Transport Options for Clue
draft-wenger-clue-transport-02

Abstract

This memo describes the assumption and the proposed options for the coding and transport of CLUE messages as outlined in version 01 of the framework draft.

Requirements Language

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

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 13, 2012.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.
1. Introduction

The CLUE WG is chartered to design a protocol to enable communication about media streams for videoconferencing and telepresence working in conjunction with the IETF's protocol suites of choice, namely SIP for basic call setup and control and RTP for media transport. (It should be noted that ITU-T Q.xx/16 has informally expressed a desire that parts or all of the work of the CLUE WG can be re-used in an H.323 environment. Therefore, occasionally, we comment on the re-use of CLUE work outside of SIP systems. This does not mean that we want to extent the charter; however, it seems sensible at least to us that if a cross protocol solution and a SIP-only solution to the CLUE problem could be devised, and both solutions are comparable in in their complexity etc etc, a solution with applicability beyond SIP may be the appropriate choice.)

This document describes options for the coding and transport of CLUE messages in a SIP / RTP environment. Specifically, three issues are addressed.

First, while the framework draft conceptually describes message flows, it does not specify how those messages are actually transferred "on the wire" and how they relate to the SIP offer/answer [RFC3264]. This document lists (hopefully all) the options that have been proposed in CLUE to date.

Second, the framework-01 draft describes three messages between the producer and the consumer in an abstract form, without specifying the details of the representation of those messages. This memo lists (some of) the options for the representation of the abstract messages of the framework draft.

Third, before any CLUE messages can be meaningfully exchanged, it is necessary to discover whether the involved systems are actually CLUE-capable. This memo discusses the proposed options for CLUE capability discovery.

In this memo we only present the options discussed to date in the working group. Deciding on the appropriate mechanism (or mechanisms, as it is not always appropriate to have a single solution for a given problem, though this is of course desirable from an interoperability viewpoint) is left for further discussion in the working group. That does not mean that the authors do not have preferences, and/or specific knowledge of certain mechanisms, and may as a result go in greater depth in describing one mechanism than another.
2. Assumptions

The Basic Clue data model is specified in the framework document. The framework defines three messages that carry the Clue data:

Consumer Capability Message

Provider Capabilities Announcement

Consumer Configure Request

(There is no clear consensus that the Consumer Capability Message is needed, but for the time being we attempt to document how it fits in the different options.)

CLUE messages may need to be sent at the initialization of a call, and possibly also at irregular intervals within a call, spaced in the order of seconds, minutes or even longer. There is also no hard real-time transmission requirement for CLUE messages; latencies in the seconds range are acceptable. More specifically, there appears not to be an issue with system reaction delay larger than the maximum round trip delay for reasonable operation of a telepresence system.

The Clue message handshake as required by the framework (independent from the issue related to the need of the Consumer Capability Message) is different from the offer/answer (o/A) exchange [RFC3264], primarily because the CLUE exchange is uni-directional, requiring a similar exchange for each side of the media flow, while one offer/answer exchange defines both sides of the media flow. (Note that asymmetry in SIP may require a second offer answer exchange, but this is not the typical case)

There is no hard requirement for synchronization of CLUE messages, though there may be a need for sequencing, (TBD).

CLUE messages may need to describe the characteristics of all endpoints in a conference (TBD), and that conference can potentially include dozens of endpoints.

It appears to be consensus within the CLUE WG that there will be an SDP offer/answer exchange as part of the solution. It further appears to be the consensus that the offer/answer will be used to establish the media channels and negotiate those SDP parameters negotiable with media types (i.e. as defined in RTP payload formats), as well as to allow interoperability with systems that do not support the CLUE protocol. It appears to be a sensible design goal that the CLUE data does not duplicate SDP attributes.
In order to achieve interoperability with systems that do not support CLUE, the first offer answer exchange could be used to negotiate CLUE support.

An open issue is whether there needs to be a final offer answer exchange, after initial o/A exchange(s) as well as CLUE exchange(s), with an SDP reflecting the negotiated media flows, in order to address requirements imposed by intermediaries like Session Border Controllers (SBC). This topic was discussed in different contexts before, and there is some text about it in RFC5939 section 3.12 [RFC5939]

The size of a CLUE message is far from final yet but when selecting a solution the issue of message size and fragmentation (if applicable) needs to be addressed.

3. Transport for CLUE messages

CLUE messages need to be conveyed from one CLUE capable system to another, i.e., there needs to be "transport" of CLUE messages. It should be clear that the message transport can be based on a transport layer (layer 4 in ISO/OSI) protocol or other layers, such as the application layer.

