Network Working Group
Request for Comments: 2776
Category: Standards Track

M. Handley
ACIRI
D. Thaler
Microsoft
R. Kermode
Motorola

February 2000

Multicast-Scope Zone Announcement Protocol (MZAP)

Status of this Memo

Copyright Notice

Abstract

This document defines a protocol, the Multicast-Scope Zone Announcement Protocol (MZAP), for discovering the multicast administrative scope zones that are relevant at a particular location. MZAP also provides mechanisms whereby common misconfigurations of administrative scope zones can be discovered.

Table of Contents

1. Introduction

The use of administratively-scoped IP multicast, as defined in RFC 2365 [1], allows packets to be addressed to a specific range of multicast addresses (e.g., 239.0.0.0 to 239.255.255.255 for IPv4) such that the packets will not cross configured administrative boundaries, and also allows such addresses to be locally assigned and hence are not required to be unique across administrative boundaries. This property of logical naming both allows for address reuse, as well as provides the capability for infrastructure services such as address allocation, session advertisement, and service location to use well-known addresses which are guaranteed to have local significance within every organization.

The range of administratively-scoped addresses can be subdivided by administrators so that multiple levels of administrative boundaries can be simultaneously supported. As a result, a "multicast scope" is defined as a particular range of addresses which has been given some topological meaning.

To support such usage, a router at an administrative boundary is configured with one or more per-interface filters, or "multicast scope boundaries". Having such a boundary on an interface means that it will not forward packets matching a configured range of multicast addresses in either direction on the interface.

A specific area of the network topology which is within a boundary for a given scope is known as a "multicast scope zone". Since the same ranges can be reused within disjoint areas of the network, there may be many "multicast scope zones" for any given multicast scope. A

scope zone may have zero or more textual names (in different languages) for the scope, for human convenience. For example, if the range 239.192/14 were assigned to span an entire corporate network, it might be given (internally) the name "BigCo Private Scope".

Administrative scope zones may be of any size, and a particular host may be within many administrative scope zones (for different scopes, i.e., for non-overlapping ranges of addresses) of various sizes, as long as scope zones that intersect topologically do not intersect in address range.

Applications and services are interested in various aspects of the scopes within which they reside:

- Applications which present users with a choice of which scope in which to operate (e.g., when creating a new session, whether it is to be confined to a corporate intranet, or whether it should go out over the public Internet) are interested in the textual names which have significance to users.

- Services which use "relative" multicast addresses (as defined in [1]) in every scope are interested in the range of addresses used by each scope, so that they can apply a constant offset and compute which address to use in each scope.

- Address allocators are interested in the address range, and whether they are allowed to allocate addresses within the entire range or not.

- Some applications and services may also be interested in the nesting relationships among scopes. For example, knowledge of the nesting relationships can be used to perform "expanding-scope" searches in a similar, but better behaved, manner to the well-known expanding ring search where the TTL of a query is steadily increased until a replier can be found. Studies have also shown that nested scopes can be useful in localizing multicast repair traffic [8].

Two barriers currently make administrative scoping difficult to deploy and use:

- Applications have no way to dynamically discover information on scopes that are relevant to them. This makes it difficult to use administrative scope zones, and hence reduces the incentive to deploy them.

- o Misconfiguration is easy. It is difficult to detect scope zones that have been configured so as to not be convex (the shortest path between two nodes within the zone passes outside the zone), or to leak (one or more boundary routers were not configured correctly), or to intersect in both area and address range.

These two barriers are addressed by this document. In particular, this document defines the Multicast Scope Zone Announcement Protocol (MZAP) which allows an entity to learn what scope zones it is within. Typically servers will cache the information learned from MZAP and can then provide this information to applications in a timely fashion upon request using other means, e.g., via MADCAP [9]. MZAP also provides diagnostic information to the boundary routers themselves that enables misconfigured scope zones to be detected.

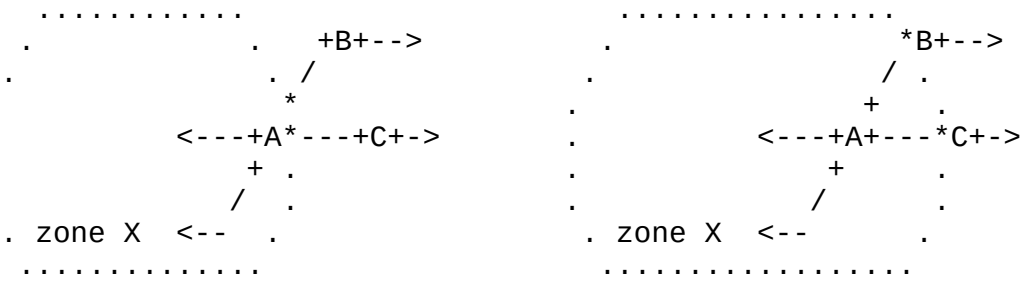
2. Terminology

The "Local Scope" is defined in RFC 2365 [1] and represents the smallest administrative scope larger than link-local, and the associated address range is defined as 239.255.0.0 to 239.255.255.255 inclusive (for IPv4, FF03::/16 for IPv6). RFC 2365 specifies:

> "239.255.0.0/16 is defined to be the IPv4 Local Scope. The Local Scope is the minimal enclosing scope, and hence is not further divisible. Although the exact extent of a Local Scope is site dependent, locally scoped regions must obey certain topological constraints. In particular, a Local Scope must not span any other scope boundary. Further, a Local Scope must be completely contained within or equal to any larger scope. In the event that scope regions overlap in area, the area of overlap must be in its own Local Scope. This implies that any scope boundary is also a boundary for the Local Scope."

A multicast scope Zone Boundary Router (ZBR) is a router that is configured with a boundary for a particular multicast scope on one or more of its interfaces. Any interface that is configured with a boundary for any administrative scope zone MUST also have a boundary for the Local Scope zone, as described above.

Such routers SHOULD be configured so that the router itself is within the scope zone. This is shown in Figure 1(a), where router A is inside the scope zone and has the boundary configuration.



A,B,C - routers * - boundary interface + - interface

(a) Correct zone boundary (b) Incorrect zone boundary

Figure 1: Administrative scope zone boundary placement

It is possible for the first router outside the scope zone to be configured with the boundary, as illustrated in Figure 1(b) where routers B and C are outside the zone and have the boundary configuration, whereas A does not, but this is NOT RECOMMENDED. This rule does not apply for Local Scope boundaries, but applies for all other boundary routers.

