

Telos

2101 SUPERIOR
CLEVELAND, OH.
44114 USA

T. 216.241.7225
F. 216.241.4103

Omnia

Telos Systems Multi-channel Splitting Algorithm for AAC and AAC-LD Encoded Audio

by

Anton Thimet and Joseph A. Zolyak

Revision 1.0, 12/22/01

**© Telos Systems
2101 Superior Avenue
Cleveland, OH 44114**

1.0 Introduction

The Telos Systems Zephyr Xstream Codec offers the end user multiple methods of transmitting MPEG4 AAC (Advanced Audio Coding) and AAC Low Delay (AAC-LD) encoded audio. These include:

- Single ISDN circuit over two B channels at 56 Kbps or 64 Kbps (i.e. 112 Kbps or 128 Kbps),
- Single ISDN circuit over a single B channel at 56 Kbps or 64 Kbps,
- Dual Port V.35 at 56 Kbps or 64 Kbps per port,
- Single Port V.35 at 56, 64, 112, 128, 256, or 384 Kbps over one port,
- AAC Ethernet streaming

This paper covers the method used by the Zephyr Xstream for transmitting and receiving AAC or AAC-LD encoded audio across multiple ISDN B channels or Dual Port V.35 channels. Transmitting an AAC or AAC-LD encoded audio stream over a single physical channel in the Zephyr Xstream **DOES NOT** use the splitting algorithm described in this document. The audio stream is transmitted over a single physical channel in exactly the same format as it was received from the AAC or AAC-LD Encoder.

2.0 Algorithm Overview

A top level block diagram of the Telos Systems Multi-channel Splitting Algorithm for AAC and AAC-LD encoded audio is shown in Figure 1. The term “Splitting”, as it is used in the context of this document, refers to the process of taking an encoded high bandwidth audio stream and transmitting (inverse multiplexing) it over multiple, lower bandwidth physical channels. The term “Unsplitting” refers to the process of receiving (demultiplexing) the multiple physical channels back into the original encoded, high bandwidth audio stream. A “physical channel” is the transmission medium that is used to send and receive encoded audio between a source and destination location. The output of the splitting process is interfaced to external hardware, which in turn is responsible for handling the protocol details associated with the ISDN B channels or V.35 channels.

The underlying hardware architecture of the Zephyr Xstream has only a minor influence on the implementation details of the splitting algorithm. The Encoder and Decoder Processes are each implemented on a separate Motorola DSP56303 processor, while the Splitting and Unsplit Processes are executed on the same Motorola DSP56303 processor. The Enhanced Synchronous Serial Interface ports of the processor are configured to transmit 24-bit wide data streams between the processors, which accounts for why the splitting algorithm is centered on 24-bit wide values. The Encoder, Decoder, Splitting, and Unsplit Processes can be implemented on any type of processor so long as the data is transmitted across multiple physical channels using the format specified in this document.

There are three key elements in the splitting process. The first key element is dividing the original encoded high bit rate audio stream into multiple, lower bit rate audio streams. The second key element is the subdivision of the lower bit rate audio streams into fixed length frames prefaced with a Telos Systems specific frame header. The third key element is to reduce the total aggregate output bit rate of the AAC or AAC-LD Encoder to provide enough space to insert the Telos Systems frame header.

When dividing the original encoded audio stream, the Splitting Process treats each consecutive sequence of 24 bits as a separate value in the multi-channel stream. If the encoded audio is to be transmitted over two ISDN B channels, then every other consecutive 24 bits are assigned to each B channel for transmission. If the encoded audio is to be transmitted over four ISDN B channels, then every other fourth consecutive set of 24 bits is assigned to each B channel for transmission. (See Section 4.0 for further details). Dual Port V.35 channels are treated in the same fashion as if they were two ISDN B channels.