In contrast to the "content representation", the transport of CLUE messages is somewhat more tightly bound to the environment. In some scenarios it may be possible to reuse most of the mechanisms defined in an option for transport between SIP and H.323, while in others this is not possible.

The selection of the transport may have some affect on the content representation, in that certain transports in the aforementioned sense are defined only to carry certain types of messages. For example, offer-answer is defined for the use in conjunction with SDP as content representation. In contrast, obviously, a CLUE-defined transport mechanism could carry any format specified by CLUE.

The CLUE protocol enables the CLUE systems to negotiate the semantic relationships of the media streams, mostly with respect to spatial relations. Another aspect that has recently risen to prominence is the negotiation of media codec settings, taking into account that in practical telepresence systems, certain combinations of codec settings may not be supported by the hardware ("codec alternatives" henceforth).

The apparently generally agreed need for interoperability with non CLUE systems requires defining an initial offer involving CLUE
support, and guidance on how to progress the call setup based on the answer. The CLUE WG discussed a couple of options including two stage offer answer, using grouping similar to [I-D.ietf-mmusic-sdp-bundle-negotiation], and using the capability negotiation of [RFC5939].

We would like to consider the following options:

3.1. Option 1 : Piggy-pack on SIP

SIP includes a number of methods that can carry (directly or through content indirection) CLUE messages. Many of these messages can be exchanged during the lifetime of a session. Piggy-packing CLUE messages on SIP has the advantage that any built-in transport and reliability mechanisms of SIP can be re-used. (Whether this is an advantage in practice is somewhat questionable, considering that the vast majority of SIP systems use UDP for the transport of SIP messages, and that their SIP messages are typically small enough to fit into an MTU—something that like is not true for some CLUE messages.) It also has the feature (advantage?) that CLUE signaling is being conveyed in the signaling plane rather than in the media plane (making things such as decomposition potentially easier and certainly more intuitive).

There are three sub-options to consider

3.1.1. Option 1.1: Using SDP (in an offer-answer context) for CLUE information

In this option, the CLUE protocol is specified through the addition of CLUE-specific SDP codepoints in the (essentially unmodified) offer/answer process, for essentially all CLUE functionalities. The stream semantics associated with spatial relations of streams are represented as new SDP attributes. Codec alternatives may be negotiated based on draft-ietf-mmusic-sdp-media-capabilities.

The nature of spatial relations currently envisioned by some CLUE participants have some simultaneous restrictions due to the limitations of physical capture devices. For this reason, it may become necessary to separate the negotiation process into a session negotiation that defines RTP sessions, and a session negotiated that deals with the spatial relations.

It is noted that, at the time of writing, there is no proposal on the table that would suggest that offer-answer only is a sensible--or even possible--design choice.
3.1.2. Option 1.2: Using an SDP MIME body to carry the CLUE information in an INVITE or UPDATE exchange

In order to separate the RTP session negotiation from the CLUE media capture selection, a clean solution appears to be to carry the CLUE information in a body separate from the classic media negotiation information, with a parallel negotiation using INVITE and UPDATE for the CLUE information. A similar approach is proposed in [I-D.ietf-siprec-protocol].

There were concerns about using re-invite, claiming that it takes too long since that commonly implies codec boxes teardown of every existing media session during re-invites. [RFC3311] suggests that although UPDATE can be used on confirmed dialogs, it is RECOMMENDED that a re-INVITE be used instead. This is because an UPDATE needs to be answered immediately, ruling out the possibility of user approval. Such approval may be needed, and is possible only with a re-INVITE.

3.1.3. Option 1.3: Using a SIP INFO package

Another option may be to define a new SIP INFO package [RFC6086]. The SIP-INFO method is very flexible in that the package can define, at least to a large extent, the semantics of a SIP-INFO exchange. However, SIP-INFO is subject to SIPoPs limitations, for example in terms of message size when SIP messages are transported over UDP (which, we understand, is the common operation point).

3.1.4. Option 1.4: SIP signaling options

There may be other options using SIP signaling, such as subscribe/notify or Message method, see [RFC6086] section 8.4.1. Note that, in those cases, a subscribe creates a separate dialog usage and is normally sent outside of existing dialog. Within this document, we are not discussing the implications of such a possible implementation path.

