



Institut
Mines-Télécom

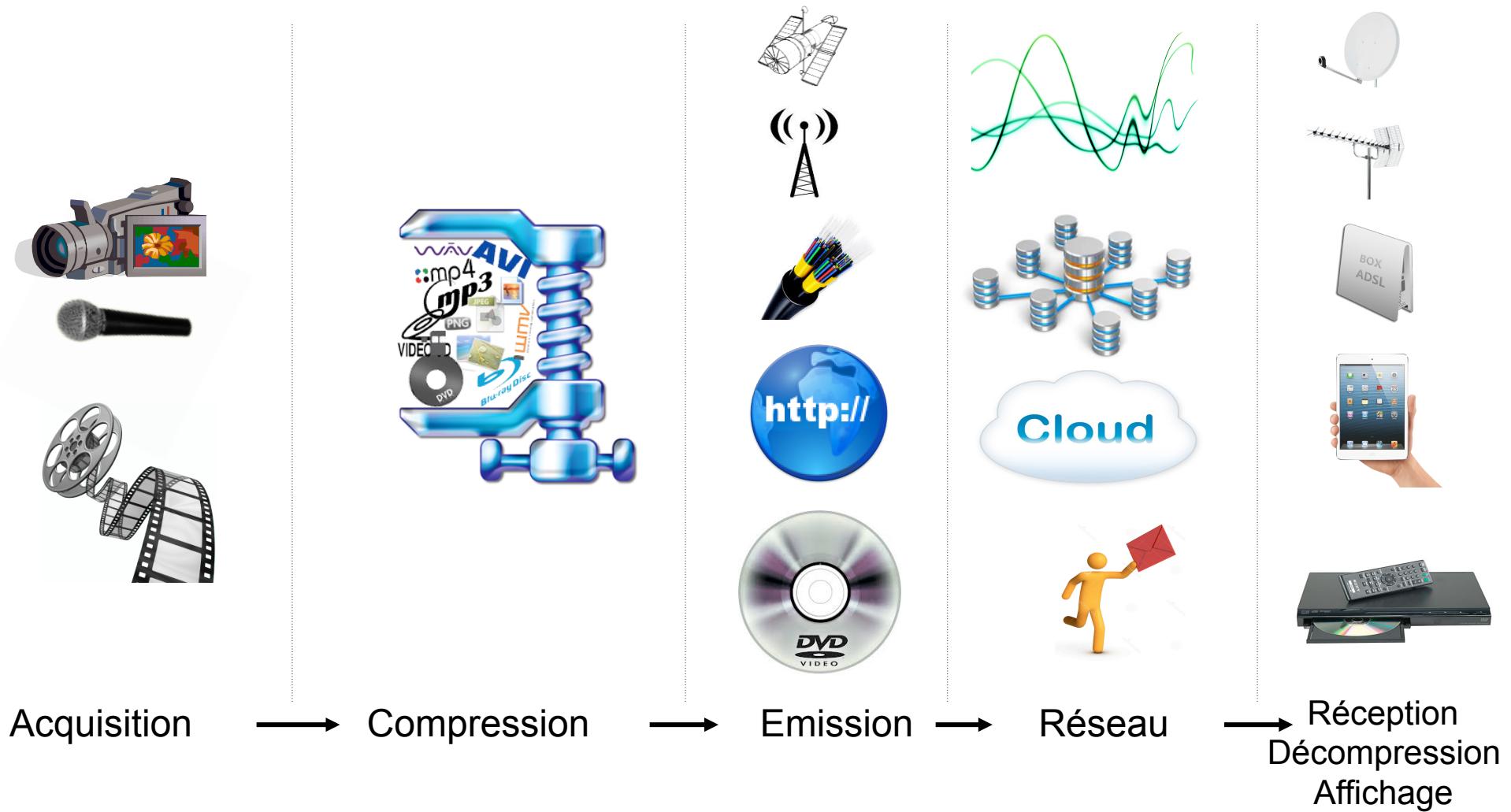
Introduction au transport de contenus multimédia

Jean Le Feuvre

SIGMA207



La chaîne de transmission média





Chaine de distribution

■ Transport Multimédia

- Acheminer le contenu de la zone de production à la zone de consommation

■ Contenu Multimédia

- Audio, Vidéo
- Sous-titres
- Graphisme Vectoriel 2D or 3D (Flash, SVG, VRML, BIFS)
- Applications (ex: applet Java, page HTML)
- Données associées
 - Capteurs: GPS, température, magnétomètre, altimètre, baromètre, ...
 - Annotations: régions d'intérêt, texte, positions 2D / 3D dans un flux vidéo
 - Instruction de programme: Java, JSON, ...



Types de Services

■ Direct ou « live »

- Radio, TV satellite ou hertzienne
- Direct internet

■ Caractéristiques

- Tous les utilisateurs reçoivent les même données
- Temps réel mais léger délai acceptable
- Données spécifiques à la session / appel
- Temps réel et très faible tolérance au délai

■ Vidéo à la demande

- Contenus utilisateurs (ex: YouTube)
- Offres VoD (Netflix, ...)
- Catch-up & Replay

- Gros volume de données
- Temps différé, tolérance au délai / mise en pause

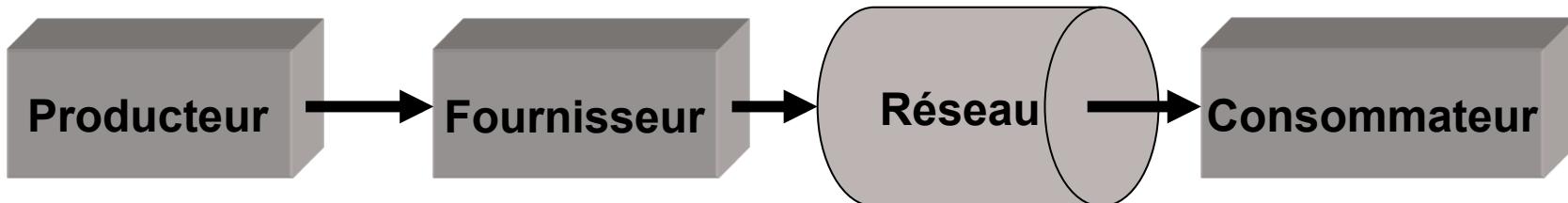
■ Stockage / Archivage

- DVD, Blu-Ray

- Aucune contrainte de temps d'acheminements
- Sans perte



Les acteurs d'une chaîne de distribution



■ Producteur

- Rôle : Création, encodage, préparation à la distribution de contenu MM
- Ex : Professionnels (Chaines TV, radio, ...), créations personnelles (User Generated Content UGC)

■ Fournisseur

- Rôle : Exploiter un réseau pour distribuer le contenu
 - Peut également faire l'encodage et la préparation à la diffusion
 - Protection de contenu
- Ex : Professionnels (CanalSat, Orange, Free, Netflix...), UGC (YouTube, P2P)

■ Réseaux

- Rôle : Transporter le contenu avec une qualité de service donnée
 - Gestion des débit, délais, taux d'erreur ...
- Protocoles et moyens physiques de transport
 - Ex : Satellite, Câble, Hertzien, Cellulaire, Internet...

