INTERNET-DRAFT
Intended Status: Standard
Expires: August 2019

T. Herbert
Quantonium
V. Siwach
Independent consultant

February 21, 2019

Address Mapping System
draft-herbert-intarea-ams-01

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 to use, as well as instructions related to using the method. The Address Mapping System has a number of potential use cases including identifier-locator protocols, network virtualization, and promotion of privacy.

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 a fast moving automobile may frequently connect to different cells). Protocols are defined to synchronize 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

  Container virtualization and Virtual Machines are popular techniques for malleable and efficient use of compute resources in datacenters. A key function in network virtualization is to map virtual addresses to physical addresses. The physical address represents the location of a virtual node. An overlay technique, such as an encapsulation protocol like VXLAN [RFC7348], GUE [GUE], Geneve [GENEVE], or GTP [GTP], is used to forward a packet to its virtual destination based on the physical address associated with a virtual address. The address mapping system provides the necessary mapping information and allows for mobility in container or VM migration.

- Identifier/locator protocols

  Identifier/locator protocols generalize the addressing model of network virtualization. These include a group of protocols and proposals that are being discussed in IETF which resolve the currently strong correlation in IP addresses between identification of a communication end point and the topological location in the network. Identifier/locator protocols include LISP [RFC6830], ILNP [RFC6740], and ILA [ILA]. These demand mechanisms for rapid lookup and notification of the correlation between identifiers of hosts and where they are located. The address mapping system provides this.

- Network function virtualization

  Network function virtualization [NFV] deployed in distributed data centers, or the cloud, requires addressing of dedicated network function instances that fulfils stringent performance requirements. This is achieved by an efficient mapping of network function (NF) logical name to an instance which the address mapping system facilitates.

- Address resolution

  Address resolution refers to the general concept of resolving a higher layer address into a lower layer address. For instance, in Ethernet, a network (IP) address is resolved to a link layer (MAC) address via IPv4 ARP (Address Resolution Protocol) or IPv6 NDP (Neighbor Discovering Protocol). The address mapping system provides an alternative system for address resolution.

- Privacy in Internet addressing

  IP addressing is a privacy concern when addresses embed information that can be used to infer the geographic location, identity, or correlations in unrelated communications of a user. Discussions on this topic and countermeasures have been scope of numerous activities at IETF ([RFC4941], [RFC6462], [RFC7721], [ADDRPRIV], [IDLOCPRIV]). An address mapping system can be used as a basis for a solution as described in section 7.

- Mobile networks

  Mobile networks, where the temporary location of a moving device is typically changing more or less rapidly, require resolution of the address of the current point of attachment (radio base station) per device identifier. During an active session the serving base station may change (handover) and the traffic is

rerouted to and from the new point of attachment's address. Whereas cellular networks so far have applied mainly proprietary procedures and 3GPP protocols [3GPP15] to mobility, forthcoming 5G architectures allow multiple heterogeneous access technologies and may employ IP-based mechanisms. The address mapping system could provide the mapping between a client address and its current point of attachment. Use of AMS in a 5G service based architecture is described in section 8.

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 an identifier/locator address split protocol such as ILA or ILNP. A network may simultaneously use different overlay protocols per its needs. Mapping information provided by the address mapping system indicates the overlay technique and overlay technique specific instructions to use 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 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 leakage of sensitive personal information. The mapping system can also foster privacy in addressing by supporting untrackable, per-flow IP addresses.

- 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.

- 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.

- Stateless mapping mode

  An address mapping system may provide a communication mode where the mapping information is carried in packets themselves. When a packet that 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 address
: A network address that is an object in the address mapping system table. Mapping addresses are typically 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 system database. 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.

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 method/overlay protocol
: A method or protocol that implements overlay forwarding. Overlay methods include encapsulation and address transformation.

Overlay instructions
: A set of instructions that are specific to an overlay method Instructions can describe how the method is used and optional protocol extensions or security parameters to use with the overlay method.

Overlay termination
: The processing done at the terminal endpoint of an 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 will often act as 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. An 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 address 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.

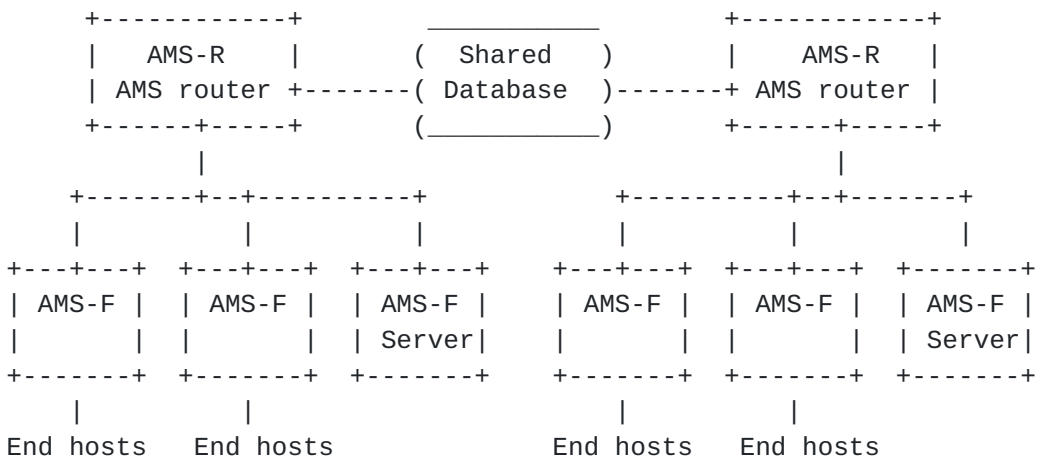
Hidden aggregation
: A method to encode aggregation in network addresses where the aggregation is visible to trusted devices within a network, but is transparent to external observers of the addresses.

2 Architecture

This section describes the architecture of the Address Mapping System.

2.1 Reference topology

The diagram below provides a generic reference topology for AMS.



2.2 Functional components

There are two fundamental types of 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 (AMRP, see section 3).

AMS routers have three primary functions:

- Serving mapping information
- Overlay forwarding
- Sending redirects

2.3.1 Serving mapping information

AMS routers serve mapping information to AMS forwarders via the Address Mapping Forwarder Protocol (AMFP, see section 4). Mapping information is provided by a request/reply protocol, a push mechanism, or mapping redirects.

2.3.2 Overlay forwarding

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 the 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 points when they are forwarding packets (using 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.

2.3.3 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 performed in the mapping system database

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

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

   - Optionally, a mapping redirect is 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, there may be some datapath optimizations that can be applied.

AMS forwarders may have two functions:

- Overlay termination which is restoring packets 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 forwarding 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 (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 address mapping system 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 a 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

BGP can be used to propagate mapping information amongst AMS routers as simple routes. [BGPOLAY] describes a scalable method for using BGP in overlay networks. [BGPILA] describes a method to distribute identifier to locator information using Multiprotocol Extensions for BGP-4.

3.3 GTP

GPRS tunneling protocol (GTP) is the primary protocol for control and user plane in 4G and has been adopted in 5G service based architecture where control and user plane is separated. GTP tunnels are data plane encapsulation programmed for subscribers as point to point segments between the network elements from enodeB to SGW (Serving Gateway) to PGW (Packet Gateway) in 4G and gnodeB (5G base station) to UPF (User Plane Function) in 5G.