We next define the term "Zone ID" to mean the lowest IP address used by any ZBR for a particular zone for sourcing MZAP messages into that scope zone. The combination of this IP address and the first multicast address in the scope range serve to uniquely identify the scope zone. Each ZBR listens for messages from other ZBRs for the same boundary, and can determine the Zone ID based on the source addresses seen. The Zone ID may change over time as ZBRs come up and down.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [2].

Constants used by this protocol are shown as [NAME-OF-CONSTANT], and summarized in section 7.

3. Overview

When a ZBR is configured correctly, it can deduce which side of the boundary is inside the scope zone and which side is outside it.

Such a ZBR then sends periodic Zone Announcement Messages (ZAMs) for each zone for which it is configured as a boundary into that scope zone, containing information on the scope zone's address range, Zone ID, and textual names. These messages are multicast to the well-

known address [MZAP-LOCAL-GROUP] in the Local Scope, and are relayed across Local Scope boundaries into all Local Scope zones within the scope zone referred to by the ZAM message, as shown in Figure 2.

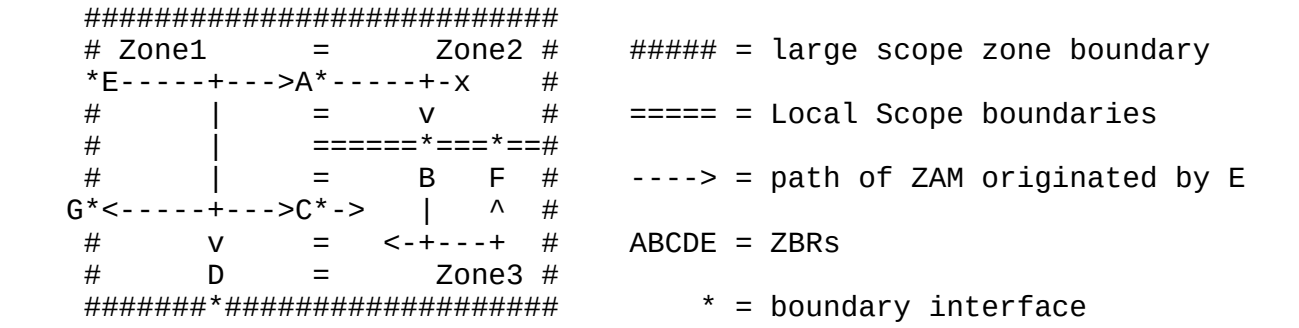
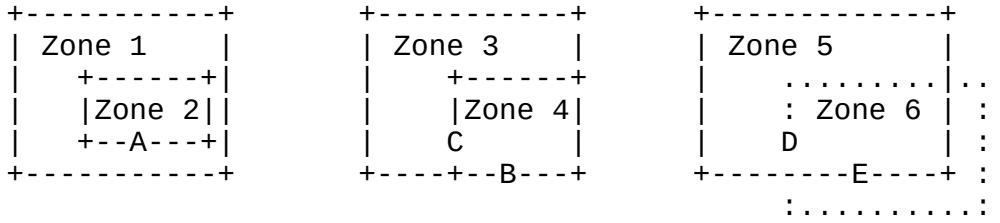


Figure 2: ZAM Flooding Example

Any entity can thus listen on a single well-known group address and learn about all scopes in which it resides.

3.1. Scope Nesting

MZAP also provides the ability to discover the nesting relationships between scope zones. Two zones are nested if one is comprised of a subset of the routers in the other, as shown in Figure 3.



(a) "Contained" Zone 2 nests inside Zone 1

(b) "Common Border" Zone 4 nests inside Zone 3

(c) "Overlap" Zones 5 and 6 do not nest

Figure 3: Zone nesting examples

A ZBR cannot independently determine whether one zone is nested inside another. However, it can determine that one zone does NOT nest inside another. For example, in Figure 3:

- ZBR A will pass ZAMs for zone 1 but will prevent ZAMs from zone 2 from leaving zone 2. When ZBR A first receives a ZAM for zone 1, it then knows that zone 1 does not nest within zone 2, but it cannot, however, determine whether zone 2 nests within zone 1.

- o ZBR B acts as ZBR for both zones 3 and 4, and hence cannot determine if one is nested inside the other. However, ZBR C can determine that zone 3 does not nest inside zone 4 when it receives a ZAM for zone 3, since it is a ZBR for zone 4 but not zone 3.

- o ZBR D only acts as ZBR zone 6 and not 5, hence ZBR D can deduce that zone 5 does not nest inside zone 6 upon hearing a ZAM for zone 5. Similarly, ZBR E only acts as ZBR zone 5 and not 6, hence ZBR E can deduce that zone 6 does not nest inside zone 5 upon hearing a ZAM for zone 6.

The fact that ZBRs can determine that one zone does not nest inside another, but not that a zone does nest inside another, means that nesting must be determined in a distributed fashion. This is done by sending Not-Inside Messages (NIMs) which express the fact that a zone X is not inside a zone Y. Such messages are sent to the well-known [MZAP-LOCAL-GROUP] and are thus seen by the same entities listening to ZAM messages (e.g., MADCAP servers). Such entities can then determine the nesting relationship between two scopes based on a sustained absence of any evidence to the contrary.

3.2. Other Messages

Two other message types, Zone Convexity Messages (ZCMs) and Zone Limit Exceeded (ZLE) messages, are used only by routers, and enable them to compare their configurations for consistency and detect misconfigurations. These messages are sent to MZAP's relative address within the scope range associated with the scope zone to which they refer, and hence are typically not seen by entities other than routers. Their use in detecting specific misconfiguration scenarios will be covered in the next section.

Packet formats for all messages are described in Section 5.

3.3. Zone IDs

When a boundary router first starts up, it uses its lowest IP address which it considers to be inside a given zone, and which is routable everywhere within the zone (for example, not a link-local address), as the Zone ID for that zone. It then schedules ZCM (and ZAM) messages to be sent in the future (it does not send them immediately). When a ZCM is received for the given scope, the sender is added to the local list of ZBRs (including itself) for that scope, and the Zone ID is updated to be the lowest IP address in the list. Entries in the list are eventually timed out if no further messages are received from that ZBR, such that the Zone ID will converge to the lowest address of any active ZBR for the scope.

Note that the sender of ZAM messages MUST NOT be used in this way. This is because the procedure for detecting a leaky Local scope described in Section 4.3 below relies on two disjoint zones for the same scope range having different Zone IDs. If ZAMs are used to compute Zone IDs, then ZAMs leaking across a Local Scope boundary will cause the two zones to converge to the same Zone ID.

4. Detecting Router Misconfigurations

In this section, we cover how to detect various error conditions. If any error is detected, the router should attempt to alert a network administrator to the nature of the misconfiguration. The means to do this lies outside the scope of MZAP.