■ Consommateur

- Rôle : Recevoir et visualiser le contenu
 - Peut nécessiter le déchiffrement du contenu
- Rôle : Interagir avec le contenu (votes, télé-achat)
 - Nécessite une voie de retour

Définitions

■ Flux Media Elémentaire (“Elementary Stream”)

- Succession temporelle des données média
- Généralement, 1 média (audio, vidéo) \Leftrightarrow 1 flux
 - Exception en codage scalable (vidéo, 3D audio): 1 média = N flux

■ Unité d'accès ou Access Unit (AU)

- \Leftrightarrow Chaque état / donnée de mise à jour du média
- « plus petit bloc de donné auquel on peut associer un temps unique»

■ Exemples d'AU

- Vidéo: 1 trame
- Audio non compressé: 1 échantillon
- Sous-titres:
 - Bitmap: une image (ex: DVD)
 - Texte: N ($>=1$) lignes

■ Fréquence des AU (ou Frame Rate)

- Constant Frame Rate (CFR): même durée pour toutes les Aus
 - Ex: TV, cinéma ...
- Variable Frame Rate (VFR): durée différentes
 - Ex: vidéosurveillance
- Principalement utilisé pour vidéo et animations

Acquisition et Formats: quelles sorties ?



Acquisition



■ Caractéristiques Vidéo

- Résolution (hauteur x largeur)
 - Ex: 640x480, 720x576, 1280x720, 1920x1080
- Format
 - Ex: RGB, YUV420, YUV422, YUV444
- Plage d'échantillonnage
 - 8, 10, 12 bit par canal couleur
 - Gamut de couleur
- Fréquence d'échantillonnage (Frame rate)
 - Ex: 23.976, 24, 25, 29.97, 30, 50, 60, 100, 120
 - Entrelacé (i) ou progressive (p)
 - Ex: 720p, 1080i

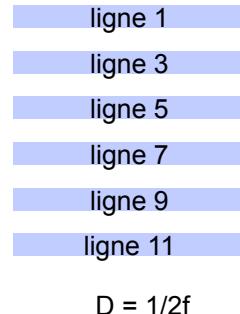
■ Quelques repères

- Trame vidéo 1080p YUV420: 3.1 Mbytes
- Vidéo 720p50 YUV420: 553 Mb/s
- DVD de 2 heures: 111 Gbytes non compressé

Acquisition : Entrelacement Vidéo

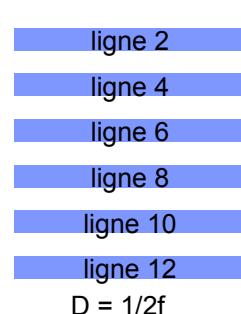


Acquisition



$$D = 1/2f$$

Champ haut
(top field)



$$D = 1/2f$$

Champ bas
(bottom field)



$$D = 1/f$$

Trame entrelacée



VS



$$D = 1/f$$

Trame progressive





Acquisition et Formats



- Exercice 1
- Exercice 2
- Exercice 3
- Exercice 4



URL du TP:

<http://perso.telecom-paristech.fr/~lefeuvre/SIGMA207/INTRO/>

Acquisition

Acquisition et Formats: quelles sorties ?



Acquisition

■ Caractéristiques audio

- Fréquence d'échantillonnage / Sample Rate (Hz)
Ex: 44100, 48k, 24k, 22050, 96k
- Format d'échantillon
 - Ex: 8 bit signé/non signé, 16 bit, 24 bit, 32 bit float
- Canaux et positions
 - Ex: stéréo, 5.1, 7.2, 22.2

■ Quelques repères

- CD audio stéréo 16 bpc 44.1 KHz: 1.4 Mbps
- Voix sur 3G:
 - Normal: 16 bps 8 KHz mono: 128 Kbps
 - HD 16 KHz: 256 Kbps



