Latest Status of UMID and its Applications in File-based Workflow

by Yoshiaki Shibata
Chair, SMPTE TC-30MR SG UMID Applications

Introduction

The UMID (Unique Material IDentifier) is a globally unique AV (audiovisual) material identifier standardized by SMPTE (Society of Motion Picture and Television Engineers) as SMPTE ST 330 and RP 205. More than a decade has passed since its initial standardization in 2000 and its presence in AV products using the MXF (Material Exchange Format), files containing the UMID as its mandatory component have been widely spread throughout the industry. But, its originally intended use as a unique material identifier to link AV material to its metadata has seldom been seen in practice.

In addition, although AV material management using a unique identifier is a common practice observed in variety of products, a problem with the material identifier often occurs when such products are combined to form a media production system. The UMID, as an industry common AV material identifier, is expected to address this sort of problem. In reality, however, most products have their own proprietary material identification schemes with little or no interoperability between them. Furthermore, although certain products have adopted the UMID for their material identifier, they have no capability to communicate with other products via the UMID, resulting in failure of their integration into a system.

Recently, we have revealed that the issue is mainly due to lack of industry common rules for the UMID applications and, in order to address the UMID application issue, a new project has just begun in the SMPTE SC (Standard Community).

In this article, the goal of UMID and its applications are discussed, together with the needs of industry common UMID application rules to achieve it, followed by possible UMID applications specifically for MXF. Then the UMID application project recently approved by the SMPTE SC is introduced to call for your participation.

Goal of UMID Applications

Figure 1 shows the structure of UMID format. The UMID is inherently a single 32-byte entity that takes the form of well-known KLV (Key-Length-Value) structure. The first 12-byte Key field, called the SMPTE UL (Universal Label), indicates that this is the UMID. The 1-byte Length field specifies the length of byte string that follows, which is fixed to 13h. The Value field is divided into two components: the 3-byte Instance Number (Inst.#) to signal its originality that is usually set to zero, and the 16-byte Material Number (Mat.#) to accommodate a globally unique value that makes the UMID as a globally unique identifier.

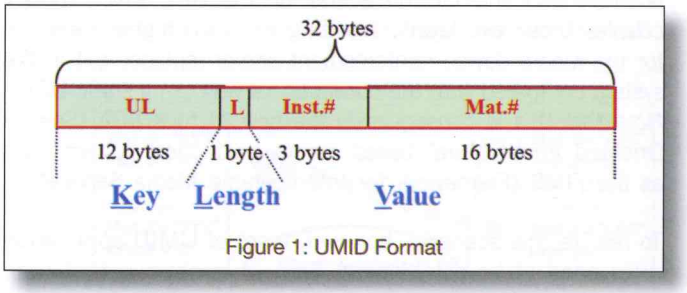

Figure 1: UMID Format

The primary use of UMID is as a globally unique AV material identifier. Because the UMID is usually attached to the AV material that needs to be globally uniquely identified with a UMID, metadata can be unambiguously associated with the AV material by specifying the reference to it via the UMID. This leads to a typical UMID application called the UMID based AV material search (Figure 2).

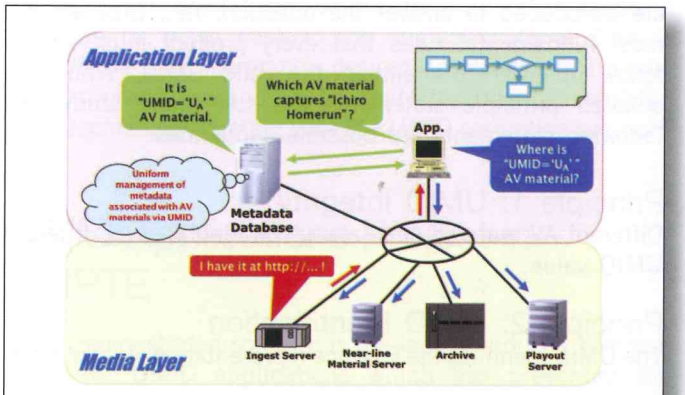

Figure 2: UMID based AV Material Search

In Figure 2, AV materials are stored in various kinds of material servers and a section of descriptive metadata associated with an AV material via its UMID are collected and separately stored altogether in a Metadata database. When an application (App.), such as video editing, wants to obtain an AV material that captures the "Ichiro Homerun" scene for example, it gives the request to the Metadata database accordingly. The Metadata database then replies to the application the requested AV material via its UMID.

Because the UMID by itself does not tell where to access the desired AV material, the application needs to resolve the UMID; i.e., to dynamically associate the UMID with its corresponding URL. In this example, when the application distributes a query such as "Do you have an AV material with this UMID value?", the Ingest server replies "Yes I have!" and provides the URL of the requested AV material, which can then be used by the application to access the material.