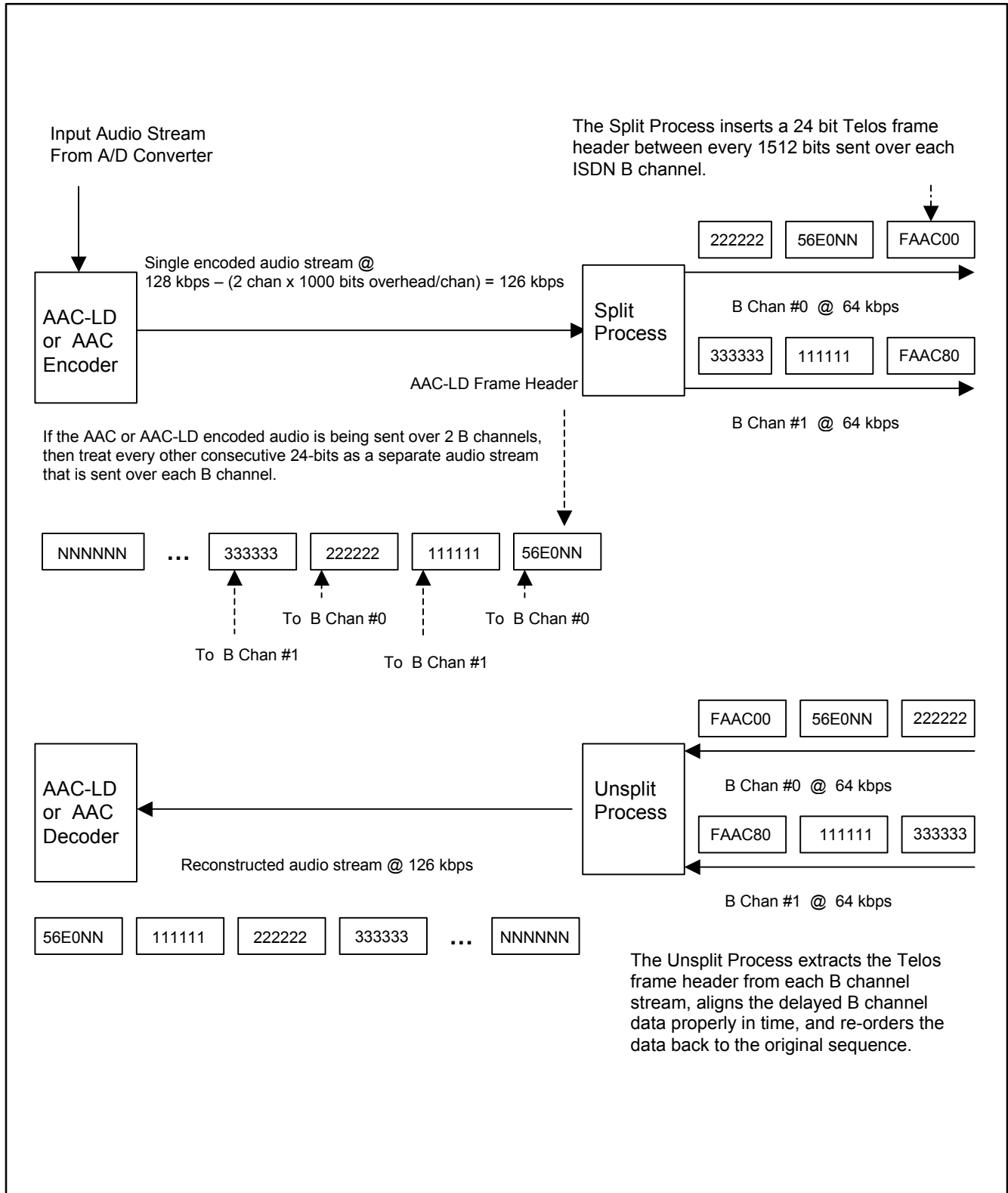


Figure 1 - Top Level Block Diagram of Telos Systems AAC and AAC-LD Splitting Algorithm

The multiple lower bit rate audio streams are subdivided into small, fixed length frames prefaced with a Telos Systems specific frame header. (See Figure 2). At the destination of the multi-channel audio transmission, the Unsplit Process uses the information in the Telos Systems specific frame header to reconstruct the original AAC or AAC-LD Encoded audio stream. The Telos Systems specific frame header contains information that identifies the sequence number of each frame (i.e. Frame Count field), as well as the relative word order in the original encoded stream (i.e. Channel ID number).

