Independent Submission
Request for Comments: 6193
Category: Informational
ISSN: 2070-1721
April 2011

M. Saito
NTT Communications
D. Wing
Cisco Systems
M. Toyama
NTT Corporation

Media Description for the Internet Key Exchange Protocol (IKE) in the Session Description Protocol (SDP)

Abstract

This document specifies how to establish a media session that represents a virtual private network using the Session Initiation Protocol for the purpose of on-demand media/application sharing between peers. It extends the protocol identifier of the Session Description Protocol (SDP) so that it can negotiate use of the Internet Key Exchange Protocol (IKE) for media sessions in the SDP offer/answer model. It also specifies a method to boot up IKE and generate IPsec security associations using a self-signed certificate.

Status of This Memo

Copyright Notice

Table of Contents

1. Applicability Statement

This document provides information about a deployed use of the Session Initiation Protocol (SIP) [RFC3261] for the Internet community. It is not currently an IETF standards track proposal. The mechanisms in this document use SIP as a name resolution and authentication mechanism to initiate an Internet Key Exchange Protocol (IKE) [RFC5996] session. The purpose of this document is to establish an on-demand virtual private network (VPN) to a home router that does not have a fixed IP address using self-signed certificates. It is only applicable under the condition that the integrity of the Session Description Protocol (SDP) [RFC4566] is assured. The method to ensure this integrity of SDP is outside the scope of this document. This document specifies the process in which a pair of SIP user agents resolve each other's names, exchange the fingerprints of their self-signed certificates securely, and agree to establish an IPsec-based VPN [RFC4301]. However, this document does not make any modifications to the specifications of IPsec/IKE. Despite the limitations of the conditions under which this document can be applied, there are sufficient use cases in which this specification is helpful, such as the following:

- Sharing media using a framework developed by Digital Living Network Alliance (DLNA) or similar protocols over VPN between two user devices.

- Accessing remote desktop applications over VPN initiated by SIP call. As an additional function of click-to-call, a customer service agent can access a customer's PC remotely to troubleshoot the problem while talking with the customer over the phone.

- Accessing and controlling medical equipment (medical robotics) remotely to monitor the elderly in a rural area (remote care services).

- Using a LAN-based gaming protocol based on peer-to-peer rather than via a gaming server.

2. Introduction

This section describes the problem in accessing home networks and provides an overview of the proposed solution.

2.1. Problem Statement

Home servers and network-capable consumer electronic devices have been widely deployed. People using such devices are willing to share content and applications and are therefore seeking ways to establish multiple communication channels with each other. However, there are several obstacles to be overcome in the case of remote home access.

It is often not possible for a device outside the home network to connect to another device inside the home network because the home device is behind a network address translation (NAT) or firewall that allows outgoing connections but blocks incoming connections. One effective solution for this problem is VPN remote access to the NAT device, which is usually a home router. With this approach, once the external device joins the home network securely, establishing connections with all the devices inside the home will become easy because popular LAN-based communication methods such as DLNA can be used transparently. However, there are more difficult cases in which a home router itself is located behind the NAT. In such cases, it is also necessary to consider NAT traversal of the remote access to the home router. In many cases, because the global IP address of the home router is not always fixed, it is necessary to make use of an effective name resolution mechanism.

In addition, there is the problem of how a remote client and a home router authenticate each other over IKE to establish IPsec for remote access. It is not always possible for the two devices to securely exchange a pre-shared key in advance. Administrative costs can make it impractical to distribute authentication certificates signed by a well-known root certification authority (CA) to all the devices. In addition, it is inefficient to publish a temporary certificate to a device that does not have a fixed IP address or hostname. To resolve these authentication issues, this document proposes a mechanism that enables the devices to authenticate each other using self-signed certificates.

2.2. Approach to Solution

This document proposes the use of SIP as a name resolution and authentication mechanism because of three main advantages:

- Delegation of Authentication to Third Party

  Devices can be free from managing their signed certificates and whitelists by taking advantage of authentication and authorization mechanisms supported by SIP.

o UDP Hole Punching for IKE/IPsec

SIP has a cross-NAT rendezvous mechanism, and Interactive Connectivity Establishment (ICE) [RFC5245] has a function to open ports through the NAT. The combination of these effective functions can be used for general applications as well as real-time media. It is difficult to set up a session between devices without SIP if the devices are behind various types of NAT.

o Reuse of Existing SIP Infrastructure

SIP servers are widely distributed as a scalable infrastructure, and it is quite practical to reuse them without any modifications.

