ISO/IEC JTC 1/SC 29/WG 03 N0274

**ISO/IEC JTC 1/SC 29/WG 03  
MPEG Systems   
Convenorship: KATS (Korea, Republic of)**

**Document type:** Output Document

**Title:** Exploration Experiments for Scene Description for MPEG Media

**Status:** Approved

**Date of document:** 2021-05-04

**Source:** ISO/IEC JTC 1/SC 29/WG 03

**No. of pages:** 1 (with cover page)

**Email of Convenor:** young.L @ samsung . com

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg3>

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 03 MPEG SYSTEMS**

**ISO/IEC JTC 1/SC 29/WG 03 *N00274***

**Online - April 2021**

|  |  |
| --- | --- |
| **Title** | Exploration Experiments for Scene Description for MPEG Media |
| **Author** | **Sungryeul Rhyu, Eric Yip** |
| **Serial** | **20401** |

# Abstract

This document provides descriptions of the Exploration Experiment on the proposed technologies for Scene Description for MPEG Media. This is a consolidation of two EEs:

* EE1 (on Carriage of Random Access Support): evaluates proposed technologies on the carriage of temporal random access support.
* EE2 (on Dynamic Scene Updates): creates dynamic scene update scenarios and proceed a evaluation of the scenarios.

# EE1: on Carriage of Random Access Support

## Introduction

The goal of EE1 is to evaluate the proposed solutions on carriage of information, more specifically glTF objects and JSON Patches to enable temporal random access of MPEG scene description. This will enable the Scene Description adhoc to validate random access capability, and to evaluate the efficiency and behavior of the proposed solutions. EE1 is expected to provide considerable insights in terms of the cumulative size of temporal random access data, and also the processing steps required, from the Presentation Engine perspective.

## Mandates

The mandates for this EE are as follows:

* To study the types of samples (including their data structure and storage format) which enable temporal random access
* To study possible coding structures to provide efficient access of required data, in terms of
  + The required number of samples to construct an independent version of a scene description
  + The total size of sample payloads to deliver the randomly accessible dynamic scene descriptions
  + The benefit of storing the required samples in either one track or two (or more) tracks
  + Processing steps to construct an independent version of a scene description for random access operation
  + Processing steps to construct an independent version of a scene description for plain playback operation (i.e., normal scene update with no trick play)
* To provide recommendations on the best practices for the composition of the various types of samples
* To mandate text to be incorporated into DIS of ISO/IEC 23090-14

## Participants

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Participant | Contact | Email | Type | - |
| Samsung Electronics | Sungryeul Rhyu | [suzz.rhyu@samsung.com](mailto:suzz.rhyu@samsung.com) | P |  |
| Nokia Corporation | Lukasz Kondrad | [lukasz.kondrad@nokia.com](mailto:lukasz.kondrad@nokia.com) | P |  |

(P = proponent, C = cross checker)

## Information about proposed technologies

### Option A. M55928

Refer detailed technology in [1].

### Option B. M56323

Refer detailed technology in [2].

### Option C. M56439

Refer detailed technology in [3].

### Option D. M56731

#### On random access support for ISO/IEC 23090-14 [4]

Proposed method suggests a mix of the two idea; based on 2 track approach, introduce a new type of sample that referencing another sample. The ISDIndependent Scene Description type sample has an independent and complete version of a scene description if the version of scene description has big difference from the v1 scene description. The DSDependent to Sample type sample has patch operations to referenced sample where the sample is either ISD type or DS type.

* Independent Scene Description (ISD): An independent sample.
* Dependent to Sample (DS): A patch sample having patch operations to another sample.

ISD type sample can be occupied when the random access point has major version changes comparing to the older version of scene descriptions. For the case, it is more efficient to bring the scene description to be independent to the other samples.

DS type sample can be occupied when a random access point has very small changes to the other samples. The DS type sample refers another sample in track 1 or 2.

Fig. 3 shows an example of proposed sample types in track 2. The first sample is a DS type which refers the first sample in track 1 which has v1. The second sample is an ISD type sample which has independent scene description v5. The third sample is a DS type which refers the second sample of track 2.



**Fig. 3 Conceptual diagram of the proposal**

Refer additional details in [4].

## Evaluation

### Test sequence

* Test set #A used in [4]
  + Use the example scene changes: 7 versions of scene description in Fig.5 of [4]
  + Proposed coding structure that provides random access point on 3rd, 5th, 7th samples
* Other test sets, coding structures if available by timeline.