While this scenario demonstrates the principle of material search and access, it may be generalized by dividing a media production system into two layers: the Application layer for various applications including workflow management, and the Media layer for media devices such as material servers and media processors.

In the upper Application layer, light weight data such as workflow information and/or material metadata, typically represented in XML (eXtensible Markup Language), need to be handled in a flexible way, while the heavy weight media data, need to be intensively processed in the lower Media layer. Then, it is the UMID and its connection via the URL which loosely couples those two layers, resulting in much higher flexibility for the media device reinforcement and/or replacement in the system compared with the conventional way using static URLs. Note that this is consequently the best fit to a SOA (Service Oriented Architecture) based media production system such as the FIMS (Framework for Interoperable Media Service).

To realize this scenario, however, a set of UMID application rules need to be established. First, it is obvious that each material server, even when supplied from a different vendor, needs to appropriately respond with the URL for a given UMID to an external application based on a defined UMID resolution protocol. Furthermore, in order for the UMID to be used as a common material identifier in the industry, it must be always *valid.*

But, what is a valid UMID? The UMID Application Principles are introduced to answer the question, i.e., they are the most fundamental rules that every product must strictly follow in order to maintain the valid UMID. While the detailed principle statements are still under study, the following statements are possible candidates:

Principle 1: UMID Integrity

Different AV material shall always be identified by different UMID value,

Principle 2: UMID Identification

The UMID identifies the Essence representation at its playout.

The Principle 1 should be obvious, i.e., different AV materials must not share the same UMID value. The Principle 2 might be controversial, but it is crucial to provide a clear boundary between *identical* and *different* AV materials. According to the proposed Principle 2, two AV materials may be said identical (and thus share the same UMID) when their A/V essence is identical at their playout and any differences occur only in their descriptive metadata such as a title or description.

Note that because the UMID Application Principles should be defined by common industry rules, they will be specified as normative part of the relevant industry standards. But because such principles are usually too abstract to be implemented into products as they state, a concept called UMID Managed Domain is introduced as an embodiment of the UMID Application Principles.

The UMID Managed Domain is defined as a logical space composed of appropriately managed AV materials via their valid UMIDs. This is implemented by introducing a material manager that manages the AV materials with their UMIDs and corresponding URLs in the Domain. To maintain the UMIDs validity, the material manager plays a crucial role in any AV material manipulation in the Domain.

For example, when AV material having its own UMID in some non-standard location is imported into the Domain, the material manager should replace its original UMID with a newly created value. This is because there is no guarantee that the original UMID is globally unique. When the material manager creates a new UMID according to the standard (SMPTE ST 330), the global uniqueness of the UMID is always guaranteed.

Another example is given for AV material modification. When an AV material containing its UMID in it is duplicated and a part of the Essence in one of the AV materials is modified by overwriting, the material manager also needs to replace its contained UMID with a newly created value in order to avoid those two AV materials of different Essence sharing the same UMID.

It should be mentioned that the material manager is also responsible for an appropriate maintenance of the UMID and its corresponding URL for each AV material, because it is in the position to know both of them. As a result, it is the material manager that must take the active role in the aforementioned UMID resolution protocol.

An interesting property of the UMID Managed Domain is that if an AV material having its own UMID is imported from another UMID Managed Domain, the material manager does not have to replace the original UMID because its global uniqueness is already guaranteed. Consequently, identical AV materials can co-exist in the source and the destination UMID Managed Domains. Effectively those two Domains are merged to form a single UMID Managed Domain.

This property of the UMID Managed Domain can be extended also in a system level, i.e., if all products participating in a media production system are correctly supporting the UMID Managed Domain, the Domains may be merged to cover an entire system, resulting in the UMID based media production system (Figure 3). Note that, in this system, the UMID is used as a common material identifier over the system even when multiple MAM (Media Asset Management) products from different vendors are involved. It should be also noted that the same protocol as used between the application and a material server in Figure 2 can be also used among products in Figure 3.

Furthermore, because of the global uniqueness of the UMID, the above scenario may be further extended in a global sense. According to the UMID Application Principle 2, two

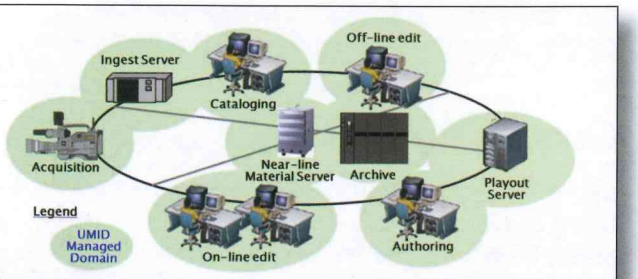

Figure 3: UMID based Media Production System

AV materials sharing the same UMID are said identical even when they are located in geographically different sites. Their descriptive metadata may be different (due to cultural difference) and/or they may use different management information (due to different material management systems). This implies that UMID plays a crucial role in establishing a globally distributed AV material management system that enables "Manage locally, Search globally".

