MMUSIC
Internet-Draft
Expires: April 9, 2007

J. Rosenberg
Cisco Systems

October 6, 2006

Interactive Connectivity Establishment (ICE): A Methodology for Network Address Translator (NAT) Traversal for Offer/Answer Protocols

draft-ietf-mmusic-ice-11

Status of this Memo

Copyright Notice

Abstract

This document describes a protocol for Network Address Translator (NAT) traversal for multimedia session signaling protocols based on the offer/answer model, such as the Session Initiation Protocol (SIP). This protocol is called Interactive Connectivity Establishment (ICE). ICE makes use of the Simple Traversal Underneath NAT (STUN) protocol, applying its binding discovery and relay usages, in addition to defining a new usage for checking connectivity between peers.

Table of Contents

1. Introduction 4
2. Overview of ICE 4
2.1. Gathering Candidate Addresses 6
2.2. Connectivity Checks 8
2.3. Sorting Candidates 10
2.4. Frozen Candidates 10
2.5. Security for Checks 11
3. Terminology 11
4. Sending the Initial Offer 13
4.1. Gathering Candidates 13
4.2. Prioritizing Candidates 16
4.3. Choosing In-Use Candidates 18
4.4. Encoding the SDP 18
5. Receiving the Initial Offer 20
5.1. Verifying ICE Support 20
5.2. Gathering Candidates 20
5.3. Prioritizing Candidates 21
5.4. Choosing In Use Candidates 21
5.5. Encoding the SDP 21
5.6. Forming the Check Lists 21
5.7. Performing Periodic Checks 23
6. Receipt of the Initial Answer 24
6.1. Verifying ICE Support 24
6.2. Forming the Check List 24
6.3. Performing Periodic Checks 24
7. Connectivity Checks 24
7.1. Applicability 24
7.2. Client Discovery of Server 25
7.3. Server Determination of Usage 25
7.4. New Requests or Indications 25
7.5. New Attributes 25
7.6. New Error Response Codes 25
7.7. Client Procedures 25
7.7.1. Sending the Request 25
7.7.2. Processing the Response 26
7.8. Server Procedures 27
7.9. Security Considerations for Connectivity Check 29
8. Completing the ICE Checks 29
9. Subsequent Offer/Answer Exchanges 30
9.1. Generating the Offer 30
9.2. Receiving the Offer and Generating an Answer 31
9.3. Updating the Check and Valid Lists 32
10. Keepalives 33
11. Media Handling 34
11.1. Sending Media 34
11.2. Receiving Media 35

12. Usage with SIP 35
12.1. Latency Guidelines 35
12.2. Interactions with Forking 37
12.3. Interactions with Preconditions 37
12.4. Interactions with Third Party Call Control 38
13. Grammar 38
14. Example 40
15. Security Considerations 46
15.1. Attacks on Connectivity Checks 46
15.2. Attacks on Address Gathering 49
15.3. Attacks on the Offer/Answer Exchanges 49
15.4. Insider Attacks 50
15.4.1. The Voice Hammer Attack 50
15.4.2. STUN Amplification Attack 50
16. IANA Considerations 51
16.1. candidate Attribute 51
16.2. remote-candidates Attribute 51
16.3. ice-pwd Attribute 52
16.4. ice-ufrag Attribute 52
17. IAB Considerations 53
17.1. Problem Definition 53
17.2. Exit Strategy 53
17.3. Brittleness Introduced by ICE 54
17.4. Requirements for a Long Term Solution 55
17.5. Issues with Existing NAPT Boxes 55
18. Acknowledgements 56
19. References 56
19.1. Normative References 56
19.2. Informative References 57
Appendix A. Design Motivations 58
A.1. Applicability to Gateways and Servers 59
A.2. Pacing of STUN Transactions 60
A.3. Candidates with Multiple Bases 61
A.4. Purpose of the Translation 63
A.5. Importance of the STUN Username 63
A.6. The Candidate Pair Sequence Number Formula 64
A.7. The Frozen State 65
A.8. The remote-candidates attribute 65
A.9. Why are Keepalives Needed? 66
A.10. Why Prefer Peer Reflexive Candidates? 67
A.11. Why Can't Offerers Send Media When a Pair Validates . . . 67
Author's Address 69
Intellectual Property and Copyright Statements 70

1. Introduction

RFC 3264 [4] defines a two-phase exchange of Session Description Protocol (SDP) messages [10] for the purposes of establishment of multimedia sessions. This offer/answer mechanism is used by protocols such as the Session Initiation Protocol (SIP) [3].

Protocols using offer/answer are difficult to operate through Network Address Translators (NAT). Because their purpose is to establish a flow of media packets, they tend to carry IP addresses within their messages, which is known to be problematic through NAT [14]. The protocols also seek to create a media flow directly between participants, so that there is no application layer intermediary between them. This is done to reduce media latency, decrease packet loss, and reduce the operational costs of deploying the application. However, this is difficult to accomplish through NAT. A full treatment of the reasons for this is beyond the scope of this specification.

Numerous solutions have been proposed for allowing these protocols to operate through NAT. These include Application Layer Gateways (ALGs), the Middlebox Control Protocol [15], Simple Traversal Underneath NAT (STUN) [13] and its revision [11], the STUN Relay Usage [12], and Realm Specific IP [17] [18] along with session description extensions needed to make them work, such as the Session Description Protocol (SDP) [10] attribute for the Real Time Control Protocol (RTCP) [2]. Unfortunately, these techniques all have pros and cons which make each one optimal in some network topologies, but a poor choice in others. The result is that administrators and implementors are making assumptions about the topologies of the networks in which their solutions will be deployed. This introduces complexity and brittleness into the system. What is needed is a single solution which is flexible enough to work well in all situations.

This specification provides that solution for media streams established by signaling protocols based on the offer-answer model. It is called Interactive Connectivity Establishment, or ICE. ICE makes use of STUN and its relay extension, commonly called TURN, but uses them in a specific methodology which avoids many of the pitfalls of using any one alone.

2. Overview of ICE

In a typical ICE deployment, we have two endpoints (known as agents in RFC 3264 terminology) which want to communicate. They are able to communicate indirectly via some signaling system such as SIP, by

which they can perform an offer/answer exchange of SDP [4] messages. Note that ICE is not intended for NAT traversal for SIP, which is assumed to be provided via some other mechanism [31]. At the beginning of the ICE process, the agents are ignorant of their own topologies. In particular, they might or might not be behind a NAT (or multiple tiers of NATs). ICE allows the agents to discover enough information about their topologies to find a path or paths by which they can communicate.

Figure Figure 1 shows a typical environment for ICE deployment. The two endpoints are labelled L and R (for left and right, which helps visualize call flows). Both L and R are behind NATs -- though as mentioned before, they don't know that. The type of NAT and its properties are also unknown. Agents L and R are capable of engaging in an offer/answer exchange by which they can exchange SDP messages, whose purpose is to set up a media session between L and R. Typically, this exchange will occur through a SIP server.

In addition to the agents, a SIP server and NATs, ICE is typically used in concert with STUN servers in the network. Each agent can have its own STUN server, or they can be the same.

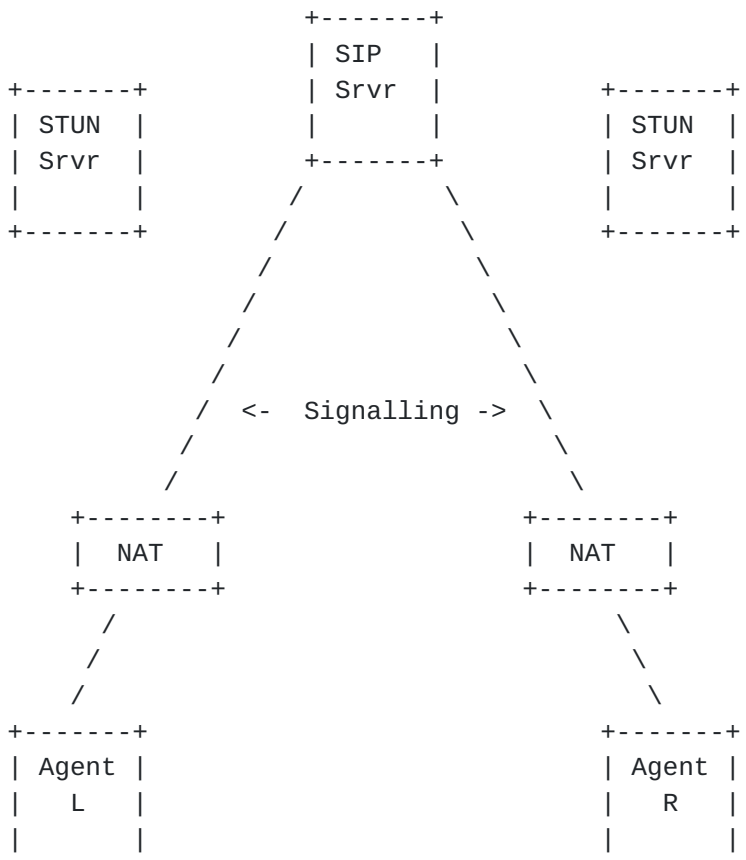


Figure 1

The basic idea behind ICE is as follows: each agent has a variety of candidate transport addresses it could use to communicate with the other agent. These might include:

- It's directly attached network interface (or interfaces in the case of a multihomed machine
- A translated address on the public side of a NAT (a "server reflexive" address)
- The address of a media relay the agent is using.

Potentially, any of L's candidate transport addresses can be used to communicate with any of R's transport addresses. In practice, however, many combinations will not work. For instance, if L and R are both behind NATs then their directly interface addresses are unlikely to be able to communicate directly (this is why ICE is needed, after all!). The purpose of ICE is to discover which pairs of addresses will work. The way that ICE does this is to systematically try all possible pairs (in a carefully sorted order) until it finds one or more that works.

2.1. Gathering Candidate Addresses

In order to execute ICE, an agent has to identify all of its address candidates. Naturally, one viable candidate is one obtained directly from a local interface the client has towards the network. Such a candidate is called a HOST CANDIDATE. The local interface could be one on a local layer 2 network technology, such as ethernet or WiFi, or it could be one that is obtained through a tunnel mechanism, such as a Virtual Private Network (VPN) or Mobile IP (MIP). In all cases, these appear to the agent as a local interface from which ports (and thus a candidate) can be allocated.

If an agent is multihomed, it can obtain a candidate from each interface. Depending on the location of the peer on the IP network relative to the agent, the agent may be reachable by the peer through one of those interfaces, or through another. Consider, for example, an agent which has a local interface to a private net 10 network, and also to the public Internet. A candidate from the net10 interface will be directly reachable when communicating with a peer on the same private net 10 network, while a candidate from the public interface will be directly reachable when communicating with a peer on the public Internet. Rather than trying to guess which interface will

work prior to sending an offer, the offering agent includes both candidates in its offer.

Once the agent has obtained host candidates, it uses STUN to obtain additional candidates. These come in two flavors: translated addresses on the public side of a NAT (SERVER REFLEXIVE CANDIDATES) and addresses of media relays (RELAYED CANDIDATES). The relationship of these candidates to the host candidate is shown in Figure 2. Both types of candidates are discovered using STUN.

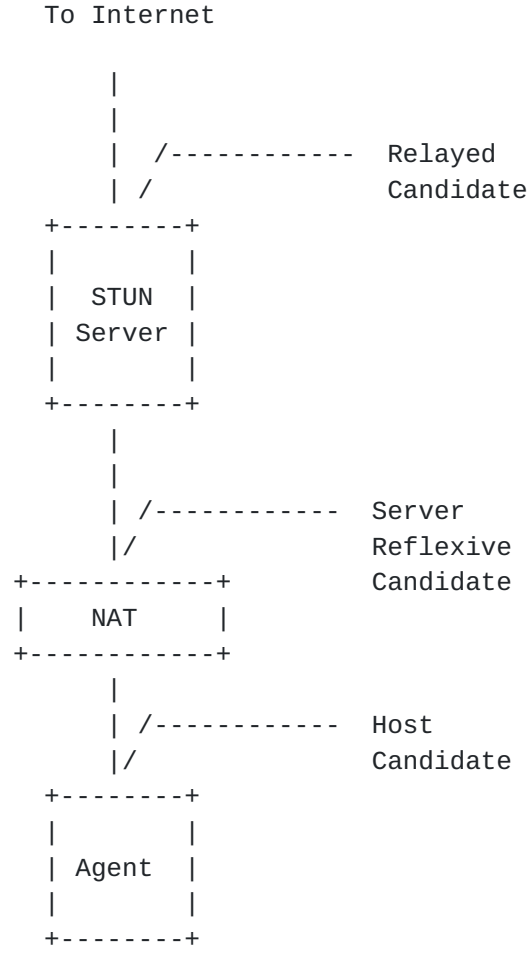


Figure 2

To find a server reflexive candidate, the agent sends a STUN Binding Request, using the Binding Discovery Usage [11] from each host candidate, to its STUN server. (It is assumed that the address of the STUN server is configured, or learned in some way.) When the agents sends the Binding Request, the NAT (assuming there is one) will allocate a binding, mapping this server reflexive candidate to the host candidate. Outgoing packets sent from the host candidate

will be translated by the NAT to the server reflexive candidate. Incoming packets sent to the server relexive candidate will be translated by the NAT to the host candidate and forwarded to the agent. We call the host candidate associated with a given server reflexive candidate the BASE.

Note

"Base" refers to the address you'd send from for a particular candidate. Thus, as a degenerate case host candidates also have a base, but it's the same as the host candidate.

When there are multiple NATs between the agent and the STUN server, the STUN request will create a binding on each NAT, but only the outermost server reflexive candidate will be discovered by the agent. If the agent is not behind a NAT, then the base candidate will be the same as the server reflexive candidate and the server reflexive candidate can be ignored.

The final type of candidate is a RELAYED candidate. The STUN Relay Usage [12] allows a STUN server to act as a media relay, forwarding traffic between L and R. In order to send traffic to L, R sends traffic to the media relay which forwards it to L and vice versa. The same thing happens in the other direction.

Traffic from L to R has its addresses rewritten twice: first by the NAT and second by the STUN relay server. Thus, the address that R knows about and the one that it wants to send to is the one on the STUN relay server. This address is the final kind of candidate, which we call a RELAYED CANDIDATE.

2.2. Connectivity Checks

Once L has gathered all of its candidates, it orders them highest to lowest priority and sends them to R over the signalling channel. The candidates are carried in attributes in the SDP offer. When R receives the offer, it performs the same gathering process and responds with its own list of candidates. At the end of this process, each agent has a complete list of both its candidates and its peer's candidates and is ready to perform connectivity checks by pairing up the candidates to see which pair works.

The basic principle of the connectivity checks is simple:

1. Sort the candidate pairs in priority order.

2. Send checks on each candidate pair in priority order.

3. Acknowledge checks received from the other agent.

A complete connectivity check for a single candidate pair is a simple 4-message handshake:

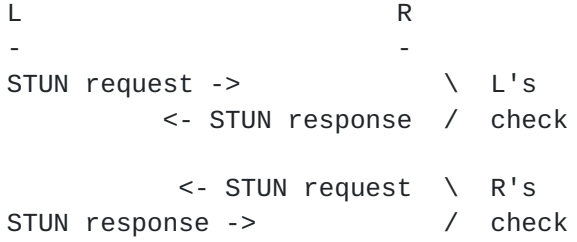


Figure 3

As an optimization, as soon as R gets L's check message he immediately sends his own check message to L on the same candidate pair. This accelerates the process of finding a valid candidate.

