

ÉCOLE NATIONALE SUPÉRIEURE LOUIS LUMIÈRE

Mémoire de fin d'études

---

**AUDIO SUR IP : PROBLÉMATIQUES  
ET PERSPECTIVES D'ÉVOLUTION**

*Partie Pratique : Projet d'Ingénierie d'un studio polyvalent  
en audio sur IP*

---

Marie-Angélique MENNECIER

*Directeur interne : Jean ROUCHOUSE*

*Directeur externe : Jean-Marc CHOUVEL*

*Rapporteur : Franck JOUANNY*

19 juin 2015

---

## **RÉSUMÉ :**

À l'heure où les réseaux audio sur IP professionnels se développent, une normalisation a été effectuée par le groupe de travail AES X192 afin de permettre leur interopérabilité. L'AES67 pose les bases nécessaires pour un fonctionnement commun, et les principales solutions existant sur le marché de l'audio sur IP professionnel s'y sont conformées depuis.

La problématique de ce mémoire est l'analyse des implications techniques de l'audio sur IP et des mécanismes nécessaires à une interopérabilité concrète, tout en proposant une ouverture sur les possibilités de développements futurs de cette technologie protéiforme.

Dans un premier temps, nous formulerons quelques rappels sur les réseaux IP qui nous serviront de socle technique pour analyser les protocoles existants, puis comprendre les fondations nécessaires à leur fonctionnement commun. Nous pourrons enfin dégager quels écueils peuvent entraver de futurs développements.

**mots clés :** AoIP, audio sur IP, AES67, Dante, Ravenna, Livewire, interopérabilité, AVB

## **ABSTRACT :**

At times when audio over IP networks have started growing and thriving, a standardization has been discussed by the AES X192 workgroup to insure that they would be interoperable. In 2013, AES67 defined such rules, and the main audio over IP professional networks have since complied.

This dissertation analyzes the technical details of such a technology, how interoperability can be effectively achieved, and how audio over IP may evolve, in the future, from its current protean form.

We shall first explain the main IP networking standards necessary for a basic understanding of audio over IP. From there, the main existing protocols will be briefly described in order to acknowledge their common denominators, and the way they can be interoperable. At last, questions about future developments and their remaining issues will be raised.

**Keywords :** AoIP, audio over IP, AES67, Dante, Ravenna, Livewire, interoperability, AVB

---

## REMERCIEMENTS

Tout d'abord, je tiens à remercier mon directeur interne, Jean Rouchouse, pour sa grande disponibilité par divers moyens de communication, sa bienveillance, et pour ses conseils éclairés pour mener à bout ce travail.

Je tiens également à remercier mon directeur externe, Jean-Marc Chauvel, pour sa sympathie, sa patience, et son accueil chaleureux à la Beudelie.

Merci à Franck Jouanny d'avoir accepté d'être rapporteur sur ce mémoire malgré son travail très prenant.

Merci à Gérard Pelé pour m'avoir permis de rencontrer Jean-Marc Chauvel, et ainsi avoir pu travailler sur une partie pratique stimulante et intéressante.

Je tiens également à saluer Léo Rossi-Roth pour ses conseils sur le sujet de l'audio sur IP.

Merci à Taieb Keraoun et Bettina Hulsmeyer pour leur aide bienveillante dans la recherche de contacts.

Merci à ma maman, Isabelle Mennecier, pour avoir supporté mes aléas psychologiques. Merci également pour son aide à la relecture (ou parfois, traduction...) minutieuse dans la belle langue de Molière.

Enfin je remercie toute ma promo pour leur moral indéfectible en ces temps éprouvants mais cathartiques du mémoire.

---

<b>PARTIE THEORIQUE : État de l'art</b> .....	<b>1</b>
<b>1. Introduction, terminologie et définitions</b> .....	<b>2</b>
1.1. Terminologie et définitions .....	3
1.1.1. Abréviations courantes.....	3
1.1.2. Définitions .....	3
1.1.3. Organismes de formation, normes principales.....	7
1.2. Principaux réseaux audionumériques et évolution.....	9
1.2.1. Bref historique et contexte.....	9
1.2.2. Rappel sur les principaux protocoles et technologies .....	10
<b>2. Réseaux Audio sur IP (AoIP)</b> .....	<b>13</b>
2.1. Technologies des réseaux informatiques, focalisation sur IP .....	13
2.1.1. Repères historiques .....	13
2.1.2. Le protocole IP : définitions.....	14
2.1.3. Rappels sur le réseau.....	23
2.1.4. Problèmes soulevés et comparaison des protocoles existants .....	26
2.2. Protocoles AoIP actuels.....	32
2.2.1. Livewire.....	33
2.2.2. Dante.....	33
2.2.3. Ravenna .....	33
2.2.4. AES67.....	34
2.2.5. AVB .....	34
2.3. Comparaison des protocoles : interopérabilité.....	36
2.3.1. configuration du réseau .....	36
2.3.2. synchronisation.....	39
2.3.3. Gestion sous Mac et Windows .....	40
2.3.4. Latence .....	42
2.3.5. Gestion des flux .....	44
2.3.6. Considérations annexes.....	46
<b>3. Conclusion et ouverture</b> .....	<b>49</b>
3.1. Question de l'interopérabilité avec les anciennes machines : le « <i>bridging</i> » .....	49
3.2. Développements informatiques .....	49
3.3. Ouverture.....	51

---

<b>PARTIE PRATIQUE : PROJET D'INGÉNIERIE D'UN STUDIO POLYVALENT</b>	
<b>EN AUDIO SUR IP .....</b>	<b>54</b>
<b>1. Présentation générale du problème.....</b>	<b>56</b>
1.1. Finalité du projet.....	56
1.2. Contexte.....	57
1.2.1. Présentation des installations .....	57
1.2.2. Suites prévues.....	59
1.2.3. Nature des prestations demandées.....	59
1.2.4. Parties concernées par le déroulement du projet .....	59
1.3. Énoncé du besoin (finalités du produit) .....	59
1.3.1. Acquisition de l'audio .....	60
1.3.2. Mise en réseau .....	60
1.3.3. Sécurité des biens et des personnes .....	61
1.3.4. Extensions fonctionnelles souhaitées .....	61
1.4. Environnement du produit recherché .....	62
1.4.1. Listes exhaustives des éléments (personnes, équipements, matières...) et contraintes (environnement) .....	62
<b>2. Cadre de réponse .....</b>	<b>64</b>
2.1. Éléments de Réseau.....	64
2.1.1. Choix des câbles.....	64
2.1.2. Diaphonie.....	65
2.1.3. Choix des switchs .....	66
2.2. Audio sur IP.....	68
2.2.1. Choix du matériel de conversion analogique-numérique.....	68
2.2.2. Réutilisation du matériel existant.....	69
2.2.3. Choix de l'horloge.....	70
2.3. Fonctionnalités secondaires .....	71
2.3.1. Choix des caméras IP .....	71
2.3.2. Choix d'un serveur NAS.....	72
2.4. Sécurité .....	73
2.4.1. sécurité électrique .....	73
2.4.2. sécurité informatique.....	74
<b>3. Propositions techniques .....</b>	<b>75</b>
3.1. Matériel non-spécifique au protocole utilisé .....	75
3.2. Proposition technique en Ravenna .....	77
3.2.1. Matériel spécifique .....	80
3.2.2. Schéma de réseau et longueur de câble nécessaire .....	81

---

3.2.3. Logiciels.....	83
3.3. Proposition technique en AVB .....	85
3.3.1. Matériel.....	86
3.3.2. Configurations de routage.....	87
3.3.3. Récupération dans un logiciel d'édition.....	88
3.4. Et en pratique ?.....	89
3.5. Conclusion de cette mise en pratique .....	92

**Bibliographie ..... 93**

Documents techniques .....	93
Normes de l'IETF (Internet Engineering Task Force).....	93
Ouvrages.....	94
Mémoires.....	94
Sources internet.....	94
PARTIE THÉORIQUE.....	94
Table des Illustrations.....	95
Partie Pratique.....	95

**ANNEXES ..... 96**

<b>1. Liste matériel studio La Beaudelie.....</b>	<b>97</b>
1.1. Salle .....	97
1.2. Studio.....	97
1.3. Régie.....	98
<b>2. ANNEXE 2 : Synoptique Ravenna solution complète.....</b>	<b>100</b>
<b>3. ANNEXE 3 : Synoptique Ravenna solution partielle.....</b>	<b>101</b>
<b>4. ANNEXE 4 : Synoptique AVB solution partielle.....</b>	<b>102</b>
<b>5. ANNEXE 5 : Synoptique AVB solution complète.....</b>	<b>103</b>
<b>6. budget Ravenna partiel.....</b>	<b>104</b>
<b>7. budget AVB partiel.....</b>	<b>105</b>

---

## PARTIE THEORIQUE : État de l'art

## 1. INTRODUCTION, TERMINOLOGIE ET DÉFINITIONS

*Dans cette partie sera exposée la problématique du mémoire, ainsi que les termes associés à l'audio sur IP, dont tous les acronymes courants dans ce domaine.*

Les réseaux de télécommunications ont connu, depuis les années 60, un fabuleux essor, notamment pour tout ce qui concerne le domaine informatique. Depuis les premiers ordinateurs, puis les premiers ordinateurs personnels, leurs réseaux propriétaires (ARPANET fut lancé en 1969 sur les réseaux téléphoniques) ont évolué, plusieurs protocoles ont émergé. Il y eut d'abord **X.25** qui fonctionne sur le principe de la commutation de paquets, l'**ATM** composé de cellules (données regroupées en paquets de taille fixe), et, plus tard, le protocole **IP** (*Internet Protocol*).

Ce protocole, qui fut inventé pour répondre à un besoin accru de compatibilité et de souplesse lorsque les réseaux locaux ont commencé à s'interconnecter, est actuellement le plus répandu, et a largement contribué à la formation d'Internet. La suite de protocoles TCP/IP a été normée en 1976 puis adoptée en 1983 par le DARPA (*Defense Advanced Research Projects Agency* : département de recherche de pointe de l'armée américaine).

En parallèle, dans les années 1970, le monde de l'audio faisait ses premiers pas dans le domaine de l'audio numérique. Plusieurs normes, que nous verrons succinctement, ont vu le jour.

Ce mémoire est un tour d'horizon actuel de la rencontre entre le monde des réseaux informatiques et celui des workflows audio numériques professionnels : la naissance, le développement, et les évolutions possibles de l'audio sur IP. Il sera question des protocoles actuels, des normes importantes, et des efforts effectués pour les rendre interopérables, ainsi que de l'avantage apporté par des solutions d'audio professionnelles sur IP, tant à court terme qu'à long terme.

## 1.1. TERMINOLOGIE ET DÉFINITIONS

Les univers des réseaux informatiques et de l'audio numérique comportant de nombreux termes techniques et abréviations, il convient de préciser les plus courants dans le cadre de ce mémoire. D'autre part, il est nécessaire de définir certains concepts liés à la problématique traitée ici : les problèmes soulevés par l'interopérabilité des protocoles les plus courants.

### 1.1.1. ABRÉVIATIONS COURANTES

AoIP : Audio Over IP. En français, on pourrait traduire par Audio sur IP.

AoE : Audio over Ethernet. En français : Audio sur Ethernet.

IP : Internet Protocol (*voir partie dédiée au protocole*)

LAN : Local Area Network ou réseau local

WAN : Wide Area Network : réseau de grande taille, pouvant comprendre plusieurs sous-réseaux

VLAN : Virtual LAN, réseau local virtuel

PTP : Precision Time Protocol, ou IEEE-1588 (*la partie 1.5 lui est dédiée*)

QoS : Quality of Service. En français, qualité de service. C'est la faculté que peut avoir un réseau dans la priorisation des paquets constituant les flux.

PoE : Power over Ethernet, ou courant porté par Ethernet.

### 1.1.2. DÉFINITIONS

**L'interopérabilité** se définit comme la capacité que possède un produit ou un système dont les interfaces sont intégralement connues à fonctionner avec d'autres produits ou systèmes existants ou futurs (Marty et Pillot, 2009).

D'après la définition du Larousse, la **compatibilité** est une "qualité relative de deux procédés de télécommunication pouvant fonctionner avec les mêmes appareils" ou encore "qualité relative de deux ordinateurs [ou dans d'autres cas, appareils], dont l'un peut exécuter des programmes écrits pour l'autre sans

traduction ni réécriture". On a donc une définition portant sur le protocole de transmission et sur le matériel.

Le domaine de l'audio sur IP reste encore très récent et mouvant, et a besoin d'un accord commun des concepteurs et constructeurs pour se développer. La situation peut rappeler l'arrivée de l'audio numérique lorsqu'aucun standard n'avait encore été établi. Il faut veiller notamment à la prise en charge de la flexibilité et la facilité d'usage, qui passe par une interopérabilité des protocoles et une compliance du matériel.

C'est cette mission que s'était donné le groupe de recherche de l'AES (Audio Engineering Society), sous le code AES-X192, et il a été récemment défini une nouvelle norme de transport : l'AES 67, qui permettrait une interopérabilité entre différents protocoles de transmission de données tels que Dante, le Ravenna, Q-Sys, AVB...

Un **standard technique ou technologique** (qui, en anglais n'est pas distingué d'une **norme**) résulte de la définition de règles d'harmonisation dans un secteur ou sur un marché. Il est effectué par un organisme reconnu au niveau national ou international, et sert de base aux différents constructeurs du secteur lorsqu'on parle de **norme**, et est proposé par un industriel pionnier ou influent dans le domaine dans le cas du **standard**.

La **scalabilité** est un critère de qualification des réseaux concernant leurs capacités d'expansion, ou plus étymologiquement, d'échelle, que ce soit en termes de ressources, de matériel, ou de capacité. Il n'existe pas réellement d'équivalent dans la langue française, outre dans certaines publications *extensibilité*, *facteur d'échelle*, *évolutivité*. Dans un réseau, c'est un critère de qualité que de prévoir des extensions futures, la technologie étant sujette à une évolution constante, et des modifications fonctionnelles pouvant être effectuées.

Dans le cadre d'un studio de prise de son, d'une chaîne de broadcast, ou de toute autre installation audio plus ou moins fixe, on peut facilement vouloir ajouter du matériel plus récent, plus puissant, ou le remplacer. La scalabilité est donc cette faculté à accueillir, en termes qualitatifs et quantitatifs, des extensions.

**Nœud** ou **hôte** : c'est un élément connecté sur le réseau, comme un ordinateur, une imprimante, une console réseau, ou un boîtier de conversion analogique-numérique AoIP.

**Modèle OSI** (Open System Interconnection) : c'est un modèle défini en 1970 et normé en 1984. Exposé dans la norme ISO-7498, puis révisé en 1994, il se constitue de 4 rubriques<sup>1</sup> :

- ISO-7498-1 (révisé en 1994): Modèle de Base
- ISO-7498-2 (datant de 1989): Référence pour la sécurité
- ISO-7498-3 (révision de 1997): Nommage et Adressage
- ISO-7498-4 : (datant de 1989) : Cadre général de Gestion

Les ordinateurs, stations, programmes, périphériques, opérateurs humains capables de transmettre et de traiter l'information sont appelés « *Real Systems* » ou « systèmes réels ». Leur appellation « *Open Systems* », ou systèmes ouverts, s'applique dans la modélisation proposée par la norme.

Il est important de noter que dans ce modèle, ou dans le modèle TCP/IP, on peut distinguer la connexion physique (qui concerne le niveau 1) des connexions virtuelles grâce aux différents niveaux.

Ces différents niveaux sont normés en tant que :

- couche 1 : niveau « physique ». C'est la modulation du signal électrique (sur médium cuivre) ou lumineux (sur fibre optique) qui va être décodée en signaux binaires : les « 0 » et les « 1 » correspondant aux « bits » informatiques.
- couche 2 : niveau « lien ». C'est la liaison élémentaire d'information entre les machines, les bits pouvant être interprétés en différents octets, qui, selon leur nature, peuvent déjà créer le « lien » d'interconnexion le plus basique entre les éléments – appelés **nœuds** (« nodes » en anglais) d'un réseau. Ces éléments dialoguent grâce à leur adresse MAC (Media Access Control), qui est l'adresse matérielle unique de leur interface (NIC, ou Network Interface Card). Le protocole « Ethernet », par exemple, se situe à ce niveau.

Une adresse MAC est codée sur 6 octets, et représentée en hexadécimal, par une suite de 12 caractères groupés par 2 et séparés par 2 points. Ces adresses sont

---

<sup>1</sup> informations disponibles sur le site de l'ISO : <http://www.iso.org/>

uniques, placées dans une ROM de la carte réseau par le constructeur. On peut éventuellement les masquer de manière logicielle avec *ipconfig* sous Windows ou *ifconfig* sous les systèmes Unix (Mac et Linux).

Par exemple :

00:1D:71:50:5D:27

Il faut savoir qu'au niveau 2, les trames adressées ne peuvent être transportées qu'au sein d'un même réseau local. Ces informations peuvent donc franchir des switches, mais pas les routeurs (voir partie sur le matériel réseau pour une approche de leur fonctionnement).

À ce niveau, l'information circulant s'appelle la « **trame** ».

- couche 3 : niveau « réseau ». C'est la couche à laquelle le réseau peut réellement commencer à être considéré comme tel : aux adresses matérielles des nœuds sont associées des adresses réseau, souvent appelées « adresses IP ».

Au niveau IP, les trames sont interprétées en « **paquets** » ou « **datagrammes** ».

- couche 4 : niveau « transport ». Cette couche concerne la façon dont sont acheminées les informations entre les deux machines. Le TCP et l'UDP font partie de cette couche. C'est à ce niveau que s'établit vraiment la communication entre les hôtes : on n'est plus dans les sous-couches d'acheminement des données, mais dans la couche où les hôtes vont commencer à communiquer. C'est également à ce niveau qu'intervient la notion de « port ».

Un port est souvent associé à un processus de niveau supérieur, qui communique un flux de données, via le réseau, vers un autre processus.

À ce niveau, les paquets sont désencapsulés en « **segments** ».

- couche 5 : niveau « session ». C'est à ce niveau que la communication est mise en forme entre les hôtes. Par exemple, le RTP (Real-Time Transport Protocol) contrôle le synchronisme des flux entre les hôtes à ce niveau.

À partir de cette couche, des données sont extraites des segments.

- couche 6 : niveau « présentation » : ce niveau, comme son nom l'indique, arrange la syntaxe des données pour la couche suivante.
- couche 7 : niveau « application ». C'est le niveau qui concerne le traitement du flux réseau par les programmes de l'ordinateur, et qui intéresse donc l'utilisateur.

Toutes ces couches sont **encapsulées** les unes dans les autres, autrement dit imbriquées, et sont interprétées à différents niveaux de transmission par les appareils ou logiciels de traitement. [9]

### **1.1.3. ORGANISMES DE FORMATION, NORMES PRINCIPALES**

#### **1.1.3.1. INTERNET**

##### **IEEE : Institute of Electrical and Electronics Engineers.**

Fondée en 1962, cette association professionnelle subdivisée en 38 sociétés a pour vocation la recherche, l'éducation, et la normation technique des nouvelles technologies. Les domaines adressés sont aussi vastes que l'océan électronique.

Dans le cadre de ce mémoire, les technologies concernant le réseau ont souvent été normées par l'IEEE. Le nommage des normes se fait par accollement du nom du groupe de travail avec des codes de lettre pour les sous-catégories, et parfois l'année de définition ou d'actualisation.

À titre indicatif, voici quelques normes importantes dans ce mémoire venant de l'IEEE [14] :

- 802.3 pour le protocole Ethernet, datant de 1983 et révisé pour la dernière fois en 2015.
- IEEE 802.1 pour le groupe de travail audio et vidéo, à l'origine de l'AVB. Plusieurs normes y sont rattachées, détaillées dans la partie traitant de ce protocole.
- IEEE-1588 pour le Precision Time Packaging, défini en 2002 pour la version 1 et actualisé en 2008 pour la version 2.

##### **IETF (Internet Engineering Task Force)**

Ce groupe de travail, dédié aux protocoles Internet, existe depuis Janvier 1986. L'IETF est, entre autres, à l'origine de la suite de protocoles TCP/IP. Sa mission est :

« [d'] améliorer le fonctionnement d'Internet en produisant des documents techniques pertinents et de qualité, propres à influencer la façon dont Internet sera pensé, utilisé, administré. »

Ses principes fondamentaux sont **l'ouverture** publique des travaux à travers la publication de « RFC » (ou *Request for Comments*, c'est-à-dire des normes appelant à l'annotation ou au commentaire), la **compétence technique**, un travail et des

membres **volontaires**, une **expérimentation technique**, et la **responsabilité pour une norme publiée**. [3]

### 1.1.3.2. AUDIO SUR IP

L'AES (Audio Engineering Society) a mis en place un groupe de travail, appelé X192, dévolu à la recherche dans le domaine de l'audio sur IP. C'est ce groupe de travail qui a, en 2013, publié une norme : l'AES67. Cette nouvelle norme, venue assez tardivement dans le monde de l'audio sur IP, est destinée à rendre possible l'interopérabilité entre divers protocoles, ce qui n'était pas le cas auparavant. Nous verrons dans la partie sur les protocoles IP pourquoi ce n'était pas encore le cas, et comment il est possible de mettre en place cette interopérabilité.

Il existe d'autres groupes de travail de l'AES, étudiant les problématiques de l'Audio sur IP. Le groupe X210 (1) travaille sur :

« la spécification d'une architecture évolutive des protocoles de contrôle pour les réseaux multimédia professionnels. La version initiale ne traitera que les aspects concernant l'audio. Elle devrait, à terme, étendre le domaine d'intérêt à la vidéo grâce à une collaboration avec un organisme de standardisation orienté vidéo tel que SMPTE. ». *[paragraphe traduit]*<sup>2</sup>

---

<sup>2</sup> présentation du groupe AES X210 : <http://www.aes.org/standards/meetings/init-projects/aes-x210-init.cfm>

## **1.2. PRINCIPAUX RÉSEAUX AUDIONUMÉRIQUES ET ÉVOLUTION**

Nous allons maintenant examiner les technologies audio numériques les plus courantes et leur évolution, afin de pouvoir les comparer avec les technologies AoIP, et éventuellement voir comment elles pourraient être intégrées dans un workflow sur réseau IP.

### **1.2.1. BREF HISTORIQUE ET CONTEXTE**

La question de l'enregistrement, de l'écriture sur un support matériel et de la conservation du son est une problématique philosophique, culturelle et technique aussi ancienne que la notation musicale, si on accepte d'y voir les prémisses d'une volonté de mémoire sonore.

Le principe d'échantillonnage et quantification du son repose sur la théorie énoncée par Nyquist en 1928 au Bell Telephone Lab. La modulation par impulsion codée (en anglais, PCM : *Pulse Code Modulation*) correspond à la discrétisation d'un signal en « échantillons » de durée  $1/f$  (avec  $f$  la fréquence d'échantillonnage), et codés en amplitude sur un nombre donné de bits appelé « quantification ». La fréquence d'échantillonnage, d'après le théorème de Nyquist-Shannon, doit être supérieure à deux fois la fréquence maximale échantillonnée afin d'éviter les repliements de spectre. [8]

Le résultat obtenu est une liste d'octets adaptée à un traitement informatique, et peut être à nouveau convertie dans le domaine analogique.

Des algorithmes de codage, tant génériques que spécifiques, ont été élaborés pour compresser ces données et faciliter le transport et l'enregistrement sur divers supports tels que les bandes magnétiques, les Compact Discs ou les disques durs. Ces algorithmes sont pertinents dans les réseaux de voix sur IP (VoIP) ou le streaming. Toutefois, le matériau sonore ne doit pas être compressé de façon à pouvoir le travailler de manière professionnelle, et ainsi garder toute la latitude de transformation et de traitement du signal possibles.

## **1.2.2. RAPPEL SUR LES PRINCIPAUX PROTOCOLES ET TECHNOLOGIES**

Dans l'audio numérique, on trouve divers médiums de transport et de codage physique du PCM. [8]

### **1.2.2.1. ADAT OPTICAL INTERFACE OU LIGHTPIPE**

L'ADAT Lightpipe, développé dans les années 1990 par Alesis en tant que format propriétaire de transport audionumérique pour ses machines, a par la suite été utilisé par de nombreux constructeurs.

En effet, l'ADAT a pour particularité d'être transmis sur 8 fibres optiques via une connectique Toslink. Cette liaison permet de transporter 8 canaux à 24 bits/48 kHz tout en évitant les éventuels problèmes de boucle de masse générés par des câbles en cuivre. Les transmissions se font en simplex, de la machine émettrice vers la machine réceptrice.

Ainsi, il est important qu'une machine envoyant un signal ADAT soit l'horloge maître, ce qui peut poser problème dans des réseaux plus complexes. Pour pallier cet inconvénient, la plupart des machines ADAT modernes permettent une entrée Word Clock ou une entrée propriétaire de synchronisation ADAT, permettant de réduire le jitter.

On peut voir sur ce schéma de l'entrée de synchronisation que le signal d'horloge est à 48 kHz, ce qui permet de synchroniser à l'échantillon près.

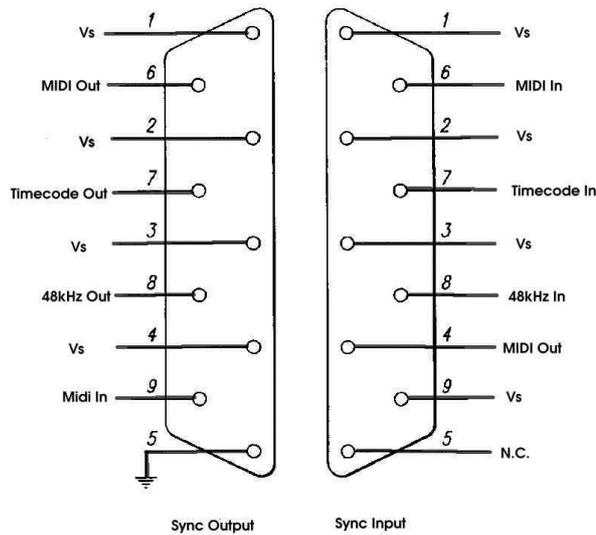


Image tirée de *Secrets of the Adat interface* par Dave MALHAM<sup>3</sup>

Une entrée MIDI et une entrée Timecode peuvent également être utilisées. Le Timecode n'est pas une horloge de précision ; on lui préférera, pour des applications audio seules le Word Clock, qui peut monter à 192 kHz si toutefois la machine synchronisée le permet.

Le SMUX (*Sample Multiplexing*), mis au point par Sonorus, est un multiplexage du signal sur plusieurs fibres. Cette technique permet de transmettre, à taux de transfert brut égal de 15 Mo/s, un flux comportant moins de canaux pour une fréquence d'échantillonnage plus élevée sur une liaison ADAT. Par exemple, à 96 kHz, on pourra transmettre 4 canaux. On a donc une augmentation du débit d'information.

### 1.2.2.2. AES/EBU (AES3)

L'AES/EBU ou AES3 permet de transporter de l'audio numérique en série à 24 bits/96 kHz, sur deux canaux séparés ou en stéréo. Son support matériel est habituellement une paire blindée associée à une connectique XLR, mais on trouve parfois du matériel capable de transmettre jusqu'à 4 canaux d'AES3 sur les 4 paires d'un câble « Ethernet » Cat5e.

Ce codage, fonctionnant par trames de 32 bits, permet de transmettre un signal d'horloge inclus dans le flux. L'audio peut être codé sur 24 bits au maximum.