### Evaluation criteria

* Test No. A-1) The required number of samples to construct each independent version (v1, v2, v3, v4, v5, v6, and v7).
* Test No. A-2) Total size of sample payloads (from all tracks).
* Test No. A-3) Random access: A simple python code (revision of Annex A of [4]) or processing steps to construct an independent version of v3, v5, v7.
* Test No. A-4) Plain playback: A simple python code or processing steps to construct an independent version of v2 from v1, v3 from v2, v4 from v3, v5 from v4, v6 from v5, v7 from v6.

### Evaluation

Evaluation experiments are to be carried out under the same evaluation criteria compared with other proposals. Option A should be included to assure participant’s common understandings on the evaluation.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Test no |  | Proposed method | Analysis on method A | Analysis on method B | Analysis on method C | Analysis on method D |
| A-1 | v1 | 0 samples |  |  |  |  |
|  | v2 | 0 samples |  |  |  |  |
|  | v3 | 0 samples |  |  |  |  |
|  | v4 | 0 samples |  |  |  |  |
|  | v5 | 0 samples |  |  |  |  |
|  | v6 | 0 samples |  |  |  |  |
|  | v7 | 0 samples |  |  |  |  |
| A-2 | all versions | 000 bytes |  |  |  |  |
| A-3 | v3, v5, v7 | Code or report |  |  |  |  |
| A-4 | v2, v3, v4, v5, v6, v7 | Code or report |  |  |  |  |

## Additional works

It is recommended to participants to propose a new test set and new evaluation criteria with the evaluation result. However, it is strongly recommended to avoid disclosing the test result with the new test conditions for fairness concern.

## Timeline

* 2021-06-02: Deadline for participants to provide a new test set and new evaluation criteria with detailed instructions
* 2021-06-09: Expected date for the release of EE1 description
* 2021-07-07: MPEG document upload deadline
* 2021-07-12: MPEG #135(online) meeting start

# EE2 on Dynamic Scene Updates

## Introduction

The goal of EE2 is to create and proceed a scenario evaluation for dynamic scene updates in MPEG Scene Description. Dynamic scene update scenarios will be collected and analysed based on the basic functionality currently supported in the DIS, with the target of identifying gaps in the support of such scenarios. Based on these gaps, possible solutions are expected to be studied according to their benefits and functionalities.

## Mandates

The mandates for this EE are as follows:

* To identify a list of different dynamic scene update scenarios.
* To study the scenarios according to the basic functionalities currently supported in the DIS by identifying the gaps in the support of these scenarios by the basic functionalities.
* To study different possible solutions which address the gaps identified, based on their benefits and suitability as extended functionalities.
* To provide recommendations on dynamic scene update scenarios and solutions to support such scenarios.
* To mandate text to be incorporated into DIS of ISO/IEC 23090-14, or its amendment/second edition.

## Participants

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Participant | Contact | Email | Type | - |
| Samsung Electronics | Eric Yip | [eric.yip@samsung.com](mailto:eric.yip@samsung.com) | P |  |
| Nokia Corporation | Lukasz Kondrad | [lukasz.kondrad@nokia.com](mailto:lukasz.kondrad@nokia.com) | P |  |

## Dynamic scene update scenarios

Scene updates are currently defined in clause 7.2 of the DIS. In particular, the triggering factors for the fetching of updates and the activation of certain nodes are defined and listed as the following:

* Wallclock time
* Presentation time
* Interaction event

It is observed that the current DIS text only supports presentation time triggered dynamic scene updates.

### Timed dynamic scene update scenarios

#### Wallclock time triggered dynamic scene updates [CD text]

The current DIS text does not contain any support for the indication of a wallclock time to be used for triggering dynamic scene updates.

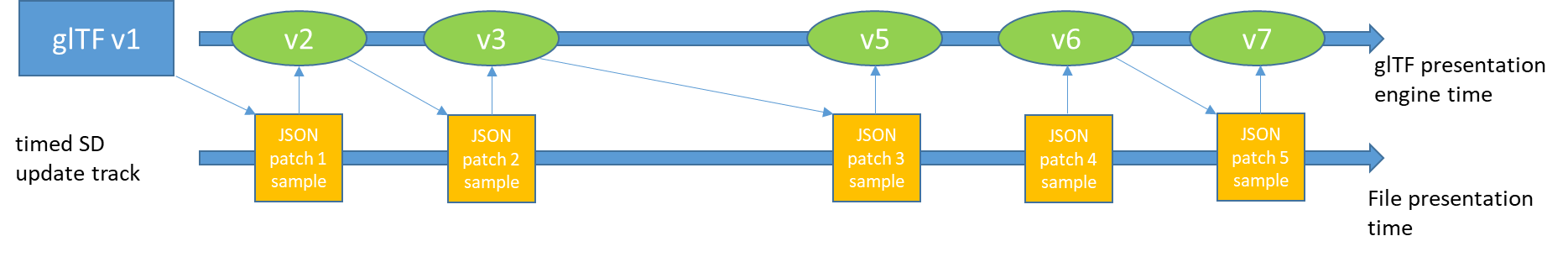
A potential solution is to include dynamic scene update attributes related to wallclock time triggers, as shown in table 1.