AMS scheme allows to migrate the GTP Anchors like PGW and UPF to open the network distributed application for mobility.

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.

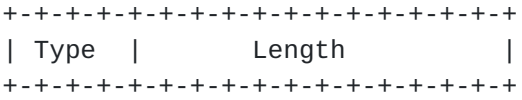
AMFP is primarily used between AMS forwarders and AMS routers. The purpose of the protocol is to populate and maintain the mapping cache in AMS forwarders.

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

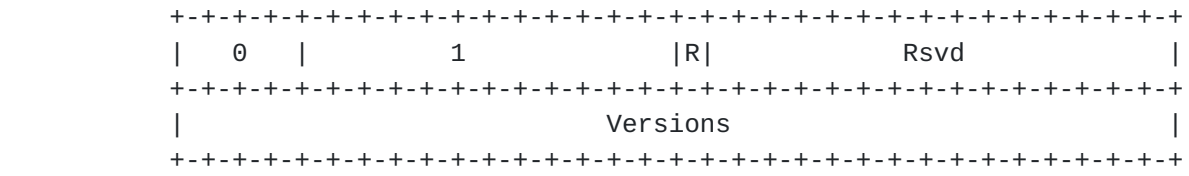
of four bytes in length and the message length includes the two bytes for the common header. The length field is computed as (message_length / 4) - 1, so the minimum message size is four and the maximum size is 16,384 bytes.

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

4.2 Hello messages

Hello messages indicate the versions of AMFP that a node supports. A 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.
- Length = 1. Indicates eight byte length.
- 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.
- Versions: A bit map of supported versions. Bit 0 refers to version 0, bit 1 refers to version 1, etc. If a bit is set then the corresponding version is supported by a node.

Version numbers are from 0 to 31. Version 0-15 will defined by IANA, and versions 16 to 31 are user defined. This document describes version 0 of AMFP.

4.3 Version negotiation

The first message sent by each side of an AMFP connection is a Hello message. Hello messages indicate the set of AMFP versions that a node supports.

When a host receives an 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 that versions 0,1,3, and 4 are supported and receives a peer Hello message with versions 1 and 2 indicated as being supported; 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 peers, that is their sets of supported versions 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.

If the first message received on a connection is not a Hello message, then that is an error so the connection MUST be terminated and an error MAY be logged. If a second Hello message is received on a connection, then that is also considered an error so the connection MUST be terminated and an error MAY be logged.

5 AMFP Version 0

This section describes the message types and operation of version 0 of AMFP.

5.1 Message types

The message types in version 0 of AMFP are:

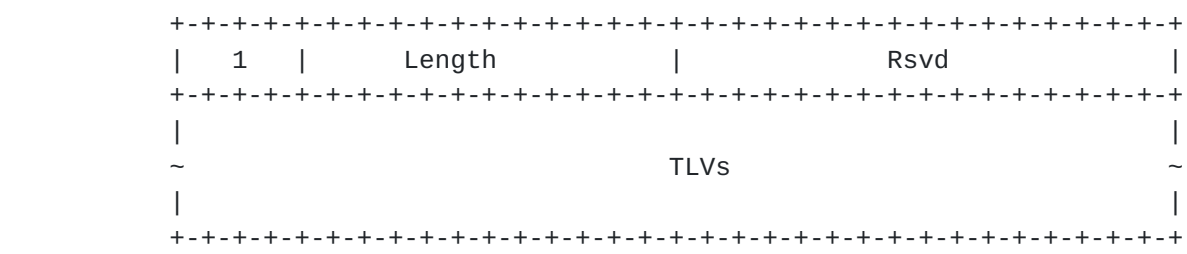
- Parameters (Type = 1)
- Map request (Type = 2)
- Map information (Type = 3)
- Compressed map information (Type = 4)
- Locator unreachable (Type = 5)
- Cache occupancy (Type = 6)

5.2 Parameters message

A Parameters message contains AMFP related parameters. The parameters

are encodes in TLVs. A Parameters message MUST be sent by each side as the first message after the AMFP version negotiation is completed.

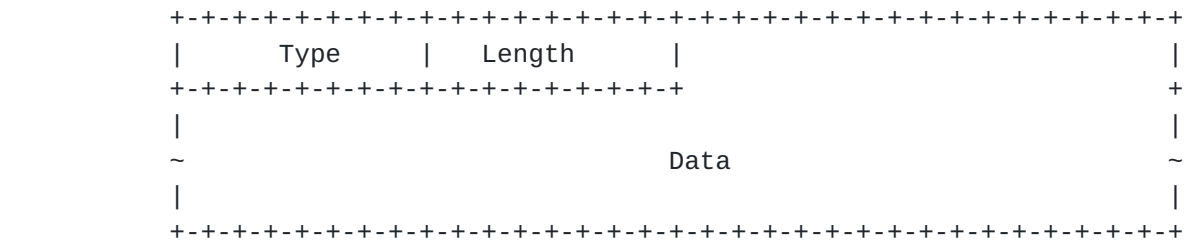
The format of an AMFP Parameters message is:



The contents of the Parameters message are:

- Type = 1. This is indicates a Parameters message.
- Length: Set to the length of the TLVs divided by four.
- Rsvd: Reserved bits. Must be set to zero when sending.
- TLVs: A list of TLVs that describe capabilities or requested options.

The format of a TLV is:



Fields are:

- Type: Type of the TLV.
- Length: Length of the TLV 32-bit units not include the first four bytes of the TLV. The minimum length of a TLV is four bytes and the maximum length is 1024 bytes.

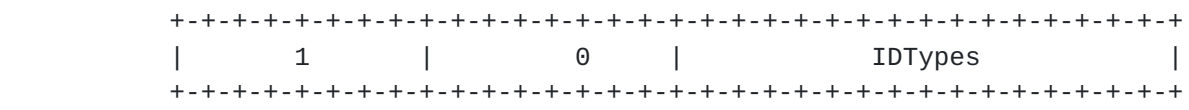
The table below lists the Parameters TLVs defined in this document. The "Length" column indicates any length requirements on TLVs, and the "Sender" column indicates whether the TLV can be sent by the router, forwarder, or both sides.

Type	Length	Sender	Meaning

0			RESERVED
1	4	Either side	Supported identifier types
2	4	Either side	Supported locator types
3	variable	Either side	Supported overlay methods
4	4	Router	Default overlay method
5	8	Router	Default timeout
6	4	Router	Default priority
7	4	Router	Default weight
8	variable	Router	Default instructions
9-127			UNASSIGNED (assignable by IANA)
128-255			User defined

5.2.1 Supported identifier types

This TLV provides the identifier types that a node supports. The TLV can be sent by either an AMS-R or an AMS-F. The format of the TLV is:



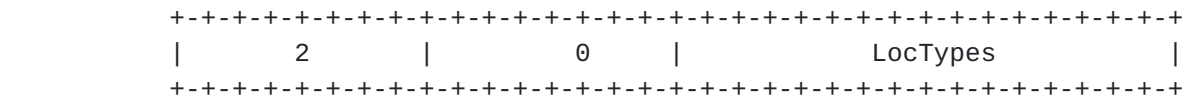
Fields are:

- Type = 1
- Length: Set to 0 to indicate four bytes length.
- IDTypes: A bitmap that indicates supported identifier types. The position in the bitmap corresponds to the defined values for identifier type. Identifier types are defined below.

