CLUE
Internet-Draft
Intended status: Standards Track
Expires: December 6, 2012

A. Romanow
Cisco Systems
A. Pepperell
Silverflare
June 4, 2012

Data model for the CLUE Framework
draft-romanow-clue-data-model-00

Abstract

This draft is for discussion in the CLUE working group. It proposes a data model for the CLUE Framework.

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 December 6, 2012.

Copyright Notice

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

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction 3
2. Terminology 3
3. Overview 4
3.1. Data model format 4
3.2. Data model namespace 4
3.3. Data Model Structure 4
4. Data Model Definition 4
4.1. <CLUE-info> 6
4.2. <capture-description> 6
4.3. <capture-id> 6
4.4. <media-type> 6
4.5. <content-type> 6
4.6. <DERIVED> 6
4.7. <switched> 6
4.8. <spatial-description> 7
4.8.1. <point-of-capture> 7
4.8.2. <capture-area> 7
4.8.3. <NATIVE-ASPECT_RATIO> 7
4.9. <audio-channel-format> 7
4.10. <encoding-group-id> 7
4.11. <future-media-capture-attribute> 7
4.12. <capture-scene> 7
4.12.1. <capture-scene-text> 8
4.12.2. <capture-scene-language> 8
4.12.3. <capture-scene-spatial-description> 8
4.12.4. <future-capture-scene-attribute> 8
4.13. <simultaneous transmission-set> 8
4.14. <stream-description> 9
4.14.1. <encoding> 9
4.14.2. <encoding-group> 10
5. Security Considerations 10
6. IANA Considerations 11
7. References 11
7.1. Normative References 11
7.2. Informative References 11
Authors' Addresses 11

1. Introduction

This document proposes a data model for the CLUE Framework [I-D.ietf-clue-framework].

The previous suggestion for a data model described the 2 CLUE messages, provider advertisement and consumer choice message. The data model proposed here differs in that it describes the basic information, or definitions that are needed. Messages can then derived from the data model, similarly to what was described in the earlier data model.

The provider advertisement and the consumer choice messages can use or not use any of the data elements defined in the data model, subject only to whether or not the element is deemed mandatory.

The model here follows the example of the xcon data model RFC 6501 [RFC6501].

Initially the data model may seem to describe only provider advertisements. But this is not the case. The consumer choice messages can also be constructed from these elements. Many of the data elements are similar for both messages, but there are also elements defined which are used only in one or the other of the messages.

One of the challenges here is that in addition to proposing a structure for the data model, we would like to also propose some additions or changes to what is in the current framework document. The problem is that any changes from what is in the current framework can detract attention from the basic information model which is the primary focus of this draft initially. We have noted where there are changes from the framework by putting the text in all capital text.

This is a rough draft, its goal is to propose a basic structure and stimulate discussion.

2. 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 [RFC2119] and indicate requirement levels for compliant implementations.

3. Overview

TEXT

3.1. Data model format

A CLUE object document is an XML [W3C.REC-xml-20081126] document. CLUE object documents MUST be based on XML 1.0 and MUST be encoded using UTF-8.

3.2. Data model namespace

This specification defines a new namespace specification for identifying the elements defined in the data model. This namespace is as follows: urn:ietf:params:xml:ns:clue-info

3.3. Data Model Structure

The information in this data model is structured in the following manner. All the information communicated between consumers and providers in a CLUE session is contained in a <CLUE-info> element. The definitions follow directly from the Framework document.

The <CLUE-info> element contains the following child elements:

- o The <capture-description> element describes the captures. It includes all the elements that have been discussed in the framework document as properties of captures document and it can be extended by adding additional elements describing a capture.
- o The <capture-scene> element is the data structure that describes the captures that the provider is capable of sending in a way that is helpfule to the consumer to decide what captures to request.
- o The <simultaneous-transmission-set> element is a data structure that describes the physical simultaneity constraints, as discussed in the framework.
- o The<stream-description> element describes the encoding characteristics of the streams.

