

1. 概要

EtherSound は導入しやすく、高音質なオーディオネットワークを供給するために、従来のテクノロジーを向上させています。EtherSound のプロトコル（特許取得済み）は決定論的でレイテンシーは非常に低く、同期したオーディオ伝送を標準の Ethernet 構造基盤のうえに実現します。

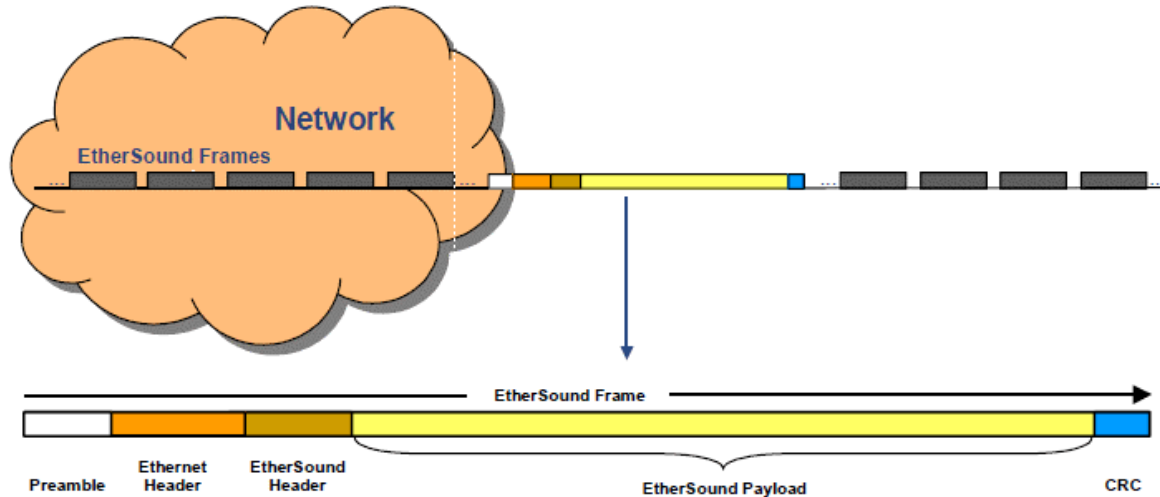
EtherSound はネットワーク化したオーディオ機器を完全にデジタルでつなぎます。100Mbps/秒のネットワークを介して、48kHz/24 ビットの PCM デジタルオーディオとコントロールデータを双方向に最高 64 チャンネル（Total=128Ch）、そして無限ともいえる台数を接続したデバイスに双方向ステータスデータを伝送します。

ネットワーク上のデバイスの台数やデバイス間の距離を増やすために、すぐに使い始められる Ethernet コンポーネント（スイッチャー）やメディアコンバーターと光ファイバーを用いることができます。テスト済み Ethernet コンポーネントのリストは、Digigram のウェブサイトをご覧ください。（ベステックオーディオ㈱にて配布いたしております）

構成とセットアップはアナログシステムよりもずっと簡単です。EtherSound は PA や固定設備、宿泊設備などで必要とされるシステムのコストを大幅に削減します。

2. EtherSound プロトコルについて

Digigram が開発した EtherSound プロトコルは Ethernet フレームに基づいており、EtherSound 対応のデバイス間でのコミュニケーション手段となります。



2.1 EtherSound フレーム

EtherSound フレームは標準の Ethernet フレームに属しており、2つのメインパートから成ります。

- EtherSound ヘッダー： プロトコルに関する重要なインフォメーションをすべて含む。
- EtherSound ペイロード： ネットワークを介して伝送された実際のデータ。

標準の Ethernet フレームと同様、EtherSound ペイロードはパケットの集合体に分けられます。各パケットは次のものから構成されます。

- パケットヘッダー： パケットのタイプとサブタイプが確定される場所。
- パケットデータ： 実際のデータ。

2.2 オーディオクロック

EtherSound ネットワークは独自のクロックを流して完全に同期します。プライマリマスターがネットワークのオーディオクロックを生成し、ダウンストリームにある各 ES デバイスのオーディオクロックはプライマリマスターのオーディオクロックに同期します。

- EtherSound フレームに由来する。
- フェイズを整合する必要があるときはいつでも、分配されたワードクロックに同期する。

クロック分配とタイミングについては第 6 章でご説明いたします。

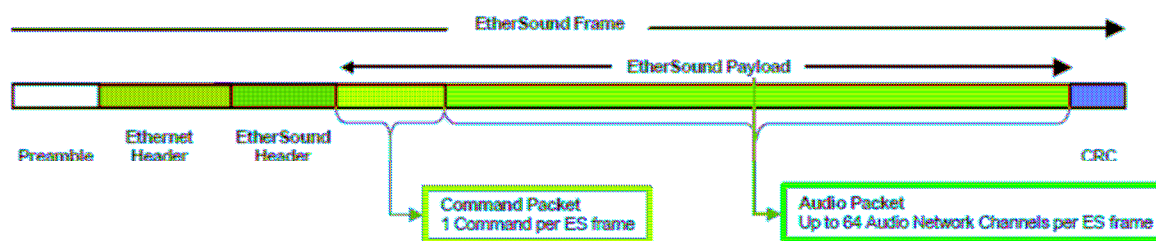
2.3 EtherSound プロトコル V1

現在、採用している EtherSound プロトコル (V1) は 100Mbps ネットワークを基本とし、2つのデータパケットを運搬します。

- ・ コマンドを一つ伝送するコマンドパケット。このコマンドを、コントロール・コマンド(書き)又はステータスリクエスト・コマンド(読み)にすることができる。
- ・ オーディオパケット。44.1kHz 又は 48kHz で 24 ビットのオーディオを最高 64 ストリーム伝送する。