Today, SIP is applied to not only Voice over IP (VoIP) but also various applications and is recognized as a general protocol for session initiation. Therefore, it can also be used to initiate IKE/IPsec sessions.

However, there is also a specification that uses a self-signed certificate for authentication in the SIP/SDP framework. "Connection-Oriented Media Transport over the Transport Layer Security (TLS) Protocol in the Session Description Protocol (SDP)" [RFC4572] (hereafter referred to as comedia-tls) specifies a method to exchange the fingerprint of a self-signed certificate to establish a Transport Layer Security (TLS) [RFC5246] connection. This specification defines a mechanism by which self-signed certificates can be used securely, provided that the integrity of the SDP description is assured. Because a certificate itself is used for authentication not only in TLS but also in IKE, this mechanism will be applied to the establishment of an IPsec security association (SA) by extending the protocol identifier of SDP so that it can specify IKE.

One easy method to protect the integrity of the SDP description, which is the premise of this specification, is to use the SIP identity [RFC4474] mechanism. This approach is also referred to in [RFC5763]. Because the SIP identity mechanism can protect the integrity of a body part as well as the value of the From header in a SIP request by using a valid Identity header, the receiver of the request can establish secure IPsec connections with the sender by confirming that the hash value of the certificate sent during IKE negotiation matches the fingerprint in the SDP. Although SIP identity does not protect the identity of the receiver of the SIP request, SIP-connected identity [RFC4916] does. Note that the possible deficiencies discussed in [RFC4474-Concerns] could affect this specification if SIP identity is used for the security mechanism.

Considering the above background, this document defines new media formats "ike-esp" and "ike-esp-udpencap", which can be used when the protocol identifier is "udp", to enable the negotiation of using IKE for media sessions over SDP exchange on the condition that the integrity of the SDP description is assured. It also specifies the method to set up an IPsec SA by exchanging fingerprints of self-signed certificates based on comedia-tls, and it notes the example of SDP offer/answer [RFC3264] and the points that should be taken care of by implementation. Because there is a chance that devices are behind NAT, this document also covers the method to combine IKE/IPsec NAT-Traversal [RFC3947][RFC3948] with ICE. In addition, it defines the attribute "ike-setup" for IKE media sessions, similar to the "setup" attribute for TCP-based media transport defined in RFC 4145 [RFC4145]. This attribute is used to negotiate the role of each endpoint in the IKE session.

2.3. Alternative Solution under Prior Relationship between Two Nodes

Under quite limited conditions, certificates signed by trusted third parties or pre-shared keys between endpoints could be used for authentication in IKE, using SIP servers only for name resolution and authorization of session initiation. Such limited cases are addressed in Section 8.

2.4. Authorization Model

In this document, SIP servers are used for authorization of each SIP call. The actual media sessions of IPsec/IKE are not authorized by SIP servers but by the remote client and the home router based on the information in SIP/SDP. For example, the home router recognizes the remote client with its SIP-URI and IP address in the SDP. If it decides to accept the remote client as a peer of a VPN session, it will accept the following IKE session. Then, during the IKE negotiation, the certificate fingerprint in the SDP is compared with the certificate exchanged in the IKE session. If they match, IKE negotiation continues. Only a successful IKE negotiation establishes an IPsec session with the remote peer.

2.5. 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 [RFC2119].

3. Protocol Overview

Figure 1 shows a case of VPN remote access from a device outside the home to a home router whose IP address is not fixed. In this case, the external device, a remote client, recognizes the Address of Record of the home router but does not have any information about its contact address and certificate. Generally, establishing an IPsec SA dynamically and securely in this situation is difficult. However, as specified in comedia-tls [RFC4572], if the integrity of SDP session descriptions is assured, it is possible for the home router and the remote client to have a prior relationship with each other by exchanging certificate fingerprints, i.e., secure one-way hashes of the distinguished encoding rules (DER) form of the certificates.

```
              REGISTRATION                REGISTRATION
                 (1)       +----------+      (1)
            +------------->|          |<---------+
            |    INVITE(2) |          |          |
            | +----------->|   SIP    |--------+ |
            | |  200 OK(2) |   Proxy  |        | |
            | | +----------|          |<-----+ | |
            | | |          |          |      | | |  _________
            | | V          +----------+      | V | /         \
         +----------+ IKE (Media Session) +---------+         \
         | Remote   |<---------(3)------->| Home    |  Home    \
         | Client   |                     | Router  | Network   |
         |         ============(4)====================          |
         |(SIP UAC) |     VPN (IPsec SA)  |(SIP UAS)|          /
         +----------+                     +---------+         /
                                                   \_________/
```

Figure 1: Remote Access to Home Network

(1) Both Remote Client and Home Router generate secure signaling channels. They may REGISTER to SIP Proxy using TLS.