At the end of this handshake, both L and R know that they can send (and receive) messages end-to-end in both directions. Note that as soon as R receives L's STUN response it knows that the R->L path works and it can start sending media on that path right away, as shown below. This allows for 'early media' to flow as fast as possible:

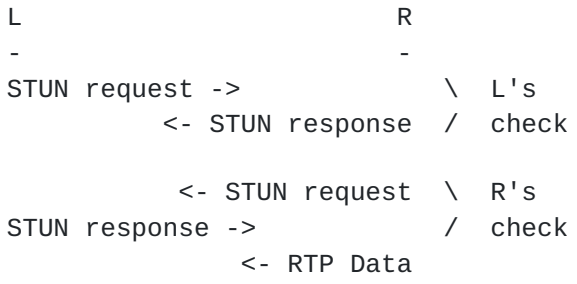


Figure 4

Once any connectivity check for a candidate for a given media component succeeds, ICE uses that candidate and immediately abandons all other connectivity checks for that component. Note that due to race conditions and packet loss, this may mean that the "best" candidate isn't selected, but it does guarantee the selection of a candidate that works, and because of the sorting process it will generally be one of the most preferred ones.

2.3. Sorting Candidates

Because the algorithm above searches all candidate pairs, if a working pair exists it will eventually find it no matter what order the candidates are tried in. In order to produce faster (and better) results, the candidates are sorted in a specified order. The algorithm is described in Section 4.2 but follows two general principles:

- Each agent gives its candidates a numeric priority which is sent along with the candidate to the peer
- The local and remote priorities are combined so that each agent has the same ordering for the candidate pairs.

The second property is important for getting ICE to work when there are NATs in front of A and B. Frequently, NATs will not allow packets in from a host until the agent behind the NAT has sent a packet towards that host. Consequently, ICE checks in each direction will not succeed until both sides have sent a check through their respective NATs.

In general the priority algorithm is designed so that candidates of similar type get similar priorities and so that more direct routes are favored over indirect ones. Within those guidelines, however, agents have a fair amount of discretion about how to tune their algorithms.

2.4. Frozen Candidates

The previous description only addresses the case where the agents wish to establish a single media component--i.e., a single flow with a single host-port quartet. However, in many cases (in particular RTP and RTCP) the agents actually need to establish connectivity for more than one flow.

The naive way to attack this problem would be to simply do independent ICE exchanges for each media component. This is obviously inefficient because the network properties are likely to be very similar for each component (especially because RTP and RTCP are typically run on adjacent ports). Thus, it should be possible to leverage information from one media component in order to determine the best candidates for another. ICE does this with a mechanism called "frozen candidates."

The basic principle behind frozen candidates is that initially only the candidates for a single media component are tested. The other media components are marked "frozen". When the connectivity checks

for the first component succeed, the corresponding candidates for the other components are unfrozen and checked immediately. This avoids repeated checking of components which are superficially more attractive but in fact are likely to fail.

While we've described "frozen" here as a separate mechanism for expository purposes, in fact it is an integral part of ICE and the the ICE prioritization algorithm automatically ensures that the right candidates are unfrozen and checked in the right order.

2.5. Security for Checks

Because ICE is used to discover which addresses can be used to send media between two agents, it is important to ensure that the process cannot be hijacked to send media to the wrong location. Each STUN connectivity check is covered by a message authentication code (MAC) computed using a key exchanged in the signalling channel. This MAC provides message integrity and data origin authentication, thus stopping an attacker from forging or modifying connectivity check messages. The MAC also aids in disambiguating ICE exchanges from forked calls.

3. Terminology

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 [1].

This specification makes use of the following terminology:

Agent: As defined in RFC 3264, an agent is the protocol implementation involved in the offer/answer exchange. There are two agents involved in an offer/answer exchange.

Peer: From the perspective of one of the agents in a session, its peer is the other agent. Specifically, from the perspective of the offerer, the peer is the answerer. From the perspective of the answerer, the peer is the offerer.

Transport Address: The combination of an IP address and port.

Candidate: A transport address that is to be tested by ICE procedures in order to determine its suitability for usage for receipt of media.

Component: A component is a single transport address that is used to support a media stream. For media streams based on RTP, there are two components per media stream - one for RTP, and one for RTCP.

Host Candidate: A candidate obtained by binding to a specific port from an interface on the host. This includes both physical interfaces and logical ones, such as ones obtained through Virtual Private Networks (VPNs) and Realm Specific IP (RSIP) [17] (which lives at the operating system level).

Server Reflexive Candidate: A candidate obtained by sending a STUN request from a host candidate to a STUN server, distinct from the peer, whose address is configured or learned by the client prior to an offer/answer exchange.

Peer Reflexive Candidate: A candidate obtained by sending a STUN request from a host candidate to the STUN server running on a peer's candidate.

Relayed Candidate: A candidate obtained by sending a STUN Allocate request from a host candidate to a STUN server. The relayed candidate is resident on the STUN server, and the STUN server relays packets back towards the agent.

Translation: The translation of a relayed candidate is the transport address that the relay will forward a packet to, when one is received at the relayed candidate. For relayed candidates learned through the STUN Allocate request, the translation of the relayed candidate is the server reflexive candidate returned by the Allocate response.

Base: The base of a server reflexive candidate is the host candidate from which it was derived. A host candidate is also said to have a base, equal to that candidate itself. Similarly, the base of a relayed candidate is that candidate itself.

Foundation: Each candidate has a foundation, which is an identifier that is distinct for two candidates that have different types, different interface IP addresses for their base, and different IP addresses for their STUN servers. Two candidates have the same foundation when they are of the same type, their bases have the same IP address, and, for server reflexive or relayed candidates, they come from the same STUN server. Foundations are used to correlate candidates, so that when one candidate is found to be valid, candidates sharing the same foundation can be tested next, as they are likely to also be valid.

Local Candidate: A candidate that an agent has obtained and included in an offer or answer it sent.

Remote Candidate: A candidate that an agent received in an offer or answer from its peer.

In-Use Candidate: A candidate is in-use when it appears in the m/c-line of an active media stream.

Candidate Pair: A pairing containing a local candidate and a remote candidate.

Check: A candidate pair where the local candidate is a transport address from which an agent can send a STUN connectivity check.

Check List: An ordered set of STUN checks that an agent is to generate towards a peer.

Periodic Check: A connectivity check generated by an agent as a consequence of a timer that fires periodically, instructing it to send a check.

Triggered Check: A connectivity check generated as a consequence of the receipt of a connectivity check from the peer.

Valid List: An ordered set of candidate pairs that have been validated by a successful STUN transaction.

4. Sending the Initial Offer

In order to send the initial offer in an offer/answer exchange, an agent must gather candidates, priorize them, choose ones for inclusion in the m/c-line, and then formulate and send the SDP. Each of these steps is described in the subsections below.

4.1. Gathering Candidates

An agent gathers candidates when it believes that communications is imminent. An offerer can do this based on a user interface cue, or based on an explicit request to initiate a session. Every candidate is a transport address. It also has a type and a base. Three types are defined and gathered by this specification - host candidates, server reflexive candidates, and relayed candidates. The base of a candidate is the candidate that an agent must send from when using that candidate.

The first step is to gather host candidates. Host candidates are

obtained by binding to ports (typically ephemeral) on an interface (physical or virtual, including VPN interfaces) on the host. The process for gathering host candidates depends on the transport protocol. Procedures are specified here for UDP.

For each UDP media stream the agent wishes to use, the agent SHOULD obtain a candidate for each component of the media stream on each interface that the host has. It obtains each candidate by binding to a UDP port on the specific interface. A host candidate (and indeed every candidate) is always associated with a specific component for which it is a candidate. Each component has an ID assigned to it, called the component ID. For RTP-based media streams, the RTP itself has a component ID of 1, and RTCP a component ID of 2. If an agent is using RTCP it MUST obtain a candidate for it. If an agent is using both RTP and RTCP, it would end up with 2*K host candidates if an agent has K interfaces.

The base for each host candidate is set to the candidate itself.

Once the agent has obtained host candidates, it obtains server reflexive and relayed candidates. The process for gathering server reflexive and relayed candidates depends on the transport protocol. Procedures are specified here for UDP.

Agents which serve end users directly, such softphones, hardphones, terminal adapters and so on, SHOULD obtain relayed candidates and MUST obtain server reflexive candidates. The requirement to obtain relayed candidates is at SHOULD strength to allow for provider variation. If they are not used, it is RECOMMENDED that it be implemented and just disabled through configuration, so that it can re-enabled through configuration if conditions change in the future. Agents which represent network servers under the control of a service provider, such as gateways to the telephone network, media servers, or conferencing servers that are targeted at deployment only in networks with public IP addresses MAY skip obtaining server reflexive and relayed candidates.

The agent next pairs each host candidate with the STUN server with which it is configured or has discovered by some means. This specification only considers usage of a single STUN server. Every Ta seconds, the agent chooses another such pair (the order is inconsequential), and sends a STUN request to the server from that host candidate. If the agent is using both relayed and server reflexive candidates, this request MUST be a STUN Allocate request from the relay usage [12]. If the agent is using only server reflexive candidates, the request MUST be a STUN Binding request using the binding discovery usage [11].

The value of Ta SHOULD be configurable, and SHOULD have a default of 50ms. Note that this pacing applies only to starting STUN transactions with source and destination transport addresses (i.e., the host candidate and STUN server respectively) for which a STUN transaction has not previously been sent. Consequently, retransmissions of a STUN request are governed entirely by the retransmission rules defined in [11]. Similarly, retries of a request due to recoverable errors (such as an authentication challenge) happen immediately and are not paced by timer Ta. Because of this pacing, it will take a certain amount of time to obtain all of the server reflexive and relayed candidates. Implementations should be aware of the time required to do this, and if the application requires a time budget, limit the amount of candidates which are gathered.

An Allocate Response will provide the client with a server reflexive candidate (obtained from the mapped address) and a relayed candidate in the RELAY-ADDRESS attribute. A Binding Response will provide the client with a only server reflexive candidate (also obtained from the mapped address). The base of the server reflexive candidate is the host candidate from which the Allocate or Binding request was sent. The base of a relayed candidate is that candidate itself. A server reflexive candidate obtained from an Allocate response is the called the "translation" of the relayed candidate obtained from the same response. The agent will need to remember the translation for the relayed candidate, since it is placed into the SDP. If a relayed candidate is identical to a host candidate (which can happen in rare cases), the relayed candidate MUST be discarded. Proper operation of ICE depends on each base being unique.

Next, redundant candidates are eliminated. A candidate is redundant if its transport address equals another candidate, and its base equals the base of that other candidate. Note that two candidates can have the same transport address yet have different bases, and these would not be considered redundant.

Finally, each candidate is assigned a foundation. The foundation is an identifier, scoped within a session. Two candidates MUST have the same foundation ID when they are of the same type (host, relayed, server reflexive, peer reflexive or relayed), their bases have the same IP address (the ports can be different), and, for reflexive and relayed candidates, the STUN servers used to obtain them have the same IP address. Similarly, two candidates MUST have different foundations if their types are different, their bases have different IP addresses, or the STUN servers used to obtain them have different IP addresses.

4.2. Prioritizing Candidates

The prioritization process results in the assignment of a priority to each candidate. An agent does this by determining a preference for each type of candidate (server reflexive, peer reflexive, relayed and host), and, when the agent is multihomed, choosing a preference for its interfaces. These two preferences are then combined to compute the priority for a candidate. That priority MUST be computed using the following formula:

```
priority = (2^24)*(type preference) +
           (2^8)*(local preference) +
           (2^0)*(256 - component ID)
```

The type preference MUST be an integer from 0 to 126 inclusive, and represents the preference for the type of the candidate (where the types are local, server reflexive, peer reflexive and relayed). A 126 is the highest preference, and a 0 is the lowest. Setting the value to a 0 means that candidates of this type will only be used as a last resort. The type preference MUST be identical for all candidates of the same type and MUST be different for candidates of different types. The type preference for peer reflexive candidates MUST be higher than that of server reflexive candidates. Note that candidates gathered based on the procedures of Section 4.1 will never be peer reflexive candidates; candidates of these type are learned from the STUN connectivity checks performed by ICE. The component ID is the component ID for the candidate, and MUST be between 1 and 256 inclusive. The local preference MUST be an integer from 0 to 65535 inclusive. It represents a preference for the particular interface from which the candidate was obtained, in cases where an agent is multihomed. 65535 represents the highest preference, and a zero, the lowest. When there is only a single interface, this value SHOULD be set to 65535. Generally speaking, if there are multiple candidates for a particular component for a particular media stream which have the same type, the local preference MUST be unique for each one. In this specification, this only happens for multi-homed hosts.

These rules guarantee that there is a unique priority for each candidate. This priority will be used by ICE to determine the order of the connectivity checks and the relative preference for candidates. Consequently, what follows are some guidelines for selection of these values.

One criteria for selection of the type and local preference values is the use of an intermediary. That is, if media is sent to that candidate, will the media first transit an intermediate server before

being received. Relayed candidates are clearly one type of candidates that involve an intermediary. Another are host candidates obtained from a VPN interface. When media is transited through an intermediary, it can increase the latency between transmission and reception. It can increase the packet losses, because of the additional router hops that may be taken. It may increase the cost of providing service, since media will be routed in and right back out of an intermediary run by the provider. If these concerns are important, the type preference for relayed candidates can be set lower than the type preference for reflexive and host candidates. Indeed, it is RECOMMENDED that in this case, host candidates have a type preference of 126, server reflexive candidates have a type preference of 100, peer reflexive have a type prefence of 110, and relayed candidates have a type preference of zero. Furthermore, if an agent is multi-homed and has multiple interfaces, the local preference for host candidates from a VPN interface SHOULD have a priority of 0.

Another criteria for selection of preferences is IP address family. ICE works with both IPv4 and IPv6. It therefore provides a transition mechanism that allows dual-stack hosts to prefer connectivity over IPv6, but to fall back to IPv4 in case the v6 networks are disconnected (due, for example, to a failure in a 6to4 relay) [22]. It can also help with hosts that have both a native IPv6 address and a 6to4 address. In such a case, lower local preferences could be assigned to the v6 interface, followed by the 6to4 interfaces, followed by the v4 interfaces. This allows a site to obtain and begin using native v6 addresses immediately, yet still fallback to 6to4 addresses when communicating with agents in other sites that do not yet have native v6 connectivity.

Another criteria for selecting preferences is security. If a user is a telecommuter, and therefore connected to their corporate network and a local home network, they may prefer their voice traffic to be routed over the VPN in order to keep it on the corporate network when communicating within the enterprise, but use the local network when communicating with users outside of the enterprise. In such a case, a VPN interface would have a higher local preference than any other interfaces.

Another criteria for selecting preferences is topological awareness. This is most useful for candidates that make use of relays. In those cases, if an agent has preconfigured or dynamically discovered knowledge of the topological proximity of the relays to itself, it can use that to assign higher local preferences to candidates obtained from closer relays.

4.3. Choosing In-Use Candidates

A candidate is said to be "in-use" if it appears in the m/c-line of an offer or answer. When communicating with an ICE peer, being in-use implies that, should these candidates be selected by the ICE algorithm, bidirectional media can flow and the candidates can be used. If a candidate is selected by ICE but is not in-use, only unidirectional media can flow and only for a brief time; the candidate must be made in-use through an updated offer/answer exchange. When communicating with a peer that is not ICE-aware, the in-use candidates will be used exclusively for the exchange of media, as defined in normal offer/answer procedures.

An agent MUST choose a set of candidates, one for each component of each active media stream, to be in-use. A media stream is active if it does not contain the a=inactive SDP attribute.

It is RECOMMENDED that in-use candidates be chosen based on the likelihood of those candidates to work with the peer that is being contacted. Unfortunately, it is difficult to ascertain which candidates that might be. As an example, consider a user within an enterprise. To reach non-ICE capable agents within the enterprise, host candidates have to be used, since the enterprise policies may prevent communication between elements using a relay on the public network. However, when communicating to peers outside of the enterprise, relayed candidates from a publically accessible STUN server are needed.

Indeed, the difficulty in picking just one transport address that will work is the whole problem that motivated the development of this specification in the first place. As such, it is RECOMMENDED that relayed candidates be selected to be in-use. Furthermore, ICE is only truly effective when it is supported on both sides of the session. It is therefore most prudent to deploy it to close-knit communities as a whole, rather than piecemeal. In the example above, this would mean that ICE would ideally be deployed completely within the enterprise, rather than just to parts of it.

