Malloc Working Group
Internet Engineering Task Force
INTERNET-DRAFT
8 November 1998
Expires 8 May 1999

Roger Kermode
Motorola

Scoped Address Discovery Protocol (SADP)
<draft-ietf-mboned-sadp-00.txt>

Status of this Memo

Abstract

This document defines a protocol, the Scoped Address Discovery Protocol (SADP), for discovering the scoped multicast address(es) associated with a session at particular scopes within a hierarchically nested set of multicast zones. SADP is designed to work within the context of Multicast Address Allocation Architecture [MAAA] consisting of the MZAP [MZAP], MASC [MASC], and AAP [AAP] protocols. It is intended that SADP will provide the necessary general services for reliable multicast and searching applications to use expanding-zone searches in lieu of the well known, but less efficient expanding-ring search.

Copyright Notice

Contents

1. Introduction

Administrative scoping [RFC2365] provide a useful means for limiting the spread of IP multicast traffic acros the Internet. Unlike Time-To-Live (TTL) scoping, administrative scoping provides the means to ensure that, for a given scope and ignoring packet loss, the same set of nodes will receive a message, regardless of which node sent the message. Thus, the use of administrative scoping greatly simplifies the design of multicast protocols that require localization, since the non-reception of sent packets is solely due to loss and not design.

The Multicast Zone Announcement Protocol (MZAP) [MZAP] will provide applications with the means for discovering the various scopes that are locally visible at each point in the Internet. In addition, MZAP will also provide the means for determining and announcing which scope zones completely encapsulate others. This additional ability will allow scope zones to be arranged into hierarchies which applications can then used expanding zone searches instead of less efficient and more problematic expanding-ring (TTL) searches. One example of how expanding-zone searches provide increased localization can be found in the Scoped Hybrid Automatic Repear reQuest with Forward Error Correction (SHARQFEC) reliable multicast protocol [SHARQFEC].

While expanding-ring searches use one multicast address and increasing TTLs, expanding-zone searches involve changing the multicast addresses for each attempt at a different scope. SADP builds upon the Multicast Address Allocation Architecture [MAAA] by adding a new service that allows applications to discover the relevant multicast address(es) associated with a session at each level in a hierarchy of scope zones. SADP does not provide the means to allocate an address should one not be present for a session in a particular zone. In this case the application should use the Address Allocation Protocol (AAP) [AAP] to allocate a new address for the scope, which can then be announced to other application instances within the scope.

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

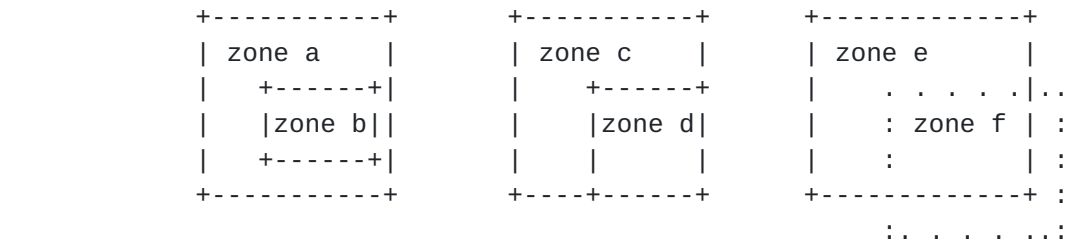
2. Overview

Administrative scoping affords the ability to create network partitions or zones in which multicast traffic addressed to one of a block of addresses assigned to that zone will be limited to that zone. The boundary of the zone is enforced by Zone Border Routers (ZBRs) that reside at the edges of the zone. ZBRs must be carefully configured so that traffic addressed within the zone does not pass outside the zone. This can be a non trivial task, and hence the Multicast Zone Announcement Protocol (MZAP) [MZAP], which is used to announce the existence of zones, also provides the mechanisms to detect ZBR misconfigurations.