If the supported identifier types TLV is not received then a node assumes that supported identifier types by a peer is unknown.

5.2.2 Supported locator types

This TLV provides the locator types that a node supports. The TLV can be sent by either and AMS-R or an AMS-F. The format of the TLV is:



Fields are:

- Type = 2

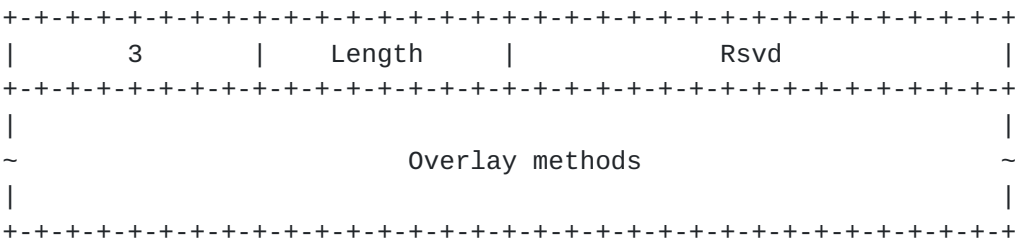
- Length: Set to 0 to indicate four bytes length.

- LocTypes: A bitmap that indicates supported locator types. The position in the bitmap corresponds to the defined values for locator types. Locator types are defined below.

If the supported locator types TLV is not received then a node assumes that supported locator types by a peer is unknown.

5.2.3 Supported overlay methods

This TLV provides the overlay methods that a node supports. The TLV can be sent by either an AMS-R or an AMS-F. The format of the TLV is:



Fields are:

- Type = 3

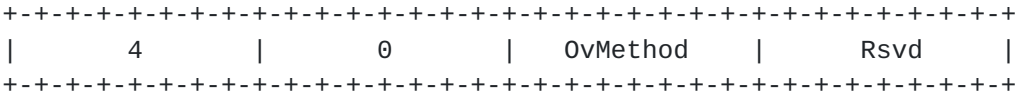
- Length: Set to length of the overlay methods bitmap divided by four. The overlay methods bitmap is padded with zeroes if necessary to align message length to four bytes.

- Overlay methods: A variable length bit map that indicates overlay methods. The position in the bitmap corresponds to the defined values for the various overlay methods. Overlay methods are defined in section 10.

If the supported overlay methods TLV is not received then a node assumes that supported overlay methods in a peer is unknown.

5.2.4 Default overlay method

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



Fields are:

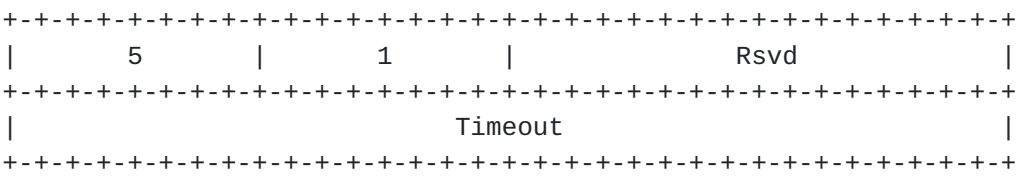
- o Type = 4
- o Length = 0, indicating four bytes length
- o OvMethod: Indicates the default overlay method to be used when sending to a locator.
- o Rsvd: Reserved bits. Must be set to zero when sending.

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

The default overlay method SHOULD be negotiated. If it's not negotiated then the default method is undefined.

5.2.5 Default timeout

This TLV provides the default timeout for reported mapping information when the timeout is not explicitly provided in a mapping information message.



Fields are:

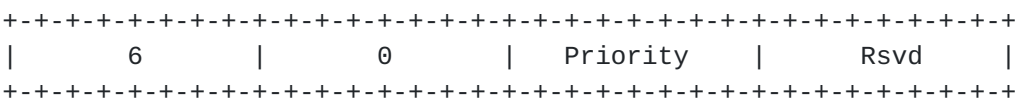
- o Type = 5
- o Length = 1, indicating eight bytes length
- o Rsvd: Reserved bits. Must be set to zero when sending.
- o Timeout: The default time to live for the identifier information in seconds.

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

If the default timeout is not negotiated then the assumed default is 300 seconds (five minutes).

5.2.6 Default priority

This TLV provides 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:

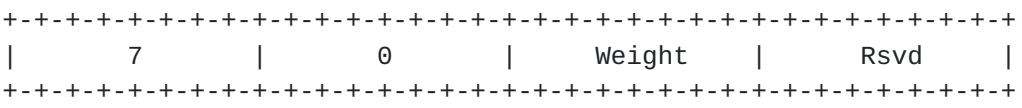
- Type = 6
- Length = 0, indicating four bytes length
- Priority: Default relative priority of a locator. Locators with higher priority values have preference to be used. Locators that have the same priority may be used for load balancing.
- Rsvd: Reserved bits. Must be set to zero when sending.

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

If the default priority is not negotiated then the assumed default value is zero.

5.2.7 Default weight

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



Fields are:

- Type = 7
- Length = 0, indicating four bytes length
- Weight: Relative weight 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.

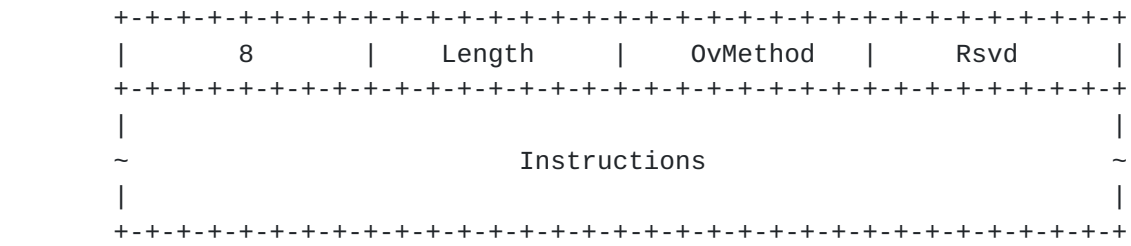
- Rsvd: Reserved bits. Must be set to zero when sending.

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

If the default weight is not negotiated then the assumed default value is zero.

5.2.8 Default instructions

This TLV provides 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:

- Type = 8
- Length: Set to length of instructions divided by four.
- OvMethod: Indicates the overlay method associated with the instructions.
- Rsvd: Reserved bits. Must be set to zero when sending.
- Instructions: Data with format and semantics that are specific to an overlay method and describe options for the method and how the overlay method is used.

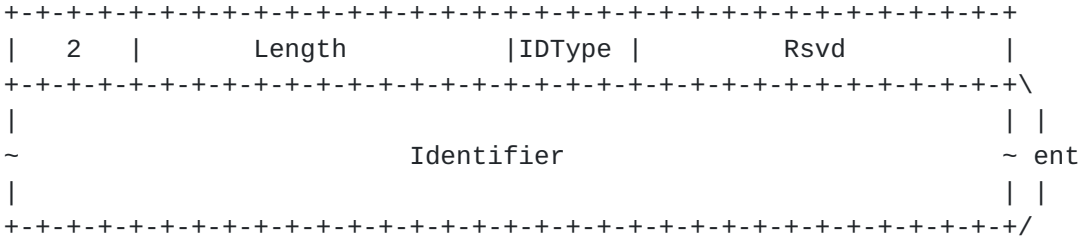
Only AMS routers send this TLV. If the TLV is received by a router it is considered an error. The TLV may sent multiple times for different overlay methods.