4.4. Encoding the SDP

The agent includes a single a=candidate media level attribute in the SDP for each candidate for that media stream. The a=candidate attribute contains the IP address, port and transport protocol for that candidate. A Fully Qualified Domain Name (FQDN) for a host MAY be used in place of a unicast address. In that case, when receiving an offer or answer containing an FQDN in an a=candidate attribute, the FQDN is looked up in the DNS using an A or AAAA record, and the resulting IP address is used for the remainder of ICE processing.

The candidate attribute also includes the component ID for that candidate. For media streams based on RTP, candidates for the actual RTP media MUST have a component ID of 1, and candidates for RTCP MUST have a component ID of 2. Other types of media streams which require multiple components MUST develop specifications which define the mapping of components to component IDs, and these component IDs MUST be between 1 and 256.

The candidate attribute also includes the priority, which is the value determined for the candidate as described in Section 4.2, and the foundation, which is the value determined for the candidate as described in Section 4.1. The agent SHOULD include a type for each candidate by populating the candidate-types production with the appropriate value - "host" for host candidates, "srflx" for server reflexive candidates, "prflx" for peer reflexive candidates (though these never appear in an initial offer/answer exchange), and "relay" for relayed candidates. The related address MUST NOT be included if a type was not included. If a type was included, the related address SHOULD be present for server reflexive, peer reflexive and relayed candidates. If a candidate is server or peer reflexive, the related address is equal to the base for that server or peer reflexive candidate. If the candidate is relayed, the related address is equal to the translation of the relayed address. If the candidiate is a host candidate, there is no related address and the rel-addr production MUST be omitted.

STUN connectivity checks between agents make use of a short term credential that is exchanged in the offer/answer process. The username part of this credential is formed by concatenating a username fragment from each agent, separated by a colon. Each agent also provides a password, used to compute the message integrity for requests it receives. As such, an SDP MUST contain the ice-ufrag and ice-pwd attributes, containing the username fragment and password respectively. These can be either session or media level attributes, and thus common across all candidates for all media streams, or all candidates for a particular media stream, respectively. However, if two media streams have identical ice-ufrag's, they MUST have identical ice-pwd's. The ice-ufrag and ice-pwd attributes MUST be chosen randomly at the beginning of a session. The ice-ufrag attribute MUST contain at least 24 bits of randomness, and the ice-pwd attribute MUST contain at least 128 bits of randomness. This means that the ice-ufrag attribute will be at least 4 characters long, and the ice-pwd at least 22 characters long, since the grammar for these attributes allows for 6 bits of randomness per character. The attributes MAY be longer than 4 and 21 characters respectively, of course.

The m/c-line is populated with the candidates that are in-use. For

streams based on RTP, this is done by placing the RTP candidate into the m and c lines respectively. If the agent is utilizing RTCP, it MUST encode the RTCP candidate into the m/c-line using the a=rtcp attribute as defined in RFC 3605 [2]. If RTCP is not in use, the agent MUST signal that using b=RS:0 and b=RR:0 as defined in RFC 3556 [5].

There MUST be a candidate attribute for each component of the media stream in the m/c-line.

Once an offer or answer are sent, an agent MUST be prepared to receive both STUN and media packets on each candidate. As discussed in Section 11.1, media packets can be sent to a candidate prior to its appearence in the m/c-line.

5. Receiving the Initial Offer

When an agent receives an initial offer, it will check if the offeror supports ICE, gather candidates, prioritize them, choose one for in-use, encode and send an answer, and then form the check lists and begin connectivity checks.

5.1. Verifying ICE Support

The agent will proceed with the ICE procedures defined in this specification if the following are both true:

- There is at least one a=candidate attribute for each media stream in the SDP it just received.
- For each media stream, at least one of the candidates is a match for its respective in-use component in the m/c-line.

If both of these conditions are not met, the agent MUST process the SDP based on normal RFC 3264 procedures, without using any of the ICE mechanisms described in the remainder of this specification, with the exception of Section 10, which describes keepalive procedures.

5.2. Gathering Candidates

The process for gathering candidates at the answerer is identical to the process for the offerer as described in Section 4.1. It is RECOMMENDED that this process begin immediately on receipt of the offer, prior to user acceptance of a session. Such gathering MAY even be done pre-emptively when an agent starts.

5.3. Prioritizing Candidates

The process for prioritizing candidates at the answerer is identical to the process followed by the offerer, as described in Section 4.2.

5.4. Choosing In Use Candidates

The process for selecting in-use candidates at the answerer is identical to the process followed by the offerer, as described in Section 4.3.

5.5. Encoding the SDP

The process for encoding the SDP at the answerer is identical to the process followed by the offerer, as described in Section 4.4.

5.6. Forming the Check Lists

Next, the agent forms the check lists. There is one check list per in-use media stream resulting from the offer/answer exchange. A media stream is in-use as long as its port is non-zero (which is used in RFC 3264 to reject a media stream). Each check list is a sequence of STUN connectivity checks that are performed by the agent. To form the check list for a media stream, the agent forms candidate pairs, computes a candidate pair priority, orders the pairs by priority, prunes them, and sets their states. These steps are described in this section.

First, the agent takes each of its candidates for a media stream (called local candidates) and pairs them with the candidates it received from its peer (called remote candidates) for that media stream. A local candidate is paired with a remote candidate if and only if the two candidates have the same component ID and have the same IP address version. It is possible that some of the local candidates don't get paired with a remote candidate, and some of the remote candidates don't get paired with local candidates. This can happen if one agent didn't include candidates for the all of the components for a media stream. In the case of RTP, for example, this would happen when one agent provided candidates for RTCP, and the other did not. If this happens, the number of components for that media stream is effectively reduced, and considered to be equal to the minimum across both agents of the maximum component ID provided by each agent across all components for the media stream.

Once the pairs are formed, a candidate pair priority is computed. Let O-P be the priority for the candidate provided by the offerer. Let A-P be the priority for the candidate provided by the answerer. The priority for a pair is computed as:

pair priority = 2^32*MIN(O-P,A-P) + 2*MAX(O-P,A-P) + (O-P>A-P:1?0)

Where O-P>A-P:1?0 is an expression whose value is 1 if O-P is greater than A-P, and 0 otherwise. This formula ensures a unique priority for each pair in most cases. One the priority is assigned, the agent sorts the candidate pairs in decreasing order of priority. If two pairs have identical priority, the ordering amongst them is arbitrary.

This sorted list of candidate pairs is used to determine a sequence of connectivity checks that will be performed. Each check involves sending a request from a local candidate to a remote candidate. Since an agent cannot send requests directly from a reflexive candidate, but only from its base, the agent next goes through the sorted list of candidate pairs. For each pair where the local candidate is server reflexive, the server reflexive candidate MUST be replaced by its base. Once this has been done, the agent MUST remove redundant pairs. A pair is redundant if its local and remote candidates are identical to the local and remote candidates of a pair higher up on the priority list. The result is called the check list for that media stream, and each candidate pair on it is called a check.

Each check is also said to have a foundation, which is merely the combination of the foundations of the local and remote candidates in the check.

Finally, each check in the check list is associated with a state. This state is assigned once the check list for each media stream has been computed. There are five potential values that the state can have:

Waiting: This check has not been performed, and can be performed as soon as it is the highest priority Waiting check on the check list.

In-Progress: A request has been sent for this check, but the transaction is in progress.

Succeeded: This check was already done and produced a successful result.

Failed: This check was already done and failed, either never producing any response or producing an unrecoverable failure response.

Frozen: This check hasn't been performed, and it can't yet be performed until some other check succeeds, allowing it to move into the Waiting state.

First, the agent sets all of the checks in each check list to the Frozen state. Then, it takes the first check in the check list for the first media stream (a media stream is the first media stream when it is described by the first m-line in the SDP offer and answer), and sets its state to Waiting. It then finds all of the other checks in that check list with the same component ID, but different foundations, and sets all of their states to Waiting as well. Once this is done, one of the check lists will have some number of checks in the Waiting state, and the other check lists will have all of their checks in the Frozen state. A check list with at least one check that is not Frozen is called an active check list.

5.7. Performing Periodic Checks

An agent performs two types of checks. The first type are periodic checks. These checks occur periodically for each media stream, and involve choosing the highest priority check in the Waiting state from each check list, and performing it. The other type of check is called a triggered check. This is a check that is performed on receipt of a connectivity check from the peer. This section describes how periodic checks are performed.

Once the agent has computed the check lists as described in Section 5.6, it sets a timer for each active check list. The timer fires every Ta/N seconds, where N is the number of active check lists (initially, there is only one active check list). Implementations MAY set the timer to fire less frequently than this. Ta is the same value used to pace the gathering of candidates, as described in Section 4.1. The first timer for each active check list fires immediately, so that the agent performs a connectivity check the moment the offer/answer exchange has been done, followed by the next periodic check Ta seconds later.

When the timer fires, the agent MUST find the highest priority check in that check list that is in the Waiting state. The agent then sends a STUN check from the local candidate of that check to the remote candidate of that check. The procedures for forming the STUN request for this purpose are described in Section 7.7.1. If none of the checks in that check list are in the Waiting state, but there are checks in the Frozen state, the highest priority check in the Frozen state is moved into the Waiting state, and that check is performed. When a check is performed, its state is set to In-Progress. If there are no checks in either the Waiting or Frozen state, then the timer for that check list is stopped.

Performing the connectivity check requires the agent to know the username fragment for the local and remote candidates, and the password for the remote candidate. For periodic checks, the remote username fragment and password are learned directly from the SDP received from the peer, and the local username fragment is known by the agent.

6. Receipt of the Initial Answer

This section describes the procedures that an agent follows when it receives the answer from the peer. It verifies that its peer supports ICE, forms the check list and begins performing periodic checks.

6.1. Verifying ICE Support

The offerer follows the same procedures described for the answerer in Section 5.1.

6.2. Forming the Check List

The offerer follows the same procedures described for the answerer in Section 5.6.

6.3. Performing Periodic Checks

The offerer follows the same procedures described for the answerer in Section 5.7.

7. Connectivity Checks

This section describes how connectivity checks are performed. Connectivity checks are a STUN usage, and the behaviors described here meet the guidelines for definitions of new usages as outlined in [11]

Note that all ICE implementations are required to be compliant to [11], as opposed to the older [13].

7.1. Applicability

This STUN usage provides a connectivity check between two peers participating in an offer/answer exchange. This check serves to validate a pair of candidates for usage of exchange of media. Connectivity checks also allow agents to discover reflexive candidates towards their peers, called peer reflexive candidates.

Finally, connectivity checks serve to keep NAT bindings alive.

It is fundamental to this STUN usage that the addresses and ports used for media are the same ones used for the Binding Requests and responses. Consequently, it will be necessary to demultiplex STUN traffic from whatever the media traffic is. This demultiplexing is done using the techniques described in [11].

7.2. Client Discovery of Server

The client does not follow the DNS-based procedures defined in [11]. Rather, the remote candidate of the check to be performed is used as the transport address of the STUN server. Note that the STUN server is a logical entity, and is not a physically distinct server in this usage.

7.3. Server Determination of Usage

The server is aware of this usage because it signaled this port through the offer/answer exchange. Any STUN packets received on this port will be for the connectivity check usage.

7.4. New Requests or Indications

This usage does not define any new message types.

7.5. New Attributes

This usage defines a new attribute, PRIORITY. This attribute indicates the priority that is to be associated with a peer reflexive candidate, should one be discovered by this check. It is a 32 bit unsigned integer, and has an attribute type of 0x0024.

7.6. New Error Response Codes

This usage does not define any new error response codes.

7.7. Client Procedures

This section defines additional procedures for the Binding Request transaction, beyond those described in [11].

7.7.1. Sending the Request

The agent acting as the client generates a connectivity check either periodically, or triggered. In either case, the check is generated by sending a Binding Request from a local candidate, to a remote candidate. The agent must know the username fragment for both

candidates and the password for the remote candidate.

A Binding Request serving as a connectivity check MUST utilize a STUN short term credential. Rather than being learned from a Shared Secret request, the short term credential is exchanged in the offer/answer procedures. In particular, the username is formed by concatenating the username fragment provided by the peer with the username fragment of the agent sending the request, separated by a colon (":"). The password is equal to the password provided by the peer. For example, consider the case where agent A is the offerer, and agent B is the answerer. Agent A included a username fragment of AFRAG for its candidates, and a password of APASS. Agent B provided a username fragment of BFRAG and a password of BPASS. A connectivity check from A to B (and its response of course) utilize the username BFRAG:AFRAG and a password of BPASS. A connectivity check from B to A (and its response) utilize the username AFRAG:BFRAG and a password of APASS.

All Binding Requests for the connectivity check usage MUST contain the PRIORITY attribute. This MUST be set equal to the priority that would be assigned, based on the algorithm in Section 4.2, to a peer reflexive candidate learned from this check. Such a peer reflexive candidate has a stream ID, component ID and local preference that are equal to the host candidate from which the check is being sent, but a type preference equal to the value associated with peer reflexive candidates.

The Binding Request by an agent MUST include the USERNAME and MESSAGE-INTEGRITY attributes. That is, an agent MUST NOT wait to be challenged for short term credentials. Rather, it MUST provide them in the Binding Request right away.

7.7.2. Processing the Response

If the STUN transaction generates an unrecoverable failure response or times out, the agent sets the state of the check to Failed. The remainder of this section applies to processing of successful responses (any response from 200 to 299).

The agent MUST check that the source IP address and port of the response equals the destination IP address and port that the Binding Request was sent to, and that the destination IP address and port of the response match the source IP address and port that the Binding Request was sent from. If these do not match, the agent sets the state of the check to Failed. The processing described in the remainder of this section MUST NOT be performed.

If the check succeeds, processing continues and the agent changes the

state for this check to Succeeded. Next, the agent sees if the success of this check can cause other checks to be unfrozen. If the check had a component ID of one, the agent MUST change the states for all other Frozen checks for the same media stream and same foundation, but different component IDs, to Waiting. If the component ID for the check was equal to the number of components for the media stream, the agent MUST change the state for all other Frozen checks for the first component of different media streams (and thus in different check lists) but the same foundation, to Waiting.

Next, the agent checks the mapped address from the STUN response. If the transport address does not match any of the local candidates that the agent knows about, the mapped address representes a new peer reflexive candidate. Its type is equal to peer reflexive. Its base is set equal to the candidate from which the STUN check was sent. Its username fragment and password are identical to the candidate from which the check was sent. It is assigned the priority value that was placed in the PRIORITY attribute of the request. Its foundation is selected as described in Section 4.1. The peer reflexive candidate is then added to the list of local candidates known by the agent (though it is not paired with other remote candidates at this time).

In addition, the agent creates a candidate pair whose local candidate equals the mapped address of the response, and whose remote candidate equals the destination address to which the request was sent. This is called a validated pair, since it has been validated by a STUN connectivity check. It is very important to note that this validated pair will often not be identical to the check itself; in many cases, the local candidate (learned through the mapped address in the response) will be different than the local candidate the request was sent from. However, the agent will know, either from the SDP or through the PRIORITY attribute that was present in a STUN request, the priorities of the local and remote candidates of the validated pair. Based on these priorities, a priority for the validated pair itself is computed if it was not already known, using the algorithm in Section 5.6, and the pair is added to the valid list.

7.8. Server Procedures

An agent MUST be prepared to receive a Binding Request on the base of each candidate it included in its most recent offer or answer. Receipt of a Binding Request on a transport address that the agent had included in a candidate attribute is an indication that the connectivity check usage applies to the request.

The agent MUST use a short term credential to authenticate the request and perform a message integrity check. The agent MUST accept

a credential if the username consists of two values separated by a colon, where the first value is equal to the username fragment generated by the agent in an offer or answer for a session in-progress, and the password is equal to the password for that username fragment. It is possible (and in fact very likely) that an offeror will receive a Binding Request prior to receiving the answer from its peer. However, the request can be processed without receiving this answer, and a response generated.