The Unsplit Process buffers the multiple audio streams and waits until every channel has received at least one frame with the same Frame Count value. Having located the most recently received frame that is common to each channel in each channel's receive buffer, the Unsplit Process determines the order in which the data is read from all the channel buffers. The Unsplit Process extracts the Telos Systems frame header, then copies the remaining frame data 24 bits at a time from each channel buffer into a single output buffer based on the Channel ID number in order to restore the original audio stream data sequence. (See Section 5.0). The data from this buffer is then sent to the AAC or AAC-LD Decoder.

The last key element is the responsibility of the end user who is configuring the AAC or AAC-LD Encoder. Any audio codec product wishing to transmit data to a Telos Systems Zephyr Xstream must configure the total aggregate bit rate output from the AAC or AAC-LD Encoder to account for the Telos Systems specific frame header. The equation for computing the total aggregate bit rate output from the Encoder is as follows:

$$\text{Total Capacity of combined channels} - (N \text{ channels} * \frac{\text{Overhead Factor}}{\text{channel}}) = \text{Encoder Programmed Bit Rate (in bps)}$$

The Overhead Factor for a 64 Kbps physical channel is 1000 bps/chan, and for a 56 Kbps channel it is 875 bps/chan. This accounts for a 24-bit Telos Systems specific frame header being inserted every 1512 bits. Table I lists the Encoder Programmed bit rates for various types and numbers of physical channels.

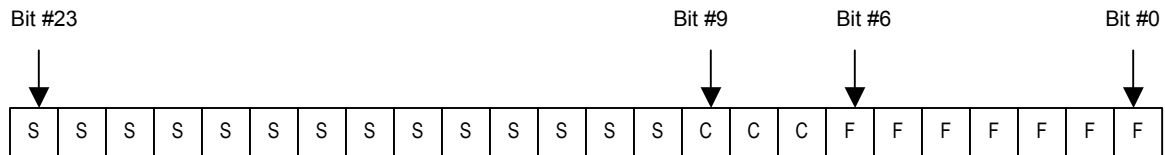
TABLE I

Number of Channels	Bit Rate per Channel	Combined Capacity of Physical Channels	Encoder Programmed Bit Rate
2	56,000 bps	112,000 bps	110,250 bps
2	64,000 bps	128,000 bps	126,000 bps
4	56,000 bps	224,000 bps	220,500 bps
4	64,000 bps	256,000 bps	252,000 bps
6	56,000 bps	336,000 bps	330,750 bps
6	64,000 bps	384,000 bps	378,000 bps

3.0 Algorithm Design Requirements

The Telos Systems Splitting Algorithm for AAC and AAC-LD encoded audio was designed to meet the following self-imposed technical requirements:

- a. The algorithm must work for both AAC and AAC-LD encoded audio using the same method.
- b. Minimize the audio delay attributable to the Splitting/Unsplit Process to less than 10 milliseconds.
- c. Minimize the reduction in channel payload capacity due to algorithm overhead to 1000 bits/sec or less per physical channel.
- d. Support transmission of encoded audio over 2 to 6 ISDN B channels. [NOTE: The Zephyr Xstream platform currently only supports transmission over two ISDN B channels].
- e. Support transmission of encoded audio over Dual Port V.35 channels at either 112 or 128 Kbps.
- f. Support inter-channel delay of transmitted audio with a maximum time difference of arrival up to 1.0 second.
- g. Does not impact or restrict the transmission of Ancillary Data in the AAC encoded audio stream.