A, B, C - Routers * - border interface + - interface . - border

Figure 1: Admin scope zone border example

Zones may be of different sizes and can also overlap. In addition to the services of zone announcement and fault detection, MZAP also provides mechanisms for determining and announcing the existence of zones that nest inside others as shown in Figure 2.



(a) "Contained" zone b nests inside zone a

(b) "Common Border" zone d nests inside zone c

(c) "Overlap" zones e and f do not nest

Figure 2: Zone nesting examples

This feature allows admin scope zones to be arranged in a hierarchy as shown in Figure 3. The ability to nest admin scope zones in hierarchies like that shown in Figure 3 is useful since it affords

localization through expanding-zone searches. For example, consider a distributed application with session members distributed evenly through out zone a. A session member in zone e, would perform a search by multicasting a query within zone e, and if unsuccessful, expand the scope to search in zone b, and eventually zone a if so needed.

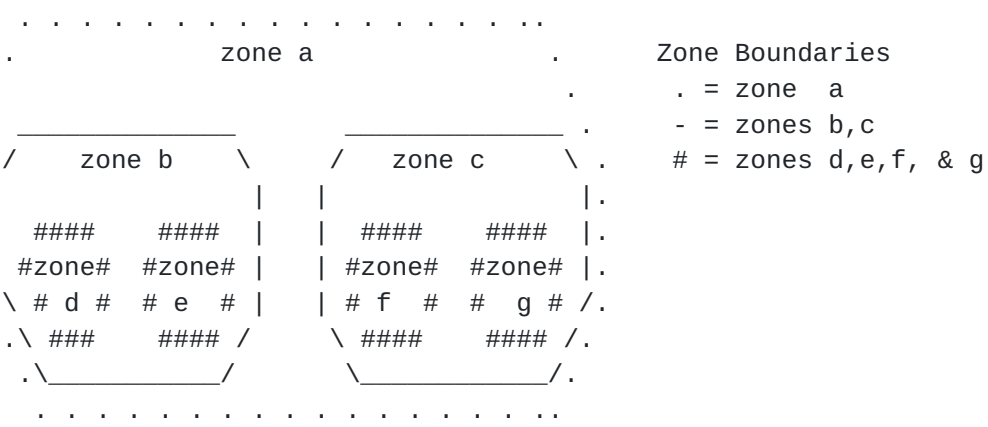


Figure 3 : Zone Nesting Hierarchy example

In order for expanding-zone searches to be feasible, session members must be able to determine two things:

o which zones are involved in the hierarchy for a particular session.

o what address(es) are to be used for communicating with other session members within the zones involved in the hierarchy.

SADP affords the ability to discover this information by using a single multicast group at each scope [SADP-RELATIVE-GROUP] for communication between SADP servers and the members of various sessions. New members to a session use the channels provided by the addresses to query existing SADP servers and session members as to which specific zones are valid and which zones to use. Since there is only one multicast address used per zone for this purpose, members of a particular session will ignore traffic intended for members of another session.

3. Usage

In this section we summarize how session members can use SADP to determine which admin zones are used by the session's hierarchy and also the address(es) within these zones that are used by the current session members should such addresses exist.

3.1 Session Identifiers

Each session that uses admin scoping will use a Globally Unique Session Identifier (GUSID) that will distinguish it from all other sessions. This GUSID will consists of a 128bit integer that is allocated dynamically using the process described in [UUID]. The GUSID will be allocated by the session creator and will be used to associate traffic with a particular session regardless of which multicast scoped address the traffic is sent to.

3.2 Session Member Operation

Several predefined administrative scopes already exist [RFC2365]:

o Link Local: Traffic is only carried across one physical link.

o Local: Traffic is restricted to a specific network region.

o Global: The entire multicast enabled network.

By definition Link Local zones nest inside Local zone which in turn nests inside the Global zone. Other zones may exist between the local and global scopes. These zones are constructed by the union of two or more local zones and are announced to routers within their confines using MZAP [MZAP].