**Table 1 – Definitions of wallclock time attributes for dynamic scene updates**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Default** | **Description** |
| absolute\_time\_UTC | DateTime | n/a | Wall clock time identifying the execution time of the scene update transaction on the glTF object. The value is denoted in UTC. |
| absolute\_time\_TAI | DateTime | n/a | Wall clock time identifying the execution time of the scene update transaction on the glTF object. The value is denoted in TAI. |

#### Presentation time triggered dynamic scene updates [DIS text]

The current DIS text supports timed dynamic scene updates which are triggered by presentation time. The presentation timestamp of the track samples containing JSON patch documents (as defined in clause 8.4 of the DIS) can be used to trigger the dynamic scene update as shown in figure 4.



**Fig. 4 Presentation time triggered dynamic scene updates**

#### Conditional timed dynamic scene updates [CD text, m56736]

Dynamic scene updates may be restricted by conditions such as the version of a scene description which is being maintained in the Presentation Engine memory at a given time. In such a scenario, the scene update sample is only applied when the version of the current scene description matches to that specified by the update sample.

One solution is to include version related dynamic scene update attributes; namely a target\_version\_id, which specifies the version of the target scene description for which the dynamic scene update is applicable, as well as a result\_version\_id, which specifies the resulting scene description when the dynamic scene update is applied. These attributes are shown in table 2.

**Table 2 – Definitions of version attributes for dynamic scene updates**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Default** | **Description** |
| target\_version\_id | String | n/a | Identifier for the version of the target scene description for which the dynamic scene update is applicable |
| result\_version\_id | String | n/a | Identifier for the version of the resulting scene description when the dynamic scene update is applied |

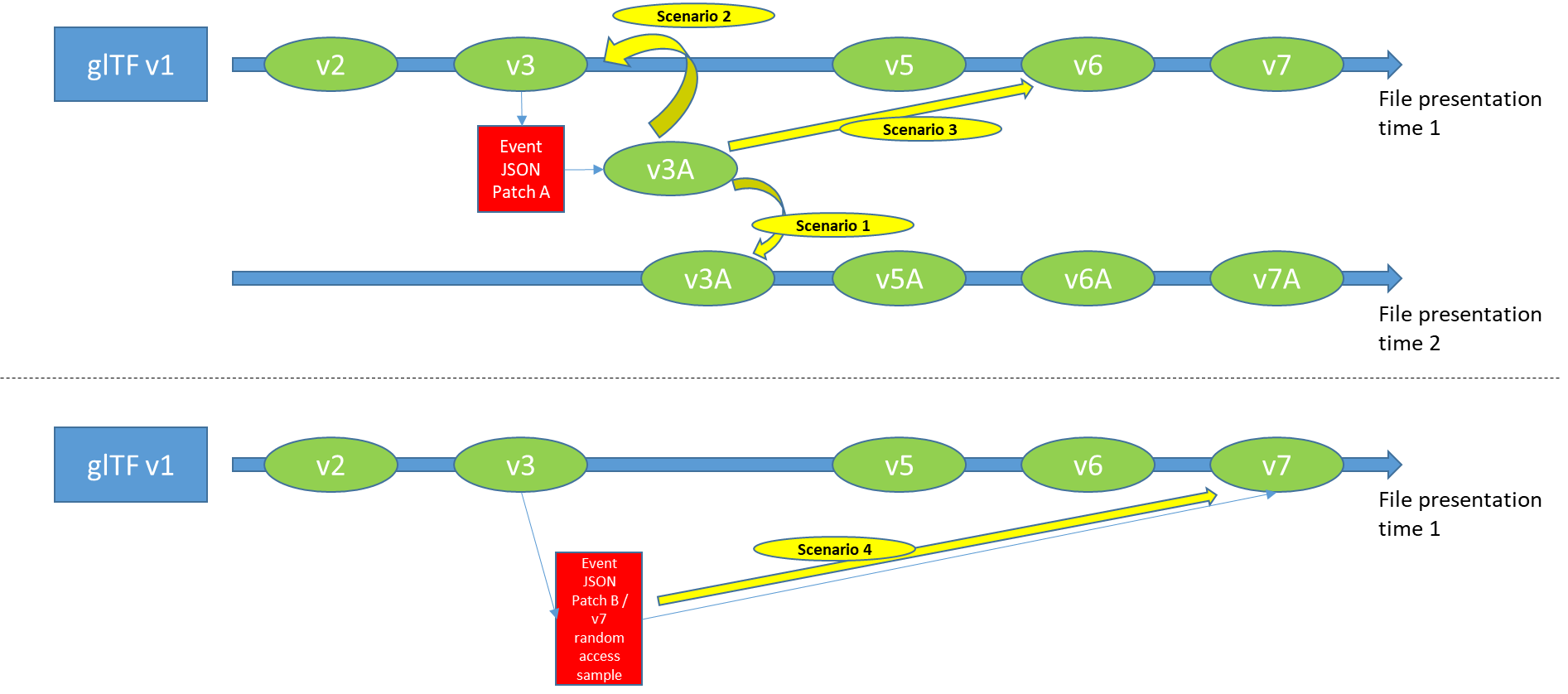
### Non-timed dynamic scene update scenarios