したがって、EtherSound フレームの周波数はサンプリング周波数に等しくなります。

EtherSound V1 のフレームは標準の Ethernet スイッチを使えるよう、IEEE 802.3 に完全に準拠するよう設計されています。



2.3.1 高サンプリングレートに対応

EtherSound フレームに保存されたオーディオデータの柔軟性により、ネットワーク周波数の正確な倍数のサンプリングレートでオーディオストリームのサンプルを伝送するために、複数のネットワークチャンネルを使用できます。

つまり、48kHz の EtherSound ネットワークは 96kHz でオーディオストリームを運搬できるということになります。ネットワークの周波数はそのまま、96kHz のオーディオストリームのオーディオデータが1つではなく2つの EtherSound ネットワークチャンネルを介して保存されます。次のように、さまざまな組み合わせが可能です。

- ・ 48kHz のオーディオストリームが 64ch
- ・ 48kHz のオーディオストリームが 62ch + 96kHz のオーディオストリームを 1ch
- ・ 48kHz のオーディオストリームが 48ch + 96kHz のオーディオストリームを 8ch
- ・ 48kHz のオーディオストリームが 32ch + 96kHz のオーディオストリームを 16ch
- ・ 96kHz のオーディオストリームが 32ch など

2.4 コントロールデータ

デバイスの内部レジスタを読み書きして特定のデバイスをコントロールするために、どの EtherSound フレームにもコントロールデータが組み込まれています。コントロールの機能は第 7 章で詳しくご説明いたします。

書きこむコマンドの最高ビットレートは 768kbits/秒、読みこみコマンドについては 384kbits/秒です。

3. EtherSound デバイスの種類

EtherSound ネットワークでの働き方によって、EtherSound のハードウェアデバイスには 4 種類あります。

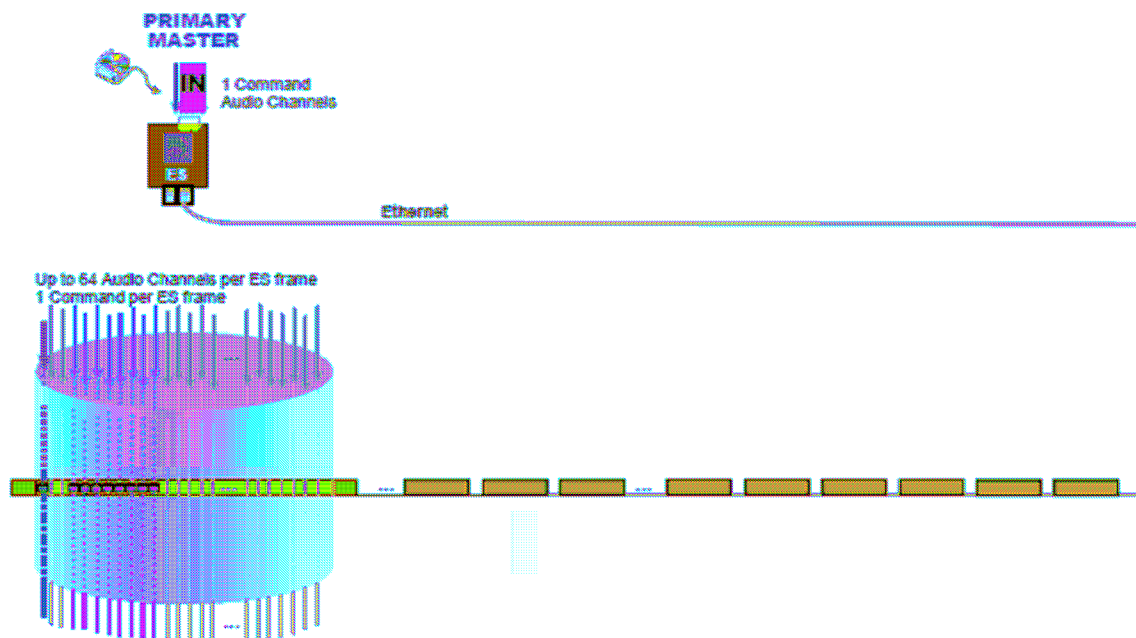
- ・ プライマリーマスター
- ・ マスター
- ・ スレーブ
- ・ マスター／スレーブ

3.1 プライマリーマスター

ネットワークで最初の EtherSound デバイスがプライマリーマスターです。ネットワークへのオーディオソースとなるだけでなく、プライマリーマスターはコマンドとオーディオクロックも伝送します。

すなわち、プライマリーマスターは EtherSound フレームをイニシャライズして構築し、オーディオデータでこれらのフレームを刺激する最初のデバイスです。“Ethernet IN”のポートをホストコンピュータ(PC)へ接続すると、ソフトウェアから EtherSound ネットワーク全体をリモートコントロールし、管理することができます。

プライマリーマスターに問題が生じた場合、そのネットワーク内で次に位置するマスターデバイスが自動的にプライマリーマスターになります。問題がなくなると、システムは自動的にもとの状態に戻ります。



