

CODECS et CONTENEURS

Principes de l'audiovisuel - Vincent MAZO - EFREI 2012/2013

I. Qu'est ce qu'un Codec ?.....	3
II. Les codecs audio et vidéo.....	4
1) DV	4
2) HDV	5
3) MPEG	6
A) MPEG1.....	6
B) MPEG2.....	6
C) MPEG4.....	7
4) DIVX.....	9
5) BETA DIGITAL	9
6) HDCAM.....	10
7) HDCAM SR.....	10
8) PRORES.....	11
9) DnxHD	12
III. Spécifications des formats vidéos courants.....	13
IV. Qu'est-ce qu'un conteneur ?.....	14
V. Conteneurs usuels	15
1) AVI.....	15
2) MOV.....	15
3) DV - HDV	16
4) MPG / MP2 / MP4	16
5) ASF / WMA / WMV	17
6) MJ2	17
7) MKV (MATROVSKA)	18
VI. Les formats d'échange	18
1) EDL	18
2) XML	19
3) OMF.....	19
4) MXF	20
5) AAF	21
VII. Comment choisir un codec ?.....	21

I. Qu'est ce qu'un Codec ?

Un codec est un procédé capable de compresser et de décompresser des données numériques. Cette fonction peut être réalisée par un circuit électronique ou par voie logicielle.

- Un codec hardware utilisera par exemple une carte spécialisée dans l'ordinateur. Ses circuits sont capables de coder et de décoder les données sans avoir besoin de mobiliser le processeur central de la machine.
- Un codec software aura en revanche l'avantage de pouvoir évoluer plus vite par l'intermédiaire de mise à jour.

Le mot « codec » signifie « compression-décompression » (ou « codage-décodage » - CODE-DECODE en anglais).

D'un côté, les codecs encodent des flux ou des signaux pour la transmission, le stockage ou le chiffrement de données. D'un autre côté, ils décodent ces flux ou signaux pour édition ou restitution.

Les différents algorithmes de compression et de décompression peuvent correspondre à différents besoins en qualité de restitution, de temps de compression ou de décompression, de limitation en termes de ressource processeur ou mémoire, de débit du flux après compression ou de taille du fichier résultant. Ils sont utilisés pour des applications comme la téléphonie, les visioconférences, la diffusion de médias sur Internet, le stockage sur CD, DVD, la télé numérique par exemple.

Il existe une grande variété de codecs. Les applications visées sont très diverses, ainsi que les moyens mis en œuvre : Stockage, streaming, faible débit, qualité maximale... On choisira ainsi le codec en tenant compte des besoins et des limitations liées au projet.

Il existe des codecs sans compression, ou dont la compression est sans perte (Lossless) et des codecs dont la compression introduira une perte (Lossy).

Les notions de codec, norme et conteneur sont souvent confondues par abus de langage :

- **La norme** décrit le format des données, les techniques de compression.
- **Le codec** est le logiciel ou le matériel qui met en œuvre un procédé capable de compresser ou décompresser les données de format normalisé.
- **Le format conteneur** contient des flux audio et vidéo respectant une quelconque norme. Ce format permet d'entrelacer les données audio et vidéo, et contient les informations permettant de les synchroniser au moment de la restitution.

II. Les codecs audio et vidéo

1) DV



Le format DV a été mis au point par un large consortium, regroupant des sociétés comme Matsushita (qui possède Panasonic), Philips, Sony, Thomson, rejoints par Hitachi, JVC, Mitsubishi, Sanyo, Sharp et Toshiba mais aussi Apple et IBM, soit au total plus d'une cinquantaine de sociétés. Cette alliance industrielle historique dans le monde de l'électronique s'est constituée pour définir les spécifications de la nouvelle génération de magnétoscopes grand public.

L'image de résolution standard de **720 × 576 pixels (25 images/seconde) en PAL et 720 x 480 pixels (30 images/seconde) en NTSC** intègre deux trames (les lignes paires composant une trame et les lignes impaires, l'autre trame) par image et est compressée en Jpeg. Le rapport d'aspect de l'image diffusée est **de 4:3 ou 16:9 après anamorphose**, un rapport S/B (Signal / Bruit) de 54 dB, le DV offre des performances bien supérieures à celles des formats analogiques Hi-8 ou S-VHS.

Le DV est un format en composantes numériques (c'est-à-dire que les signaux de luminance, de différentiel de bleu et de différentiel de rouge sont traités séparément). Dans les systèmes à 625 lignes (PAL, SECAM), l'échantillonnage est de type **4:2:0**, c'est-à-dire qu'ils ont un nombre identique de pixels **au 4:1:1** mais que la perte de définition est répartie en horizontal et en vertical. Dans les systèmes à 525 lignes (NTSC), il a été choisi une structure d'échantillonnage de type 4:1:1. Ce choix s'explique par le fait que les systèmes à 525 lignes ayant déjà une résolution verticale assez faible, auraient plutôt mal accepté d'être amputés de la moitié de leurs points de chrominance en vertical.

La structure d'échantillonnage du DV est aussi l'une des raisons pour laquelle on évite de l'utiliser dans des applications telles que le chromakey.

La quantification de l'image DV est réalisée avec **8 bits par couche**. Ainsi, chaque couche R,V et B est susceptible de coder 2^8 (soit 256) niveaux dont 220 utiles. Le mélange de ces trois couches permet théoriquement au DV de coder 2203 (soit 10 648 000) couleurs différentes. Ce taux de quantification permet d'obtenir un assez bon rapport signal/bruit (74 dB théoriques), mais il limite les possibilités d'étalonnage.

La compression qui utilise les outils du MJPEG : le principe de compression est du type JPEG, elle s'effectue selon le contenu de l'image **en intra-trame ou en intra-image**. Cela permet de tenir compte, quand elles existent, des redondances entre les deux trames d'une image et d'obtenir une meilleure efficacité de compression. En aucun cas, il n'est tenu compte de la redondance temporelle entre les images (principe du MPEG avec qui le DV diffère donc), chacune d'entre-elles est codée séparément et ne dépend d'aucune autre ce qui permet un montage à l'image près. Le débit initial du signal vidéo 4:2:0 ou 4:1:1 est de 125 Mbit/s. Il est réduit, après compression à **25 Mbit/s** d'où un taux de compression de l'ordre de 5:1.

Il existe une variante appelée **DVCPRO** reprenant les mêmes principes mais avec un débit de **50 Mb/s**. Le **DVCPRO HD** gère la haute définition et offre des débits de **100Mb/s**.

Principes de l'audiovisuel - Vincent MAZO - EFREI 2012/2013

Le format DV est très populaire pour les tournages au niveau semi-professionnel, c'est une solution numérique à faible coût. On dispose par exemple du transfert natif via **firewire** qui permet de transmettre les données en conservant la mise en forme du DV.

En plus des spécifications vidéos, le format DV peut traiter 4 pistes au format PCM non compressé, échantillonnées à 32 kHz et codées en 12 bits, ou 2 pistes en 48 kHz et codées sur 16 bits.

Le débit total enregistré (vidéo + audio + signaux auxiliaires) avoisine les 28,8 Mbit/s. On considère usuellement qu'une heure de DV équivaut à environ 13 Go.