(2) Remote Client sends an offer SDP with an INVITE request to Home Router, and Home Router returns an answer SDP with a reliable response (e.g., 200 OK). Both exchange the fingerprints of their self-signed certificates in SDP during this transaction. Remote Client does not accept an answer SDP with an unreliable response as the final response.

(3) After the SDP exchange, Remote Client, which has the active role, initiates IKE with Home Router, which has the passive role, to establish an IPsec SA. Both validate that the certificate presented in the IKE exchange has a fingerprint that

matches the fingerprint from SDP. If they match, IKE negotiation proceeds as normal.

(4) Remote Client joins the Home Network.

By this method, the self-signed certificates of both parties are used for authentication in IKE, but SDP itself is not concerned with all the negotiations related to key-exchange, such as those of encryption and authentication algorithms. These negotiations are up to IKE. In many cases where IPsec is used for remote access, a remote client needs to dynamically obtain a private address inside the home network while initiating the remote access. Therefore, the IPsec security policy also needs to be set dynamically at the same time. However, such a management function of the security policy is the responsibility of the high-level application. SDP is not concerned with it. The roles of SDP here are to determine the IP addresses of both parties used for IKE connection with c-line in SDP and to exchange the fingerprints of the certificates used for authentication in IKE with the fingerprint attribute in SDP.

4. Protocol Identifiers

This document defines two SDP media formats for the "udp" protocol under the "application" media type: "ike-esp" and "ike-esp-udpencap". The format "ike-esp" indicates that the media described is IKE for the establishment of an IPsec security association as described in IPsec Encapsulating Security Payload (ESP) [RFC4303]. In contrast, "ike-esp-udpencap" indicates that the media described is IKE, which is capable of NAT traversal for the establishment of UDP encapsulation of IPsec packets through NAT boxes as specified in [RFC3947] and [RFC3948]. Even if the offerer and answerer exchange "ike-esp-udpencap", IKE conforming to [RFC3947] and [RFC3948] can end up establishing a normal IPsec tunnel when there is no need to use UDP encapsulation of IPsec. Both the offerer and answerer can negotiate IKE by specifying "udp" in the "proto" field and "ike-esp" or "ike-esp-udpencap" in the "fmt" field in SDP.

In addition, this document defines a new attribute "ike-setup", which can be used when the protocol identifier is "udp" and the "fmt" field is "ike-esp" or "ike-esp-udpencap", in order to describe how endpoints should perform the IKE session setup procedure. The "ike-setup" attribute indicates which of the end points should initiate the establishment of an IKE session. The "ike-setup" attribute is charset-independent and can be a session- or media-level attribute. The following is the ABNF of the "ike-setup" attribute.

```
ike-setup-attr = "a=ike-setup:" role
role           = "active" / "passive" / "actpass"

'active':   The endpoint will initiate an outgoing session.
'passive':  The endpoint will accept an incoming session.
'actpass':  The endpoint is willing to accept an incoming
            session or to initiate an outgoing session.
```

Both endpoints use the SDP offer/answer model to negotiate the value of "ike-setup", following the procedures determined for the "setup" attribute defined in Section 4.1 of [RFC4145]. However, "holdconn", as defined in [RFC4145], is not defined for the "ike-setup" attribute.

Offer	Answer
active	passive
passive	active
actpass	active / passive

The semantics for the "ike-setup" attribute values of "active", "passive", and "actpass" in the offer/answer exchange are the same as those described for the "setup" attribute in Section 4.1 of [RFC4145], except that "ike-setup" applies to an IKE session instead of a TCP connection. The default value of the "ike-setup" attribute is "active" in the offer and "passive" in the answer.

5. Normative Behavior

In this section, a method to negotiate the use of IKE for media sessions in the SDP offer/answer model is described.

