Open Community Specification
POCS-1

Using OPTIONS to Query for Operational Status in the Session Initiation Protocol (SIP)

March 2012

Summary

This document describes a procedure for using the Session Initiation Protocol (SIP) OPTIONS method in order to allow one SIP entity to query the operational status of another SIP entity. Through discovery of the status of a SIP entity, it is possible to expedite session establishment or provide diagnostic information to network administrators.

Author(s)

Paul E. Jones <paulej@packetizer.com>
Gonzalo Salgueiro <gsalguei@cisco.com>
Vivek Bhargava <vbharga@cisco.com>

Acknowledgements

Table of Contents

Using OPTIONS to Query for Operational Status in the Session Initiation Protocol (SIP)

1   Introduction

In order to build efficient and robust networks that utilize the Session Initiation Protocol (SIP) [1], it is sometimes necessary for one SIP entity to know the status of another SIP entity.   Having this knowledge can better enable intelligent routing of messages when establishing a dialog.   This knowledge may also be used to alert network administrators to potential problems with SIP entities so that corrective action can be taken to maintain healthy SIP infrastructure with expected service performance and experience.

The SIP specification defines the REGISTER method that enables a SIP user agent to register with a registrar and to send periodic registration ("refresh") messages.   The purpose for the REGISTER message is to add and remove bindings, though it may also serve as a type of "keep-alive" mechanism to assist in determining whether a user agent is presently available.   However, a REGISTER message is not the appropriate method to communicate SIP entity status between any two SIP entities, such as between two SIP proxy servers or between a Back-to-Back User Agent (B2BUA) and any number of peer B2BUAs that help facilitate communication between administrative domains.

One use of the OPTIONS method is to enable a SIP entity to query for the capabilities of a remote SIP entity in advance of establishing a dialog, which may help in selecting an appropriate target user agent.   By extending the scope of this method it is possible to enable any SIP entity to query for the operational status of another SIP entity, thereby achieving the objective of equipping SIP entities with more knowledge about the operational status of peer entities.   Presently, there is no other means to determine the operational status of a peer SIP entity.

This document defines the usage of the OPTIONS message in order to determine the operational status of any SIP entity.

It is worth noting that OPTIONS is widely used today by numerous equipment manufacturers to determine operational status.   Unfortunately, use of OPTIONS in this way is not always implemented consistently from one manufacturer to another.   This specification aims to address this consistency issue and to provide implementers with a reference against which to implement.

2   Conventions used in this document

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].

3   Required Support

Transmission of the OPTIONS message between any two entities is optional.

The SIP specification requires that all user agents support the OPTIONS message. To comply with procedures described in this document, all SIP entities, including SIP user agents and proxy servers, shall support the OPTIONS message when received outside of a dialog in accordance with this specification.

4   Problem Statement

A SIP-based communication network relies on call signaling messages exchanged between communicating SIP entities. A robust communications network provides for redundancy such that failure of no one entity will interrupt communications through the network. SIP provides a way to find the next hop entity and redirect messages dynamically through the use of DNS. Each SIP entity may query DNS for a list of addresses of next hop entities to which to forward a message.

In cases where DNS is not used, a SIP entity may be statically provisioned with information relating to next hop entities.

Either with the use of DNS or through static provisioning, there are no mechanisms to prevent delays that may result from messages sent to unresponsive SIP entities. To prevent unnecessary delay due to timeouts and subsequent message re-transmissions, a method by which each entity may discover the operational status of its communicating peers is needed. Note that communication delays may still exist due to network congestion, but such issues are outside the scope of this specification.

This specification defines the use of OPTIONS to determine the operational status of peer SIP entities to reduce delays and improve overall operational efficiency.

This specification does not replace the use of OPTIONS as described in RFC 3261, namely to discover the communicating entity's capabilities, but does further expand on use of OPTIONS to determine the basic state of a SIP entity as per RFC 3261 Section 11.2. The use of OPTIONS as described in this document is strictly as a mechanism to determine the operational status of a neighboring entity. This specification also does not impose any new requirements on the underlying transport protocol.