2) HDV



Le HDV (High Definition Video) est le premier format vidéo haute définition (par opposition à la vidéo standard telle celle de la télévision à 625 ou 525 lignes) pour caméscopes grand public et semi-professionnel.

Tandis que le format SDTV (standard-definition television) est en 720×576 pixels, le format HDV fait 1 440×1 080 pixels (une interpolation numérique permet de restituer les 1 920×1 080 pixels annoncés par anamorphose).

C'est l'évolution du format DV pour répondre aux besoins de la haute définition. Les cassettes utilisées pour enregistrer en HDV sont les mêmes qu'en DV.

- Format vidéo : 720/60p, 720/30p, 720/50p, 720/25p, 1080/60i, 1080/50i, 1080/25p (i pour entrelacé, et p pour progressif)
- Rapport d'aspect des pixels : 1,0 (1280x720) et 1,33 (1440x1080)
- Compression : MPEG-2 Video (profile & level: MP@H-14)
- Fréquence d'échantillonnage de la luminance : 74,25 MHz (1280x720), 55,7 MHz (1440x1080)
- Structure d'échantillonnage : 4:2:0
- Quantification : 8 bits (pour la luminance et la chrominance)
- Débit après compression : environ 19 Mbit/s (1280x720) ou 25 Mbit/s (1440x1080)

De même que pour le DV, HDV prend en charge l'audio :

- Compression audio : MPEG-1 Audio Layer II
- Fréquence d'échantillonnage audio : 48 kHz
- Quantification audio : 16 bits
- Débit audio après compression : 384 kbit/s
- Nombre de pistes audio : Stéréo (2 canaux)
- Type de flux : Transport Stream Packetized Elementary Stream
- Interface du flux : IEEE1394 FireWire (MPEG2-TS)

Ce codec populaire est lui aussi assez répandu. Cependant les compromis importants au niveau de la qualité générale (taux de compression élevé) et de la définition réduite donnent des résultats mitigés.

3) MPEG

La famille MPEG regroupe un certains nombres de techniques héritées de la compression JPEG en ajoutant la notion de compression temporelle. En effet, en vidéo les images successives comportent de nombreuses ressemblances. Les algorithmes MPEG en tirent partie pour réduire le poids des données.

A) MPEG1

On trouve ainsi le MPEG1 avec les spécifications suivantes :

- 352×240 pixels à 30 images par seconde en NTSC
- 352×288 pixels à 25 images par seconde en PAL/SECAM

Le MPEG-1 permet d'obtenir des débits de l'ordre de 1,2 Mbit/s (exploitable sur un lecteur de CD-ROM).

Le MPEG-1 permet d'encoder une vidéo grâce à plusieurs techniques :

- *Intra coded frames (Frames I, correspondant à un codage interne)* : les images sont codées séparément sans faire référence aux images précédentes
- *Predictive coded frames (frames P ou codage prédictif)* : les images sont décrites par différence avec les images précédentes
- *Bidirectionally predictive coded frames (Frames B)* : les images sont décrites par différence avec l'image précédente et l'image suivante
- *DC Coded frames* : les images sont décodées en faisant des moyennes par bloc

Le format audio MP3 repose sur la compression adoptée par le format MPEG1 layer 3.

Le codec MPEG1 est aujourd'hui de moins en moins utilisé du fait des importantes limitations qu'il impose.

B) MPEG2

MPEG-2 est la norme de seconde génération (1994) du Moving Picture Experts Group qui fait suite à MPEG-1. MPEG-2 définit les aspects compression de l'image et du son et le transport à travers des réseaux pour la télévision numérique.

Les aspects Systèmes (synchronisation, transport, stockage) sont définis dans la norme ISO/CEI 13818-1 (Codage générique des images animées et du son associé - Partie Systèmes). Les aspects compression, quant à eux, sont définis dans les normes ISO/CEI 13818-2 et 3 (Codage générique des images animées et du son associé - Parties vidéo, audio).

Ce format vidéo est utilisé pour les DVD et SVCD avec différentes résolutions d'image. Ce format est également utilisé dans la diffusion de télévision numérique par satellite, câble, réseau de télécommunications ou hertzien (TNT).

De manière générale, c'est un codec qui est encore utilisé aujourd'hui du fait de sa mise en œuvre très simple compte tenu des besoins actuels.

Principes de l'audiovisuel - Vincent MAZO - EFREI 2012/2013

Le format MPEG2 supporte une grande variété de réglages possibles pour permettre de s'adapter au mieux à l'utilisation recherchée (qualité élevée en HD, diffusion en SD, etc...). On définit ainsi des profils et des niveaux de référence.

Le profil indique les techniques particulières mises en œuvre (algorithme de compression, sous échantillonnage de la chrominance...)

MPEG-2 Profiles

Abbr.	Name	Picture Coding Types	Chroma Format	Aspect Ratios	Scalable modes	Intra DC Precision
SP	Simple profile	I, P	4:2:0	square pixels, 4:3, or 16:9	none	8, 9, 10
MP	Main profile	I, P, B	4:2:0	square pixels, 4:3, or 16:9	none	8, 9, 10
SNR	SNR Scalable profile	I, P, B	4:2:0	square pixels, 4:3, or 16:9	SNR (signal-to-noise ratio) scalable	8, 9, 10
Spatial	Spatially Scalable profile	I, P, B	4:2:0	square pixels, 4:3, or 16:9	SNR- or spatial-scalable	8, 9, 10
HP	High profile	I, P, B	4:2:2 or 4:2:0	square pixels, 4:3, or 16:9	SNR- or spatial-scalable	8, 9, 10, 11
422	4:2:2 profile	I, P, B	4:2:2 or 4:2:0	square pixels, 4:3, or 16:9	none	8, 9, 10, 11
MVP	Multi-view profile	I, P, B	4:2:0	square pixels, 4:3, or 16:9	Temporal	8, 9, 10

Le niveau précise la taille de l'image et le débit maximum notamment.

MPEG-2 Levels

Abbr.	Name	Frame rates (Hz)	Max horizontal resolution	Max vertical resolution	Max luminance samples per second (approximately height x width x framerate)	Max bit rate in Main profile (Mbit/s)
LL	Low Level	23.976, 24, 25, 29.97, 30	352	288	3,041,280	4
ML	Main Level	23.976, 24, 25, 29.97, 30	720	576	10,368,000, except in High profile, where constraint is 14,475,600 for 4:2:0 and 11,059,200 for 4:2:2	15
H-14	High 1440	23.976, 24, 25, 29.97, 30, 50, 59.94, 60	1440	1152	47,001,600, except that in High profile with 4:2:0, constraint is 62,668,800	60
HL	High Level	23.976, 24, 25, 29.97, 30, 50, 59.94, 60	1920	1152	62,668,800, except that in High profile with 4:2:0, constraint is 83,558,400	80

On parle par exemple d'une vidéo compressée en mpeg2 en MP@ML pour indiquer que l'on a utilisé le profil **MP** et le niveau **ML**.

Le Mpeg2 est très répandu:

BETACAM SX, IMX, XDCAM HDet EX, HDCAM-SR,BLU-RAY, DVD, VCD, Diffusion TNT SD (DVB T), Télévision par ADSL... Tous ces formats utilisent le codec Mpeg2 en tirant profit des nombreux profils disponibles.