<sup>3</sup> ADAT reverse engineering et signal: ADAT reverse engineering et signal : <http://www.dmalham.freemove.co.uk/adat.html>

### 1.2.2.3. S/PDIF

Le S/PDIF est souvent présenté comme la version « non professionnelle » de l'AES/EBU. En effet, le protocole est très semblable, mais les connectiques diffèrent : le transport se fait sur un câble asymétrique, et utilise une connectique dédiée à ce type de signal. Les trames transmises sont très similaires à celles de l'AES/EBU.

### 1.2.2.4. MADI ou AES10

Le MADI, ou AES10, peut être considéré comme une évolution de l'AES3 dans le sens où il permet de transporter, via un câble coaxial, jusqu'à 64 canaux à 96 kHz. Un multiplexage temporel est effectué, le *Time Division Multiplexing*.

### 1.2.2.5. AES50, SUPERMAC, HYPERMAC

L'AES50, SuperMAC et HyperMAC sont des protocoles similaires de transport sur un câble Ethernet, mais fonctionnant à la couche 1. Ils utilisent la modulation de la couche physique pour transmettre les données audio d'un appareil à l'autre.

Ces protocoles ont des caractéristiques intéressantes pour la sonorisation : un codage de canal robuste permettant une récupération des erreurs, une vérification de redondance cyclique (*Cyclic Redundancy Check*) pour la correction d'erreur, un débit élevé et une latence faible et prévisible (on peut parler de *déterminisme*).

SuperMAC permet de transporter de 6 canaux bidirectionnels de PCM à 24 bits et 384 kHz, à 48 canaux bidirectionnels à 24 bits et 44.1 ou 48 kHz, avec une latence de 61,5 µs. HyperMAC permet jusque 192 canaux avec 41,66 µs de latence. Ils peuvent tous les deux intégrer des données de contrôle via TCP/IP.

Ces signaux ne sont pas destinés à être routés sur un réseau standard, ni à être acquis via la carte réseau d'un ordinateur, des interfaces de conversion existent. Toutefois, l'idée d'un véritable réseau, où chaque machine peut communiquer avec les autres, est déjà sous-jacente grâce à l'existence de « routeurs » dédiés, qui ressemblent plus à des patchs de connexion. Ils permettent également de router les canaux séparément. Le transport se fait également sur un médium d'usage courant dans un réseau informatique traditionnel, Cette connectique similaire peut encore induire en erreur avec certaines consoles de sonorisation « en réseau ».

Par exemple, les consoles de sonorisation Aviom ont un standard propriétaire appelé A-net, qui n'est pas routable, destiné à recevoir les signaux d'un boîtier de scène numérique (« *digital snake* »). Il est intéressant de noter que Aviom a conclu un accord avec Dante, et ce matériel peut donc être intégré dans un réseau AoIP.

## **2. RÉSEAUX AUDIO SUR IP (AOIP)**

La naissance de l'audio en réseau est née d'un besoin grandissant de transporter l'audio non-plus « point à point », comme les protocoles de transport audionumériques cités dans ce mémoire, mais d'une source vers plusieurs destinations.

Nous allons commencer par examiner les technologies des réseaux traditionnels, afin de comprendre comment fonctionnent les réseaux audio dans un tel environnement.

### **2.1. TECHNOLOGIES DES RÉSEAUX INFORMATIQUES, FOCALISATION SUR IP**

Il existe, depuis l'invention du réseau informatique, plusieurs protocoles destinés à la communication entre les machines. Le protocole IP est né d'une volonté d'harmonisation de ces flux : en effet, sans reconnaissance de l'information d'une machine à l'autre, le réseau ne peut pas exister, et c'est une véritable Babel qui naît de cette absence de compréhension.

#### **2.1.1. REPÈRES HISTORIQUES**

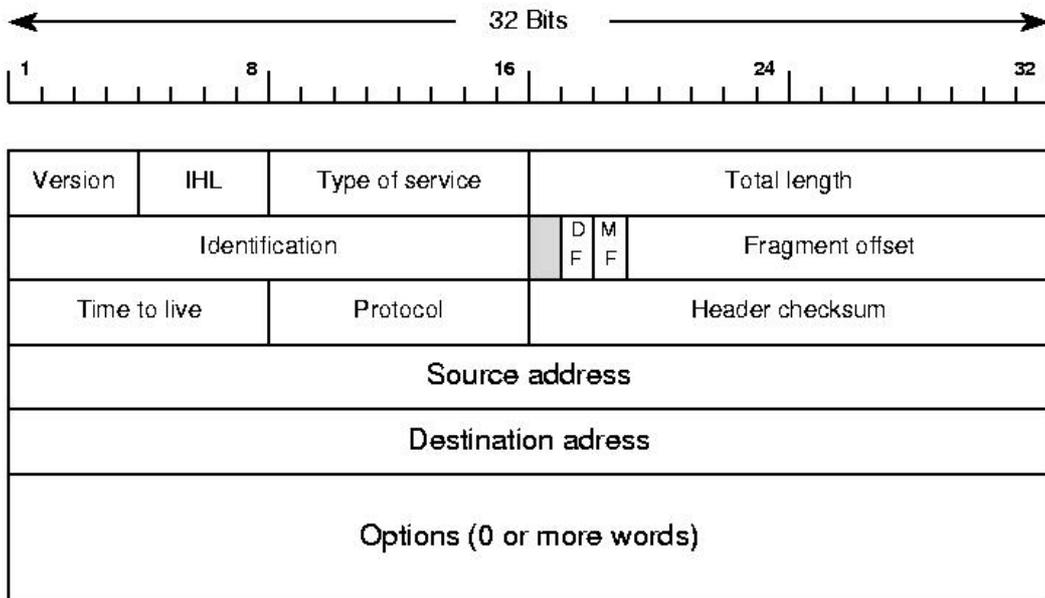
Initialement prévu pour le transfert asynchrone de données, le protocole IP avait pour but la normalisation des réseaux. En effet, l'interconnexion des réseaux qui a donné naissance à Internet repose sur la possibilité de transmettre des informations entre des machines sur des réseaux différents.

Ce n'est pas à la base un protocole optimisé pour les flux sensibles au temps, mais l'ajout de protocoles sur les couches logiques supérieures, ainsi que l'augmentation des débits de transmission, permettent aujourd'hui de développer des réseaux audio fonctionnant sur ce protocole.

### 2.1.2. LE PROTOCOLE IP : DÉFINITIONS

**IP** : Internet Protocol. C'est un protocole de communication en réseau situé au **niveau 3** du **modèle OSI**. Une modélisation spécifique au protocole IP a été proposée en parallèle et complément du modèle OSI. Toutes les couches sont également encapsulées les unes dans les autres :

- la couche « lien », correspondant au niveau **1** «**link**» du modèle OSI, et donc, par exemple, aux trames Ethernet
- la couche « réseau » aussi appelée couche « Internet » : c'est le niveau de la couche IP en tant que telle, **niveau 3** sur le modèle OSI.
- la couche « transport » : c'est à ce niveau que les données sont supervisées pour vérifier qu'elles sont arrivées à bon port. TCP et UDP en font partie, et, par exemple, les trames d'ACK (acknowledgement ou « reconnaissance » en TCP) vont se trouver.
- la couche « application » qui regroupe toutes les couches supérieures du modèle OSI. On remarque ainsi que le modèle IP ne tient compte ni de la couche physique, ni de la précision des strates au-dessus de la couche transport.



*En tête d'une trame IP. Source : le site du Netanya College*

Le protocole IP existe actuellement en 2 versions : IP version 4, appelé **IPv4**, codée sur 32 bits et représentée par une suite de 4 nombres chacun compris entre 0 et 255 séparés par des points, et une version 6, sur 128 bits, représentée par 16 groupes de 4 chiffres hexadécimaux (base 16, utilisation des caractères de 0 à f) séparés par deux points. Le caractère répété « :: » permet de remplacer, une seule fois dans l'adresse, des zéros consécutifs.

Exemple :

1001:c03d::1337:4311:dead:f001 :5e11:c001:e995

La version 6 a été implémentée pour répondre au besoin croissant de nouvelles adresses IP, utilisées pour communiquer sur Internet : on passe ainsi de  $2^{32}$  adresses possibles en IPv4, ce qui équivaut à 4 294 967 296 adresses, à  $2^{128}$  adresses possibles, soit plus de  $340 \times 10^{33}$  (décilliards) périphériques connectés par mètres carrés de surface terrestre habitable. Dans un monde où les « réseaux de nouvelle génération », comprenant les smartphones, les objets connectés, et les appareils dédiés à la domotique se développent, cela laisse une marge quasi-illimitée de progression avant d'avoir à développer une nouvelle norme d'adressage.

Elle n'est toutefois pas encore implémentée pour les réseaux AoIP, mais serait en réflexion pour de futurs développements. En effet, au sein d'un LAN aussi gros soit-il (comme dans un complexe de broadcast avec de nombreuses stations), le système d'adressage IPv4 est largement suffisant en termes numériques : de par le concept de « classes » de sous-réseau, on pourra sans trop de soucis adresser de manière unique chaque poste.

Il est toutefois intéressant de noter que la version 6 permet de nouvelles fonctionnalités, comme l'étiquetage des flux dans l'en-tête des trames (*Label Flow*) approprié à un service de qualité, (ou QoS), plus efficace, des algorithmes de détection de la taille du MTU entre deux machines, ou une plus grande sécurité des flux avec IPSec.

### 2.1.2.1. RÉSEAUX ET CLASSES, SOUS-RÉSEAUX, NOTATION CIDR

Dialoguer en permanence avec un très grand nombre de machines peut poser des problèmes à cause du grand nombre de paquets circulant en même temps, et risquant donc d'être perdus ou supprimés. Pour être fonctionnels, on subdivise souvent les réseaux en réseaux plus petits, ou « sous-réseaux ».

Il existe ce qu'on appelle des « **classes** » de sous-réseau, qui vont théoriquement de A jusque E. Elles sont, dans les faits, utilisées jusqu'à C, les classes D et E étant respectivement dévolues au multicast et à la recherche.

Différentes plages d'adresses sont donc allouées à chaque classe différente en IPv4 :

Classe A : 1 octet pour le réseau (moins un bit réservé, 0.0.0.0 est la « passerelle par défaut, et 127.0.0.0 est l'adresse de la machine locale), 3 octets pour les nœuds. On a donc, en plage d'adresse :

1.0.0.0 à 127.255.255.255 : soit 256 réseaux et 33 554 430 hôtes.

Classe B : 2 octets pour le réseau (14 bits + 2 bits réservés) et 2 octets pour les nœuds : les adresses vont donc :

de 128.0.0.0 à 191.255.255.255, soit 3 276 réseaux et 32767 hôtes.

Classe C : 3 octets réseau (21 bits + 2 bits réservés), 1 octet pour les nœuds. On a donc en adresses possibles :

de 192.0.0.0 à 223.255.255.255, soit 4 194 303 réseaux différents avec 254 nœuds chacun.

On remarque donc que pour la classe A, on peut faire moins de sous-réseaux mais accueillir plus d'hôtes, tandis qu'en classe C, l'inverse se produit. Sur Internet, l'allocation en classe A se fait hiérarchiquement de la classe A vers la classe C.

### 2.1.2.2. NOTIONS DE « MASQUE » ET « PASSERELLE »

Un **masque de sous-réseau** sert à préciser dans quelle classe de réseau on se trouve. En effet, le nombre de bits positifs indique l'adressage alloué au sous-réseau, et le reste est disponible pour ses hôtes.

La notation CIDR (*Classless Inter-domain Routing*), utilisée tant en IPv4 qu'en IPv6, correspond au nombre de bits positifs du masque de sous-réseau.

Elle est représentée par « /n » où « n » est le nombre de bits positifs en partant depuis le début de l'adresse.

Par exemple :

/16 correspond à un masque de sous réseau

« 11111111.11111111.00000000.00000000 », soit en décimal : 255.255.0.0

Cette notation est quasi-systématique dans un réseau IPv6, dont les adresses sont beaucoup plus longues que celles en IPv4. [15]

On appelle « **passerelle** » (*gateway* en anglais) un nœud intermédiaire entre deux réseaux différents. Il existe des passerelles à tous les niveaux du modèle OSI. Par exemple, un pont (aussi appelé *bridge*), qui convertit un protocole en un autre, comme un routeur Wifi, fonctionne au niveau « liaison ». Un routeur pourra servir de passerelle au niveau « réseau », et donc changer l'adressage IP des hôtes en fonction des réseaux à traverser. [9]

### 2.1.2.3. MULTICAST

La plupart des adressages sur les réseaux traditionnels se font en **unicast**. C'est une communication entre deux hôtes.

Il existe, dans le protocole IP, des fonctionnalités d'adressage collectif : la première est le **broadcast**, où un hôte s'adresse à tous les hôtes du réseau ; et la seconde, qui pourrait nous intéresser dans le cadre de l'AoIP, est le **multicast**.

L'adressage en multicast permet en effet à plusieurs hôtes de récupérer les mêmes paquets : ainsi, on économise de la bande passante. Pour ce faire, une adresse MAC spéciale existe dans la trame Ethernet, et une plage d'adresse IP, la « **classe D** », est réservée à cet usage.

Un protocole spécial permet aux hôtes de rejoindre des groupes multicast : l'IGMP (*Internet Group Management Protocol*). Ce protocole peut actuellement être implémenté directement dans les switches.

L'IGMP (*Internet Group Management Protocol*) existe en 3 versions. Grâce à l'écoute IGMP (« snooping ») un switch peut écouter les trafics entre les machines, et déterminer quelles doivent être les connexions multicast. Sans ce protocole, un switch va envoyer un multicast en broadcast et encombrer le réseau. [9] [18]

#### **2.1.2.4. CHOIX DE TCP OU UDP EN AOIP**

En IP, il existe plusieurs protocoles de transmission des données et d'informations entre les machines : le **TCP** et l'**UDP**. On entendra souvent parler de TCP/IP pour désigner le premier. Ces protocoles sont à la couche 4 du modèle OSI, et viennent donc se superposer à IP.

TCP: Transmission Control Protocol. Ce protocole concerne l'échange de **données**, c'est-à-dire des informations identifiables par les ordinateurs comme telles, contrairement aux couches inférieures du modèle OSI.

Comme son nom l'indique, ce protocole de transmission requiert une réponse du receveur vers l'expéditeur pour valider l'arrivée d'un segment de données. On appelle ce mode de transmission le « mode connecté ».

UDP: User Datagram Protocol. Ce protocole est également une possibilité de communication sur la couche IP. Contrairement au TCP, il fonctionne en mode « non-connecté » : l'arrivée à bon port des segments n'est pas validée, c'est-à-dire que la source ne sait pas s'ils sont arrivés jusqu'à l'hôte de destination. C'est un avantage concernant le trafic réseau : moins de paquets circulent, et la transmission réelle des données se fait de manière plus rapide, car elle ne requiert pas de confirmation. En revanche, les paquets qui se seraient perdus ne seront pas retransmis : c'est là l'inconvénient de l'UDP.

L'UDP permet uniquement un contrôle de type « checksum » pour déterminer si les en-têtes ou les données utiles ne sont pas corrompues, mais ces données ne seront pas retransmises.

Dans les faits, les réseaux IP ont été développés pour permettre l'interconnexion des réseaux de manière compatible. Ce besoin de compatibilité est adressé au détriment de l'optimisation : des fonctionnalités annexes doivent être définies. Tout le problème du transit de l'audio sur protocole IP réside dans l'absence par défaut de synchronisation entre les machines, d'un débit suffisant, et de la fragmentation de l'information en paquets asynchrones et à l'arrivée incertaine.

Le développement matériel au niveau des capacités de traitement des contrôleurs de l'interface réseau, et donc de l'envoi et de la réception de paquets à une plus haute fréquence, a permis d'envisager des envois de données plus importants.

L'audio étant un flux avec des contraintes temporelles, il y avait également nécessité d'avoir des paquets transitant de manière synchrone entre les machines concernées, cadencés par une horloge, et arrivant de manière certaine à destination. Les mécanismes de QoS permettent de relever ce défi, ainsi que le **protocole PTP** ou (« Precision Time Packaging ») normé dans l'IEEE 1588.

Le groupe FNS/ACIP de l'EBU (*European Broadcasting Union*) a préconisé, plutôt que d'utiliser le TCP, d'utiliser RTP (Real-time Transfer Protocol) sur UDP. ACIP signifie *Audio Contribution over IP*, soit la contribution de l'audio sur IP. On peut trouver cette recommandation dans la Fiche Technique 3326 révisée en 2014, concernant l'interopérabilité. En effet, la confirmation nécessaire d'un lien TCP entre deux hôtes du réseau engendre une multiplication des paquets ; de plus, pour un réseau audio comportant un ou plusieurs flux de multiples canaux non-compressés, la latence de la réponse d'un hôte à l'autre pourrait compromettre la rapidité de traitement de l'audio et augmenter la latence. [12]

Cette recommandation concernant le RTP/UDP est reprise dans l'AES67-2013. Avant cette normalisation, tous les réseaux AoIP avaient déjà intégré une utilisation de l'UDP, préférée à celle de TCP pour les raisons pratiques inhérentes aux flux audio sur IP.

### 2.1.2.5. RTP (*REAL-TIME TRANSPORT PROTOCOL*)

Le RTP est un protocole au niveau 5 du modèle OSI : il vient établir des « sessions » entre les hôtes, de façon unidirectionnelle. C'est une couche nécessaire pour transporter les flux audio, au dessus de la couche UDP, elle-même au-dessus de la couche IP.

Défini par l'IETF dans la RFC 3550 pour les applications « temps réel » et complétée par la RFC 3551 dédiée aux profils multimédia audio et vidéo, ce protocole fournit des recommandations pour renseigner le type des données utiles (« payload »), la numérotation des paquets pour pouvoir les remettre en ordre ou déterminer si certains ne sont pas arrivés, et un horodatage en vue de vérifier leur arrivée à bon port.

Il constitue plus un cadre commun pour bâtir différents profils de session possibles qu'une spécification arrêtée et précise de l'implémentation des flux à contrainte de temps, et fait donc appel à plusieurs protocoles complémentaires.

Par exemple, le **RTCP** (*Real-time Transport Control Protocol*), qui est son pendant protocolaire, et sert à obtenir un retour d'information de la source vers l'hôte, concernant l'horodatage ou les statistiques d'arrivée des paquets. Il n'est pas obligatoire dans l'AES67. Il a en effet une certaine utilité pour donner à l'hôte source un retour qualitatif, mais étant sur la même file prioritaire que le RTP sur le réseau, il peut donc potentiellement engendrer des interférences de gestion des paquets de flux utiles au sein des switchs, et donc des pertes. [3] [4] [5] [7]

### 2.1.2.6. SDP (*SESSION DESCRIPTION PROTOCOL*)

Pour établir une session d'un hôte vers un ou plusieurs autres, il est nécessaire d'échanger des métadonnées concernant la nature et les paramètres de transmission via le protocole « *Session Description Protocol* ». On trouvera sa spécification dans la RFC4566. [6]

Voici un exemple, tiré de la partie 8.5.2 de l'AES67-2013, illustrant sa syntaxe. Ce sont les données nécessaires à la création d'un flux audio à 24 bits et 48 kHz entre deux machines :

```

«v=0
o=audio 1311738121 1311738121 IN IP4 192.168.1.1
c=IN IP4 192.168.1.1/32
s=Stage left I/O
t=0
m=audio 5004 RTP/AVP 96
i=Channels 1-8
a=rtpmap:96 L24/48000/8
a=sendonly
a=ptime:0.250
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:0
a=mediaclock:direct=2216659908 »

```

On peut observer, à travers cet exemple, que la définition de paramètres du SDP (défini dans la RFC4566 au paragraphe 9) est représentée par des lettres.

- la variable « v » représente la **version** du protocole, ici « 0 ».
- « o » est le paramètre **d'origine** du flux, décrivant le nom d'utilisateur, puis celui de la session, le numéro de session, le type de réseau (ici : « IN » est pour Internet), le type et la valeur de l'adresse de provenance (unicast, puisque venant d'une machine unique).
- « c » concerne les données de **connexion** : type de réseau, type d'adresse et adresse unicast ou multicast. On rajoute souvent la « durée de vie » des paquets derrière le symbole « / », comprise entre 0 et 255. Cette fonction est obligatoire pour une adresse multicast et permet d'éviter l'engorgement du réseau.
- « i » : concerne les **informations** sur la session. Ici, les canaux 1 à 8.
- « m » : concerne le **media** transporté au cours de la session (ici « audio »), le port d'arrivée sur le récepteur, ici 5004, puis la façon dont il est transporté (RTP/AVP *Real-time Transport Protocol/Audio-Video Profile*)
- « a » définit les **attributs** du flux, notamment, comme montré ci-dessus, l'horloge de référence utilisée via le paramètre « *ts-refclk* ». C'est ici le PTPv2 (IEEE 1588-2008). La séquence hexadécimale qui suit représente l'identifiant de l'horloge Grand Maître (GMID).

Paramètre *rtptime* : donne le code correspondant au type des données, le nom du type d'encodage (L24, soit PCM linéaire sur 24 bits) suivi de la fréquence d'échantillonnage, puis des paramètres d'encodage : ici le nombre de canaux.

Paramètre *ptime* : indique le « temps de paquet » en millisecondes, c'est-à-dire la durée réelle de l'audio inclus dans le paquet. En multipliant par la fréquence d'échantillonnage, on obtient donc le nombre d'échantillons contenus dans le paquet.

Paramètre *mediaclk* : ne fait pas vraiment partie du standard officiel, mais a été implémenté dans « draft-ietf-avtcore-clksrc » pour pouvoir indiquer l'époque de début de la session. En effet, le RTP fournit une horloge « absolue », datant du 1er janvier 1970. On indique ainsi la translation d'horloge à effectuer par rapport au PTP. [6] [7]

**MTU** : Maximum Transfer Unit, ou « unité maximale de transfert ». C'est la taille maximale, en bits, d'une trame physique transitant sur un ou plusieurs réseaux, sans qu'elles aient besoin d'être fractionnées. Le MTU définit par conséquent la taille optimale d'une **trame**. On peut donc définir le **MSS** (*Maximum Segment Size*) comme étant la taille maximale d'un segment IP. Si le segment est plus long qu'un MTU, il risque d'être fragmenté, et s'il est plus court, des bits lui seront ajoutés. De manière courante, un MTU est de 1500 octets. Pour un en-tête Ethernet de 40 octets, on aura donc un MSS de 1460 octets.

En Audio sur IP, il peut y avoir une utilisation complète de la trame Ethernet, mais pour un seul canal, les données concernant l'audio peuvent n'être que de quelques centaines d'octets. On peut aussi faire l'usage du champ DF (*Don't Fragment*) de l'en-tête IP pour éviter la fragmentation, qui pourrait introduire des désynchronisations à l'arrivée. Ce champ peut être dangereux : un paquet marqué « DF » sera détruit s'il est nécessaire de le fragmenter tout de même. Une bonne gestion du MTU sur les appareils du réseau est donc importante. [15]

### 2.1.3. RAPPELS SUR LE RÉSEAU

En s'appuyant sur les mémoires de William Leveugle et Léo Rossi-Roth, ne seront développées que les principales architectures de réseaux (câblés, à la différence de leur sujet de mémoire traitant du Wifi) et les problèmes qui peuvent s'y poser (en termes d'horloge, de routing, de qualité du son, d'intégration des anciens éléments, et de latence).

#### 2.1.3.1. DÉBITS ET LIAISONS

Comme évoqué dans la partie de définitions, une liaison informatique peut être effectuée à divers niveaux protocolaires, tant au niveau physique ou au niveau applicatif.

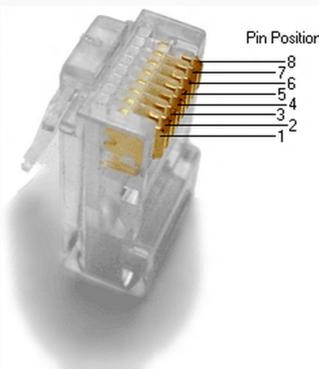
Souvent, en Audio sur IP, on utilise Ethernet comme un protocole de liaison physique et de données.

Au fur et à mesure que l'Ethernet s'est développé, le débit théorique d'une liaison s'est amélioré. On est passé du standard **Ethernet** (*10base2 sur 4 câbles, puis 10Base-T à 10 Mbps*) au **Fast Ethernet** (*100Base-TX : transmission duplex à 100 Mbps*), puis à du **Gigabit Ethernet** (*1000Base-T, fonctionnant en full duplex à un débit de 1Gbps*) de manière courante. Actuellement, le 10G Ethernet et même le 100G Ethernet existent, même s'ils requièrent des connectiques spécifiques telles que la fibre multimode pour transmettre les hyper-fréquences de la couche physique.

Sur une liaison câblée, le plus souvent par câble « Ethernet » avec une prise RJ45 (connecteur 8P8C, 8 points - 8 contacts), la carte réseau interprète les signaux électriques arrivant par les différentes paires. Il existe des câbles droits et des câbles croisés. Deux normes de branchement des embases existent : T568A et T568B, permutant les paires de réception et de transmission. Un câble droit, qui sera adapté pour la plupart des utilisations, aura donc ses deux connecteurs suivant la même norme au choix.

## Wiring

See [modular connector](#) for numbering of the pins<sup>[7]</sup>

Pin	T568A Pair	T568B Pair	1000BASE-T Signal ID	Wire	T568A Color	T568B Color	Pins on plug face (socket is reversed)
1	3	2	DA+	tip	 white/green stripe	 white/orange stripe	
2	3	2	DA-	ring	 green solid	 orange solid	
3	2	3	DB+	tip	 white/orange stripe	 white/green stripe	
4	1	1	DC+	ring	 blue solid	 blue solid	
5	1	1	DC-	tip	 white/blue stripe	 white/blue stripe	
6	2	3	DB-	ring	 orange solid	 green solid	
7	4	4	DD+	tip	 white/brown stripe	 white/brown stripe	
8	4	4	DD-	ring	 brown solid	 brown solid	

Pour des réseaux en 10baseT ou 100baseTx, seules 2 paires sont utilisées, laissant 2 autres paires à disposition. Toutefois, sur les réseaux 1000baseT, aussi appelés Gigabit Ethernet (correspondant à la norme IEEE 802.3ab), les quatre paires du câble sont utilisées pour 4 liaisons full-duplex.

Sur un réseau Ethernet Gigabit cuivré, le moins coûteux à implémenter, le signal est une modulation PAM-5 (Pulse Amplitude Modulation with 5 levels, ou Modulation d'Amplitude d'Impulsions à 5 niveaux). L'horloge utilisée pour générer la modulation a une fréquence de 125 MHz. En utilisant les quatre paires de transmission parallélisées, on obtient un taux brut de 500 Mbps.

La PAM-5 permet donc de définir cinq niveaux électriques, dont quatre correspondent chacun à deux bits, le dernier servant de contrôle et de correction d'erreur.

La transmission Gigabit est donc sensible au bruit, toute erreur d'interprétation de niveau électrique entraînant la perte simultanée de deux bits, par exemple en cas de dérive d'horloge.

IP est une couche « agnostique » du médium traversé. La sous-couche Ethernet, qui sert de liaison entre la modulation électrique et le paquet IP, est-elle la meilleure solution ?

En termes d'infrastructures et de coût, Ethernet est très répandu et bon-marché. Toutefois, il existe d'autres protocoles de transport sur la couche « liaison », tels que l'ATM (*Asynchronous Transfer Mode*).