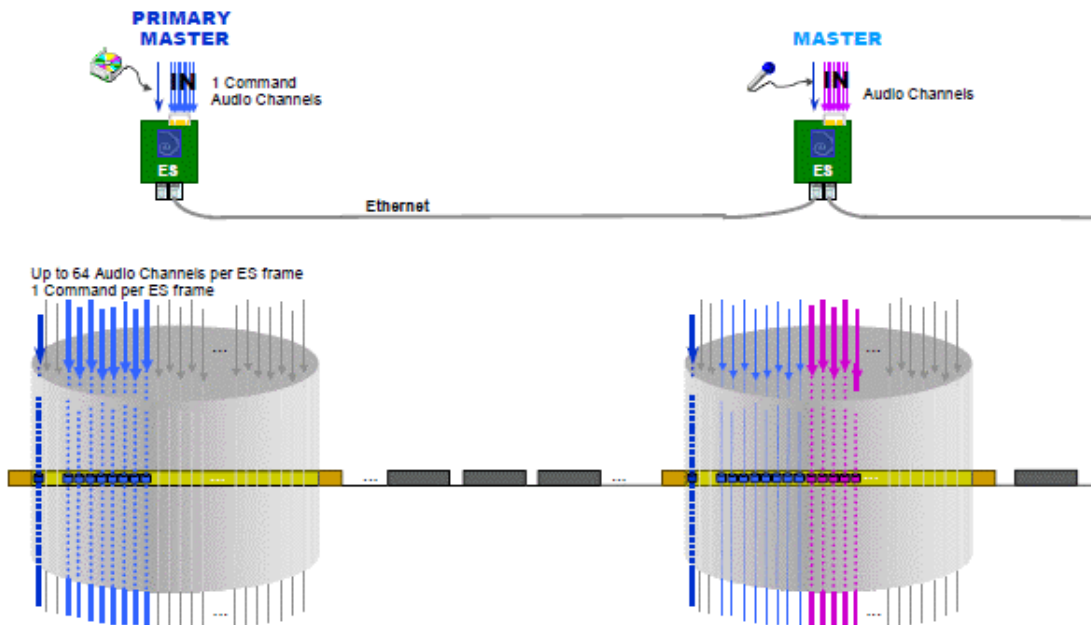
2つのデータの流は EtherSound の所有する規則によって決められます。

プライマリーマスターによって作られた通信データフローをダウンストリームと呼び、これは EtherSound ネットワークのいかなるデバイスでも有効です。

そして、プライマリーマスターに送り返されたユニキャストのデータフローをアップストリームと呼びます。

3.2 マスター

ネットワーク内でプライマリマスターよりダウンストリームにあり、EtherSound のストリームにオーディオチャンネルを追加する役目を果たすデバイスをマスターと呼びます。

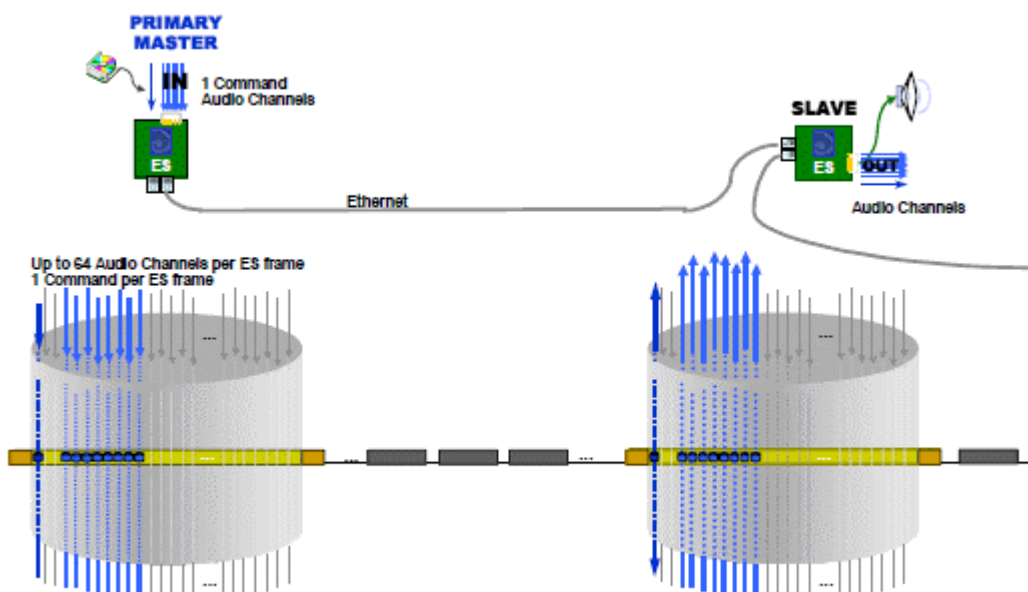


マスターデバイスはプライマリマスターのステータスリクエストとコマンドに応答します。

双方向のループで使用した場合、オーディオデータをダウンストリームに書き込むかアップストリームに書き込むかの設定を決めます。

3.3 スレーブ

EtherSound ストリームを受信し、スタンダードオーディオを復元する EtherSound デバイスがスレーブです。スレーブデバイスはプライマリマスターのステータスリクエストとコマンドに応答します。



双方向ループ内で使用しているときは、オーディオデータをダウンストリームから読み取るかアップストリームから読むかを設定が決めます。

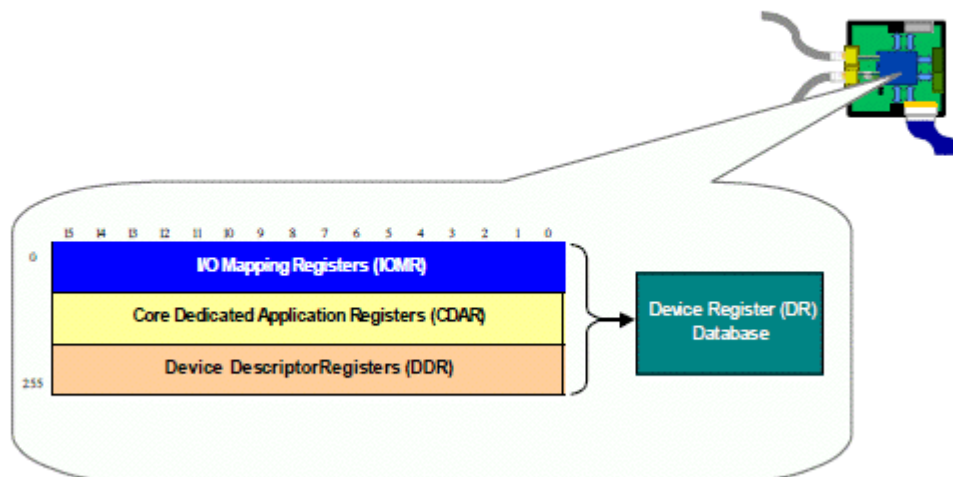
3.4 マスター／スレーブ

マスター／スレーブデバイスは両方の機能を持っており、EtherSound ネットワークからオーディオデータを読み、EtherSound ネットワークへオーディオデータを書き込みます。マスター／スレーブデバイスはプライマリマスターのステータスリクエストとコマンドに応答します。