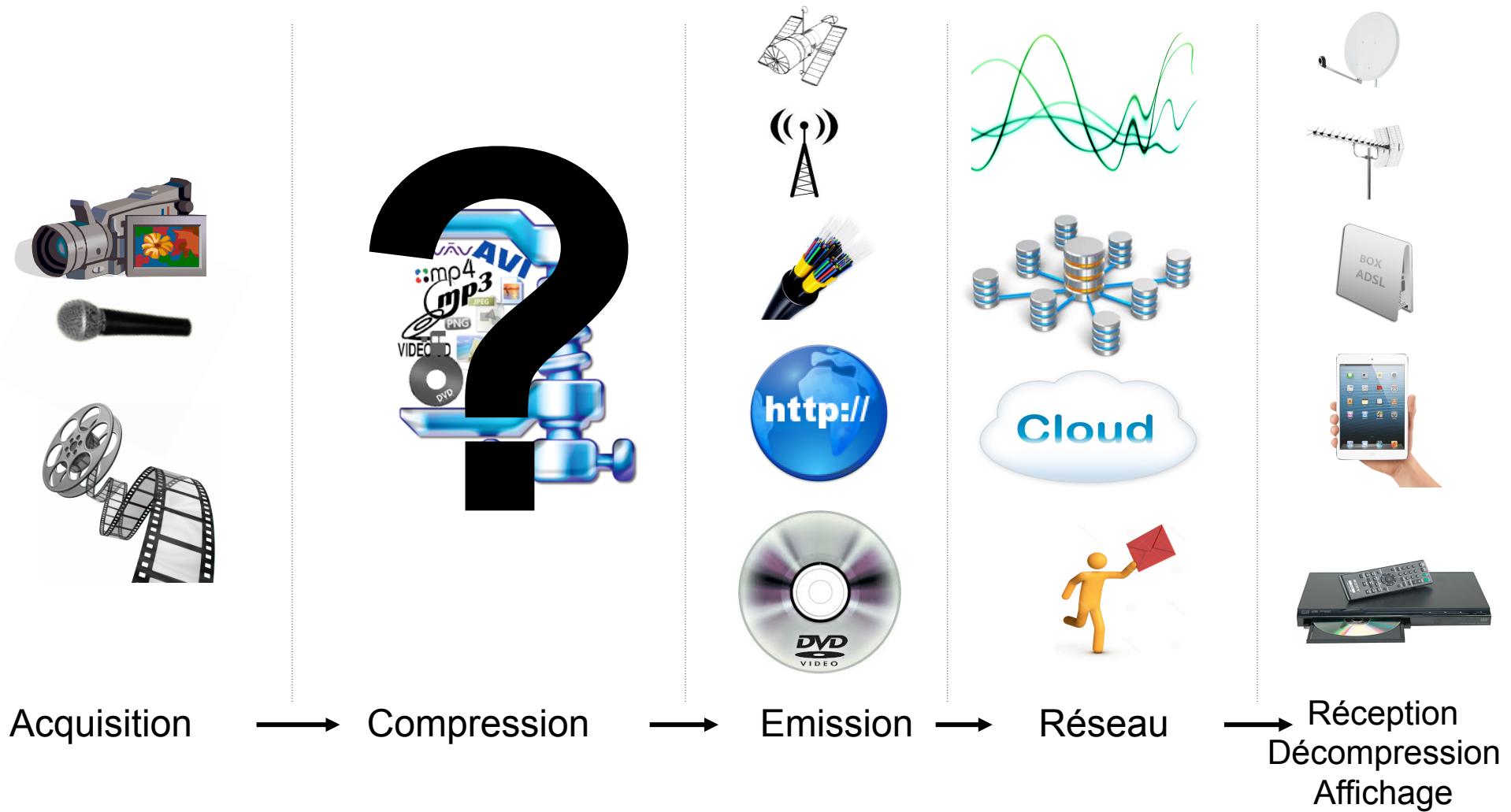
Acquisition et Formats



Acquisition

- Exercice 5
- Exercice 6
- Exercice 7

La chaîne de transmission média



Caractéristiques des réseaux



Emission

■ Diffusion Satellite et Hertzienne

- Pertes sans retransmissions
 - FEC « Forward Error Correction »
- Débit fixe
- Latence fixe

■ Réseau IP

- Managé:
 - peu de pertes, peu de gigue réseau
 - Débit garanti
- Général:
 - pertes, gigue importante
 - Débit variable

■ Support physique

- Pas de pertes, Débit fixe,



Caractéristiques des réseaux

■ ISP speed in France by netflix

- <https://ispspeedindex.netflix.com/country/france/>

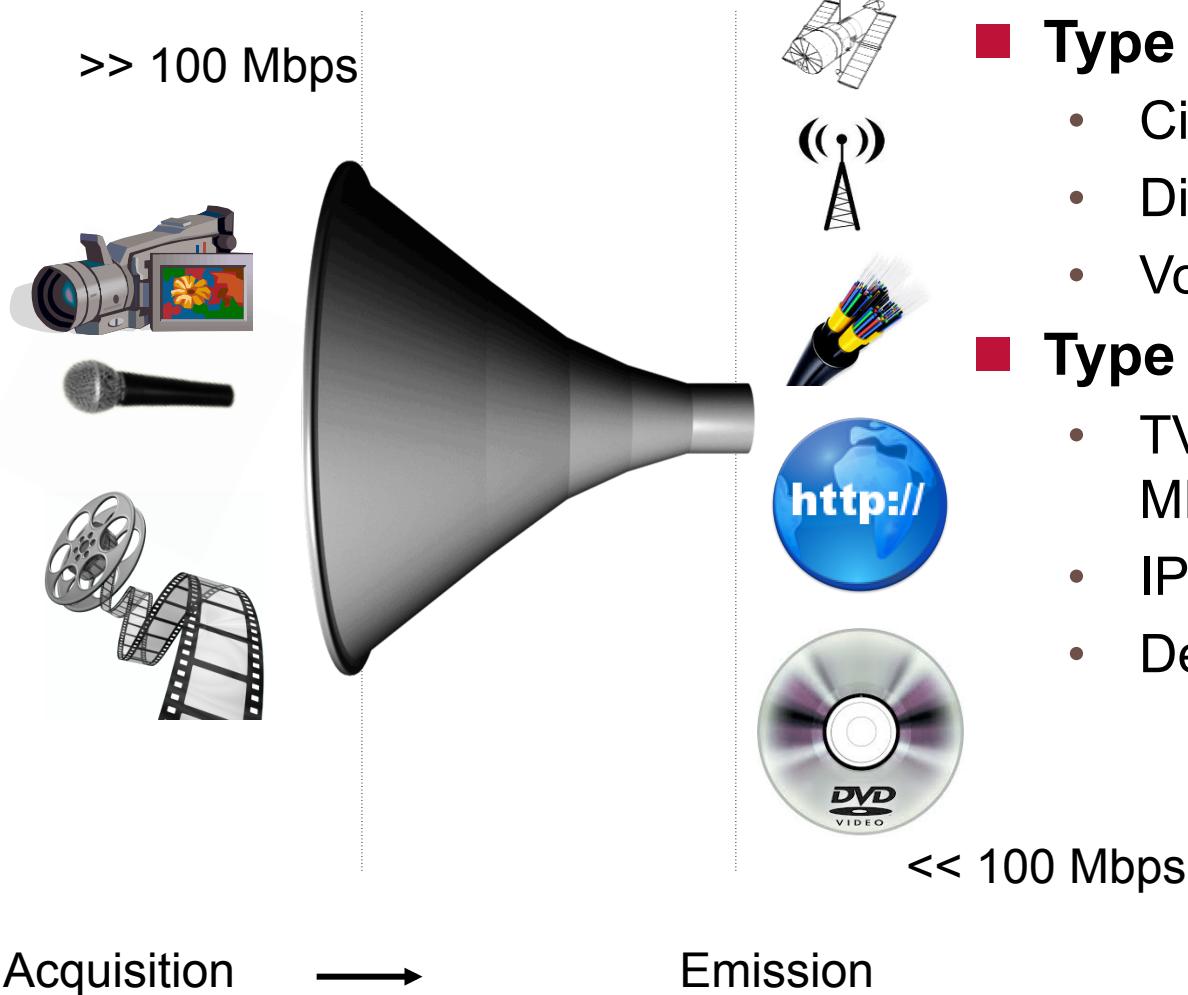
■ SpeedTest

- <http://www.speedtest.net/fr/>

■ DTV

- <http://www.digitalbitrate.com/>

Quelle compression choisir ?



