MULTIMAGIX

AUDIO BASICS

Presenting by Phani

Agenda

- Motivation
- Audio
- Sampling
- Digital Audio Terminologies
- Nyquest theorem
- Need for Codecs
- Different types of Audio codecs
- FFMPEG API’s for Audio Processing

MOTIVATION

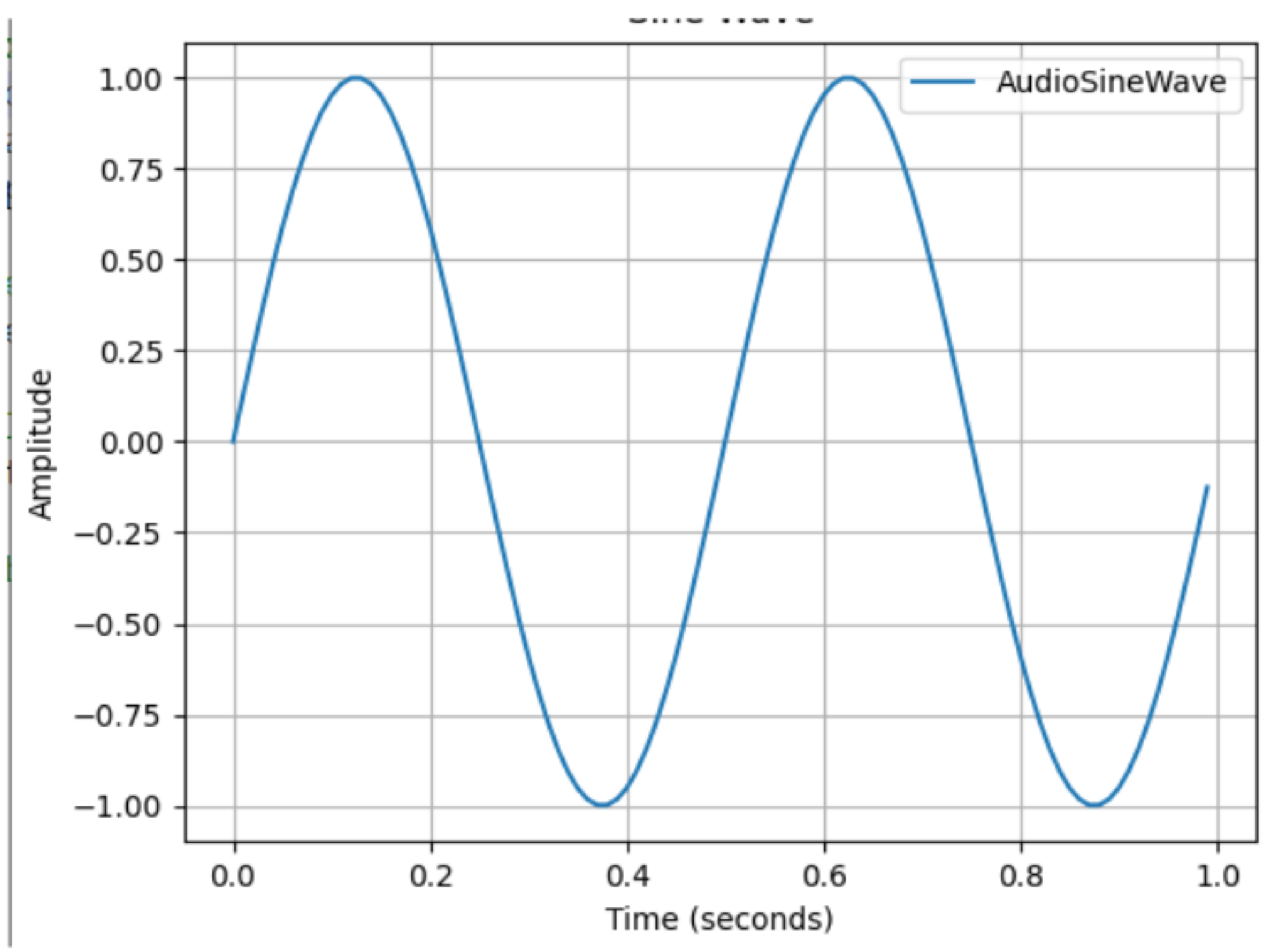
- Feedback
  - From subscribers
- Breaking learning myths on audio/video
  - Minimal domain knowledge
  - More of SW approaches or SW point of view
- Mini projects
  - All challenges upfront
  - Integrate to actual product

