SPEERMINT
Internet-Draft
Intended status: BCP
Expires: September 5, 2009

T. Creighton
J. Livingood
Comcast

March 4, 2009

Use of DNS SRV and NAPTR Records for SPEERMINT
draft-ietf-speermint-srv-naptr-use-05

Status of this Memo

Copyright Notice

Abstract

The objective of this document is to specify the Best Current Practice (BCP) adopted by a service provider or other organization providing multimedia communication services such as Voice over Internet Protocol(VoIP) in order to locate another service provider to peer with in the context of Session PEERing for Multimedia INTerconnect. This document attempts to fill the gaps in information from RFC 3261, RFC 3263, and other documents, in order to more assist service providers in more easily performing SIP peering.

[EDITORIAL NOTE: XREF ERROR GENERATED IN ABSTRACT, XREFS REMOVED]

Table of Contents

1. Requirements Language

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

2. Introduction

A service provider needs to identify the ingress Session Initiation Protocol (SIP) (RFC 3261 [RFC3261]) server of a peering network before it can signal and route SIP-based real-time communications sessions. This function of locating the ingress SIP server of a peering network is typically performed by the egress SIP server of the service provider originating the SIP session. Also, the ingress server in the peering network needs to locate the originating service provider's egress server in situations where the peering connection is terminated after receiving a SIP request or if the egress SIP server of originating service provider fails. The SIP servers at the originating and peering sides use DNS procedures, using both SRV [RFC3261] and NAPTR [RFC3404] records, in order to locate each other. It should be noted that

3. Session Peering Setup

SIP systems are represented by user agents (UA). The diagram below shows the case of Direct Peering where a user agent (UA1), hosted by a service provider (SP1), initiates a SIP session to a User Agent

(UA2), hosted by another service provider (SP2). The egress SIP server of SP1 is a SIP Signaling Path Border Element (SBE) as defined in Section 3 of [SPEERMINT-Terminology], called SBE1, that interfaces with session peering service provider SP2. The SIP session initiated by UA1 is received by this network element, SBE1. The resource to which the SIP request needs to be routed by SBE1 is identified by a SIP or SIPS URI. For example, this could be the SIP URI of UA2 found in the Request-URI of the SIP request received by SBE1. Alternatively, for example, this could also be the next hop from SBE1 found in the topmost Route header of SIP request. In order to determine the resource to route the request to, SP1 MAY make use of ENUM [RFC3761] lookup services or some other internal lookup to determine the SIP URI of the resource. Such an ENUM lookup service may use e164.arpa or one or more other private ENUM zones. This lookup MAY be performed by SBE1 or another network element of SP1.