■ Type d'application

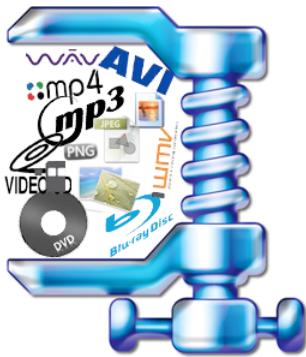
- Cinéma (~= sans perte)
- Diffusion TV
- VoD (débit moindre)

■ Type de réseaux

- TV Numérique: compression MPEG-2 ou MPEG-4
- IP/HTTP: choix plus large
- Dépend du parc en place



Définitions



Compression

■ Codec

- Entité en charge de **coder** et **décoder** le signal entre sa représentation compressée et sa représentation non compressée

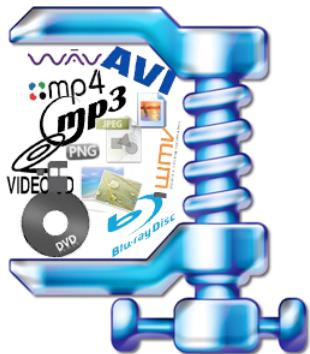
■ Gestion de débit

- Constant Bit Rate (CBR): débit de sortie de l'encodeur constant
- Variable Bit Rate (VBR): débit de sortie de l'encodeur peut varier
- Capped VBR: VBR avec débit maximal imposé



Compression Vidéo

- Exercice 8
- Exercice 9
- Exercice 10
- Exercice 11



Compression



Différents types de fichiers multimédia



Stockage

- **Fichiers pour le stockage de données multimédia**
 - Fichiers de stockage mono-flux
 - Stockage d'un seul flux élémentaire
 - Fichiers de stockage multi-flux
- **Approche historique : Stockage sous forme de flux de transport**
 - Élémentaires ou non
 - Exemples: mp3, VOB, mpeg-2 transport stream ...
 - Avantages
 - Adapté aux problématiques de transport sur un réseau
- **Problèmes**
 - Inutilité des mécanismes de correction d'erreurs
 - Inefficacité/Inutilité de la fragmentation
 - Inefficacité de la recherche ou de la lecture aléatoire
 - Difficulté d'édition

Stockage simple d'un seul flux élémentaire



Stockage

- Stockage consécutif des données codées
- Séparation difficile des Access Units
 - Nécessité d'interpréter le bitstream
 - Dépend du format du média
 - Nécessité de rechercher des 'start codes'
 - Difficulté d'édition, d'insertion de données ...
- Exemples:
 - mp3: entêtes dans chaque AU
 - H263, H264: présence de start codes systématique
- Ajout de mécanismes de séparation externe au codage
 - AAC: stockage ADTS ou stockage LATM

Stockage de plusieurs flux élémentaires



Stockage

■ Facilité d'édition

- Ajout/Suppression d'un flux, d'Access Units
- Edition temporelle (chapitrage, ...)
- Flux potentiellement désynchronisés

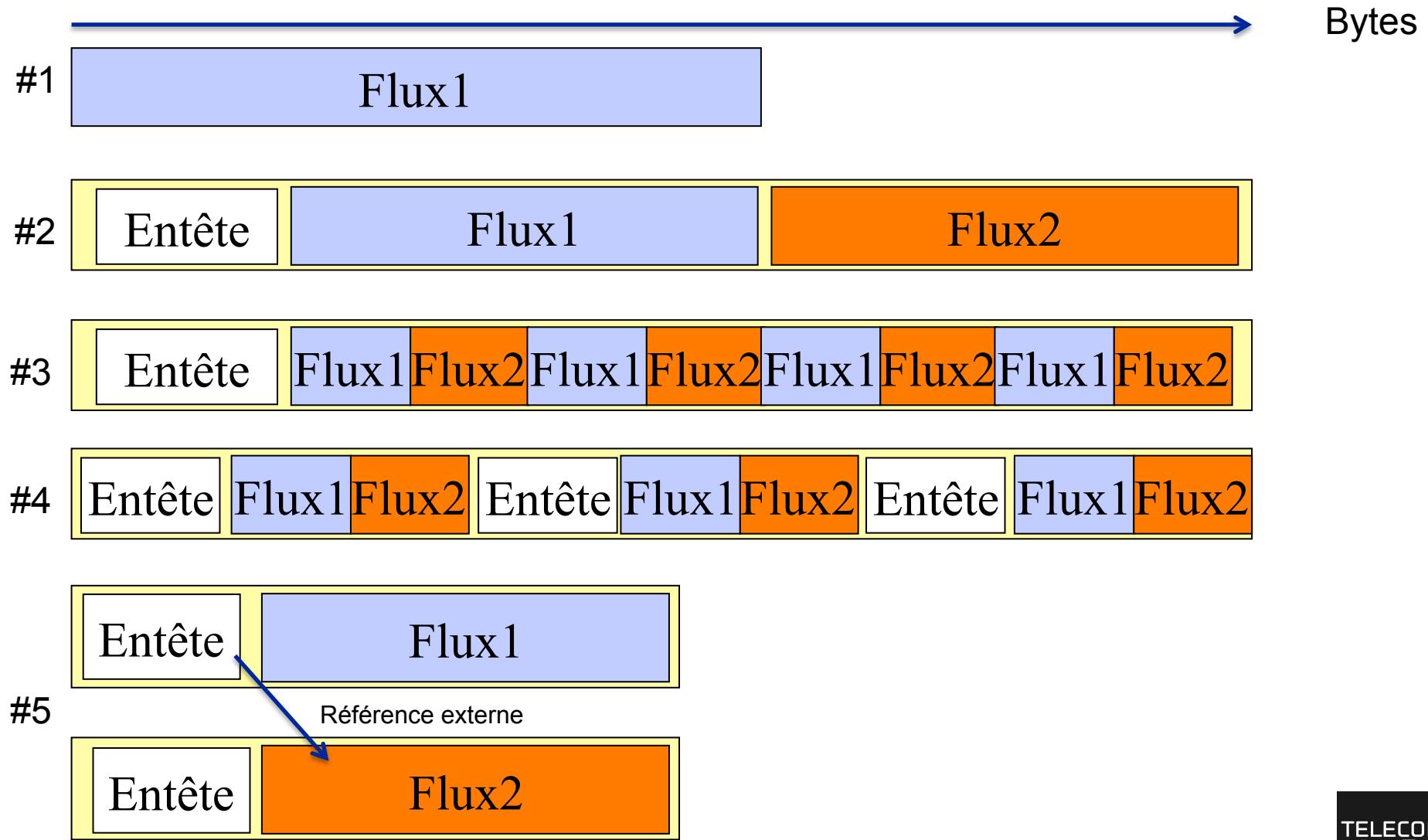
■ Stockage séparé entre données (Access Unit) et métadonnées/information de signalisation