#### Event triggered dynamic scene updates [m56736]

The current DIS text does not contain any support for the indication of dynamic updates which are triggered by interaction events.

The triggering of a scene update by an interaction event is highly content dependent, and can be expressed as content metadata through event scene update attributes. Unlike a timed scene update which typically progresses the timeline of the content, an event triggered scene update may allow for deviations from the main content timeline, hence allowing for loops or skips in the content timeline.

Depending on the content author’s intent, there may be four different event update scenarios as shown in figure 5.



**Fig. 5 Event triggered dynamic scene update scenarios**

Scenario 1

* An event A triggers the dynamic scene update of glTF v3 to glTF v3A. This update utilises the event update sample containing JSON Patch document A.
* The updated scene description (v3A) is executed by the presentation engine, after which further updates may be enabled through timed or non-timed scene updates.

Scenario 2

* An event A triggers the dynamic scene update of glTF v3 to glTF v3A. This update utilises the event update sample containing JSON Patch document A.
* The updated scene description (v3A), as a result of, and through the nature of the event, has a certain playout period (a certain presentation time period).
* After the playout of scene description v3A, the presentation engine returns the scene to the scene description version before the occurrence of the event (v3).

Scenario 3

* An event A triggers the dynamic scene update of glTF v3 to glTF v3A. This update utilises the event update sample containing JSON Patch document A.
* The updated scene description (v3A), as a result of, and through the nature of the event, has a certain playout period (a certain presentation time period).
* After the playout of scene description v3A, the presentation engine skips to a specified scene description version (v6).

Scenario 4 (a special case of scenario 1, mixed with scenario 3)

* An event B triggers the dynamic scene update of glTF v3 to glTF v7.
* glTF v7 is a scene description version which exists within the same presentation timeline of glTF v3, and so the update can also be considered a direct scene skip.
* This update may utilize an event update sample containing JSON Patch document B, or may be realised by the execution of a random access scene description sample containing glTF v7.

The support of these scenarios can be specified by event scene update attributes shown in table 3.

**Table 3 – Definitions of event related attributes for dynamic scene updates**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Default** | **Description** |
| event\_id | String | n/a | Identifier for the event triggering the dynamic scene update |
| skip\_version\_id | String | n/a | Identifier for the version of the skip to scene description which is executed by the presentation engine, after the playout of the event updated scene graph |
| skip\_time | integer | n/a | Specifies the skip to time (referenced to the presentation time) of the scene graph version to skip to by the presentation engine, after the playout of the event updated scene graph |
| return\_event | integer | n/a | Flag to indicate whether the scene graph version is returned after the event scene playout. When set to 1, this flag indicates that the scene graph version is returned to the version previous to the event, after the playout of the scene graph version triggered by the event scene update. |
| playout\_time | integer | n/a | Specifies the playout time of the updated scene graph version triggered by the event scene update (in seconds, or any other time unit) |

## Timeline

* 2021-06-02: Deadline for participants to provide new scenarios for dynamic scene updates
* 2021-06-09: Expected date for the release of EE2 description
* 2021-07-07: MPEG document upload deadline
* 2021-07-12: MPEG #135(online) meeting start

# References

1. (41.1)[SD] On Random Access Support for Scene Description with Scene Updates, M55928, Online, Jan 2021.
2. [SD] On Random Access Support for Dynamic Scenes, M56323, Online, Feb 2021
3. [41.1] Carriage of glTF JSON documents and JSON patch documents, M56439, Online, Apr 2021
4. [41.1] On random access support for ISO/IEC 23090-14, M56731, Apr 2021
5. [41.1] On dynamic scene updates in ISO/IEC 23090-14, M56736, Apr 2021