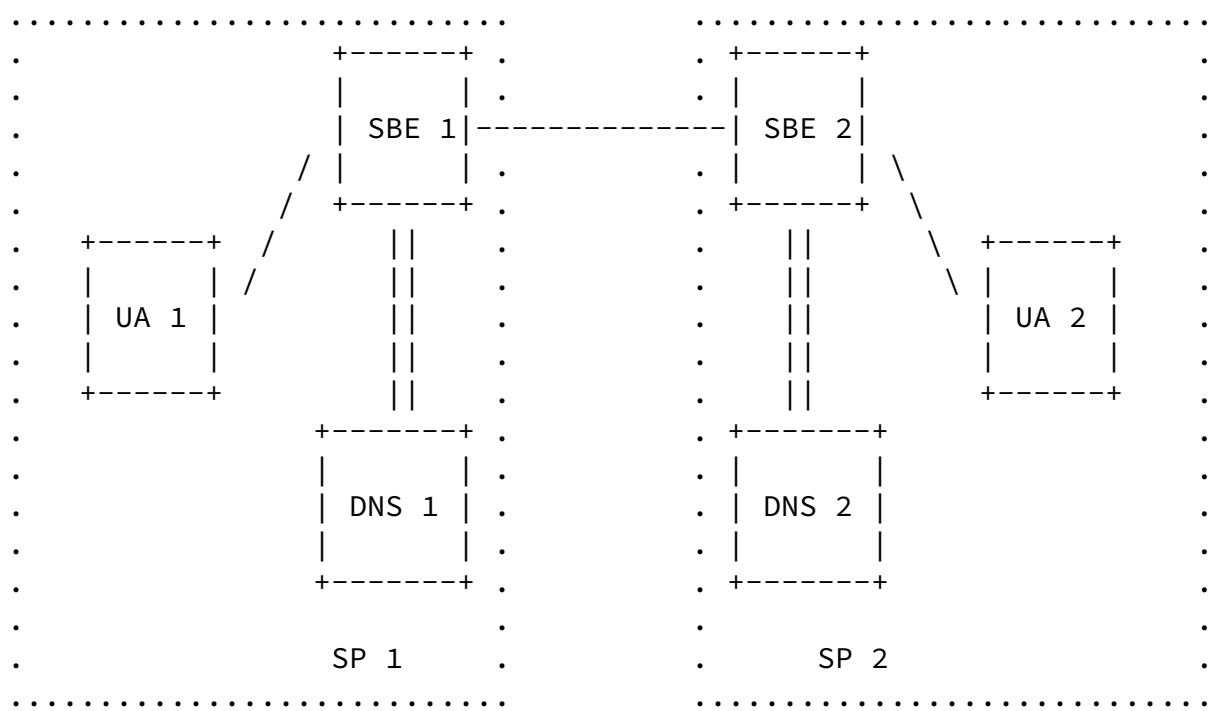


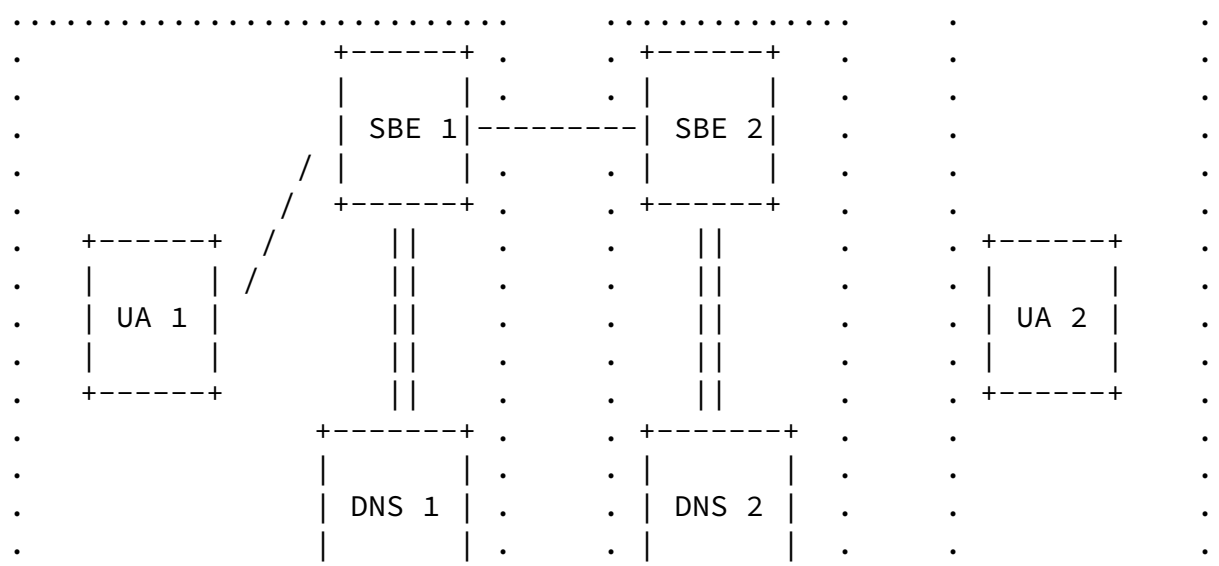
Figure 1

In order to route the SIP request to this resource in SP2, SBE1 needs to determine the ingress SIP signaling path border element for SP2, called SBE2, by resolving the SIP or SIPS URI using DNS. SBE1 makes use of the NAPTR and DNS SRV mechanism defined in [RFC3263] to determine the IP address, port, and transport protocol for peering with the SP2 ingress SIP proxy server (i.e. SBE2). SBE1 and SBE2 which are involved in the session peering, support a set of protocols and have list of preferences for these protocols. UDP, TCP and TLS MUST be supported by these proxies [EDITORIAL NOTE: QUESTION FOR WORKING GROUP ON WHAT SHOULD BE A MUST HERE - JUST TLS OR ALL?].