- [S] Bits #23 .. #10 → * Frame Sync field in Telos AAC or AAC-LD frame header. Fixed value of “11111010101011” binary.
- [C] Bits #9 .. #7 → * Channel Identification Number field in Telos AAC or AAC-LD frame header.
- * Range is from “0” to “7” decimal. If the audio stream is split across two ISDN B channels, then the first B channel is designated as Channel #0 and the second B channel is designated as Channel #1.
 - * If the data is split across four B channels, then first B channel is Channel #0 and the last B channel is designated as Channel #3.
 - * The Channel ID Number field is used to reconstruct the original sequence of 24-bit values across all the frames with the same Frame Count field value.
- [F] Bits #6 .. #0 → * Frame Count field in Telos AAC or AAC-LD frame header.
- * Range is from “0” to “127” decimal. After the frame count value reaches 127, it wraps around to zero again.
 - * The Frame Count field is used to align the received frame packets across multiple channels with respect to each other in time.

Figure 2 - Telos Systems AAC and AAC-LD Frame Header Format

4.0 Multi-channel Splitting Process

This section outlines the sequence of steps involved in the Splitting Process.

Configure the AAC or AAC-LD Encoder output bit rate to match the capacity of the physical transmission channels minus the overhead bit rate to support the Telos Systems frame header.

Configure the Splitting Process to parse for AAC or AAC-LD encoded audio.

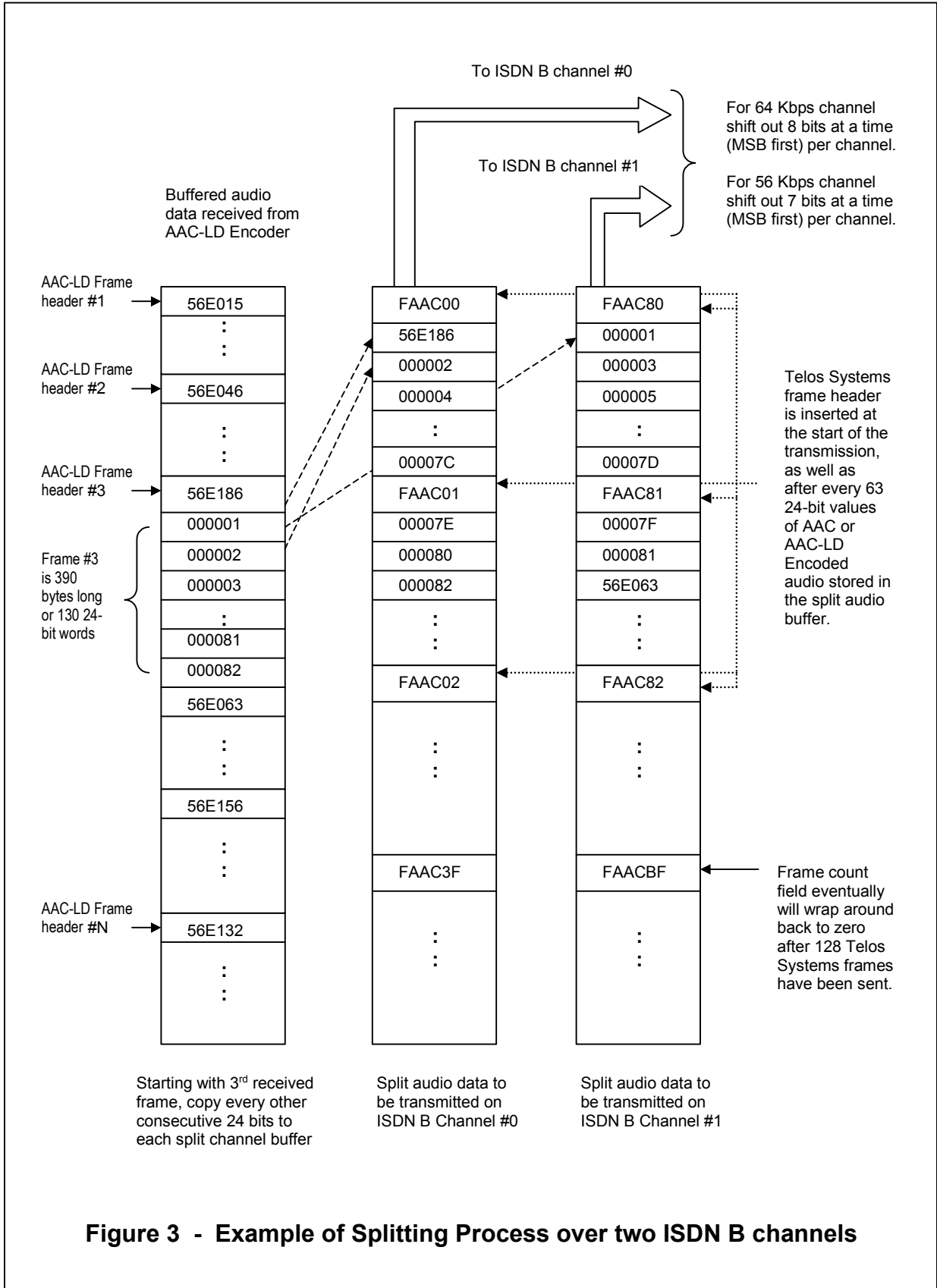
The Splitting Process monitors the encoded audio stream and confirms that it matches the format it expects. The Splitter does not start any data transmission over the physical channels until it detects three frame headers in a row of the expected encoded audio format.