4.1. Detecting non-convex scope zones

Zone Convexity Messages (ZCMs) are used by routers to detect non-convex administrative scope zones, which are one possible misconfiguration. Non-convex scope zones can cause problems for applications since a receiver may never see administratively-scoped packets from a sender within the same scope zone, since packets travelling between them may be dropped at the boundary.

In the example illustrated in Figure 4, the path between B and D goes outside the scope (through A and E). Here, Router B and Router C send ZCMs within a given scope zone for which they each have a boundary, with each reporting the other boundary routers of the zone from which they have heard. In Figure 4, Router D cannot see Router B's messages, but can see C's report of B, and so can conclude the zone is not convex.

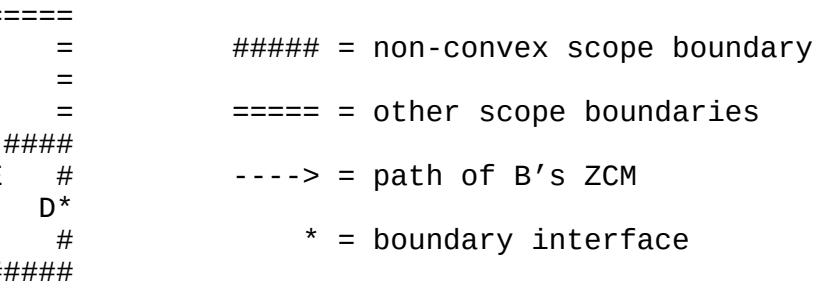


Figure 4: Non-convexity detection

Non-convex scope zones can be detected via three methods:

(1) If a ZBR is listed in ZCMs received, but the next-hop interface (according to the multicast RIB) towards that ZBR is outside the scope zone,

(2) If a ZBR is listed in ZCMs received, but no ZCM is received from that ZBR for [ZCM-HOLDTIME] seconds, as illustrated in Figure 4, or

(3) ZAM messages can also be used in a manner similar to that for ZCMs in (1) above, as follows: if a ZAM is received from a ZBR on an interface inside a given scope zone, and the next-hop interface (according to the multicast RIB) towards that ZBR is outside the scope zone.

Zone Convexity Messages MAY also be sent and received by correctly configured ordinary hosts within a scope region, which may be a useful diagnostic facility that does not require privileged access.

4.2. Detecting leaky boundaries for non-local scopes

A "leaky" boundary is one which logically has a "hole" due to some router not having a boundary applied on an interface where one ought to exist. Hence, the boundary does not completely surround a piece of the network, resulting in scoped data leaking outside.

Leaky scope boundaries can be detected via two methods:

(1) If it receives ZAMs originating inside the scope boundary on an interface that points outside the zone boundary. Such a ZAM message must have escaped the zone through a leak, and flooded back around behind the boundary. This is illustrated in Figure 5.

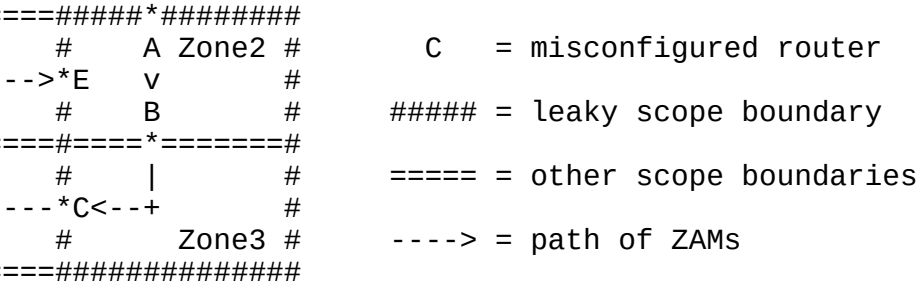


Figure 5: ZAM Leaking

(2) If a Zone Length Exceeded (ZLE) message is received. The ZAM packet also contains a Zones Traveled Limit. If the number of Local Scope zones traversed becomes equal to the Zones Traveled Limit, a ZLE message is generated (the suppression mechanism for preventing implosion is described later in the Processing Rules section). ZLEs detect leaks where packets do not return to another part of the same scope zone, but instead reach other Local Scope zones far away from the ZAM originator.

In either case, the misconfigured router will be either the message origin, or one of the routers in the ZBR path list which is included in the message received (or perhaps a router on the path between two such ZBRs which ought to have been a ZBR itself).

4.3. Detecting a leaky Local Scope zone

A local scope is leaky if a router has an administrative scope boundary on some interface, but does not have a Local Scope boundary on that interface as specified in RFC 2365. This can be detected via the following method:

- If a ZAM for a given scope is received by a ZBR which is a boundary for that scope, it compares the Origin's Scope Zone ID in the ZAM with its own Zone ID for the given scope. If the two do not match, this is evidence of a misconfiguration. Since a temporary mismatch may result immediately after a recent change in the reachability of the lowest-addressed ZBR, misconfiguration should be assumed only if the mismatch is persistent.

The exact location of the problem can be found by doing an mtrace [5] from the router detecting the problem, back to the ZAM origin, for any group within the address range identified by the ZAM. The router at fault will be the one reporting that a boundary was reached.

4.4. Detecting conflicting scope zones

Conflicting address ranges can be detected via the following method:

- If a ZBR receives a ZAM for a given scope, and the included start and end addresses overlap with, but are not identical to, the start and end addresses of a locally-configured scope.

Conflicting scope names can be detected via the following method:

- If a ZBR is configured with a textual name for a given scope and language, and it receives a ZAM or ZCM with a name for the same scope and language, but the scope names do not match.