- Possibilité de parcourir le fichier/de comprendre la structure du fichier sans décoder ou interpréter les données codées
- Détermination facile des RAP (Points d'accès aléatoires)

■ Stockage avancé des données

- Possibilité de répartir les données sur plusieurs fichiers
- Préparation au streaming et au téléchargement progressif

Exemples de répartition des données





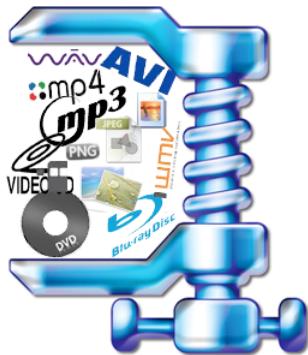
Exemples de fichiers de stockage A/V

- **Fichiers DVD: MPEG-2 Program Stream (VOB) + Information de navigation (IFO)**
- **Fichiers Flash Video: FLV, F4V**
- **Fichiers AVI**
- **Fichiers Microsoft: ASF**
- **Fichiers musiques: AAC, MP3, M4A, AC3**
- **Fichiers vidéo DV**
- **Fichiers ISO: MP4, 3GP, AVC, PSP, M4A ...**
- **...**



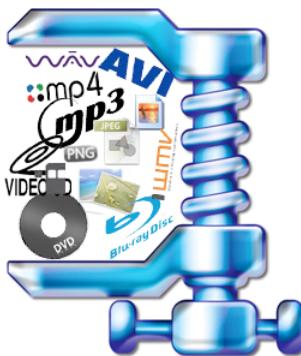
Compression Vidéo

■ Exercice 12



Compression

Random Access lors du décodage



Compression

■ Décoder un flux compressé à partir de n'importe quel instant

- Dépendances de codages entre AU(T) et AU(t < T)
- Identification des points d'accès aléatoires (Random Access Points)
 - Sans dépendances de codage avec autres AUs
 - Après lesquels (en ordre de présentation) les AUs peuvent être décodées
- Pour les autres temps != RAP
 - VOD: Revenir au RAP précédent
 - Live: attendre le prochain RAP

■ Audio, Texte, Images, Meta-data

- AU \Leftrightarrow RAP (la plupart du temps)

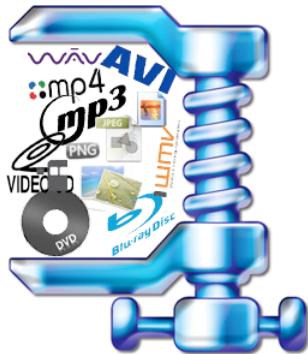
■ Vidéo, graphismes compressés

- Toutes les AUs ne sont pas RAP
- La rapidité d'accès au média dépend de la fréquence des RAP



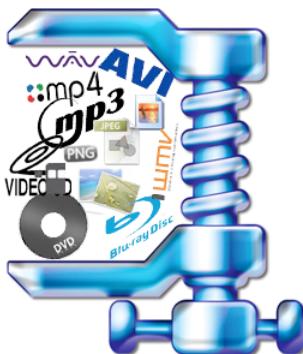
Compression Vidéo

■ Exercice 13



Compression

Particularités du codage vidéo



Compression

■ Principes de base

- Supprimer les redondances en trouvant les pixels similaires au sein d'un trame ou entre trames
- Certaines trames ne dépendent pas des autres pour le décodage (intra frames)
- Meilleur efficacité si les références viennent du passé et du futur de la trame à compresser
 - Les trames ne sont pas toujours en ordre de présentation
 - Les « trames futures » demande un mise en mémoire de plusieurs trames donc augmente la latence

■ Group of picture (or GOP)

- Séquence d'images, en ordre de décodage, démarrant par une trame I et se terminant avant la prochaine trame I

■ Trames B ou trames Bidirectionnelle

- Images utilisant des images du passé et du futur pour leur reconstruction
- Dans AVC et HEVC, les trames B peuvent servir de références
- **Si trames B, ordre décodage != ordre présentation**

Encodeur HEVC

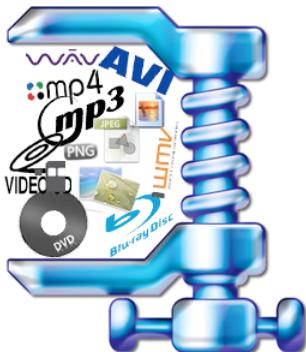
```

POC      0 LId: 0 TId: 0 ( I-SLICE, nQP 37 QP 37 )      1856 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 ]
[L1 ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      8 LId: 0 TId: 0 ( B-SLICE, nQP 38 QP 38 )      712 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 0 ]
[L1 0c ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      4 LId: 0 TId: 0 ( B-SLICE, nQP 39 QP 39 )      712 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 0 ]
8 ] [L1 8c 0 ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      2 LId: 0 TId: 0 ( B-SLICE, nQP 40 QP 40 )      712 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 0 ]
4 ] [L1 4c 8 ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      1 LId: 0 TId: 0 ( B-SLICE, nQP 41 QP 41 )      712 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 0 ]
2 ] [L1 2c 4 ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      3 LId: 0 TId: 0 ( B-SLICE, nQP 41 QP 41 )      712 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 2c ]
0 ] [L1 4 8 ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      6 LId: 0 TId: 0 ( B-SLICE, nQP 40 QP 40 )      712 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 4c ]
2 ] [L1 8 4 ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      5 LId: 0 TId: 0 ( B-SLICE, nQP 41 QP 41 )      712 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 4
0 ] [L1 6c 8 ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      7 LId: 0 TId: 0 ( B-SLICE, nQP 41 QP 41 )      712 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 6c ]
4 ] [L1 8 6 ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      16 LId: 0 TId: 0 ( B-SLICE, nQP 38 QP 38 )     5312 bits [Y 42.5379 dB   U 62.8706 dB   V 53.0399 dB] [ET      2 ] [L0 8 6
4 0 ] [L1 8c 6 4 0 ] [MD5:902ff2cabd4bb4f1e518d46e86c11531,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      12 LId: 0 TId: 0 ( B-SLICE, nQP 39 QP 39 )     720 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 8
6 ] [L1 16c 8 ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      10 LId: 0 TId: 0 ( B-SLICE, nQP 40 QP 40 )     712 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 8
6 ] [L1 12c 16 ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      9 LId: 0 TId: 0 ( B-SLICE, nQP 41 QP 41 )      712 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 8
10 ] [L1 10c 12 ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      11 LId: 0 TId: 0 ( B-SLICE, nQP 41 QP 41 )     712 bits [Y 70.9638 dB   U 67.9535 dB   V 53.0399 dB] [ET      1 ] [L0 10c
8 ] [L1 12 16 ] [MD5:5a9c37c8d625786c87914263dd0a76de,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      14 LId: 0 TId: 0 ( B-SLICE, nQP 40 QP 40 )     728 bits [Y 42.3554 dB   U 62.8050 dB   V 53.0399 dB] [ET      1 ] [L0 12c
10 ] [L1 16 12 ] [MD5:386ad9db9960cb6417cd703b870a4012,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      13 LId: 0 TId: 0 ( B-SLICE, nQP 41 QP 41 )     728 bits [Y 42.1979 dB   U 62.8050 dB   V 53.0399 dB] [ET      1 ] [L0 12
8 ] [L1 14c 16 ] [MD5:df4e7700a5f0feac07618abf3c24d9d8,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      15 LId: 0 TId: 0 ( B-SLICE, nQP 41 QP 41 )     712 bits [Y 42.4325 dB   U 62.8050 dB   V 53.0399 dB] [ET      1 ] [L0 14c
12 ] [L1 16 14 ] [MD5:06843f30f94b8b538d1d158e387329d6,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56a3f79d26a88a]
POC      24 LId: 0 TId: 0 ( B-SLICE, nQP 38 QP 38 )     912 bits [Y 42.5571 dB   U 63.1694 dB   V 53.0399 dB] [ET      2 ] [L0 16
14 12 8 ] [L1 16c 14 12 8 ] [MD5:d9d8ed907bb55145deaa2b5f6e0c9351,be6517fdf1830325b589db97398a4e59,e741dc9e7ef2adc37d56