C) MPEG4

La norme MPEG-4 spécifie d'abord des techniques pour gérer le contenu de scènes comprenant un ou plusieurs objets audio-vidéo. Contrairement à MPEG-2 qui visait uniquement des usages liés à la télévision numérique (diffusion DVB et DVD), les usages de MPEG-4 englobent toutes les nouvelles applications multimédias comme le téléchargement et le streaming sur Internet, le multimédia sur téléphone mobile, la radio numérique, les jeux vidéo, la télévision et les supports haute définition. La technologie a bien évolué depuis 1994 !

La norme spécifie de nouveaux codecs audio et vidéo et enrichit les contenus multimédia, en ajoutant de nouvelles applications comme le VRML (étendu), la prise en charge de la gestion des droits numériques et de plusieurs types d'interactivités.

La norme est divisée en plusieurs « parts », ou parties, qui spécifient un type de codage particulier. Dans chaque partie plusieurs « profils » (collection d'algorithmes) et « niveaux »

Principes de l'audiovisuel - Vincent MAZO - EFREI 2012/2013

(contraintes quantitatives) sont définis. Un consortium industriel désirant utiliser MPEG-4 choisit une ou plusieurs parties de la norme et, pour chaque partie, il peut sélectionner un ou plusieurs profils et niveaux correspondant à ses besoins.

- La partie 1 décrit la synchronisation et le multiplexage de la vidéo et de l'audio.
- La partie 2 est un codec de compression pour les signaux vidéo. L'un des nombreux profils de ce codec est l'ASP (Advanced Simple Profile).
- La partie 3 est une norme de compression pour le codage perceptuel et les signaux audio ; elle spécifie notamment le format audio AAC.
- La partie 4 décrit les procédures pour les tests de conformité.
- La partie 5 fournit des logiciels de référence des autres parties de la norme.
- La partie 6 décrit le Delivery Multimedia Integration Framework (DMIF).
- La partie 7 fournit des implémentations optimisées (cf. partie 5)
- La partie 8 décrit les méthodes de transport du MPEG-4 sur IP.
- La partie 9 fournit des implémentations matérielles des autres parties à titre d'illustration.
- La partie 10 est une norme avancée de compression vidéo appelée aussi H.264 ou AVC (Advanced Video Codec), et qui comporte une extension appelée SVC (Scalable Video Coding).
- La partie 11 spécifie la description de scène et moteur d'application.
- La partie 12 spécifie le format de fichier ISO Base media.
- La partie 13 fournit les extensions de gestion et de protection de la propriété intellectuelle (IPMP).
- La partie 14 spécifie le format de fichier MP4.
- La partie 15 spécifie le format de fichier du codec AVC (partie 10), sur base de la partie 12.
- La partie 16 fournit l'extension du cadre d'animation (AFX).
- La partie 17 spécifie le format de sous-titrage Timed Text.
- La partie 18 spécifie la compression et transmission de polices de caractères.
- La partie 19 décrit le flux de texture synthétisé.
- La partie 20 spécifie la représentation « allégée » de description de scène (pour applications mobiles).
- La partie 21 spécifie MPEG-J GFX.
- La partie 22 spécifie le format Open Font, basé sur OpenType.

Les outils AVC ou H.264 sont très utilisés de nos jours. Les codecs matériels et logiciels sont de plus en plus performants. MPEG4 remplace progressivement MPEG2.

On considère qu'à qualité comparable le codec MPEG4 est 2 fois plus léger que MPEG2.

4) DIVX

DivX 3.11 et les versions précédentes réfèrent généralement à une version « piratée » du codec de l'époque de Microsoft (MPEG-4 modifié).

DivX a été créé aux alentours de 1999 à Montpellier. Le codec de Microsoft, initialement prévu pour l'encapsulation dans un format de streaming ASF a été transformé pour permettre son stockage à l'intérieur de fichiers AVI — c'est le codec DivX 3.11.

DivXNetworks, Inc. a par la suite produit une version 4 du codec totalement indépendante du codec de Microsoft (et de sa version 3.11) afin d'éviter des problèmes avec la firme de Redmond. DivXNetworks a demandé le dépôt d'un brevet pour son nouveau codec, qui suit la certification de la norme MPEG4 Advanced Simple Profile.

Le codec DivX est disponible en téléchargement depuis le site de DivXNetworks, Inc. pour les systèmes d'exploitation Windows (Vista, XP, 2000, ME, 98), GNU/Linux et Mac OS X.

Ce n'est pas un logiciel libre et son code source n'est pas disponible, mais une version publique — appelée OpenDivX — a été mise à disposition par DivXNetworks au début de l'année 2001.

Xvid fut alors créé à la suite d'une controverse sur les réelles intentions qu'avait DivXNetworks quand il publia OpenDivX, connu aussi sous le nom de code Project Mayo (également nom donné au site web hébergeant le projet) : certains l'accusèrent en effet d'avoir lancé le projet dans le seul but de récolter des idées pour les réutiliser ensuite dans le codec DivX 4. Xvid, logiciel libre cette fois, parvenu sans conteste à la maturité, est à présent maintenu par un groupe indépendant et ouvrant sur plusieurs ramifications d'études et de travail. Aujourd'hui, OpenDivX et Project Mayo ont disparu.

DIVX et les codecs dérivés reprennent les mêmes outils que MPEG4. Ces codecs grands public sont peu utilisés dans le monde professionnel.

5) BETA DIGITAL



Le format Betacam numérique enregistre le signal vidéo en utilisant une compression intra-image de type MJPEG (soit 50 trames compressées JPEG par seconde) et donc - entre autres - une compression de type DCT sur le signal au format composantes, avec une quantification sur 10 bits (YUV de type 4:2:2). Les modes d'enregistrement 625 et 525 lignes utilisent respectivement 720 x 576 et 720 x 486 points, avec un débit vidéo moyen de l'ordre de 130 Mbit/s.

La bande audio est enregistrée sur quatre canaux, échantillonnés à 48 kHz, quantifiés sur 20 bit, codage PCM.

C'est le format professionnel par excellence pour ce qui est de la définition standard.

6) HDCAM



Le HDCAM est une évolution de 1997 qui prend en charge la haute définition avec les caractéristiques suivantes :

- Taille de l'image : 1440 x 1080 (1920x1080 sous-échantillonnée)
- Compression : De type M-JPEG à 135 Mbit/s
- Quantification : 8 bits
- Structure vidéo YCrCb : 3:1:1
- Débit enregistrement :
 - 111.863 Mbit/s en 23,976PsF,
 - 111.975 Mbit/s en 24PsF,
 - 116.640 Mbit/s en 25PsF et 50i,
 - 139.828 Mbit/s en 29,97PsF et 59,94i
- Taux de compression : 7:1
- Cadence enregistrement :
 - mode progressif : 23.98p, 24p, 25p, 30p
 - mode entrelacé : 50i, 59.94i, 60i sur CINEALTA (HDW-F)
- Audio : 4 canaux d'audio numérique AES/EBU 20 bit-48 kHz

Le HDCAM est une amélioration du format BETA mais la prise en charge de la HD n'est pas optimale du fait de l'anamorphose et de la faible quantification.

7) HDCAM SR



Ce format créé en 2003 par Sony répond à la problématique d'enregistrement en HD.