[EDITORIAL NOTE: ALEX SUGGESTS ADDING A REFERENCE TO TLS IN FIRST USE ABOVE]

As a best current practice, SBE1 and SBE2 SHOULD be deployed in a highly scalable and highly available manner, such as a cluster of servers. These servers are of different prioritization and weight, to ensure capacity-based load balancing. [EDITORIAL NOTE: CONSIDER REWORDING THIS LAST SENTENCE]

The figure below shows the case of Indirect Peering where SBE2 is the ingress SIP server of a transit service provider. The mechanism to locate SBE2 is the same as described for Direct Peering scenario.



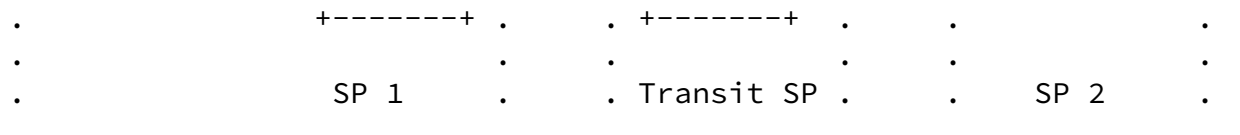
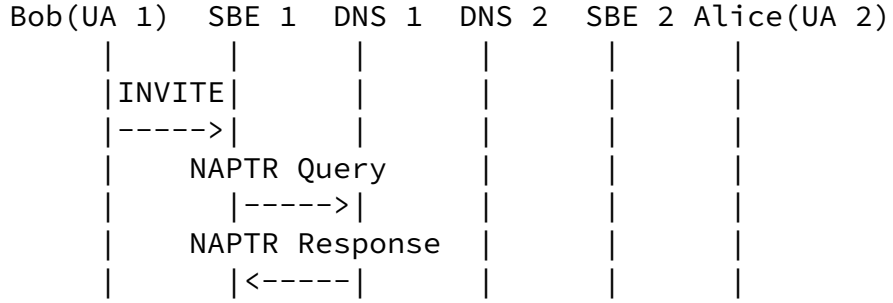


Figure 2

The figure below shows a high level SIP call flow setting up a Direct Peering session between SP1 and SP2. In this call flow a VoIP session is established between a caller, Bob (sip:bob@example1.com), in SP1 and callee, Alice(sip:alice@example2.com), in SP2 using SIP INVITE request. All SIP signaling MUST go through both SBE1 and SBE2, as these are the ingress and egress points in SP1 and SP2 networks, respectively.