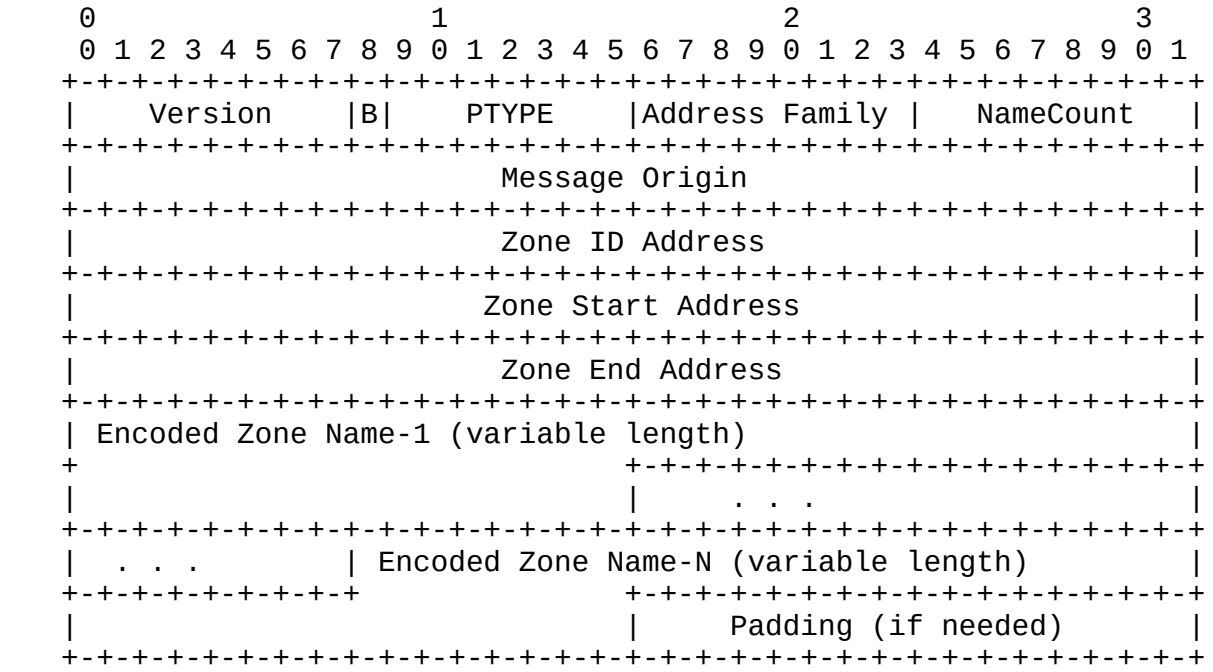
Detecting either type of conflict above indicates that either the local router or the router originating the message is misconfigured. Configuration tools SHOULD strip white space from the beginning and end of each name to avoid accidental misconfiguration.

5. Packet Formats

All MZAP messages are sent over UDP, with a destination port of [MZAP-PORT] and an IPv4 TTL or IPv6 Hop Limit of 255.

When sending an MZAP message referring to a given scope zone, a ZBR MUST use a source address which will have significance everywhere within the scope zone to which the message refers. For example, link-local addresses MUST NOT be used.

The common MZAP message header (which follows the UDP header), is shown below:



Version:
The version defined in this document is version 0.

"Big" scope bit (B):
If clear, indicates that the addresses in the scoped range are not subdividable, and that address allocators may utilize the entire range. If set, address allocators should not use the entire range, but should learn an appropriate sub-range via another mechanism (e.g., AAP [7]).

Packet Type (PTYPE):
The packet types defined in this document are:
0: Zone Announcement Message (ZAM)
1: Zone Limit Exceeded (ZLE)
2: Zone Convexity Message (ZCM)
3: Not-Inside Message (NIM)

Address Family:
The IANA-assigned address family number [10,11] identifying the address family for all addresses in the packet. The families defined for IP are:
1: IPv4
2: IPv6

Name Count:
The number of encoded zone name blocks in this packet. The count may be zero.

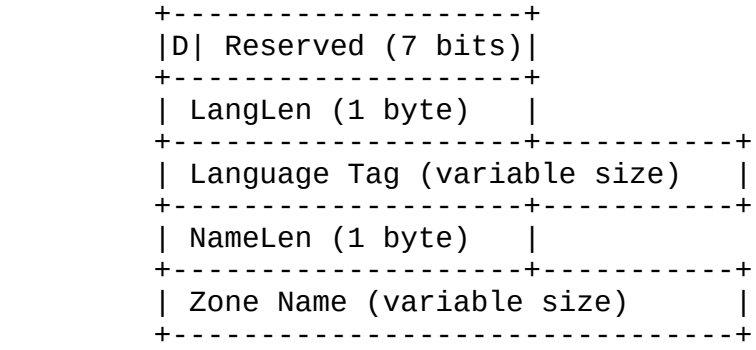
Zone Start Address: 32 bits (IPv4) or 128 bits (IPv6)
This gives the start address for the scope zone boundary. For example, if the zone is a boundary for 239.1.0.0 to 239.1.0.255, then Zone Start Address is 239.1.0.0.

Zone End Address: 32 bits (IPv4) or 128 bits (IPv6)
This gives the ending address for the scope zone boundary. For example, if the zone is a boundary for 239.1.0.0 to 239.1.0.255, then Zone End Address is 239.1.0.255.

Message Origin: 32 bits (IPv4) or 128 bits (IPv6)
This gives the IP address of the interface that originated the message.

Zone ID Address: 32 bits (IPv4) or 128 bits (IPv6)
This gives the lowest IP address of a boundary router that has been observed in the zone originating the message. Together with Zone Start Address and Zone End Address, it forms a unique ID for the zone. Note that this ID is usually different from the ID of the Local Scope zone in which the origin resides.

Encoded Zone Name:



The first byte contains flags, of which only the high bit is defined. The other bits are reserved (sent as 0, ignored on receipt).

"Default Language" (D) bit:
If set, indicates a preference that the name in the following language should be used if no name is available in a desired language.

Language tag length (LangLen): 1 byte
The length, in bytes, of the language tag.

Language Tag: (variable size)
The language tag, such as "en-US", indicating the language of the zone name. Language tags are described in [6].

Name Len:
The length, in bytes, of the Zone Name field. The length MUST NOT be zero.

Zone Name: multiple of 8 bits
The Zone Name is an ISO 10646 character string in UTF-8 encoding [4] indicating the name given to the scope zone (eg: "ISI-West Site"). It should be relatively short and MUST be less than 256 bytes in length. White space SHOULD be stripped from the beginning and end of each name before encoding, to avoid accidental conflicts.