- Taille de l'image : 1920x1080
- Compression : MPEG4 Studio Profile/600 Mbit/s ou 800 Mbit/s
- Quantification : 10 bits
- Structure vidéo YCrCb : 4:2:2 ou 4:4:4 (RVB)
- Débit enregistrement : ~440Mbit/s ou 880Mbit/s
- Taux de compression :
 - 2,7:1 (HDCAM SR à 440 Mbits en 4:2:2) ;
 - 4:1 (HDCAM SR à 440 Mbits en 4:4:4);
 - 2:1 (HDCAM SR à 880 Mbits en 4:4:4).
- Cadence enregistrement :
 - mode progressif : 23,98p, 24p, 25p, 30p, 50p, 60p
 - mode entrelacé : 50i, 59.94i, 60i
- Débit du flux de données : 593Mbit/s ou 1186Mbit/s
- Audio : 12 canaux d'audio numérique AES/EBU 24-bit/48 kHz

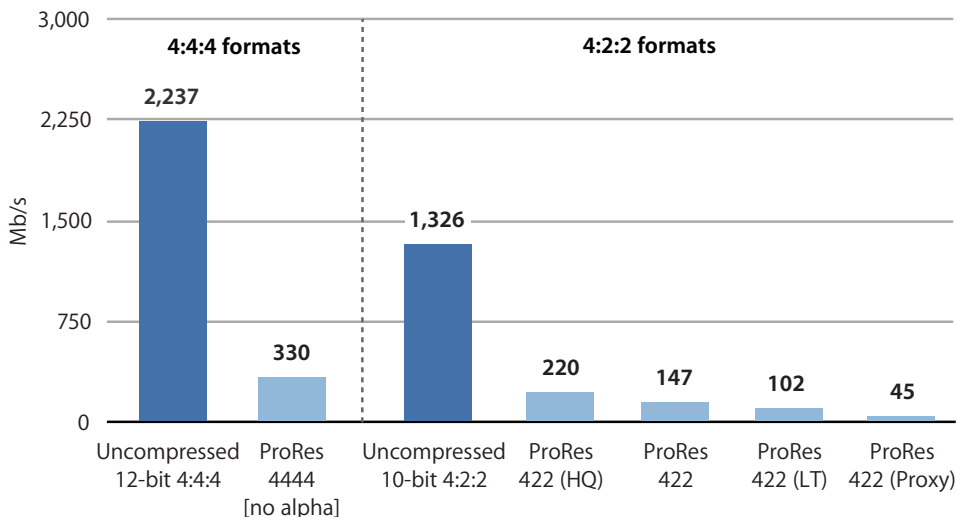
8) PRORES

Prores est le codec promulgué par Apple pour ses solutions de montage Final cut. Il est destiné aux étapes de fabrication et d'édition d'un programme plutôt qu'à sa diffusion. On dispose alors d'un CODEC proposant une qualité proche du flux non compressé pour un poids et un débit largement inférieurs.

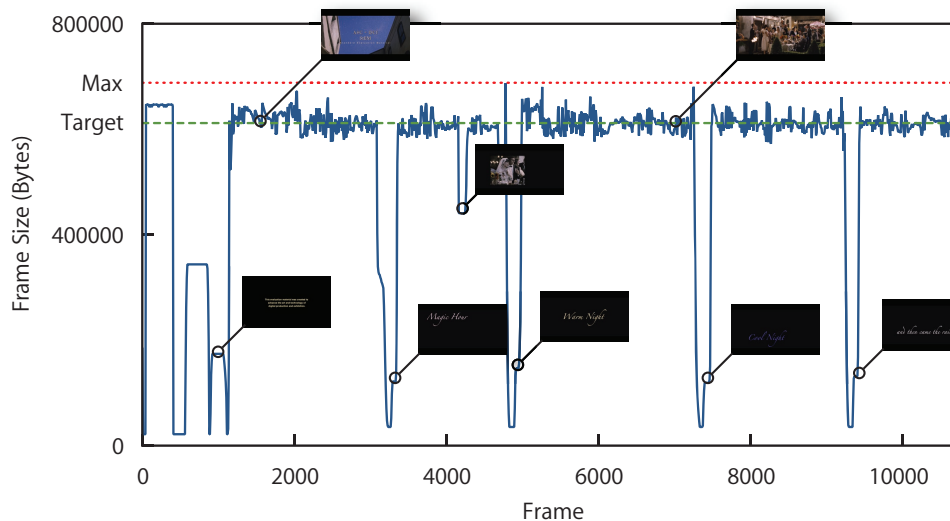
- Résolution 2K, 1080 , 1280, 720
- Structure d'échantillonnage en 4:2:2, 4:4:4 ou 4:4:4:4
- Quantification sur 10-bit
- Encodage intra-image à débit variable
- Débits HD : 147 Mbit/s (PRORES) et 220 Mbit/s (PRORES HQ) en HD @60i
- Débits SD : 42 Mbit/s (PRORES) et 63 Mbit/s (PRORES HQ) en SD @59,94i

Le codage et le décodage sont très rapides (avec la possibilité d'afficher en demi résolution). Des caméras récentes sont capables d'enregistrer directement en Prores, ce qui facilite le workflow en supprimant l'étape de transcodage des fichiers.

Data Rates - Uncompressed and Apple ProRes at 1920 x 1080, 29.97 fps



Compressed Frame Sizes - Apple ProRes 422



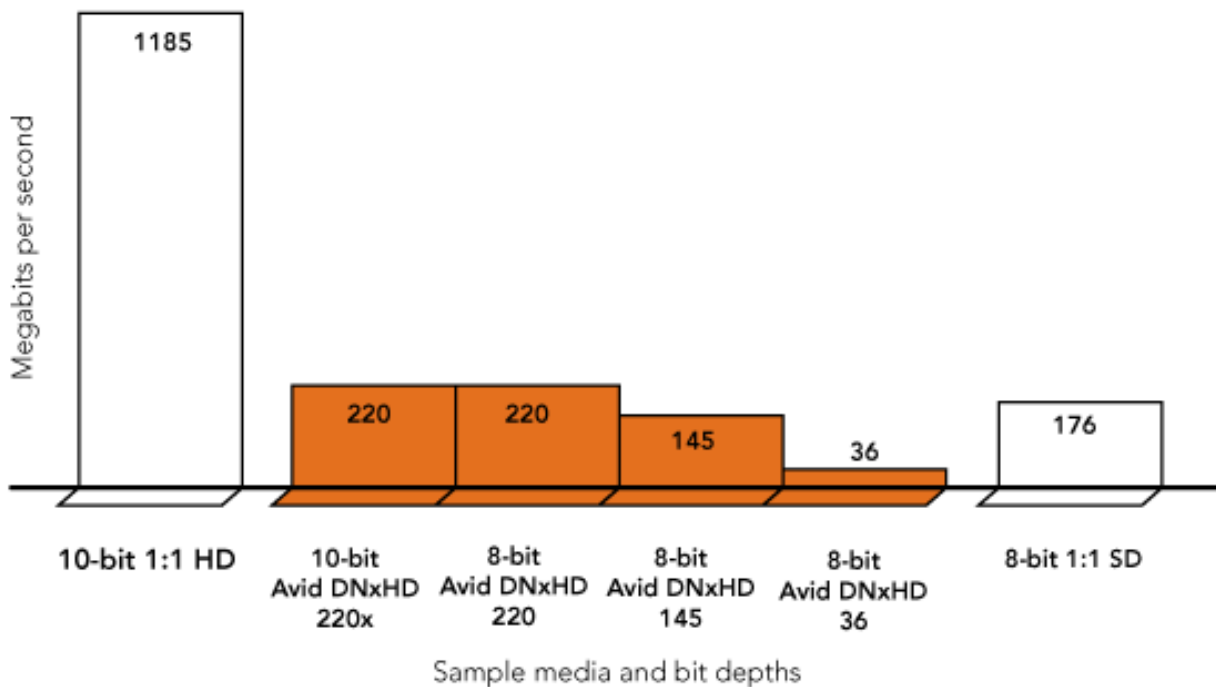
9) DnxHD

C'est le codec propriétaire intermédiaire proposé par AVID pour ses stations de montage. DnxHD est l'abréviation de Digital Nonlinear Extensible High Definition.