```

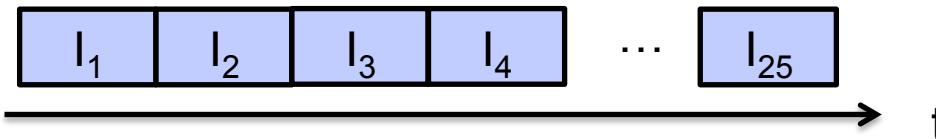


Exemples de GOPs

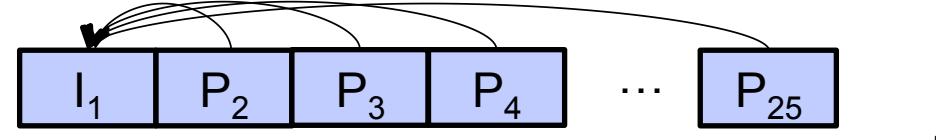


Compression

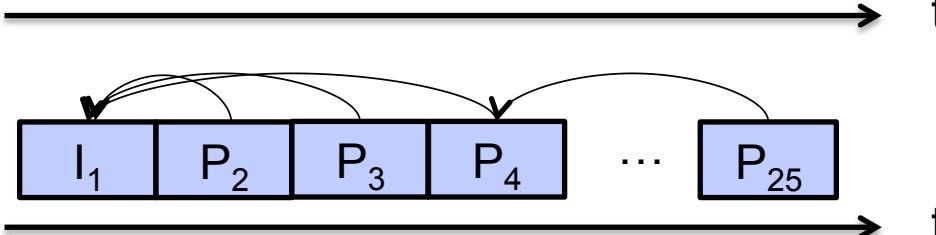
All Intra
(I only)



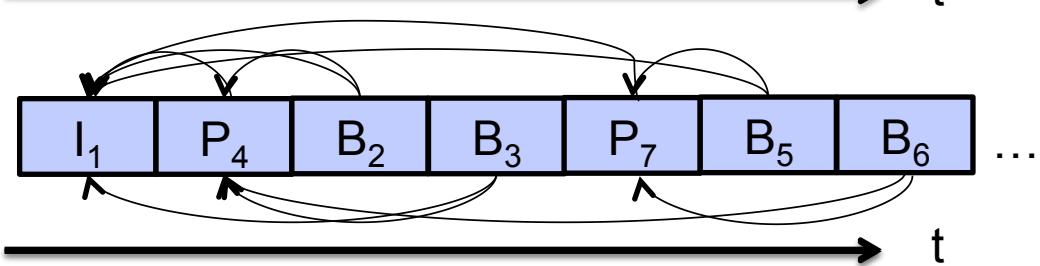
Low-delay
(I & P)



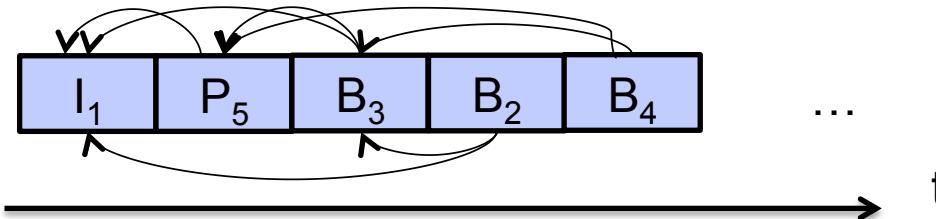
Low-delay
(I & P)



IPB



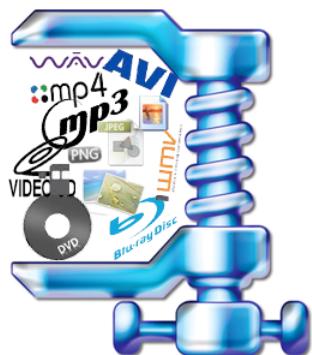
IPB
Hierarchical
B-frames





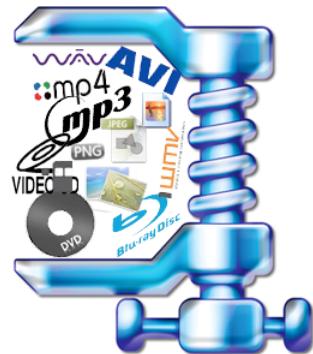
Compression Vidéo

- Exercice 14
- Exercice 15
- Exercice 16



Compression

Particularités du codage audio



Compression

■ Principes de bases

- Supprimer les signaux non perceptibles
- Supprimer les redondances
- Pour cela il faut travailler avec beaucoup plus qu'un échantillon!