5.1. SDP Offer and Answer Exchange

An offerer and an answerer negotiate the use of IKE following the usage of the protocol identifiers defined in Section 4. If IPsec NAT-Traversal is not necessary, the offerer MAY use the media format "ike-esp" to indicate an IKE session.

If either of the endpoints that negotiate IKE is behind the NAT, the endpoints need to transmit both IKE and IPsec packets over the NAT. That mechanism is specified in [RFC3947] and [RFC3948]: both endpoints encapsulate IPsec-ESP packets with a UDP header and multiplex them into the UDP path that IKE generates.

To indicate this type of IKE session, the offerer uses "ike-esp-udpencap" media lines. In this case, the offerer MAY decide their transport addresses (combination of IP address and port) before starting IKE, making use of the ICE framework. Because UDP-encapsulated ESP packets and IKE packets go through the same UDP hole of a NAT, IPsec NAT-Traversal works if ICE reserves simply one UDP path through the NAT. However, those UDP packets need to be multiplexed with Session Traversal Utilities for NAT (STUN) [RFC5389] packets if ICE is required to use STUN. A method to coordinate IPsec NAT-Traversal and ICE is described in Sections 5.4 and 5.5.

The offer MAY contain media lines for media other than "ike-esp" or "ike-esp-udpencap". For example, audio stream may be included in the same SDP to have a voice session when establishing the VPN. This may be useful to verify that the connected device is indeed operated by somebody who is authorized to access it, as described in Section 9. If that occurs, the negotiation described in this specification occurs only for the "ike-esp" or "ike-esp-udpencap" media lines; other media lines are negotiated and set up normally. If the answerer determines it will refuse the IKE session without beginning the IKE negotiation (e.g., the From address is not on the permitted list), it SHOULD reject the "ike-esp" or "ike-esp-udpencap" media line in the normal manner by setting the port number in the SDP answer to 0 and SHOULD process the other media lines normally (only if it is still reasonable to establish that media without VPN).

If the offerer and the answerer agree to start an IKE session by the offer/answer exchange, they will start the IKE setup. Following the comedia-tls specification [RFC4572], the fingerprint attribute, which may be either a session- or a media-level SDP attribute, is used to exchange fingerprints of self-signed certificates. If the fingerprint attribute is a session-level attribute, it applies to all IKE sessions and TLS sessions for which no media-level fingerprint attribute is defined.

Note that it is possible for an offerer to become the IKE responder and an answerer to become the IKE initiator. For example, when a Remote Access Server (RAS) sends an INVITE to an RAS client, the server may expect the client to become an IKE initiator. In this case, the server sends an offer SDP with ike-setup:passive and the client returns an answer SDP with ike-setup:active.

5.2. Maintenance and Termination of VPN Session

If the high-level application recognizes a VPN session as the media session, it MAY discard the IPsec SA and terminate IKE when that media session is terminated by a BYE request. Therefore, the application aware of the VPN session MUST NOT send a BYE request as long as it needs the IPsec SA. On the other hand, if the high-level application detects that a VPN session is terminated, it MAY terminate the media associated with the VPN or the entire SIP session. Session timers in SIP [RFC4028] MAY be used for the session maintenance of the SIP call, but this does not necessarily ensure that the VPN session is alive. If the VPN session needs session maintenance such as keep-alive and rekeying, it MUST be done utilizing its own maintenance mechanisms. SIP re-INVITE MUST NOT be used for this purpose. Note that each party can cache the certificate of the other party as described in the Security Considerations section of comedia-tls [RFC4572].

5.3. Forking

Forking to multiple registered instances is outside the scope of this document. At least, it is assumed that a User Agent Client (UAC) establishes a session with only one User Agent Server (UAS). Encountering forked answers should be treated as an illegal process, and the UAC should cancel the session.

5.4. Port Usage

IKE generally uses local UDP port 500, but the IPsec NAT-Traversal specification requires a port transition to local UDP port 4500 during IKE negotiation because IPsec-aware NAT may multiplex IKE sessions using port 500 without changing the port number. If using ICE for IPsec Nat-Traversal, this port transition of IKE means ICE has to generate an additional UDP path for port 4500, and this would be unnecessary overhead. However, IPsec NAT-Traversal allows an IKE session to use local UDP port 4500 from the beginning without using port 500. Therefore, the endpoints SHOULD use their local UDP port 4500 for an IKE session from the beginning, and ICE will only need to generate a UDP path of port 4500.