To minimize audio throughput delay, the first two parsed frames are discarded and data transmission starts with the third parsed frame. (It was deemed that there is no useful audio content within the first second of establishing an audio connection with the receiving end, thus the data can be discarded and the third frame which is being received from the Encoder in real-time can be transmitted almost immediately). There is a tradeoff, however, between minimizing the audio throughput delay versus buffering a sufficient number of bits to prevent running out of data to transmit. Because the AAC and AAC-LD frames have variable length, there is the potential that the Splitting Process could run out of data while transmitting the third frame before the next frame arrives. Actually, this could statistically occur between any two consecutive frames. As a result, the Splitter waits until it receives a fixed number of bits beyond detecting the third frame header before it starts data transmission.

Once the Splitting Process has detected the third AAC or AAC-LD frame, it begins the process of dividing the encoded audio stream into multiple, lower bit rate streams based on the number of physical channels that are used to transmit the data. For example, let's assume that the encoded audio is going to be sent across two ISDN B channels at 64 kbps per channel. The Splitting Process allocates memory space for two buffers, one buffer for each B channel. The first 24-bit value written into each buffer is the Telos Systems AAC frame header with the frame count field value set equal to zero. (Refer to Figure 3). Next the AAC or AAC-LD frame header value for the third frame received from the Encoder is stored in the buffer designated as Channel #0 in the Telos Systems frame header. The next consecutive 24 bits of the original encoded audio stream are then copied into the other buffer (designated as Channel #1).

The process of copying every other consecutive set of 24 bits to each buffer continues until sixty-three 24-bit values of the original encoded audio stream has been copied into each buffer. At this point the frame count value is incremented and the next Telos Systems specific frame header value is stored in each respective output buffer. (In general, this step is concurrent with receiving new data from the Encoder and transmitting data over the physical channel). The process of inserting a Telos Systems specific frame header after every sixty-three 24-bit values is the same for 56 Kbps or 64 Kbps physical channels.

Once the Splitter has parsed the third consecutive AAC or AAC-LD frame header, and the necessary number of words has been buffered up to prevent drop outs, the Splitting Process begins transmission over each physical channel. For a 64 Kbps capacity channel, eight bits at a time are taken out of each buffer and transmitted over each corresponding physical channel (most significant bit first). For a 56 Kbps capacity channel, seven bits at a time are taken out of each buffer.