4. EtherSound のレジスタ

各 EtherSound の FPGA ファームウェアは、256 の 16bit デバイス・レジスタ (DR) のデータベースを内蔵しています。これらのデバイス・レジスタ (DR データベースとも呼ばれる) は、構成情報ステータスを供給したり、EtherSound デバイスの動きを管理したりします。DR には 3 つのカテゴリーがあります。

- DDR: デバイス・ディスクリプター・レジスタ (48レジスタ)。デバイスの EtherSound カーネルのステータスと構成を表します。全 EtherSound デバイスにこのレジスタが入っています。
- IOMR: I/O マッピング・レジスタ (最大 128 レジスタ)。EtherSound ネットワークのチャンネルアサインメントを表す。IOMR の数は EtherSound デバイスのオーディオ容量によります。
- CDAR: コア・デディケイテッド・アプリケーション・パラメータ (最大 80 レジスタ)。CDAR は特定のオプション機能のためにリザーブされており、その機能 (RS232、GPIO、アナログゲイン、 μ コントローラー又は DSP ソフトウェア用のパラメータ) はすべての EtherSound デバイスにデフォルトで入っているわけではありません。



5. EtherSound ネットワークとオーディオの伝送

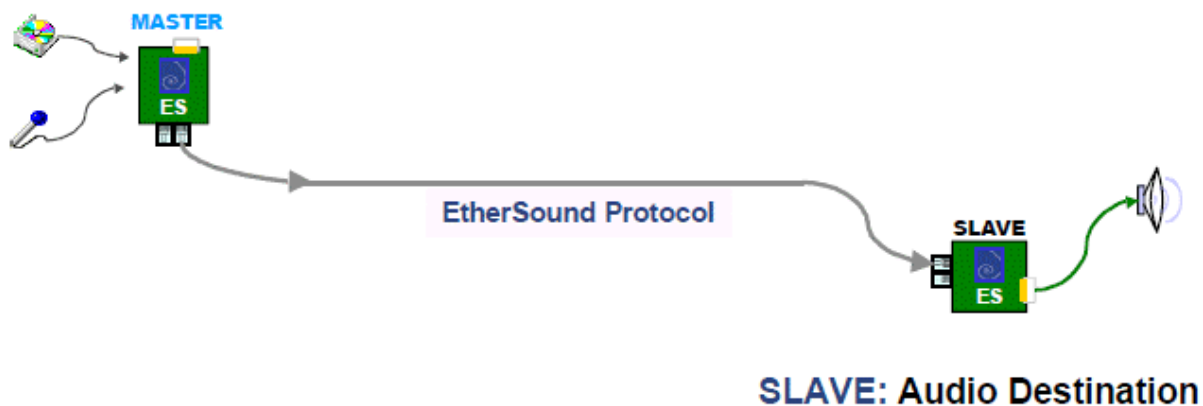
5.1 Ethernet/IEEE 802.3 互換性

EtherSound 技術は標準の IEEE 802.3x に互換性があり、フルデュプレックス・スイッチドのファスト Ethernet ネットワークで動作します。データは大抵、100Mbps/秒 (100Base Tx)フルデュプレックスの最小バンド幅で専用の LAN(ローカル・エリア・ネットワーク)を介して伝送されます。

5.2 最もシンプルな構成

最も簡単な EtherSound の構成は、一方向にプライマリマスターとスレーブを1台ずつつなげる方法です。

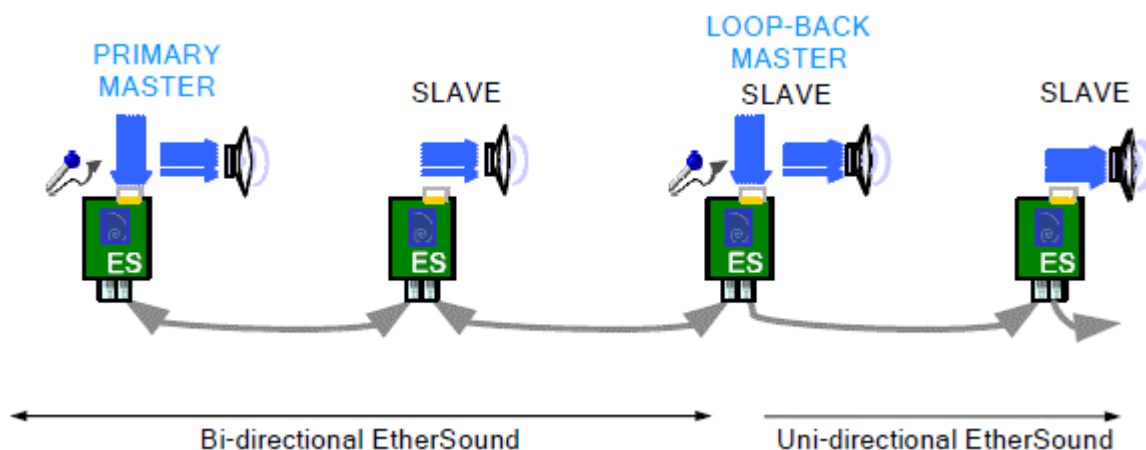
PRIMARY MASTER: Audio Source



この場合、オーディオはダウンストリームへのみ伝送されるので、単一方向になります。

5.3 双方向の EtherSound

双方向の EtherSound とは EtherSound の遂行を意味し、至るところにオーディオが存在するようにセットアップすることをいいます。つまり、どこへでもオーディオをインサートでき、どこからでもオーディオを使え