Il existe en différente qualité dont les débits sont respectivement de 36, 145 et 220 Mbits/s. La version à 220 Mbit/s autorise une quantification sur 8 ou 10 bits alors que les autres versions fonctionnent exclusivement en 8 bits.

Le tableau ci-dessous compare les caractéristiques du codec Avid DNxHD et celles d'autres codecs HD courants :

Format	Avid DNxHD 36	Avid DNxHD 145	Avid DNxHD 220	DVCPRO HD	HDCAM	HDCAM SR
Bit Depth	8-bit	8-bit	8- and 10-bit	8-bit	8-bit	10-bit
Sampling	4:2:2	4:2:2	4:2:2	1280 Y samples 4:2:2	1440 Y samples 3:1:1	4:2:2
Bandwidth	36 Mb/sec	145 Mb/sec	220 Mb/sec	100 Mb/sec	135 Mb/sec	440 Mb/sec



III. Spécifications des formats vidéos courants

Spécifications Techniques								
Format	Constructeur	Taille média	Echantillonnage	Ratio	Débit	codec	media	Transport
HDCAM SR	SONY		4 :2 :2 ou 4 :4 :4	2,7 :1 ou 4 :1	440 - 880 Mbps			HDSDI
DVCpro HD	Panasonic	6.35mm	4:2:2/ 8 bits	6,7:1	100 Mbps	4DV	metal particule	SDTI
HDcam	SONY	½-pouce	17:6:6/ 8-10 bits	4,4:1	144 Mbps		metal particule	HDSDI/SDTI
HDV	SONY		4 :2 :0		25 Mbps	IBP GOP 6		IEEE1394
IMX	Sony	½-pouce	4:2:2 / 8 bits	3.3:1	50Mbits/s	MPEG-2		SDTI-CP
D1	Sony/BTS	19mm	4:2:2 / 8 bits	1:1	216 Mbits/s	ITU-R 601	metal	SDI
D2	Sony	19mm	composite digital	1:1	142 Mbits/s		metal	composite digital I/O
D3	Panasonic	½-pouce	composite digital	1:1	142 Mbits/s		metal	composite digital I/O
D5 / D5 HD	Panasonic	½-pouce	4:2:2 / 10 bits	4:1 @ 8 bit 5:1 @ 10 bit 1:1 @ 576/50i	288 Mbits/s		metal	SDI
D6 (voodoo)		3/4pouce	22:11:11 /10-8 bits	1:1	1 Gbits/s		metal particule	
D7(DVCPR O50)	Panasonic	6.35mm	4:2:2 / 8 bits	3.3:1	50Mbits/s	Double DV	metal particule	SDI
DCT	Ampex	19mm	4:2:2 / 10 bits	2:1	137 Mbits/s	DCT	metal	SDI
Digital Betacam	Sony/Thomson	½-pouce	4:2:2 / 10 bits	2.2:1	128 Mbits/s	DCT	metal	SDI
BetacamSX	Sony	½-pouce	4:2:2@ML	10:1 / Gop : IB	18Mbits/s	MPEG-2		QSDI/SDI
DVCAM	Sony	6.35mm	4:1:1@ML	5:1 / I frames	25Mbits/s	DV	metal	SDI/QSDI/1394
DVCPRO	Panasonic	6.35mm	4:1:1@ML	5:1 I frames	25Mbits/s	DV	metal particule	SDI/1394
DVCPRO50	Panasonic	6.35mm	4:2:2@ML	3.3:1 I frames	50Mbits/s	Double -DV	metal particule	SDI
Digital-S(D9)	JVC	½-pouce	4:2:2	3.3:1	50Mbits/s	Double -DV		SDI
miniDV et D8	consortium	6.35mm	4:1:1	5:1 I frames	25Mbits/s	DV	metal	S-Video/1394
D-VHS	consortium	½-pouce		variable	14.1Mbits/s	MPEG-2		S-Video/1394
D-VHS HD	JVC	½-pouce			19-28Mbits/s	720p/1080i		S-Video/1394
DVD	consortium	12cm	4:2:0 P@ML	variable / longs GOP	6Mbits/s environ	MPEG-2	polycarbonate	S-Video/1394
VideoCD 2.0	consortium	12cm	4 :1 :1	variable / N:12 M:3	1.2Mbits/s	MPEG-1	polycarbonate	composite
Laserdisc	consortium	12-pouces	Composite			NTSC/ Pal	PMMA	composite

IV. Qu'est-ce qu'un conteneur ?

Un fichier multimédia contenant des éléments audio et vidéo peut se décomposer en plusieurs parties bien distinctes :

- Le flux vidéo encodé (essence)
- Le flux audio encodé (essence)
- Les métadonnées
- Le conteneur

Ainsi, lorsqu'on parle d'un fichier AVI ou MOV, il s'agit du conteneur. Il contient en général un flux vidéo, encodé avec un codec vidéo spécifique et un flux audio encodé avec un codec audio spécifique.

Un conteneur (wrapper ou container en anglais) est un format de fichier qui peut contenir divers types de données compressées à l'aide de codecs normalisés. Ce fichier pourra contenir différents types de données, incluant des fonctionnalités supplémentaires. Dans les cas les plus simples, on se limitera à une seule piste vidéo, alors que des conteneurs plus évolués pourront gérer plusieurs pistes audios, des sous-titres, des chapitres, des métadonnées additionnelles...

- un ou plusieurs flux vidéo (ce qui permet par exemple de regarder une scène filmée sous plusieurs angles différents) ;
- un ou plusieurs flux audio (ce qui permet d'obtenir une version multilingue du média) ;
- des sous-titres (ce qui permet également le multilingue) ;
- des éléments de chapitrage (de la même manière que sur les DVD) ;
- des métadonnées (par exemple le titre du média, le nom du réalisateur, la date, etc.) ;
- une description des flux que contient le conteneur ;
- éventuellement d'autres données.

Le rôle du conteneur est alors d'assurer la synchronisation de tous ces éléments lors de leur présentation.