L'ATM, qui ne sera pas traité dans le cadre de ce mémoire, diffère du modèle IP sur Ethernet : le protocole utilise des « circuits virtuels » implémentés directement dans le matériel (*circuit switching*), qui orientent des « cellules » unitaires de taille constante entre deux machines par multiplexage temporel.

La régularité de cette conception est multiple : la connexion est robuste et fiable, la qualité de service est assurée, et le jitter est très faible. À titre d'exemple, l'AES47 est un protocole de transport audio sur ATM existant depuis 2009.

### **2.1.3.2. TRAITEMENT DE L'INFORMATION ARRIVANT SUR LE CONTRÔLEUR DE L'INTERFACE RÉSEAU (NIC : NETWORK INTERFACE CONTROLLER)**

Certains contrôleurs réseau sont intégrés directement sur la carte mère, d'autres sont implantés sur des cartes communiquant avec l'ordinateur par le bus PCI (*Peripheral Component Interconnect*), PCIe (*PCI express*), ou encore – de manière pratiquement obsolète - ISA (*Industry Standard Architecture*). Certains fonctionnent également sur USB.

Le contrôleur réseau est piloté par un logiciel, capable de gérer les couches supérieures d'information.

Globalement, cette interface réseau, bien souvent intégrée sur une seule puce, est constituée de plusieurs queues de réception (Rx) et de transmission (Tx). Il y a souvent plus de queues de transmission que de réception, car le protocole Ethernet sur un médium physique utilise souvent la règle d'écouter avant d'émettre. Cette règle, la **CSMA/CD** (Carrier Sense Multiple Access with Collision Detection) permet à l'interface réseau d'écouter ce qui transite sur le médium avant d'émettre, d'analyser les flux pour émettre en évitant les collisions, et de réémettre dans cette éventualité.

La meilleure façon d'éviter les collisions, d'ailleurs, est de **segmenter** le réseau, c'est-à-dire éviter que tous les hôtes soient sur la même ligne de transmission. La topologie la plus adaptée pour éviter les collisions est par conséquent la topologie en **étoile** : c'est celle-là qui sera préconisée pour les réseaux AoIP.

Les queues Tx et Rx peuvent être prises en charge de différentes manières. Les nouvelles générations de contrôleurs réseau, par exemple, intègrent un module de désencapsulation IP, ce qui permet de décharger le processeur et le bus PCI dans le cas d'un contrôleur réseau sur carte d'extension : le TOE (*TCP offload engine* ou moteur de décongestion TCP).

Les données reçues peuvent être transmises via accès direct en mémoire pour un traitement plus rapide par l'application récupérant le flux audio.

## **2.1.4. PROBLÈMES SOULEVÉS ET COMPARAISON DES PROTOCOLES EXISTANTS**

### **2.1.4.1. LATENCE MINIMALE**

La latence, dans plusieurs domaines de l'audio tels que la prise de son ou la sonorisation, est un phénomène très gênant. Pour réaliser une prise de son dans de bonnes conditions, des tests perceptifs ont été effectués sur plusieurs instrumentistes professionnels, et le chiffre de **10 ms** a été défini comme valeur à ne pas dépasser dans la spécification de l'AES67. En effet, plusieurs tests psychoacoustiques ont été effectués pour des personnes en configurations de concert ou de prise de son séparée, et un délai compris entre 8 et 25 ms revient à les écarter de 2,7 à 8,5m. C'est un écartement qui peut rester de l'ordre du naturel pour des musiciens. Au-delà, la latence peut devenir gênante et empêcher les artistes de jouer de manière fluide.

La spécification de l'AVB (*Audio Video Bridging*, voir partie 2.2.5 pour une présentation) prévoit deux classes de trafic. La classe A spécifie **2 ms** de latence entre deux points du réseau, et à travers sept « sauts ». La classe B, qui prévoit une éventuelle extension du protocole sur les réseaux sans-fil, spécifie **50 ms**.

La classe A de l'AVB est en deçà des recommandations de l'AES67. Elle est donc adaptée à de l'audio dans les configurations requérant une latence basse. [13]

### **2.1.4.2. QUALITÉ DU SIGNAL**

[12] La redondance d'une transmission permet de limiter la perte de données, mais demande un doublage des infrastructures réseau. En cas de panne d'un des réseaux, le 2ème pourra prendre en charge les informations transmises. Toutefois, cette sécurité a un coût de déploiement non négligeable.

Les réseaux AoIP, pour la plupart, permettent d'instaurer de la redondance : les appareils conçus par Digigram ou Merging Technologies ont deux ports réseau disponibles afin d'implémenter deux réseaux séparés physiquement. En cas de panne sur l'un des deux, la redondance permet d'éviter la perte d'information.

### 2.1.4.3. PROBLÉMATIQUES DE QoS

La QoS (aussi appelée Quality of Service) est une fonctionnalité indispensable pour un réseau convoyant des flux audio de haute qualité (la qualité audio étant fixée à 16 ou 24 bits et 44,1 ou 48 kHz minimum, avec, éventuellement, plusieurs canaux), en PCM (Pulse Code Modulation).

Canaux	Quantification (bits)	Échantillonnage (kHz)	Débit (kilo-octets par seconde ou Ko/s)
1	24	44,1	132,3
1	24	192	576
2	24	192	1152
4	24	192	2304
8	24	192	4608

*Tableau représentant le débit en fonction des canaux audionumériques transportés, de leur quantification et de leur fréquence d'échantillonnage*

La QoS a donc pour but d'assurer la meilleure exploitation de la bande passante pour tous les usagers. Elle assure également que les usagers ayant des contraintes de temps réel puissent effectivement obtenir une priorité suffisante sur les autres pour que ce débit soit respecté.

En IPv4, c'est un champ spécial dans le paquet, le champ TOS (*Type of Service*) qui indique son degré de priorité.

En IPv6, même si cette version du protocole n'est pas encore totalement implémentée dans les réseaux audio sur IP, c'est un système de *flow label* qui est utilisé (ou « étiquetage de flux »).

L'IETF propose deux catégories pour effectuer une QoS :

- IntServ (*Integrated Services*) : services gérés indépendamment, par exemple un flux sera géré en tant que tel.
- DiffServ (*Differentiated Services*) : plusieurs applications peuvent être regroupées et gérées par ce service.

En AoIP, plusieurs applications différentes sont souvent utilisées. C'est donc DiffServ qui est employé, surtout sur les LAN.

Avec DiffServ, il existe 3 classes de services :

- EF (*Expedited Forwarding*), aussi appelé « Service Premium », utilisé pour les transferts de paquets d'horloge en AoIP.

- AF (*Assured Forwarding*) : bonne qualité, regroupant plusieurs types de sous-classes. Pour l'audio sur IP, il est recommandé de mettre une qualité de classe AF41 sur les flux. Cette sous-classe a une basse probabilité de largage de paquet.
- best-effort : c'est la priorité la plus basse, celle où l'arrivée des paquets n'est pas assurée, et dépendra de la gestion des autres transmissions. Les messages de contrôle seront dans cette classe pour les réseaux AoIP. Ils seront également transmis en TCP, pour éviter tout phénomène de perte de paquet sur une donnée de contrôle.

DiffServ permet de déployer une QoS efficace sur un grand réseau IP. Elle sépare l'architecture en 2 composants :

- la technique de transfert
- la configuration des paquets lors du transfert

« Per Hop Behaviour », comportement à la traversée des sauts (PHB)

Le PHB, qui concerne le comportement de routage aux nœuds du réseau, est le nom donné à la configuration des nœuds pour déterminer le traitement des flots différenciés dans le réseau.

DiffServ définit la sémantique des PHB, c'est-à-dire les résultats que l'on souhaite obtenir, mais non les mécanismes spécifiques les régissant ou leur implémentation effective.

Les PHB sont souvent établis une bonne fois pour toutes, alors que les mécanismes peuvent évoluer, soit en étant modifiés ou améliorés. Ils doivent être facilement déployés dans les réseaux IP, par des mécanismes **d'ordonnement**, de **mise en forme**, de **gestion de la perte des paquets**.

DiffServ agrège les flots en « classe », appelées « agrégats ».

Les comportements des routeurs et switches qui le permettent sont déterminés par les **champs DSCP** (DiffServ Code Points), sur 6 bits. Ces 6 bits sont répartis comme tel : CCCDD0, où C est un bit dédié à la classe, D à un numéro de classe.

Pour le Service Premium, cela correspond à un champ DHCP 101 110, et pour l'AF41, à un champ de 100 010.

#### 2.1.4.4. ROUTAGE ET COMMUTATION

Sur les réseaux, on peut rencontrer à la fois des **routeurs** et des **switchs (commutateurs)**, ainsi que des « **hubs** » ou **concentrateurs**.

Les concentrateurs, où l'on peut connecter différents nœuds de réseau, typiquement des ordinateurs, ou interfaces audionumériques avec un port réseau, servent, comme leur nom l'indique, à concentrer les flux de données et les rassembler sur une même ligne, ou **segment**. Ils peuvent donc entraîner des réductions de bande passante.

Ce sont des appareils passifs, qui relayent à tous les ports de sortie les trames entrantes. Ils fonctionnent au niveau 1 du modèle OSI. Ils ne sont plus très utilisés depuis que la performance des switchs, ou commutateurs, s'est améliorée, et que leur coût a baissé.

En effet, un switch est un élément permettant de connecter différents nœuds d'un même réseau, traditionnellement au niveau 2 « physique ». Sur un réseau local, il peut être suffisant de ne fonctionner qu'avec ces appareils.

Un switch comporte une **table de commutation** intégrée avec laquelle il compare les adresses IP des hôtes branchés. Il existe un très grand nombre de techniques de mise en tampon (*buffering*), de gestion de la priorité des paquets dans des queues de traitement (*queuing*), et aussi de gestion au niveau 3 (notamment pour le multicast avec la gestion des groupes par le protocole IGMP, ou *Internet Group Management Protocol*). Certains switchs comportent même des petits serveurs DHCP (Dynamic Host Configuration Protocol, qui est un protocole d'attribution d'adresses IP habituellement implanté sur les routeurs).

Un routeur s'occupe des opérations de **roulage** des paquets IP. Cet appareil fonctionne au « niveau 3 » du modèle OSI. Son fonctionnement simplifié consiste à recevoir un paquet, comparer son adresse à sa table de roulage en mémoire (ou en cache pour une gestion plus rapide), et le renvoyer sur le bon port de sortie.

De nos jours, la frontière devient de plus en plus floue entre switchs et routeurs avec l'apparition des switchs administrables (« *managed switches* »). Ces machines peuvent gérer des VLANs (Virtual LANs), administrer des groupes de multicast, attribuer des adresses IP dynamiques via DHCP. Ces fonctions n'étaient jusqu'ici implémentées que dans les routeurs.

L'apparition des réseaux audio sur IP, et plus globalement des réseaux à flux chrono-sensibles, a entraîné une migration de l'intelligence des réseaux depuis les

machines hôtes vers les machines dédiées au réseau que sont les switches et les routeurs. En effet, une optimisation des machines intermédiaires permet de soulager les hôtes d'une partie de la gestion réseau, et ainsi de leur donner la possibilité de développer d'autres fonctionnalités.

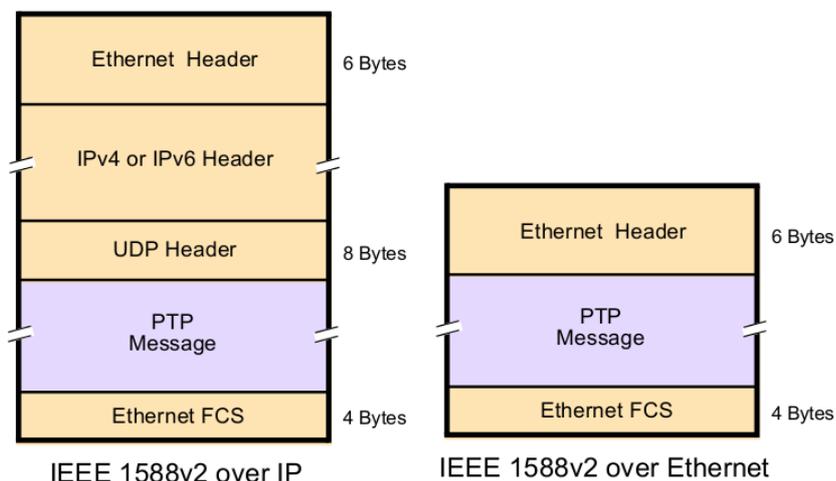
### 2.1.4.5. HORLOGE ET CADENCEMENT, SYNCHRONISATION

#### 2.1.4.5.1. PTP : FONCTIONNEMENT

Anciennement, la synchronisation des flux transitant sur les réseaux IP se faisait avec le **NTP** (*Network Time Protocol*). Cette technique ne permettait qu'une latence minimale allant de 1 à 2 ms sur un LAN, ou de 1 à 20 ms sur un WAN.

Or, des flux audio nécessitent la plus faible latence possible : du microphone aux enceintes de monitoring, plusieurs conversions et traitements peuvent être opérés tout en introduisant de la latence supplémentaire. Il est donc impératif que la latence imputable au réseau seul soit la plus basse possible.

Le protocole de gestion de la précision temporelle (PTP : *Precision Time Protocol*), existant actuellement en trois versions, permet d'ajouter des informations de synchronisation extrêmement précises encapsulées soit dans un paquet Ethernet (niveau 2) soit dans une trame IP (niveau 3).



*IP and Ethernet encapsulations for PTP messages.*

*Illustration des encapsulations possibles des messages PTP tirée d'un document explicatif rédigé par Albedo Telecom*

Le PTP, ou IEEE 1588, définit une hiérarchie d'horloges :

- **Horloge Grand Maître** (*Grand Master*), sur laquelle sont synchronisées toutes les horloges du réseau.
- **Horloge ordinaire**, souvent située en bout de chaîne IEEE 1588. Ce type d'horloge ne dispose que d'un seul port réseau.
- **Horloge frontière** (*Boundary Clock*) : ce type d'horloge possède deux ports réseau, qui peuvent être de type différent. Par exemple, un routeur Wifi.
- **Horloge de transparence** (*Transparent Clock*) : ce type d'horloge est dédié à la gestion de la remise en phase des flux en tenant compte des temps de trajet inhérents au réseau, et surtout au non-déterminisme des hôtes, qui ont une latence potentiellement variable de traitement des paquets.

Les appareils remplissant cette fonction doivent avoir plusieurs ports réseau. Deux types de transparence existent : la transparence de « **pair à pair** », c'est-à-dire entre deux hôtes, et la transparence « **de bout-en-bout** » du réseau.

La **sélection de l'horloge**, qui est la première phase de fonctionnement du protocole PTP, se fait de manière algorithmique selon des critères de qualité, de stabilité, et de fiabilité des horloges. L'algorithme s'appelle le *Best Master Clock Algorithm*, ou BMCA. Toutes les horloges présentes sur le réseau envoient des messages spéciaux, appelés « annonces », pour décrire le type d'horloge qu'elles peuvent fournir.

À titre indicatif, une horloge « atomique » a une précision de l'ordre de 15 ns, une horloge GPS, comme celle utilisée pour la synchronisation sur WAN de Ravenna, de l'ordre de 100 ns.

Après cette première phase, la phase de **synchronisation** à proprement parler peut commencer : des messages sont envoyés à intervalles de quelques secondes pour recalculer les délais entre les modules du réseau. Les hôtes envoient des requêtes de synchronisation auxquelles répond l'Horloge Grand Maître.

Plusieurs types de messages de synchronisation existent. Il y a des messages dont la priorité est haute, les plus critiques pour la synchronisation. D'autres ne nécessitent pas une prise en charge de priorité absolue.

On peut citer, pour les messages ayant une priorité haute (*Expedited Forwarding*), tous les messages d'annonce (*Announce*), de synchronisation (*Sync*), ainsi que toutes les requêtes de délai et les réponses qu'elles ont généré.

Pour les messages non-prioritaires transitant avec la qualité la plus basse (*Best Effort*), ce sont des messages de gestion et de signalement (comme, par exemple, un hôte signalant le type d'horloge qu'il génère).

Tous les messages prioritaires sont horodatés avec une précision de l'ordre de la nanoseconde. Les temps d'entrée sur l'interface réseau et leur passage à la couche 1 sont mesurés, et le paquet sortant, répondant à une requête de synchronisation, l'est également.

#### **2.1.4.5.2. RECOMMANDATIONS MATÉRIELLES**

Le PTP a besoin, pour fonctionner, d'un matériel compliant. Il doit pouvoir horodater de manière fiable les données transmises. D'autre part, les switchs reconnaissant ce protocole peuvent servir d'horloge de transparence ou d'horloge de frontière.

#### **2.1.4.5.3. COMPATIBILITÉ**

Le PTPv1 (version de 2003) et le PTPv2 (version de 2008) fonctionnent sur les mêmes bases, mais la version 2 n'est pas rétro-compatible avec la version 1 : les fonctionnalités « unicast » ont été rajoutées dans cette version pour une synchronisation entre deux appareils seulement.

Travailler dans un environnement réseau peut induire la nécessité d'un système de stockage directement accessible depuis tous les hôtes.

## **2.2. PROTOCOLES AOIP ACTUELS**

Une fois établies les bases techniques de l'audio sur IP, on comprend que de nombreux mécanismes et protocoles sont nécessaires pour pouvoir faire transiter de l'audio sur un médium tel qu'un réseau IP.

Plusieurs implémentations techniques d'audio sur IP existent actuellement. Nous n'allons présenter que les principales, ainsi que l'AVB, qui semble proposer des méthodes de transport multimédia très prometteuses pour l'avenir.

### 2.2.1. LIVEWIRE

**Livewire** a été un des premiers protocoles audio sur IP inventés, en 2003, par Telos de chez Axia. Révolution à l'époque, ce protocole permet des fonctionnalités que d'autres n'ont pas implémenté : l'alimentation de certains périphériques par Ethernet (Power over Ethernet, ou PoE), les consoles pouvant servir de routeurs et permettant donc le multicast.

Depuis, le protocole est devenu **Livewire+** en décembre 2013 : il a été retravaillé pour être interopérable avec l'AES67, notamment la synchronisation. Livewire utilisait une synchronisation propriétaire, mais s'est conformé au PTP v2<sup>4</sup>.

### 2.2.2. DANTE

Depuis 2006, Audinate, une entreprise Australienne, développe **Dante**. Ce protocole AoIP propriétaire a eu un grand succès. Il présentait, et présente encore, des solutions de routage et de latence jusqu'à présent inégalées.

Audinate développe également les logiciels pilotes Mac et Windows pour permettre une acquisition informatique, ainsi que les logiciels de contrôle et de routage.

De nombreux constructeurs sont partenaires de Dante, dans divers domaines allant de la sonorisation au broadcast, en passant par les consoles de mixage, les systèmes d'interphonie, de conférence, les préamplificateurs. On peut citer Yamaha, Allen and Heath, Nexo ou Soundcraft.

Ils détiennent une importante part du marché de l'audio sur IP, avec 200 constructeurs partenaires.<sup>5</sup>

### 2.2.3. RAVENNA

**Ravenna**, cavalièrement nommé d'après la ville où Dante Alighieri est enterré, a fait son apparition en 2010 à l'International Broadcasting Convention d'Amsterdam. Il était prévu que ce protocole soit une nouvelle révolution dans le domaine de l'audio sur IP.

---

<sup>4</sup> Site de Axia, présentant le livewire <http://www.telosalliance.com/Axia/Livewire-AoIP-Networking>

<sup>5</sup> La liste des constructeurs partenaires est consultable sur le site d'Audinate <https://www.audinate.com/about/news-activity/press/audinate-announces-200th-licensee>

En effet, il propose une solution « ouverte » par rapport à Dante, qui était initialement propriétaire, jusqu'à ce que l'AES67 ne fasse surface.

Ravenna a actuellement de nombreux partenaires, dont Archwave, Lawo, AETA, Genelec ou Merging Technologies.<sup>6</sup>

#### **2.2.4. AES67**

**AES67**, inventé par le groupe de travail X192 de l'AES (Audio Engineering Society) en décembre 2013, définit des normes d'interopérabilité pour les réseaux audio sur IP. Les professionnels développant l'audio sur IP qui ont sponsorisé ce groupe de recherche sont :

- Axia de la Telos Alliance, qui développe Livewire
- QSC Audio Products
- AVA Networks
- l'AES (Audio Engineering Society)
- ALC NetworX, qui a créé Ravenna.

C'est donc en analysant principalement les normes élaborées par ces développeurs et les besoins des constructeurs pour leurs réseaux AoIP que la version 2013 de l'AES67 a été rédigée.

#### **2.2.5. AVB**

L'**AVB** (*Audio Video Bridging*) n'est pas à proprement parler de l'« Audio sur IP », cette appellation regroupe deux profils d'utilisation distincts. C'est un protocole niveau 2 pour la version IEEE 1722, et niveau 3 pour IEEE 1733.

La question de l'inclure dans les protocoles présentés dans ce mémoire s'est posée. L'AES67 proposant une interopérabilité avec ce protocole, il a été décidé de le présenter tout de même.

---

<sup>6</sup> La liste des constructeurs implémentant le Ravenna est consultable sur le site d'ALCNetworx [16] <http://ravenna.alcnetworx.com/partners/partner-companies.html>

L'AVB est un ensemble de normes et de recommandations définies par l'IEEE, pour l'audio et la vidéo. On le trouve dans l'industrie automobile haut de gamme, où des petits réseaux audio-vidéo peuvent être utilisés à l'intérieur des véhicules.

L'IEEE 802.1 est le groupe de travail associé à ce nouveau protocole. Les normes principales dont il est question avec l'AVB sont :

- IEEE802.1BA : *Audio Video Bridging Systems*. Ce standard propose une norme pour les flux audio et vidéo sur un LAN ou un MAN.
- IEEE 802.1qav : FQTS ( *Forwarding and Queuing Enhancement for Time-Sensitive Streams* : Amélioration de la mise en queues et de la réexpédition pour les flux chrono-sensibles). Ce protocole définit des « *talkers* », qui envoient des données audio sur le réseau, et des « *listeners* » qui les « écoutent ».
- IEEE802.1AS : *Timing and synchronization for time-sensitive applications*. Ce protocole, aussi appelé gPTP ( *generalized Precision Time Packaging* ), propose une extension du PTP. Il faut, pour pouvoir le mettre en oeuvre, du matériel plus avancé que celui du PTP : l'horodatage des paquets est plus précis.
- IEEE 802.1qat : *Stream Reservation Protocol* (protocole de réservation des flux), permettant de créer des VLAN (réseaux locaux virtuels) et ainsi d'assurer une certaine qualité de service.

De plus, deux profils principaux d'AVB existent :

- 1722 : protocole « niveau 2 » de l'AVB.
- 1733 : extension du RTCP ( *Real-time Streaming Protocol* ) pour permettre aux flux RTP de traverser les réseaux AVB. C'est ce profil qui est donc prévu pour l'interopérabilité avec les réseaux audio sur IP.

On peut voir l'AVB comme un VLAN, où les flux audio (*streams* selon la terminologie) sont prioritaires sur tous les autres flux. C'est ainsi que ce protocole appelle des modifications matérielles importantes dans les différents éléments de réseau : les switches ont un CPU intégré pour permettre la réservation de la bande passante, les cartes réseau doivent pouvoir gérer non seulement le Gigabit Ethernet, mais la réservation de flux, l'horodatage précis, et la prise en charge prioritaire des paquets AVB.

L'interopérabilité de l'AVB et de l'AES67 serait théoriquement possible. Ce serait une possibilité intéressante : dans la spécification de l'AES67, on trouve une méthode de « *mapping* » d'une adresse IP multicast à une adresse Ethernet multicast, détaillée dans la RFC1112 paragraphe 6.4 [3] .

« Le protocole Ethernet permet de créer des groupes locaux, car il autorise une adresse multicast dans le champ « destinataire » de ses trames. Pour pouvoir utiliser le multicast IP sur cette couche, il faut définir une procédure de report des groupes d'adresses IP sur les adresses multicast Ethernet. »

« L'adresse groupée d'hôtes IP est reportée sur une adresse Ethernet multicast en copiant les 23 LSB (*Least Significant Bits*) de l'adresse IP dans les 23 LSB de l'adresse multicast Ethernet, qui est une adresse MAC égale à 01-00-05-00-00-00 (hexadécimal). D'ailleurs, puisqu'il y a 28 bits significatifs dans une adresse IP multicast, on peut en faire correspondre plusieurs à la même adresse multicast Ethernet. » [traduction]

Il y a donc possibilité de transmettre des multicasts de la norme AES67 vers des appareils AVB. Toutefois, il reste un besoin matériel de l'AVB très spécifique qui ne permet pas encore cette compliance.

Sur le site d'Audinate, dans les questions fréquemment posées par les utilisateurs, on peut apprendre que Dante pourrait implémenter l'AVB lorsque cette technologie sera plus couramment implémentée et testée.

## **2.3. COMPARAISON DES PROTOCOLES : INTEROPÉRABILITÉ**

### **2.3.1. CONFIGURATION DU RÉSEAU**

#### **2.3.1.1. DÉCOUVERTE ET ADRESSAGE**

L'étape de configuration d'un réseau audio sur IP est le premier pas vers son fonctionnement. Pour cela, il faut énumérer les différentes machines sur un réseau, leur attribuer une adresse IP unique, et configurer des groupes multicast pour permettre ce type de trafic. Il faut aussi énumérer précisément le type de service que chaque hôte peut fournir.

De manière habituelle, un réseau IP peut se configurer via **DHCP** (*Dynamic Host Control Protocol*) ou **DNS** (*Domain Name Server*). De façon simplifiée, pour le DHCP, chaque hôte du réseau envoie une requête DHCP en broadcast sur le port 67 lors de sa connexion. Si un serveur DHCP est présent, il leur fera une proposition d'adresse IP disponible via le port 68, et attendra la confirmation du client. Le serveur

confirmera alors l'adresse IP du client, et fournira aussi les indications de masque de sous-réseau, de passerelle, et de nom de l'hôte.

Le serveur **DNS**, lui, est une base de données hiérarchique : les hôtes deviennent des « domaines », auxquels sont attribués des noms. Internet est structuré ainsi : les noms séparés par des points sont les « noms de domaine », ce qui présente un avantage pour accéder aux ressources du point de vue de l'utilisateur final.

Toutefois, sur des petits réseaux AoIP, ces protocoles ne sont pas forcément implémentés : ils requièrent des serveurs spécialisés dans ces opérations, avec une gestion régulière de la base de données. D'autres techniques permettent des services de découverte des hôtes et de configuration réseau pour des petits LANs. L'AES67 n'a pas émis d'obligation en cette matière, mais a consacré une annexe complète à la comparaison des différents protocoles utilisés pour la découverte des appareils et l'attribution des adresses IP.

Le protocole **Zeroconf**, pour « *zero configuration networking* », est un service de découverte et d'adressage automatique des machines d'un réseau. Il est utilisé par Ravenna. Il est très proche du protocole **Bonjour**, inventé par Apple. Ces deux protocoles ne requièrent pas de serveurs spécialisés, et permettent la détection de tous les hôtes ainsi que leurs caractéristiques.

#### Quelques techniques de découverte :

MTDiscovery (*Merging Technologies Discovery*), basé sur le protocole **Bonjour** développé par Apple, est une application de découverte de réseau. Développée par Merging Technologies pour permettre de lister les appareils du réseau, elle utilise zeroconf et DNS-SD (*Dynamic Name System – Service Discovery*) pour fonctionner.