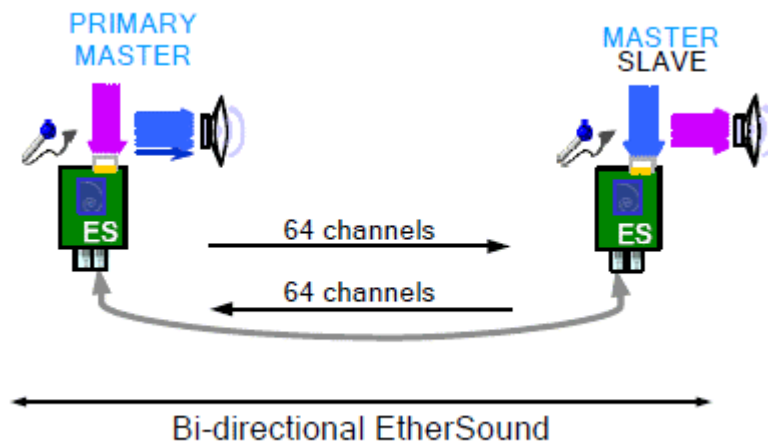
るということです。ダイジェンチェーン内でループバック・デバイスを決めればこの機能を使え、ファスト

Ethernetのフルデュプレックス機能を用いてEtherSoundフレームをプライマリマスターに送り戻します。従って、ダウンストリームに挿入されたどのネットワークチャンネルも、ループバック・デバイスのアップストリームにあるすべてのデバイスで使えます。

ループバック・デバイスのダウンストリームにあるデバイスは、単一方向のオーディオしか供給されません。よって、追加されたオーディオはダウンストリームのデバイスでのみ有効です。

5.3.1 双方向にハイスピード、ハイキャパシティのブリッジを実現

双方向 EtherSound の特定のアプリケーションは、2 台のオーディオデバイス間でハイスピード、ハイキャパシティブリッジで構成します。細かく言いますと、プライマリマスターがダウンストリームへオーディオデータを書きこみ、このオーディオデータは、アップストリームにオーディオデータを書き込むマスター／スレーブデバイスで読まれます。この構成では、24 ビット／48kHz のオーディオを 64 チャンネル両方向で使えますので、合わせると 128 チャンネルのキャパシティーがあることになります。



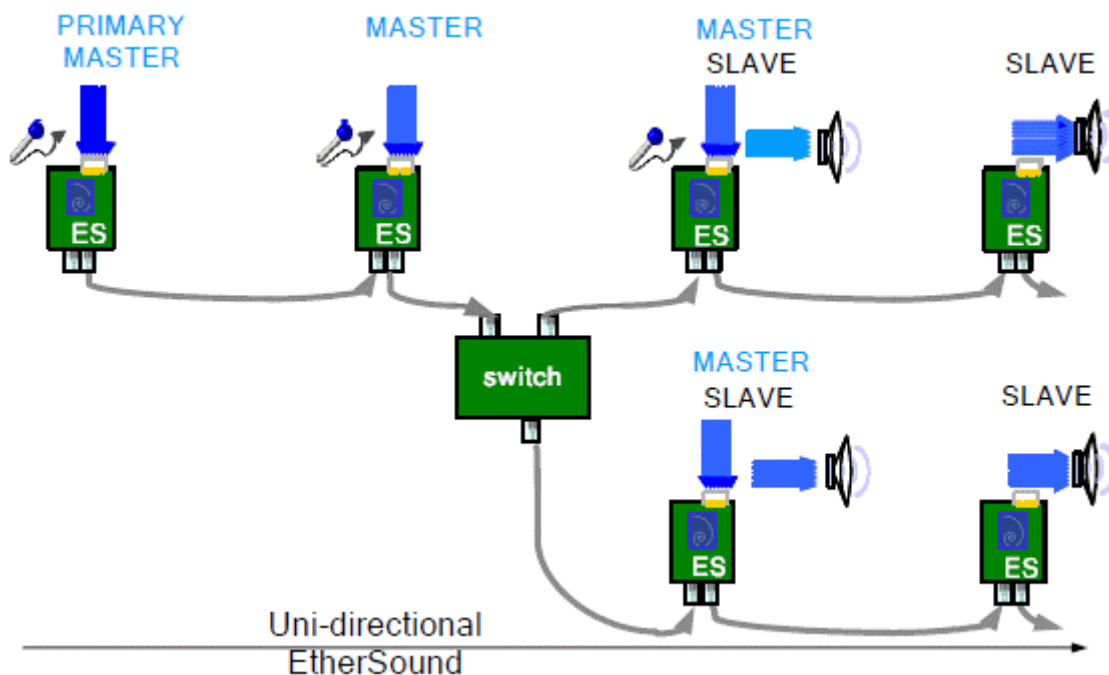
高いサンプルレートで伝送すると、このブリッジは各方向に 96kHz で 32 ストリームのオーディオを扱うこともできます。

5.4 単一方向の EtherSound

一方で、EtherSound ネットワークをシンプルに単一方向にすることもできます。

単一方向の EtherSound はさまざまなネットワーク・トポロジーと互換をもつので、デージーチェーンでもスターでも、両方を組み合わせても使えます。

オーディオの伝送は一方にに限られます。あるオーディオチャンネルが EtherSound ストリームの 1 台のデバイスによって“引き抜かれた”とき、そのチャンネルはマスターデバイスがこのチャンネルにオーディオを挿入するまで、その先のダウンストリームにあるスレーブデバイスすべてに有効であり続けます。