If default instructions are not negotiated then the assumed default value is no instructions.

5.3 Map Request message

A Map Request message 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 message is:



The contents of the Map Request message are:

- Type = 2. This is indicates a Map Request message
- 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) / 4.
- 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).

This message MUST only be sent by an AMS forwarder. If an AMS forwarder receives a Map Request message it is considered an error.

5.4 Map Information message

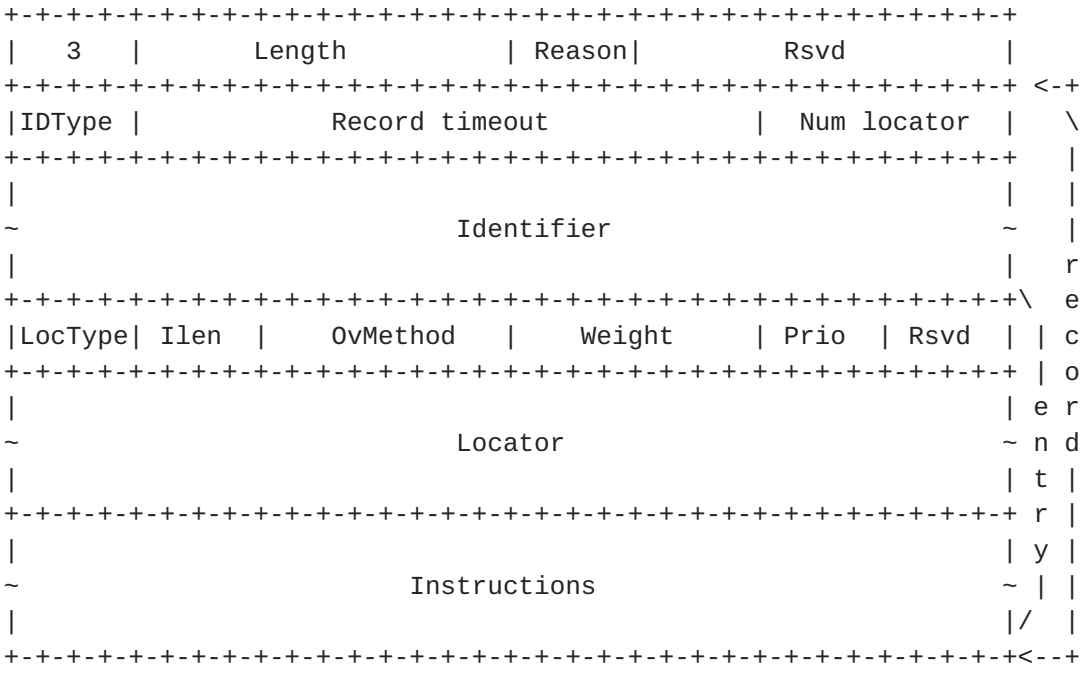
A Map Information message is sent by an AMS router to provide mapping information. In addition to providing locators for an identifier, the message also contains the overlay method to use and related instructions for sending to an identifier.

A Map Information message is composed of a 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 the identifier. A locator entry is composed of a four byte header that includes the overlay method to use, the

locator, and optional 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. Both records and entries are variable length. 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 = 3. This indicates a Map Information message
- Length: Set to the sum of lengths of all the identifier records in the message divided by four. The length of an identifier 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 the instruction field.
- Reason: Specifies the reason that the message was sent. Reasons are:
  - 0: Map reply to a map request
  - 1: Redirect

    o 2: Push map information

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

The contents of an identifier record are:

  o IDType: Identifier type. Specifies the identifier type. This also implies the length of each identifier in the list. Identifier types are defined below.

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

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

  o Identifier: An identifier of type specified in IDType.

The contents of a locator entry are:

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

  o 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 how the overlay is used. The instructions length is from zero to sixty bytes.

  o 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 section 10.

  o 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.

  o Prio: 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.

  o 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 and how the overlay method is used. Ilen indicates the length of the field.

This message MUST only be sent by an AMS router. If an AMS router receives a Map Information message it is considered an error.

5.5 Compressed Map Information message

The Compressed Map Information message may be sent as an efficient alternative to the Map Information message. The Compressed Map Information can be used when all these conditions are met:

- There is only locator provided for each identifier

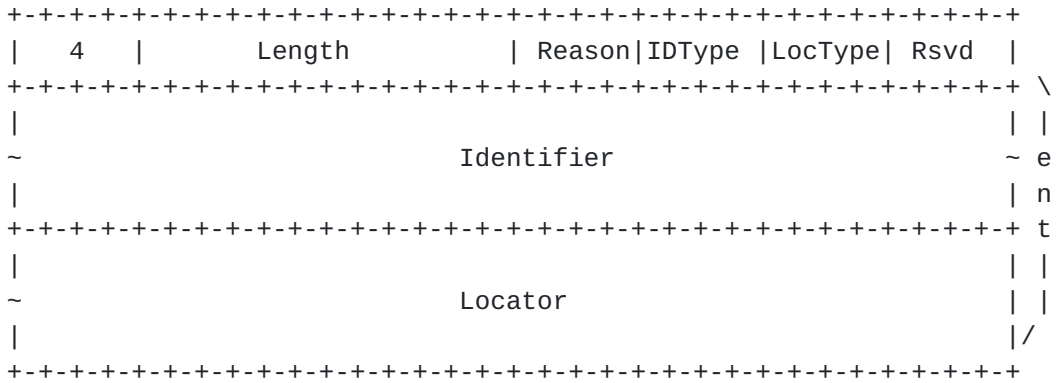
- The identifier type and locator type are common for all the mappings reported in the message

- The priority, weight, overlay method, record timeout, and overlay instructions are the default values negotiated for the AMFP session

A Compressed Map Information message is composed of a four byte header followed by a list of identifier/locator pairs.

The identifier/locator pairs are repeated for each mapping being reported. The total number of identifiers being reported can computed as (message_length - 4) / (identifier_size + locator_size).

The format of the Compressed Map Information message header is:



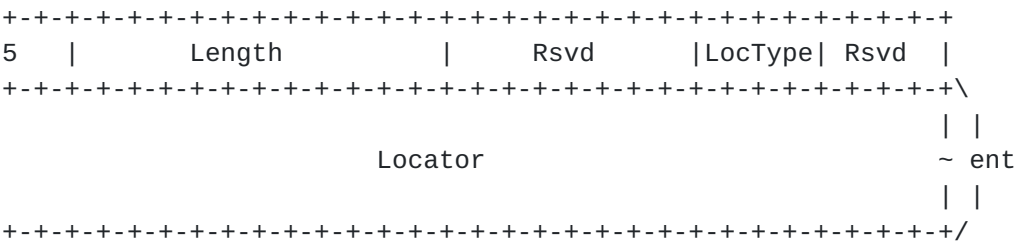
The contents of the Compressed Map Information message header are:

- Type = 4. This indicates a Compressed Map Information message
- Length: Set to the sum of lengths of the identifier/locator pairs in the message divided by four
- 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.
- Rsvd: Reserved bits. Must be set to zero when sending
- Identifier: An identifier of type specified in IdType.
- Locator: A locator of type specified in LocType.

This message MUST only be sent by an AMS router. If an AMS router receives a Compressed Map Information message it is considered an error.