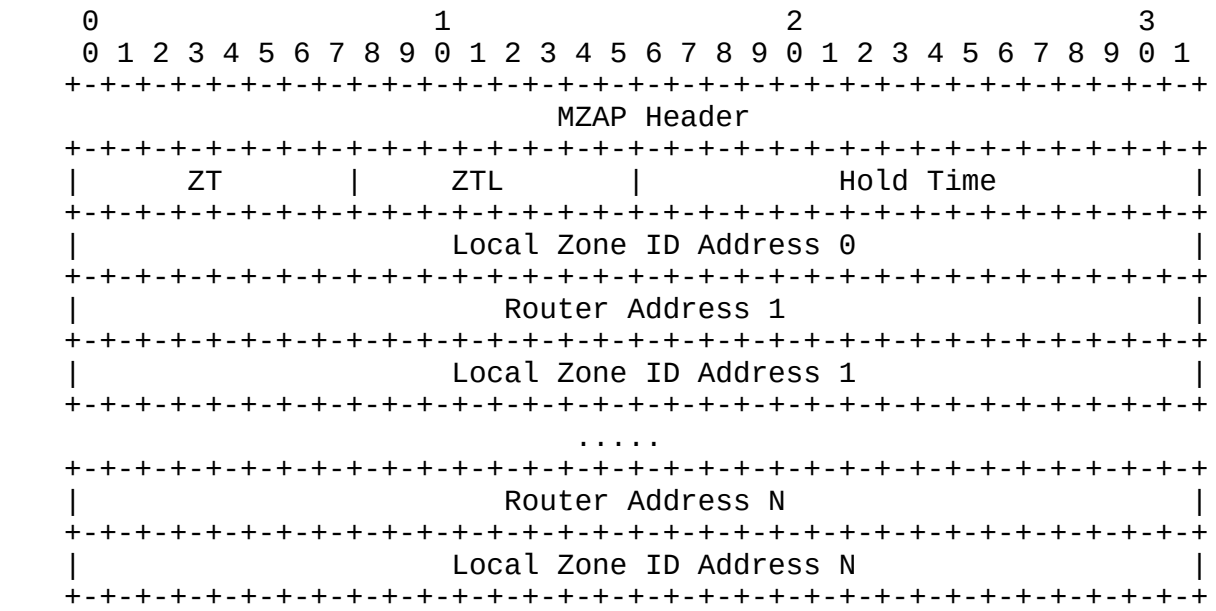
Padding (if needed):
The end of the MZAP header is padded with null bytes until it is 4-byte aligned.

5.1. Zone Announcement Message

A Zone Announcement Message has PTYPE=0, and is periodically sent by a ZBR for each scope for which it is a boundary, EXCEPT:

- the Local Scope
- the Link-local scope

The format of a Zone Announcement Message is shown below:



The fields are defined as follows:

Zones Traveled (ZT): 8 bits
: This gives the number of Local Zone IDs contained in this message path.

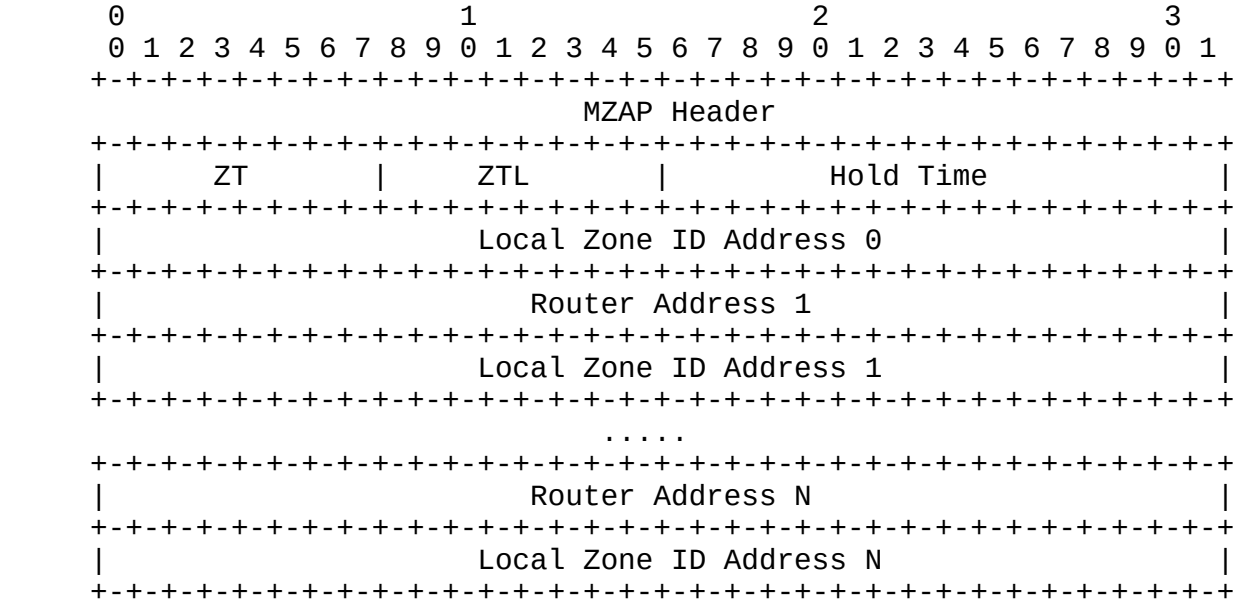
Zones Traveled Limit (ZTL): 8 bits
: This gives the limit on number of local zones that the packet can traverse before it MUST be dropped. A value of 0 indicates that no limit exists.

Hold Time:
: The time, in seconds, after which the receiver should assume the scope no longer exists, if no subsequent ZAM is received. This should be set to [ZAM-HOLDTIME].

Zone Path: multiple of 64 bits (IPv4) or 256 bits (IPv6)
The zone path is a list of Local Zone ID Addresses (the Zone ID Address of a local zone) through which the ZAM has passed, and IP addresses of the router that forwarded the packet. The origin router fills in the "Local Zone ID Address 0" field when sending the ZAM. Every Local Scope router that forwards the ZAM across a Local Scope boundary MUST add the Local Zone ID Address of the local zone that the packet of the zone into which the message is being forwarded, and its own IP address to the end of this list, and increment ZT accordingly. The zone path is empty which the ZAM is first sent.

5.2. Zone Limit Exceeded (ZLE)

The format of a ZLE is shown below:

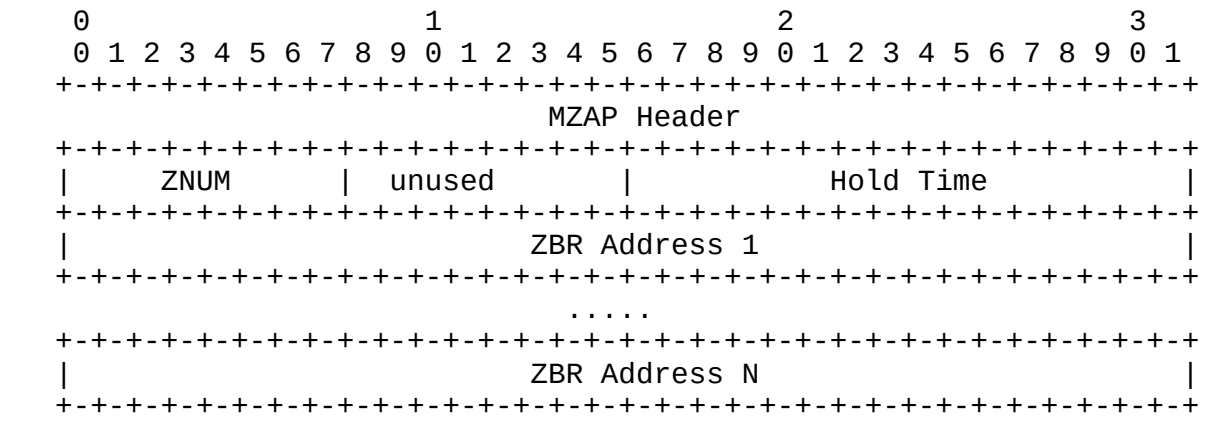


All fields are copied from the ZAM, except PTYPE which is set to one.