The general algorithm that new members to a session should use to determine which zones and addresses are involved in the hierarchy for a particular session is as follows:

1) Determine the GUSID, largest zone, and addresses for the largest zone for the session. (this task is beyond the scope of this document, but can be assumed to involve some kind of out-of-band communication.)

2) Starting with the SADP group [SADP-RELATIVE-GROUP] for the local scope, issue a SADP Request (SADP_REQ) message containing the GUSID address.

3) Wait for a response on the SADP [SADP-RELATIVE-GROUP] address for at least [SADP-REQ-TIMEOUT] seconds. If no response is heard increase the scope to the next largest zone and repeat step 2. In cases where there are two non-nesting zones larger than the current try one zone and then the other, should the first zone not result in a reply.

4) Continue steps 2) and 3) until the largest zone has been queried or a response has been heard.

In cases where the scope must be increased in order to find a session member that can reply, the new session member MAY decide to add levels to the hierarchy in order to increase localization for future session members. New session members that decide to take this step will use the existing addresses as discovered using SADP and request new ones using AAP [AAP].

SADP servers and existing session members, upon hearing an SADP_REQ message from a new session member will issue an SADP Response (SADP_RESP) after waiting for a random amount of time (T) that is calculated as follows:

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

Should a member receive a SADP_RESP before its timer it expires it SHALL suppress its own response. This method ensures that close to one session member will respond.

3.2 SADP Server Operation

Were SADP to be deployed in a wide scale session with the members of various sessions to use SADP between each other it would quickly cause catastrophic congestion. The reason for this is that whenever a new node joined a sparsely populated session with a large maximum scope, it would inevitably end up sending SADP_REQs to every scope up until the largest scope. Thus the highly likely occurrence of having a global and continental scope zones combined with numerous sparse sessions (probably on the order of 10,000 to 100,000) would quickly cause SADP_REQ flooding at the continental scope.

To address this shortcoming SADP allows, and in fact encourages, the deployment of SADP servers. These servers subscribe to the [SADP-RELATIVE-GROUP] for each zone they are in and cache the SADP_RESP messages they receive at each scope. Having cached and merged the responses for sessions at various scopes, they can then respond to SADP_REQs heard at lower scopes using the information heard at the larger scope(s). Should a SADP server hear a SADP_REQ at some intermediate scope it MUST NOT announce address information for scopes smaller than one on which the SADP_REQ was received.

The effect of allowing larger-scoped information to be announced at lower scopes by SADP servers significantly reduces the number of scopes a new session will have to query. New session members now need only expand the scope until a SADP server is found. This is a marked improvement over the case where no SADP servers exist and the search must continue until an existing session member is found.