When using ICE, a responder's IKE port observed by an initiator is not necessarily 500 or 4500. Therefore, an IKE initiator MUST allow any destination ports in addition to 500 and 4500 for the IKE packets that it sends. An IKE initiator just initiates an IKE session to the port number decided by an SDP offer/answer or ICE.

5.5. Multiplexing UDP Messages When Using ICE

Conforming to ICE, an offerer and an answerer start a STUN connectivity check after SDP exchange. Then the offerer initiates the IKE session making use of the UDP path generated by STUN packets. In addition, UDP-encapsulated ESP packets are multiplexed into the same UDP path as IKE. Thus, it is necessary to multiplex the three different packets, STUN, IKE, and UDP-encapsulated ESP, into the same UDP path. This section describes how to demultiplex these three packets.

At the first step, the endpoint that received a UDP packet at the multiplexed port MUST check the first 32 bits (bits 0-31) of the UDP payload. If they are all 0, which is defined as a non-ESP marker, that packet MUST be treated as an IKE packet.

Otherwise, it is judged as an ESP packet in the IPsec NAT-Traversal specification. It is furthermore necessary to distinguish STUN from ESP. Therefore, the bits 32-63 from the beginning of the UDP payload MUST be checked. If the bits do not match the magic cookie of STUN 0x2112A442 (most packets do not match), the packet is treated as an ESP packet because it is no longer a STUN packet.

However, if the bits do match the magic cookie, an additional test is necessary to determine if the packet is STUN or ESP. The magic cookie field of STUN overlaps the sequence number field of ESP, so a possibility still remains that the sequence number of ESP coincides with 0x2112A442. In this additional test, the validity of the fingerprint attribute of the STUN message MUST be checked. If there is a valid fingerprint in the message, it is judged as a STUN packet; otherwise, it is an ESP packet.

The above logic is expressed as follows.

```
if SPI-field-is-all-zeros
     { packet is IKE }
  else
     {
     if bits-32-through-63 == stun-magic-cookie-value and
        bits-0-through-1 == 0 and
        bits-2-through-15 == a STUN message type and
        bits-16-through-31 == length of this UDP packet
        {
         fingerprint_found == parse_for_stun_fingerprint();
         if fingerprint_found == 1
            { packet is STUN }
         else
            { packet is ESP }
        }
     else
        { packet is ESP }
     }
```

6. Examples

6.1. Example of SDP Offer and Answer Exchange without IPsec NAT-Traversal

If IPsec NAT-Traversal is not necessary, SDP negotiation to set up IKE is quite simple. Examples of SDP exchange are as follows.

(Note: Due to RFC formatting conventions, this document splits SDP across lines whose content would exceed 72 characters. A backslash character marks where this line folding has taken place. This backslash and its trailing CRLF and whitespace would not appear in actual SDP content.)

offer SDP

```
...
m=application 500 udp ike-esp
c=IN IP4 192.0.2.10
a=ike-setup:active
a=fingerprint:SHA-1 \
4A:AD:B9:B1:3F:82:18:3B:54:02:12:DF:3E:5D:49:6B:19:E5:7C:AB
...
```

answer SDP

```
...
m=application 500 udp ike-esp
c=IN IP4 192.0.2.20
a=ike-setup:passive
a=fingerprint:SHA-1 \
D2:9F:6F:1E:CD:D3:09:E8:70:65:1A:51:7C:9D:30:4F:21:E4:4A:8E
...
```

Figure 2: SDP Example When Offerer Is an IKE Initiator

```
offer SDP
   ...
   m=application 500 udp ike-esp
   c=IN IP4 192.0.2.10
   a=ike-setup:passive
   a=fingerprint:SHA-1 \
   4A:AD:B9:B1:3F:82:18:3B:54:02:12:DF:3E:5D:49:6B:19:E5:7C:AB
   ...

answer SDP
   ...
   m=application 500 udp ike-esp
   c=IN IP4 192.0.2.20
   a=ike-setup:active
   a=fingerprint:SHA-1 \
   D2:9F:6F:1E:CD:D3:09:E8:70:65:1A:51:7C:9D:30:4F:21:E4:4A:8E
   ...
```