For requests being received on a relayed candidate, the source transport address used for STUN processing (namely, generation of the XOR-MAPPED-ADDRESS attribute) is the transport address as seen by the relay. That source transport address will be present in the REMOTE-ADDRESS attribute of a STUN Data Indication message, if the Binding Request was delivered through a Data Indication. If the Binding Request was not encapsulated in a Data Indication, that source address is equal to the current active destination for the STUN relay session.

When the agent receives a STUN Binding Request for which it generates a successful response, the agent checks the source transport address of the request. If this transport address does not match any existing remote candidates, it represents a new peer reflexive remote candidate. This candidate is given a priority equal to the PRIORITY attribute from the request. The type of the candidate is equal to peer reflexive. Its foundation is set to an arbitrary value, different from the foundation for all other remote candidates. The username fragment for this candidate is equal to the bottom half (the part after the colon) of the username in the Binding Request that was just received. The password for this username fragment is taken from the SDP from the peer. If agent has not yet received this SDP (a likely case for the offerer in the initial offer/answer exchange), it MUST wait for the SDP to be received, and then proceed with rest of the processing described in the remainder of this section. This candidate is then added to the list of remote candidates. However, it is not paired with any local candidates.

Next, the agent MUST generate a triggered check in the reverse directon if it has not already sent such a check. The triggered check has a local candidate equal to the candidate on which the STUN request was received, and a remote candidate equal to the source transport address where the request came from (which may be a newly formed peer reflexive candidate). The agent knows the priorities for the local and remote candidates of this check, and so can compute the priority for the check itself. If there is already a check on the check list with this same local and remote candidates, and the state of that check is Waiting or Frozen, its state is changed to In-Progress and the check is performed. If there was already a check on

the check list with this same local and remote candidates, and its state was In-Progress, the agent SHOULD generate an immediate retransmit of the Binding Request. This is to facilitate rapid completion of ICE when both agents are behind NAT. If there was a check in the list already and its state was Succeeded or Failed, nothing further is done. If there was no matching check on the check list, it is inserted into the check list based on its priority, its state is set to In-Progress, and the check is performed.

7.9. Security Considerations for Connectivity Check

Security considerations for the connectivity check are discussed in Section 15.

8. Completing the ICE Checks

When a pair is added to the valid list, and the agent was the offeror in the most recent offer/answer exchange, the agent MUST check to see if there is a pair on the validated list for each component of each media stream. If there is, the offeror MUST stop timer Ta, and MUST cease retransmitting any Binding Requests for transactions in progress. It MUST ignore any responses which may subsequently arrive to transactions previously in progress. The offeror MUST generate an updated offer as described in Section 9. It does this regardless of whether the highest priority pairs in the check list match the current in-use candidate pairs.

When a pair is aded to the valid list, and the agent was the answerer in the most recent offer/answer exchange, the agent MAY begin sending media using that candidate pair, as described in Section 11.1. In addition, if there is a candidate pair on the valid list for each component of each media stream, the answerer MUST stop timer Ta, and MUST cease retransmitting any Binding Requests for transactions in progress. It MUST ignore any responses which may subsequently arrive to transactions previously in progress.

Note that only agent that was the answerer in the most recent offer/answer exchange gets to send media right away. The offeror must wait for a subsequent offer/answer exchange if the valid candidates don't match those in the m/c-line.

> OPEN ISSUE: It is possible that higher priority checks may still succeed, if we allowed things to continue. This can happen for several reasons. First, an in-progress check of higher priority had some packet loss and thus hasn't completed. Timer Tws was meant to handle this (I removed this timer from -10 to simplify). More interestingly, higher priority checks may have not been done

because a triggered check of lower priority succeeded. This happens in cases where the number of checks at each agent are assymetric. It is possible to fix both of these problems by delaying the completion of the ICE procedures for a bit more time. This adds complexity and latency. The basic algorithm would be this. You take the lowest priority pair in the valid list. You keep doing checks as long as there are higher priority checks on the list in the Waiting state. If there are none, you wait a brief time (say 50ms) and then consider ICE finished.

9. Subsequent Offer/Answer Exchanges

An agent MAY generate a subsequent offer at any time. However, the rules in Section 7.7.2 will cause the offerer to generate an updated offer when the candidates in the valid list are not all in-use.

9.1. Generating the Offer

When an agent generates an updated offer, the set of candidate attributes to include depend on the state of ICE processing. If ICE is "done", which occurs when the valid list includes a candidate pair for each component of each media stream, the agent MUST include a candidate attribute for each local candidate amongst the pairs in the valid list (including peer reflexive candidates), and SHOULD NOT include any others. This will cause STUN keepalives to be sent for the in-use candidates, and thats it.

If, however, the valid list does not yet include a candidate pair for each component of each media stream, the agent SHOULD include all current candidates, including any peer reflexive candidates it has learned since the last offer or answer it sent. This MAY include candidates it did not offer previously, but which it has gathered since the last offer/answer exchange.

If a candidate was sent in a previous offer/answer exchange, it SHOULD have the same priority. For a peer reflexive candidate, the priority SHOULD be the same as determined by the processing in Section 7.7.2. The foundation SHOULD be the same. The username fragments and passwords for a media stream SHOULD remain the same as the previous offer or answer.

Population of the m/c-lines also depends on the state of ICE processing. If, for a particular media stream, the valid list has candidate pairs for all of the components of that media stream, those pairs are used. In particular, the m/c-line would be constructed by from the local candidate from each of those candidate pairs. In addition, the agent MUST include the a=remote-candidates attribute

for that media stream, and include in it the remote candidates for each of the pairs that were used.

If, for a particular media stream, the valid list does not have pairs for all of the components of the stream, the agent SHOULD populate the m/c-line for that media stream based on the considerations in Section 4.3.

The agent MUST use the same ice-pwd and ice-ufrag for a media stream as its previous offer or answer. Note that it is permissible to use a session-level attribute in one offer, but to provide the same password as a media-level attribute in a subsequent offer. This is not a change in password, just a change in its representation.

9.2. Receiving the Offer and Generating an Answer

When the answerer generates its answer, it must decide what candidates to include in the answer, and how to populate the m/c-line.

For each media stream in the offer, the agent checks to see if the stream contained the remote-candidates attribute. If it did, it means that the offerer believed that ICE processing has completed for that media stream. In this case, the remote-candidates attribute contains the candidates that the answerer is supposed to use. It is possible that the agent doesn't even know of these candidates yet; they will be discovered shortly through a response to an in-progress check. The agent MUST populate the m/c-line with the candidates from the a=remote-candidates attribute. In addition, it MUST include an a=candidate attribute in its answer for each candidate in the a=remote-candidates attribute. If the agent is not aware of the candidate yet, it will need to generate a priority value for it. The type preference in the computation is peer-reflexive, and the stream ID and component ID are known from the offer. The agent chooses an arbitrary local preference value if it is multi-homed, since it won't yet know the interface associated with this candidate.

If a media stream does not yet contain the a=remote-candidates attribute, it means that the offerer believes that ICE checks are still in progress for that media stream. In this case, the answerer SHOULD include an a=candidate attribute for all of the candidates for that media stream it knows about (including peer-reflexive candidates). The m/c-line is populated based on the considerations in Section 4.3.

Construction of the ice-pwd and ice-ufrag are identical to the procedures followed by the offerer, as described in Section 9.1.

Note that the a=remote-candidates attribute SHOULD NOT be included in the answer, and if included, will just be ignored by the offerer, since it is not used in any processing of the answer.

9.3. Updating the Check and Valid Lists

Once the subsequent offer/answer exchange has completed, each agent needs to compute the new check lists resulting from this exchange, and then remove any pairs from the valid list which are no longer usable. Once these adjustments are made, ICE processing continues using these new lists.

Each agent recomputes the check lists using the procedures described in Section 5.6. If a check on the new check lists was also on the previous check lists, and its state was Waiting, In-Progress, Succeeded or Failed, its state is copied over. If a check on the new check lists does not have a state (because its a new check on an existing check list, or a check on a new check list, or the check was on an old check list but its state was not copied over) its state is set to Frozen.

If none of the check lists are active (meaning that the checks in each check list are Frozen), the agent sets the first check in the check list for the first media stream to Waiting, and then sets the state of all other checks in that check list for the same component ID and with the same foundation to Waiting as well.

Next, the agent goes through each check list, starting with the highest priority check. If a check has a state of Succeeded, and it has a component ID of 1, then all Frozen checks in the same check list with the same foundation whose component IDs are not one, have their state set to Waiting. If, for a particular check list, there are checks for each component of that media stream in the Succeeded state, the agent moves the state of all Frozen checks for the first component of all other media streams (and thus in different check lists) with the same foundation to Waiting.

If a check was on the old check list, but was not on the new check list, and had a state of In-Progress, the corresponding STUN transaction is abandoned. No further retransmits will be sent for the STUN request, and any response that might be received is ignored.

Next, the agent prunes the valid list. For each pair on the valid list, the agent examines each candidate in the pair. If the candidate was not peer reflexive, and was not present in the most recent offer/answer exchange, the candidate pair is removed from the valid list.

OPEN ISSUE: This means that you cannot forcefully remove a peer reflexive candidate. This feature was possible, at much complexity, in previous versions of the spec. An alternative is to remove a peer reflexive candidate if it was not present in the offer/answer, and was discovered more than 500ms ago.

10. Keepalives

STUN connectivity checks are also used to keep NAT bindings open once a session is underway. This is accomplished by periodically re-starting the check process, as described in this section.

Once the initial offer/answer exchange has taken place, the agent sets a timer to fire in Tr seconds. Tr SHOULD be configurable and SHOULD have a default of 15 seconds. When Tr fires, the agent MUST reset the states for all of the checks in the check list using the procedures defined in Section 5.6 and then begin performing periodic checks as described in Section 5.7. By the time the timer fires for the first time, the check list will include only the in-use candidates. Reperforming these checks will therefore performing a period keepalive.

OPEN ISSUE: ICE isn't saying anything about what happens if these periodic keepalives should fail. It they do, something really bad has happened, like a NAT reboot or failure. I think we should keep that out of scope.

When an ICE agent is communicating with an agent that is not ICE-aware, keepalives still need to be utilized. Indeed, these keepalives are essential even if neither endpoint implements ICE. As such, this specification defines keepalive behavior generally, for endpoints that support ICE, and those that do not.

All endpoints MUST send keepalives for each media session. These keepalives MUST be sent regardless of whether the media stream is currently inactive, sendonly, recvonly or sendrecv. The keepalive SHOULD be sent using a format which is supported by its peer. ICE endpoints allow for STUN-based keepalives for UDP streams, and as such, STUN keepalives MUST be used when an agent is communicating with a peer that supports ICE. An agent can determine that its peer supports ICE by the presence of the a=candidate attributes for each media session. If the peer does not support ICE, the choice of a packet format for keepalives is a matter of local implementation. A format which allows packets to easily be sent in the absence of actual media content is RECOMMENDED. Examples of formats which readily meet this goal are RTP No-Op [27] and RTP comfort noise [23]. If the peer doesn't support any formats that are particularly well

suited for keepalives, an agent SHOULD send RTP packets with an incorrect version number, or some other form of error which would cause them to be discarded by the peer.

STUN-based keepalives will be sent periodically every Tr seconds as described above. If STUN keepalives are not in use (because the peer does not support ICE), an agent SHOULD ensure that a media packet is sent every Tr seconds. If one is not sent as a consequence of normal media communications, a keepalive packet using one of the formats discussed above SHOULD be sent.

11. Media Handling

11.1. Sending Media

Agents always send media using a candidate pair. An agent will send media to the remote candidate in the pair (setting the destination address and port of the packet equal to that remote candidate), and will send it from the local candidate. When the local candidate is server or peer reflexive, media is originated from the base. Media sent from a relayed candidate is sent through that relay, using procedures defined in [12].

If an agent was the offerer in the most recent offer/answer exchange, when it sends media, it MUST use the candidates in the m/c-line for each media stream. However, it MUST only send media once those candidates also appear in the valid list. If the candidates in the m/c-line are not the ones that are ultimately selected by ICE, this implies that the offerer will need to wait for the subsequent offer/answer exchange to complete before it can send media.

If an agent was the answerer in the most recent offer/answer exchange, the rules are different. When the agent wishes to send media, and the candidate pairs in the m/c-lines are also the highest priority ones in the valid list for each media stream, it uses those candidate pairs. If, however, the highest priority pairs in the valid list for a media stream are not the same as the ones in the m/c-lines, the agent MUST use the highest priority pairs in the valid list. However, the agent MUST discontinue using those candidate pairs Tlo seconds after the next opportunity its peer would have to send an updated offer. In the case of an answer delivered in a 200 OK to an offer in a SIP INVITE (regardless of whether that same answer appeared in an earlier unreliable provisional response), this would be Tlo seconds after receipt of the ACK. Tlo SHOULD be configurable and SHOULD have a default of 5 seconds. This time represents the amount of time it should take the offerer to perform its connectivity checks, arrive at the same conclusion about the

candidate pair, and then generate an updated offer. If, after Tlo seconds, no updated offer arrives, the answerer MUST cease sending media, and will need to wait for the updated offer.

OPEN ISSUE: In previous versions of ICE, once this timer fired, you just sent media to the one in the m/c-line. This causes the media streams to flip back and forth between addresses, which I am trying to avoid. Since this timer should never go off anyway, I removed this feature.

ICE has interactions with jitter buffer adaptation mechanisms. An RTP stream can begin using one candidate, and switch to another one, though this happens rarely with ICE. The newer candidate may result in RTP packets taking a different path through the network - one with different delay characteristics. As discussed below, agents are encouraged to re-adjust jitter buffers when there are changes in source or destination address. Furthermore, many audio codecs use the marker bit to signal the beginning of a talkspurt, for the purposes of jitter buffer adaptation. For such codecs, it is RECOMMENDED that the sender change the marker bit when an agent switches transmission of media from one candidate pair to another.

11.2. Receiving Media

ICE implementations MUST be prepared to receive media on any candidates provided in the most recent offer/answer exchange.

It is RECOMMENDED that, when an agent receives an RTP packet with a new source or destination IP address for a particular media stream, that the agent re-adjust its jitter buffers.

RFC 3550 [20] describes an algorithm in Section 8.2 for detecting SSRC collisions and loops. These algorithms are based, in part, on seeing different source transport addresses with the same SSRC. However, when ICE is used, such changes will sometimes occur as the media streams switch between candidates. An agent will be able to determine that a media stream is from the same peer as a consequence of the STUN exchange that proceeds media transmission. Thus, if there is a change in source transport address, but the media packets come from the same peer agent, this SHOULD NOT be treated as an SSRC collision.

12. Usage with SIP

12.1. Latency Guidelines

ICE requires a series of STUN-based connectivity checks to take place

between endpoints. These checks start from the answerer on generation of its answer, and start from the offerer when it receives the answer. These checks can take time to complete, and as such, the selection of messages to use with offers and answers can effect perceived user latency. Two latency figures are of particular interest. These are the post-pickup delay and the post-dial delay. The post-pickup delay refers to the time between when a user "answers the phone" and when any speech they utter can be delivered to the caller. The post-dial delay refers to the time between when a user enters the destination address for the user, and ringback begins as a consequence of having succesfully started ringing the phone of the called party.

To reduce post-dial delays, it is RECOMMENDED that the caller begin gathering candidates prior to actually sending its initial INVITE. This can be started upon user interface cues that a call is pending, such as activity on a keypad or the phone going offhook.

If an offer is received in an INVITE request, the callee SHOULD immediately gather its candidates and then generate an answer in a provisional response. When reliable provisional responses are not used, the SDP in the provisional response is the answer, and that exact same answer reappears in the 200 OK. To deal with possible losses of the provisional response, it SHOULD be retransmitted until some indication of receipt. This indication can either be through PRACK [9], or through the receipt of a successful STUN Binding Request. Even if PRACK is not used, the provisional response SHOULD be retransmitted using the exponential backoff described in [9]. Furthermore, once the answer has been sent, the agent SHOULD begin its connectivity checks. Once candidate pairs for each component of a media stream enter the valid list, the callee can begin sending media on that media stream.