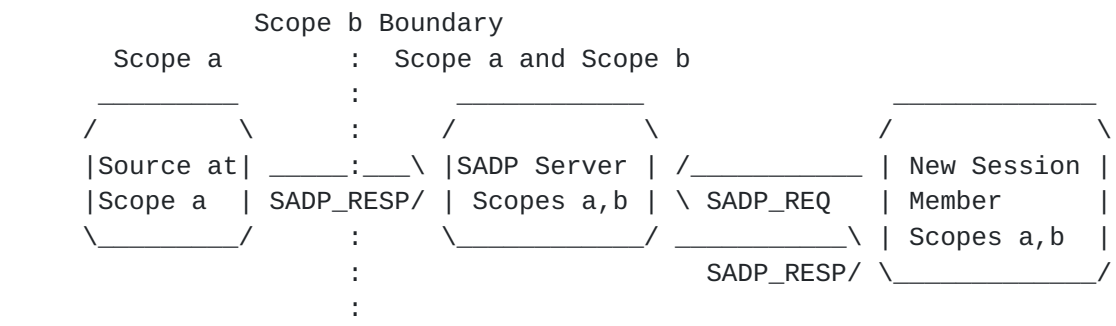
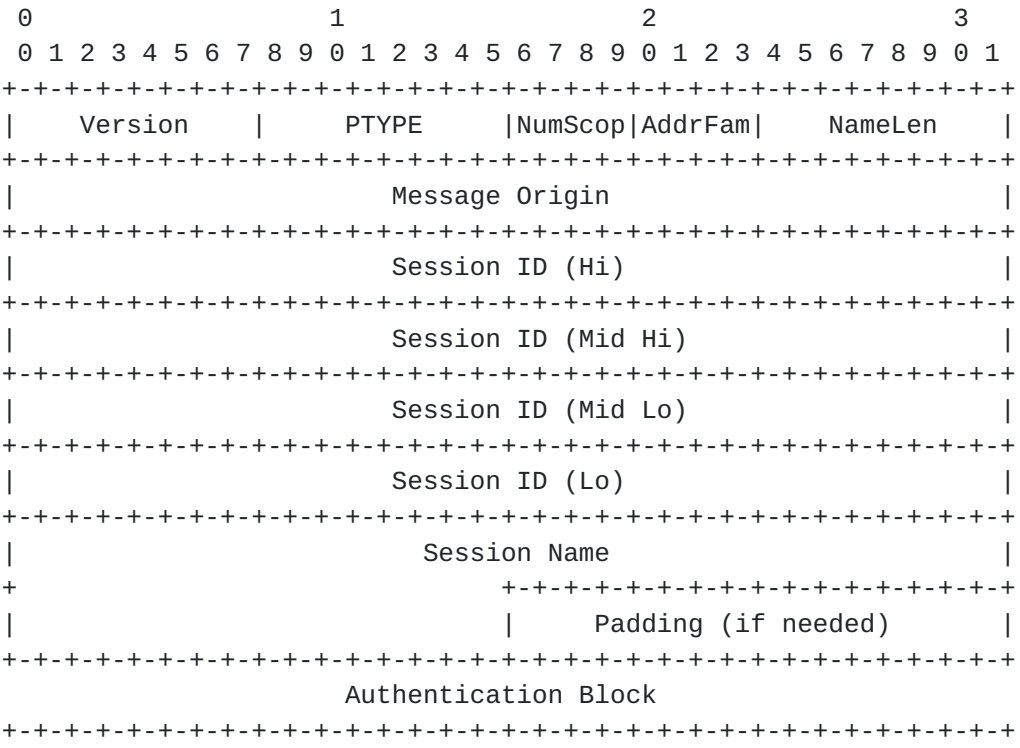


Figure 4 : SADP Server acting as proxy session member

4. Packet Formats

All SADP messages are sent over UDP, with a destination port of [SADP-PORT]. THe common SADP message header (which follows the UDP header), is shown below,



Version: 8 bits
The version defined in this document is version 0.

Packet Type (PTYPE): 8 bits
The packet types defined in this document are:
0: SADP Request
1: SADP Response

Number of Scope Entries (NumScop) : 4 bits
The number of scope entries present within a SADP_RESP message. This field should be set to zero for SADP_REQ messages.

Address Family (AddrFam): 4 bits
This indicates the format of the following packet. The following values are defined by this document:
0: IPv4
1: IPv6

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

Session ID Address: 128 bits
This 128 bit number uniquely identifies a session.

Name Len:
The length, in bytes, of the Session Name field.

Session Name: multiple of 8 bits
The Zone Name is an ISO 10646 character string in UTF-8 encoding [RFC2279] indicating the name given to the session (e.g.: ``42ndIETF-Chicago''). It should be relatively short and MUST be less than 256 bytes in length. All the session members SHOULD be configured to give the same Session Name, or a zero length string MAY be given. A zero length string is taken to mean that another session member is expected to be configured with the session name. Having ALL the members of a session announce zero length names should be considered an error.