新しいオーディオチャンネルとしてネットワークへ影響を与える前に、DSP プロセッシングとオーディオミキシングを適用するようにするため、この例ではマスター／スレーブデバイスがネットワークからオーディオチャンネルを引き出して構成されています。

各デジチェーンのマスター／スレーブによってインサートされたオーディオデータは、ダウンストリームに位置するスレーブデバイスでのみ使用可能です。

5.5 ネットワークの接続

Ethernet 規格のアンシールド・ツイステッド・ペア (UTP) CAT5 ケーブルを使えば、ハードウェアを追加することなく、無限といっても過言ではない台数のデバイスをデジチェーンできます。チェーンにした最初と最後のデバイス間の距離に制限はなく、各デバイス間は最高で 100m にすることが可能です。長距離の場合にはメディアコンバーターや光ファイバーを用いてカバーします。

一般の Ethernet スイッチを使って、ネットワーク化したオーディオデバイスへ EtherSound を単一方向で伝送分配したり、さらに複雑なネットワーク構造にしたり、デバイス間の距離を長くしたりすることもできます。

ネットワーク構造についてのドキュメントと試験済みコンポーネントのリストは、Digigram ウェブサイトからダウンロードが可能です。

6. クロックの分配とタイミング

6.1 オーディオクロック

EtherSound デバイスのオーディオクロックは、ネットワークから引き出すことも組み入れることもできます。ネットワーククロックは EtherSound ネットワークで最初に接続したデバイス(=プライマリマスター。詳しい説明は次章)が生成します。

6.1.2 オーディオクロックを引き出したネットワーク

入ってくる EtherSound フレームからオーディオクロックを引き出します。



組み込まれた PLL は、オーディオクロックジッターを低くします。

全デバイスはプライマリマスターに同期します。伝搬遅延によってフェイズが変わります。

6.1.3 外部クロック

必要とされる何台かのデバイスのインプットとアウトプットでフェイズが正しく同期していれば、外部の同期機能が使えます。ネットワーク構造とサンプリング周波数によって、連続する最高で 8 台のデバイスを同期させ、フェイズをそろえられます。もっと多くのデバイスが接続されている場合には、続くデバイスが 1 サンプルずつレイテンシーを加えます。

明白な理由から、外部の同期クロックはプライマリマスターと同期していなければなりません。

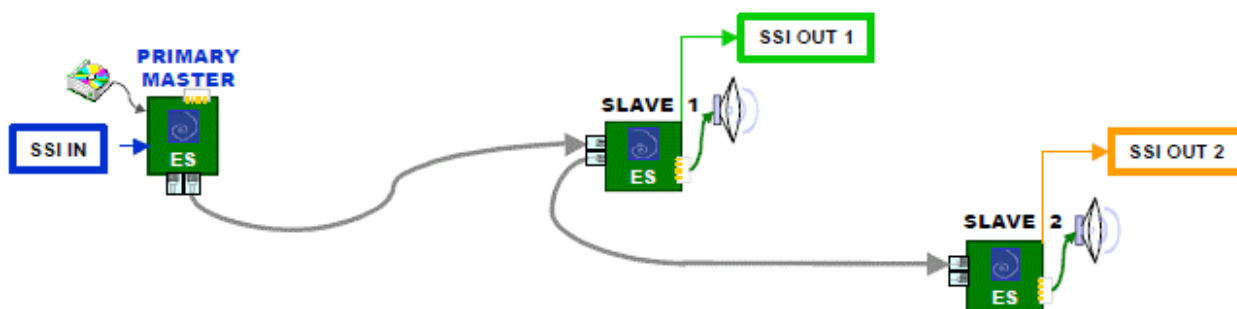
6.2 タイミングとレイテンシー(参考例)

事実として:

- ・ 一般的な端から端までの EtherSound 伝送時間 (SSI インから SSI アウト) は 6 サンプルである。
- ・ ネットワークの出力と入力のバッファリングのため、デジチェーン接続の場合はオーディオをダウンストリームに流す EtherSound デバイス 1 台ごとに 1.6μ 秒以下のレイテンシーが加わる。

6.2.1 標準のデジチェーンを使用した例

ここでの例には A/D と D/A コンバーターによるタイミングは含まれません。



48kHz で 100m、デジチェーンしたネットワークの場合：

- SSIOUT1 が持つレイテンシーは、 $125 \mu\text{秒}$ ([6Sample@48kHz](#)) + $1.5 \mu\text{秒}$ ($\pm 0.1 \mu\text{秒}$) + $0.48 \mu\text{秒}$ (100m のケーブル) = $127 \mu\text{秒}$ ($\pm 0.1 \mu\text{秒}$)
- SSIOUT2 が持つレイテンシーは、 $125 \mu\text{秒}$ ([6Sample@48kHz](#)) + $2 \times 1.5 \mu\text{秒}$ ($\pm 0.1 \mu\text{秒}$) + $2 \times 0.48 \mu\text{秒}$ (100m のケーブル) = $129 \mu\text{秒}$ ($\pm 0.2 \mu\text{秒}$)
- SSIOUT1 と SSIOUT2 間の時間差は $2 \mu\text{秒}$ ($\pm 0.1 \mu\text{秒}$) で、48kHz で約 10 分の 1 サンプル。ということは、時間差が 1 サンプルになるまでに、およそ 10 台の EtherSound ユニートをデジチェーンできることになる。