La version pour Mac de la *Dante Virtual Soundcard* installe également le protocole Bonjour. Par contre, la version Windows utilise une version propriétaire, le *Dante Discovery*. Les données de contrôle sont ensuite régulièrement mises à jour via le Dante ConMon, abréviation pour « contrôle et monitoring ». Sans ce petit logiciel, il peut y avoir des dysfonctionnements de détection d'horloge ou d'hôtes.

Il existe d'autres techniques possibles de configuration du réseau.

Livewire utilise une méthode de configuration propriétaire : *Axia Discovery Protocol*. Pour que tous les hôtes aient un accès aux flux, une adresse multicast dédiée aux données de contrôle est utilisée par tous les appareils. Une fois le réseau configuré, chaque machine communique des informations sur les flux qu'elle a établis, et enregistre les informations de toutes les autres machines communiquant.

L'AES67 renseigne d'autres techniques possibles pour la découverte du réseau et des services dans l'annexe E, comme le protocole SAP (*Session Announcement Protocol*) déjà utilisé dans les applications de voix sur IP (VoIP). Ce protocole annonce les caractéristiques nécessaires à l'établissement d'une session, et est donc en lien avec le SDP (*Session Description Protocol*) qui crée les flux audio.

### **2.3.1.2. ROUTAGE SUR LES RÉSEAUX AOIP**

La plupart des réseaux audio sur IP comportent des fonctionnalités de routage virtuel applicatif, accessible via un navigateur.

Ainsi, depuis un ordinateur, on peut paramétrer tout le routing des flux audio via une application. Celle-ci peut avoir un format de grille comme le *Dante Controller* dans l'onglet « Device view », ou l'application de routage développée par MOTU pour contrôler ses interfaces AVB. Elle peut également avoir un format liste, montrant les machines avec leurs entrées et sorties disponibles, comme le *Ravenna Discovery* de Merging Technologies ou *PathfinderPC* d'Axia : toutes les entrées et toutes les sorties disponibles sont répertoriées, et il suffit de cocher les cases souhaitées pour établir la connexion voulue.

Le système « grille » est le plus ergonomique : on peut voir d'un coup d'œil les routages audio sur le réseau.

Le *Dante Controller* permet, de plus, d'autres fonctionnalités de gestion de la réception et de la transmission ; il permet aussi d'importer ou de sauvegarder des presets, et de gérer précisément tout appareil du réseau : l'encodage, la fréquence d'échantillonnage, et le contrôle de la latence.

**PathfinderPC** est l'application de routage de Livewire. C'est un logiciel fonctionnant comme serveur, avec toutes les machines Livewire en clients. Des options intéressantes de routing prises en charge sont les flux vidéo, les données de contrôle des machines et de GPIO (*General Purpose Input/Output*, ou entrée-sortie

programmable). On peut assigner des fonctions diverses à ce dernier, telles que l'allumage du rouge d'enregistrement.

Dante a annoncé un logiciel, **Dante Via**, qui permet de créer un réseau local ou l'on peut accéder à toutes les cartes son et appareils audio branchés sur les différents ordinateurs. C'est une extension puissante du Dante Virtual Soundcard. Sa performance dépendra des caractéristiques des ordinateurs.

Merging Technologies, implémentant le Ravenna, propose également un logiciel de routage dédié, destiné à fonctionner avec ses machines, mais pouvant assurer une interopérabilité avec d'autres hôtes AoIP par le biais de l'AES67 : le **Ravenna Easy Connect**.

### 2.3.2. SYNCHRONISATION

La synchronisation des différents appareils audio est effectuée de plusieurs façons sur les réseaux Audio sur IP. Les recommandations d'harmonisation de l'AES67 ont généré une entente sur l'utilisation du STP pour les flux, synchronisé sur le PTPv2, ou IEEE1588-2008. En effet, ce sont ces éléments qui sont les plus critiques dans la transmission au niveau de l'interopérabilité. Dante fonctionnait avec le protocole PTPv1, et Livewire, avant de devenir Livewire+, possédait un système de synchronisation propriétaire. Les deux protocoles ont été ajustés pour garantir l'interopérabilité.

Il y a aussi, en termes d'horloge, le 803.1AS pour l'AVB. Ce système de synchronisation est beaucoup plus précis que le PTP. Toutefois, l'annexe D de l'AES67 ouvre vers le « *bridging* » des horloges PTPv2, utilisée dans la norme, et l'IEEE802.1AS, dans la perspective d'une possibilité de les rendre interopérables.

La synchronisation d'un réseau étendu doit tenir compte des temps de transmission des paquets sur le réseau, ainsi que de leur acquisition. De plus, si le réseau se trouve réparti en divers endroits du globe, le système GPS permet une synchronisation intervenant de l'extérieur.

#### Acquisition informatique des flux audio

L'intérêt de pouvoir travailler avec de l'audio transitant sur un réseau, c'est de le rendre très facilement accessible depuis les ordinateurs. Des configurations matérielles et logicielles doivent être respectées pour pouvoir rendre cette prise en charge possible.

## 2.3.3. GESTION SOUS MAC ET WINDOWS

### 2.3.3.1. PRISE EN CHARGE LOGICIELLE

La différence entre des standards techniques ouverts et des standards commerciaux réside dans l'accessibilité de ce standard (AES67 se veut ouvert, ainsi que Ravenna. Dante est une technologie fermée soumise à un Non-Disclosure Agreement).

#### Core Audio

Core Audio, sous Mac, est responsable de la gestion du son à partir d'OS X. À proprement parler, c'est une interface de développement intégrée à l'OS, comportant les outils nécessaires à l'utilisation des périphériques, la récupération et le routage des flux audio, de l'enregistrement ou de la lecture.

Théoriquement, les développeurs d'Apple avaient accès à des outils de développement dédiés à l'AVB sous OS X Frameworks depuis Mac OS X 10.8 Mountain Lion. Toutefois, la prise en charge des flux audio dans Core Audio n'est permise qu'à partir d'OS X 10.10 chez MOTU : on constate ici un décalage entre l'implémentation théorique d'une fonctionnalité et son développement effectif.

Ravenna Core Audio Driver\_: c'est le driver utilisé sur Mac à partir de la version OS X

La Dante Virtual Soundcard permet aussi une prise en charge par OS X, à partir de OS X 10.7.5, et également OS X 10.8.5, 10.9.5 et 10.10. L'architecture Intel est la seule fonctionnant, les anciens Mac fonctionnant sur PowerPC ne pourront pas utiliser la Dante Virtual Soundcard.

#### Sous Windows : ASIO

Depuis Windows Vista, la gestion logicielle du son a été refondue sous Windows. En effet, initialement développé par Steinberg, ASIO (*Audio Stream Input/Output*) est un des drivers audio les plus utilisés sur PC de nos jours. Il est souvent optimisé par chaque constructeur pour correspondre aux spécifications du matériel et assurer une prise en charge plus rapide.

Dans le cadre de l'audio sur IP, les pilotes développés par les divers constructeurs permettent de récupérer les flux audio sortant de la carte réseau, et ainsi l'utiliser comme une carte son à de nombreuses entrées.

Ravenna étant un standard ouvert, ils n'ont pas développé, au contraire de Dante, leurs propres logiciels. Nous allons examiner le driver ASIO de Merging Technologies.

Le **Merging Ravenna ASIO Driver** permet d'utiliser Ravenna sur les PC à partir de Windows 7 Professionnel 32 ou 64 bits.

Dante Network Controller\_:

Ce logiciel propriétaire nécessite une installation préalable du Dante Controller, le logiciel de routage audio. Il peut lui-même être installé un système Windows 7 SP1, 8, ou 8.1. Il fonctionne également sur Windows Server 2008 R2, ou 2012 R2.

### **2.3.3.2. CONFIGURATION MATÉRIELLE MINIMALE**

En matière de préconisation matérielle, la plupart des protocoles AoIP développés sont sensés fonctionner sur les installations réseau existantes. Dans les faits, le développement matériel fait par les constructeurs se trouve optimisé pour certaines solutions, et n'a pas été forcément testé pour tout le matériel du marché.

Ainsi, sur le site d'ALCNetworX (Ravenna), on peut lire que n'importe quel switch fonctionnera, mais sur le site de Merging Technologies, qui a implémenté Ravenna dans ses interfaces de conversion Horus et Hapi, un type de switch précis est recommandé : Dell Powerconnect. Ce switch est entièrement paramétrable, permet une gestion avancée de la QoS, et comprend un serveur DHCP. Merging a également prévu une configuration du switch téléchargeable sur son site, déjà préparée pour plus de facilité.

Merging Technologies a également créé une page de configurations matérielles d'ordinateur, régulièrement mise à jour. Ces configurations spécifient les possibilités techniques en fonction du processeur et du chipset.

AVB :

Contrairement à ses collègues, l'AVB induit une utilisation de hardware extrêmement spécifique et avancé. Les normes concernant le temps réel, la qualité, et la réservation de flux par VLAN demandent des switchs et des interfaces réseau capables de gérer matériellement ces contraintes, et les firmwares adaptés : un pilote générique ne sera pas à même de prendre en charge une utilisation aussi poussée des buffers, des queues Rx et Tx, avec une latence admissible en audio.

Il existe assez peu de switchs sur le marché capables de comprendre et gérer les paquets AVB. Ils comportent des développements matériels spécifiques : Par exemple, XMOS <sup>7</sup>, une entreprise spécialisée dans le développement de processeurs DSP/ multicœurs capables d'hyperthreading. Ils peuvent gérer les contraintes temporelles de synchronisation et de priorité des flux utilisés pour l'AVB.

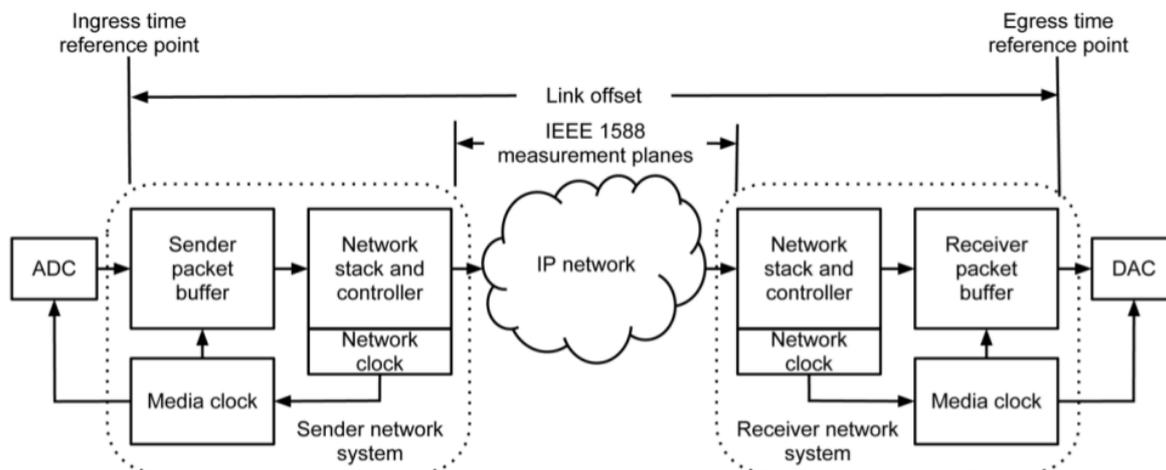
#### **2.3.4. LATENCE**

La latence sur un réseau AVB est extrêmement faible. Selon la taille du réseau et le nombre de switchs à traverser, elle sera au maximum de 2 ms. En effet, le nombre maximal de sauts pour un protocole 1722 est de sept, et la vitesse de transmission d'un port d'entrée à un port de sortie d'un switch est de l'ordre de 150 µs, ce qui n'est possible que dans un switch administré.

Pour ce qui est de la latence sur un réseau AoIP, l'AES67 normalise la latence maximale à 10 ms de point à point sur le réseau. Une transmission « point à point » est considérée par l'AES67 comme étant le temps de trajet sur le réseau, entre l'envoyeur et le récepteur, y compris après le temps de mise en buffer à la réception.

---

<sup>7</sup> Pour en savoir plus sur le développement d'une interface pour l'AVB, voir le site de XMOS : <http://www.xmos.com/products/reference-designs/avb-lc>



**Figure 1 - Example network illustrating link offset and ingress and egress reference points**

*Image tirée du document normatif de l'AES67, illustrant le «link offset», ou décalage réseau, pris comme référence pour la normation de la latence.*

La latence dépend aussi du « temps de paquet » ou « *packet time* ». En effet, plus la durée de l'audio contenu dans un paquet est faible, plus vite celui-ci sera traité. En contrepartie, un temps de paquet très court provoquera l'introduction de bits de bourrage, et donc de gaspillage de la bande passante.

La mémoire tampon (*buffer*), assignée par le pilote de la carte réseau, est un troisième paramètre qui peut être facteur de latence. En effet, plus un buffer est court, plus l'audio sera rapidement pris en charge par l'application. En contrepartie, il sera plus sensible au « jitter », et donc à la perte de paquets, qui se manifeste par des craquements audibles.

La taille recommandée pour un buffer est la taille minimale permettant d'éviter les conséquences du jitter tout en restant acceptable concernant la latence. Merging Technologies recommande une taille de buffers de 256 échantillons dans le manuel du Ravenna Core Audio.<sup>8</sup>

<sup>8</sup> Manuel du Ravenna Core Audio, page 11  
[http://www.merging.com/uploads/assets/Installers/Pyramix\\_8\\_1/Ravenna\\_ASIO\\_and\\_CoreAudio\\_Guide.pdf](http://www.merging.com/uploads/assets/Installers/Pyramix_8_1/Ravenna_ASIO_and_CoreAudio_Guide.pdf)

### 2.3.5. GESTION DES FLUX

#### DiffServ

Pour les réseaux opérant sous Ravenna ou AES/67, la qualité de service est assurée avec DiffServ. Les valeurs des champs DSCP (*DiffServ Coding Point*) ont été définies, dans l'AES67, en trois rubriques :

- Expedited Forwarding, ou « Service Premium », pour l'horloge du PTP
- Assured Forwarding avec la sous-classe « AF41 » pour une prise en charge prioritaire des paquets RTP
- Best-Effort pour le trafic annexe, comme ceux du contrôle et du routage.

#### Encodage, « payload »

L'encodage obligatoirement implémenté pour assurer l'interopérabilité des réseaux AoIP est au minimum le **L16** et le **L24**, respectivement PCM linéaire à 16 et à 24 bits, avec une fréquence d'échantillonnage de 48 kHz. Ces formats sont détaillés dans les RFC 3551 (rubrique 4.5.11) et 3190 (rubrique 4).

Il est également suggéré dans le document de l'AES67 de pouvoir fonctionner à 44.1 ou 96 kHz.

Dans les faits, les réseaux AoIP peuvent dépasser cette recommandation. Ravenna, par exemple, peut transporter de l'audio à 192 kHz, et également les formats DSD (*Direct Stream Digital*, appelé sigma-delta, suréchantillonné jusqu'à 64 fois la fréquence d'échantillonnage unitaire  $F=48$  kHz, et codé sur 1 bit) et DXD (*Digital Extreme Definition*, type d'encodage PCM à 24 bits/352,8 kHz) tant que le réseau est assez robuste pour assurer cette possibilité.

À l'inverse, Livewire+ peut également gérer des formats de média compressés, notamment le MPEG-4 sous ses différentes versions d'*Advanced Audio Coding* (AAC), afin de faciliter l'envoi de flux audio sur un WAN (*Wide Area Network*, qui dépasse donc le réseau local). Il peut être bienvenu dans le cadre de streaming sur des événements sportifs, de journalisme de terrain, de directs ou de duplex dont la latence doit être réduite au maximum, ou encore de radio numérique.

« Choose AAC-LD for delay-sensitive applications, AAC-HE and AAC-HEv2 for low bitrate requirements, standard MPEG AAC for best quality and resilience to packet loss at higher bitrates, MP3 and MP2 for legacy applications. »

Ravenna a prévu une possibilité de transport de données au format AES3, qui n'est pas encore formalisée dans l'AES67, et qui est utilisée dans l'implémentation des machines de Merging Technologies, Horus et Hapi.

#### Métadonnées :

Pour une interopérabilité, il faut qu'à tous les niveaux de désencapsulation des paquets, les hôtes communiquent similairement pour qu'à la fin l'audio soit reconnu comme tel. Certaines métadonnées associées aux flux, comme le SDP (*Session Description Protocol*) sont utilisées.

Ce protocole, décrit par la RFC 4566 de l'IETF, est l'ensemble de métadonnées nécessaires à la description d'une « session », c'est-à-dire l'envoi d'un flux. Il est donc utilisé par les hôtes pour la découverte des appareils réseau et pour la gestion du multicast (en relation avec l'IGMP, *Internet Group Management Protocol*).

Plusieurs informations sont données par le SDP, au format texte (UTF-8). Certaines sont obligatoires, et d'autres facultatives.

La description de la session est primordiale pour la bonne reconnaissance des divers protocoles utilisés par les machines, et l'établissement des flux entre elles. Les préconisations d'interopérabilité définissent les échanges minimaux de métadonnées pour une reconnaissance mutuelle des flux.

Si l'on voulait conclure de manière synthétique sur l'apport de l'AES67, on pourrait dire que tout ce qui concerne le transport et la synchronisation de l'audio a été normalisé, et permet une interopérabilité théorique des hôtes. On pourrait critiquer l'homogénéisation des protocoles AoIP, qui ainsi, perdent un peu de leur optimisation pour un domaine spécialisé.

Le contrôle et la découverte du réseau n'ont pas été normés afin de permettre des différences orientées des protocoles : Livewire reste optimal pour un réseau de type « broadcast », Dante est très polyvalent et bien conçu pour une utilisation

confortable, et Ravenna tente de se loger dans le secteur niche de l'audio de très haute qualité par sa gestion des formats DSD et DXD.

Le groupe X210 s'occupe d'une standardisation du contrôle et de la découverte, en travaillant sur l'*Open Control Architecture*, ou OCA. Le protocole *Open Control Protocol* ou OCP fonctionnerait par-dessus la couche TCP/IP, comme la plupart des protocoles de contrôle actuel. Les événements n'ayant pas trait au contrôle seraient en UDP, par rapport aux problèmes générés par le mode connecté évoqué plus haut.

L'OCP proposerait aussi notoirement une norme de sécurité utilisant la couche TLS (*Transport Layer Security*).

### **2.3.6. CONSIDÉRATIONS ANNEXES**

#### Sécurité

SNMP : petits réseaux (simple network management protocols)

Un réseau ayant une vocation de communication entre les différentes machines, avec des échanges de données ou des messages de contrôle, il est important de le protéger.

L'AES67 ne prévoit pas de norme spéciale concernant la sécurisation des flux audio ou de ceux de contrôle. Pourtant, le *Real-time Transport Protocol* permet l'encryption en TLS (*Transport Layer Security*). Malgré cela, le TLS rajoute en effet des calculs supplémentaires aux machines qui l'utilisent, et pourrait ajouter de la latence dans le traitement de l'audio.

#### Pare-feu et passage de l'audio

Les protocoles AoIP, afin de fonctionner convenablement, obligent à ouvrir des ports dans les pare-feu des ordinateurs connectés sur le réseau. Certains constructeurs proposent de désactiver le pare-feu pour un fonctionnement optimal, comme pour le Ravenna.

Le *Dante Controller* peut s'occuper de l'ouverture des ports du pare-feu intégré à Windows ou Mac OS X. En cas d'un pare-feu externe, il faut ouvrir manuellement les ports spécifiés dans la notice du *Dante Controller*.

La liste des ports couramment utilisés sur un réseau AoIP est donnée et normée dans l'AES67 pour les divers protocoles, tant pour les données audio que pour les flux de contrôle en TCP ou l'horloge PTP. C'est un paramètre primordial pour pouvoir utiliser les applications du réseau, envoyer et recevoir les flux.

Par exemple, le *Real-time Transport Protocol* utilise le port UDP 5004, et le protocole de contrôle associé, le RTCP, utilise le port UDP 5005. Les signaux d'horloge du PTP empruntent le port 319, en multicast, à l'adresse 224.0.1.129, également en UDP. Tous les flux de contrôle peuvent être assignés sur des ports de 1024 à 65535.

Sur un petit LAN, la sécurité ne devient problématique que s'il y a un accès à Internet prévu. En effet, un attaquant venant de l'extérieur doté d'un scanner de réseau comme nmap peut obtenir de nombreuses informations sur les numéros des ports ouverts, les programmes utilisant ces ports, la version de l'OS, et trouver une vulnérabilité dans un des hôtes.

Une fois dans le réseau, il est possible à un attaquant de forger des paquets IP en falsifiant la source ou la destination (attaque « Man in the Middle »). Il peut devenir dangereux, dans un réseau AoIP où la rapidité des flux est primordiale, d'avoir un accès à Internet, surtout dans un réseau important en taille et en influence (une radio, une chaîne de télévision...).

En résumé, les critères principaux de sécurisation sont **l'authentification** des hôtes émetteurs et récepteurs, le contrôle de **l'intégrité** et de la **validité** des en-têtes et du contenu des paquets. On peut y ajouter la **confidentialité** des flux.

Tous ces critères sont compromis dans le cadre de flux audio : en effet, l'UDP est un mode « non-connecté » entre les utilisateurs, ce qui pose des problèmes d'authenticité de la connexion. Pour ce qui est de la confidentialité, le chiffrement des flux, effectuée le plus souvent au niveau application, pourrait ajouter de la latence dans le traitement, ce qui n'est pas bienvenu dans une installation audio.

Il existe plusieurs solutions pour éviter que des attaquants de l'extérieur puissent pénétrer malgré une connexion à Internet : la DMZ (*Demilitarized Zone*), peut sérieusement retarder une attaque.

Elle est toutefois compliquée à mettre en œuvre pour une petite structure, et n'est utile que pour les services sensibles comme un serveur DNS (*Domain Name System*, servant à l'attribution d'une hiérarchisation des hôtes et de leur nommage),

un serveur de messagerie électronique, ou même un proxy pour l'accès à Internet. Il faut au moins un appareil servant de pare-feu disposant de 3 interfaces réseau : une pour le LAN, une pour la DMZ, et une en communication avec l'extérieur.

De plus, elle demande un contrôle et un monitoring régulier des serveurs situés à l'intérieur de la DMZ.

Toutefois, face à toutes ces méthodes de sécurité, il ne faut pas oublier que dès qu'on utilise un protocole dédié à la communication, il y aura forcément des solutions pour le détourner de manière malveillante, et que la sécurité est une éternelle escalade de patchs correctifs contre des vulnérabilités exploitables.

### **3. CONCLUSION ET OUVERTURE**

#### **3.1. QUESTION DE L'INTEROPÉRABILITÉ AVEC LES ANCIENNES MACHINES : LE « BRIDGING »**

Le changement paradigmatique apporté par les réseaux AoIP est énorme. On n'a plus de patchs analogiques avec des câbles physiques, les connexions se font toutes via des câbles Ethernet. On ne peut donc pas dire, par un simple coup d'œil aux branchements, quelles connexions sont établies, d'autant plus qu'un LAN peut permettre une interconnexion facile et souple en unicast ou en multicast.

Toutefois, les anciennes machines de studio peuvent avoir des sorties ADAT, MADI, S/PDIF, Toslink ou autres, et on peut vouloir, pour réduire le budget, les inclure dans le workflow. En effet, la plupart des appareils AoIP peuvent intégrer d'autres flux audionumériques au réseau via ces connectiques et protocoles habituels. Ainsi, il n'est pas nécessaire de renouveler d'emblée l'intégralité du parc matériel d'une installation, cela peut être fait progressivement.

#### **3.2. DÉVELOPPEMENTS INFORMATIQUES**

En informatique se développent actuellement les systèmes d'exploitation temps réel et les systèmes d'exploitation distribués.

Un système d'exploitation temps réel permet de rendre déterministe et prioritaire la prise en charge des flux audio, et ainsi rendre parfaitement transparent un réseau AoIP. La polyvalence de la plupart des systèmes d'exploitation actuels est basée sur leur capacité à passer d'un processus à l'autre, de changer de tâche : ils disposent d'un ordonnanceur qui décide d'assigner aux processus de l'ordinateur un ordre et une durée de calcul allouée en fonction de leur priorité.

Les systèmes temps réel permettent de prendre en charge les processus de manière déterministe et maîtrisée dans le temps. Ainsi, le jitter, responsable de pertes d'échantillons, et la latence s'en trouvent considérablement réduits.

Un exemple actuel de gestion temps réel des applications audio est le système MassCore de Merging Technologies. Ce logiciel, bien que n'étant pas à proprement parler un système d'exploitation complet, permet de dédier un ou plusieurs cœurs d'un processeur Intel Quad Core au traitement du signal, et de contourner les mécanismes habituels de la gestion de l'audio dans un système d'exploitation non-dédié. Ainsi, par une gestion directe du processeur, on ne risque plus une interruption de processus due à un programme à haute priorité, ou une latence de gestion induite par le traitement de l'audio par l'OS.

Les flux audio gérés sont ainsi très nombreux (jusqu'à 384 entrées et sorties simultanées à 44,1 ou 48 kHz, ou 96 I/O à 4 fois ces fréquences d'échantillonnage), alors que la latence de l'entrée à la sortie n'est que de 1,33 ms.

Le cloud computing est également une tendance de stockage et d'utilisation de processus dématérialisée. Des machines virtuelles dont on peut choisir la puissance sont stockées sur des serveurs, et peuvent être dédiées à des programmes applicatifs, ou à du stockage.

Chez AVID notamment, le *AVID Cloud Collaboration* prévoit de permettre à des artistes de collaborer « où qu'ils soient dans le monde, comme s'ils étaient dans le même studio ». Seul le stockage des sessions sera en réseau et potentiellement accessible en temps réel par plusieurs personnes.

A propos du Cloud, le traitement audio centralisé sur machine virtuelle ne sera peut-être pas pour tout de suite. Dante prévient, sur son site, des problèmes potentiels de l'audio sur VM (*Virtual Machine*) : l'accès et l'utilisation élevée de l'interface réseau sont entièrement gérées par le CPU. Les performances ne sont sans doute pas, aujourd'hui, suffisantes pour un traitement en temps réel et un retour avec peu de latence.

### 3.3. OUVERTURE

Ce mémoire peut avoir de nombreux développements et considérations d'un ordre plus philosophique que la simple communication de machines entre elles, aussi évoluée soit-elle. En effet, le développement technologique se fait à une vitesse telle que d'ici quelques années, des progrès matériels et logiciels auront rendu obsolète la majorité de la partie théorique présentée.

Durant ce tour d'horizon de l'audio sur IP, nous avons pu tirer quelques constats.

La première remarque intéressante est la latence temporelle entre le concept et la forme. En effet, les normes présentées, quelles qu'elles soient, ont toutes été conçues bien avant d'être effectivement implémentées par les constructeurs, puis adoptées par les utilisateurs. Bien que certaines normes telles que le protocole IP (dont la version 4 date de 1981, il y a maintenant 34 ans) aient traversé le temps avec assez peu de modifications (4 modifications dont la dernière en 2013 pour l'IPv6), nombreux sont les protocoles oubliés, peu implémentés car trop spécifiques ou trop spécialisés.

La surabondance des normes de communication inter-machine nous rappelle la diversité des langages humains : les plus complexes et les plus spécialisés finissent par évoluer ou disparaître au profit de langages plus efficaces ou moins obsolètes, à mesure que la culture ou les mœurs évoluent.

La démarche sous-jacente de l'interopérabilité est un combat contre cette obsolescence prématurée de la technologie, qui rend frileux à la fois les développeurs et les constructeurs – qui investissent beaucoup de temps et d'effort dans un produit qu'ils espèrent rentable – et les utilisateurs – qui voudraient pouvoir rentabiliser cet investissement, ce pari qu'ils ont fait en achetant un matériel professionnel spécialisé.