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

Authentication Block:
The Authentication Block provides information which can be used to confirm that the sender of the SADP message is a valid member of the session. Session Members that cannot confirm that the sender of an SADP Request Message MAY ignore it, while new session members that receiver an SADP Response Message MUST ignore it. (the format of the authentication block is to be decided)

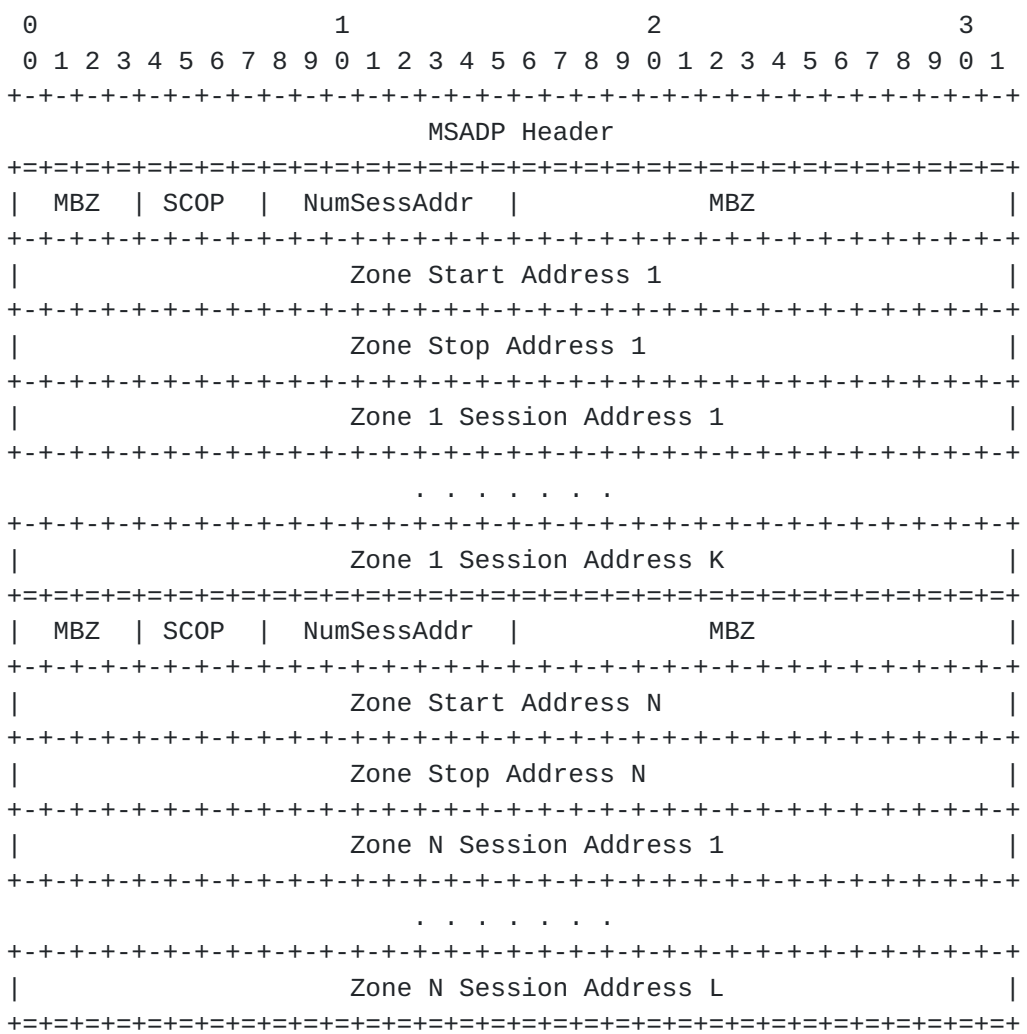
4.1 SADP Request

SADP Request (SADP_REQ) Messages have PTYPE=0, and are sent by new session members that wish to learn which administrative scopes and multicast addresses to use within a particular session. SADP_REQ Messages are sent according to the algorithm described in 3.2.

4.2 SADP Response

The SADP Response (SADP_RESP) Message has PTYPE=1, and is sent in response to a SADP_REQ Message. It contains the list of address that are to be used by a session within each scope. Session members that transmit SADP Response Messages MUST NOT include zone and address information for scopes known to be smaller that of the address upon which the triggering SADP Request Message was received.