■ Fenêtre d'analyse - Audio Coding Window

- Encodeur audio encode N échantillons d'un coup
- Ceci redéfinit l'AU du *flux compressé*: 1 AU Audio \Leftrightarrow N échantillons audio
- Exemples
 - MP3: 576 or 1152 échantillons par AU
 - AAC: 1024 échantillons per AU

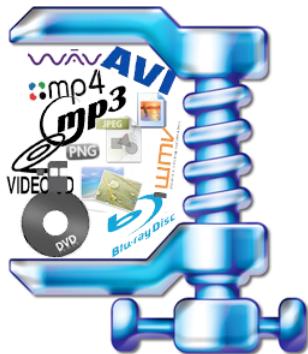
■ Random Access

- (presque tous les codecs) Chaque AU est RAP
- Il peut exister des dépendances entre deux ou trois AU consécutives mais chacune peut être décodée indépendamment, même si le signal reconstruit ne sera pas parfait



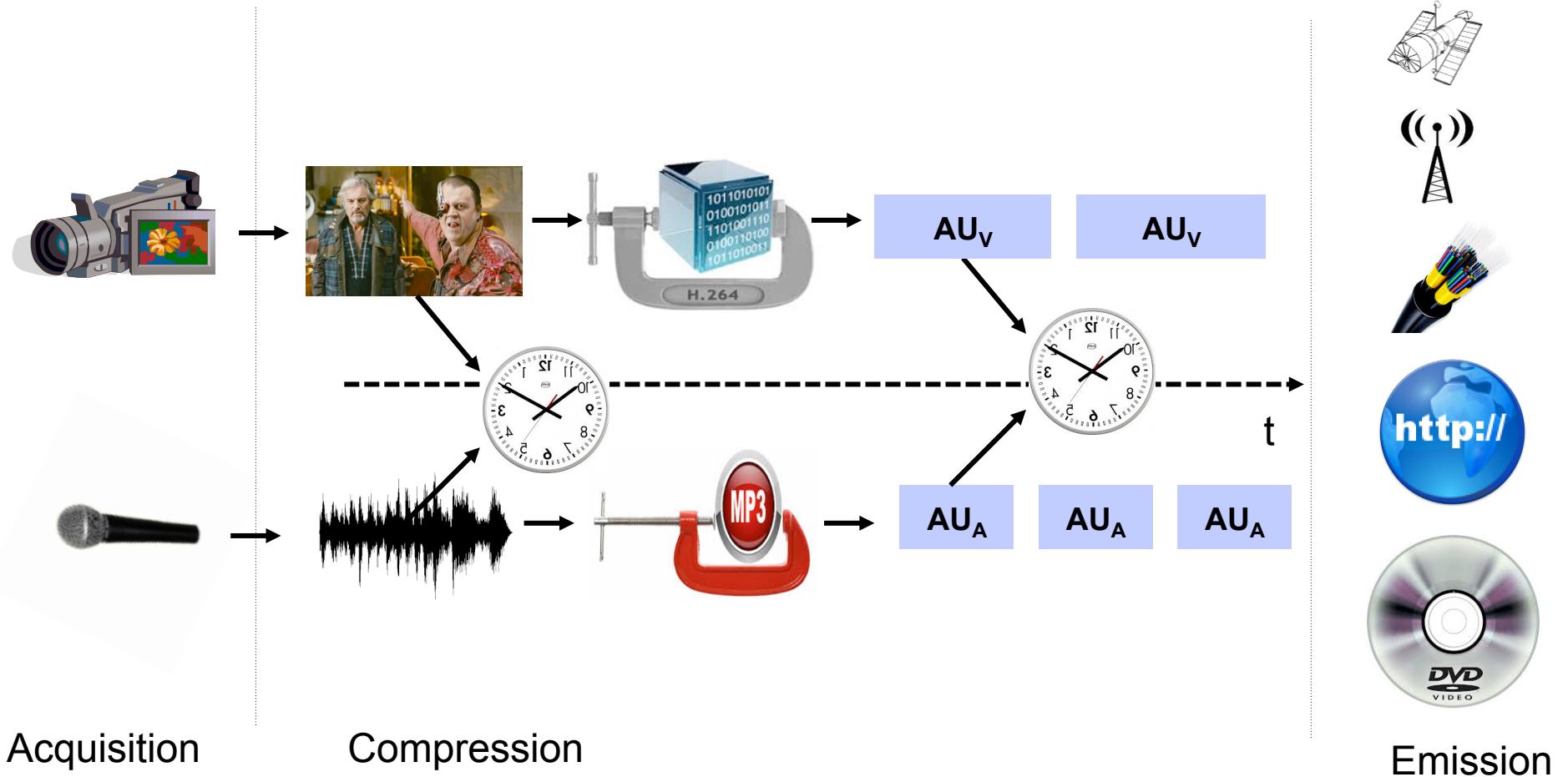
Compression Audio

- Exercice 17
 - Exercice 18

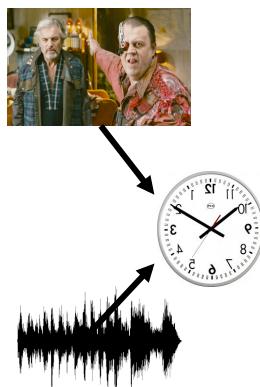


Compression

Synchronisation des média

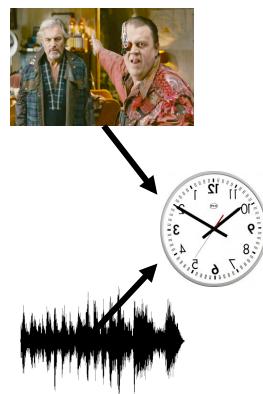


Synchronisation intra-flux



- Assure le rendu correct d'un média
 - Audio, vidéo, texte, ...
- Présenter les trames échantillonnées à temps
 - Fréquence échantillonnage fixe
 - Ex: vidéo 25 FPS -> une trame toutes les 40 ms
- Prévoir les pertes de trames
 - Reconstruire le temps de présentation si perte
 - Numéroter les trames ou associer un temps
- Prévoir les dérives du système d'acquisition
 - Fréquences d'échantillonnage non fixes
 - Ex: vidéo, graphismes vectoriels, sous-titres
 - Numérotation des trames non suffisante
 - => associer **un temps ou une durée** par trame

Synchronisation inter-flux



■ Assure le rendu correct d'un ensemble de média

- Audio + vidéo
- Audio + texte, Vidéo + texte
- AV + application, ...