Ainsi, une économie naît et vit : il ne faut pas oublier que la technologie dans un domaine professionnel est un fruit et une source de travail pour de nombreuses personnes, et que toute une écologie de niche peut naître ou être détruite par une innovation technique.

On pourrait modéliser l'évolution technologique par une chaîne causale ayant des impacts successifs : pour que des machines deviennent interopérables, il faut qu'une norme commune soit établie. Une fois cette norme établie, il faut vérifier que le matériel soit compliant à la légère modification apportée, notamment le matériel non-dédié, censé être générique.

On peut à titre d'exemple citer les interfaces réseau : une modification au niveau programmation – analogie humaine d'un apprentissage – peut être faite plus facilement qu'un changement matériel, qui est une péremption fonctionnelle due à l'évolution de l'environnement. Avec M. Chouvel, nous avons constaté, après des tests d'interfaces de conversion analogique-numérique, que certains flux AVB nécessitaient une prise en charge matérielle très évoluée, que nous ne suspicions pas car la configuration au niveau logiciel répondait au cahier des charges.

D'une part, certains progrès techniques, qui ne sont pas encore appliqués au domaine de l'audio, peuvent faire porter l'imagination et l'expérimentation vers de nouveaux horizons.

Les RTOS (*Real-Time Operation System*) commencent déjà à se développer dans des domaines aussi variés que l'industrie, la défense ou l'aéronautique. Or, l'audio sur IP permet, de manière quasi-instantanée, de communiquer à travers un réseau. Nous n'avons parlé que des réseaux Ethernet et principalement sur medium cuivré, mais d'autres milieux de propagation plus rapides et plus fiables existent déjà, tels que la fibre optique. Il est donc naturel de s'orienter vers des systèmes d'exploitation dédiés à l'audio, optimisés pour recevoir des flux avec une prise en charge logicielle conçue pour cet usage – ce qui n'est pas du tout le cas dans les systèmes d'exploitation généralistes et polyvalents, qui coûtaient moins cher à développer, et étaient plus rentables que des systèmes dédiés à l'heure où l'ordinateur personnel se lançait dans l'aventure qui le mène aujourd'hui aux limites de ses capacités, en termes d'intégration des composants, de taille des gravures, de rapidité des bus de transmission internes, de mémoire vive ou de supports inscriptibles...

Cette “crise matérielle” va devoir déboucher sur une mutation de fonctionnement informatique, qui va peut être également mener à une mutualisation des ressources de calcul, de stockage, de puissance. Il existe des “systèmes d'exploitation répartis”, encore très expérimentaux, où des groupes d'ordinateurs peuvent fonctionner et organiser une macro-structure informatique. Cette organisation existe déjà dans la recherche, où on peut allouer un peu de temps et d'énergie de son processeur à un calcul réparti en réseau. Les pirates informatiques, par une utilisation plus « négative » de la possibilité technologique, détournent une grande quantité de machines pour les utiliser à mauvais escient, comme par exemple pour une attaque DDOS (*Distributed Denial of Service*) figurant parmi les techniques les plus connues et redoutées.

On peut d'ailleurs observer l'émergence, ces dernières années, de délocalisation des stockages et services informatiques sur de puissants groupes de serveurs, appelés “*datacenters*”. Dématérialiser et délocaliser les données était déjà une tendance sur Internet, où, au fur et à mesure que le maillage s'est resserré et les débits ont augmenté, il est devenu possible d'accéder partout et de manière quasi instantanée à des serveurs situés à l'autre bout du globe.

Cette émergence récente du *cloud* et des services de streaming multimédia, de réseau social, de délocalisation des services informatiques des entreprises entraîne également de vertigineuses perspectives si on peut y voir un terrain praticable pour l'audio.

Au final, ce mémoire est une question plus qu'une réponse à la technologie. Peut-on parler de *techno -philosophie*, réflexion née de et par notre propre création, tels des pygmalions émerveillés par notre statue mouvante ? Nous avons créé ces machines à notre image, tel un miroir de notre réflexion. Nous les avons fait évoluer vers des aspirations humaines que nous ne saurons peut être jamais nous-mêmes atteindre : l'affranchissement du temps, de l'espace, de la matière ?

PARTIE PRATIQUE : PROJET  
D'INGÉNIERIE D'UN STUDIO POLYVALENT  
EN AUDIO SUR IP

## RÉSUMÉ

Le studio de la Beaudelie est, dans l'état actuel, pourvu d'une **régie** (appelée « régie » dans ce cahier des charges) et d'une **cabine de prise de son** (appelée « studio »). Il sert principalement à des prises de son de musique acoustique et électroacoustique.

Un agrandissement physique du studio est prévu : l'adjonction d'une **grange** augmentera la capacité musicale du lieu, et en sera une extension fonctionnelle. Cette nouvelle salle pourra servir à la **prise de son** ou à des **concerts**, voire **des projections**.

Pour garantir la souplesse de configuration, nous allons étudier la possibilité de mettre en oeuvre un réseau Audio sur IP, ses avantages et les limitations techniques que l'on peut rencontrer.

**Mots clés** : AoIP, étude d'ingénierie, audio sur IP, studio

## ABSTRACT

The "La Beaudelie" Studio currently consists in a **control room** and a **studio room**. It is mainly used for recording electroacoustic or acoustic music.

A physical extension of the studio is under way : a greater room, the **barn**, will extend the facility, both spatially and functionally. This new room will serve for sound recording purposes, for housing concerts, and for film projections.

In order to ensure a greater flexibility, we will study the feasibility of an **audio over IP** network system, its advantages and the technical limitations that can be encountered.

**Keywords** : AoIP, engineering, audio over IP, studio

## 1. PRÉSENTATION GÉNÉRALE DU PROBLÈME

Le studio de la Beudelie est dans l'état actuel pourvu d'une **régie** (appelée « régie » dans ce cahier des charges) et d'une **cabine de prise de son** (appelée « studio »).

Actuellement, il sert principalement à des prises de son de musiques acoustique et électroacoustique.

### 1.1. FINALITÉ DU PROJET

Un agrandissement physique du studio est prévu : l'adjonction d'une **grange** augmentera la capacité musicale du lieu, et en sera une extension fonctionnelle. Cette nouvelle salle pourra servir à la **prise de son** ou à des **concerts**, voire **des projections**.

D'autre part, il est prévu que d'autres salles puissent assurer la fonction de **régie** initialement tenue par la régie du studio. Il y en aurait potentiellement trois :

- Le **studio**, qui sert actuellement à la prise de son, et où se trouvent des écouteurs KEF concerto.
- Une salle au 2ème étage du bâtiment, appelée « **bureau** »
- La **régie** actuelle, attenante au studio, et qui a une vue sur celui-ci via une fenêtre prévue à cet effet.

## **1.2. CONTEXTE**

Techniquement, le studio se présente dans sa forme actuelle, sous un jour assez traditionnel :

### **1.2.1. PRÉSENTATION DES INSTALLATIONS**

#### **1.2.1.1. STUDIO**

Il y a, dans le studio, un système d'écoute (pouvant également assurer la fonction de talkback) constitué d'un amplificateur Rotel RA-162 et d'une paire d'enceintes KEF Concerto SP 1006 8 Ohm 30W.

Un câble multipaire permet de brancher jusqu'à 20 microphones. Il est branché sur la MOTU 896 mk2 du studio, ainsi que sur la MOTU Traveler.

Le parc de microphones disponible est consultable dans la liste exhaustive de matériel fournie en annexe 1.

#### **1.2.1.2. RÉGIE**

La régie, accolée au studio, compte plusieurs cartes son sur lesquelles est branché le multipaire.

Cartes son disponibles :

- MOTU 896 MKII : branchée à l'ordinateur en Firewire,
- MOTU Traveler, branchée en ADAT sur la 896,
- Console : M-Audio Project Mix, branchée en Firewire.

Ecoutes : M-Audio BX8A.

Ordinateur : **iMac** datant de 2008, 20 pouces, sous Mac OS X 10.8

- 2,66 GHz Intel Core 2 Duo
- 4GB 800MHz DDR2 SDRAM
- disque de 320GB en Serial ATA
- 3 ports USB 2.0
- carte réseau Gigabit Ethernet

L'ordinateur compte plusieurs logiciels d'édition installé, les principaux étant Pro Tools 10 et Digital Performer 7.

### **1.2.1.3. GRANGE**

Enceintes:

- 6 enceintes « Cabasse » Antigua NT 3/2
- 2 Focal « chorus » 716V colonne
- Subwoofer BX 10S de M-Audio

Amplificateurs :

- Un Pioneer VSX 923, pourvu essentiellement d'entrées HDMI, d'une entrée LAN et d'entrées S/PDIF. Il dispose de 6 sorties pour enceinte et 2 sorties dédiées aux subwoofers.
- Le T.Amp S-75MK2 permet d'accroître le nombre d'enceintes utilisées en diffusion multicanal.

D'autres appareils sont présents dans la grange, on peut en trouver la liste sur l'annexe 1, mais ils ne seront pas intégrés dans cette démarche d'ingénierie pour le moment.

### **1.2.1.4. BUREAU**

Le bureau équipé d'un iMac aux caractéristiques similaires à celui de la régie, et d'une paire d'enceintes Yamaha BS50M.

### 1.2.2. SUITES PRÉVUES

Une fois l'installation d'un réseau AoIP effectuée, une **étude acoustique** pourrait être menée pour optimiser le système de sonorisation de la salle. En effet, au vu des différentes configurations de concert possible, il peut être intéressant de voir les divers placements possibles des enceintes pour un rendu sonore optimal dans chaque cas de figure proposé.

La première installation d'un réseau AoIP peut également trouver des suites et développements matériels pour une éventuelle **expansion**. Il faut réfléchir à cette option dans la conception actuelle du projet.

### 1.2.3. NATURE DES PRESTATIONS DEMANDÉES

Dans ce projet, la prestation demandée est une aide à la conception d'un studio plus polyvalent, et fonctionnellement souple. La **prise de son** doit pouvoir se faire indifféremment dans la grange ou dans la cabine actuelle, et la **régie** doit pouvoir se faire dans la régie actuelle, le studio, la grange, et le bureau du 2<sup>ème</sup> étage.

### 1.2.4. PARTIES CONCERNÉES PAR LE DÉROULEMENT DU PROJET

Demandeur et utilisateur final : Jean-Marc CHOUVEL

Aide à la conception : Marie-Angélique MENNECIER

Directeur interne : Jean ROUCHOUSE

Consultant d'ingénierie : Franck JOUANNY

## 1.3. ÉNONCÉ DU BESOIN (FINALITÉS DU PRODUIT)

La **souplesse et la simplicité d'utilisation** du studio sont les principaux besoins demandés en même temps que sa pluralité : on tâchera de trouver une solution la plus « plug-and-play » possible, avec des configurations et routages pré-enregistrés, afin d'avoir le minimum de configuration à effectuer.

### 1.3.1. ACQUISITION DE L'AUDIO

Pour l'acquisition de l'audio, plusieurs solutions techniques peuvent être envisagées.

Choisir des modules de conversion, de type « stage box » ou « carte son sur IP », capables de transporter le signal directement via un protocole IP serait la meilleure option. Les interfaces de type stage box sont souvent destinées à fonctionner avec une console. Un grand nombre de consoles proposent des solutions propriétaires de « stage box ».

Il y a également possibilité d'utiliser l'infrastructure existante en numérisant le flux grâce à une **interface de conversion** destinée à transformer l'audio en signal numérique porté par le protocole IP. L'Horus ou Hapi de Merging Technologies utilisant le Ravenna, ou d'autres cartes telles que la MOTU 8M en AVB sont capables d'intégrer un signal ADAT, AES3 ou MADI, et de le paqueter sur IP.

Plusieurs boîtiers de scène ou appareils de conversion existent sur le marché, il conviendra d'étudier les possibilités techniques offertes avec attention.

Il sera nécessaire d'avoir une **horloge** de qualité, afin d'avoir le minimum de jitter et de latence possible. En situation de prise de son, la latence est un problème très inconfortable pour les musiciens, et en situation de concert, il est très problématique également, voire handicapant, d'avoir de la latence dans les retours.

### 1.3.2. MISE EN RÉSEAU

Un **serveur de fichiers** accessible depuis toutes les régies, pour centraliser la prise de son, offrirait une facilité pour l'unification des sauvegardes. La question se pose alors : ce serveur servira-t-il de stockage en temps réel, ou en temps différé ? En temps réel, il faudra une certaine rapidité de la machine.

Par rapport au système de stockage, il pourrait être utile d'avoir un serveur NAS (Network Attached Storage) afin que des appareils tant Mac que Windows puissent indifféremment y accéder sans problème de format de fichier. Le NAS n'a pas non plus de contrainte de mise sous tension permanente, ni besoin d'un logiciel dédié pour fonctionner. Une configuration en **RAID** (niveau à déterminer) peut être utile

pour une plus grande sécurité des sauvegardes, toutefois le « mirroring » des fichiers peut s'avérer coûteux en supports de stockage, ce qui fera préférer le Jbod (Just A Bunch of Disks), moins sécurisé ou rapide.

Un **switch** ou un **routeur** capable de transporter et distribuer les flux audio adressés au bon endroit, et de gérer les très hauts débits (Gigabit Ethernet) ainsi que la QoS (Quality of Service) sera essentiel.

Il y aura nécessité de réfléchir à un **réseau d'ordre**, tant audio que vidéo. En effet, à cause de l'épaisseur et de la solidité des murs, il est très compliqué d'y percer des trous, et un système de caméras IP permettrait d'avoir un visuel sur ce qui se passe dans la grange (et éventuellement dans la cabine de prise de son actuelle si ce n'est pas la régie adjacente, disposant d'une visibilité sur la cabine, qui est utilisée).

Un système de **retour vidéo** sera donc envisagé entre la grange et les différentes régies, avec possibilité de captation. Une qualité « HD » serait requise à des fins de réutilisation de l'image.

Placer un couple stéréophonique à demeure, mais sur un bras articulé pour permettre de le placer à un point optimal de la salle, est également envisagé.

### **1.3.3. SÉCURITÉ DES BIENS ET DES PERSONNES**

Concernant la **sécurité des personnes et des biens**, il conviendra de s'assurer de la bonne **ventilation** des locaux, que l'**installation électrique** est de qualité, et prévoir de raccorder sur onduleur les ordinateurs, le NAS et les machines.

### **1.3.4. EXTENSIONS FONCTIONNELLES SOUHAITÉES**

Plusieurs configurations sont possibles pour de la **prise de son** dans la grange :

#### **Orchestre de musiciens :**

- Placés côté « écran »
- Placés côté « escalier »

Dans ces deux configurations, la régie pour la prise de son pourra être celle de l'étage, la cabine de prise de son où se trouvent les KEF, ou encore la régie initiale.

**Pour une configuration « concert » :**

- Musiciens devant l'écran, public autour
- Musiciens côté escalier, public en face
- Configuration « face au Mana », dans la longueur de la salle

Dans ces 3 configurations, il sera nécessaire de pouvoir moduler le système de diffusion (stereo, 5.1, 7.1 ou 8 points).

Il est également bien de noter que des artistes « phonogenistes » voudront se placer face au système d'écoute.

Une utilisation « **auditorium de mixage image** » ou cinéma, de la grange serait à étudier également.

Ces besoins fonctionnels requièrent un nombre optimal de 32 entrées microphoniques, et la possibilité de travailler au moins à 24 bits / 48 kHz. Il faut penser aux besoins en retours casque : de quoi accommoder confortablement un ensemble de musiciens, au moins huit simultanément au cours d'un enregistrement.

## **1.4. ENVIRONNEMENT DU PRODUIT RECHERCHÉ**

### **1.4.1. LISTES EXHAUSTIVES DES ÉLÉMENTS (PERSONNES, ÉQUIPEMENTS, MATIÈRES...) ET CONTRAINTES (ENVIRONNEMENT)**

**On trouvera en annexe 1 de ce document une liste complète de tout le matériel présent à la Beudelie.**

#### **1.4.1.1. CONTRAINTES PHYSIQUES**

Le studio a des murs de schiste épais et difficiles à percer, générant également des problèmes de transmission de signaux électromagnétiques. L'implémentation d'un Wireless LAN s'en trouverait compliquée. Il faudrait mener une étude pour évaluer la possibilité d'un tel déploiement technique.

La Beudelie est un lieu éloigné des grandes villes : le studio doit être assez autonome et facilement dépannable, des solutions « de secours » doivent être suffisamment rapides à mettre en place. En particulier, on réfléchira aux différentes

possibilités de réutilisation du matériel existant en conjonction du matériel plus récent dévolu à la création d'un réseau AoIP, à la fois pour alléger le budget si nécessaire, mais également pour prévoir des solutions de secours.

#### **1.4.1.2. CARACTÉRISTIQUES POUR CHAQUE ÉLÉMENT DE L'ENVIRONNEMENT**

La **MOTU 896 mk2** accueille huit convertisseurs analogique-numérique pour des microphones, des instruments (haute impédance) ou des entrées ligne à 24 bits et 192 kHz. Elle possède des entrées ou sorties **ADAT**, 8 canaux, à 24 bits et 44.1 kHz ou 48 kHz, et une liaison **AES/EBU**, soit deux canaux à 24 bits et maximum 192 kHz. Il serait intéressant d'explorer ces fonctionnalités en tant que sorties possibles vers une interface les convertissant vers l'AoIP ou l'AVB, bien qu'elles brident potentiellement la fréquence d'échantillonnage à 96 kHz ou 48 kHz selon le nombre de pistes désirées. Toutefois, cette limitation technique tient encore compte du minimum qualitatif souhaité de 24 bits/48kHz.

La **MOTU Traveler** présente 4 entrées microphone ou ligne en connectique « combo » jack ou XLR, ainsi que 4 entrées ligne jack TRS symétrique ou asymétrique. Elle compte également les mêmes entrées et sorties en ADAT et la liaison AES/EBU que la MOTU 896 mk3, avec en plus, en ADAT, le choix d'utiliser l'interface avec 16 canaux à 24/48kHz, 8 canaux à 24 bits/96 kHz ou en TosLink (qui transporte deux canaux à 24 bits/96kHz maximum avec le protocole S/PDIF).

## 2. CADRE DE RÉPONSE

### 2.1. ÉLÉMENTS DE RÉSEAU

#### 2.1.1. CHOIX DES CÂBLES

Le choix des **câbles** est une problématique extrêmement importante, que l'installation soit analogique ou numérique. Les câbles réseau se répartissent en catégories selon la bande passante qu'ils doivent pouvoir passer. Il existe également des types de blindage assez variés.

Un câble Ethernet est constitué de 4 paires de câbles associées et torsadées ; c'est une des mesures possibles pour réduire la diaphonie. Il existe des câbles où chaque paire a une tresse, ou bien une feuille de blindage en aluminium. La gaine extérieure, contenant toutes les paires, peut également être blindée par une tresse. L'avantage du blindage est indiscutable concernant les problèmes éventuels d'interférences (diaphonie) et parasites externes et de diaphonie entre les câbles.

Les protocoles Ravenna, Dante, Livewire ou l'AES67 passent sur les réseaux Gigabit standards (d'après les constructeurs, certains cas de figure doivent être évités pour un fonctionnement optimal), ou même 100baseTX pour un nombre de pistes relativement faible.

Du câble Cat5e ou plus du Cat6, sont indiqués pour les transmissions Gigabit Ethernet.

Désignation	Appellation (ISO/IEC 11801)	Blindage global	Blindage par paire
Câble à paires torsadées non écrantée	U/UTP	Aucun	Aucun
Câble à paires torsadées avec écran général	F/UTP	Feuillard	Aucun
Câble à paires torsadées avec écran et tresse général	SF/UTP	Tresse. Feuillard	Aucun
Câble à paires torsadées écrantée par paire	U/FTP	Aucun	Feuillard
Câble à paires torsadées avec écran général et écrantée par paire	F/FTP	Feuillard	Feuillard
Câble à paires torsadées avec tresse général et écrantée par paire	S/FTP	Tresse	Feuillard
Câble à paires torsadées avec écran et tresse général et écrantée par paire	SF/FTP	Tresse. Feuillard	Feuillard

*Appellation des câbles à paires torsadées selon la norme ISO 11801*

*Image tirée de l'article de Mohammed Elliq et Alain Delhaise,*

*Conducteurs et Câbles en Audio (Partie 1)*

Les câbles choisis seront dévolus à une installation fixe. Par conséquent, on peut choisir le meilleur type de blindage possible, tout en sachant que le rayon de courbure de ces câbles s'en trouvera diminué, et leur prix, leur poids et leur rayon augmentés. Ainsi, on pourra être certain que les parasitages externes ou internes aux câbles entre les pistes sera minimisé.

On choisira donc des câbles de type Cat5e F/FTP ou S/FTP au moins, mais du Cat6 blindé est également possible.

### **2.1.2. DIAPHONIE**

Les réseaux câblés, lorsque les conducteurs sont proches les uns des autres, sont souvent soumis à des problèmes de *diaphonie*. En effet, la diaphonie est un parasitage qui peut être dû à une induction d'une piste proche, ou bien encore à une réflexion sur une terminaison ou si le câble est mal dimensionné en hyperfréquences.

Il existe plusieurs solutions pour remédier aux problèmes de diaphonie.

Concernant les phénomènes dûs à l'induction, un blindage ou une protection des câbles, qui n'est autre qu'une mise en œuvre du principe de Faraday et une isolation électromagnétique des conducteurs, peut suffire.

La torsion des conducteurs par paires symétriques permet aussi d'atténuer la diaphonie significativement. En effet, sur des paires symétriques, la réception se fait sur un amplificateur opérationnel en mode commun et le bruit est alors théoriquement annulé.

Toutefois, si une des deux lignes symétriques de la paire présente plus de bruit que l'autre, le bruit ne pourra pas totalement être annulé à l'arrivée ; la torsion des câbles d'une même paire permet donc d'homogénéiser le bruit d'un parasite externe sur les deux lignes, et par conséquent de mieux l'annuler à la réception.

### 2.1.3. CHOIX DES SWITCHS

Il est nécessaire de faire attention à la topologie du réseau. En cas de boucles, les paquets peuvent transiter plus longtemps que nécessaire et engorger le réseau. Pour ce faire, une **topologie en étoile** est la plus simple à mettre en œuvre, et recommandée pour les réseaux AoIP, mais une topologie en arbre bien aménagée, ou une topologie mixte, peuvent également être envisagées.

Dans une topologie en étoile, tous les appareils sont connectés à un switch central ; ainsi, en cas de panne d'un nœud du réseau, le reste continue à fonctionner. Le grand nombre de câbles qu'elle requiert, et sa dépendance totale sur le switch central, sont ses inconvénients principaux.

Il convient d'éviter les switches ou routeurs « Green Ethernet ». Cette technologie, détaillée dans la norme IEEE 802.3az, inventée et mise en œuvre par D-link, consiste à mettre en veille les ports du switch lorsque aucun périphérique actif n'est détecté.

En effet, une modification du signal d'inactivité (« Idle ») des nœuds du réseau permet une signalisation prioritaire de l'allumage ou extinction des ports d'un switch ou routeur, et ainsi de réduire la consommation énergétique. Le « temps d'allumage » d'un port à la détection de données y transitant est de 16,5  $\mu$ s sur une ligne 100-base-T.

Or, le protocole « *Precision Time Protocol* », version 1 pour Dante, ou version 2 pour Ravenna/AES67, peut avoir une précision de l'ordre de la nanoseconde. La précision de l'horloge via GPS proposée dans le cadre de l'AES67 est de l'ordre de plus ou moins 100 ns.

Le site de Yamaha Audio préconise donc de désactiver la fonction Energy Efficient Ethernet dont peuvent être dotés certains switchs pour éviter d'éventuels problèmes d'horloge et de synchronisation des flux.

Sur les sites de Audinate, créateurs de Dante, ou ALC NetworX qui ont développé Ravenna, il n'y a désormais plus de marque de switch recommandée<sup>9</sup>. Il est donc laissé à l'appréciation de l'installateur de choisir un switch laissant passer des flux préférentiellement en Gigabit, en fonction de son nombre de ports.

Il est de plus fortement recommandé sur ces sites de désactiver l'EEE (*Energy-efficient Ethernet*), d'avoir un switch capable de créer quatre files d'attente de QoS (voir partie Théorique) et Diffserv. Audinate recommande par ailleurs, si possible, des switchs configurables par l'utilisateur.

Sur la notice de l'Horus de Merging Technologies, il est toutefois spécifié que seuls certains switchs sont certifiés pour fonctionner avec leurs machines : les DellPowerConnect 2816 et ceux de la même famille, 2808, 2848...

Cette famille de switchs est également configurable et permet une QoS avec Diffserv et 4 files de priorité.

Chez Axia Audio de la Telos Alliance, inventeurs de Livewire, il existe encore cependant un switch spécifiquement recommandé pour un fonctionnement optimal du réseau, le xSwitch, ainsi que quelques switchs approuvés<sup>10</sup>, comme la série des Cisco Catalyst 2960-C et 3560-C, qui en plus de toutes les fonctionnalités de configuration et QoS, peuvent également fournir du Power over Ethernet, un prérequis propre à certains appareils d'Audio sur IP d'Axia Audio. Les switchs les plus compacts de cette série n'ont pas, non plus, de ventilateur intégré, source de bruit, et n'ont donc pas besoin d'être isolés acoustiquement d'une régie ou d'un lieu de prise de son<sup>11</sup>.

---

<sup>9</sup> <https://www.audinate.com/resources/networks-switches>

<sup>10</sup> Switchs approuvés par Axia pour Livewire : <http://www.telosalliance.com/Axia/What-Ethernet-Switches-has-Axia-Approved>

Pour le projet du studio à La Beaudelie, selon que le choix se porte sur des appareils communiquant en Ravenna ou en AVB, il conviendra donc de choisir un switch (voir plans de réseau proposés) Gigabit, ou choisir un ou plusieurs switches spécifiques à l'AVB.

## 2.2. AUDIO SUR IP

### 2.2.1. CHOIX DU MATÉRIEL DE CONVERSION ANALOGIQUE-NUMÉRIQUE

Il est possible d'utiliser les ordinateurs comme Digital Audio Workstation plutôt que d'utiliser un matériel dédié, comme une console IP, et plutôt utiliser une console USB ou Firewire comme surface de commande pour piloter la console du logiciel d'acquisition et de traitement du son utilisé.

L'autre option serait d'utiliser une console à part entière ou transiterait l'audio, avec une sortie réseau pour distribuer les flux d'audio sur IP, en unicast ou en multicast vers les régies. L'avantage de cette technique serait de pouvoir mixer directement la prise de son, et avoir un accès au routing plus manuel que sur l'ordinateur. Ce n'est toutefois pas forcément la meilleure solution d'un point de vue pratique, car un changement de régie pourrait devenir compliqué en termes de déplacement de matériel, bien que des consoles légères existent, à moins d'acheter une console par régie. Cette dépense budgétaire n'est pas obligatoire, et on privilégiera donc la première solution consistant en l'utilisation d'une surface de commande intégrée aux régies, et de matériel de conversion dissocié.

Cette option permettrait d'utiliser les différents périphériques de conversion analogique/numérique en tant que « cartes son déportées » via le réseau.