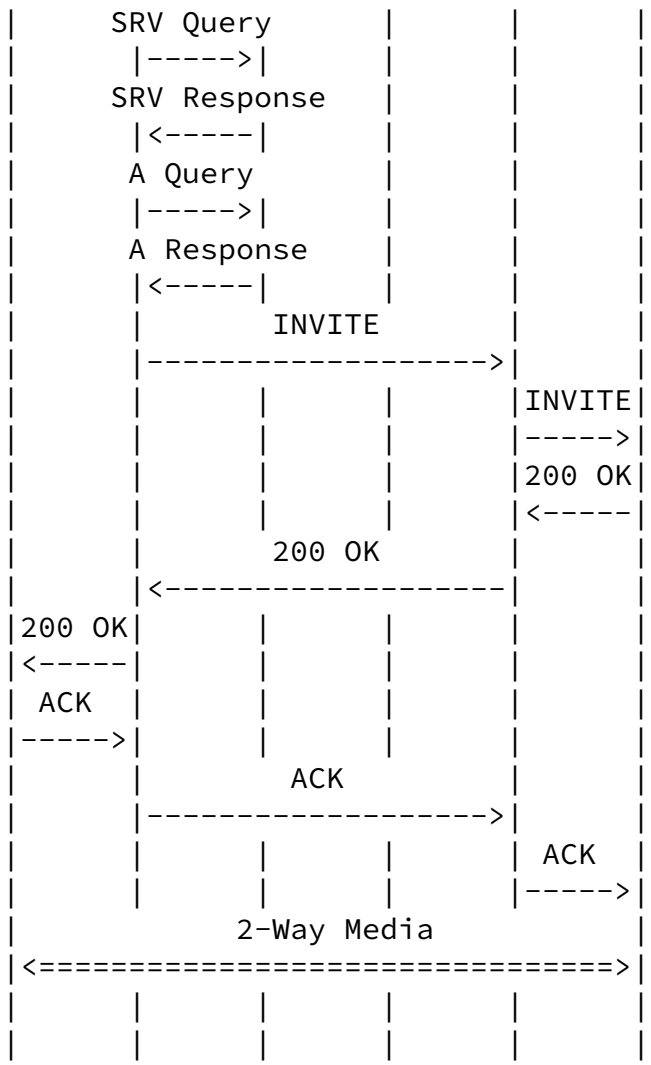


Figure 3

[EDITORIAL NOTE: WG Q - DO WE NEED A CALL FLOW FOR Indirect Peering?]

[EDITORIAL NOTE: DO WE NEED A SUB-SECTION NUMBER PRIOR TO NEXT SENTENCE? ALSO, DO WE NEED TO EXPAND IT WITH MORE TEXT?]

The target, to which the request is sent, is determined by SBE1 as follows:

3.1. Target Determination

The target resource is identified with a SIP or SIPS URI. This is the URI in the Route header, if present, or the URI from the request URI of the SIP request received by SBE1. The host value of the hostport component of the URI is the TARGET [EDITORIAL NOTE: VERIFY / CLARIFY AS NEEDED]. This TARGET is the domain to be contacted. The NAPTR/SRV/A resource record lookup as described in the following section should be skipped if the transport/port/IP address is already specified for the target URI.

3.2. NAPTR Lookup

Next, SBE1 determines the transport protocol of the TARGET, SBE2, by performing a NAPTR query for the TARGET. NAPTR processing as described in [RFC2915] will result in the discovery of the most preferred transport protocol [EDITORIAL NOTE: CONSIDER REWORDING / VALIDATE PREVIOUS 3 WORDS] of a server instance of SBE2 and SRV records.

Considering our example call flow above [EDITORIAL NOTE: INCL LOCAL REFERENCE HERE], SBE1 wishes to resolve sip:alice@example.com and performs a NAPTR query for that TARGET domain example.com, and the following NAPTR records are returned:

```
;          order   pref  flags   service         regexp   replacement
IN NAPTR  50   50  "s"    "SIPS+D2T"    ""  _sips._tcp.example.com
IN NAPTR  90   50  "s"    "SIP+D2T"     ""  _sip._tcp.example.com
IN NAPTR 100   50  "s"    "SIP+D2U"     ""  _sip._udp.example.com
```

Figure 4

DNS MUST return at least three records - one with "SIP+D2T", one with "SIP+D2U" and one with "SIPS+D2T" service type for the case of Direct and Indirect Peering (section 4.3 in [SPEERMINT-Terminology]). For Indirect Peering (section 4.4 in [SPEERMINT-Terminology]) since domain validation as specified in Section 26.3.2.2 of [RFC3261] for TLS at layer 5 will not work, SIPS over TLS cannot be used. [EDITORIAL NOTE: J.ELWELL CONCERN RAISED - IF THIS IS CASE, FEELS WE DO NOT HAVE A SECURE SOLUTION. WG FEEDBACK?]