Il existe une multitude de conteneurs, voici les plus répandus :

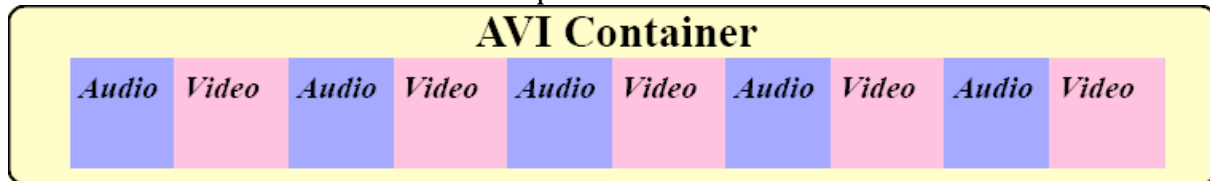
- Audio Video Interleave (extension .avi) : développé par Microsoft.
- MPEG 1/2/4 (extensions .mpg, .mpeg) : *codecs* et *conteneurs* défini par le Moving Picture Experts Group.
- Matroska (extensions .mkv, .mka, .mks) : développé par CoreCodec, Inc..
- Advanced Systems Format (extensions .asf, .wmv, .wma (uniquement lorsqu'il ne contient que de l'audio)) : développé par Microsoft.
- QuickTime (extension .mov) : développé par Apple.
- Ogg (extensions .ogg, .ogv, .oga, .ogx), ainsi que Ogg Media (extension .ogm) : développé par Xiph.org.
- 3gp (extensions .3gp, .3g2) : défini par le 3GPP.
- WebM (extensions .webm, .weba) : basé sur *Matroska*, développé par Google.
- NUT (extension .nut) : développé par des développeurs de MPlayer et FFmpeg.
- RealMedia (extension .rm) : développé par RealNetworks.
- Material eXchange Format (extension .mxf) : standard de la SMPTE.
- Advanced Stream Redirector (extension .asx) : développé par Microsoft.
- Transport Stream (extension .ts) : protocole de communication développé par le Moving Picture Experts Group.

Principes de l'audiovisuel - Vincent MAZO - EFREI 2012/2013

Il existe également des conteneurs spécifiques pour l'audio :

- WAVEform audio format (extension .wav) : développé par Microsoft et IBM, largement utilisé sur la plate-forme Windows.
- Broadcast Wave Format (extension .wav) : extension du *WAVEform audio format*, développé par Microsoft.
- Audio Interchange File Format (extensions .aif, .aiff, .aifc) : développé par Apple, largement utilisé sur la plate-forme Mac OS.
- Au (extensions .au, .snd) : développé par Sun Microsystems.
- Core Audio Format (extension .caf) : également développé par Apple.

Les différents contenus sont ainsi encapsulés à l'intérieur du conteneur :



Les données sont découpées en paquets qui seront décodés par le codec approprié.

V. Conteneurs usuels

1) AVI

AVI (Audio Video Interleaved) est un conteneur mis au point par Microsoft, c'est le plus utilisé dans le mode Windows. Il permet de rassembler et de synchroniser des flux audios et vidéos en utilisant une très grande variété de codecs différents : DV, MJPEG, DIVX, XVID, MP3, PCM...

Un fichier AVI accepte aussi bien les contenus en définition standard ou en haute définition

Le format AVI permet de réunir en un seul fichier une piste vidéo et jusqu'à 99 pistes audio, ce qui permet de bénéficier, par exemple, de plusieurs langues pour un même film, mais il ne prend pas en charge le sous-titrage ni le chapitrage.

.AVI	CODEC VIDEO	DV, MV1, MJPEG, DIVX, Indeo, WMV9,...
	CODEC AUDIO	Sans compression, MP3...

2) MOV

C'est l'équivalent Dans le monde Mac, il repose sur les technologies développées par Apple au sein du logiciel Quicktime. Comme pour AVI, il est possible d'utiliser une multitude codecs.

Le formalisme est cependant différent. Un fichier Quicktime contient une ou plusieurs pistes ; chacune comporte un type de données particulier : audio, vidéo, effet ou texte (pour les sous-titres) ou une référence vers un média situé dans un autre fichier.

Principes de l'audiovisuel - Vincent MAZO - EFREI 2012/2013

Par exemple un fichier « référence.mov » peut contenir une piste vidéo faisant référence au contenu du fichier « source.mov ». référence.mov ne contient pas réellement les données mais uniquement un lien.

Ce système offre une grande souplesse d'utilisation et permet d'économiser une place importante en mémoire. On ne copie plus les données, on utilise des liens pour les retrouver.

.MOV	CODEC VIDEO	MJPEG, MPEG4, DV, SORENSON, DnxHD, PRORES, Blackmagic,...
	CODEC AUDIO	Sans Compression, AIC, AAC...

3) DV – HDV

Le DV est un CODEC mais également un conteneur.

Il contiendra des essences utilisant le codec DV

.DV	CODEC VIDEO	DV
	CODEC AUDIO	DV

De la même manière HDV est également un conteneur

.HDV	CODEC VIDEO	HDV
	CODEC AUDIO	HDV

Ces conteneurs ont l'avantage de permettre une utilisation en natif. Le montage se fait alors directement dans le codec utilisé pour l'enregistrement.

4) MPG / MP2 / MP4

MPEG dispose également de conteneurs spécifiques pour le mpeg1, le mpeg2 et le mpeg4 :

.MPG	CODEC VIDEO	MPEG1 (mpv)
	CODEC AUDIO	MPEG LAYER I, II, ou III (mpa)

.MP2	CODEC VIDEO	MPEG2 (m2v)
	CODEC AUDIO	MPEG LAYER III, MPEG2 AAC (m2a)

.MP4	CODEC VIDEO	MPEG4 (m4v) SP, ASP, AVC (H264) ou H263
	CODEC AUDIO	MPEG LAYER III, MPEG4 AAC... (m4a)

Principes de l'audiovisuel - Vincent MAZO - EFREI 2012/2013

Il n'existe pas de norme de MPEG3, celle-ci a été englobée dans la version MPEG4.

Les conteneurs MPEG sont des standards universels qui ne s'appuient pas sur des contraintes propriétaires, de plus ils offrent une grande souplesse quand aux contenus pris en charge... Il existe une infinité de combinaisons en MPEG.

Les fichiers 3GP utilisent par exemple, une partie du MPEG4 destinée à la diffusion pour les téléphones portables avec une vidéo en H263 et un son encodé en AMR.

5) ASF / WMA / WMV

Windows media est un conteneur développé par Microsoft.

ASF ne spécifie pas les formats de compression des flux audio et vidéo qu'il contient, mais uniquement la structure de ces flux. Les fichiers au format ASF peuvent être encodés avec un très grand nombre de codecs. Parmi les objectifs assignés à ce format par Microsoft, on peut noter la lecture en réseau via le streaming (HTTP, MMS...), ou sur des supports de stockages locaux.

Le format ASF est basé sur des objets sérialisés qui prennent ici la forme de séquences d'octets identifiées par un marqueur GUID (Globally Unique Identifier). Les fichiers ASF peuvent aussi contenir des métadonnées telles que le nom de l'artiste, le titre, l'album, l'année, le genre, le réalisateur... C'est l'équivalent des tags ID3 des fichiers MP3. Le conteneur ASF fournit également un ensemble complet d'outils pour la Gestion numérique des droits (DRM) dans les formats WMA et WMV.

Au début de son histoire, les fichiers de ce format se présentaient avec l'extension .asf, mais elle a été depuis remplacée par les extensions .wma (Windows Media Audio, audio seulement) et .wmv (Windows Media Video, audio et vidéo). Les fichiers ASF ont le type MIME application/vnd.ms-asf ou video/x-ms-asf

.ASF	CODEC VIDEO	MPEG4, WMV1-2-3, VC1(WMV9)
.WMV .WMA	CODEC AUDIO	WMA1-2-3

6) MJ2

C'est le conteneur des vidéos encodées en MJPEG 2000.