However, prior to this point, any media that needs to be sent towards the caller (such as SIP early media [25] cannot be transmitted. For this reason, implementations SHOULD delay alerting the called party until candidates for each component of each media stream have entered the valid list. In the case of a PSTN gateway, this would mean that the setup message into the PSTN is delayed until this point. Doing this increases the post-dial delay, but has the effect of eliminating 'ghost rings'. Ghost rings are cases where the called party hears the phone ring, picks up, but hears nothing and cannot be heard. This technique works without requiring support for, or usage of, preconditions [6], since its a localized decision. It also has the benefit of guaranteeing that not a single packet of media will get clipped, so that post-pickup delay is zero. If an agent chooses to delay local alerting in this way, it SHOULD generate a 180 response once alerting begins.

Based on the rules in Section 11.1, the offerer will not be able to send media until the highest priority valid candidates match the m/c-line. When used with SIP, if the initial offer is sent in the INVITE, and the answer is sent in both the provisional and final 200 OK response, the offerer will generally not be able to send media until it sends a re-INVITE and receives the 200 OK response to that re-INVITE. This can take several hundred milliseconds. If this latency is an issue (it is generally not considered an issue for voice systems), reliable provisional responses [9] MAY be used, in which case an UPDATE [24] can be used to send an updated offer prior to the call being answered.

As discussed in Section 15, offer/answer exchanges SHOULD be secured against eavesdropping and man-in-the-middle attacks. To do that, the usage of SIPS [3] is RECOMMENDED when used in concert with ICE.

12.2. Interactions with Forking

ICE interacts very well with forking. Indeed, ICE fixes some of the problems associated with forking. Without ICE, when a call forks and the caller receives multiple incoming media streams, it cannot determine which media stream corresponds to which callee.

With ICE, this problem is resolved. The connectivity checks which occur prior to transmission of media carry username fragments, which in turn are correlated to a specific callee. Subsequent media packets which arrive on the same 5-tuple as the connectivity check will be associated with that same callee. Thus, the caller can perform this correlation as long as it has received an answer.

12.3. Interactions with Preconditions

Quality of Service (QoS) preconditions, which are defined in RFC 3312 [6] and RFC 4032 [7], apply only to the transport addresses listed in the m/c lines in an offer/answer. If ICE changes the transport address where media is received, this change is reflected in the m/c lines of a new offer/answer. As such, it appears like any other re-INVITE would, and is fully treated in RFC 3312 and 4032, which apply without regard to the fact that the m/c lines are changing due to ICE negotiations ocurring "in the background".

Indeed, an agent SHOULD NOT indicate that Qos preconditions have been met until the ICE checks have completed and selected the candidate pairs to be used for media.

ICE also has (purposeful) interactions with connectivity preconditions [26]. Those interactions are described there. Note that the procedures described in Section 12.1 describe their own type

of "preconditions", albeit with less functionality than those provided by the explicit preconditions in [26].

12.4. Interactions with Third Party Call Control

ICE works with Flows I and IV as described in [16]. Flow I works without the controller supporting or being aware of ICE. Flow IV will work as long as the controller passes along the ICE attributes without alteration. Flow III may disrupt ICE processing, since it will distort the stream ID values used in the computation of priorities. When there is but a single media stream, Flow III will work as long as the controller passes through the ICE attributes unmodified. Flow II is fundamentally incompatible with ICE; each agent will believe itself to be the answerer and thus never generate a re-INVITE.

> OPEN ISSUE: Its really too bad flow III doesn't work with multimedia; should consider ways to make it work. There are several ways.

The flows for continued operation, as described in Section 7 of RFC 3725, require additional behavior of ICE implementations to support. In particular, if an agent receives a mid-dialog re-INVITE that contains no offer, it MUST go through the process of gathering candidates, prioritizing them and generating an offer, as if this was an initial offer for a session. Furthermore, that list of candidates SHOULD include the ones currently in-use.

13. Grammar

This specification defines four new SDP attributes - the "candidate", "remote-candidates", "ice-ufrag" and "ice-pwd" attributes.

The candidate attribute is a media-level attribute only. It contains a transport address for a candidate that can be used for connectivity checks.

The syntax of this attribute is defined using Augmented BNF as defined in RFC 4234 [8]:

```
candidate-attribute   = "candidate" ":" foundation SP component-id SP
                        transport SP
                        priority SP
                        connection-address SP     ;from RFC 4566
                        port         ;port from RFC 4566
                        [SP cand-type]
                        [SP rel-addr]
                        [SP rel-port]
                        *(SP extension-att-name SP
                             extension-att-value)

foundation            = 1*ice-char
component-id          = 1*DIGIT
transport             = "UDP" / transport-extension
transport-extension   = token              ; from RFC 3261
priority              = 1*DIGIT
cand-type             = "typ" SP candidate-types
candidate-types       = "host" / "srflx" / "prflx" / "relay" / token
rel-addr              = "raddr" SP connection-address
rel-port              = "rport" SP port
extension-att-name    = byte-string    ;from RFC 4566
extension-att-value   = byte-string
ice-char              = ALPHA / DIGIT / "+" / "/"
```

The foundation is composed of one or more ice-char. The component-id is a positive integer, which identifies the specific component for which the transport address is a candidate. It MUST start at 1 and MUST increment by 1 for each component of a particular candidate. The connect-address production is taken from RFC 4566 [10], allowing for IPv4 addresses, IPv6 addresses and FQDNs. The port production is also taken from RFC 4566 [10]. The token production is taken from RFC 3261 [3]. The transport production indicates the transport protocol for the candidate. This specification only defines UDP. However, extensibility is provided to allow for future transport protocols to be used with ICE, such as TCP or the Datagram Congestion Control Protocol (DCCP) [28].

The cand-type production encodes the type of candidate. This specification defines the values "host", "srflx", "prflx" and "relay" for host, server reflexive, peer reflexive and relayed candidates, respectively. The set of candidate types is extensible for the future. Inclusion of the candidate type is optional. The rel-addr and rel-port productions convey information the related transport addresses. Rules for inclusion of these values is described in Section 4.4.

The a=candidate attribute can itself be extended. The grammar allows

for new name/value pairs to be added at the end of the attribute. An implementation MUST ignore any name/value pairs it doesn't understand.

The syntax of the "remote-candidates" attribute is defined using Augmented BNF as defined in RFC 4234 [8]. The remote-candidates attribute is a media level attribute only.

```
remote-candidate-att = "remote-candidates" ":" remote-candidate
                       0*(SP remote-candidate)
remote-candidate = component-ID SP connection-address SP port
```

The attribute contains a connection-address and port for each component. The ordering of components is irrelevant. However, a value MUST be present for each component of a media stream.

The syntax of the "ice-pwd" and "ice-ufrag" attributes are defined as:

```
ice-pwd-att           = "ice-pwd" ":" password
ice-ufrag-att         = "ice-ufrag" ":" ufrag
password              = 22*ice-char
ufrag                 = 4*ice-char
```

The "ice-pwd" and "ice-ufrag" attributes can appear at either the session-level or media-level. When present in both, the value in the media-level takes precedence. Thus, the value at the session level is effectively a default that applies to all media streams, unless overriden by a media-level value.

14. Example

Two agents, L and R, are using ICE. Both agents have a single IPv4 interface. For agent L, it is 10.0.1.1, and for agent R, 192.0.2.1. Both are configured with a single STUN server each (indeed, the same one for each), which is listening for STUN requests at an IP address of 192.0.2.2 and port 3478. This STUN server supports both the Binding Discovery usage and the Relay usage. Agent L is behind a NAT, and agent R is on the public Internet. The NAT has an endpoint independent mapping property and an address dependent filtering property. The public side of the NAT has an IP address of 192.0.2.3.

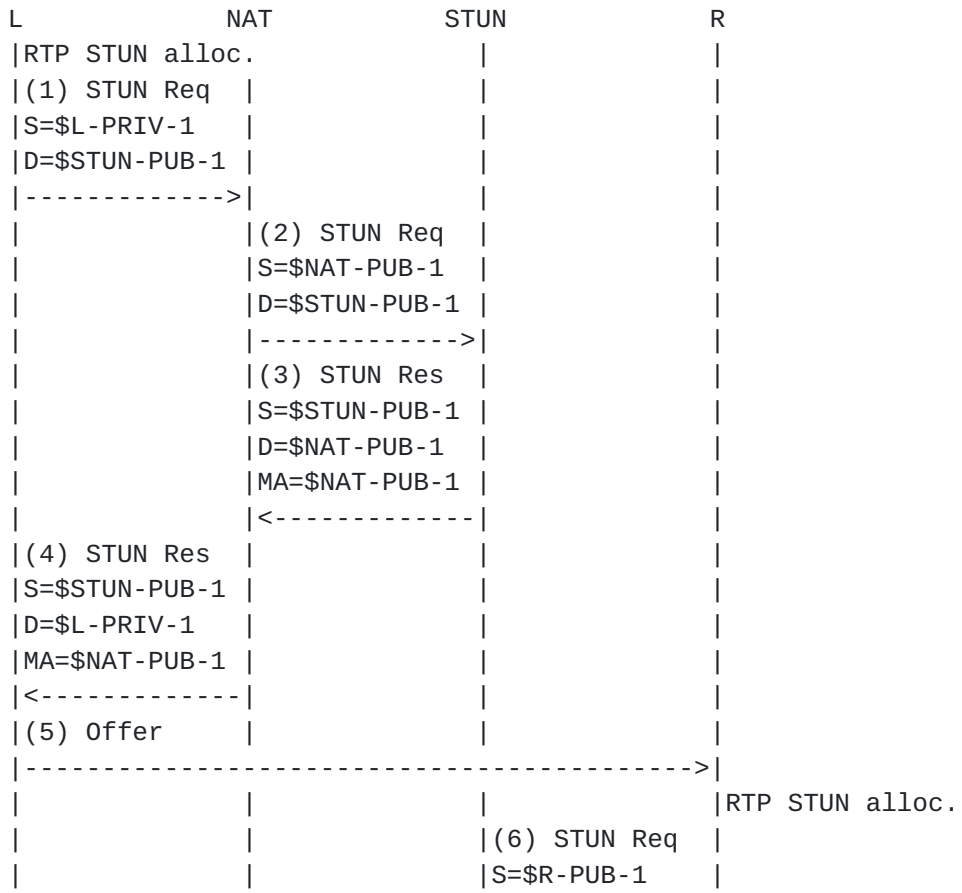
To facilitate understanding, transport addresses are listed using variables that have mnemonic names. The format of the name is entity-type-seqno, where entity refers to the entity whose interface

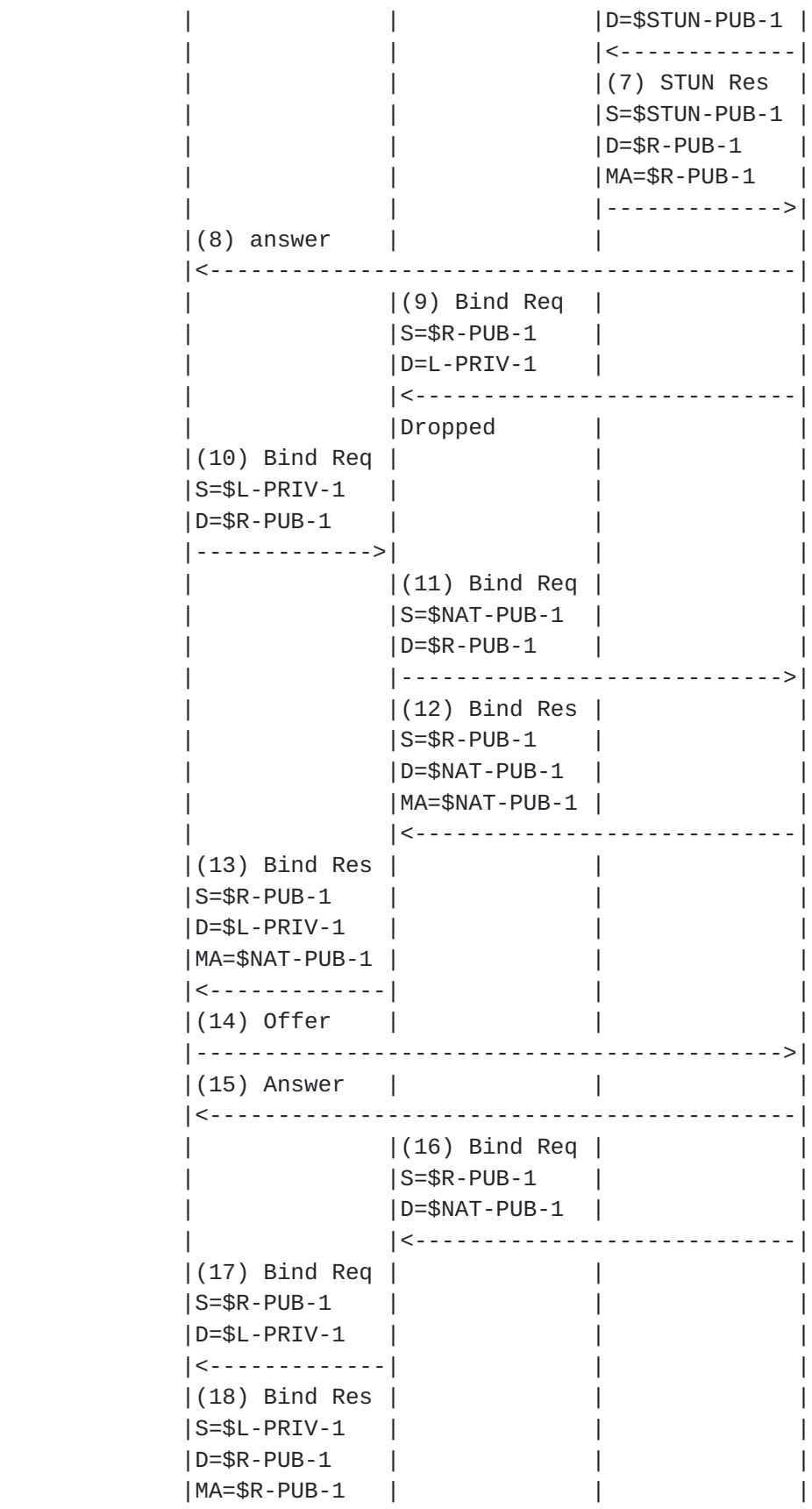
the transport address is on, and is one of "L", "R", "STUN", or "NAT". The type is either "PUB" for transport addresses that are public, and "PRIV" for transport addresses that are private. Finally, seq-no is a sequence number that is different for each transport address of the same type on a particular entity. Each variable has an IP address and port, denoted by varname.IP and varname.PORT, respectively, where varname is the name of the variable.

The STUN server has advertised transport address STUN-PUB-1 (which is 192.0.2.2:3478) for both the binding discovery usage and the relay usage. However, neither agent is using the relay usage.

In the call flow itself, STUN messages are annotated with several attributes. The "S=" attribute indicates the source transport address of the message. The "D=" attribute indicates the destination transport address of the message. The "MA=" attribute is used in STUN Binding Response messages and refers to the mapped address.

The call flow examples omit STUN authentication operations and RTCP, and focus on RTP for a single media stream.