3.2. Option 2: CLUE control channel on the media plane over UDP

During the initial SIP handshake, a secure(?) CLUE channel is established (if both systems are CLUE capable). This channel may be UDP or TCP based. Using UDP may require an additional reliability mechanism, perhaps using a mechanism similar to BFCP over UDP, and addressing fragmentation is likely to be necessary due to message size. These complications are not required for a TCP based solution. On the other hand, using ICE to address firewall and NAT traversal as well as working with intermediaries like SBCs works better with UDP. Note that even under this option, we assume that the actual protocol exchange to negotiate and open media channels is being conducted.
using an SDP content representation, quite possibly through a "facial" offer-answer exchange that nails down the actual media flows to be used, for the benefit of SBCs and similar middleboxes.

3.3. Option 3: Other Work

At least three other individual submissions address similar topics as this section, and the reader's attention is drawn to those. Specifically:

[I-D.hansen-clue-protocol-choices-evaluation] goes into some detail in analyzing the pros and cons of a previous version of our document. The authors arrive at a conclusion that can be summarized as that there is a need for a transport mechanism that is not based on SIP, but using a UDP session negotiated using SIP and Offer/Answer for the transport of CLUE messages. CLUE messages in this case probably ought to be interpreted narrowly in that they relate to spatial relationships and related issues, in contrast to codec parameter negotiation.

[I-D.romanow-clue-sdp-usage] arrives at a similar conclusion. The draft lists those codepoints that could be conveyed using SDP in an offer-answer setting: video properties (bandwidth and resolution), and bandwidth-related group settings. Everything else, including spatial relationship of captures, is suggested to be conveyed over a CLUE-specific protocol, conveyed over a UDP (?) session negotiated in SIP during the early (first) offer-answer exchange.

[I-D.cazeaux-clue-sip-signaling] signaling appears to advocate a solution in which SDP based O/A is used to negotiate media. The negotiated media appear to be a superset of the media later being used. CLUE specific information, such as spatial relationships, but also the details of the media sessions (including restrictions of provider content selection based on consumer capabilities), appear to be relegated to signaling conveyed over a SIP/OA negotiated CLUE channel.

All three aforementioned drafts appear to acknowledge the need for a CLUE signaling channel, possibly conveyed directly over UDP (in contrast to being conveyed over SIP-info or something similar), although these drafts vary in the degree to which they use the CLUE signaling channel.

4. Content Representation

The data model in the framework-03 draft does not include a specification of the representation of the data. Many different
representation languages, for example XML, possibly SDP, ASN.1, and others can be used, and we need to decide on one, possibly for each data structure defined in a CLUE solution (that is, for example, it's possible that some data points of CLUE can be conveyed in SDP, whereas others use XML). Depending on the transport decision, we may be restricted to certain representations for certain data structures, or we may have freedom of choice. Referring to the options suggested above, it is clear that option 3.1.1 mandates SDP for representing CLUE. However, all other options appear not to require any pre-defined choice, at least for some (though not necessarily for all) of the CLUE-defined codepoints.

One observation that has to be made at this point (described in greater detail above) is that the framework-01 draft's message exchange system requires more than one end to end exchange due to the asymmetry. Another observation is that the advertisement describes the sending options, which makes the CLUE exchange different from the offer/answer mechanism SIP videoconferencing endpoints use today. For this reason we do not think that the option in 3.1.1 is a good direction. Therefore, there appears not to be a hard requirement to use SDP exclusively for the representation of CLUE messages. For some messages, SDP may be an appropriate choice, but for others, there is no precedence: We have a freedom of choice here, which is why this section exists.

It is very well possible that even moderately complex CLUE messages may exceed MTU sizes commonly found in today's Internet. There has been discussion in CLUE of sessions with thousands of participants—a very real requirement for at least one of the authors of this draft, who routinely participates in multipoint videoconferences with 200+ participants. Even if a CLUE message can be compressed into a few bytes for each endpoint, such sessions will violate the commonly found Ethernet 1492-byte MTU. Accordingly, message transport protocols will have to be prepared to split CLUE messages into fragments, which has implications on the design complexity of those protocols. This problem is especially an issue for verbose representations, such as XML.

### 4.1. Option 1: SDP

SDP and its various extensions are used in SIP based systems for the offer/answer exchange, and, therefore, those systems include SDP parsers that could probably be extended to support CLUE messages. SDP is also a fairly compact, but still (though barely) human readable. Even though it does not appear to us to make overly much sense to use SDP for CLUE, since it will require a separate blob for describing the CLUE relations between the media captures, it still viable to use text based representation for CLUE if using any of the
options which is not 3.1.1. [I-D.romanow-clue-sdp-usage] suggests that an SDP-only representation of CLUE based parameters is an (impossible/suboptimal) bad choice. We concur. As mentioned before, though, those parameters that can reasonably be negotiated using SDP o/A (with however many round trip it takes) should in our opinion be represented in SDP. We shouldn't be in SDP-ng's business.