5.3. Zone Convexity Message

A Zone Announcement Message has PTYPE=2, and is periodically sent by a ZBR for each scope for which it is a boundary (except the Link-local scope). Note that ZCM’s ARE sent in the Local Scope.

Unlike Zone Announcement Messages which are sent to the [MZAP-LOCAL-GROUP], Zone Convexity Messages are sent to the [ZCM-RELATIVE-GROUP] in the scope zone itself. The format of a ZCM is shown below:



The fields are as follows:

Number of ZBR addresses (ZNUM): 8 bits
This field gives the number of ZBR Addresses contained in this message.

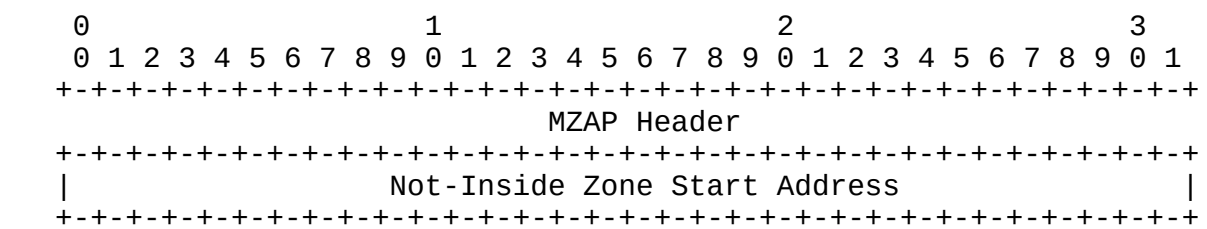
Hold Time:
The time, in seconds, after which the receiver should assume the sender is no longer reachable, if no subsequent ZCM is received. This should be set to [ZCM-HOLDTIME].

ZBR Address: 32 bits (IPv4) or 128 bits (IPv6)
These fields give the addresses of the other ZBRs from which the Message Origin ZBR has received ZCMs but whose hold time has not expired. The router should include all such addresses which fit in the packet, preferring those which it has not included recently if all do not fit.

5.4. Not-Inside Message

A Not-Inside Message (NIM) has PTYPE=3, and is periodically sent by a ZBR which knows that a scope X does not nest within another scope Y ("X not inside Y"):

The format of a Not-Inside Message is shown below:



The fields are as follows:

MZAP Header: Header fields identifying the scope X. The Name Count may be 0.

Not-Inside Zone Start Address: 32 bits (IPv4) or 128 bits (IPv6) This gives the start address for the scope Y.

6. Message Processing Rules

6.1. Internal entities listening to MZAP messages

Any host or application may join the [MZAP-LOCAL-GROUP] to listen for Zone Announcement Messages to build up a list of the scope zones that are relevant locally, and for Not-Inside Messages if it wishes to learn nesting information. However, listening to such messages is not the recommended method for regular applications to discover this information. These applications will normally query a local Multicast Address Allocation Server (MAAS) [3], which in turn listens to Zone Announcement Messages and Not-Inside Messages to maintain scope information, and can be queried by clients via MADCAP messages.

An entity (including a MAAS) lacking any such information can only assume that it is within the Global Scope, and the Local Scope, both of which have well-known address ranges defined in [1].

An internal entity (e.g., an MAAS) receiving a ZAM will parse the information that is relevant to it, such as the address range, and the names. An address allocator receiving such information MUST also use the "B" bit to determine whether it can add the address range to the set of ranges from which it may allocate addresses (specifically, it may add them only if the bit is zero). Even if the bit is zero, an MAAS SHOULD still store the range information so that clients who use relative- addresses can still obtain the ranges by requesting them from the MAAS.

An internal entity (e.g., an MAAS) should assume that X nests within Y if:

a) it first heard ZAMs for both X and Y at least [NIM-HOLDTIME] seconds ago, AND

b) it has not heard a NIM indicating that "X not inside Y" for at least [NIM-HOLDTIME] seconds.

6.2. Sending ZAMs

Each ZBR should send a Zone Announcement Message for each scope zone for which it is a boundary every [ZAM-INTERVAL] seconds, +/- 30% of [ZAM-INTERVAL] each time to avoid message synchronisation.

The ZAM packet also contains a Zones Traveled Limit (ZTL). If the number of Local Zone IDs in the ZAM path becomes equal to the Zones Traveled Limit, the packet will be dropped. The ZTL field is set when the packet is first sent, and defaults to 32, but can be set to a lower value if a network administrator knows the expected size of the zone.

6.3. Receiving ZAMs

When a ZBR receives a ZAM for some scope zone X, it uses the following rules.

If the local ZBR does NOT have any configuration for scope X:

(1) Check to see if the included start and end addresses overlap with, but are not identical to, the start and end addresses of any locally-configured scope Y, and if so, signal an address range conflict to a local administrator.

(2) Create a local "X not inside" state entry, if such an entry does not already exist. The ZBR then restarts the entry's timer at [ZAM-HOLDTIME]. Existence of this state indicates that the ZBR knows that X does not nest inside any scope for which it is a boundary. If the entry's timer expires (because no more ZAMs for X are heard for [ZAM-HOLDTIME]), the entry is deleted.

If the local ZBR does have configuration for scope X:

(1) If the ZAM originated from OUTSIDE the scope (i.e., received over a boundary interface for scope X):

   a) If the Scope Zone ID in the ZAM matches the ZBR's own Scope Zone ID, then signal a leaky scope misconfiguration.

   b) Drop the ZAM (perform no further processing below). For example, router G in Figure 2 will not forward the ZAM. This rule is primarily a safety measure, since the placement of G in Figure 2 is not a recommended configuration, as discussed earlier.

2) If the ZAM originated from INSIDE the scope:

   a) If the next-hop interface (according to the multicast RIB) towards the Origin is outside the scope zone, then signal a non- convexity problem.

   b) If the Origin's Scope Zone ID in the ZAM does not match the Scope Zone ID kept by the local ZBR, and this mismatch continues to occur, then signal a possible leaky scope warning.

   c) For each textual name in the ZAM, see if a name for the same scope and language is locally-configured; if so, but the scope names do not match, signal a scope name conflict to a local administrator.

   d) If the ZAM was received on an interface which is NOT a Local Scope boundary, and the last Local Zone ID Address in the path list is 0, the ZBR fills in the Local Zone ID Address of the local zone from which the ZAM was received.