Il est très utilisé dans le cinéma numérique où on pourra exploiter toute la souplesse et la qualité intrinsèque du CODEC basé sur JPEG2000.

Il autorise de nombreux formats audios

.MJ2	CODEC VIDEO	Motion JPEG2000
	CODEC AUDIO	PCM, MPG4 audio ...

7) MKV (MATROVSKA)

Ce conteneur peut regrouper au sein d'un même fichier (généralement avec l'extension .mkv) plusieurs pistes vidéo et audio ainsi que des sous-titres et des chapitres.

.MKV	CODEC VIDEO	DIVX, XVID, REALVIDEO, H264,...
	CODEC AUDIO	OGG, AAC, MP2, MP3, AC3, DTS, PCM...
	Sous-titres	SRT, SSA, ASS, SUB/IDX...

- support de la plupart des flux vidéo et audio existants ;
- grande capacité pistes de sous-titres de différents formats ;
- support des chapitres ;
- possibilité d'étiquetage des pistes audio, vidéo et les sous-titres ;
- possibilité d'intégration de pièces jointes, comme les images d'illustration (couverture/jaquette).

On considère MATROSKA comme le successeur du conteneur AVI.

VI. Les formats d'échange

1) EDL

EDL (Edit Decision List) est une liste regroupant les time codes associés à un montage. Elle contient les TC de début et de fin pour chaque plan ainsi que le nom de la source associée.

Le cas typique de l'utilisation d'une EDL est lors d'une conformation, lorsque l'on souhaite remplacer des medias en basse définition par des médias de meilleure qualité.

					SOURCE TIME- CODE IN	SOURCE TIME- CODE OUT	EDIT TIME- CODE IN	EDIT TIME- CODE OUT
TITLE: KENNYV1 FCM: NON-DROP FRAME								
001	XFER01	V	C		01:21:39:26	01:21:40:09	01:00:00:00	01:00:00:13
002	XFER01	V	C		01:20:16:10	01:20:20:12	01:00:00:13	01:00:04:15
003	XFER01	V	C		01:07:28:28	01:07:34:17	01:00:04:15	01:00:10:04
004	XFER01	V	C		01:21:32:27	01:21:35:03	01:00:10:04	01:00:12:10
004	BL	V	D	015	00:00:00:00	00:00:00:15	01:00:12:10	01:00:12:25
005	BL	V	C		00:00:00:00	00:00:00:00	01:00:12:27	01:00:12:27
005	XFER01	V	D	015	01:04:26:27	01:04:31:13	01:00:12:27	01:00:17:13
006	XFER02	V	C		02:33:36:25	02:33:38:03	01:00:17:13	01:00:18:21
007	XFER02	V	C		02:37:13:01	02:37:13:15	01:00:18:21	01:00:19:05

2) XML

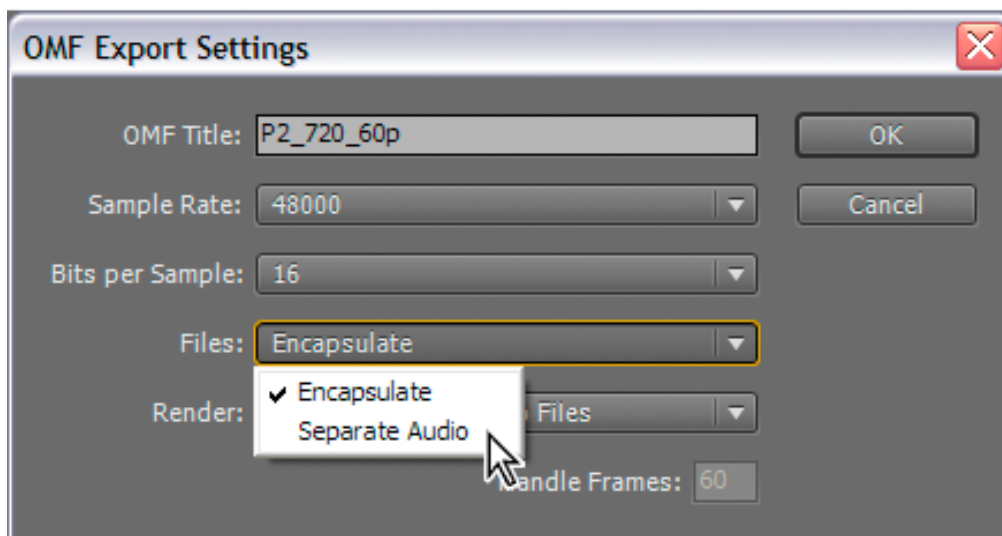
XML est avant tout une famille de langage informatique reposant sur le principe de la description de contenu en utilisant des balises. Les fichiers XML permettront de décrire un montage en faisant référence aux médias utilisés en précisant les time code des points de coupe...etc. C'est une sorte de conteneur vide qui ne fait que pointer vers les médias en décrivant la manière de les utiliser ainsi que différentes métadonnées (chapitres, noms...)