AUDIO

All sounds that we hear are just air pressure variations hitting our ears at different rates. The high and low pressure repeating is called sound wave. Sound waves in the air can be turned into an electrical signal called an analog signal using a microphone

AUDIO

- Frequency is Cycles per second and measured in Hertz (Hz).
- Amplitude (db) is the intensity of the wave, often termed as volume

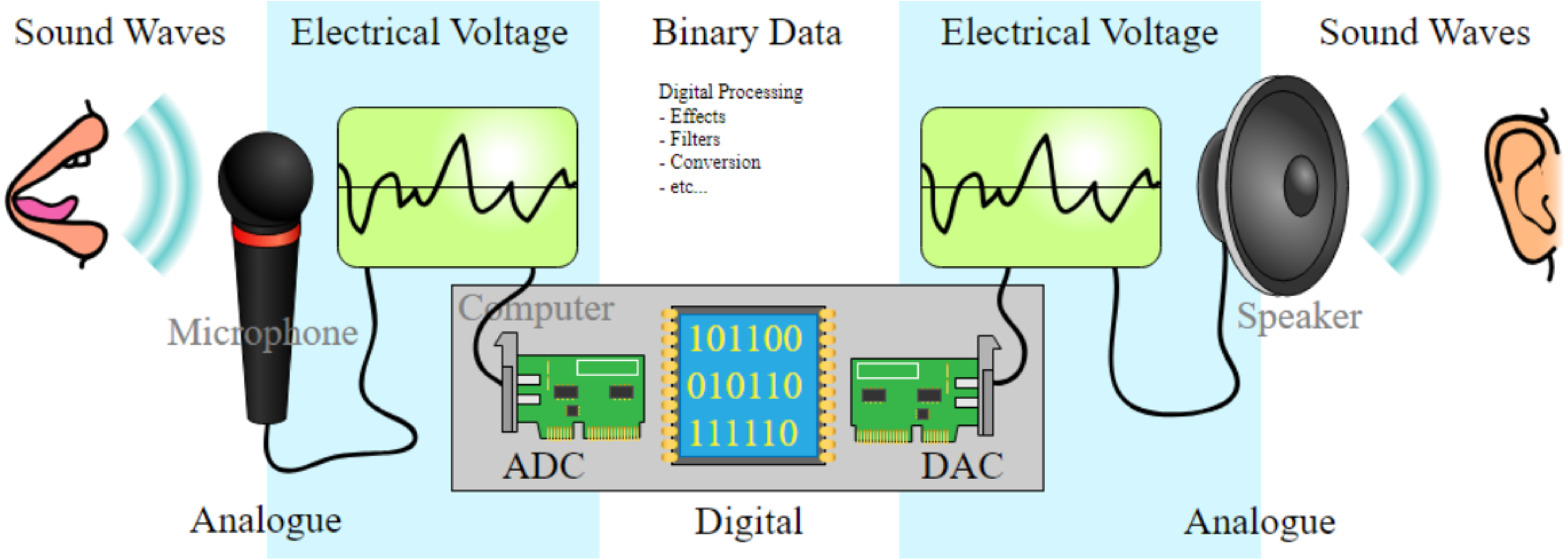


f = 2Hz

AUDIO

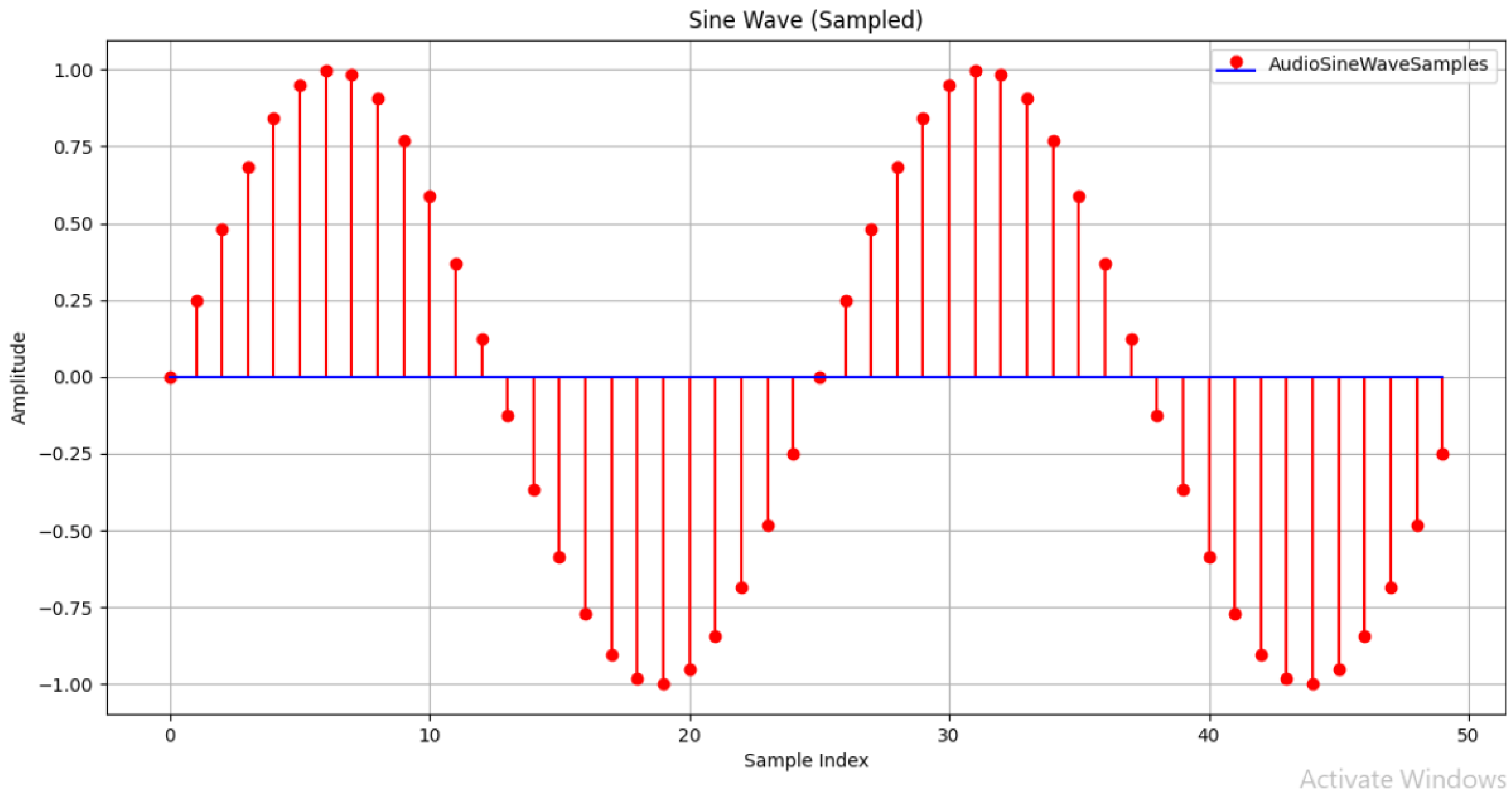
The ADC must accomplish two (2) tasks:

1) sampling
2) quantization



SAMPLING

A sample is the smallest unit to represent sound, it represents a single position of a audio sinewave. The process of taking samples of analog waveform at regularly spaced time intervals is called Sampling



DIGITAL AUDIO TERMINOLOGIES

Sample Rate: The sample rate is how many sample there are per second like the 44.1kHz is a common sample rate. Thats 44,100 samples per second. You are losing sound information by converting the signal from an analog signal to digital signal, **but most people can't hear the difference, which is why most audio work is done digitally**

Bit-Depth: A samples resolution is defined by it's bit-depth which is how many 1's and 0's are required to describe each sample. A higher bit-depth means each individual sample has more possible positions it could be in.