3.3. SRV Lookup

Depending on what transport protocols SBE1 supports, SBE1 selects one from the preference list of NAPTR results and performs the SRV lookup to obtain a list of available server instances for SBE2. TLS SHOULD be the preferred transport protocol for peering between SBE1 and SBE2.

In our example, SBE1 uses TCP, the SRV lookup for _sip._tcp.example.com would return this list of available servers :

```
;;    Priority Weight   Port   Target
IN SRV  0        1      5060   server1.example.com
IN SRV  0        2      5060   server2.example.com
```

Figure 5

Alternatively, if no NAPTR records are found, then SBE1 uses the preferred transport protocol and issues an SRV query for that specific transport using "sips" for SIPS URI and SIP URI with TLS and "sip" for SIP URI as the SRV domain prefix.

In our example, SBE1 prefers to use TCP and target SIP URI of SP2 is sip:alice@example2.com, it sends a SRV query for _sip._tcp.example2.com.

The SRV responses MAY also include A and/or AAAA records with it.

3.4. Using SRV Results

If A records or AAAA records are not returned with the SRV responses, procedures from [RFC2782] describes how to use and interpret the results obtained from the SRV query. The target entry of the SRV RRs is looked up by querying the DNS for address records. If the SRV response from DNS includes A or AAAA records with it, it will cut down on round trips and lookup of DNS again for target entry. On determining the transport protocol, service, port and address record from the SRV RRs as described above, the SBE1 will try to connect to the (protocol, address, service). Once the connection is established to an available instance of SBE2, SBE1 sends the SIP request to SBE2. SBE1 MUST act in a stateful manner and any retransmission of SIP requests for a specific SIP transaction, including ACKS for non-2xx response or CANCEL for that SIP transaction MUST go to the same server instance of SBE2.

When SBE1 sends the SIP request to SBE2, it SHOULD set the sent-by parameter of the topmost Via header in the SIP request to a domain that identifies SBE1. It MUST NOT specify the port.

[EDITORIAL NOTE: SHOULD THE ABOVE SENTENCE SAY MUST NOT OR SOMETHING ELSE?]

[EDITORIAL NOTE: ALEX SUGGESTS HAVING A SIP EXPERT REVIEW 4.2 ON SENT-BY]

[EDITORIAL NOTE: ALEX SUGGESTS A CALL FLOW FOR 4.2 - DOES WG AGREE?]

4. High Availability

High Availability is ensured by detecting failures in the ability to connect to SBE1 and SBE2 server instances. In the event of a failure, when SBE1 tries to send SIP INVITE to SBE2, the following failures could occur:

4.1. SBE1 Fails to Reach SBE2

A 503 error response is reported by the transaction layer, or failure can occur at the transport layer due to TCP disconnect in connection, ICMP error in UDP or time out at transport layer or SIP layer timeout when its not receiving any SIP response. In such situations, SBE1 tries a new SIP request transaction to the next available server instance of SBE2 as determined by SRV RRs entry. The SIP T1 timer on SBE1 SHOULD be configurable with a upper limit value of 500ms. A shorter value of T1, say 100ms, reflects a faster fail-over support.

[EDITORIAL NOTE: ALEX FINDS FIRST SENTENCE ABOVE CONFUSING - CONSIDER REWORDING IT]

[EDITORIAL NOTE: ALEX SUGGESTS ADDING A REF TO PROPER SECTION OF SRV DOC, AS WELL AS REF FOR SIP TIMERS]

4.2. Using SRV Results

Failure may also occur after the request is received by SBE2 from SBE1 due to closure of the transport connection the request came in on at SBE2, before the response can be sent back to SBE1. In this situation, SBE2 uses the domain value present in the 'sent-by' parameter in the top most Via header of the received SIP INVITE, and queries for SRV records at this domain name using the service identifier "_sips" if the Via transport is "TLS", "_sip" otherwise. The sorted list of SRV RRs are obtained and used as described in [RFC2782] to send the response back to SBE1. If the topmost element in the list of server instances of SBE1 fails, the next available one is tried.