5.6 Locator Unreachable message

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 = 5. This indicates a Locator Unreachable message.
- Length: Set to the size of the locator times the number of locators in the list divided by four.
- LocType: Specifies the locator type. This also implies the length of each locator in the list. Locator types are defined below.
- Rsvd: Reserved bits. Must be set to zero when sending.
- 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).

This message MUST only be sent by an AMS router. If an AMS router receives a Locator Unreachable message it is considered an error.

5.7 Identifier and locator types

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

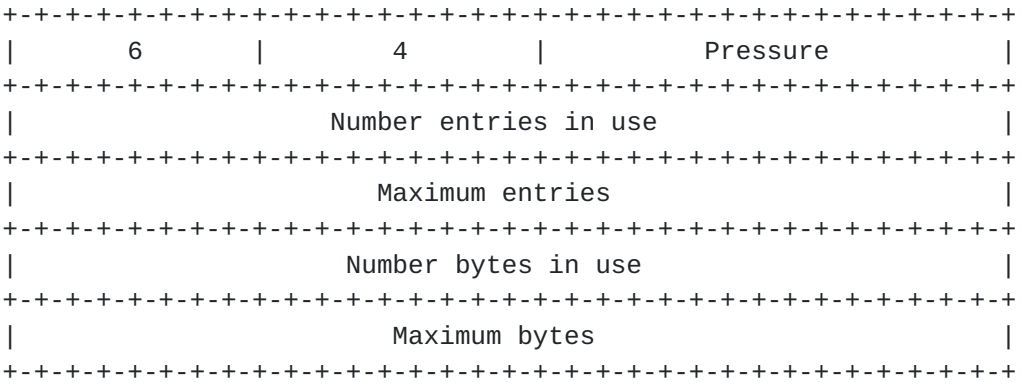
- 0: Null value, 0 bit value. This indicates that absence of locator or identifier information.
- 1: IPv6 address, 128 bit value
- 2: IPv4 address, 32 bit value
- 3: 32 bit index
- 4: 64 bit index
- 5: ILA value. A 64 bit value that represent a canonical ILA identifier when used in an IDType field and a canonical ILA locator when used in a LocType field.

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

5.8 Cache Occupancy message

This message provides the mapping cache size and occupancy of an AMS

forwarder. This serves as a hint that a router can use when pushing cache entries. The format of a Cache Occupancy message is:



Fields are:

- Type = 6. This indicates a Cache Occupancy message.
- Length = 4, indicating twenty bytes length
- Pressure: Indicates relative pressure the cache is under. The higher the number, the greater the pressure.
- Number entries in use: Approximate number of cache entries in the forwarder cache.
- Maximum entries: Approximate maximum number of cache entries in the forwarder cache. Zero indicates no reported information.
- Number bytes in use: Approximate number of bytes used in the forwarder cache.
- Maximum bytes: Approximate maximum number of bytes in the forwarder cache. Zero indicates no reported information.

If the cache size is not reported by a forwarder, then a router may assume local default values configured for the domain. Note that the protocol allows the forwarder to report cache occupancy and limits in several ways. Routers MAY use this information to modify the rate of pushing mapping entries or sending redirects.

This message is only sent by AMS forwarders. If an AMS forwarder receives a Cache Occupancy message then it is considered an error.

5.9 Operation

This section describes the operation of AMFP.

5.9.1 Populating an mapping cache

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

- 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.

5.9.2 Redirects

An AMS router can send redirects in conjunction with forwarding packets. Redirects are Mapping Information or Compressed Mapping Information messages 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 the address of the AMS forwarder. Note that this technique assumes a symmetric path towards the source.

5.9.2.1 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 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.

5.9.2.2 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 to a forwarder per second per redirected identifier. If a mapping change is detected the the rate limiting SHOULD be reset so that redirects for a new mapping can be sent immediately.

5.9.3 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 or Compressed Map Information message. If the router does not have a mapping entry for the requested identifier, it MAY reply with 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 require 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.

5.9.4 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 an AMS 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 a 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.

5.9.5 Cache maintenance

This section describes maintenance of a mapping cache.

5.9.5.1 Timeouts

A node SHOULD apply the timeout for a mapping entry that was indicated in a Map Information message or as negotiated default. 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.

5.9.5.2 Cache refresh

In order to avoid cycling a mapping entry with a redirect after 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.

5.9.6 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 the 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 using an overlay method. 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).

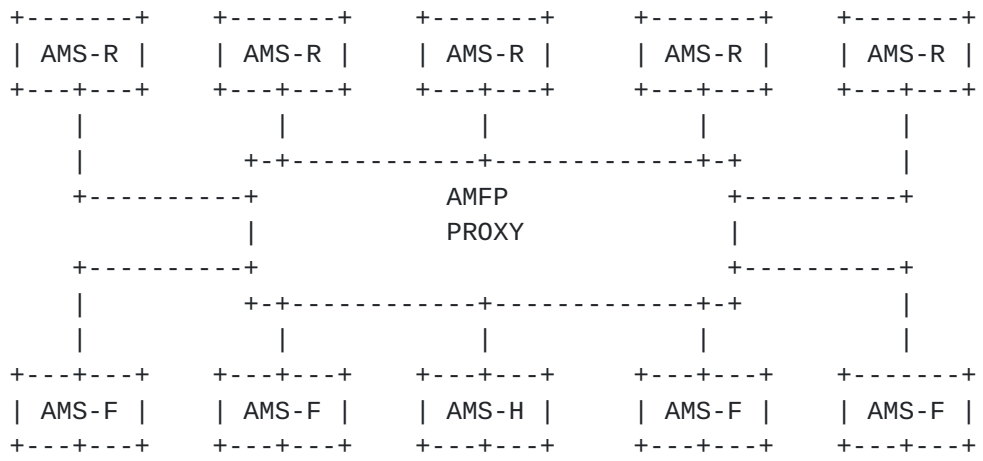
5.9.7 Locator unreachable handling

When connectivity to a locator is loss, the address 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.

5.9.8 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 AMFP.

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 in AMS forwarders is problematic, AMS router proxies could be used to 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.

5.9.9 Protocol errors

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

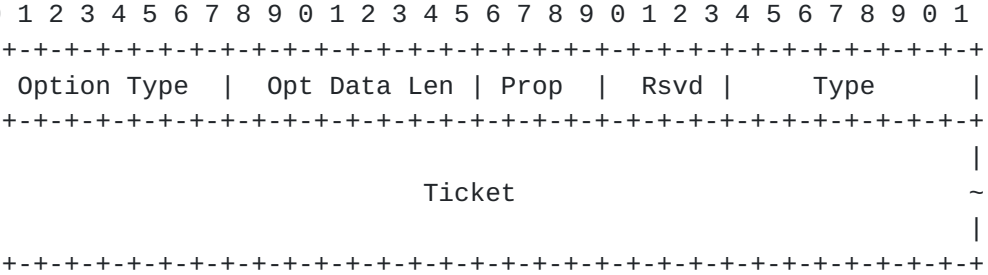
Protocol errors include mismatch of length for the message type or a Parameters TLV, unknown message type or Parameters TLV type, reserved bits not set to zero, unknown identifier type or locator type, unknown reason, unknown overlay method or instructions, loss of message synchronization in a TCP stream, or a message or parameters TLV was received that is inappropriate for the AMFP role. Note that if the end of a message does not end on field or record or message boundary this also considered a protocol error.