5.0 Multi-channel Unsplit Process

This section outlines the sequence of steps involved in the Unsplit Process.

Configure the Unsplit Process to parse for the Telos Systems frame header. Allocate a memory buffer for each physical channel used to receive data. The size of the buffer is very important to the Unsplit Process. The buffer should be able to store half the total number of frames supported by the Telos Systems frame header. In particular, the size of each buffer should be:

$$0.5 * (128 \text{ max frames} * 64 \text{ words/frame} * 24 \text{ bits/word}) = 98304 \text{ bits}$$

Construct 24-bit values from data received on each 56 Kbps or 64 Kbps physical channel. Store the constructed 24-bit values in the corresponding buffer associated with each physical channel.

Parse the received data from each physical channel and search for the Telos Systems frame header. Create a table with the buffer address and bit shift used to locate the Telos Systems frame header within the buffered data from each physical channel.

Wait until at least two Telos Systems frame headers have detected in each physical channel buffer.

Figures 4 and 5 illustrate the Unsplit Process for two physical channels. (The same concept can be extended for more than two channels). The first step in the Unsplit Process is to compare the Frame Count field in the last Telos Systems frame header detected for each channel. If the current Frame Count field values are equal, then the frames are already synchronized and the buffered data can be sent along to the Decoder once the Telos Systems frame header is extracted and the proper data ordering is determined. (More about this in a moment). Figures 4 and 5 show two examples of when the Frame Count field values are unequal, and how to determine which channel has connected at the receiving side and received data first.

Figure 4 illustrates the situation where the difference in the Frame Count field is less than 64 frames (i.e. the channel buffer size). Physical Channel #1 has connected first and has already five frames worth of data in its buffer (i.e. current Frame Count = 4), while Physical Channel #0 has just received its second frame (i.e. current Frame Count = 1) as shown in this figure. The Unsplit Process computes the difference between the current Frame Count values, and determines that Physical Channel #1 has connected first because it has the larger Frame Count value but the difference is less than 64 frames. The Unsplit Process uses the Frame Count difference to compute a negative offset into the Physical Channel #1 Frame Header Look Up Table to locate the position of the frame with the same Frame Count value as Physical Channel #0 (i.e. Frame Count = 1). The Unsplit Process would follow the same steps if Physical Channel #0 had the larger Frame Count value, but the difference is less than 64 frames, to locate the Physical Channel #0 frame with the matching current Frame Count value as Physical Channel #1.

Figure 5 illustrates the situation where the Frame Count field value for Physical Channel #0 has wrapped back around to zero, thus the difference between the two Frame Count fields is more than or equal to 64 frames. (This can occur when the transmission is disrupted and reconnected on one physical channel). When the Unsplit Process computes the difference between the current Frame Count values and sees the value is 64 frames or greater, the Unsplit Process determines that physical channel with the lower Frame Count field connected first. As shown in Figure 5, Physical Channel #1 has actually connected at the receiving end after Physical Channel #0, even though Frame Count field value (i.e. 127) is greater than the Frame Count field value for Physical Channel #0 (i.e. 1). The Unsplit Process uses the Frame Count difference to compute a negative offset into the Physical Channel #0 Frame Header Look Up Table to locate the position of the frame with the same Frame Count value as Physical Channel #1 (i.e. Frame Count = 127).