La question des retours audio et vidéo pour les musiciens, dans le cas où ils souhaiteraient enregistrer dans des lieux séparés, nécessite le minimum de latence possible. Des solutions existent chez Merging Technologies (**Horus** ou une machine similaire plus compacte, **Hapi**) où les retours sont directement envoyés dans les casques avant numérisation du signal. Toutefois, si les musiciens sont dans des pièces séparées, ce système ne pourra pas être exploité, il faudra nécessairement transiter par le réseau avant restitution de l'audio.

### **2.2.2. RÉUTILISATION DU MATÉRIEL EXISTANT**

On pourrait, pour alléger le budget, songer à réutiliser le matériel déjà présent. Le cahier des charges stipulant un besoin de 2x16 entrées respectivement dans la grange et dans le studio, on pourrait utiliser la MOTU 896 mk2 et la MOTU Traveler pour disposer d'entrées supplémentaires, que l'on brancherait sur une des machines capables de communiquer en AoIP.

Selon le choix du module de conversion vers l'AoIP ou le protocole AVB, exposé plus bas, on pourrait utiliser les sorties des MOTU 896 mk3 ou Traveler et les convertir sur le réseau IP, ou utiliser leurs entrées pour les transformer en système de retour audio.

Il faudra utiliser une autre machine pour le retour audio depuis l'ordinateur de la régie : une entrée micro de la MOTU 896 servait remplir cette fonction. Il y a, dans la liste de matériel de La Beaudelie, une petite carte de conversion avec deux entrées microphoniques et se branchant en USB : l'USB Dual Pre Series de chez ART Pro. L'iMac de la régie, ainsi que celui du bureau, disposent tous deux de ports USB 2.0, ce qui serait suffisant pour ce type d'utilisation.

Toutefois, il faudra être prudent par rapport à l'utilisation simultanée du système en Ravenna, l'adjonction d'une carte supplémentaire pourrait monopoliser des ressources informatiques précieuses pour l'enregistrement, et introduire plus de problèmes qu'elle ne viendrait en résoudre. Il faudrait faire des essais pour vérifier la capacité de gestion du système d'exploitation de l'ordinateur. Se reporter à la partie 1.4.2 pour la description de ces machines.

### 2.2.3. CHOIX DE L'HORLOGE

Pour de l'audio numérique à haute fréquence d'échantillonnage, il faut impérativement une horloge stable sur tout le réseau. Certaines machines peuvent générer un signal de synchronisation très fiable.

Là encore, plusieurs solutions techniques se présentent : certains appareils tels que Horus ou Hapi de Merging Technologies peuvent fournir une horloge extrêmement fiable, et peuvent être horloge-maître sur un réseau via le protocole PTP (*Precision Time Protocol*, protocole de précision temporelle, aussi appelé protocole d'**horloges distribuées**).

L'AES67 prévoit l'utilisation d'un algorithme associé au PTP (*Precision Time Packaging*) pour trouver l'horloge la plus fiable d'un petit réseau, et l'utiliser comme Master Clock : le Best Master Clock Algorithm (BMCA).

Pour une solution en AVB (*Audio Video Bridging*), la synchronisation est également automatique, la meilleure horloge du réseau est sélectionnée.

Concernant le matériel non-AoIP branché sur les interfaces Ravenna ou AVB, il convient de s'assurer d'une bonne synchronisation, sous peine de graves dysfonctionnements.

L'ADAT peut poser des problèmes au niveau synchronisation : un canal Word Clock est fourni à l'arrière de la machine, en BNC, pour éviter les problèmes. En cas de sélection d'une des deux interfaces MOTU, on pourra les synchroniser via Word Clock. L'ADAT, chez MOTU, fonctionne via le timecode SMPTE ou l'ABS, mécanisme propriétaire d'horloge. Il faut donc vérifier que l'horloge de la machine ADAT soit correctement synchronisée avec l'horloge maître<sup>12</sup>.

---

<sup>12</sup> Configuration ADAT des interfaces MOTU :  
[http://www.motu.com/techsupport/technotes/tt\\_opticalexpansion](http://www.motu.com/techsupport/technotes/tt_opticalexpansion)

## 2.3. FONCTIONNALITÉS SECONDAIRES

### 2.3.1. CHOIX DES CAMÉRAS IP

Même dans le cas d'un simple retour visuel, une certaine qualité d'image est bienvenue : une caméra IP en full HD (1920 x 1080) serait une option valable. On aura besoin d'une seconde caméra si on veut un retour visuel bidirectionnel.

Une façon d'implémenter un réseau vidéo qui soit, d'une part, sur un réseau séparé, et d'autre part économise des câbles, est l'utilisation d'un réseau Wifi. En effet, certaines caméras IP que l'on peut trouver sur le marché permettent une transmission sans fil. Le problème principal de ce choix résiderait dans l'épaisseur des murs de schiste à la Beaudelie, risquant de bloquer les ondes électromagnétiques.

En termes de solution matérielle, les Novodio Smartcam+ HD sont motorisées, et peuvent fonctionner en Wifi 802.11b/g/n, ou sur câble Ethernet. Elles ne fournissent pas du vrai HD à 1920 par 1080 pixels, mais une version de 1280 par 960, et encode le flux vidéo en mpeg-4 « advanced video coding » avec H264, ce qui donnerait un débit de l'ordre de 7Mo/s.

On peut piloter à distance leur zoom optique, leur rotation complète, ou leur inclinaison jusqu'à 120° en tout, et les flux fournis sont disponibles via le navigateur d'un ordinateur ou d'une tablette connectée sur le réseau.

Le seul inconvénient est leur branchement électrique permanent sans possibilité d'extinction.

On peut les connecter sur le réseau généré par une antenne Wifi Airport Express. Cette petite borne qui permet de créer un Wireless LAN sous la norme 802.11n. Cette norme permet d'effectuer des échanges MIMO (*Multiple Input, Multiple Output*), et de choisir la fréquence d'émission à 2,4 ou 5 Ghz. La borne Airport Express émet sur les deux bandes de fréquences en même temps et permet un débit maximal théorique de 300Mo/s.

La borne serait à priori placée au-dessus de la grange, dans le bureau.

### 2.3.2. CHOIX D'UN SERVEUR NAS

Les serveurs NAS, ou Network Attached Storage, sont des serveurs accessibles via de multiples protocoles réseau, et dédiés au stockage de données. Pour un studio avec plusieurs régies, il peut être utile de centraliser les sauvegardes sur un stockage en RAID.

Ainsi, après une session de travail, celle-ci peut être copiés sur le serveur, et sa sauvegarde pourra être accessible depuis n'importe quelle autre régie.

ProTools10 permet de gérer une session sur un NAS<sup>13</sup> ce qui peut rendre le travail encore plus rapide. En revanche, il est également spécifié, sur la page 53 de ce même document, que les possibilités de travail directement sur un tel serveur sont dépendantes « de multiples facteurs » comme la bande passante du réseau ou la performance du NAS.

Ajouter un contrôleur RAID directement sur l'ordinateur de travail aurait pu être une solution, mais avec 3 régies, il est plus intéressant d'avoir un stockage directement sur le réseau afin de pouvoir partager les ressources, d'autant plus que si les ordinateurs sont fabriqués par Mac, toute adjonction de matériel est délicate.

Une autre possibilité intéressante du NAS est l'absence de problèmes de compatibilité dans le cas de transferts entre différents systèmes de fichiers : un Mac ou un PC peuvent indifféremment y écrire ou y récupérer des données. Dans le cas d'artistes venant avec leur matériel informatique, ils pourraient également accéder aux données du NAS ou y écrire.

Un Nas avec de la RAM suffisante semblerait capable d'écrire un flux audio temps réel sur un RAID1-0 sans problème. Il faut effectuer des tests de bande passante pour déterminer si un VLAN ne serait pas nécessaire à l'acheminement des flux audio, depuis les ordinateurs jusqu'au serveur de fichiers, pour être sûr que tous les paquets arrivent à destination.

Les sessions ProTools à partir de la version 10 HDx ou « Complete Toolkit » étant lues à partir de la RAM de l'ordinateur en fonction, il ne devrait pas y avoir de souci de fluidité de lecture ou édition de session.

---

<sup>13</sup> voir page 52 des nouveautés ProTools10 :  
<http://www.percorsiaudio.com/WhatsNewinPT10.pdf>

Pour un serveur NAS servant uniquement de backup, tout NAS pourrait suffire à créer des sauvegardes de sessions post-enregistrement.

## **2.4. SÉCURITÉ**

### **2.4.1. SÉCURITÉ ÉLECTRIQUE**

Il est essentiel que les machines, surtout les machines vouées à l'enregistrement et le matériel informatique, soient protégées par un onduleur. En effet, une mise hors tension inopinée peut conduire à la détérioration du matériel ou à la perte d'information.

Ainsi, il convient de mettre les ordinateurs du réseau sur onduleur, ainsi que le serveur NAS éventuel et les interfaces son. Il existe trois catégories d'onduleurs, que l'on pourrait résumer ainsi :

- les onduleurs passifs, qui se déclenchent quand ils détectent une coupure électrique (*offline*). Une micro-coupure ne suffira généralement pas à faire fonctionner cet appareil.
- les onduleurs interactifs (*line-interactive*), dotés d'un processeur pour surveiller la qualité de l'alimentation fournie aux machines, en plus d'assurer une fonction de batterie en cas de panne.
- les onduleurs en ligne (*online*), qui filtrent et régulent la tension en permanence.

L'onduleur en ligne fournit la meilleure protection, mais son prix est très élevé pour une petite infrastructure.

## 2.4.2. SÉCURITÉ INFORMATIQUE

Un réseau IP est vulnérable à de nombreux types d'attaques : capture de paquets, surveillance des ports ouverts, communication indésirable avec des applications, ou même prise de contrôle. Il convient d'assurer une sécurisation du réseau et des ports utilisés, notamment par un pare-feu entre le réseau local et Internet.

La plupart des routeurs domestiques permettent des fonctionnalités de NAT (*Network Address Translation*, ou translation d'adresses réseau). C'est, comme son nom l'indique, une réassignation des adresses d'un sous-réseau depuis le routeur. La box obtient une adresse IP unique depuis Internet, et tous les hôtes connectés peuvent avoir une adresse individuelle différente de celle de la box. La box sert ainsi de **passerelle** à Internet. Une passerelle (en anglais, *gateway*) interconnecte deux réseaux différents, et peut aussi servir de **pare-feu**, qui est un filtre logiciel et/ou matériel des connexions entrantes et/ou sortantes. Dans cet exemple, la passerelle relie le LAN et Internet.

Le NAT est un premier degré de sécurité : en effet, les appareils ne sont pas directement accessibles depuis Internet. D'après la RFC1918 qui a trait aux adresses privées, plusieurs plages IP sont adressables en fonction de l'utilisation requise sur ces postes, détaillées à la page 3 :

10.0.0.0 - 10.255.255.255

172.16.0.0 - 172.31.255.255

192.168.0.0 - 192.168.255.255

Un logiciel utilisant NAT-PMP (*NAT-Port Mapping Protocol*, ou NAT-Adressage de Ports), tel que Bonjour d'Apple ou Zeroconf utilisé sur les réseaux Ravenna pour la découverte de matériel, peut, sous certaines conditions, parvenir à résoudre les règles de translation d'adresses et entraîner des vulnérabilités.

### **3. PROPOSITIONS TECHNIQUES**

Les deux propositions techniques exposées ont été choisies selon le cahier des charges : il fallait répondre au problème de la flexibilité de l'installation tout en garantissant une mise en route rapide et efficace.

Deux solutions seront retenues : l'implémentation d'un réseau Ravenna, et l'implémentation d'un réseau AVB, avec des variantes de réutilisation du matériel déjà en place à La Beaudelie.

Pour ce faire, les deux propositions techniques seront organisées selon le matériel utilisé, puis les configurations logicielles possibles pour optimiser le fonctionnement de l'installation.

#### **3.1. MATÉRIEL NON-SPÉCIFIQUE AU PROTOCOLE UTILISÉ**

Selon la topologie du réseau, il faudra différentes longueurs de câble blindé, au moins de type Cat5e. On peut également étudier la possibilité de raccorder le routeur du studio et créer deux LAN séparés.

Par rapport au système d'amplification Pioneer de la grange, utilisé pour les configurations de concert, de projection ou de mixage cinéma, les entrées disponibles sont très nombreuses en HDMI. La fonctionnalité « Aggregate » de Pro Tools permet d'utiliser cette sortie depuis un ordinateur.

La longueur maximale d'un câble HDMI se situe entre 5 et 15m, selon la qualité du câble. Il faut donc pouvoir déplacer un ordinateur directement dans la grange, utiliser un répéteur, ou encore une rallonge HDMI sur câble Ethernet.

Le Pioneer dispose également d'une entrée LAN. Il semblait intéressant d'interroger la possibilité d'y faire passer de l'audio. Le support technique de Pioneer a répondu :

« Nous vous informons que le VSX-923 est compatible afin de partager de la musique avec une connexion sur le réseau via un câble Ethernet en utilisant la fonction DLNA [*Digital Living Network Alliance, qui cherche à créer une norme d'interopérabilité pour des produits multimédia grand public*] à configurer sur Windows media Player 11 depuis un PC, et la fonction Air-Play à configurer sur iTunes depuis un produit Apple. »

Il semble donc peu probable que des protocoles AoIP spécifiques fonctionnent actuellement avec cette norme, à moins de rajouter un petit serveur de conversion DLNA, qui risque de rajouter de la latence dans la chaîne de restitution.

Il y aurait une possibilité, via la spécification du *Real-time Transport Protocol*, de récupérer les flux audio grâce à la description donnée par le *Session Description Protocol* (SDP). Livewire décrit la méthode sur leur site internet<sup>14</sup>, qui consiste à télécharger une petite extension dédiée à la récupération des flux par Direct Show (grâce auquel fonctionne le lecteur Windows Media Player), à créer un fichier SDP (extension .sdp) spécifiant les caractéristiques du flux, puis à l'ouvrir avec le Windows Media Player. Ce logiciel étant compatible DLNA, il pourrait convenir au streaming vers l'amplificateur.

Sous Mac, une démarche analogue est possible, en récupérant les flux grâce à *QuickTime Streaming Server*, qui peut les rendre accessibles via iTunes, également compatible DLNA.<sup>15</sup>

Ces deux solutions présentées, bien que techniquement possibles, sont longues à mettre en œuvre. De plus, ces solutions seraient difficilement possibles à implémenter sur le même ordinateur au niveau de la gestion de ses ressources et de l'encombrement du réseau : il faudrait à la fois pouvoir récupérer plusieurs flux audio non-compressés et les réexpédier, le tout avec une latence acceptable, ce qui

---

<sup>14</sup><http://www.telosalliance.com/images/Axia%20Products/Support%20Documents/Tech%20Tips/Playing-Livewire-Streams-Using-SDP-Rev-3-2015.pdf>

<sup>15</sup>[https://manuals.info.apple.com/MANUALS/0/MA579/en\\_US/QuickTimeStreamingSrvrAdminGuide.pdf](https://manuals.info.apple.com/MANUALS/0/MA579/en_US/QuickTimeStreamingSrvrAdminGuide.pdf)

poserait des problèmes de surcharge des queues Rx et Tx (réception et transmission) et une dégradation très probable de l'audio.

Il est d'ailleurs conseillé de fermer les autres applications audio pour un travail sur Pro Tools ou tout autre logiciel d'édition, à fortiori en configuration « temps réel ».

Le Pioneer dispose également de deux entrées audio numériques, respectivement en coaxial et en optique. Il est peut être possible d'utiliser ces entrées également. Toutefois, la longueur maximale préconisée par l'AES pour une liaison audio sur un câble coaxial est de 15m, et les connectiques Toslink ne dépassent pas 10m pour de la fibre plastique. Il faut par conséquent que l'ordinateur source soit dans la grange, ou adapter les câbles afin de les rallonger.

### 3.2. PROPOSITION TECHNIQUE EN RAVENNA

Merging Technologies a sorti, en 2014, **Hapi**, un convertisseur analogique-numérique capable de transmettre jusqu'à 16 canaux en entrée ou sortie à une fréquence d'échantillonnage de 192 kHz, dépendant du nombre de cartes de conversion utilisées et de leur fonction.

Pour les besoins de l'installation, il faut 16 entrées micro dans la grange, et 16 entrées micro dans le studio.

Pour une utilisation avec un Pro Tools Native sur Mac OS X 10.8 minimum, il est possible d'enregistrer jusqu'à 32 pistes à 192 kHz, ce qui est la capacité maximale requise pour l'installation en utilisant toutes les entrées « grange » et « régie » en même temps.

Sous Mac OS X avec Core Audio, le pilote *Ravenna Core Audio Driver* permet de n'avoir que 32 samples de latence d'entrée, ce qui fait, à 192 kHz, 5,2  $\mu$ s environ, ou 10,4  $\mu$ s à 96 kHz. On constate que le driver ASIO de Windows entraîne le double de cette latence incompressible.

Au studio de la Beaudelie, on travaillera le plus souvent à 24 bits et 48 kHz. La recommandation de Merging Technologies au sujet des buffers est de 256 samples<sup>16</sup>. Il est possible d'ajuster cette taille, mais il faut prendre garde aux erreurs d'enregistrements dues à un buffer trop petit. Ces erreurs peuvent être dues à une

---

<sup>16</sup> source : <http://www.merging.com/products/networked-audio/for-avid-pro-tools>)

utilisation élevée du processeur pour traiter l'audio venant de la carte réseau, mais également au jitter sur le réseau.

En termes de contrôle, les niveaux des préamplificateurs de la machine sont directement pilotables via ProTools. L'accès à tous ses réglages (phase des préamplificateurs, gain, alimentation phantom...) se fait par un navigateur standard, donc via le protocole HTTP, fonctionnant sur couche TCP. Un accès direct depuis Pro Tools est également possible grâce à la synchronisation MIDI.

À titre d'exemple, l'interface de contrôle des préamplificateurs se présente ainsi depuis un navigateur :

#### Hapi PreAmp Web menu

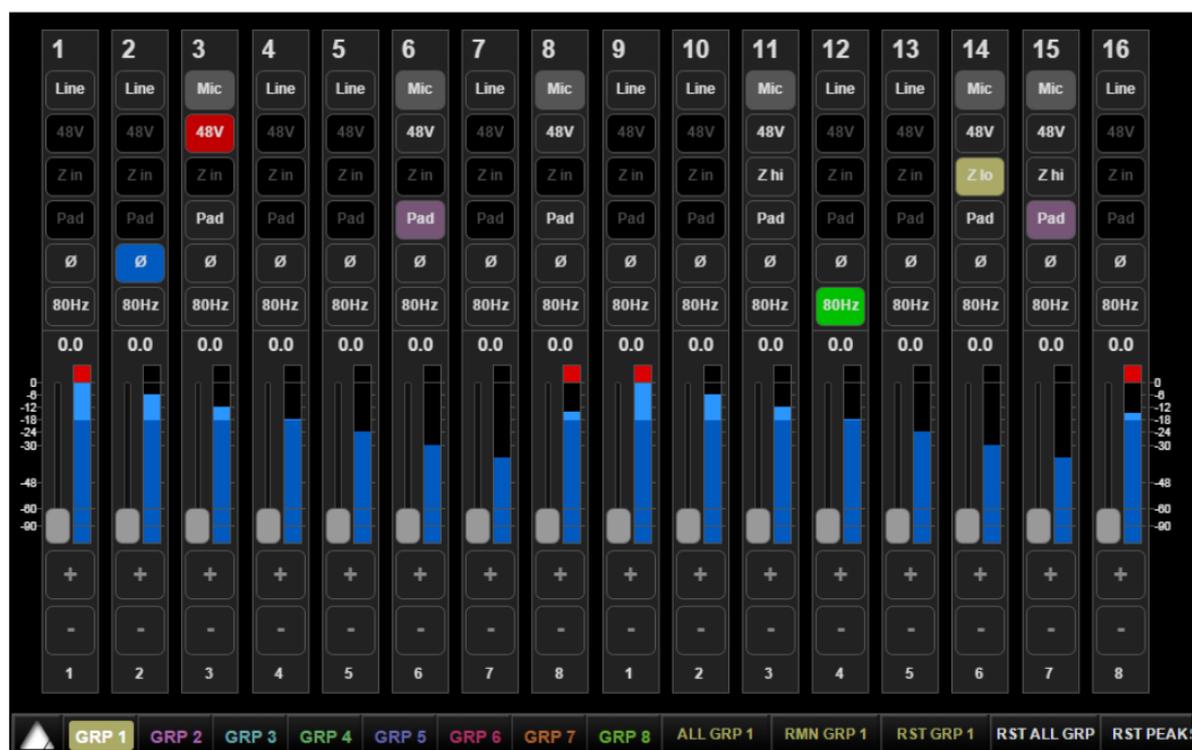


Image tirée du manuel du Hapi.

Si, en théorie, les données de contrôle, non-prioritaires, n'influent pas sur les flux audio, il faut rester prudent. En traversant le réseau en TCP, tout paquet perdu appellera sa retransmission, jusqu'à ce qu'il parvienne à destination.

Il faut donc être certain que les mécanismes de *Quality of Service* soient correctement implémentés pour que le trafic UDP, et donc l'audio, n'en soit pas dégradé.

Parmi le matériel non-spécifique au réseau AoIP, on peut compter l'achat d'une surface de commande adaptée. Là encore, les options sont vastes sur le marché, quelque soit le nombre de potentiomètres linéaires assignables ou la connectique, il existe un marché dédié aux surfaces de contrôle.

L'**Artist Mix** d'Avid est une option intéressante en termes de compacité et de fonctionnalités. Elle peut d'ailleurs être branchée n'importe où sur le réseau : ses commandes peuvent circuler via le protocole *Eucon* sur un câble Ethernet. Là encore, il faudra veiller à ce que la *Quality of Service* permette d'utiliser cet outil sans impacter le trafic audio.

Toutefois, il y aurait besoin d'un **switch** supplémentaire pour l'utiliser dans une régie ou dans le studio, le switch principal se trouvant dans la grange. On peut donc réfléchir à une autre option.

Il faudrait mener des tests pour vérifier que l'ordinateur puisse recevoir des données de contrôle via un contrôleur standard, fonctionnant via USB ou Firewire, sans dégrader la réception de l'audio.

Traditionnellement, dans la hiérarchie des interruptions des périphériques (*Interrupton Requests*, ou IRQ : le matériel peut demander un accès prioritaire au processeur pour un traitement immédiat), le contrôleur USB a la priorité sur le contrôleur réseau. Il est faudrait pouvoir aller voir plus en détail dans l'ordinateur utilisé la façon dont est gérée cette hiérarchie.

Pour des données de contrôle, ponctuelles, et un ordinateur de puissance suffisante, le contrôleur USB ne devrait pas poser de problème. Toutefois, seul un système d'exploitation temps réel dépourvu d'interruptions non-déterministes et non autorisées de matériel pourrait mettre à l'abri le flux audio.

### 3.2.1. MATÉRIEL SPÉCIFIQUE

#### 3.2.1.1. CONFIGURATION « TOTALE » EN RAVENNA

Il faudrait donc, pour une configuration « totale » indépendante de l'infrastructure actuelle :

- deux interfaces Hapi <sup>17</sup> de Merging Technologies
- quatre cartes de conversion AD8D
- éventuellement un NADAC à utiliser comme retour casque, ou comme sortie pour des enceintes (jusqu'à 8 sorties), en fonction de la configuration requise
- des plaques pour rack 19 pouces destinées à recevoir des embases XLR, afin de faciliter une utilisation « scène » ou studio : 4 x 8 XLR en plaques 1U, ou 2 x 16 XLR en 1 ou 2U
- quatre câbles de conversion de la connectique DB-25 (Tascam) à 8 entrées XLR chacun, couramment appelé « snake ».
- éventuellement un câble multipaire pour une extension dans la grange, pour assurer un meilleur confort d'installation des musiciens.
- le switch conseillé pour cette utilisation est le DellPowerConnect 2816 (ou apparenté)
- une surface de commande, telle que l'Artist Mix

#### 3.2.1.2. CONFIGURATION « PARTIELLE » EN RAVENNA

Toutefois, ce genre de configuration est assez coûteux. Pour un premier investissement, il peut être intéressant de proposer une configuration « partielle », qui peut se prêter à une extension ultérieure.

On pourrait imaginer un Hapi principal dans la grange doté d'une carte de conversion analogique-numérique à 8 entrées AD8D, et une carte de conversion numérique-analogique à 8 sorties DA8P pour accommoder les retours casque.

Il y aurait éventuellement un Hapi d'extension, pour le moment doté d'une seule carte d'entrée, dans la régie. Il pourrait être déplacé selon les besoins vers la grange.

Ces deux machines peuvent accueillir chacune une entrée ADAT, qui ajouterait donc 8 canaux à 24/48 kHz. On peut envisager d'utiliser les MOTU 896 et Traveler

---

<sup>17</sup> Brochure disponible sur le site de Merging Technologies : [http://www.merging.com/uploads/assets/Merging\\_pdfs/Brochures/HapiBrochure-2014.pdf](http://www.merging.com/uploads/assets/Merging_pdfs/Brochures/HapiBrochure-2014.pdf)

branchées en ADAT sur chaque Hapi et avoir le total des 32 canaux possibles souhaités. On prendra soin de synchroniser les horloges des MOTU par le Word clock, avec un câble BNC branché du « Word Clock out » des interfaces Hapi vers le « Word Clock in » des MOTU.

En prévision de **retours casque**, on sortirait de la carte de conversion numérique/analogique du Hapi vers les entrées analogiques d'un amplificateur pour casques.

Une option possible d'amplificateur serait le **SM Pro HP6E**, disposant de 10 entrées jack stéréophonique TRS et 6 sorties « casque », mais il semble présenter des problèmes de souffle, d'après les retours des utilisateurs sur le site de Thomann.

On pourrait également choisir le **LD systems HPA**, disposant de 6 entrées jack stéréo et 6 canaux indépendants pour des sorties « casque ».

Quelque soit le choix de machine pour le retour casque, on branchera la carte de conversion numérique/analogique du Hapi à l'aide d'un câble de conversion Tascam db25 vers des connectiques jack stéréo.

### 3.2.2. SCHÉMA DE RÉSEAU ET LONGUEUR DE CÂBLE NÉCESSAIRE

Le Dell Power Connect 2816 dispose de 16 ports. On pourrait donc théoriquement n'utiliser que ce switch pour brancher tout le réseau. De plus, ce switch ne disposant pas de ventilateur, il est possible de le placer dans un environnement dédié à la prise de son ou au monitoring sans générer de nuisances sonores.

Les longueurs de câble nécessaires à cette opération sont dessinées dans le schéma associé, présenté en Annexe 2 pour la solution « totale », et en Annexe 3 pour la solution « partielle ». La longueur requise n'est pas une distance exacte, il faut laisser un rayon d'action autour de chaque point d'arrivée pour pouvoir déplacer le matériel.

Le switch se trouverait dans la grange, avec le rack, les amplificateurs, la console lumière, etc... Proche de l'entrée.

- De la grange à un PC branché dans la grange : **5 m** est un rayon d'action raisonnable autour de la zone de branchement possible du switch.
- pour une caméra IP placée sur un bras télescopique ou pied mobile dans la grange, il faudrait une longueur optimale de **7m** de câble (3,15m de plafond, et un rayon de 3m environ autour du rack). Ce câble n'est nécessaire que si le Wifi n'est pas possible à utiliser, des tests seraient nécessaires.