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE xmeml>
<xmeml version="3">
  <clip id="Jeremy Solo">
    <name>Jeremy Solo</name>
    <duration>188</duration>
    <rate>
      <ntsc>TRUE</ntsc>
      <timebase>30</timebase>
    </rate>
    <file id="Jeremy Solo1">
      <name>Jeremy Solo1</name>
      <pathurl>file://localhost/Jeremy%20Solo</pathurl>
      <rate>
        <timebase>30</timebase>
        <ntsc>TRUE</ntsc>
      </rate>
      <duration>188</duration>
    </file>
  </clip>
</xmeml>
```

On utilise par exemple ce format pour exporter un montage depuis final cut.

3) OMF

OMF est un conteneur multiplateforme très utilisé pour échanger des médias entre différentes plateformes de production. Typiquement on exporte un fichier OMF depuis le banc de montage pour l'envoyer au mixage son etc...



On pourra au choix encapsuler les médias dans le fichier OMF ou utiliser un système de référence vers des médias indépendants.

4) MXF

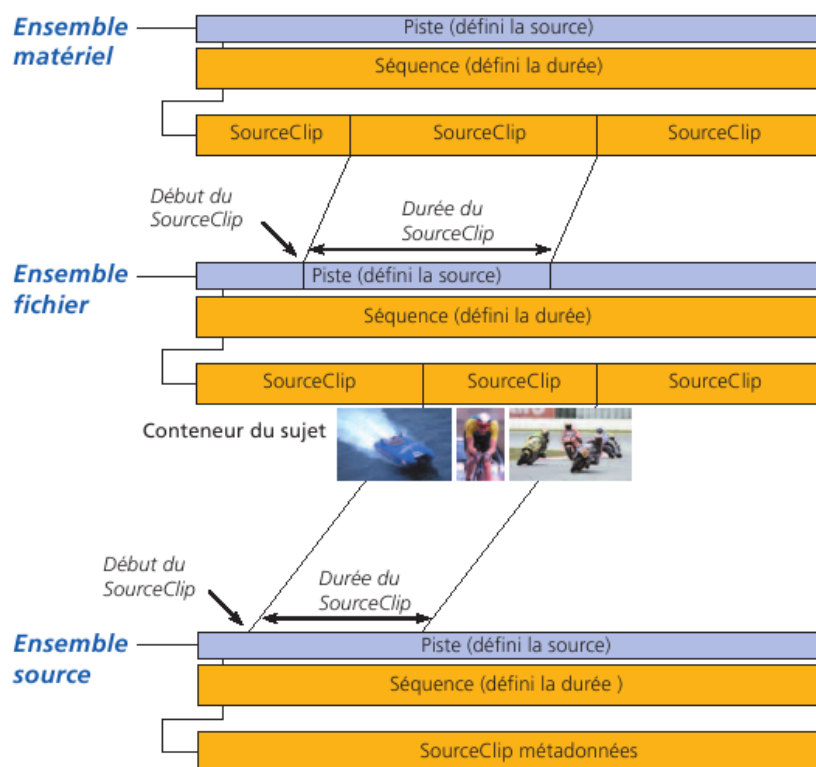
MXF (Material eXchange Format) est un conteneur défini par la SMPTE. Il est de plus en plus employé dans les systèmes haut de gamme. Il permet de contourner les problèmes de communication entre les différentes applications et plateformes, mais les différentes variantes ne sont pas toujours correctement prise en charge. Par exemple les fichiers XDCAM (MXF OP1A) et P2 (MXF OP Atom) ne sont pas compatibles entre eux.

Le conteneur peut contenir les médias et des métadonnées pour enrichir le contenu.

Les métadonnées sont donc des données sur les données, c'est à dire des informations sur les essences contenues dans le fichier MXF et qui vont permettre leur exploitation. Elles peuvent être générées à la source, dès la prise de vue par l'intermédiaire des équipements de prise de vues et/ou ajoutées ensuite tout au long de la chaîne de traitement, à la production, pour la diffusion ou pour l'archivage.

Elles sont de deux catégories :

- Les métadonnées structurelles : c'est à dire concernant la structure même des essences. Elles définissent la nature même de l'essence ainsi que des paramètres «techniques» (format, taux d'échantillonnage, codecs utilisés...) et sont souvent créées automatiquement avec le fichier. Elles sont nécessaires aux équipements pour l'exploitation des contenus.
- Les métadonnées descriptives : ce sont des données informatives et non pas techniques qui répondent à des préoccupations éditoriales ou descriptives. Elles fournissent des informations ou des annotations sur le contenu audiovisuel : titre du document, lieu de la prise de vue, nom des participants, droits, indications du script, évènements...



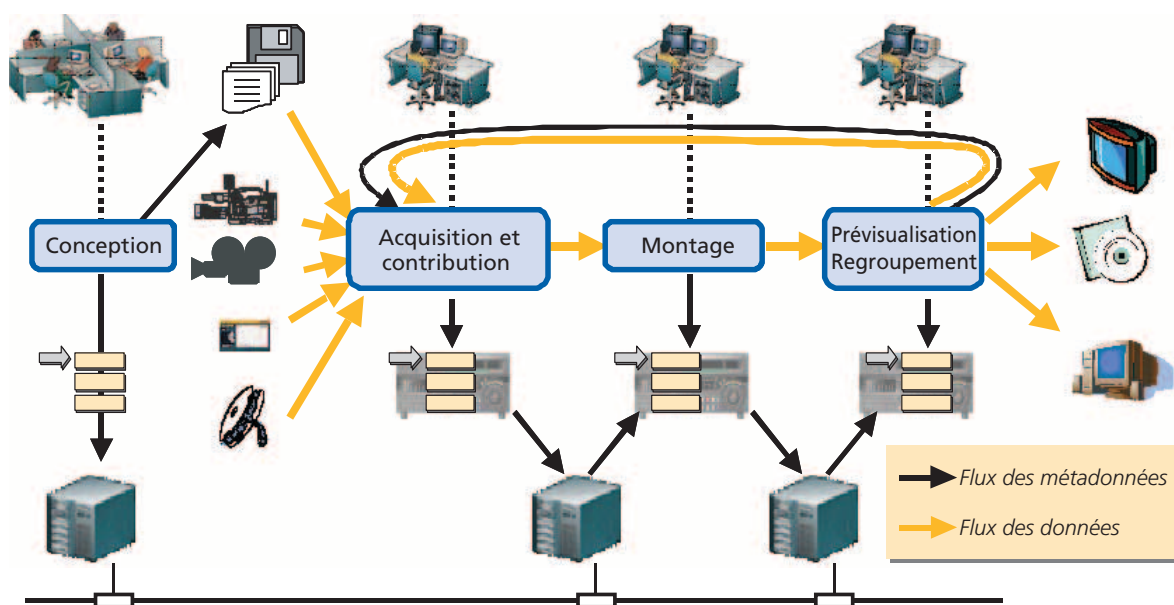
Le conteneur du sujet se compose d'ensembles et de pistes

5) AAF

Le format AAF (Advanced Authoring Format) est un format d'échange utilisé dans le monde professionnel dans les structures de postproduction. Il permet de contourner les problèmes d'interopérabilité des différentes plateformes informatiques.

Il contient des essences (médias audios, vidéos, images...) ainsi que des données permettant d'agencer les contenus (métadonnées) en cours d'élaboration. Il reprend les mêmes principe que MXF mais décrit de manière bien plus précise (et parfois complexe) la progression du montage (effets, transitions...) et la manière de le finaliser.

Typiquement on pourra exporter une timeline en AAF pour l'ouvrir dans un autre logiciel de montage.



VII. Comment choisir un codec ?

En fonction de l'étape de la chaîne de production où l'on se trouve les besoins ne sont pas les mêmes :

Lors du tournage, on favorise des formats dont la compression est simple et rapide car on ne dispose pas (la plupart du temps) d'ordinateurs très performants. On préfère des codecs à GOP long pour optimiser la redondance temporelle et gagner de l'espace. La nécessité de traitement en temps réel implique souvent une compression à débit constant. (MPEG4 et tous les formats constructeurs)

Pendant la post-production, on peut utiliser des espaces de stockage plus importants, le poids des fichiers n'est plus à proprement parler à un problème (attention au débit !). On va devoir décompresser les vidéos pour les afficher et les monter. Les Gops courts et formats intras demanderont beaucoup moins de puissance de calcul et de temps processeur. De plus certains codecs sont pris en charge directement par des cartes propriétaires ce qui rend leur utilisation d'autant plus efficace. (DV, DVCPRO, DNX, PRORES...)

Principes de l'audiovisuel - Vincent MAZO - EFREI 2012/2013

Les contraintes liées à la diffusion dépendent de la cible. On n'exporte pas le même fichier pour une projection en cinéma numérique ou un visionnage sur internet. Il faudra à chaque fois adapter le codec choisi. (JPEG2000 pour le cinéma numérique, MPEG4 pour le reste en général)

Il faudra bien souvent transformer le fichier d'un format à un autre pour faciliter son traitement. Cette opération de transcodage peut être particulièrement longue et compliquée si elle n'a pas été correctement anticipée. Il n'y a donc pas un seul bon choix possible. Il dépend souvent du matériel à disposition. En cas de doute, rien de mieux qu'un test en grandeur nature en utilisant des rushes (tournés lors d'un repérage par exemple) et en leur faisant suivre toutes les étapes de la chaîne.