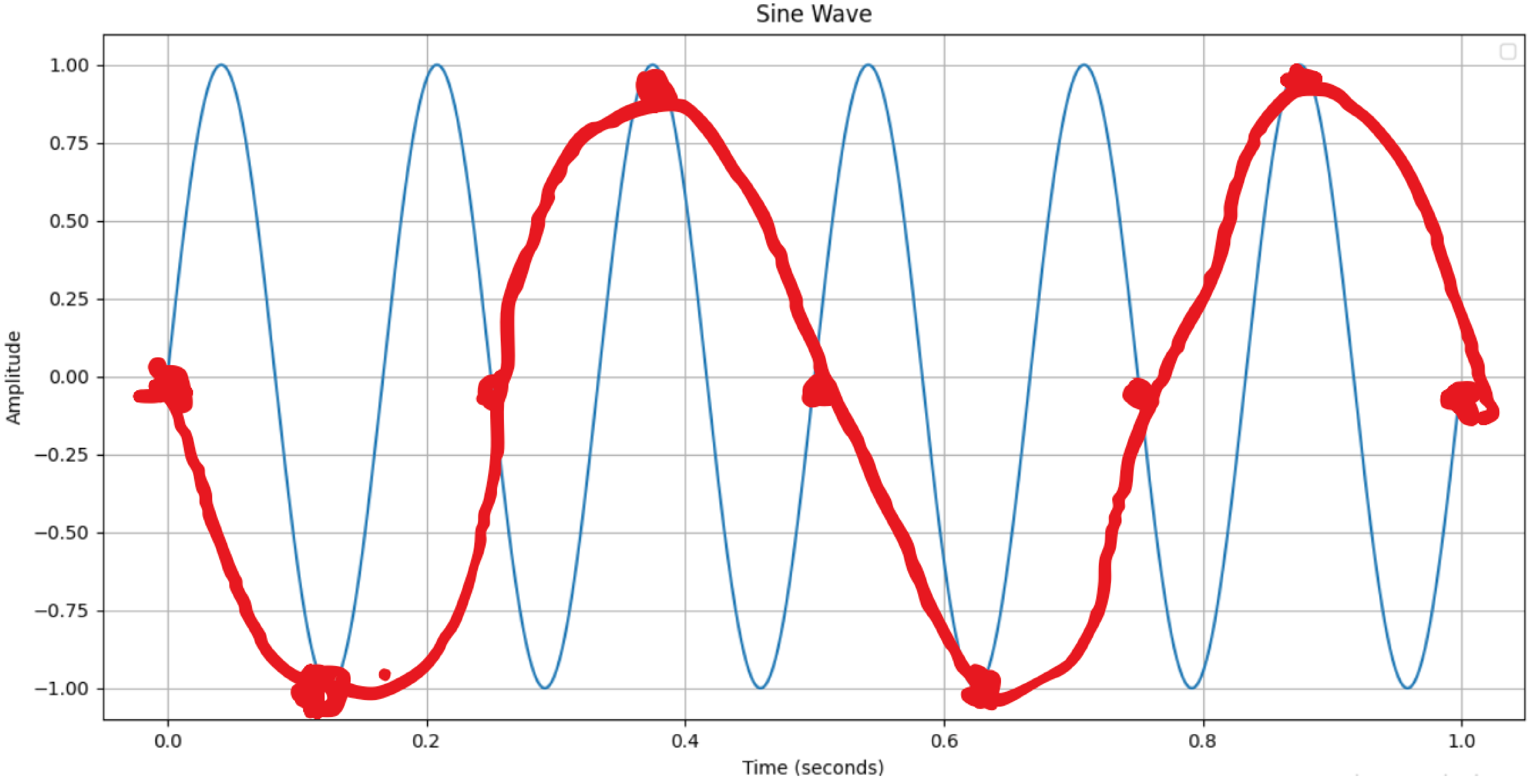
Bit Rate: Number of bits per second

File Type: Once recorded, these 1's and 0's are stored as a file called raw data.

CAN EVERY SINE WAVE BE RECONSTRUCTED FORM ITS SAMPLES?