- Pour un Hapi racké avec les autres appareils : **0.5m ou moins**
- pour un Hapi dans le studio : **10m** (en supposant que 32 entrées dans le studio ne seront pas nécessaires, mais en conservant l'option de pouvoir déplacer le Hapi dans la grange pour 32 entrées à cet endroit : prévoir un petit câble Ethernet Cat5e supplémentaire à l'intérieur de la grange dans l'éventualité où les deux Hapi seraient mis en rack avec les autres machines)
- Pour un câble ordinateur dans le studio (branché ou non) : **10m**

Ce câble pourra être le même que celui prévu pour le Hapi. En effet, la pièce peut assurer deux fonctions successivement, mais pas en simultané. On n'aura donc besoin de cette pièce qu'en « régie » ou en « studio », et donc économiser l'installation d'un câble.

- pour une caméra IP optionnelle dans le studio : **10m**, sauf si on la connecte sur WLAN (Wireless LAN)
- de la grange au bureau : **8m**
- de la grange à la régie : **15m**
- de la grange au NAS : **15m** s'il est branché dans la régie, ou alors **10m** vers le bureau

Un total de 90 à 100m de câble sera donc nécessaire pour la solution « totale » en Ravenna.

Pour la solution « partielle », il faut compter sensiblement la même longueur de câbles, hormis vers le NADAC, ce qui ôte 5m de la longueur totale. Il faut également prévoir deux câbles Toslink et deux câbles BNC pour les connectiques ADAT et Word Clock. Il faudra de plus 20 embases pour confectionner les câbles, et la pince à sertir adaptée aux câbles Cat5e.

On peut aussi acheter des câbles déjà confectionnés si on n'est pas sûr de pouvoir sertir les embases correctement : un mauvais sertissage peut entraîner des difficultés de propagation sur le câble et une perte de qualité du signal.

### 3.2.3. LOGICIELS

#### 3.2.3.1. DÉCOUVERTE ET ROUTAGE

Le logiciel d'aide à la découverte des appareils réseau s'appelle MTDDiscovery (*Merging Technologies Discovery*). Grâce au protocole Bonjour, il détecte tous les éléments du réseau prenant en compte les Services Bonjour, et les regroupe selon leurs catégories :

- appareils Ravenna (Horus, Hapi, CoreAudio, Asio...)
- serveurs Emotion, Pyramix, Vcube ; Ovation...
- autres périphériques réseau détectés, mais non-reconnus par le service Bonjour, tels que des imprimantes.

Selon un code couleur, deux appareils surlignés de la même couleur sont sur le même réseau. Pour fonctionner en Ravenna, il faut que les appareils soient idéalement sur le même réseau.

On peut accéder aux réglages des appareils listés via le clic droit ou ctrl+clic sous Mac. Un navigateur web standard permet de les contrôler, via le protocole HTTP.

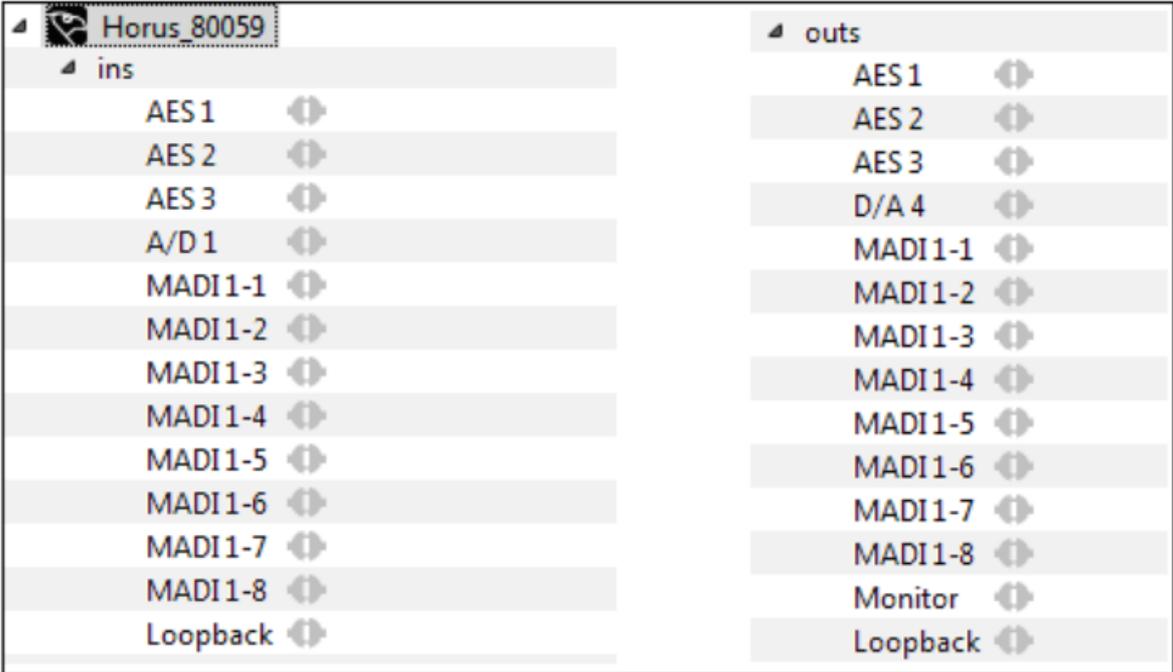
#### 3.2.3.2. RÉCUPÉRER L'AUDIO DANS UN LOGICIEL D'ÉDITION

Pour récupérer l'audio dans un logiciel d'édition, on peut résumer le processus d'acquisition par les étapes suivantes :

- - ouverture de session entre le Hapi et l'ordinateur, et transmission de l'audio ;
- - récupération du flux par l'interface réseau, qui s'occupe de traduire les paquets reçus en données ;
- - récupération des données par le pilote adapté entre la carte et Core Audio ;
- - utilisation de l'audio par une application.

Dans le cas d'un streaming audio depuis un Hapi jusqu'à Pro Tools, le Ravenna Easy Connect, application dédiée à l'interconnexion des appareils, permet de relier les sorties des hôtes sur le réseau vers les entrées d'autres hôtes.

On routera donc les flux à récupérer dans le logiciel vers les « entrées Core Audio ».



The screenshot displays a software interface for a device named 'Horus\_80059'. It is divided into two main sections: 'ins' (inputs) on the left and 'outs' (outputs) on the right. Each section contains a list of channels, each with a small icon to its right. The 'ins' section lists: AES 1, AES 2, AES 3, A/D 1, MADI1-1, MADI1-2, MADI1-3, MADI1-4, MADI1-5, MADI1-6, MADI1-7, MADI1-8, and Loopback. The 'outs' section lists: AES 1, AES 2, AES 3, D/A 4, MADI1-1, MADI1-2, MADI1-3, MADI1-4, MADI1-5, MADI1-6, MADI1-7, MADI1-8, Monitor, and Loopback.

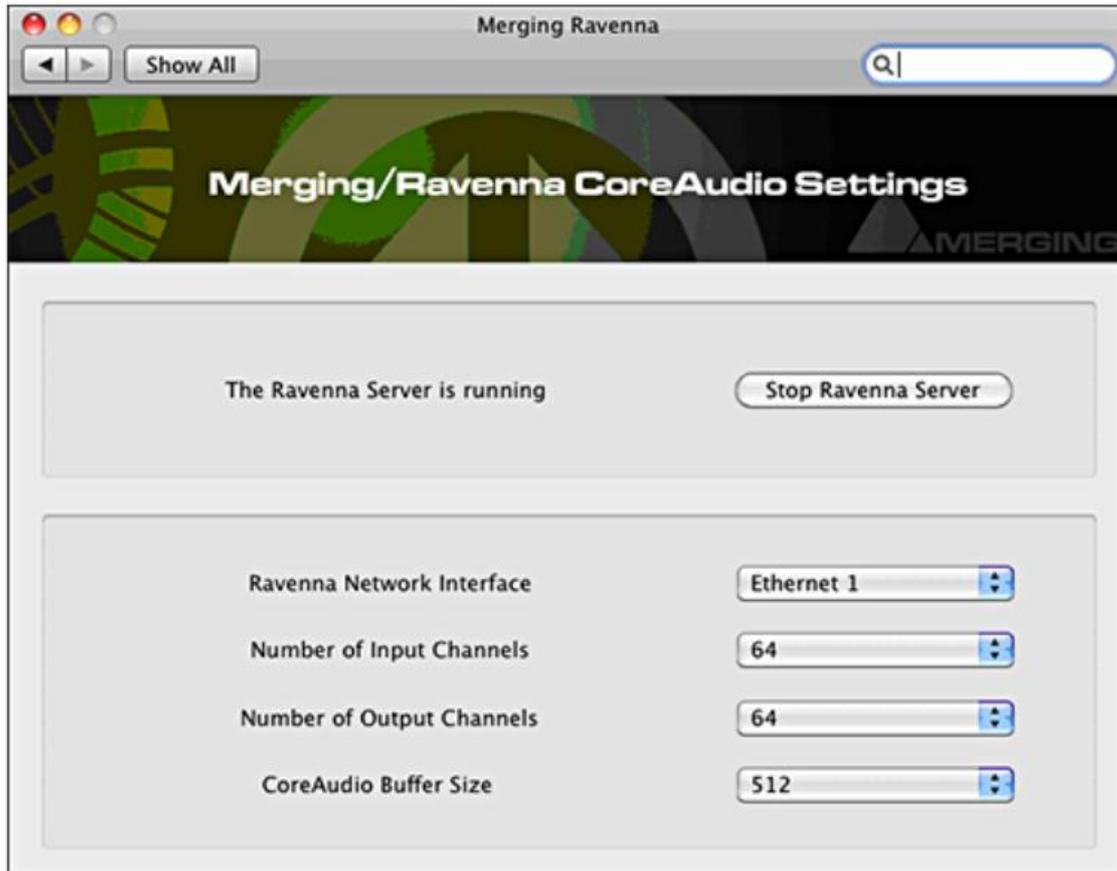
ins	outs
AES 1	AES 1
AES 2	AES 2
AES 3	AES 3
A/D 1	D/A 4
MADI1-1	MADI1-1
MADI1-2	MADI1-2
MADI1-3	MADI1-3
MADI1-4	MADI1-4
MADI1-5	MADI1-5
MADI1-6	MADI1-6
MADI1-7	MADI1-7
MADI1-8	MADI1-8
Loopback	Monitor
	Loopback

**Easy Connect ins and outs**

*Interface du Ravenna Easy Connect*<sup>18</sup>

Il faut ensuite ouvrir l'interface utilisateur de Core Audio, et récupérer les flux entrants. On peut paramétrer le choix de l'interface réseau, le nombre d'entrées et de sorties, et la taille du buffer. On peut aussi interrompre la transmission en arrêtant le Ravenna Server, qui se trouve sur le Hapi.

<sup>18</sup> Image extraite du Ravenna Network Guide, sur le site de Merging Technologies, page 14



**Merging/Ravenna Core Audio Settings Panel**

*Interface du Ravenna Core Audio Driver, où l'on peut régler les paramètres de réception*

Il ne reste plus qu'à définir le pilote Ravenna en entrées/sorties de ProTools pour recevoir l'audio.

ProTools est capable de gérer directement les données de contrôle du Hapi, comme le gain de ses préamplificateurs, via le Midi Time Code (MTC). La procédure est également détaillée dans le *Ravenna Network Guide*.

### **3.3. PROPOSITION TECHNIQUE EN AVB**

L'AVB est une solution envisageable pour construire un réseau audionumérique souple dans une installation. Toutefois, c'est un protocole de niveau 2 dans la modélisation OSI. Il n'est donc normalement pas routable tel quel.

Il existe cependant des routeurs dédiés, plus chers que des routeurs standards, capables de gérer de tels flux et de les diriger comme souhaité dans l'installation.

L'AVB peut être envisagé dans le cadre de ce mémoire, car la spécification de l'AES67 datant de 2013 prévoit une interopérabilité avec ce protocole. Pour des futurs développements de l'installation, c'est une problématique importante.

### 3.3.1. MATÉRIEL

En termes de matériel, MOTU a sorti plusieurs interfaces et switchs compatibles avec Mac OS 10.10 minimum. Il est également requis de posséder des machines récentes disposant d'un port Thunderbolt, comme l'a prouvé un test matériel de connectivité effectué avec Jean-Marc Chouvel en utilisant un Mac sous OS X 10.10 Yosemite, mais dépourvu de port Thunderbolt, et une carte MOTU Ultralite AVB.

Avec OS X Yosemite, la fonctionnalité AVB ne permet pas encore de faire du 44,1 kHz, du 88,2 kHz, ou du 176,4 kHz. Les multiples de 48 kHz sont possibles jusqu'à trois fois la fréquence d'échantillonnage (suréchantillonnage à 3F).

On peut donc proposer la configuration « totale » suivante :

- la MOTU 8M possède 8 entrées combo jack/XLR, jusqu'à 16 entrées ou sorties ADAT (à 48 kHz maximum), 8 sorties Jack, une sortie casque stéréo, un port « AVB » à connecteur RJ45, un port USB, un port Thunderbolt, une entrée et une sortie pour l'horloge.

On dispose par conséquent de 24 entrées simultanées possibles par interface. Il faut toutefois, pour en bénéficier, une machine capable de leur faire parvenir de l'ADAT et les connectiques associées. On peut envisager la 896 déjà présente ou la 8pre de MOTU pour cette utilisation, et ainsi étendre les entrées XLR à 32 au total, ou utiliser la Traveler et disposer d'un total de 28 entrées XLR réparties entre la grange et le studio.

- le Monitor 8 de MOTU est un appareil dédié aux retours casques, sur lequel on peut brancher jusqu'à 12 sorties casques. Pour les besoins de retours casques, les 8M pourraient suffire : elles disposent de 8 sorties TRS. Si on choisit par exemple la MOTU Traveler ou la 8pre branchée en ADAT, on dispose de 2 sorties TRS supplémentaires.
- Il existe un switch dédié à l'AVB également vendu par MOTU : il dispose de 5 ports réseau dévolus à des appareils fonctionnant sur AVB, et d'un port réseau sur lequel l'AVB ne fonctionne pas. On peut y connecter tout autre appareil prenant en charge le protocole Ethernet.

- Cependant, il pourrait être nécessaire de prendre plusieurs switchs de MOTU, au moins deux pour fonctionner confortablement, ce qui serait assez coûteux. On pourrait aussi choisir un plus gros switch, par exemple chez Netgear, selon que l'on branche les caméras sur des câbles RJ45 ou non.
- Si les caméras sont sur un WLAN (*Wireless LAN*), un seul switch MOTU AVB Switch pourra être utilisé pour ce réseau, en débranchant alternativement les appareils suivant les configurations souhaitées.

MOTU recommande de désactiver ou éteindre tous les appareils avant de changer les branchements, afin de permettre aux machines de toutes être reconnues sur le réseau.

Concernant la longueur des câbles nécessaire, l'utilisation d'un second switch réduit d'une vingtaine de mètres ce besoin, il faut entre 70 et 80m de câbles, selon la place du NAS.

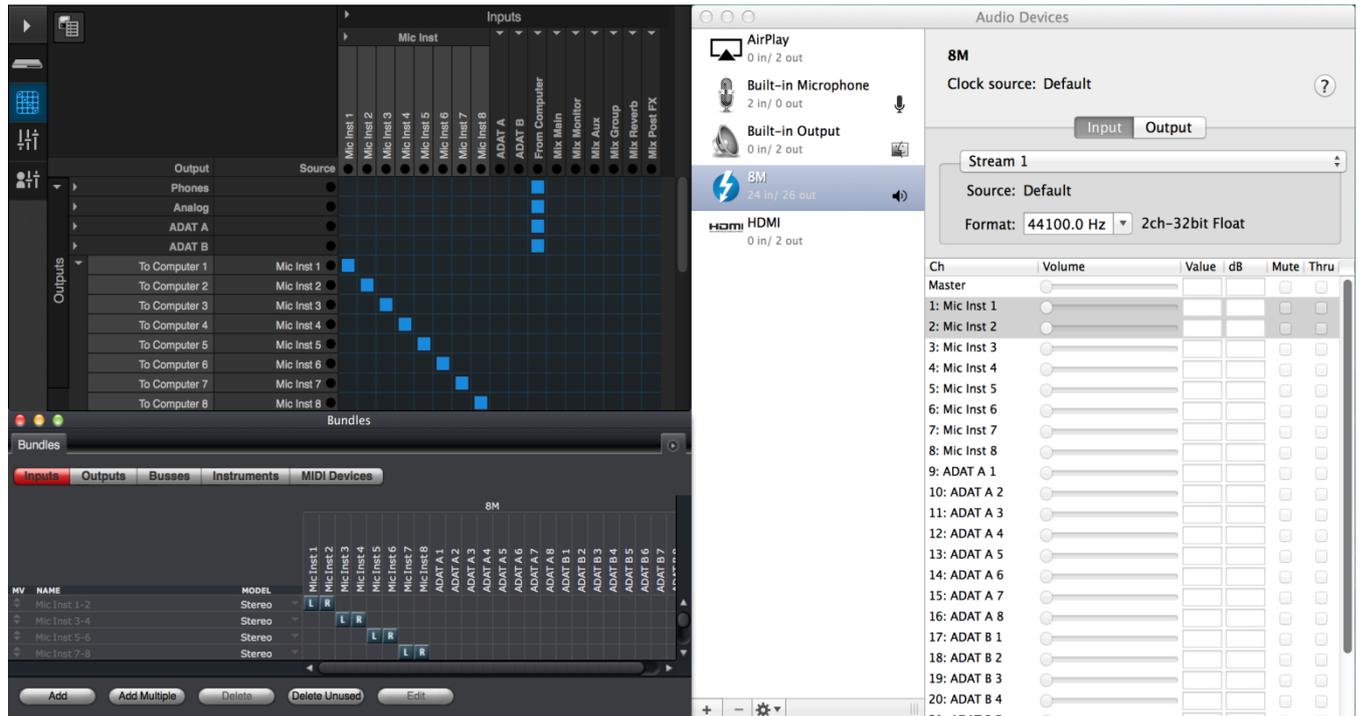
### **3.3.2. CONFIGURATIONS DE ROUTAGE**

Depuis l'ordinateur, l'application web de routage de MOTU est accessible depuis un navigateur standard, comme Safari ou Firefox. On peut y accéder via l'adresse IP de la machine. Elle est constituée d'une grille. Les entrées disponibles sont en haut, disposées en colonnes, et les sorties à gauche, réparties en lignes. Il suffit de cocher les cases correspondantes pour assigner une entrée vers une ou plusieurs sorties.

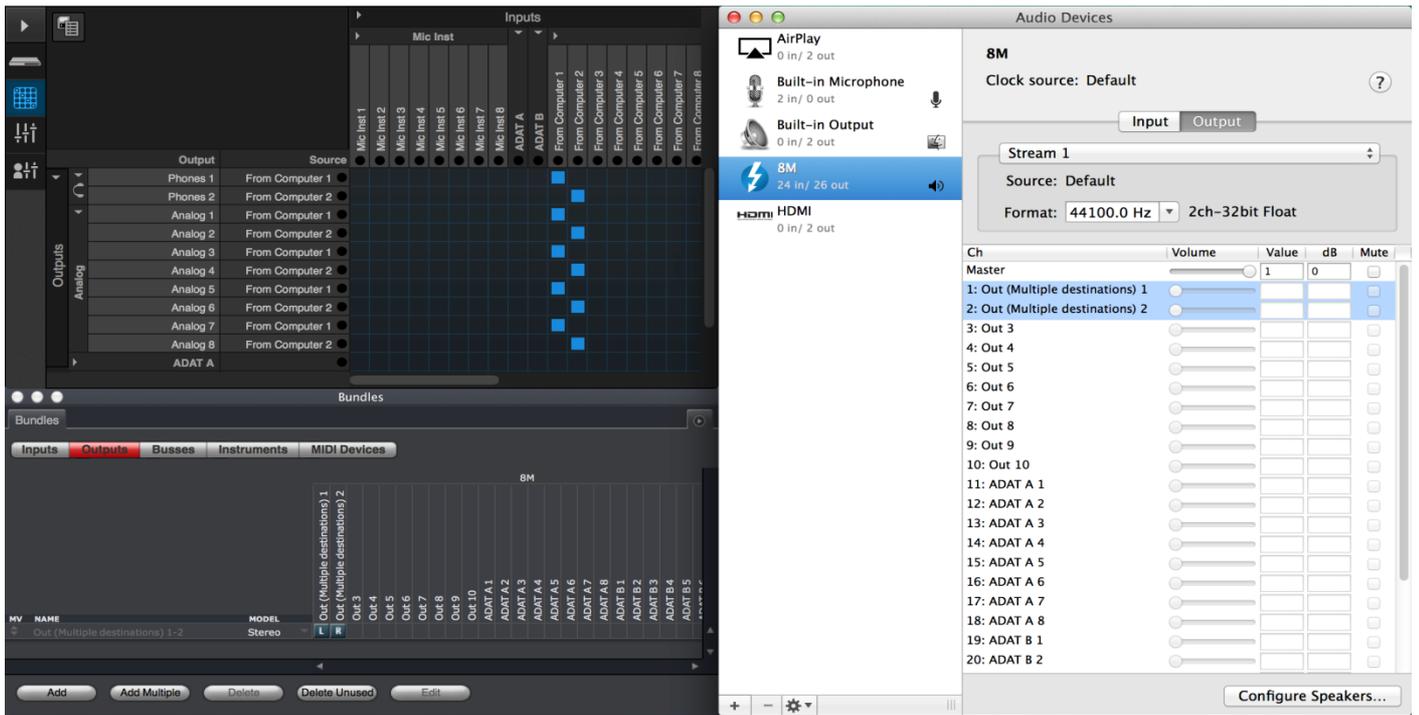
On peut ainsi configurer des « presets » et les enregistrer pour une utilisation plus rapide. Il y aurait donc différents presets à la Beaudelie : pour la prise de son dans la grange seulement, dans la grange et dans le studio, pour les concerts selon leur disposition, pour la projection cinéma, ou pour le mixage, ajustables en fonction des entrées microphones nécessaires, et des besoins en retours, ou de la régie sélectionnée.

### 3.3.3. RÉCUPÉRATION DANS UN LOGICIEL D'ÉDITION

Pour recevoir les flux AVB dans ProTools ou Digital Performer (ou tout autre logiciel d'édition compatible avec Core Audio), il faut récupérer les flux dans l'onglet « routing » de l'application web, et router les entrées voulues vers « to computer » en cochant les cases appropriées. Core Audio récupère le flux AVB, et il sera disponible pour l'envoyer vers ProTools ou Digital Performer, via le gestionnaire Core Audio :



De même, pour renvoyer un flux audio depuis un logiciel d'édition vers le réseau AVB, on peut trouver l'onglet « from Computer » dans l'application web, et router la sortie voulue vers l'entrée de l'appareil ciblé.



Dans cet exemple, on a routé les mêmes sorties 1 et 2 de l'ordinateur, provenant ici de Digital Performer, vers la sortie casque de la 8M en stéréo, et vers toutes les sorties analogiques de la 8M également.

### 3.4. ET EN PRATIQUE ?

Nous allons essayer, dans cette partie, de nous mettre en situation par rapport à l'utilisation d'un réseau sur IP. Pour cette partie, nous imaginerons un réseau Ravenna, dans le cas de la « configuration partielle » représentée à l'Annexe 3.

#### Prise de son dans plusieurs lieux

Imaginons un trio jazz à enregistrer, piano, batterie, et chant. On voudrait placer le percussionniste dans la grange pour disposer de son acoustique naturelle, tout en isolant acoustiquement le pianiste et le chanteur de la batterie.

Dans cette configuration, on aurait une contrainte de latence dans les retours à gérer, sous peine que les musiciens ne se retrouvent complètement désynchronisés, ne disposant que de leur retour casque pour entendre le batteur.

Il faut également penser à un système d'interphonie, depuis la régie vers tous les musiciens, éventuellement avec une composante vidéo bidirectionnelle, qui ne sera pas détaillée ici.

En termes de disposition spatiale on placera le pianiste et le chanteur dans le studio. On pourrait, à titre d'exemple, utiliser le matériel suivant :

- un couple stéréophonique sur le piano
- un microphone pour le chanteur
- dix microphones pour la batterie.

On utiliserait donc:

- 3 entrées XLR de la carte de conversion AD8D du Hapi « studio »,
- huit entrées XLR du Hapi « grange »,
- deux entrées XLR sur la MOTU 896, elle-même branchée sur l'entrée ADAT du Hapi « grange ». La MOTU peut, au total, fournir 8 entrées ou sorties à 24 bits/48 kHz.

Il faudra veiller à ce que l'horloge de la MOTU soit esclave de celle du Hapi, via le câble Word Clock (WC) partant du connecteur « Word Clock Out » du Hapi vers le « Word Clock In » de la MOTU, et vérifier sur l'appareil que la synchronisation soit effective.

Pour le retour du percussionniste, nous utiliserons la liaison analogique au niveau ligne, sortant de la carte de conversion numérique-analogique du Hapi « grange » vers l'amplificateur « casque ».

Pour les retours du chanteur et du pianiste, on pourra utiliser la sortie « casque en façade » du Hapi « studio », dont le gain est ajustable par un potentiomètre. On pourra également utiliser la sortie « casque » de la MOTU Traveler, branchée en ADAT.

On pourra router le signal de Talkback provenant de la régie (via l'interface USB dual Pre) dans tous les retours des musiciens, et assigner une piste du logiciel d'édition à cet effet. La piste Talkback sera routée vers toutes les sorties casque, et, la plupart du temps, mise en position « muette » (*mute*).

On reportera cette configuration dans le Ravenna Easy Connect. On peut sauvegarder jusqu'à cinq presets de configuration par machine Hapi.

### Merging Easy Connect – configuration

L'interface Easy Connect, accessible depuis l'ordinateur, et présentée en partie 3.2.3.2, permet d'effectuer les routages des flux audio.

Pour le Hapi « grange », il faudra assigner :

#### **Entrées**

- « A/D » avec les entrées Core Audio de l'ordinateur (audio de 0 à 7)
- « ADAT » avec les entrées Core Audio (8 à 15)

#### **Sortie :**

- « D/A » avec les sorties Core Audio (0 à 7)

Pour le Hapi « studio », on devra connecter :

#### **Entrée :**

- « A/D » avec les entrées Core Audio (de 16 à 23)

#### **Sorties :**

- « ADAT » avec les sorties Core Audio (de 24 à 31)
- « casque » avec les sorties Core Audio (8 et 9)

Une fois cette configuration effectuée, il faut s'assurer que tous les témoins de connexion soient verts.

Il est ensuite possible de récupérer l'audio dans Pro Tools, comme indiqué dans la partie 3.2.3.2. Il faudra veiller à router l'audio venant de l'USB dual Pre pour permettre la communication depuis la régie (cette possibilité reste néanmoins à tester : la capacité de Core Audio à agréger les entrées et sorties du pilote Ravenna Core audio et celles dédiées à l'USB dual Pre n'est pas assurée).

### 3.5. CONCLUSION DE CETTE MISE EN PRATIQUE

L'avantage d'un réseau « partiel » réutilisant les appareils existants, en Ravenna comme en AVB, est de réduire considérablement le budget par rapport à une solution « complète », tout en gardant possible l'extension AoIP future.

D'autre part, grâce à l'interopérabilité proposée par l'AES67, on pourra bientôt faire des réseaux composés de machines de constructeurs différents, communiquant par des protocoles différents mais interopérables, et pouvant fonctionner ensemble sans problèmes.

On n'est pas encore à cette étape au niveau matériel, bien que des tests publics appelés « plugfests » soient régulièrement organisés par l'AES pour vérifier que l'AES67 permette l'application matérielle des normes et théories de l'interopérabilité, appelée « **compliance** ».