The format for a SADP Response Message is shown below:



SCOP : 4 bits
The SCOP value associated with the zone as defined in RFC 1884 [RFC1884] for IPv6 and RFC 2365 [RFC2365] for IPv4.

NumSessAddr : 8 bits
The number of session address per scope zone that are included. Addresses will be listed in ascending order. The correspondence between address and channel function is the responsibility of the session application.

MBZ :
Must Be Zero, these bits must be set to zero, but may be used for other functions later revision of the protocol.

Zone X Start Address : 32 bits (IPv4) or 128 bits (IPv6)
The smallest address for the block of multicast addresses associated with a zone. If a zone X is valid for the range 239.128.0.0 to 239.128.255.255, this field will be set to 239.128.0.0.

Zone X Stop Address : 32 bits (IPv4) or 128 bits (IPv6)
The largest address for the block of multicast addresses associated with a zone. If a zone X is valid for the range 239.128.0.0 to 239.128.255.255, this field will be set to 239.128.255.255.

Zone X Session Address Y : 32 bits (IPv4) or 128 bits (IPv6)
Up to Y address may be included for a zone address entry, where Y is equal to the NumSessAddr value for entry X.

5. Constants

[SADP-RELATIVE-GROUP]: The relative group with each scope zone, to which session members send SADP Requests and Responses. All sessions that use administratively scoped multicast MUST subscribe to this address.

[SADP-REQ-TIMEOUT]: The time after which a session member that sends SADP Request should wait before concluding that no session members are present at the current scope. Default value is 3 seconds.

[SADP-SUPPRESSION-INTERVAL]: The interval over which a session member chooses a random delay before responding to SADP Request. Default value 2 seconds.

6. Security Considerations

SADP employs distributed mechanisms to allow new session members to learn of the existence of session-specific admin scoped multicast address. This fact lay SADP open to attack by malicious hosts that could potentially mis-inform new session members of incorrect addresses, thereby affecting a man-in-the-middle attack.

To prevent attacks of this nature by non-session members from occurring all SADP messages are signed by the sender. However, this measure does not prevent malicious hosts from joining a session and then performing the same attack. Hence, SADP's security depends upon a suitable gating process for new-member admittance combined with (as yet to be determined) mechanisms that allow spoofed SADP messages to be identified for removal before processing.

7. Acknowledgments

The Author would like to acknowledge Mark Handley and Dave Thaler for the helpful discussions and feedback which helped shape and refine this document.

8. References

[AAP] Handley, M., "The Address Allocation Protocol", Internet Draft, August 1998.

[MAAA] Handley, M., Thaler, D., and D. Estrin, "The Internet Multicast Address Allocation Architecture", Internet Draft, December 1997.

[MZAP] Handley, M., Thaler, D., "Multicast-Scope Zone Announcement Protocol (MZAP)", draft-ietf-mboned-mzap-02.txt, Internet-Draft, August, 1998.

[RFC1884] Hinden, R., Deering, S., "IP Version 6 Addressing Architecture", RFC 1884, December 1995.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP, RFC 2119, March 1997.

[RFC2279] Yergeau, F., "UTF-8, a transformation format of ISO 10646", RFC 2279, January 1998.

[RFC2365] Meyer, D., "Administratively Scoped IP Multicast", BCP, RFC 2365, July 1998.

[SHARQFEC] Kermode, R., "Scoped Hybrid Automatic Repeat reQuest with Forward Error Correction (SHARQFEC)", ACM SIGCOMM98, Vancouver Canada, September 1998.

[UUID] Leach, J., Salz, R., "UUIDs and GUIDs", draft-leach-uuids-guids-01.txt, Internet-Draft, February, 1998.

9. Author's Address

Roger Kermode
Motorola
Chicago Corporate Research Laboratories
1301 East Algonquiin Rd, MS IL02-2712
Schaumburg, IL 60196

Phone: (847) 538 4587
Email: ark008@email.mot.com

10. Full Copyright Statement