[EDITORIAL NOTE: FOR NEXT REV - SHOULD WE ADD CALL FLOW FOR FAILURE

SCENARIOS DESCRIBED IN 4.1 AND 4.2?]

[EDITORIAL NOTE: ALEX RECOMMENDS WE CLARIFY WHETHER THE DESCRIBED FUNCTIONALITY HERE IS 'ON TOP' OF STANDARD SIP OR NEW FUNCTIONALITY, AND THAT A REF TO THE PROPER SECTION OF RFC3261 IS CONSIDERED]

5. Caching/TTL

5.1. Caching

SBE SHOULD use caching of DNS results to eliminate unnecessary DNS queries.

5.2. TTL

SRV RRs have a TTL value based on which the SBE1 caches the entry for that duration, if it supports caching, and any further requests to the same TARGET domain are delivered to the cached server instance. The TTL recommended for SRV is about 1 hr. The TTL for NAPTR is much higher, about 1 day (24hrs) since the NAPTR records do not vary that often as compared to SRV.

6. Security Considerations

This document introduces no new security considerations.

7. IANA Considerations

There are no IANA considerations in this document.

8. Acknowledgements

Special thanks go to Yiu Lee for his valuable input to this document, as well as John Elwell, Alexander Mayrhofer, and Chris Griffiths for their detailed reviews of this document.

9. References

9.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2782] Gulbrandsen, A., Vixie, P., and L. Esibov, "A DNS RR for specifying the location of services (DNS SRV)", RFC 2782, February 2000.

[RFC2915] Mealling, M. and R. Daniel, "The Naming Authority Pointer (NAPTR) DNS Resource Record", RFC 2915, September 2000.

[RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, June 2002.

[RFC3263] Rosenberg, J. and H. Schulzrinne, "Session Initiation Protocol (SIP): Locating SIP Servers", RFC 3263, June 2002.

[RFC3404] Mealling, M., "Dynamic Delegation Discovery System (DDDS) Part Four: The Uniform Resource Identifiers (URI)", RFC 3404, October 2002.

[RFC3667] Bradner, S., "IETF Rights in Contributions", RFC 3667, February 2004.

[SPEERMINT-Terminology] Malas, D. and D. Meyer, "SPEERMINT Terminology", November 2008.

9.2. Informative References

[RFC2434] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.

[RFC3761] Faltstrom, P. and M. Mealling, "The E.164 to Uniform Resource Identifiers (URI) Dynamic Delegation Discovery System (DDDS) Application (ENUM)", RFC 3761, April 2004.

Appendix A. Document Change Log

[RFC Editor: This section is to be removed before publication]

draft-ietf-speermint-srv-naptr-use-05:

- jason: addressed some of John and Alex's questions and issues
- jason: moved ref to rfc3761 from normative to informative
- jason: highlighted open editorial questions and issues that need to be closed before WGLC

draft-ietf-speermint-srv-naptr-use-04:

- jason: addressed feedback from several people received on -02 version of draft

- jason: still have about 15 discrete pieces of feedback to include
- jason: also highlighted in [BRACKETS] several areas that need minor work

draft-ietf-speermint-srv-naptr-use-03:

- jason: converted from MS Word template to XML

Appendix B. Open Issues

Decide what we want to do in Using SRV Results section

Benny Rodrig suggests adding some more description of how this works in the indirect case

Several open issues with John Elwell

Several open issues with Alex Mayrhofer

Several editorial notes to close out

Authors' Addresses

Tom Creighton
Comcast Cable Communications
One Comcast Center
1701 John F. Kennedy Boulevard
Philadelphia, PA 19103
US

Email: tom_creighton@cable.comcast.com
URI: http://www.comcast.com

Jason Livingood
Comcast Cable Communications
One Comcast Center
1701 John F. Kennedy Boulevard
Philadelphia, PA 19103
US

Email: jason_livingood@cable.comcast.com
URI: http://www.comcast.com