Figure 3: SDP Example When Offerer Is an IKE Responder

6.2. Example of SDP Offer and Answer Exchange with IPsec NAT-Traversal

We consider the following scenario here.

```
                    +----------------------+
                    |                      |
                    |       Internet       |
                    |                      |
                    +----------------------+
                      |                  |
                      |                  |(192.0.2.20:45664)
                      |             +----------+
                      |             |   NAT    |
                      |             +----------+
                      |                  |
   (192.0.2.10:4500)|                  |(192.0.2.100:4500)
                 +----------+       +-----------+
                 | offerer  |       | answerer  |
                 +----------+       +-----------+
```

Figure 4: NAT-Traversal Scenario

As shown above, an offerer is on the Internet, but an answerer is behind the NAT. The offerer cannot initiate an IKE session unless the answerer prepares a global routable transport address that accepts IKE packets. In this case, the following offer/answer exchange will take place.

```
   offer SDP
      ...
      a=ice-pwd:YH75Fviy6338Vbrhrlp8Yh
      a=ice-ufrag:9uB6
      m=application 4500 udp ike-esp-udpencap
      c=IN IP4 192.0.2.10
      a=ike-setup:active
      a=fingerprint:SHA-1 \
      4A:AD:B9:B1:3F:82:18:3B:54:02:12:DF:3E:5D:49:6B:19:E5:7C:AB
      a=candidate:1 1 udp 2130706431 192.0.2.10 4500 typ host
      ...

   answer SDP
      ...
      a=ice-pwd:asd88fgpdd777uzjYhagZg
      a=ice-ufrag:8hhY
      m=application 45664 udp ike-esp-udpencap
      c=IN IP4 192.0.2.20
      a=ike-setup:passive
      a=fingerprint:SHA-1 \
      D2:9F:6F:1E:CD:D3:09:E8:70:65:1A:51:7C:9D:30:4F:21:E4:4A:8E
      a=candidate:1 1 udp 2130706431 192.0.2.100 4500 typ host
      a=candidate:2 1 udp 1694498815 192.0.2.20 45664 typ srflx \
      raddr 192.0.2.100 rport 4500
      ...
```

Figure 5: SDP Example with IPsec NAT-Traversal

7. Application to IKE

After the fingerprints of both parties are securely shared over the SDP exchange, the IKE initiator MAY start the IKE session with the other party. To follow this specification, a digital signature MUST be chosen as an authentication method in IKE phase 1. In this process, a certificate whose hash value matches the fingerprint exchanged over SDP MUST be used. If the certificate used in IKE does not match the original fingerprint, the endpoint MUST terminate the IKE session by detecting an authentication failure.

In addition, each party MUST present a certificate and be authenticated by each other.

The example described in Section 3 is for tunnel mode IPsec used for remote access, but the mode of negotiated IPsec is not limited to tunnel mode. For example, IKE can negotiate transport mode IPsec to encrypt multiple media sessions between two parties with only a pair of IPsec security associations. The only thing for which the SDP offer/answer model is responsible is to exchange the fingerprints of certificates used for IKE; therefore, the SDP offer/answer is not responsible for setting the security policy.

8. Specifications Assuming Prior Relationship between Two Nodes

This section describes the specification for the limited cases in which certificates signed by trusted third parties or pre-shared keys between endpoints can be used for authentication in IKE. Because the endpoints already have a prior relationship in this case, they use SIP servers for only name resolution and authorization. However, even in this case, the integrity of the SDP description MUST be assured.

8.1. Certificates Signed by Trusted Third Party

The protocol overview in this case is the same as in Section 3. The SDP offer/answer procedure is also the same as in Sections 5 and 6. Both endpoints have a prior relationship through the trusted third parties, and SIP servers are used for name resolution and authorization of session initiation. Even so, they MAY exchange fingerprints in the SDP because one device can have several certificates and it would be necessary to specify in advance which certificate will be used for the following IKE authentication. This process also ensures that the certificate offered in the IKE process is the same as that owned by the peer that has been authorized at the SIP/SDP layer. By this process, authorization in SIP and authentication in IKE become consistent with each other.