■ Fréquence d'échantillonnage différentes

- vidéo 25 FPS -> une trame toutes les 40 ms
- AAC 44.1kHz -> une trame toutes les ~23.22 ms

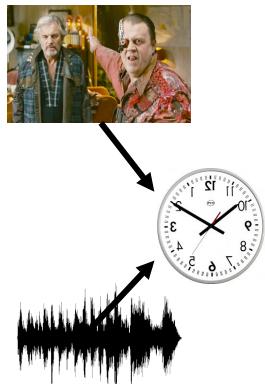
■ Utilisation d'une base de temps commune

- Synchronisation obtenue en comparant les temps associés à chaque trame de chaque média
 - Origine des temps commune
- Trouver l'échelle de temps introduisant le moins d'erreurs possible (PGMC, ...)



Synchro Media

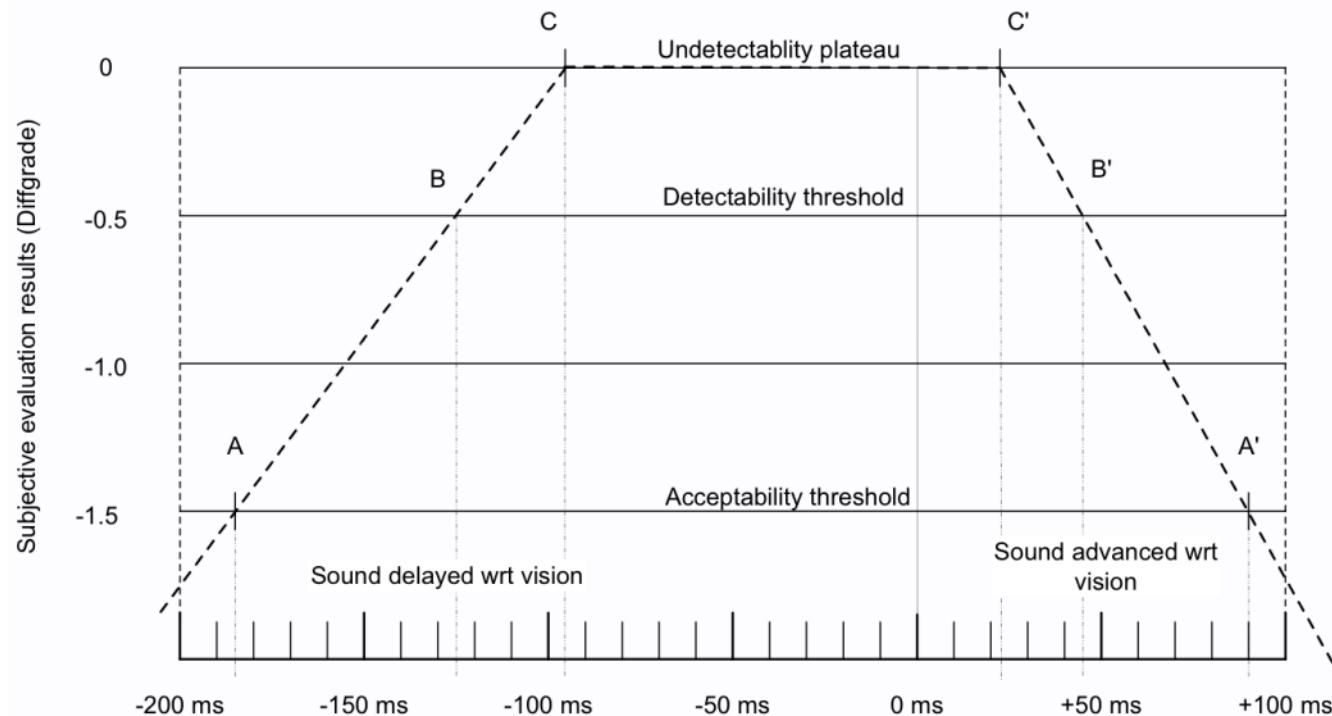
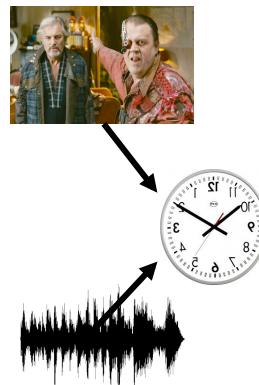
■ Exercice 19



Tolérance à la désynchronisation

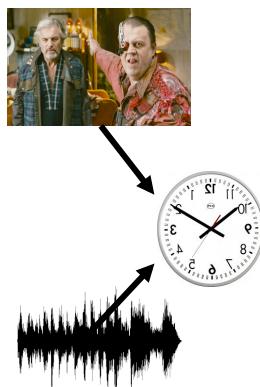
■ Fenêtre de synchro acceptable

- (audio – vidéo) timing
- +90ms – 185ms



ITU-R BT.1359 Figure 2

Tolérance à la désynchronisation



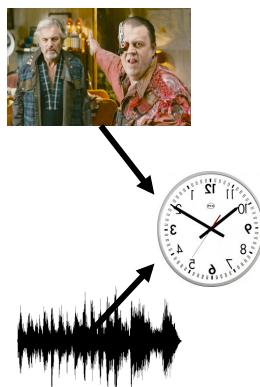
■ Dans l'industrie TV

- EBU Recommandation R37
 - Avance audio max 40 ms
 - Retard audio max 60 ms
- ATSC IS-191
 - Avance audio max 45 ms
 - Retard audio max 75 ms

■ Pour la visioconférence

- même chiffres
- Sensibilité varie en fonction de la résolution !!

Tolérance Synchro: Cas avancés



■ Nouveaux types de services:

- Codage scalable
 - Par couche (eg SVC), passage SD -> HD, SNR
 - Par description multiple (MDC)
- Services TV3D
 - Vues codées et transportées séparément

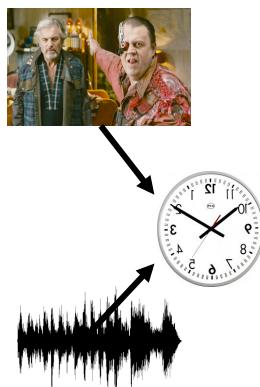
■ Nouveaux besoins de synchro:

- Synchronisation de décodage = tolérance Zéro !
- Si délai sur un flux
 - Décodage cassé (codage scalable)
 - Reconstruction cassée



Synchro Media

■ Exercice 20





Institut
Mines-Télécom



jean.lefeuvre@telecom-paristech.fr

SIGMA207