6 Stateless mapping optimization using FAST

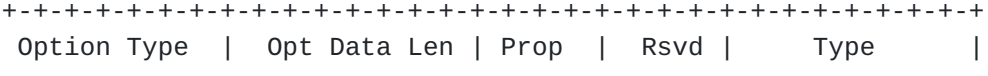
An alternative to requiring a mapping lookup on each packet is to encode the mapping information in packets themselves. This can be achieved by encoding mapping information in Firewall and Service Tickets [FAST]. The basic concept is that mapping information is encoded in FAST tickets which are attached in packets at 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.

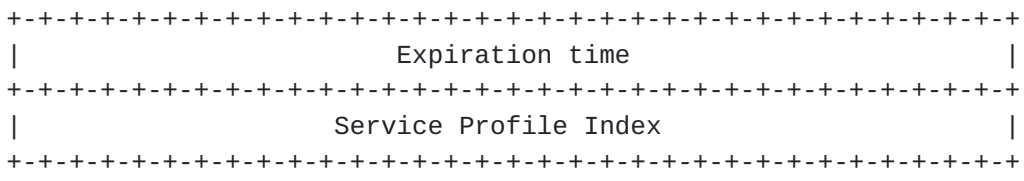
6.1 Firewall and Service Tickets encoding

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:

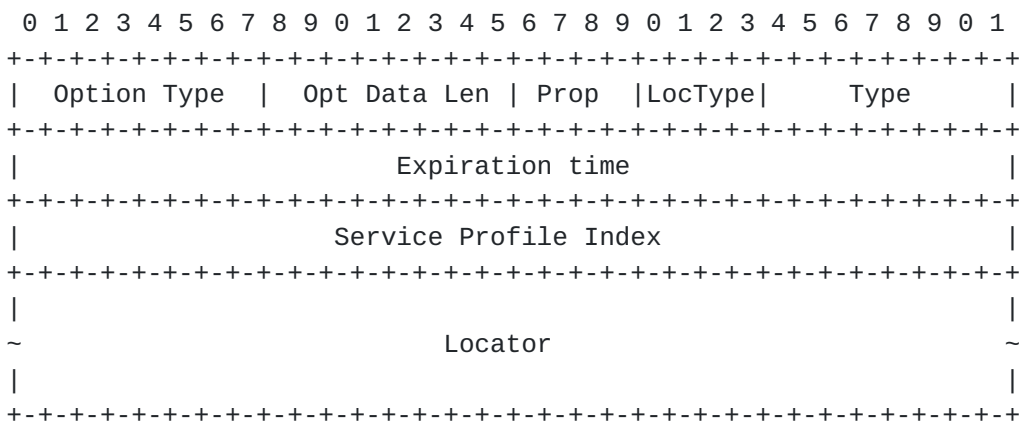




This format can be amended to include address mapping encoding.

6.2 Address mapping encoding

A locator address can be directly encoded 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: 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.

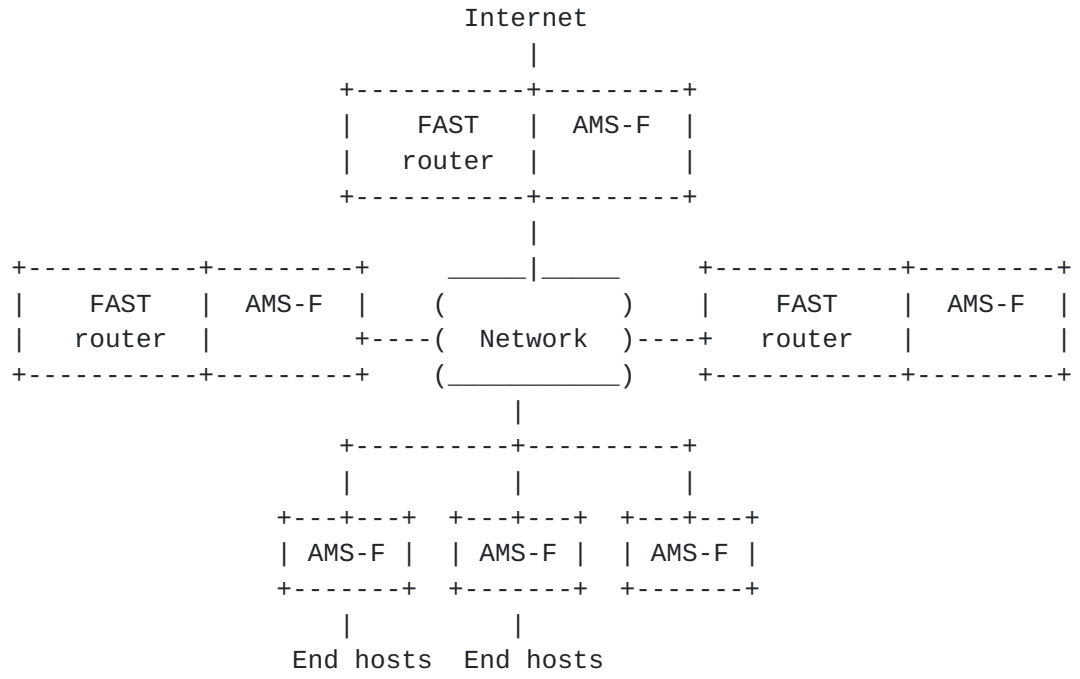
- Locator: A locator of type indicated by LocType. The size of a locator is specified by the type.

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 thousand of these. In this case, the border routers might maintain a 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.

6.3 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.



6.4 Operation

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

6.4.1 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.

6.4.2 Qualified locators

There are two possibilities for locator information in a ticket:

- The locator is fully qualified.
- The locator is not qualified.

6.4.2.1 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.

6.4.2.2 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 type is an 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 device 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.

6.4.3 AMS forwarder processing and FAST

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.

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.

6.4.4 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.

6.4.5 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.

6.4.6 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.

6.4.7 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.

6.4.8 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.

6.4.9 Expired tickets

FAST typically expects ticket to have an expiration time. If a ticket is received before the expiration time and 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 packets.

For tickets that are expired and contain mapping information, a FAST

node 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.

7 Privacy in Internet addresses

This section discusses the interaction between the address mapping system and privacy in Internet addressing. The address mapping system can facilitate strong privacy in Internet addressing. [ADDRPRIV] discusses privacy in addressing.

7.1 Criteria for privacy in addressing

Per [ADDRPRIV], 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.

7.2 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.

7.3 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 Carrier Grade NAT [RFC2663]-- 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.

7.3.1 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 address generation function may have the form:

Address = Func(key, group_ident, gen)

Where "key" is secret to network, "group_ident" is a group identifier for an aggregated set of addresses (for instance, the set of addresses for a device), 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.

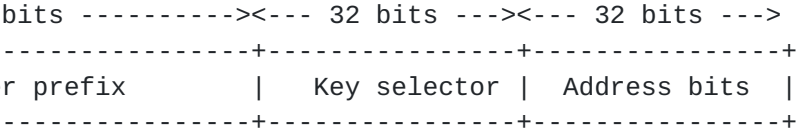
The inverse function has the form:

(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.

7.3.2 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).

7.3.3 Practicality of hidden aggregation methods

The premise of hidden aggregation is that only trusted devices in the network are able to 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. 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 a group identifier 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.