加えて、スイッチをクロスさせると、決定論的でないフェイズシフトに変わりうる重大なレイテンシーを招くことがあります。

6.2.2 ワードクロックを同期したデジチェーンの例

オーディオインプットとアウトプットのフェイズをそろえる必要があるならば、EtherSound デバイスにワードクロックを流します。

上記と同様、48kHz でデバイス間が 100m のワードクロックで同期させたデバイスをデジチェーンしたネットワークの場合：

- SSIOUT1 がもつレイテンシーは $125 \mu\text{秒}$ ([6Sample@48kHz](#))
- SSIOUT2 がもつレイテンシーは $125 \mu\text{秒}$ ([6Sample@48kHz](#))
- SSIOUT10 がもつレイテンシーは $146 \mu\text{秒}$ ([7Sample@48kHz](#))

7. ホストポートを使ったリモートコントロール

プライマリーマスターのホストポートから、ネットワークに接続した全デバイスのレジスタへアクセス可能です。

7.1 ESM API を使用した、ソフトウェアによるリモートコントロール

外付けコンピュータにつないだ“Ethernet IN”インターフェイスを使用すると、Ethernet プロトコルによりソフトウェア管理ができます。この方法だと、どのローカルデバイスレジスタへも読み書きが可能です。これはプライマリーマスターでしか使えませんが、どのデバイスのどのレジスタでもアドレスすることができます。

この API は基礎を成すネットワークや下のネットワークの抽出レベルが強いため、Ethernet プロトコルの知識は必要とされません。

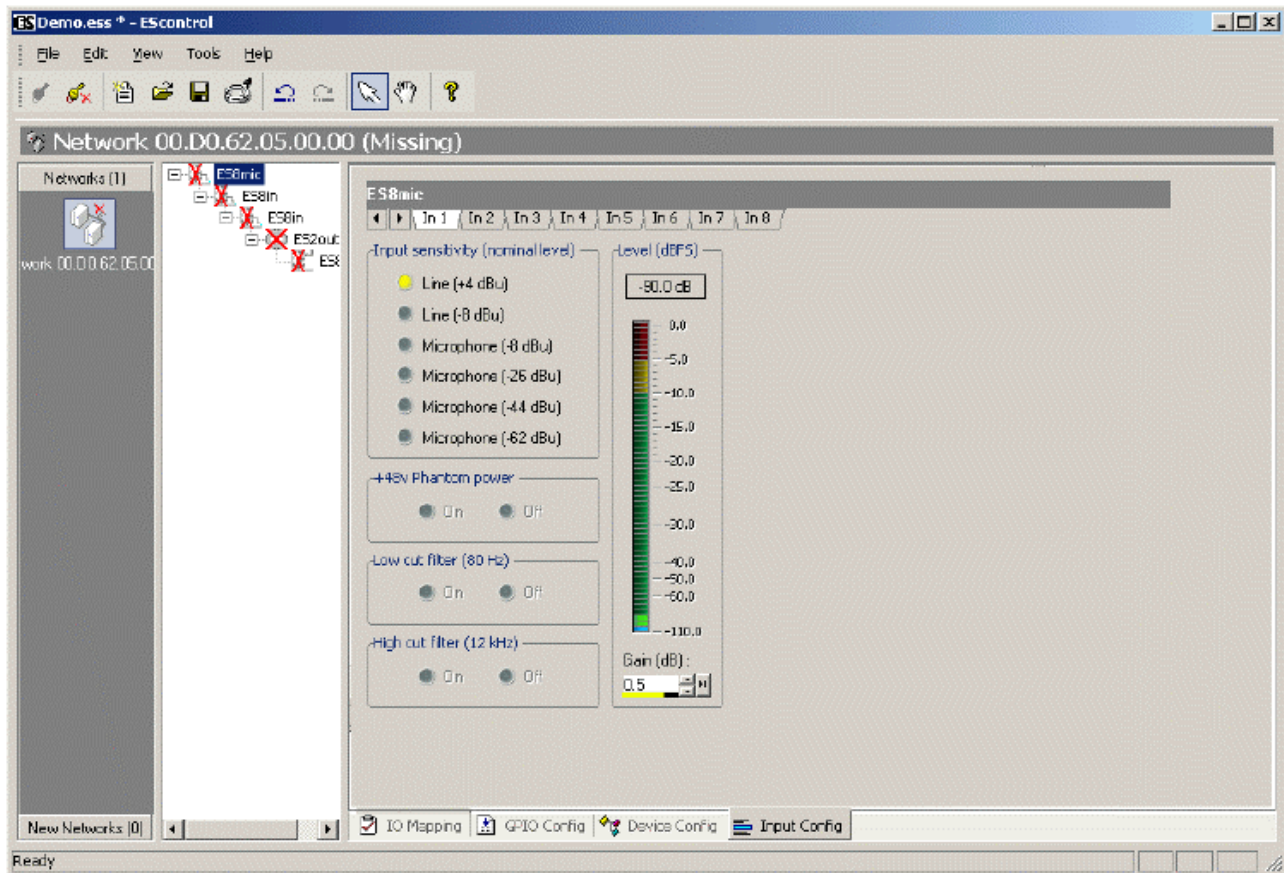
- ・ ネットワーク一覧表
- ・ デバイス一覧表
- ・ マルチベンダー・デバイスのネットワークの管理
- ・ デバイスの認識
- ・ ルーティングのコントロール
- ・ GPIO、RS232 の管理
- ・ 特定の機能へのアクセス

7.2 EScontrol を使った、ソフトウェアによるリモートコントロール

Digigram は API を使ったアプリケーションを開発しました。Microsoft Windows をベースにした EScontrol のグラフィカルインターフェイスは、EtherSound ネットワークの簡単なコントロールを可能にします。

EScontrol がアクセス、コントロールできるもの：

- ・ EtherSound ルーティング
- ・ EtherSound 標準機能の管理 (GPIO、ネットワークエラー)
- ・ 拡張機能がインストールされた場合の EtherSound 製品の特定機能