5 Procedure for Using OPTIONS to Query for Operational Status

All SIP entities are required to respond to OPTIONS messages, though not all SIP entities are required to transmit OPTIONS messages. Implementations that adhere to this specification shall transmit OPTIONS messages for the purpose of determining operational status as described in subsequent sections. Further, implementations that adhere to this specification shall respond to OPTIONS messages that query for operational status as defined below.

These procedures are often informally referred to as an "OPTIONS ping".

5.1 Transmitting OPTIONS Messages

A SIP entity transmits an OPTIONS message outside of a dialog periodically or as desired in order to determine the status of another SIP entity. For the purposes of this specification, the transmitting and receiving SIP entities may be SIP User Agents, including B2BUAs, or SIP proxy servers. The syntax of the OPTIONS message is described in RFC 3261.

The reason for sending OPTIONS messages as per this specification is to discover whether one or more neighboring SIP entities are presently operable and able to process signaling messages. This includes basic IP reachability, SIP application-level reachability, ability to deliver advanced SIP-based services (e.g., video), and so on. Having this information in advance can help reduce the time required to establish a SIP session and might help avoid session establishment failures.

While any addressing mechanism might be used to transmit OPTIONS messages, use of IP addresses is recommended in this specification. When a SIP entity is configured to send OPTIONS messages to peer SIP entities and is configured with a hostname, the transmitting SIP entity shall first resolve the name using DNS, looking for both multiple address records and SRV records, which may in turn be associated with multiple address records.

The OPTIONS message SHOULD be addressed to a peer SIP entity using a SIP URI of the form "sip:hostport" (see Section 25.1 of RFC 3261 for definitions). For example, "sip:192.168.1.5" is preferred, whereas "sip:bob@example.com" is not preferred since it might resolve to a plurality of addresses. The request URI required when addressing a proxy is described in Section 11 of RFC 3261.

The OPTIONS message MUST be transmitted directly to the peer's IP address and not to an intermediary SIP entity. Further, SIP entities SHOULD set the Max-Forwards header field value to 01.

To avoid unduly taxing a receiving SIP entity, transmitters of OPTIONS messages shall honor the Retry-After header field if received.

A SIP entity may transmit an OPTIONS message to a peer SIP entity, even when there are other ongoing message exchanges. The reason is that, though a receiving SIP entity may be responsive to other SIP messages, it may have been put into a maintenance state, meaning it

should not facilitate the establishment of new SIP sessions. Use of OPTIONS messages as per this specification can help facilitate the graceful shutdown of equipment.

5.2 Processing OPTIONS messages

A SIP entity that receives an OPTIONS message MUST respond with a status code indicative of its present ability to process SIP signaling messages. Refer to Section 6 for response codes.

6 SIP Response Codes Handling

A SIP entity shall respond to an OPTIONS method with a 200 response code when it is willing and able to process SIP messages, unless one of the following response codes is more appropriate.

When a receiving SIP entity is unable to process additional SIP messages, it should respond with a 503 response code. A 503 response code is also used when a SIP entity is placed into a maintenance mode to indicate that it should not accept new SIP dialogs. A receiving SIP entity may include a Retry-After header in a response to the OPTIONS message to avoid further requests for a desired period of time.

Note that a SIP entity may respond with other 5xx or 4xx error codes as appropriate and as recommended in RFC 3261.

7 Frequency of Message Transmission

Any two entities that communicate with each other on a regular basis may be configured to transmit an OPTIONS message from time to time as provisioned by the network administrator. The frequency with which an OPTIONS message is transmitted is outside the scope of this document. Having said that, the frequency with which OPTIONS messages are transmitted should not place an undue burden on SIP entities.

Transmission of an OPTIONS message for the purpose of learning the operational status of a remote SIP entity may seem unnecessary if there is already active communications with the remote entity. However, using OPTIONS, even when there is already active communication, allows a querying SIP entity to learn when another SIP entity is reaching an overload state or when it has been put into a maintenance mode.

If a requesting entity fails to receive a response to an OPTIONS message, it may retransmit that message following the procedures defined in RFC 3261. If a requesting SIP entity receives a 486 or 503 response code, it can send a subsequent OPTIONS messages in order to detect a change in operational status, but it should, as per RFC 3261, honor the Retry-After header field received in the previous response.

8 References

[1]   Rosenberg, J., et al, "SIP: Session Initiation Protocol", RFC 3261, June 2002.

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