7.4 Scaling bulk address assignment

Assigning multiple addresses without aggregation is difficult to scale. Conceptually, 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]:

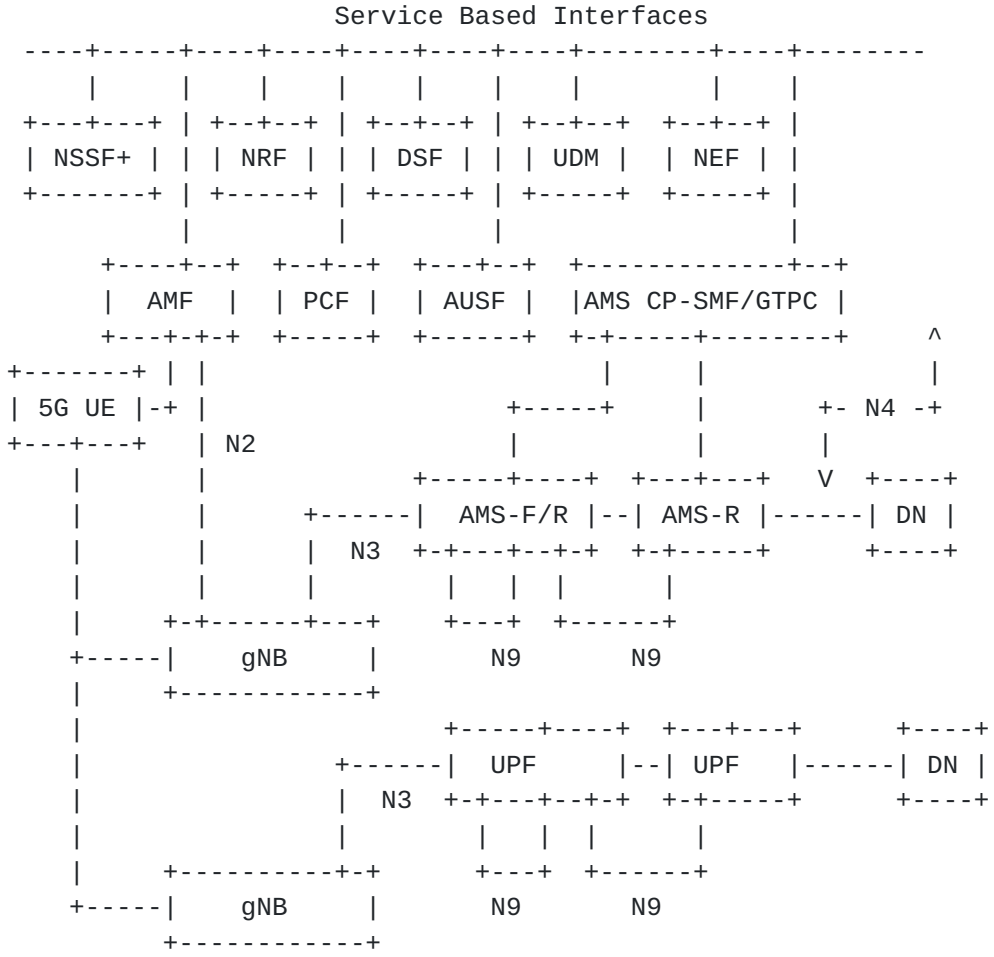
> Most DHCPv6 clients only ask for one non-temporary address, but the protocol allows requesting multiple temporary and even multiple non- temporary addresses, and the server could choose to provide multiple addresses. ... The maximum number of IPv6 addresses that can be provided in a single DHCPv6 packet, given a typical MTU of 1500 bytes or smaller, is approximately 30.

8 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 [3GPP15].

8.1 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.



AMS is used over the N3 and N9 interface. Address mappings in the

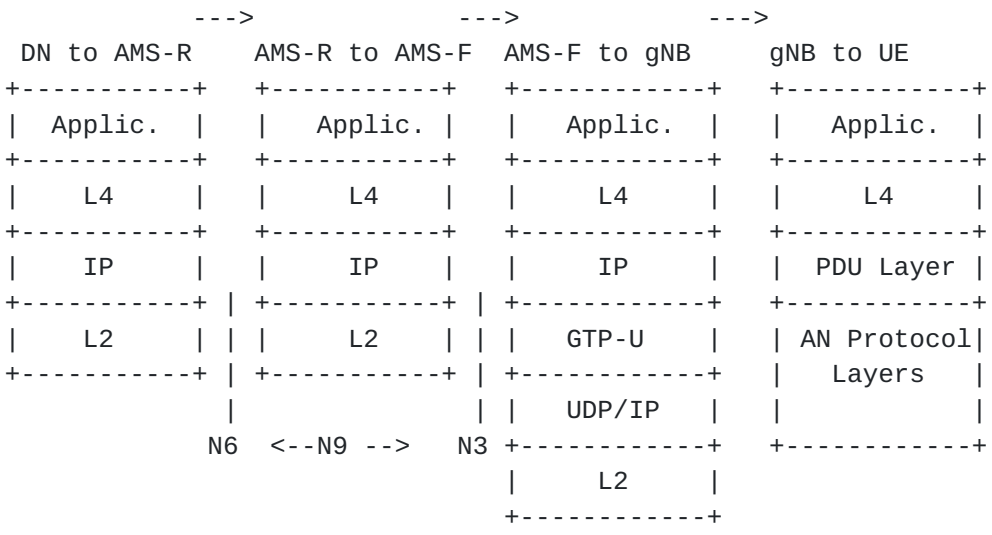
downlink from the data network are done by an AMS-R. Transformations for edge traffic can be done by an AMS-F close to the gNB or by an AMS-R in the case of a cache miss.

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 8.3). 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.

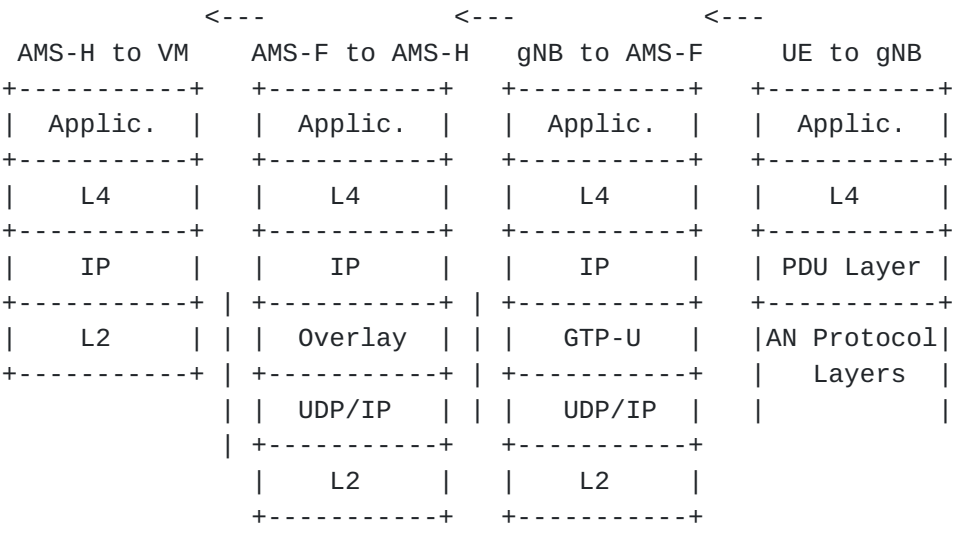
An AMS-CP communicates with other AMS-CPs, AMS-Fs, and AMS-Rs in the same address system mapping domain via control protocols that are independent of the 5G control plane. The address mapping database is shared amongst AMS-CP and AMS-Rs utilizing underlying distributed database technology deployed.