UMID Applications in MXF

The MXF is a SMPTE standard container format that can contain any kinds and any forms of AV Essences. MXF also carries metadata, which is associated with the whole, and/or particular points of the Essence it contains.

Among other valuable characteristics of MXF, its dual structure is one of the most unique characteristics of MXF. In short, what is to be played out is not always the same as what is actually contained in a file (Figure 4). This is realized by introducing the structural metadata called the File Package (FP) and the Material Package (MP). As demonstrated in Figure 4, the File Package is used to describe the temporal information of the Essence contained in a file together with its technical properties such as codec, frame size, and so on, while the Material Package is used for presentation of the Essence at its playout. All packages in MXF are identified by their own UMID under the names File Package UID (FpUmid) and Material Package UID (MpUmid), respectively.

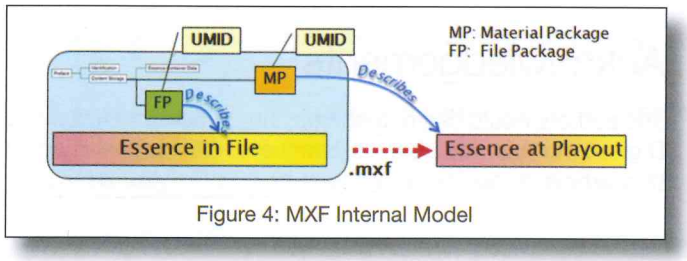

Figure 4: MXF Internal Model

Because the Material Package represents the Essence at the playout of a MXF file, the MpUmid of a MXF file is the unique material identifier that globally uniquely identifies the MXF file according to the UMID Application Principle 2, which needs to be carefully managed to constitute the UMID Managed Domain.

The FpUmid, on the other hand, must also be unique. But because it is not used by the UMID Managed Domain, the scope of its uniqueness may be only to within a file or relevant files. This enables a new secondary MXF file to be associated with its original via the UMID, while both the original and the secondary created MXF files are independently managed via their MpUmids. Figure 5 demonstrates this sort of UMID application in MXF when a secondary MXF file is created as a result of a partial retrieval of the original MXF file.

In Figure 5, the resulting MXF file "Result.mxf" is obtained by partially retrieving an original MXF file "Source.mxf" with its In/Out points as illustrated. Because "Result.mxf" is a newly created MXF file, its MpUmid must be a newly created UMID value with a zero Instance Number (Inst.#) according to SMPTE ST330. But because the FpUmid in "Result.mxf" is permitted not to be globally unique, it can be used as a linking tool to the original MXF file. Specifically, because the Essence contained in "Result.mxf" is created from the output of "Source.mxf", it is reasonable for the value of FpUmid in "Result.mxf" to be also created from the value of MpUmid in "Source.mxf" by the Material Number (Mat.#) of the FpUmid in "Result.mxf" being inherited from that of the MpUmid in "Source.mxf", while its Instance Number being replaced with a non-zero value to indicate that it is not original.

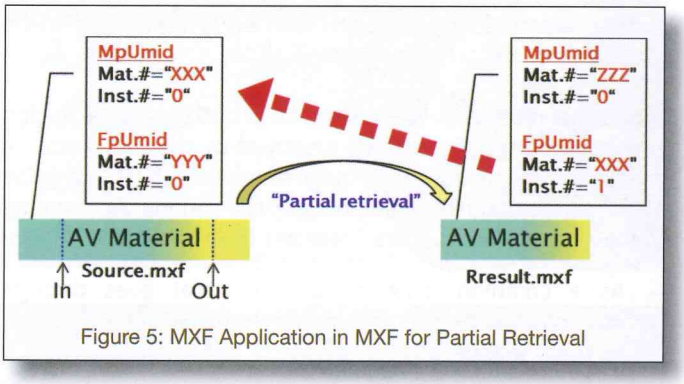

Figure 5: MXF Application in MXF for Partial Retrieval

It should be noted that when the FpUmid in "Result.mxf" is resolved to a URL after zero-masking the Instance Number, the URL will resolve to "Source.mxf". In other words, a link can be made back to the original material. This is one of the most frequently desired operations in a media production system, and can be easily implemented using the MXF technology together with the UMID resolution protocol.

UMID Applications Project in SMPTE

As is demonstrated so far, it is the lack of industry common rules for UMID applications which has prevented the UMID from being widely used in the industry as originally intended. To address this issue, a new project has just started in SMPTE.

The SMPTE TC-30MR SG UMID Application, chaired by the author, has been set up since April 2012, with its scope:

1. To explore the best practices of UMID application,
2. To identify typical UMID Application Principles,
3. To identify relevant technologies needed to be additionally standardized.