D=$STUN-PUB-1
(7) STUN Res
S=$STUN-PUB-1
D=$R-PUB-1
MA=$R-PUB-1
(8) answer
(9) Bind Req
S=$R-PUB-1
D=L-PRIV-1
Dropped
(10) Bind Req
S=$L-PRIV-1
D=$R-PUB-1
(11) Bind Req
S=$NAT-PUB-1
D=$R-PUB-1
(12) Bind Res
S=$R-PUB-1
D=$NAT-PUB-1
MA=$NAT-PUB-1
(13) Bind Res
S=$R-PUB-1
D=$L-PRIV-1
MA=$NAT-PUB-1
(14) Offer
(15) Answer
(16) Bind Req
S=$R-PUB-1
D=$NAT-PUB-1
(17) Bind Req
S=$R-PUB-1
D=$L-PRIV-1
(18) Bind Res
S=$L-PRIV-1
D=$R-PUB-1
MA=$R-PUB-1

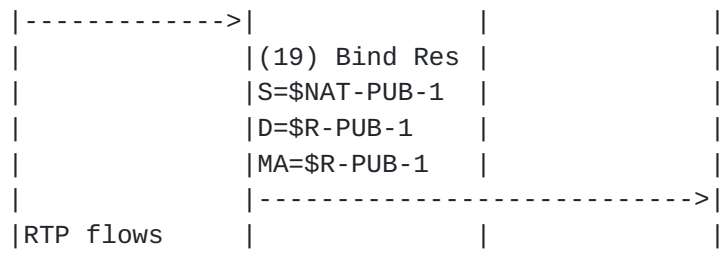


Figure 9

First, agent L obtains a host candidate from its local interface (not shown), and from that, sends a STUN Binding Request to the STUN server to get a server reflexive candidate (messages 1-4). Recall that the NAT has the address and port independent mapping property. Here, it creates a binding of NAT-PUB-1 for this UDP request, and this becomes the server reflexive candidate for RTP.

Agent L sets a type preference of 126 for the host candidate and 100 for the server reflexive. The local preference is 65535. Based on this, the priority of the host candidate is 2130706178 and for the server reflexive candidate is 1694498562. The host candidate is assigned a foundation of 1, and the server reflexive, a foundation of 2. It chooses its server reflexive candidate as the in-use candidate, and encodes it into the m/c-line. The resulting offer (message 5) looks like (lines folded for clarity):

```
    v=0
    o=jdoe 2890844526 2890842807 IN IP4 $L-PRIV-1.IP
    s=
    c=IN IP4 $NAT-PUB-1.IP
    t=0 0
    a=ice-pwd:asd88fgpdd777uzjYhagZg
    a=ice-ufrag:8hhY
    m=audio $NAT-PUB-1.PORT RTP/AVP 0
    a=rtpmap:0 PCMU/8000
    a=candidate:1 1 UDP 2130706178 $L-PRIV-1.IP $L-PRIV-1.PORT typ local
    a=candidate:2 1 UDP 1694498562 $NAT-PUB-1.IP $NAT-PUB-1.PORT typ srflx
raddr
$L-PRIV-1.IP rport $L-PRIV-1.PORT
```

The offer, with the variables replaced with their values, will look like (lines folded for clarity):

```
v=0
o=jdoe 2890844526 2890842807 IN IP4 10.0.1.1
s=
c=IN IP4 192.0.2.3
t=0 0
a=ice-pwd:asd88fgpdd777uzjYhagZg
a=ice-ufrag:8hhY
m=audio 45664 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=candidate:1 1 UDP 2130706178 10.0.1.1 8998 typ local
a=candidate:2 1 UDP 1694498562 192.0.2.3 45664 typ srflx raddr
10.0.1.1 rport 8998
```

This offer is received at agent R. Agent R will obtain a host candidate, and from it, obtain a server reflexive candidate (messages 6-7). Since R is not behind a NAT, this candidate is identical to its host candidate, and they share the same base. It therefore discards this candidate and ends up with a single host candidate. With identical type and local preferences as L, the priority for this candidate is 2130706178. It chooses a foundation of 1 for its single candidate. Its resulting answer looks like:

```
v=0
o=bob 2808844564 2808844564 IN IP4 $R-PUB-1.IP
s=
c=IN IP4 $R-PUB-1.IP
t=0 0
a=ice-pwd:YH75Fviy6338Vbrhrlp8Yh
a=ice-ufrag:9uB6
m=audio $R-PUB-1.PORT RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=candidate:1 1 UDP 2130706178 $R-PUB-1.IP $R-PUB-1.PORT typ local
```

With the variables filled in:

```
v=0
o=bob 2808844564 2808844564 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=ice-pwd:YH75Fviy6338Vbrhrlp8Yh
a=ice-ufrag:9uB6
m=audio 3478 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=candidate:1 1 UDP 2130706178 192.0.2.1 3478 typ local
```

Agents L and R both pair up the candidates. They both initially have two. However, agent L will prune the pair containing its server reflexive candidate, resulting in just one. At agent L, this pair (the check) has a local candidate of $L_PRIV_1 and remote candidate of $R_PUB_1, and has a candidate pair priority of 4.57566E+18 (note that an implementation would represent this as a 64 bit integer so as not to lose precision). At agent R, there are two checks. The highest priority has a local candidate of $R_PUB_1 and remote candidate of $L_PRIV_1 and has a priority of 4.57566E+18, and the second has a local candidate of $R_PUB_1 and remote candidate of $NAT_PUB_1 and priority 3.63891E+18.

Agent R begins its connectivity check (message 9) for the first pair (between the two host candidates). The host candidate from agent L is private and behind a different NAT, and thus this check is discarded.

When agent L gets the answer, it performs its one and only connectivity check (messages 10-13). This will succeed. This causes agent L to create a new pair, whos local candidate is from the mapped address in the binding response (NAT-PUB-1 from message 13) and whose remote candidate is the destination of the request (R-PUB-1 from message 10). This is added to the valid list. At this point, agent L examines the valid list and sees that there is a candidate there for each component of each media stream (which is just RTP for the single audio stream). It therefore considers ICE checks complete and sends an updated offer (message 14). This offer serves only to remove the candidate that was not selected and indicate the remote candidates; the m/c-line remains unchanged. This offer looks like:

```
    v=0
    o=jdoe 2890844528 2890842809 IN IP4 10.0.1.1
    s=
    c=IN IP4 192.0.2.3
    t=0 0
    a=ice-pwd:asd88fgpdd777uzjYhagZg
    a=ice-ufrag:8hhY
    m=audio 45664 RTP/AVP 0
    a=remote-candidates 1 192.0.2.1 3478
    a=rtpmap:0 PCMU/8000
    a=candidate:2 1 UDP 1694498562 192.0.2.3 45664 typ srflx raddr
10.0.1.1 rport 8998
```

Agent R can construct the answer. Since the remote-candidates listed in the offer match the ones that agent R had already selected for the m/c-line in the previous answer, there is no change there. Its answer therefore looks like:

```
v=0
o=bob 2808844565 2808844566 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=ice-pwd:YH75Fviy6338Vbrhrlp8Yh
a=ice-ufrag:9uB6
m=audio 3478 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=candidate:1 1 UDP 2130706178 192.0.2.1 3478 typ local
```

Upon receipt of the check from agent L (message 11), agent R will generate its triggered check. This check happens to match the next one on its check list - from its host candidate to agent L's server reflexive candidate. This check (messages 16-19) will succeed. Consequently, agent R constructs a new candidate pair using the mapped address from the response as the local candidate (R-PUB-1) and the destination of the request (NAT-PUB-1) as the remote candidate. This pair is added to the valid list. Since this pair matches the pair in the m/c-lines, agent R can send media as well.

15. Security Considerations

There are several types of attacks possible in an ICE system. This section considers these attacks and their countermeasures.

15.1. Attacks on Connectivity Checks

An attacker might attempt to disrupt the STUN connectivity checks. Ultimately, all of these attacks fool an agent into thinking something incorrect about the results of the connectivity checks. The possible false conclusions an attacker can try and cause are:

False Invalid: An attacker can fool a pair of agents into thinking a candidate pair is invalid, when it isn't. This can be used to cause an agent to prefer a different candidate (such as one injected by the attacker), or to disrupt a call by forcing all candidates to fail.

False Valid: An attacker can fool a pair of agents into thinking a candidate pair is valid, when it isn't. This can cause an agent to proceed with a session, but then not be able to receive any media.

False Peer-Reflexive Candidate: An attacker can cause an agent to discover a new peer reflexive candidate, when it shouldn't have. This can be used to redirect media streams to a DoS target or to the attacker, for eavesdropping or other purposes.

False Valid on False Candidate: An attacker has already convinced an agent that there is a candidate with an address that doesn't actually route to that agent (for example, by injecting a false peer reflexive candidate or false server reflexive candidate). It must then launch an attack that forces the agents to believe that this candidate is valid.

Of the various techniques for creating faked STUN messages described in [11], many are not applicable for the connectivity checks. Compromises of STUN servers are not much of a concern, since the STUN servers are embedded in endpoints and distributed throughout the network. Thus, compromising the STUN server is equivalent to comprimising the endpoint, and if that happens, far more problematic attacks are possible than those against ICE. Similarly, DNS attacks are usually irrelevant since STUN servers are not typically discovered via DNS, they are signaled via IP addresses embedded in SDP. Injection of fake responses and relaying modified requests all can be handled in ICE with the countermeasures discussed below.

To force the false invalid result, the attacker has to wait for the connectivity check from one of the agents to be sent. When it is, the attacker needs to inject a fake response with an unrecoverable error response, such as a 600. However, since the candidate is, in fact, valid, the original request may reach the peer agent, and result in a success response. The attacker needs to force this packet or its response to be dropped, through a DoS attack, layer 2 network disruption, or other technique. If it doesn't do this, the success response will also reach the originator, alerting it to a possible attack. Fortunately, this attack is mitigated completely through the STUN message integrity mechanism. The attacker needs to inject a fake response, and in order for this response to be processed, the attacker needs the password. If the offer/answer signaling is secured, the attacker will not have the password.

Forcing the fake valid result works in a similar way. The agent needs to wait for the Binding Request from each agent, and inject a fake success response. The attacker won't need to worry about disrupting the actual response since, if the candidate is not valid, it presumably wouldn't be received anyway. However, like the fake invalid attack, this attack is mitigated completely through the STUN message integrity and offer/answer security techniques.

Forcing the false peer reflexive candidate result can be done either

with fake requests or responses, or with replays. We consider the fake requests and responses case first. It requires the attacker to send a Binding Request to one agent with a source IP address and port for the false candidate. In addition, the attacker must wait for a Binding Request from the other agent, and generate a fake response with a XOR-MAPPED-ADDRESS attribute containing the false candidate. Like the other attacks described here, this attack is mitigated by the STUN message integrity mechanisms and secure offer/answer exchanges.

Forcing the false peer reflexive candidate result with packet replays is different. The attacker waits until one of the agents sends a check. It intercepts this request, and replays it towards the other agent with a faked source IP address. It must also prevent the original request from reaching the remote agent, either by launching a DoS attack to cause the packet to be dropped, or forcing it to be dropped using layer 2 mechanisms. The replayed packet is received at the other agent, and accepted, since the integrity check passes (the integrity check cannot and does not cover the source IP address and port). It is then responded to. This response will contain a XOR-MAPPED-ADDRESS with the false candidate, and will be sent to that false candidate. The attacker must then intercept it and relay it towards the originator.

The other agent will then initiate a connectivity check towards that false candidate. This validation needs to succeed. This requires the attacker to force a false valid on a false candidate. Injecting of fake requests or responses to achieve this goal is prevented using the integrity mechanisms of STUN and the offer/answer exchange. Thus, this attack can only be launched through replays. To do that, the attacker must intercept the check towards this false candidate, and replay it towards the other agent. Then, it must intercept the response and replay that back as well.

This attack is very hard to launch unless the attacker themself is identified by the fake candidate. This is because it requires the attacker to intercept and replay packets sent by two different hosts. If both agents are on different networks (for example, across the public Internet), this attack can be hard to coordinate, since it needs to occur against two different endpoints on different parts of the network at the same time.

If the attacker themself is identified by the fake candidate the attack is easier to coordinate. However, if SRTP is used [21], the attacker will not be able to play the media packets, they will only be able to discard them, effectively disabling the media stream for the call. However, this attack requires the agent to disrupt packets in order to block the connectivity check from reaching the target.

In that case, if the goal is to disrupt the media stream, its much easier to just disrupt it with the same mechanism, rather than attack ICE.

15.2. Attacks on Address Gathering

ICE endpoints make use of STUN for gathering candidates rom a STUN server in the network. This is corresponds to the Binding Discovery usage of STUN described in [11]. As a consequence, the attacks against STUN itself that are described in that specification can still be used against the binding discovery usage when utilized with ICE.

However, the additional mechanisms provided by ICE actually counteract such attacks, making binding discovery with STUN more secure when combined with ICE than without ICE.

Consider an attacker which is able to provide an agent with a faked mapped address in a STUN Binding Request that is used for address gathering. This is the primary attack primitive described in [11]. This address will be used as a server reflexive candidate in the ICE exchange. For this candidate to actually be used for media, the attacker must also attack the connectivity checks, and in particular, force a false valid on a false candidate. This attack is very hard to launch if the false address identifies a third party, and is prevented by SRTP if it identifies the attacker themself.

If the attacker elects not to attack the connectivity checks, the worst it can do is prevent the server reflexive candidate from being used. However, if the peer agent has at least one candidate that is reachable by the agent under attack, the STUN connectivity checks themselves will provide a peer reflexive candidate that can be used for the exchange of media. Peer reflexive candidates are generally preferred over server reflexive candidates. As such, an attack solely on the STUN address gathering will normally have no impact on a session at all.

15.3. Attacks on the Offer/Answer Exchanges

An attacker that can modify or disrupt the offer/answer exchanges themselves can readily launch a variety of attacks with ICE. They could direct media to a target of a DoS attack, they could insert themselves into the media stream, and so on. These are similar to the general security considerations for offer/answer exchanges, and the security considerations in RFC 3264 [4] apply. These require techniques for message integrity and encryption for offers and answers, which are satisfied by the SIPS mechanism [3] when SIP is used. As such, the usage of SIPS with ICE is RECOMMENDED.

15.4. Insider Attacks

In addition to attacks where the attacker is a third party trying to insert fake offers, answers or stun messages, there are several attacks possible with ICE when the attacker is an authenticated and valid participant in the ICE exchange.

15.4.1. The Voice Hammer Attack

The voice hammer attack is an amplification attack. In this attack, the attacker initiates sessions to other agents, and includes the IP address and port of a DoS target in the m/c-line of their SDP. This causes substantial amplification; a single offer/answer exchange can create a continuing flood of media packets, possibly at high rates (consider video sources). This attack is not specific to ICE, but ICE can help provide remediation.

Specifically, if ICE is used, the agent receiving the malicious SDP will first peform connectivity checks to the target of media before sending it there. If this target is a third party host, the checks will not succeed, and media is never sent.

Unfortunately, ICE doesn't help if its not used, in which case an attacker could simply send the offer without the ICE parameters. However, in environments where the set of clients are known, and limited to ones that support ICE, the server can reject any offers or answers that don't indicate ICE support.

15.4.2. STUN Amplification Attack

The STUN amplification attack is similar to the voice hammer. However, instead of voice packets being directed to the target, STUN connectivity checks are directed to the target. This attack is accomplished by having the offerer send an offer with a large number of candidates, say 50. The answerer receives the offer, and starts its checks, which are directed at the target, and consequently, never generate a response. The answerer will start a new connectivity check every 50ms, and each check is a STUN transaction consisting of 9 retransmits of a message 65 bytes in length (plus 28 bytes for the IP/UDP header) that runs for 7.9 seconds, for a total of 105 bytes/ second per transaction on average. In the worst case, there can be 158 transactions in progress at once (7.9 seconds divided by 50ms), for a total of 132 kbps, just for STUN requests.

It is impossible to eliminate the amplification, but the volume can be reduced through a variety of heuristics. For example, agents can limit the number of candidates they'll accept in an offer or answer, they can increase the value of Ta, or exponentially increase Ta as

time goes on. All of these ultimately trade off the time for the ICE exchanges to complete, with the amount of traffic that gets sent.

OPEN ISSUE: Need better remediation for this. Especially an issue if we reduce Ta to be as fast as media packets themselves, in which case this attack is as equally devastating as the voice hammer.

16. IANA Considerations

This specification defines four new SDP attributes per the procedures of Section 8.2.4 of [10]. The required information for the registrations are included here.

16.1. candidate Attribute

Contact Name: Jonathan Rosenberg, jdrosen@jdrosen.net.

Attribute Name: candidate

Long Form: candidate

Type of Attribute: media level

Charset Considerations: The attribute is not subject to the charset attribute.