8.2. Configured Pre-Shared Key

If a pre-shared key for IKE authentication is installed in both endpoints in advance, they need not exchange the fingerprints of their certificates. However, they may still need to specify which pre-shared key they will use in the following IKE authentication in SDP because they may have several pre-shared keys. Therefore, a new attribute, "psk-fingerprint", is defined to exchange the fingerprint of a pre-shared key over SDP. This attribute also has the role of making authorization in SIP consistent with authentication in IKE. Attribute "psk-fingerprint" is applied to pre-shared keys as the "fingerprint" defined in [RFC4572] is applied to certificates. The following is the ABNF of the "psk-fingerprint" attribute. The use of "psk-fingerprint" is OPTIONAL.

```
attribute                 =/ psk-fingerprint-attribute

psk-fingerprint-attribute = "psk-fingerprint" ":" hash-func SP
                             psk-fingerprint
```

```
hash-func                  = "sha-1" / "sha-224" / "sha-256" /
                             "sha-384" / "sha-512" / token
                             ; Additional hash functions can only come
                             ; from updates to RFC 3279

psk-fingerprint            = 2UHEX *(":" 2UHEX)
                             ; Each byte in upper-case hex, separated
                             ; by colons.

UHEX                       = DIGIT / %x41-46 ; A-F uppercase
```

An example of SDP negotiation for IKE with pre-shared key authentication without IPsec NAT-Traversal is as follows.

```
offer SDP
   ...
   m=application 500 udp ike-esp
   c=IN IP4 192.0.2.10
   a=ike-setup:active
   a=psk-fingerprint:SHA-1 \
   12:DF:3E:5D:49:6B:19:E5:7C:AB:4A:AD:B9:B1:3F:82:18:3B:54:02
   ...

answer SDP
   ...
   m=application 500 udp ike-esp
   c=IN IP4 192.0.2.20
   a=ike-setup:passive
   a=psk-fingerprint:SHA-1 \
   12:DF:3E:5D:49:6B:19:E5:7C:AB:4A:AD:B9:B1:3F:82:18:3B:54:02
   ...
```

Figure 6: SDP Example of IKE with Pre-Shared Key Authentication

9. Security Considerations

This entire document concerns security, but the security considerations applicable to SDP in general are described in the SDP specification [RFC4566]. The security issues that should be considered in using comedia-tls are described in Section 7 in its specification [RFC4572]. This section mainly describes the security considerations specific to the negotiation of IKE using comedia-tls.

Offering IKE in SDP (or agreeing to one in the SDP offer/answer model) does not create an obligation for an endpoint to accept any IKE session with the given fingerprint. However, the endpoint must engage in the standard IKE negotiation procedure to ensure that the chosen IPsec security associations (including encryption and authentication algorithms) meet the security requirements of the higher-level application. When IKE has finished negotiating, the decision to conclude IKE and establish an IPsec security association with the remote peer is entirely the decision of each endpoint. This procedure is similar to how VPNs are typically established in the absence of SIP.

In the general authentication process in IKE, subject DN or subjectAltName is recognized as the identity of the remote party. However, by using SIP identity and SIP-connected identity mechanisms in this spec, certificates are used simply as carriers for the public keys of the peers and there is no need for the information about who is the signer of the certificate and who is indicated by subject DN.

In this document, the purpose of using IKE is to launch the IPsec SA; it is not for the security mechanism of RTP and RTCP [RFC3550] packets. In fact, this mechanism cannot provide end-to-end security inside the VPN as long as the VPN uses tunnel mode IPsec. Therefore, other security methods such as the Secure Real-time Transport Protocol (SRTP) [RFC3711] must be used to secure the packets.

When using the specification defined in this document, it needs to be considered that under the following circumstances, security based on SIP authentication provided by SIP proxy may be breached.

- o If a legitimate user's terminal is used by another person, it may be able to establish a VPN with the legitimate identity information. This issue also applies to the general VPN cases based on the shared secret key. Furthermore, in SIP we have a similar problem when file transfer, IM, or comedia-tls where non-voice/video is used as a means of communication.

- o If a malicious user hijacks the proxy, he or she can use whatever credential is on the Access Control List (ACL) to gain access to the home network.

For countermeasures to these issues, it is recommended to use unique information such as a password that only a legitimate user knows for VPN establishment. Validating the originating user by voice or video before establishing VPN would be another method.