Comme le dit Nicolas Sturmel, directeur de recherche chez Digigram, suite au Plugfest de Munich qui a eu lieu fin 2014 (traduction) <sup>19</sup>

« L'adoption d'une nouvelle norme entraîne un mouvement majeur de l'industrie. Au fur et à mesure de l'avancement dans l'année 2015, nous pouvons nous attendre à des nouvelles implémentations de l'AES67/Ravenna, ainsi qu'au progrès dans l'établissement de « workflows » suivant ces normes. Ce progrès permettra à nos clients d'optimiser l'implantation de ces solutions techniques. »

En effet, il reste encore à intégrer les normes dans le matériel, à intégrer le matériel aux installations, et intégrer les installations aux habitudes de travail actuelles. Les besoins de compliance matérielle, d'intégration à l'industrie, puis de formation professionnelle sont les conséquences de ce changement de mode opératoire.

---

<sup>19</sup> source : nouvelles du Plugfest sur le site de Digigram : <http://www.digigram.com/company/article.php?id=646>

---

# Bibliographie

## DOCUMENTS TECHNIQUES

- [1] **HILDEBRAND, Andreas**, *AES67-2013 : AES standard for audio applications of networks – High-performance streaming audio-over-IP interoperability*, publié par Audio Engineering Society Inc, 2013
- [2] [AES 7944, 2009] *Audio-Over-IP Acceptance Test Strategy*, Matthew J. O'Donnell, publié par Audio Engineering Society Inc, 2009

## NORMES DE L'IETF (INTERNET ENGINEERING TASK FORCE)

- [3] Site de l'IETF avec de nombreuses normes en ligne :  
<http://www.ietf.org/about/standards-process.html>
- [4] **SCHULZRINNE, H, RAO, A, LANPHIER, R**, *Real Time Streaming Protocol*, publié par l'IETF, daté d'avril 1998  
<http://tools.ietf.org/html/rfc2326>
- [5] **SCHULZRINNE, H, CASNER, S, FREDERICK, R, JACOBSON, V**, *RTP : A Transport Protocol for Real-Time Applications*, publié par l'IETF, daté de Juillet 2003  
<http://tools.ietf.org/html/rfc3550>
- [6] **HANDLEY, M, JACOBSON, V, PERKINS, C**, *SDP : Session Description Protocol*, publié par l'IETF, daté de juillet 2006  
<http://tools.ietf.org/html/rfc4566>
- [7] **WILLIAMS, A, GROSS, K, VAN BRANDENBURG, R, STOKKING, H**, *RTP Clock Source Signalling draft-ietf-avtcore-clksrc-11*, publié par l'IETF, daté 25 mars 2014  
<https://tools.ietf.org/html/draft-ietf-avtcore-clksrc-11>

## OUVRAGES

[8] **ROADS Curtis**, *L'Audio numérique, Musique et Informatique*, Paris, Dunod, Collection Audio-Photo-Vidéo, 2007, trad. Jean de Reydellet.

[9] **PUJOLLE, Guy**, *Les Réseaux, Édition 2008*, Paris, Eyrolles, 2007.

## MÉMOIRES

[10] **ROSSI-ROTH, Léo**, *Étude des possibilités d'utilisation d'un réseau sans-fil pour le transport d'un réseau audio numérique*, mémoire sous la dir. de ROUCHOUSE, Jean et COSQUER, Ronan, École Nationale Supérieure Louis Lumière, section son, soutenu en juin 2014

[11] **LEVEUGLE, William**, *Des possibilités d'évolution vers le sans fil des réseaux audio numériques pour la sonorisation- Cas des Wireless Lan*, mémoire sous la dir. de ROUCHOUSE, Jean et QUINIQU, Benoît, École Nationale Supérieure Louis Lumière, section son, soutenu en mai 2012

## SOURCES INTERNET

### PARTIE THÉORIQUE

[12] **FNS-ACIP**, *Audio Contribution over IP : Requirements for Interoperability*, Fiche technique publiée par l'European Broadcasting Union, Genève, daté Novembre 2014  
<https://tech.ebu.ch/docs/tech/tech3326.pdf>

[13] **CHAFE, C. CACERES, J-P. GUREVITCH, M.** « Effects en temporal separation on synchronization in rhythmic performance », *Perception*, 2010, vol. 39, pg 982-992  
<https://ccrma.stanford.edu/~cc/pub/pdf/temporalSep.pdf> )

[14] **HYUNG-TAEK L. HERRSCHER, D. WALT, J W. CHAARI, F.**, *Performance Analysis of the IEEE 802.1 Ethernet Audio/Video Bridging Standard*, publié par Simutools en Italie, 2012  
<http://www.ieee.org/>

Consultations multi-URL, sites des constructeurs, sites de ressources

Base de Données Cisco :

*Le site internet de Cisco, constructeur de matériel informatique pour le réseau, abonde de documentation précise.*

[15] **Cisco Systems**, Cisco IOS IP Configuration Guide, Release 12.2  
[http://www.cisco.com/c/en/us/td/docs/ios/12\\_2/ip/configuration/guide/fipr\\_c.pdf](http://www.cisco.com/c/en/us/td/docs/ios/12_2/ip/configuration/guide/fipr_c.pdf)

[16] Consultation du site de Ravenna :  
<http://ravenna.alcnetworx.com/>

[17] Consultation du site de Merging Technologies :  
<http://www.merging.com/support/>

[18] Consultation du site d'Audinate :  
<https://www.audinate.com/resources>

[19] Consultation du site d'Axia :  
<http://www.telosalliance.com/Axia>

[20] Consultation de la documentation du site d'Apple :  
[https://developer.apple.com/library/mac/documentation/MacOSX/Conceptual/OSX\\_Technology\\_Overview/About/About.html#//apple\\_ref/doc/uid/TP40001067-CH204-TPXREF101](https://developer.apple.com/library/mac/documentation/MacOSX/Conceptual/OSX_Technology_Overview/About/About.html#//apple_ref/doc/uid/TP40001067-CH204-TPXREF101)

[21] Consultation des Fiches techniques du site de MOTU :  
<http://www.motu.com/techsupport>

## TABLE DES ILLUSTRATIONS

Connecteur ADAT :

[22] Dave MALHAM, Secrets of the Adat interface, mis à jour le 20 janvier 2000  
<http://www.dmalham.freemove.co.uk/adat.html>

Image du datagramme IP :

[23] ZAFRANI, Sami, Establishing and Running Large Scale Internet Based (TCP/IP) Services, (dir netanya college  
<http://mars.netanya.ac.il/~unesco/cdrom/booklet/HTML/NETWORKING/node020.html>

## PARTIE PRATIQUE

[24] [ELLIQ, Mohammed, DELHAISE, Alain, « Conducteurs et Câbles en Audio \(Partie 1\) »](#),

# ANNEXES

## **1. LISTE MATÉRIEL STUDIO LA BEAUDELIE**

### **1.1. SALLE**

#### **Enceintes**

- 6 enceintes Cabasse modèle Antigua MT32
- 2 enceintes Focal 716V colonne
- Un subwoofer M-Audio BX 10S

#### **Amplificateurs**

- Pioneer VSX 923
- T.Amp S75 MK2

#### **Autres**

- Projecteur Epson full HD
- Lecteur Blu-Ray Panasonic DMP BDT 161
- Console Lumière Starville DJ-X16

### **1.2. STUDIO**

#### **Enceintes**

- KEF Concerto SP1006, 8 Ohm, 30W

#### **Amplificateur**

- Rotel RA 162

#### **Microphones**

- 2 x Audix TR40 Omnidirectionnels
- 2 x AKG C414B commutables omni, infracardio, cardio, hypercardio ou bidirectionnel
- 2 x Neumann KM185 hypercardioïdes
- 2 x Audiotecnica AT4040 cardioïdes
- 2 x Beyerdynamic M88TGA
- 2 x Jomeek JM37DP
- 2 x T-bone 2SC140
- 1 x microphone sterophonique Audiotecnica AT825

**Casques**

- 2 x Audiotechnica ATHM40 FS
- Sony MDR 7509 HD
- Prodipe

**Divers**

- 2 écrans d'isolation microphonique T-Bone
- Multipaire 20 XLR
- 2 x préamplificateur à Tubes ART Tube MP
- Interface MIDI-USB Midisport UNO
- Interface USB Dual pre project series

**Instruments**

- Clavier XBoard61 de E-Mu à 61 touches
- Piano Bernstein modèle M, muni d'une 3ème pédale, sortie MIDI

### **1.3. RÉGIE**

**Ecoutes**

- M-Audio BX8a préamplifiées
- Yamaha BS50M

**Cartes son**

- MOTU 896 MK3, (en firewire) 8 analog, 2 ADAT, SPDIF et AES
- MOTU Traveler

**Console**

- M-Audio Project Mix I/O (connecté en Firewire. Possède des sorties SPDIF ou ADAT
  - o 16 « ins » simultanés : 8 mic / 8 ligne
  - o 12 sorties simultanées : 4 ligne

**Autres**

- Préamplificateur et compresseur Twin-Q Jomeek

**Software (sous Mac)**

- Digital performer 7
- Mach five 2
- Pro Tools 10
- Final Cut Studio

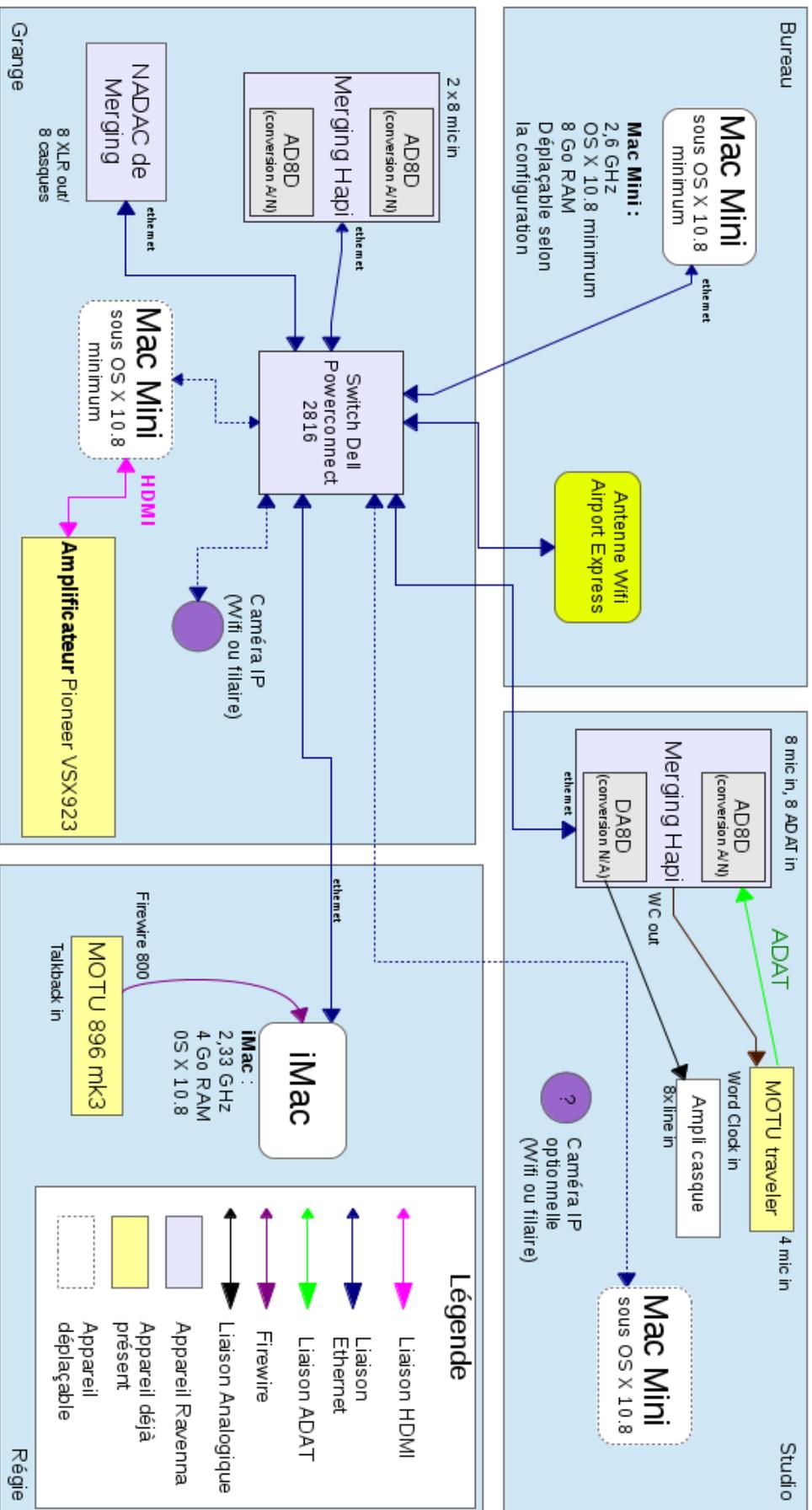
**Ordinateur iMac, 20 pouces (acheté en 2008) :**

- 2.66GHz Intel Core 2 Duo 065-7619
- 4GB 800MHz DDR2 SDRAM
- disque de 320GB en Serial ATA
- 3 ports USB 2.0
- un port Firewire 400
- un port Firewire 800
- carte réseau Gigabit Ethernet

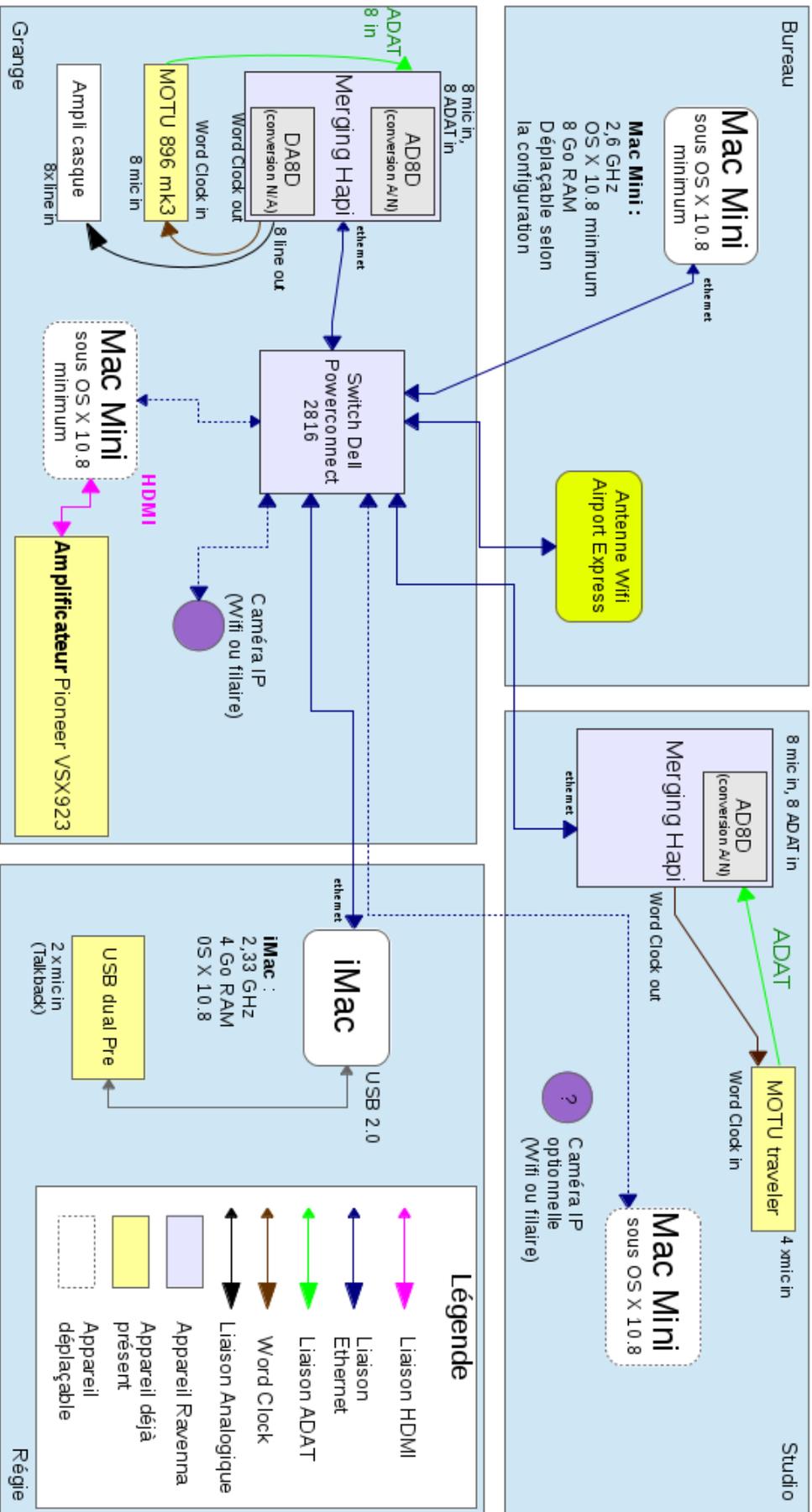
**Système d'exploitation :**

- OS X 10.8

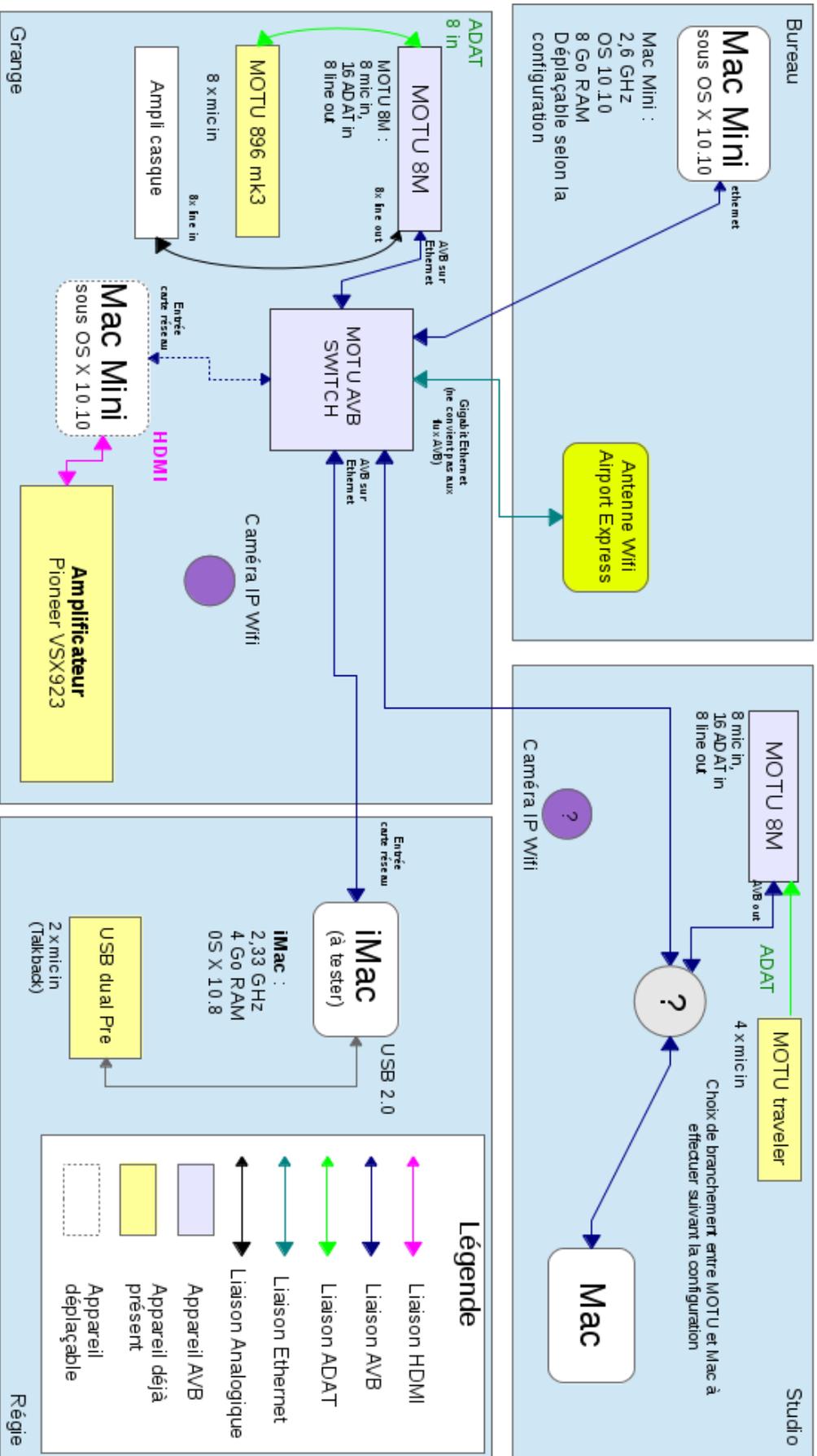
## Annexe 2 : Synoptique Ravenna solution complète – studio la Beaudelie



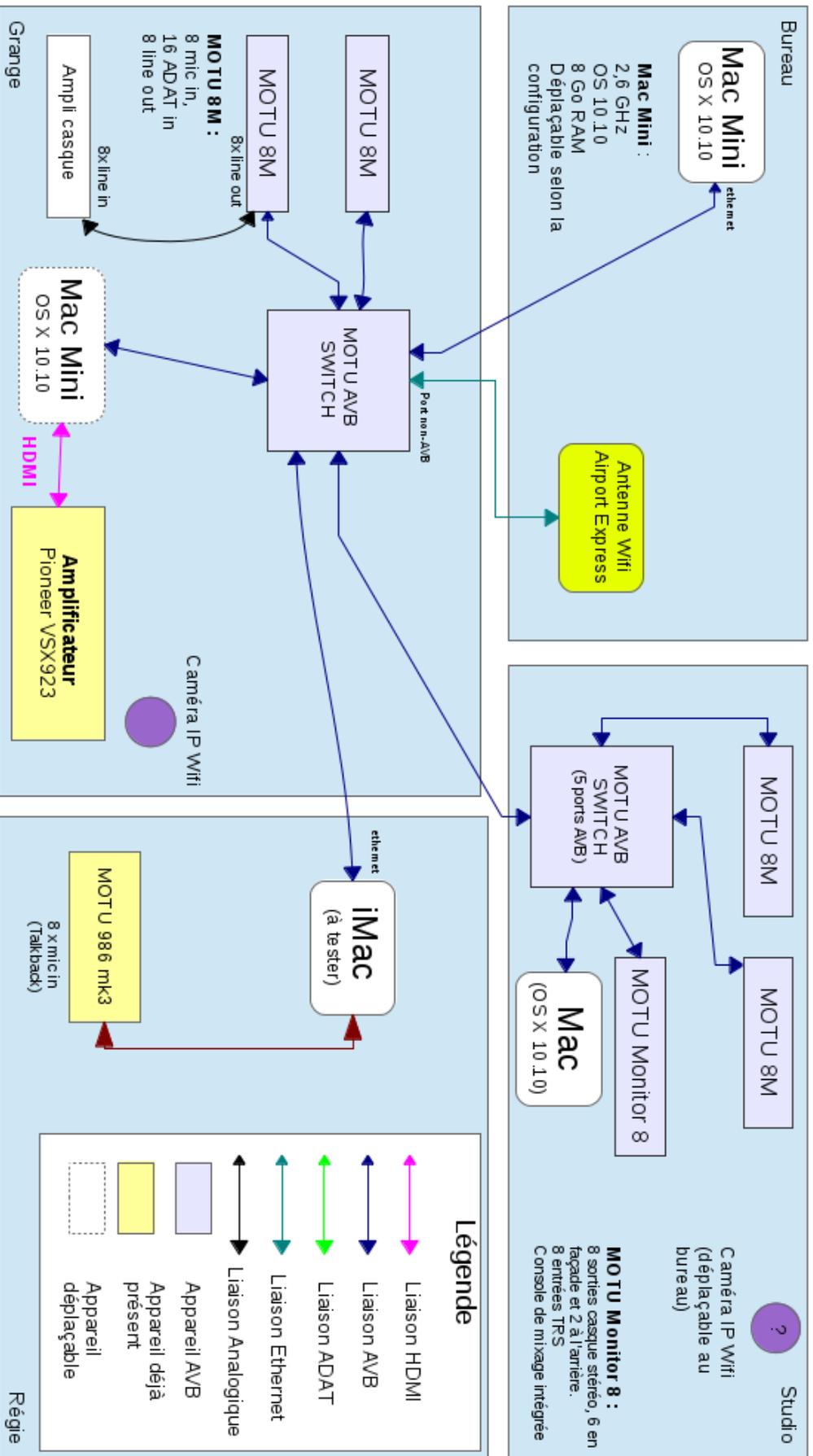
### Annexe 3: Synoptique Ravenna, solution "legacy" – studio la Beaudelle



### Annexe 4 : Synoptique AVB solution "legacy" – studio la Beaudelle



## Annexe 5 : Synoptique AVB solution complète – studio la Beaudelie



Choix de branchement entre une MOTU 8M et Mac Mini à effectuer selon la configuration requise

## 6. BUDGET RAVENNA PARTIEL

Dénomination	Quantité	Prix unitaire	Total
Câble ethernet RJ45 catégorie 6 blindé gris (10 m)	5	9,90	49,50
Câble ethernet RJ45 catégorie 6 blindé gris (15 m)	1	9,90	9,90
Câble ethernet RJ45 catégorie 6 blindé gris (20 m)	2	24,90	49,80
Pack de 2 x Novodio SmartCam HD + Caméra IP HD 960p Wi-Fi motorisée	1	249,00	249,00
Mac mini i5 bicoeur 2,6 GHz 8 Go 1 To Iris Graphics	1	699,00	699,00
Avid Artist Mix	1	1390,00	1390,00
Netgear READYNAS 104	1	224,90	224,90
Disque dur interne Toshiba DT01ACA200 2 To	2	82,95	165,90
Merging Hapi	2	2290	4580,00
AD8D ou DA8D	3	1475,00	4425,00
Dell Powerconnect 2816	1	240	240,00
rack pour 16 embases	2	5	10,00
snake XLR - db25	3	90	270,00
câble BNC 75 ohm	2	12,00	24,00
Ampli casque SM Pro Audio HP6E	1	145,00	145,00
<b>TOTAL</b>			<b>12532,00</b>

## 7. BUDGET AVB PARTIEL

Dénomination	Quantité	Prix unitaire	Total
Câble ethernet RJ45 catégorie 6 blindé gris (10 m)	5	9,90	49,50
Câble ethernet RJ45 catégorie 6 blindé gris (15 m)	1	9,90	9,90
Câble ethernet RJ45 catégorie 6 blindé gris (20 m)	2	24,90	49,80
Pack de 2 x Novodio SmartCam HD + Caméra IP HD 960p Wi-Fi motorisée	1	249,00	249,00
Mac mini i5 bicoeur 2,6 GHz 8 Go 1 To Iris Graphics	1	699,00	699,00
Avid Artist Mix	1	1390,00	1390,00
Netgear READYNAS 104	1	224,90	224,90
Disque dur interne Toshiba DT01ACA200 2 To	2	82,95	165,90
Motu 8M	2	1549	3098,00
Motu 8Pre USB	1	549,00	549,00
Motu AVB Switch	1	315	315,00
Cordial Câble optique ADAT SPDIF Toslink 0.5m	2	10	20,00
Ampli casque SM Pro Audio HP6E	1	145	145,00
câble BNC 75 ohm	2	12,00	24,00
<b>TOTAL</b>			<b>6989,00</b>