives. GUE is a generic and extensible encapsulation protocol with good performance, ILA is identifier/locator split protocol that works with IPv6 and has very good performance.

8 Security Considerations

AMFP must have protection against message forgery. In particular secure redirects and mapping information message are required to prevent and attacked from spoofing messages and illegitimately redirecting packets. This security is provided by using TCP connections so that origin of the messages is never ambiguous.

Transport Layer Security (TLS) [RFC5246] MAY be used to provide secrecy, authentication, and integrity check for AMFP messages.

The TCP Authentication Option [RFC5925] MAY be used to provide authentication for AMFP messages.

9 IANA Considerations

10 References

10.1 Normative References

[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>.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

[ILA] Herbert, T., and Lapukhov, P., "Identifier Locator Addressing for IPv6" draft-herbert-intarea-ila-00

10.2 Informative References

[RFC5246]] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, DOI 10.17487/RFC5246, August 2008, <https://www.rfc-editor.org/info/rfc5246>.

[RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, DOI 10.17487/RFC5925, June 2010, <https://www.rfc-editor.org/info/rfc5925>.

[BGPILA] Lapukhov, P., "Use of BGP for dissemination of ILA mapping information" draft-lapukhov-bgp-ila-afi-02

Author's Address

Tom Herbert
Quantonium
Santa Clara, CA
USA

Vikram Siwach

Email: tom@quantonium.net