Channel #0 Frame Header Lookup Table

Frame Header Start Address	Frame Header Bit Shift
0x40C0	0
0x4100	0

Channel #1 Frame Header Lookup Table

Frame Header Start Address	Frame Header Bit Shift
0x5000	0
0x5040	0
0x5080	0
0x50C0	0
0x5100	0

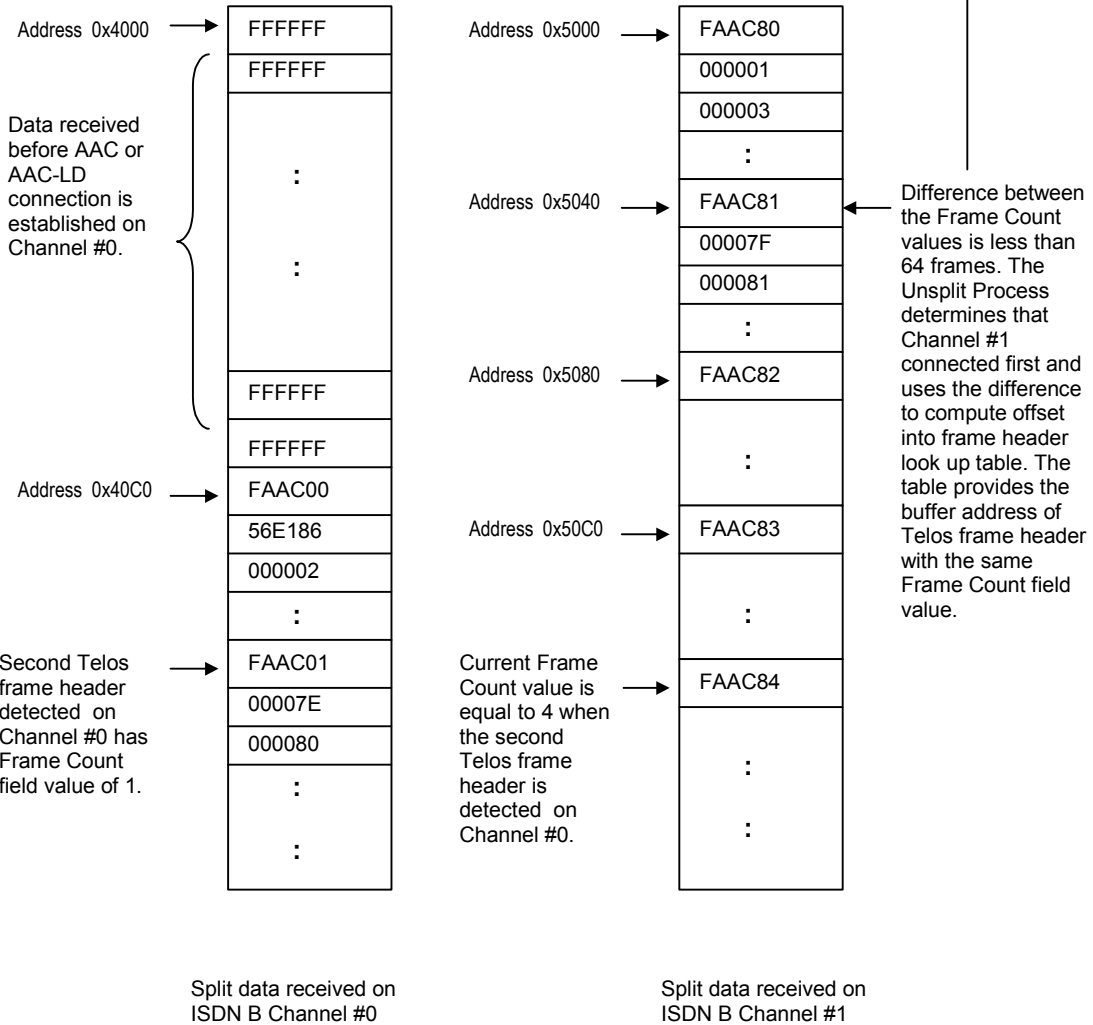


Figure 4 - Unsplit Process example when Frame Count difference < 64 frames

Channel #0 Frame Header Lookup Table

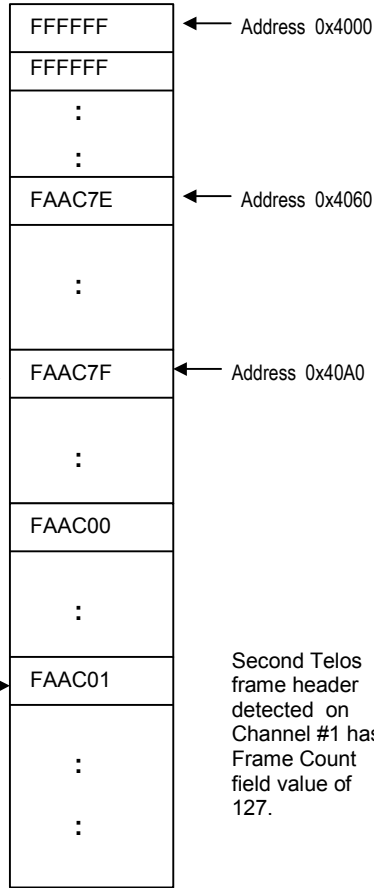
Frame Header Start Address	Frame Header Bit Shift
0x4060	0
0x40A0	0
0x40E0	0
0x4120	0

Channel #1 Frame Header Lookup Table

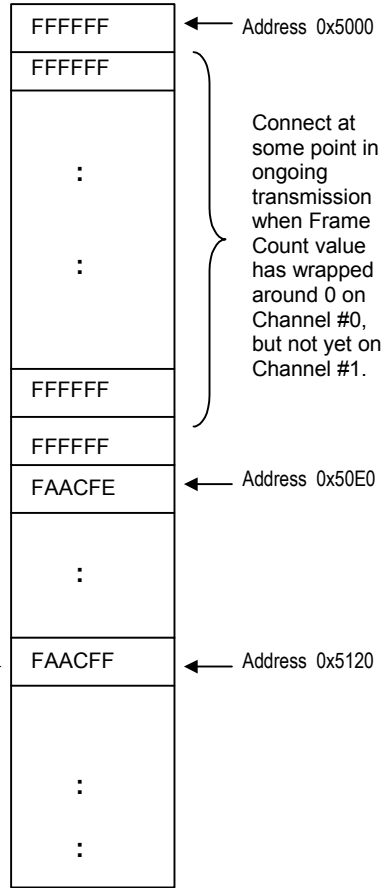
Frame Header Start Address	Frame Header Bit Shift
0x50E0	0
0x5120	0

Difference between the Frame Count values is more than 64 frames. The Unsplit Process determines that Channel #0 connected first and uses the difference to compute offset into frame header look up table. The table provides the buffer address of Telos frame header with the same Frame Count field value.

Current Frame Count value is equal to 1 when the second Telos frame header is detected on Channel #1.



Split data received on ISDN B Channel #0



Split data received on ISDN B Channel #1

Connect at some point in ongoing transmission when Frame Count value has wrapped around 0 on Channel #0, but not yet on Channel #1.

Second Telos frame header detected on Channel #1 has Frame Count field value of 127.

Figure 5 - Unsplit Process example when Frame Count difference > 64 frames

Once the current common frame is located in each physical channel buffer, it is time to reconstruct the original encoded audio stream. The Unsplit Process extracts the Telos Systems specific frame header from the split audio stream in each receive channel buffer, and examines the Channel ID number in the frame header. The Unsplit Process then uses the Channel ID number to restore the original sequence of 24-bit values.

The Unsplit Process copies one 24-bit value at a time from each channel buffer into a separate output buffer which holds the data being sent to the AAC or AAC-LD Decoder. After copying sixty-three 24-bit values from each channel buffer, the next value in each channel buffer should be a Telos Systems specific frame header. If it has a Frame Count field value one increment higher than the previous frame header, then continue copying data from the channel buffer to the common output buffer. If the Telos Systems specific frame header is not found, or it does not have the expected Frame Count field value, then the Unsplit Process must re-synchronize the physical channel data streams again before sending data to the AAC or AAC-LD Decoder.