Purpose: This attribute is used with Interactive Connectivity Establishment (ICE), and provides one of many possible candidate addresses for communication. These addresses are validated with an end-to-end connectivity check using Simple Traversal Underneath NAT (STUN).

Appropriate Values: See Section 13 of RFC XXXX [Note to RFC-ed: please replace XXXX with the RFC number of this specification].

16.2. remote-candidates Attribute

Contact Name: Jonathan Rosenberg, jdrosen@jdrosen.net.

Attribute Name: remote-candidates

Long Form: remote-candidates

Type of Attribute: media level

Charset Considerations: The attribute is not subject to the charset attribute.

Purpose: This attribute is used with Interactive Connectivity Establishment (ICE), and provides the identity of the remote candidates that the offerer wishes the answerer to use in its answer.

Appropriate Values: See Section 13 of RFC XXXX [Note to RFC-ed: please replace XXXX with the RFC number of this specification].

16.3. ice-pwd Attribute

Contact Name: Jonathan Rosenberg, jdrosen@jdrosen.net.

Attribute Name: ice-pwd

Long Form: ice-pwd

Type of Attribute: session or media level

Charset Considerations: The attribute is not subject to the charset attribute.

Purpose: This attribute is used with Interactive Connectivity Establishment (ICE), and provides the password used to protect STUN connectivity checks.

Appropriate Values: See Section 13 of RFC XXXX [Note to RFC-ed: please replace XXXX with the RFC number of this specification].

16.4. ice-ufrag Attribute

Contact Name: Jonathan Rosenberg, jdrosen@jdrosen.net.

Attribute Name: ice-ufrag

Long Form: ice-ufrag

Type of Attribute: session or media level

Charset Considerations: The attribute is not subject to the charset attribute.

Purpose: This attribute is used with Interactive Connectivity Establishment (ICE), and provides the fragments used to construct the username in STUN connectivity checks.

Appropriate Values: See Section 13 of RFC XXXX [Note to RFC-ed: please replace XXXX with the RFC number of this specification].

17. IAB Considerations

The IAB has studied the problem of "Unilateral Self Address Fixing", which is the general process by which a agent attempts to determine its address in another realm on the other side of a NAT through a collaborative protocol reflection mechanism [19]. ICE is an example of a protocol that performs this type of function. Interestingly, the process for ICE is not unilateral, but bilateral, and the difference has a signficant impact on the issues raised by IAB. Indeed, ICE can be considered a B-SAF (Bilateral Self-Address Fixing) protocol, rather than an UNSAF protocol. Regardless, the IAB has mandated that any protocols developed for this purpose document a specific set of considerations. This section meets those requirements.

17.1. Problem Definition

From RFC 3424 any UNSAF proposal must provide:

> Precise definition of a specific, limited-scope problem that is to be solved with the UNSAF proposal. A short term fix should not be generalized to solve other problems; this is why "short term fixes usually aren't".

The specific problems being solved by ICE are:

> Provide a means for two peers to determine the set of transport addresses which can be used for communication.
>
> Provide a means for resolving many of the limitations of other UNSAF mechanisms by wrapping them in an additional layer of processing (the ICE methodology).
>
> Provide a means for a agent to determine an address that is reachable by another peer with which it wishes to communicate.

17.2. Exit Strategy

From RFC 3424, any UNSAF proposal must provide:

> Description of an exit strategy/transition plan. The better short term fixes are the ones that will naturally see less and less use as the appropriate technology is deployed.

ICE itself doesn't easily get phased out. However, it is useful even in a globally connected Internet, to serve as a means for detecting whether a router failure has temporarily disrupted connectivity, for example. ICE also helps prevent certain security attacks which have nothing to do with NAT. However, what ICE does is help phase out other UNSAF mechanisms. ICE effectively selects amongst those mechanisms, prioritizing ones that are better, and deprioritizing ones that are worse. Local IPv6 addresses can be preferred. As NATs begin to dissipate as IPv6 is introduced, server reflexive and relayed candidates (both forms of UNSAF mechanisms) simply never get used, because higher priority connectivity exists to the native host candidates. Therefore, the servers get used less and less, and can eventually be remove when their usage goes to zero.

Indeed, ICE can assist in the transition from IPv4 to IPv6. It can be used to determine whether to use IPv6 or IPv4 when two dual-stack hosts communicate with SIP (IPv6 gets used). It can also allow a network with both 6to4 and native v6 connectivity to determine which address to use when communicating with a peer.

17.3. Brittleness Introduced by ICE

From RFC3424, any UNSAF proposal must provide:

> Discussion of specific issues that may render systems more "brittle". For example, approaches that involve using data at multiple network layers create more dependencies, increase debugging challenges, and make it harder to transition.

ICE actually removes brittleness from existing UNSAF mechanisms. In particular, traditional STUN (as described in RFC 3489 [13]) has several points of brittleness. One of them is the discovery process which requires a agent to try and classify the type of NAT it is behind. This process is error-prone. With ICE, that discovery process is simply not used. Rather than unilaterally assessing the validity of the address, its validity is dynamically determined by measuring connectivity to a peer. The process of determining connectivity is very robust.

Another point of brittleness in traditional STUN and any other unilateral mechanism is its absolute reliance on an additional server. ICE makes use of a server for allocating unilateral addresses, but allows agents to directly connect if possible. Therefore, in some cases, the failure of a STUN server would still

allow for a call to progress when ICE is used.

Another point of brittleness in traditional STUN is that it assumes that the STUN server is on the public Internet. Interestingly, with ICE, that is not necessary. There can be a multitude of STUN servers in a variety of address realms. ICE will discover the one that has provided a usable address.

The most troubling point of brittleness in traditional STUN is that it doesn't work in all network topologies. In cases where there is a shared NAT between each agent and the STUN server, traditional STUN may not work. With ICE, that restriction is removed.

Traditional STUN also introduces some security considerations. Fortunately, those security considerations are also mitigated by ICE.

Consequently, ICE serves to repair the brittleness introduced in other UNSAF mechanisms, and does not introduce any additional brittleness into the system.

17.4. Requirements for a Long Term Solution

From RFC 3424, any UNSAF proposal must provide:

> Identify requirements for longer term, sound technical solutions -- contribute to the process of finding the right longer term solution.

Our conclusions from STUN remain unchanged. However, we feel ICE actually helps because we believe it can be part of the long term solution.

17.5. Issues with Existing NAPT Boxes

From RFC 3424, any UNSAF proposal must provide:

> Discussion of the impact of the noted practical issues with existing, deployed NA[P]Ts and experience reports.

A number of NAT boxes are now being deployed into the market which try and provide "generic" ALG functionality. These generic ALGs hunt for IP addresses, either in text or binary form within a packet, and rewrite them if they match a binding. This interferes with traditional STUN. However, the update to STUN [11] uses an encoding which hides these binary addresses from generic ALGs. Since [11] is required for all ICE implementations, this NAPT problem does not impact ICE.

Existing NAPT boxes have non-deterministic and typically short expiration times for UDP-based bindings. This requires implementations to send periodic keepalives to maintain those bindings. ICE uses a default of 15s, which is a very conservative estimate. Eventually, over time, as NAT boxes become compliant to behave [30], this minimum keepalive will become deterministic and well-known, and the ICE timers can be adjusted. Having a way to discover and control the minimum keepalive interval would be far better still.

18. Acknowledgements

The authors would like to thank Flemming Andreasen, Rohan Mahy, Dean Willis, Eric Cooper, Dan Wing, Douglas Otis, Tim Moore, and Francois Audet for their comments and input. A special thanks goes to Bill May, who suggested several of the concepts in this specification, Philip Matthews, who suggested many of the key performance optimizations in this specification, Eric Rescorla, who drafted the text in the introduction, and Magnus Westerlund, for doing several detailed reviews on the various revisions of this specification.

19. References

19.1. Normative References

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

[2] Huitema, C., "Real Time Control Protocol (RTCP) attribute in Session Description Protocol (SDP)", RFC 3605, October 2003.

[3] 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.

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

[5] Casner, S., "Session Description Protocol (SDP) Bandwidth Modifiers for RTP Control Protocol (RTCP) Bandwidth", RFC 3556, July 2003.

[6] Camarillo, G., Marshall, W., and J. Rosenberg, "Integration of Resource Management and Session Initiation Protocol (SIP)", RFC 3312, October 2002.

[7] Camarillo, G. and P. Kyzivat, "Update to the Session Initiation Protocol (SIP) Preconditions Framework", RFC 4032, March 2005.

[8] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", RFC 4234, October 2005.

[9] Rosenberg, J. and H. Schulzrinne, "Reliability of Provisional Responses in Session Initiation Protocol (SIP)", RFC 3262, June 2002.

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

[11] Rosenberg, J., "Simple Traversal Underneath Network Address Translators (NAT) (STUN)", draft-ietf-behave-rfc3489bis-04 (work in progress), July 2006.

[12] Rosenberg, J., "Obtaining Relay Addresses from Simple Traversal of UDP Through NAT (STUN)", draft-ietf-behave-turn-01 (work in progress), June 2006.

19.2. Informative References

[13] Rosenberg, J., Weinberger, J., Huitema, C., and R. Mahy, "STUN - Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)", RFC 3489, March 2003.

[14] Senie, D., "Network Address Translator (NAT)-Friendly Application Design Guidelines", RFC 3235, January 2002.

[15] Srisuresh, P., Kuthan, J., Rosenberg, J., Molitor, A., and A. Rayhan, "Middlebox communication architecture and framework", RFC 3303, August 2002.

[16] Rosenberg, J., Peterson, J., Schulzrinne, H., and G. Camarillo, "Best Current Practices for Third Party Call Control (3pcc) in the Session Initiation Protocol (SIP)", BCP 85, RFC 3725, April 2004.

[17] Borella, M., Lo, J., Grabelsky, D., and G. Montenegro, "Realm Specific IP: Framework", RFC 3102, October 2001.

[18] Borella, M., Grabelsky, D., Lo, J., and K. Taniguchi, "Realm Specific IP: Protocol Specification", RFC 3103, October 2001.

[19] Daigle, L. and IAB, "IAB Considerations for UNilateral Self-Address Fixing (UNSAF) Across Network Address Translation", RFC 3424, November 2002.

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

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

[22] Carpenter, B. and K. Moore, "Connection of IPv6 Domains via IPv4 Clouds", RFC 3056, February 2001.

[23] Zopf, R., "Real-time Transport Protocol (RTP) Payload for Comfort Noise (CN)", RFC 3389, September 2002.

[24] Rosenberg, J., "The Session Initiation Protocol (SIP) UPDATE Method", RFC 3311, October 2002.

[25] Camarillo, G. and H. Schulzrinne, "Early Media and Ringing Tone Generation in the Session Initiation Protocol (SIP)", RFC 3960, December 2004.

[26] Andreasen, F., "Connectivity Preconditions for Session Description Protocol Media Streams", draft-ietf-mmusic-connectivity-precon-02 (work in progress), June 2006.

[27] Andreasen, F., "A No-Op Payload Format for RTP", draft-ietf-avt-rtp-no-op-00 (work in progress), May 2005.

[28] Kohler, E., Handley, M., and S. Floyd, "Datagram Congestion Control Protocol (DCCP)", RFC 4340, March 2006.

[29] Hellstrom, G. and P. Jones, "RTP Payload for Text Conversation", RFC 4103, June 2005.

[30] Audet, F. and C. Jennings, "NAT Behavioral Requirements for Unicast UDP", draft-ietf-behave-nat-udp-07 (work in progress), June 2006.

[31] Jennings, C. and R. Mahy, "Managing Client Initiated Connections in the Session Initiation Protocol (SIP)", draft-ietf-sip-outbound-04 (work in progress), June 2006.

Appendix A. Design Motivations

ICE contains a number of normative behaviors which may themselves be simple, but derive from complicated or non-obvious thinking or use

cases which merit further discussion. Since these design motivations are not neccesary to understand for purposes of implementation, they are discussed here in an appendix to the specification. This section is non-normative.

A.1. Applicability to Gateways and Servers

Section 4.1 discusses procedures for gathering candidates, including host, server reflexive and relayed. In that section, recommendations are given for when an agent should obtain each of these three types. In particular, for agents embedded in PSTN gateways, media servers, conferencing servers, and so on, ICE specifies that an agent can stick with just host candidates, since it has a public IP address.

This leads to an important question - why would such an endpoint even bother with ICE? If it has a public IP address, what additional value do the ICE procedures bring? There are many, actually.

First, doing so greatly facilitates NAT traversal for clients that connect to it. Consider a PC softphone behind a NAT whose mapping policy is address and port dependent. The softphone initiates a call through a gateway that implements ICE. The gateway doesn't obtain any server reflexive or relayed candidates, but it implements ICE, and consequently, is prepared to receive STUN connectivity checks on its host candidates. The softphone will send a STUN connectivity check to the gateway, which passes through the intervending NAT. This causes the NAT to allocate a new binding for the softphone. The connectivity is received by the gateway, and will cause it gateway to send a check back to the softphone, at this newly created candidate. A successful response confirms that this candidate is usable, and the gateway can send media immediately to the softphone. This allows direct media transmission between the gateway and softphone, without the need for relays, even though the softphone was behind a 'bad' NAT.

Second, implementation of the STUN connectivity checks allows for NAT bindings along the way to be kept open. Keeping these bindings open is essential for continued communications between the gateway and softphone.

Third, ICE prevents a fairly destructive attack in multimedia systems, called the voice hammer. The STUN connectivity check used by an ICE endpoint allows it to be certain that the target of media packets is, in fact, the same entity that requested the packets through the offer/answer exchange. See Section 15 for a more complete discussion on this attack.

A.2. Pacing of STUN Transactions

STUN transactions used to gather candidates and to verify connectivity are paced out at an approximate rate of one new transaction every Ta seconds, where Ta has a default of 50ms. Why are these transactions paced, and why was 50ms chosen as default?

Sending of these STUN requests will often have the effect of creating bindings on NAT devices between the client and the STUN servers. Experience has shown that many NAT devices have upper limits on the rate at which they will create new bindings. Furthermore, transmission of these packets on the network makes use of bandwidth and needs to be rate limited by the agent. As a consequence, the pacing ensures that the NAT devices does not get overloaded and that traffic is kept at a reasonable rate.

Another aspect of the STUN requests is their bandwidth usage. In ICE, each STUN request contains the STUN 20 byte header, in addition to the USERNAME, MESSAGE-INTEGRITY and PRIORITY attributes. The USERNAME attribute contains a 4-byte attribute overhead, plus the username value itself. This username is the concatenation of the two fragments, plus a colon. Each fragment is supposed to be at least 4 bytes long, making the total length of the USERNAME attribute (4*2 + 1 + 4) = 13 bytes. The MESSAGE-INTEGRITY attribute is 4 bytes of overhead plus 20 bytes value, for 24 bytes. The PRIORITY attribute is 4 bytes of overhead plus 4 bytes of value, for 8 bytes. Thus, the total length of the STUN Binding Request is (20 + 13 + 24 + 8) = 65 bytes, with 28 bytes of overhead for IP and UDP for a total of 93 bytes. The response contains the STUN 20 byte header, the XOR-MAPPED-ADDRESS, and MESSAGE-INTEGRITY attributes. XOR-MAPPED-ADDRESS has 4 bytes overhead plus an 8 byte value, for a total of 12 bytes. Thus, each STUN response is (20 + 12 + 24) = 56 bytes plus 28 bytes of UDP/IP overhead for a total of 84 bytes. Checks typically fall into one of two cases. If a check works, each transaction has a single request and a single response, for a total of 2 packets and 177 bytes over one RTT interval. Assuming a fairly agressive RTT of 70ms, this produces 20.23 kbps, but only briefly. If a check fails because the pair is invalid, there will be nine requests and no responses. This produces 837 bytes over 7.9s, for a total of 105.9 bps, but over a long period of time.

> OPEN ISSUE: The bandwidth computations are pretty complex because ICE is not a CBR stream, and its bandwidth utilization depends on how many transactions it ends up generating before it finishes. Need to work this model more.