4. Data Model Definition

The following non-normative diagram shows the structure of CLUE elements. The symbol "!" preceding an element indicates that the element is REQUIRED in the data model. The symbol "*" following an element indicates that multiple instances of the element are valid at the point at which it appears. "!" in combination with "*" indicates that at least one instance of the element is required, but that multiple instances are valid.

```
   !<CLUE info>
       |
       |--!<capture-description>*
       |      |--!<capture-id>
       |      |--!<media-type>
       |      |--!<content-type>
       |      |--<DERIVED>
       |      |--<switched>
       |      |--<spatial-description>
       |      |        |--<point-of-capture>
       |      |        |--<capture-area>
       |      |        |--<NATIVE-ASPECT-RATIO>
       |      |--!<audio-channel-format>
       |      |--!<encoding-group-ID>
       |      |--<future-media-capture-attribute>*
       |--!<capture-scene>*
       |      |--!<entry>*
       |      |        |--<capture-id>*
       |      |        |--<switch-policy
       |      |--<capture-scene-text>
       |      |--<capture-scene-language>
       |      |--<capture-scene-spatial-description>
       |      |        |--<capture-scene-area>
       |      |        |--<capture-scene-scale>
       |      |--<future-capture-scene-attribute>*
       |--<simultaneous-transmission-set>*
       |      |--!<entry>*
       |      |        |--<capture-id>*
       |--!<stream-description>
       |      |--<encoding>*
       |      |        |--!<encoding-id>
       |      |        |--!<multiplex-id>
       |      |        |--<AUDIO-RENDERING-ID>
       |      |        |--<max-audio-bandwidth>
       |      |        |--<max-video-bandwidth>
       |      |        |--<max-H264-Mbps>
       |      |        |--<max-width>
       |      |        |--<max-height>
       |      |        |--<max-frame-rate>
       |      |        |--<future-encoding-attribute>*
       |      |--<encoding-group>*
       |      |        |--!<encoding-group-id>
       |      |        |--<entry>*
       |      |                |<video encoding>*
       |      |        |--><entry>*
       |      |                |<audio encoding>*
       |      |        |--<max-group-bandwidth>
       |      |        |--<max-group-H264-<Mbps>
```

```
       |      |        |--<future-encoding-group-attribute>*
```

Data Model Definition

The following sections describe these elements in more detail.

4.1. <CLUE-info>

4.2. <capture-description>

The capture-description describes all the aspects of a capture, as discussed in the framework document.

4.3. <capture-id>

This element is a unique numeric value assigned to each capture by the provider. It is used by the consumer in choosing captures and streams.

4.4. <media-type>

This element indicates support by the consumer for a single capture media type, for instance "audio" or "video".

4.5. <content-type>

This per-capture attribute element indicates whether the media capture includes "main media", typically motion video of a real scene, or whether it includes "slides media", typically a computer-generated video stream rather than real people.

This is following the content attribute RFC [ref].

4.6. <DERIVED>

Instead of the boolean composed, we are suggesting an attribute that indicates whether media capture has been changed from its original form. It can take different values. As of now, the only vlaue defined is composed. Composed means a video image it is formed of other, more primary, captures mixed together. For audio it refers to mixed audio.[this needs to follow whatever the WG decides as to whether there is an audio composed, etc.]

4.7. <switched>

Instead of a boolean, we propose this attribute indicates the algorithm or method by which content is chosen form media sources.

The current value is loudest speaker. [this depends entirely on what the WG decides]

4.8. <spatial-description>

Spatial-description is a general element with sub-elements that describe spatial relationships. As of now, the framework defines point-of-capture and capture-area. We propose a third one here-NATIVE-ASPECT-RATIO.

4.8.1. <point-of-capture>

Use framework text.

4.8.2. <capture-area>

Use framework text.

4.8.3. <NATIVE-ASPECT_RATIO>

If a capture has a native aspect ratio (for instance, it corresponds to a camera that generates 4:3 video) then it can be supplied here. This can help rendering.

4.9. <audio-channel-format>

Use framework text.

4.10. <encoding-group-id>

This element is a unique numeric value assigned to each encoding group by the provider. It is included in a capture element because it links the capture to the group of encodings possible to use for the capture.

4.11. <future-media-capture-attribute>

This is a placeholder for media capture attributes that have yet to be defined.

4.12. <capture-scene>

A capture-scene is a description of the captures that the provider is capable of sending. It includes entries which are rows of audio or video captures.

Information concerning the relationship between captures is communicated by the capture-scene, that is, an entry in the capture-

scene signifies a complete representation of the total scene. When there are multiple entries of the same media type in the capture scene, this means that each entry is an alternate representation of the total scene.