8. EtherSound テクノロジーの評価

ESnet 評価ボードは、EtherSound テクノロジーをテスト、評価するための、すぐに使えるプラットフォームです。EtherSound ハードウェアは、マスターモード、スレーブモード、あるいはその両方でオペレートするように設計することができます。デバイスは EtherSound ストリームへオーディオチャンネルや再生オーディオチャンネル、もしくはその両方へオーディオチャンネルを送れます。

Digigram の ESnet エバリュエーションキットには ESnet エバリュエーションボード×2 枚、ソフトウェア、設定と最低限の EtherSound 環境について書かれた説明書が含まれます。

十分にテスト、評価できるように、エバリュエーションボードには様々なコネクションが含まれています。

9. 付録 A: 用語解説

Broadcast: すべての機材に対して同時に信号を送る構成。

CDAR: コア・デディケイテッド・アプリケーション・レジスタの略。これら EtherSound レジスタは Digigram、あるいはユーザーが特定した周辺機器機能 (RS232、GPIO、アナログゲインなど) のためにリザーブされている。

DAIO: デディケイテッド・アプリケーション・インプット・アウトプットの略。Digigram リファレンスデザインのなかで有効な 10 ピンのセットで、FPGA の特別なロジックプログラミングが必要。

デージーチェーン: 全デバイスがお互いに直列でリンクされているネットワークポロジ。

DDR: デバイス・ディスクリプター・レジスタの略。EtherSound カーネルのステータスと構成を表す。

Ethernet: 最もよく使われるローカル・エリア・ネットワーク(LAN)。元来 デジタル・イクイップメント・コーポレーション、インテル・コーポレーション、ゼロックスコーポレーション(DIX)で開発され、IEEE コミッティで標準化後、ISO が規格化。標準のコンポーネントはすぐ使いはじめられ、リーズナブルな価格であるために広く使われている。

FPGA: フィールド・プログラマブル・ゲート・アレー。プログラム可能なデバイスで、ソフトウェアのアップデートとともに柔軟にデザインすることが可能。

フレーム: 明確にしたフォーマットに従って、実在するものとして伝送される文字の 1 セットが 1 フレームとなる。伝送される前に、フレームはフィジカルレベルでのコーディング順序に従う。EtherSound のフレームは基準の Ethernet 802.3 と完全に互換する。

フルデュプレックス: データを同時に送受信する伝送のこと。

GPIO: ジェネラル・パーパス・インプット・アウトプットの略。ESnet MS88 Eeprom 300K Reference Design と ESnet MSx 88 Eeprom 200K Reference Design には 8 つの GPIO がある。リレーやスイッチなどの外付けデバイスをリモートコントロールするために、ロー・ボルテージ・ディフアレンシャル・シグナリング(LVDS)レベルを取り出すか出力するように、完全にこれを構成することが可能。2 種類を切り換えるスイッチがついたマイクを GPIO インプットへ接続し、GPIO を例として使用可能。

ホストポート: アプリケーションボード上の μ コントローラー又は DSP とのコミュニケーション用に、ESnet MS88 Eeprom 300K Reference Design と ESnet MSx 88 Eeprom 200K Reference Design で供給されるインターフェイス。

コミュニケーションは EtherSound デバイスレジスタを通して行われる。

IEEE 802: 情報科学機器を接続するために、コミッティーが 1980 年に設けたスタンダード。IEEE 802.3-4-5 は ISO OSI の規範モデルからのフィジカルとリンクレイヤー(MAC)を表す。異なるフィジカルレイヤーは IEEE 802.2 ノームと連動でき、これはリンクレイヤー(LLC)の上部を意味する。

レイテンシー: システムに加えるときに測定されるインサクション・ディレイ。

マスター: EtherSound ストリームにオーディオチャンネルを追加するプライマリーマスターからダウンストリームにあたるネットワーク内のデバイスがマスターとなる。

マスターはプライマリーマスターのコマンドとステータスリクエストに回答する。マスター／スレーブとプライマリーマスターも参照のこと。

マスター／スレーブ: オーディオチャンネルを与え、引き出せもするように構成できるデバイス。

マスター／スレーブはプライマリーマスターのコマンドとステータスリクエストに回答する。

マルチキャスト: 機材から同時にグループや他機器へ伝送するように組んだ構成。

ネットワークチャンネル: EtherSound フレームのオーディオパケットの初期スロット。EtherSound プロトコル V1 内のオーディオパケットには、64 チャンネルの 24 ビットネットワークが含まれる。

プライマリマスター: ネットワーク内で最初の EtherSound デバイスがプライマリマスター。ネットワークへオーディオを流すソースとなるだけでなく、コマンドとオーディオクロックも送信する。

スレーブ: EtherSound のストリームを受信し、スタンダードオーディオに戻す EtherSound デバイスがスレーブ。

スレーブはプライマリマスターのコマンドとステータスリクエストに回答する。

SSI: シンクロナス・シリアル・インターフェイス。スタンダードオーディオの DAC や ADC とデータを送受信するための、もっとも一般的な方法。このインターフェイスは、8 のデータワイヤー（4 インは 8 チャンネルのアップロードが可能、4OUT は 8 チャンネルのダウンロードが可能）とクロックコントロールの形式で、ESnet MS 88 Eeprom 300K Reference Design と ESnet MSx 88 Eeprom 200K Reference Design に供給されている。

スター: すべてのコミュニケーションを処理している同一ユニット（下図のスイッチ）へ全デバイスが接続されているネットワークポロジ。

ユニキャスト: 1 台の機材からもう 1 台の機材へ伝送する構成。



日本地区 EtherSound システム

MasterDistributor

ベストエックオーディオ㈱