4.2. Option 2 : XML

XML is very flexible, and the representation of choice for many IETF technologies not bound to a certain legacy. It certainly allows for all flexibility needed to represent all CLUE messages currently considered. It also is naturally extensible in a way SDP is not. On the downside, XML is fairly verbose, which has implications on the transport. Even considering this verboseness, we believe that XML may be an appropriate representation for CLUE messages that cannot be represented in SDP.

4.3. Option 3 : ASN.1

ASN.1 is similarly flexible and extensible as XML, and (in its binary representation) fairly compact. While it is commonly used in H.323, and while the video conferencing industry certainly has access to the tools necessary to deploy ASN.1 (a major obstacle in other industries), it is not widely used by SIP implementations.

4.4. Option 4 : Clue Defined Format

It is, of course, possible that the CLUE WG defines its own format, possibly compact, possibly binary and possibly extensible representation language or format for CLUE messages.

4.5. Examples

An example or examples should be added here when possible

4.6. Proposal

The preferred solution can be XML-based for codepoints not easily (currently?) representable in SDP, and SDP based for everything else. With respect to XML's verboseness, fragmentation support in the transport protocol may be needed and the transport probably should include a fragmentation and reassembly support beyond IP fragmentation/re-assembly. Such support may require an encapsulation of the message with headers that will allow fragmentation and reassembly support.
5. **Clue Discovery**

This section summarizes ways to discover whether systems involved are CLUE-capable. For simplicity, point-to-point scenarios are assumed. Multipoint scenarios are similar since we are considering centralized conference models only.

Discovery appears to be necessarily bound to the capability exchange of the involved systems.

5.1. **Option 1 : CLUE discovery as a side effect of opening a CLUE control channel**

If, for the transport of CLUE messages (or at least a subset thereof), a media plane control channel were used (section 3.2), then the discovery of CLUE capability would be a side effect of the opening of this control channel during the initial offer/answer exchange. At this point in time, there is no proposal on the table that suggest that we can avoid a CLUE control channel.

5.2. **Option 2 : SIP Message Transport**

Very roughly speaking, if we use the INFO message for the transport of all CLUE messages, then by using the Recv-Info header field the support for the CLUE package can be signaled. If using a second MIME body the support of the MIME body in the offer answer can be used.

6. **IANA Considerations**

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

7. **Security Considerations**

Any method for bypassing NAT/Firewall protections of course brings security issues, which need to be dealt with.

8. **Acknowledgements**

The list of authors needs to grow.
9. Informative References

[I-D.cazeaux-clue-sip-signaling]
Cazeaux, S. and E. Bertin, "Requirements for Controlling multiple streams for Telepresence (CLUE) signaling.",
draft-cazeaux-clue-sip-signaling-00 (work in progress), March 2012.

[I-D.hansen-clue-protocol-choices-evaluation]
Hansen, R. and A. Romanow, "Evaluation of using SIP or an independent protocol for CLUE messaging",
draft-hansen-clue-protocol-choices-evaluation-00 (work in progress), November 2011.

[I-D.ietf-mmusic-sdp-bundle-negotiation]
Holmberg, C. and H. Alvestrand, "Multiplexing Negotiation Using Session Description Protocol (SDP) Port Numbers",
draft-ietf-mmusic-sdp-bundle-negotiation-00 (work in progress), February 2012.

[I-D.ietf-siprec-protocol]
draft-ietf-siprec-protocol-03 (work in progress), March 2012.

[I-D.romanow-clue-sdp-usage]
Romanow, A., Andreasen, F., and A. Krishna, "Investigation of Session Description Protocol (SDP) Usage for Controlling multiple streams for Telepresence (CLUE)",
draft-romanow-clue-sdp-usage-00 (work in progress), March 2012.


Internet-Draft               clue-transport                   March 2012


Authors' Addresses

Dr. Stephan Wenger
Vidyo
433 Hackensack Ave
Hackensack, NJ  07601
USA

Email: stewe@stewe.org

Marshall Eubanks
AmericaFree.TV
P.O. Box 141
Clifton, Virginia  20124
USA

Phone: +1-703-501-4376
Email: marshall.eubanks@gmail.com

Roni Even
Huawei

Email: ron.even.tlv@gmail.com

Gonzalo Camarillo
Ericsson

Email: Gonzalo.Camarillo@ericsson.com