Some elements of the capture-scene charactize the scene, as for example, capture-text.

An entry is a list of video captures or a list of audio captures, and an indication of the switching policy. Every <entry> element contains the following child elements:

o <capture-id> See above.
o <switch-policy> text from framework.

4.12.1. <capture-scene-text>

Text from framework

4.12.2. <capture-scene-language>

Text from framework

4.12.3. <capture-scene-spatial-description>

4.12.3.1. <capture-scene-area>

Indicates an overall encompassing co-ordinate "bounding box" at a per-capture set level.

4.12.3.2. capture-scene-scale>

On a per-capture scene level, this indicates whether the co-ordinates used for its media captures are in real-world units, relative units, or whether no scale information can be inferred.

4.12.4. <future-capture-scene-attribute>

This is a placeholder for capture set attributes that have yet to be defined.

4.13. <simultaneous transmission-set>

Indicates physical simultaneity constraints. Use framework text. Every <entry> element contains the following child elements:

o <capture-id> As above

4.14. <stream-description>

4.14.1. <encoding>

4.14.1.1. <encoding-id>

Use framework text. A unique value for the set of encoding parameters that constitute an encoding stream..

4.14.1.2. < MULTIPLEX-ID>

This element is the means by which the consumer, the receiver of a stream, tells the provider which multiplex id to generate that stream with, so that the consumer can demultiplex the stream correctly when receiving it.

4.14.1.3. <AUDIO-RENDERING-TAG>

Optional id for rendering audio supplied by consumer to the provider. The provider then uses this tag to to associate an audio stream with a video stream.

4.14.1.4. <max-audio-bandwidth>

This element gives the bandwidth limit for any audio strea. It can be used in provider advertisements with respect to potential streams, or in consumer request messages, describing a stream choice.

4.14.1.5. <max-video-bandwidth>

This element gives the provider's bandwidth limit for any video stream. It can be used in provider advertisements with respect to potential streams, or in consumer request messages, describing a stream choice.

4.14.1.6. <max-H264-Mbps>

This element gives the provider's maximum macroblock per second rate for any video stream.

4.14.1.7. <max-width>

Framework text

4.14.1.8. <max-height>

Framework text

4.14.1.9. <max-frame-rate>

This element gives a maximum frame rate for a video stream.

4.14.1.10. <future-encoding-attribute>

4.14.2. <encoding-group>

Encoding group is an element used by the provider to describe its capabilities. It groups together the potential encodings that are available for a particular capture. Therefore, a capture has associated with with it an encoding-id.

Every <entry> element contains the following child elements:

- o A <video-encoding> is an encoding for video.
- o An <audio-encoding> is an encoding for audio.

4.14.2.1. <encoding-group-id>

A unique id for the group of encodings that can be used for a particular capture.

4.14.2.2. <max-group-bandwidth>

This indicates the maximum bandwidth that the provider can transmit for its combined set of active encodings.

4.14.2.3. <max-group-H264-Mbps>

The maximum number of macroblocks per second that the provider can transmit for its combined set of active encodings.

4.14.2.4. <future-encoding-group-attribute>

This is a placeholder for encoding group attributes that have yet to be defined.

5. Security Considerations

TBD

6. IANA Considerations

TBD

7. References

7.1. Normative References

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

[RFC6501] Novo, O., Camarillo, G., Morgan, D., and J. Urpalainen, "Conference Information Data Model for Centralized Conferencing (XCON)", RFC 6501, March 2012.

7.2. Informative References

[I-D.ietf-clue-framework]
Romanow, A., Duckworth, M., Pepperell, A., and B. Baldino, "Framework for Telepresence Multi-Streams", draft-ietf-clue-framework-05 (work in progress), May 2012.

[I-D.lennox-clue-rtp-usage]
Lennox, J., Witty, P., and A. Romanow, "Real-Time Transport Protocol (RTP) Usage for Telepresence Sessions", draft-lennox-clue-rtp-usage-04 (work in progress), June 2012.

Authors' Addresses

Allyn Romanow
Cisco Systems
San Jose, CA 95134
USA

Email: allyn@cisco.com

Andy Pepperell
Silverflare

Email: andy.pepperell@silverflare.com