Regarding item 1, because more than a decade has passed since the initial introduction of UMID to the industry, there exist various trials of UMID applications so far. Current uses of the UMID have been limited only within a particular products and/or a systems despite its original intention. But many of the lessons learned from such trials still should be applicable even to use of the UMID as a globally unique material identifier. Therefore, we propose to collect existing UMID application practices used so far by the industry and explore the best practices through analysis of the responses.

Regarding item 2, analysis of existing practices also leads to the identification and collation of UMID Application Principles, i.e., the candidate statements for the principles will be carefully tested against the existing practices in order to refine them.

Regarding item 3, identification of relevant technologies that needs to be standardized to enhance UMID applications is an important task of this project. While plausible technologies already identified include the UMID resolution protocol and basic rules for the UMID application in MXF, some aspects of the UMID Managed Domain need to be further examined to identify those parts that need to be standardized to achieve interoperability.

While the most well-known use of a UMID is as a linking tool between the AV material identified by the UMID and its metadata referencing to the material via the UMID, using the UMID as an industry common globally unique AV material identifier in practice is also relevant to the following topics:

- As a common material identifier that goes through multiple MAM systems from different vendors, resulting in their seamless integration,
- As a common material identifier for a distributed AV material management system, enabling to achieve "Manage locally, Search globally",
- As a dynamic linking tool between the upper Application layer and the lower Media layer (See Figure 2), realizing their loose coupling.

Finally, it should be noted that this project was initially started as a study group. This does not conduct any standardization process itself but submits one or more reports to the parent Technology Committee (TC-30MR) with any recommendations for standardizations. The work of the study group is due to be concluded by March 2013. The actual standardization activities can then start based on the outcome of the study group reports.

Conclusions

In this article, the UMID and its applications are discussed based on the observation that it has not been used so much as originally intended because of a lack of supplementary rules for the UMID application. The provision of such rules will significantly enhance the usefulness of UMID within our industry. In particular, because the UMID has been adopted as a core component of MXF, the UMID has the potential to be the core of a media IT infrastructure for the future file-based media production system.

To address the UMID application issue, a new project called SMPTE TC-30MR SG UMID Application has just started within SMPTE. The project has initially been set up as a study group, whose guidance will lead to the relevant standardization activities in the near future.

Because this project has just started, it is a good time for those involved in AV material management, based on the UMID or any other unique material identifier, to join the project. Furthermore, even though the UMID was designed specifically for AV material identification, those having knowledge and/or experience of unique identifier based entity management are also welcome to join because their expertise could be also applicable to UMID based AV material management.

At the time of writing, there are more than thirty participants in the project. But, unfortunately, none of them is from the Asia-Pacific region (except Japan). Therefore, experts from this region are especially welcome to join the project and to contribute their expertise to the future-proof activity. Historically, most of standardizations for the file-based workflow have been led by the US and EU experts. But now, this is your turn to take part in developing the future file-based media production system!

For more information, please don't hesitate to contact the author at info@metafrontier.jp

Acknowledgements

The authors would like to thank Mr. Jim Wilkinson (Wellspring Digital) for his valuable assistance in the preparation of this article.

author

Yoshiaki Shibata
metaFrontier.jp, LLC - Japan

Yoshiaki Shibata is the President and Chief Consultant of metaFrontier.jp, LLC. He started his career at Sony in 1991. In 1998, Shibata joined the MPEG-7 standardization activity, in charge of the MPEG-7 schema design. Since 2001, he started working in the M&E industry, where his initial contributions include successful implementations of UMID and EssenceMarkTM for professional VTRs. In 2002, he joined the XDCAMTM project, played a crucial role in the metadata part of the product development, including applying for more than forty patents on metadata related technology. In 2011, Shibata has left Sony to be the Japan's first independent consultant for the media and metadata technology, and founded metaFrontier.jp, LLC in 2012. He is an active member of SMPTE, FIMS, AMWA, EBU Technical, ITE and IPSJ.

ABU TECHNOLOGY
UPCOMING EVENTS

14-16 May	Regional Workshop on Satellite Broadcasting	Colombo
2-4 July	HBF-ABU Preparatory Seminar on WRC-15	Kuala Lumpur
9-11 July	Workshop of Radio Studio Automation	Bangkok
15-16 July	In-country Workshop in Kiribati	Kiribati
5-9 August	ABU-HFC E-mail Coordination for B13 Season	Kuala Lumpur
6-8 August	ABU Workshop on HDTV Studio Planning	Hanoi
20-22 August	Regional Workshop on OTT and HBB Technology for Broadcasters	Kuala Lumpur
26-28 August	ABU-Central Asia Media Forum	Bishkek