Given that these numbers are close to, if not greater than, the bandwidths utilized by many voice codecs, this seems a reasonable

value to use.

OPEN ISSUE: There is some debate about whether to reduce this pacing interval smaller, say 20ms, to speed up ICE, or perhaps make it equal to the bandwidth that would be utilized by the media streams themselves.

A.3. Candidates with Multiple Bases

Section 4.1 talks about merging together candidates that are identical but have different bases. When can an agent have two candidates that have the same IP address and port, but different bases? Consider the topology of Figure 16:

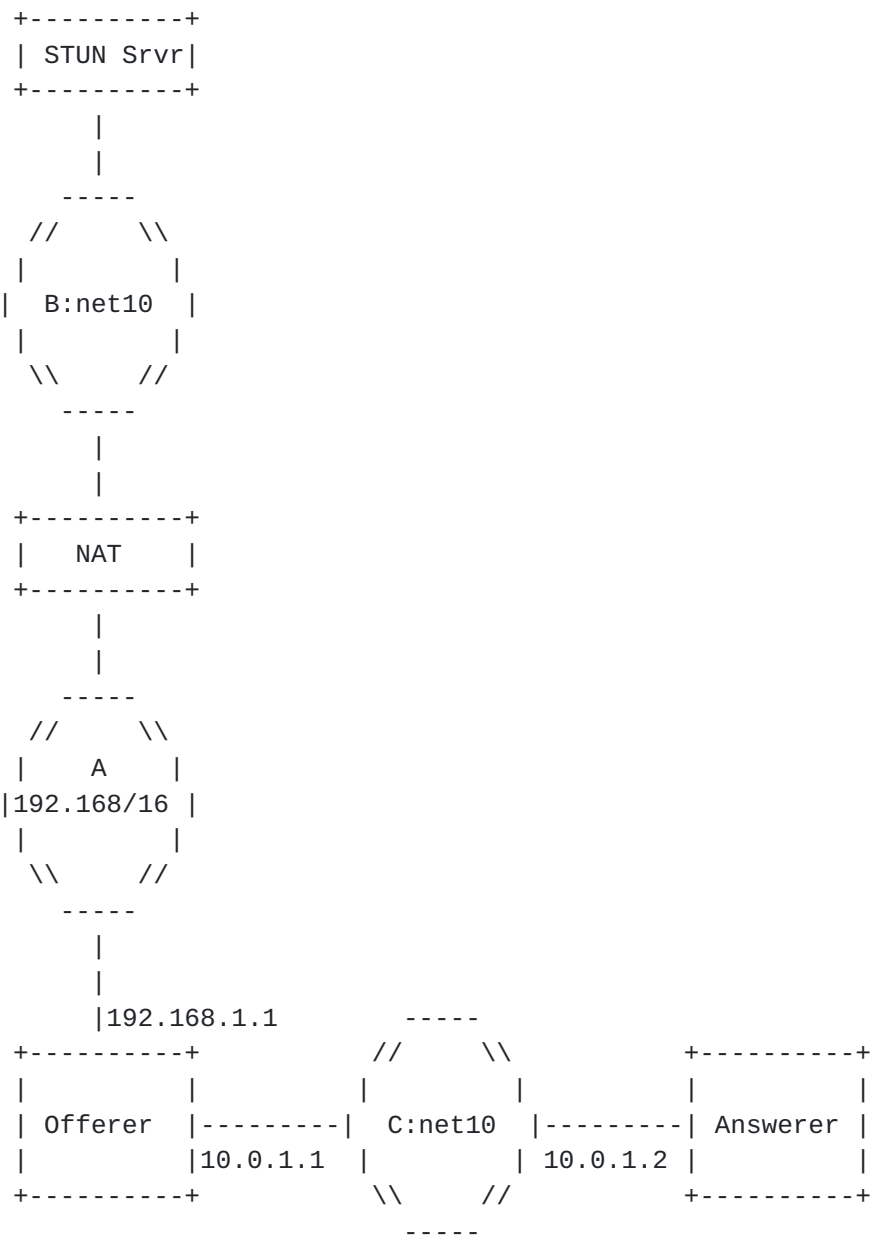


Figure 16

In this case, the offerer is multi-homed. It has one interface, 10.0.1.1, on network C, which is a net 10 private network. The Answerer is on this same network. The offerer is also connected to network A, which is 192.168/16. The offerer has an interface of 192.168.1.1 on this network. There is a NAT on this network, natting into network B, which is another net10 private network, but not connected to network C. There is a STUN server on network B.

The offerer obtains a host candidate on its interface on network C

(10.0.1.1:2498) and a host candidate on its interface on network A (192.168.1.1:3344). It performs a STUN query to its configured STUN server from 192.168.1.1:3344. This query passes through the NAT, which happens to assign the binding 10.0.1.1:2498. The STUN server reflects this in the STUN Binding Response. Now, the offerer has obtained a server reflexive candidate with a transport address that is identical to a host candidate (10.0.1.1:2498). However, the server reflexive candidate has a base of 192.168.1.1:3344, and the host candidate has a base of 10.0.1.1:2498.

A.4. Purpose of the Translation

When a candidate is relayed, the SDP offer or answer contain both the relayed candidate and its translation. However, the translation is never used by ICE itself. Why is it present in the message?

There are two motivations for its inclusion. The first is diagnostic. It is very useful to know the relationship between the different types of candidates. By including the translation, an agent can know which relayed candidate is associated with which reflexive candidate, which in turn is associated with a specific host candidate. When checks for one candidate succeed and not the others, this provides useful diagnostics on what is going on in the network.

The second reason has to do with off-path Quality of Service (QoS) mechanisms. When ICE is used in environments such as PacketCable 2.0 [[TODO: need PC2.0 reference]], proxies will, in addition to performing normal SIP operations, inspect the SDP in SIP messages, and extract the IP address and port for media traffic. They can then interact, through policy servers, with access routers in the network, to establish guaranteed QoS for the media flows. This QoS is provided by classifying the RTP traffic based on 5-tuple, and then providing it a guaranteed rate, or marking its Diffserv codepoints appropriately. When a residential NAT is present, and a relayed candidate gets selected for media, this relayed candidate will be a transport address on an actual STUN relay. That address says nothing about the actual transport address in the access router that would be used to classify packets for QoS treatment. Rather, the translation of that relayed address is needed. By carrying the translation in the SDP, the proxy can use that transport address to request QoS from the access router.

A.5. Importance of the STUN Username

ICE requires the usage of message integrity with STUN using its short term credential functionality. The actual short term credential is formed by exchanging username fragments in the SDP offer/answer exchange. The need for this mechanism goes beyond just security; it

is actual required for correct operation of ICE in the first place.

Consider agents A, B, and C. A and B are within private enterprise 1, which is using 10.0.0.0/8. C is within private enterprise 2, which is also using 10.0.0.0/8. As it turns out, B and C both have IP address 10.0.1.1. A sends an offer to C. C, in its answer, provides A with its host candidates. In this case, those candidates are 10.0.1.1:8866 and 10.0.1.1:8877. As it turns out, B is in a session at that same time, and is also using 10.0.1.1:8866 and 10.0.1.1:8877 as host candidates. This means that B is prepared to accept STUN messages on those ports, just as C is. A will send a STUN request to 10.0.1.1:8866 and and another to 10.0.1.1:8877. However, these do not go to C as expected. Instead, they go to B! If B just replied to them, A would believe it has connectivity to C, when in fact it has connectivity to a completely different user, B. To fix this, the STUN short term credential mechanisms are used. The username fragments are sufficiently random that it is highly unlikely that B would be using the same values as A. Consequently, B would reject the STUN request since the credentials were invalid. In essence, the STUN username fragments provide a form of transient host identifiers, bound to a particular offer/answer session.

An unfortunate consequence of the non-uniqueness of IP addresses is that, in the above example, B might not even be an ICE agent. It could be any host, and the port to which the STUN packet is directed could be any ephemeral port on that host. If there is an application listening on this socket for packets, and it is not prepared to handle malformed packets for whatever protocol is in use, the operation of that application could be affected. Fortunately, since the ports exchanged in SDP are ephemeral and usually drawn from the dynamic or registered range, the odds are good that the port is not used to run a server on host B, but rather is the agent side of some protocol. This decreases the probability of hitting a port in-use, due to the transient nature of port usage in this range. However, the possibility of a problem does exist, and network deployers should be prepared for it. Note that this is not a problem specific to ICE; stray packets can arrive at a port at any time for any type of protocol, especially ones on the public Internet. As such, this requirement is just restating a general design guideline for Internet applications - be prepared for unknown packets on any port.

A.6. The Candidate Pair Sequence Number Formula

The sequence number for a candidate pair has an odd form. It is:

PAIR-SN = 10000*MAX(O-SN,A-SN) + MIN(O-SN,A-SN) + O-IP/SZ

Why is this? When the candidate pairs are sorted based on this

value, the resulting sorting has the MAX/MIN property. This means that the pairs are first sorted based on increasing value of the maximum of the two sequence numbers. For pairs that have the same value of the maximum sequence number, the minimum sequence number is used to sort amongst them. If the max and the min sequence numbers are the same, the IP address of the offerers candidate serves as a tie breaker. The factor of 1000 is used since there will always be fewer than a 1000 candidates, and thus the largest value a sequence number (and thus the minimum sequence number) can have is always less than 1000. This creates the desired sorting property.

Recall that candidate sequence numbers are assigned such that, for a particular set of candidates of the same type, the RTP components have lower sequence numbers than the corresponding RTCP component. Also recall that, if an agent prefers host candidates to server reflexive to relayed, sequence numbers for host candidates are always lower than server reflexive which are always lower than relayed. Because of this,

A.7. The Frozen State

The Frozen state is used for two purposes. Firstly, it allows ICE to first perform checks for the first component of a media stream. Once a successful check has completed for the first component, the other components of the same type and local preference will get performed. Secondly, when there are multiple media streams, it allows ICE to first check candidates for a single media stream, and once a set of candidates has been found, candidates of that same type for other media streams can be checked first. This effectively 'caches' the results of a check for one media stream, and applies them to another. For example, if only the relayed candidates for audio (which were the last resort candidates) succeed, ICE will check the relayed candidates for video first.

A.8. The remote-candidates attribute

The a=remote-candidates attribute exists to eliminate a race condition between the updated offer and the response to the STUN Binding Request that moved a candidate into the Valid list. This race condition is shown in Figure 17. On receipt of message 4, agent A adds a candidate pair to the valid list. If there was only a single media stream with a single component, agent A could now send an updated offer. However, the check from agent B has not yet generated a response, and agent B receives the updated offer (message 7) before getting the response (message 10). Thus, it does not yet know that this particular pair is valid. To eliminate this condition, the actual candidates at B that were selected by the offerer (the remote candidates) are included in the offer itself.

Note, however, that agent B will not send media until it has received this STUN response.

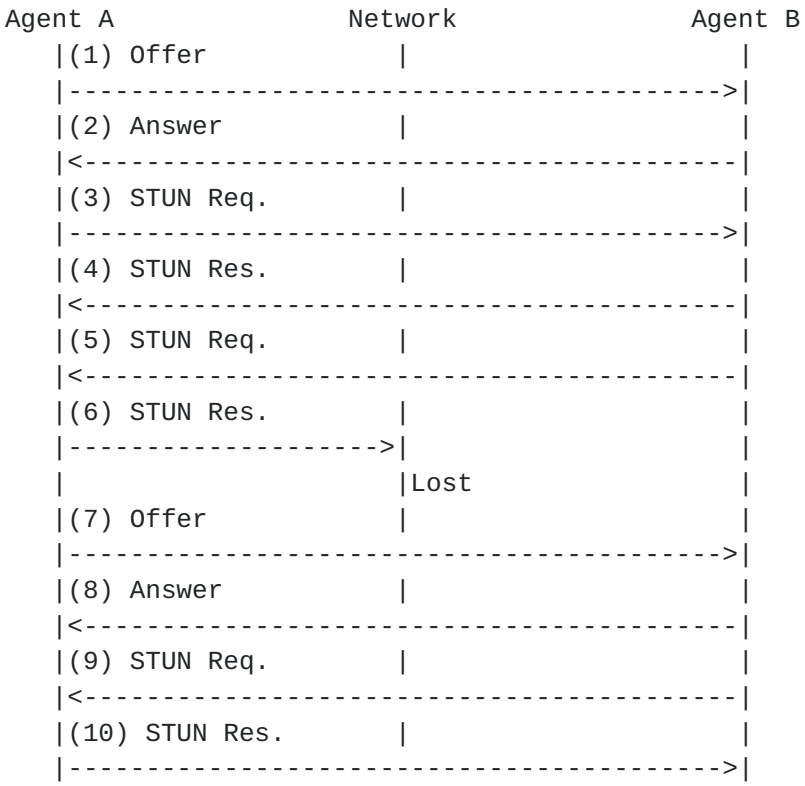


Figure 17

A.9. Why are Keepalives Needed?

Once media begins flowing on a candidate pair, it is still necessary to keep the bindings alive at intermediate NATs for the duration of the session. Normally, the media stream packets themselves (e.g., RTP) meet this objective. However, several cases merit further discussion. Firstly, in some RTP usages, such as SIP, the media streams can be "put on hold". This is accomplished by using the SDP "sendonly" or "inactive" attributes, as defined in RFC 3264 [4]. RFC 3264 directs implementations to cease transmission of media in these cases. However, doing so may cause NAT bindings to timeout, and media won't be able to come off hold.

Secondly, some RTP payload formats, such as the payload format for text conversation [29], may send packets so infrequently that the interval exceeds the NAT binding timeouts.

Thirdly, if silence suppression is in use, long periods of silence

may cause media transmission to cease sufficiently long for NAT bindings to time out.

For these reasons, the media packets themselves cannot be relied upon. ICE defines a simple periodic keepalive that operates indpendently of media transmission. This makes its bandwidth requirements highly predictable, and thus amenable to QoS reservations.

A.10. Why Prefer Peer Reflexive Candidates?

Section 4.2 describes procedures for computing the priority of candidate based on its type and local preferences. That section requires that the type preference for peer reflexive candidates always be lower than server reflexive. Why is that? The reason has to do with the security considerations in Section 15. It is much easier for an attacker to cause an agent to use a false server reflexive candidate than it is for an attacker to cause an agent to use a false peer reflexive candidate. Consequently, attacks against the STUN binding discovery usage are thwarted by ICE by preferring the peer reflexive candidates.

A.11. Why Can't Offerers Send Media When a Pair Validates

Section 11.1 describes rules for sending media. The rules are asymmetric, and not the same for offerers and answerers. In particular, an answerer can send media right away to a candidate pair once it validates, even if it doesnt match the pairs in the m/c-line. THe offerer cannot - it must wait for an updated offer/answer exchange. Why is that?

This, in fact, relates to a bigger question - why is the updated offer/answer exchange needed at all? Indeed, in a pure offer/answer environment, it would not be. The offerer and answerer will agree on the candidates to use through ICE, and then can begin using them. As far as the agents themselves are concerned, the updated offer/answer provides no new information. However, in practice, numerous components along the signaling path look at the SDP information. These include entities performing off-path QoS reservations, NAT traversal components such as ALGs and Session Border Controllers (SBCs) and diagnostic tools that passively monitor the network. For these tools to continue to function without change, the core property of SDP - that the m/c-lines represent the addresses used for media - must be retained. For this reason, an updated offer must be sent.

To ensure that an updated offerer is sent, ICE purposefully prevents the offerer from sending media until that offer is sent. It furthermore restricts the answerer in how long it can send media

until an updated offer is received. This provides protocol incentives for sending the updated offer.

The updated offer also helps ensure that ICE did the right thing. In very unusual cases, the offerer and answerer might not agree on the candidates selected by ICE. This would be detected in the updated offer/answer exchange, allowing them to restart ICE procedures to fix the problem.

Author's Address

Jonathan Rosenberg
Cisco Systems
600 Lanidex Plaza
Parsippany, NJ 07054
US

Phone: +1 973 952-5000
Email: jdrosen@cisco.com
URI: http://www.jdrosen.net

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

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