8.2 Protocol layering

The diagram below illustrates the protocol layers of 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 above where GTP-U is used over N3 and IP routing is used on N9.



AMS and protocol layers in the Downlink core are depicted below.



8.3 Control plane between AMS and the network

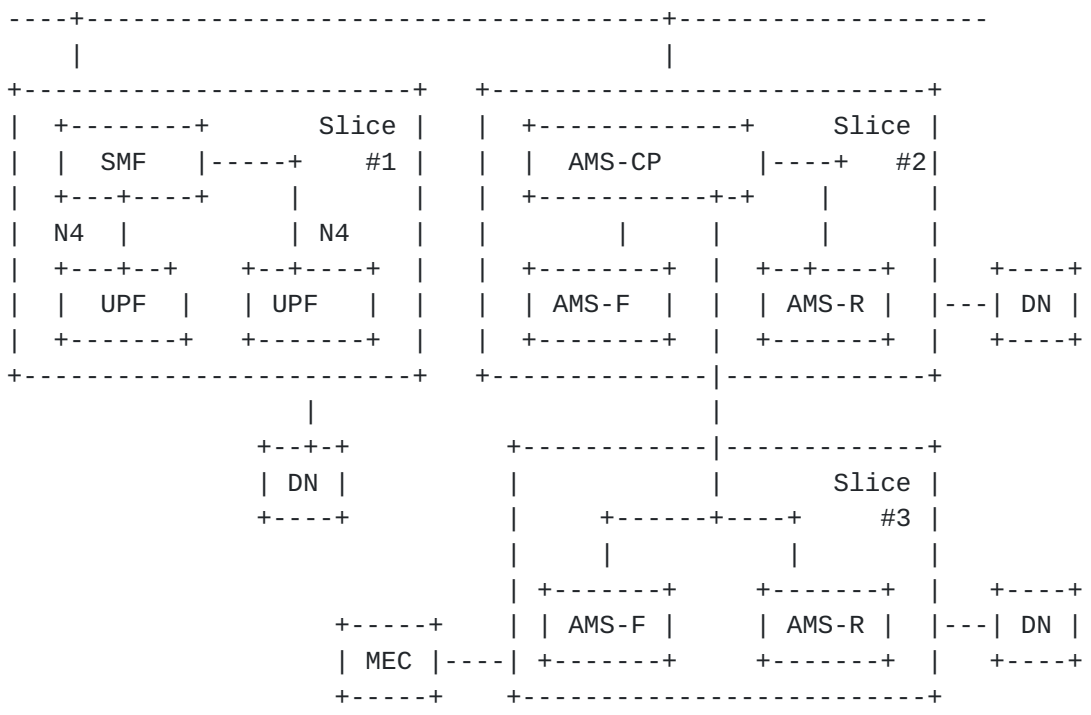
AMS is a consumer of several 5G network services. The service operations of interest to AMS are:

- 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.

8.4 AMS and network slices

The figure below illustrates the use of network slices with AMS.



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 of AMS-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 address mapping domain, is shared between slice #2 and slice #3. Alternatively, the two slices could each have their own AMS-CP and define separate address mapping domains.

8.5 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-C is the AMS routing protocol in 3GPP. 3GPP is based on an anchored routing model,

however the protocol can be augmented with AMS forwarders to achieve anchorless routing bypass. 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-U, require no change. The addition of AMS forwarders and mapping caches is done as an optimization for handling critical, low latency applications.

8.6 Overlay forwarding methods in 5G networks

As described in section 2.4, 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.

9 Security Considerations

AMFP must have protection against message forgery. In particular secure redirects and mapping information message are required to prevent attacks by 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.

10 IANA Considerations

TBD

11 Acknowledgments

The authors would like to thank Dirk von Hugo for contributions to this document.

12 References

12.1 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.

12.2 Informative References

[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>.

[RFC6830] Farinacci, D., Fuller, V., Meyer, D., and D. Lewis, "The Locator/ID Separation Protocol (LISP)", RFC 6830, DOI 10.17487/RFC6830, January 2013, <https://www.rfc-editor.org/info/rfc6830>.

[RFC6740] RJ Atkinson 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>.

[RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 4941, DOI 10.17487/RFC4941, September 2007, <https://www.rfc-editor.org/info/rfc4941>.

[RFC6462] Cooper, A., "Report from the Internet Privacy Workshop", RFC 6462, DOI 10.17487/RFC6462, January 2012, <https://www.rfc-editor.org/info/rfc6462>.

[RFC7721] Cooper, A., Gont, F., and D. Thaler, "Security and Privacy Considerations for IPv6 Address Generation Mechanisms", RFC 7721, DOI 10.17487/RFC7721, March 2016, <https://www.rfc-editor.org/info/rfc7721>.

[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>.

[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>.

[RFC7934] Colitti, L., Cerf, V., Cheshire, S., and D. Schinazi, "Host Address Availability Recommendations", BCP 204, RFC 7934, DOI 10.17487/RFC7934, July 2016, <https://www.rfc-editor.org/info/rfc7934>.

[RFC5246]] Dierks, T. and E. Rescorl, "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>.

[GUE] Herbert, T., Yong, L., and Zia, O., "Generic UDP Encapsulation" draft-ietf-intarea-gue-06

[GENEVE] Gross, J., Ed., Ganga, I. Ed., and Sridhar, T., "Geneve: Generic Network Virtualization Encapsulation", draft-ietf-nvo3-geneve-08

[GTP] 3rd Generation Partnership Project (3GPP), "3GPP TS 29.060", <www.3gpp.org/dynareport/29060.htm>

[ILA] Herbert, T., and Lapukhov, P., Privacy issues in ID/locator separation systems <draft-nordmark-id-loc-privacy-00>

[NFV] ETSI Industry Specification Group (ISG) NFV, "ETSI GS NFV 003 V1.2.1: Network Functions Virtualisation (NFV); Terminology for Main Concepts in NFV," 2014. <http://www.etsi.org/deliver/etsi gs/NFV/001

099/003/01.02.01 60/gs NFV003v010201p.pdf>.

[ADDRPRIV] Herbert, T., "Privacy in IPv6 Network Prefix Assignment", draft-herbert-ipv6-prefix-address-privacy-00

[IDLOCPRIV] Nordmark, E., "Privacy issues in ID/locator separation systems", draft-nordmark-id-loc-privacy-00

[3GPP15] 3rd Generation Partnership Project (3GPP), "3GPP - Release 15", <http://www.3gpp.org/release-15>

[BGPOLAY] Templin, F., Saccone, G., Dawra, G., Lindem, A., Moreno, V., "A Simple BGP-based Mobile Routing System for the Aeronautical Telecommunications Network", draft-templin-atn-bgp-08.txt

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

[FAST] Herbert, T., "Firewall and Service Tickets", draft-herbert-fast-03

Authors' Addresses

Tom Herbert
Quantonium
Santa Clara, CA
USA

Email: tom@quantonium.net

Vikram Siwach

Email: vsiwach@gmail.com