If a ZAM for the same scope (as identified by the origin Zone ID and first multicast address) was received in the last [ZAM-DUP-TIME] seconds, the ZAM is then discarded. Otherwise, the ZAM is cached for at least [ZAM-DUP-TIME] seconds. For example, when router C in Figure 2 receives the ZAM via B, it will not be forwarded, since it has just forwarded the ZAM from E.

The Zones Travelled count in the message is then incremented, and if the updated count is equal to or greater than the ZTL field, schedule a ZLE to be sent as described in the next subsection and perform no further processing below.

If the Zone ID of the Local Scope zone in which the ZBR resides is not already in the ZAM's path list, then the ZAM is immediately re-originated within the Local Scope zone. It adds its own address and the Zone ID of the Local Scope zone into which the message is being forwarded to the ZAM path list before doing so. A ZBR receiving a ZAM with a non-null path list MUST NOT forward that ZAM back into a Local Scope zone that is contained in the path list. For example, in Figure 2, router F, which did not get the ZAM via A due to packet loss, will not forward the ZAM from B back into Zone 2 since the path list has { (E,1), (A,2), (B,3) } and hence Zone 2 already appears.

In addition, the ZBR re-originates the ZAM out each interface with a Local Scope boundary (except that it is not sent back out the interface over which it was received, nor is it sent into any local scope zone whose ID is known and appears in the path list). In each such ZAM re-originated, the ZBR adds its own IP address to the path

list, as well as the Zone ID Address of the Local Scope Zone into which the ZAM is being sent, or 0 if the ID is unknown. (For example, if the other end of a point-to-point link also has a boundary on the interface, then the link has no Local Scope Zone ID.)

6.4. Sending ZLEs

This packet is sent by a local-zone boundary router that would have exceeded the Zone Traveled Limit if it had forwarded a ZAM packet. To avoid ZLE implosion, ZLEs are multicast with a random delay and suppressed by other ZLEs. It is only scheduled if at least [ZLE-MIN-INTERVAL] seconds have elapsed since it previously sent a ZLE to any destination. To schedule a ZLE, the router sets a random delay timer within the interval [ZLE-SUPPRESSION-INTERVAL], and listens to the [MZAP-RELATIVE-GROUP] within the included scope for other ZLEs. If any are received before the random delay timer expires, the timer is cleared and the ZLE is not sent. If the timer expires, the router sends a ZLE to the [MZAP-RELATIVE-GROUP] within the indicated scope.

The method used to choose a random delay (T) is as follows:

```
Choose a random value X from the uniform random interval [0:1]
Let C = 256
Set T = [ZLE-SUPPRESSION-INTERVAL] log( C*X + 1) / log(C)
```

This equation results in an exponential random distribution which ensures that close to one ZBR will respond. Using a purely uniform distribution would begin to exhibit scaling problems as the number of ZBRs rose. Since ZLEs are only suppressed if a duplicate ZLE arrives before the time chosen, two routers choosing delays which differ by an amount less than the propagation delay between them will both send messages, consuming excess bandwidth. Hence it is desirable to minimize the number of routers choosing a delay close to the lowest delay chosen, and an exponential distribution is suitable for this purpose.

A router SHOULD NOT send more than one Zone Limit Exceeded message every [ZLE-MIN-INTERVAL] regardless of destination.

6.5. Receiving ZLEs

When a router receives a ZLE, it performs the following actions:

(1) If the router has a duplicate ZLE message scheduled to be sent, it unschedules its own message so another one will not be sent.

(2) If the ZLE contains the router’s own address in the Origin field, it signals a leaky scope misconfiguration.

6.6. Sending ZCMs

Each ZBR should send a Zone Convexity Message (ZCM) for each scope zone for which it is a boundary every [ZCM-INTERVAL] seconds, +/- 30% of [ZCM-INTERVAL] each time to avoid message synchronisation.

ZCMs are sent to the [ZCM-RELATIVE-GROUP] in the scoped range itself. (For example, if the scope range is 239.1.0.0 to 239.1.0.255, then these messages should be sent to 239.1.0.252.) As these are not Locally-Scoped packets, they are simply multicast across the scope zone itself, and require no path to be built up, nor any special processing by intermediate Local Scope ZBRs.

6.7. Receiving ZCMs

When a ZCM is received for a given scope X, on an interface which is inside the scope, it follows the rules below:

(1) The Origin is added to the local list of ZBRs (including itself) for that scope, and the Zone ID is updated to be the lowest IP address in the list. The new entry is scheduled to be timed out after [ZCM-HOLDTIME] if no further messages are received from that ZBR, so that the Zone ID will converge to the lowest address of any active ZBR for the scope.

(2) If a ZBR is listed in ZCMs received, but the next-hop interface (according to the multicast RIB) towards that ZBR is outside the scope zone, or if no ZCM is received from that ZBR for [ZCM-HOLDTIME] seconds, as in the example in Figure 4, then signal a non-convexity problem.

(3) For each textual name in the ZCM, see if a name for the same scope and language is locally-configured; if so, but the scope names do not match, signal a scope name conflict to a local administrator.

6.8. Sending NIMs

Periodically, for each scope zone Y for which it is a boundary, a router originates a Not-Inside Message (NIM) for each "X not inside" entry it has created when receiving ZAMs. Like a ZAM, this message is multicast to the address [MZAP-LOCAL-GROUP] from one of its interfaces inside Y.

Each ZBR should send such a Not-Inside Message every [NIM-INTERVAL] seconds, +/- 30% of [NIM-INTERVAL] to avoid message synchronization.

6.9. Receiving NIMs

When a ZBR receives a NIM saying that "X is not inside Y", it is forwarded, unmodified, in a manner similar to ZAMs:

(1) If the NIM was received on an interface with a boundary for either X or Y, the NIM is discarded.

(2) Unlike ZAMs, if the NIM was not received on the interface towards the message origin (according to the Multicast RIB), the NIM is discarded.

(3) If a NIM for the same X and Y (where each is identified by its first multicast address) was received in the last [ZAM-DUP-TIME] seconds, the NIM is not forwarded.

(4) Otherwise, the NIM is cached for at least [ZAM-DUP-TIME] seconds.

(5) The ZBR then re-originates the NIM (i.e., with the original UDP payload) into each local scope zone in which it has interfaces, except that it is not sent back into the local scope zone from which the message was received, nor is it sent out any interface with a boundary for either X or Y.