10. IANA Considerations

IANA has registered the following new SDP attributes and media formats.

```
Attribute name:          ike-setup
Long form name:          IKE setup extensions
Type of attribute:       Session-level and media-level
Subject to charset:      No
Purpose:                 Attribute to indicate initiator and responder
                         of IKE-based media session
Appropriate values:      See Section 4 of RFC 6193
Contact name:            Makoto Saito, ma.saito@nttv6.jp

Media format name:       ike-esp
Long form name:          IKE followed by IPsec ESP
Associated media:        application
Associated proto:        udp
Subject to charset:      No
Purpose:                 Media format that indicates IKE and IPsec ESP
                         as a VPN session
Reference to the spec:   See Section 5 of RFC 6193
Contact name:            Makoto Saito, ma.saito@nttv6.jp


Media format name:       ike-esp-udpencap
Long form name:          IKE followed by IPsec ESP or UDP encapsulated
                         IPsec ESP
Associated media:        application
Associated proto:        udp
Subject to charset:      No
Purpose:                 Media format that indicates IKE that
                         supports NAT-Traversal and IPsec ESP or UDP
                         encapsulation of IPsec ESP packets as a VPN
                         session
Reference to the spec:   See Section 5 of RFC 6193
Contact name:            Makoto Saito, ma.saito@nttv6.jp


Attribute name:          psk-fingerprint
Long form name:          Fingerprint of pre-shared key extensions
Type of attribute:       Session-level and media-level
Subject to charset:      No
Purpose:                 Attribute to indicate a pre-shared key that
                         will be used in the following media session
Appropriate values:      See Section 8.2. of RFC 6193
Contact name:            Makoto Saito, ma.saito@nttv6.jp
```

11. Acknowledgments

We would like to thank Remi Denis-Courmont, Dale Worley, Richard Barnes, David Hancock, Stuart Hoggan, Jean-Francois Mule, Gonzalo Camarillo, and Robert Sparks for providing comments and suggestions contributing to this document. Eric Rescorla especially gave insightful comments from a security point of view. Shintaro Mizuno and Shida Schubert also contributed a lot of effort to improving this document.

12. References

12.1. Normative References

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

[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.

[RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.

[RFC3947] Kivinen, T., Swander, B., Huttunen, A., and V. Volpe, "Negotiation of NAT-Traversal in the IKE", RFC 3947, January 2005.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, January 2005.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, December 2005.

[RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)", RFC 4303, December 2005.

[RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.

[RFC4572] Lennox, J., "Connection-Oriented Media Transport over the Transport Layer Security (TLS) Protocol in the Session Description Protocol (SDP)", RFC 4572, July 2006.

[RFC5245] Rosenberg, J., "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols", RFC 5245, April 2010.

[RFC5389] Rosenberg, J., Mahy, R., Matthews, P., and D. Wing, "Session Traversal Utilities for NAT (STUN)", RFC 5389, October 2008.

[RFC5996] Kaufman, C., Hoffman, P., Nir, Y., and P. Eronen, "Internet Key Exchange Protocol Version 2 (IKEv2)", RFC 5996, September 2010.

12.2. Informative References

[RFC4474-Concerns] Rosenberg, J., "Concerns around the Applicability of RFC 4474", Work in Progress, February 2008.

[RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.

[RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.

[RFC4028] Donovan, S. and J. Rosenberg, "Session Timers in the Session Initiation Protocol (SIP)", RFC 4028, April 2005.

[RFC4145] Yon, D. and G. Camarillo, "TCP-Based Media Transport in the Session Description Protocol (SDP)", RFC 4145, September 2005.

[RFC4474] Peterson, J. and C. Jennings, "Enhancements for Authenticated Identity Management in the Session Initiation Protocol (SIP)", RFC 4474, August 2006.

[RFC4916] Elwell, J., "Connected Identity in the Session Initiation Protocol (SIP)", RFC 4916, June 2007.

[RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, August 2008.

[RFC5763] Fischl, J., Tschofenig, H., and E. Rescorla, "Framework for Establishing a Secure Real-time Transport Protocol (SRTP) Security Context Using Datagram Transport Layer Security (DTLS)", RFC 5763, May 2010.

Authors' Addresses

Makoto Saito
NTT Communications
1-1-6 Uchisaiwai-Cho, Chiyoda-ku
Tokyo 100-8019
Japan

EMail: ma.saito@nttv6.jp

Dan Wing
Cisco Systems
170 West Tasman Drive
San Jose, CA 95134
United States

EMail: dwing@cisco.com

Masashi Toyama
NTT Corporation
9-11 Midori-Cho 3-Chome, Musashino-Shi
Tokyo 180-8585
Japan

EMail: toyama.masashi@lab.ntt.co.jp