7. Constants

[MZAP-PORT]: The well-known UDP port to which all MZAP messages are sent. Value: 2106.

[MZAP-LOCAL-GROUP]: The well-known group in the Local Scope to which ZAMs are sent. All Multicast Address Allocation servers and Zone Boundary Routers listen to this group. Value: 239.255.255.252 for IPv4.

[ZCM-RELATIVE-GROUP]: The relative group in each scope zone, to which ZCMs are sent. A Zone Boundary Router listens to the relative group in each scope for which it is a boundary. Value: (last IP address in scope range) - 3. For example, in the Local Scope, the relative group is the same as the [MZAP-LOCAL-GROUP] address.

[ZAM-INTERVAL]: The interval at which a Zone Boundary Router originates Zone Announcement Messages. Default value: 600 seconds (10 minutes).

[ZAM-HOLDTIME]: The holdtime to include in a ZAM. This SHOULD be set to at least 3 * [ZAM-INTERVAL]. Default value: 1860 seconds (31 minutes).

[ZAM-DUP-TIME]: The time interval after forwarding a ZAM, during which ZAMs for the same scope will not be forwarded. Default value: 30 seconds.

[ZCM-INTERVAL]: The interval at which a Zone Boundary Router originates Zone Convexity Messages. Default value: 600 seconds (10 minutes).

[ZCM-HOLDTIME]: The holdtime to include in a ZCM. This SHOULD be set to at least 3 * [ZCM-INTERVAL]. Default value: 1860 seconds (31 minutes).

[ZLE-SUPPRESSION-INTERVAL]: The interval over which to choose a random delay before sending a ZLE message. Default value: 300 seconds (5 minutes).

[ZLE-MIN-INTERVAL]: The minimum interval between sending ZLE messages, regardless of destination. Default value: 300 seconds (5 minutes).

[NIM-INTERVAL]: The interval at which a Zone Boundary Router originates Not-Inside Messages. Default value: 1800 seconds (30 minutes).

[NIM-HOLDTIME]: The holdtime to include the state within a NIM. This SHOULD be set to at least 3 * [NIM-INTERVAL]. Default value: 5460 (91 minutes)

8. Security Considerations

While unauthorized reading of MZAP messages is relatively innocuous (so encryption is generally not an issue), accepting unauthenticated MZAP messages can be problematic. Authentication of MZAP messages can be provided by using the IPsec Authentication Header (AH) [12].

In the case of ZCMs and ZLEs, an attacker can cause false logging of convexity and leakage problems. It is likely that is would be purely an annoyance, and not cause any significant problem. (Such messages could be authenticated, but since they may be sent within large scopes, the receiver may not be able to authenticate a non-malicious sender.)

ZAMs and NIMs, on the other hand, are sent within the Local Scope, where assuming a security relationship between senders and receivers is more practical.

In the case of NIMs, accepting unauthenticated messages can cause the false cancellation of nesting relationships. This would cause a section of the hierarchy of zones to flatten. Such a flattening would lessen the efficiency benefits afforded by the hierarchy but would not cause it to become unusable.

Accepting unauthenticated ZAM messages, however, could cause applications to believe that a scope zone exists when it does not. If these were believed, then applications may choose to use this non-existent administrative scope for their uses. Such applications would be able to communicate successfully, but would be unaware that their traffic may be traveling further than they expected. As a result, any application accepting unauthenticated ZAMs MUST only take scope names as a guideline, and SHOULD assume that their traffic sent to non-local scope zones might travel anywhere. The confidentiality of such traffic CANNOT be assumed from the fact that it was sent to a scoped address that was discovered using MZAP.

In addition, ZAMs are used to inform Multicast Address Allocation Servers (MAASs) of names and address ranges of scopes, and accepting unauthenticated ZAMs could result in false names being presented to users, and in wrong addresses being allocated to users. To counter this, MAAS's authenticate ZAMs as follows:

(1) A ZBR signs all ZAMs it originates (using an AH).

(2) A ZBR signs a ZAM it relays if and only if it can authenticate the previous sender. A ZBR MUST still forward un-authenticated ZAMs (to provide leak detection), but should propagate an authenticated ZAM even if an un-authenticated one was received with the last [ZAM-DUP-TIME] seconds.

(3) A MAAS SHOULD be configured with the public key of the local zone in which it resides. A MAAS thus configured SHOULD ignore an unauthenticated ZAM if an authenticated one for the same scope has been received, and MAY ignore all unauthenticated ZAMs.

9. Acknowledgements

This document is a product of the MBone Deployment Working Group, whose members provided many helpful comments and suggestions, Van Jacobson provided some of the original ideas that led to this protocol. The Multicast Address Allocation Working Group also provided useful feedback regarding scope names and interactions with applications.

10. References

[1] Meyer, D., "Administratively Scoped IP Multicast", BCP 23, RFC 2365, July 1998.

[2] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[3] Thaler, D., Handley, M. and D. Estrin, "The Internet Multicast Address Allocation Architecture", Work in Progress.

[4] Yergeau, F., "UTF-8, a transformation format of ISO 10646", RFC 2279, January 1998.

[5] Fenner, W. and S. Casner, "A ‘traceroute’ facility for IP Multicast", Work in Progress.

[6] Alvestrand, H., "Tags for the Identification of Languages", RFC 1766, March 1995.

[7] Handley, M. and S. Hanna. "Multicast Address Allocation Protocol (AAP)", Work in Progress.

[8] Kermode, R. "Scoped Hybrid Automatic Repeat reQuest with Forward Error Correction (SHARQFEC)", ACM SIGCOMM 98, September 1998, Vancouver, Canada.

[9] Hanna, S., Patel, B., and M. Shah. "Multicast Address Dynamic Client Allocation Protocol (MADCAP)", RFC 2730, December 1999.

[10] Reynolds, J. and J. Postel, "Assigned Numbers", STD 2, RFC 1700, October 1994.

[11] IANA, "Address Family Numbers", http://www.isi.edu/in-notes/iana/assignments/address-family-numbers

[12] Kent, S. and R. Atkinson, "IP Authentication Header", RFC 2402, November 1998.

11. Authors' Addresses

Mark Handley
AT&T Center for Internet Research at ICSI
1947 Center St, Suite 600
Berkely, CA 94704
USA

EMail: mjh@aciri.org

David Thaler
Microsoft
One Microsoft Way
Redmond, WA 98052
USA

EMail: dthaler@microsoft.com

Roger Kermode
Motorola Australian Research Centre
12 Lord St,
Botany, NSW 2019
Australia

EMail: Roger.Kermode@motorola.com

12. Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.