No if we take less samples during sampling, aliasing effect

f=6hz
samples = 9
(less than 2*6)



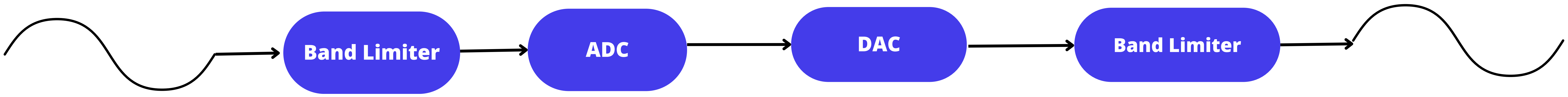
Nyquist Sampling Theorem

To avoid aliasing or to perfectly reconstruct from its samples, we need to sample the waveform over twice as fast as it's highest frequency component.

$$f_s > 2f_{\max}$$

For example, the standard audio CD has a sampling rate of 44.1 kHz, which is more than twice the typical human hearing range (20 Hz to 20,000 Hz). This ensures that audio signals can be accurately represented and reconstructed.

The Nyquist theorem is particularly applicable to bandlimited signals only.



Need for Audio Codec

- As mentioned the digital audio data or raw data usually takes a lot of disk space which makes it in-efficient to use it across different applications or transfer the data.

- Audio codecs play a crucial role in efficiently representing, compressing, and decompressing audio signals while maintaining acceptable quality and widely used in Streaming services and Multi-media applications

- An audio codec is software that compresses and decompresses audio data, it consumes only 10% space of its original size. If you have a 1GB audio file, Audio codecs compress, and it becomes a 10 MB file. more importantly, it doesn't affect the audio quality of the file though the size has reduced.

BLOCK DIAGRAM

TYPES OF AUDIO CODEC

major types of Audio codecs.

- Uncompressed Audio codec
- Compressed Audio codec

Uncompressed Audio formats

- WAV (Waveform Audio File)
- PCM (Pulse Code Modulation)

Compressed audio codecs categories

- Lossless compression
- Lossy compression

Lossless Audio Formats

- FLAC (Free Lossless Audio Codec)
- ALAC (Apple’s Lossless Audio Codec)

Lossy Audio Formats

- MP3 , AAC, Ogg Vorbis

TYPES OF AUDIO CODEC

Audio Format	File Extensions	Compression Ratio (%)	Decoding Speed (samples/mcs)	Encoding Speed (samples/mcs)	Software
Lossless:					
WAVE	.wav	100.00%	∞	∞	-
FLAC	.flac	61.05%	66.72	22.86	libFLAC
MP4 ALAC	.mp4, .m4a	63.18%	14.26	-	libALAC
WavPack	.wv	60.83%	23.21	-	libwavpack
APE	.ape	59.61%	9.6	-	libMAC
OptimFROG	.ofr	57.86%	5.62	3.77	ofr
Lossy:					
OGG Vorbis	.ogg	N/A	26.97	4.95	libvorbis
OGG Opus	.opus	N/A	17.55	4.55	libopus
MPEG-1 Layer-3	.mp3	N/A	48.89	4.88	libmp3lame, libmpg123
MP4 AAC LC	.mp4, .m4a	N/A	17.41	3.73	libfdk-aac
Musepack	.mpc	N/A	46.14	-	libmpc

FFMPEG

- **FFmpeg is the leading multimedia framework, able to decode, encode, transcode, mux, demux, stream, filter and play**

libavutil:
- Core multimedia utility library.
- Includes functions for programming, math, and more.

libavcodec:
- Audio/video codec library.
- Contains encoders and decoders.

libavformat:
- Multimedia container format library.
- Includes demuxers and muxers.

libavdevice:
- Input/output device library.
- Supports various multimedia frameworks.

libavfilter:
- Media filter library.
- Contains filters for multimedia processing.

libswscale:
- Image scaling and color space conversion.
- Highly optimized operations

libswresample:
- Audio resampling and conversion.
- Highly optimized operations.

FFMPEG

conatiner - mpeg4

video stream - h264

audio stream - mp3

inputfile.mp4
(all the digital video or audio files are encoded by default)

Basic Structures Needed

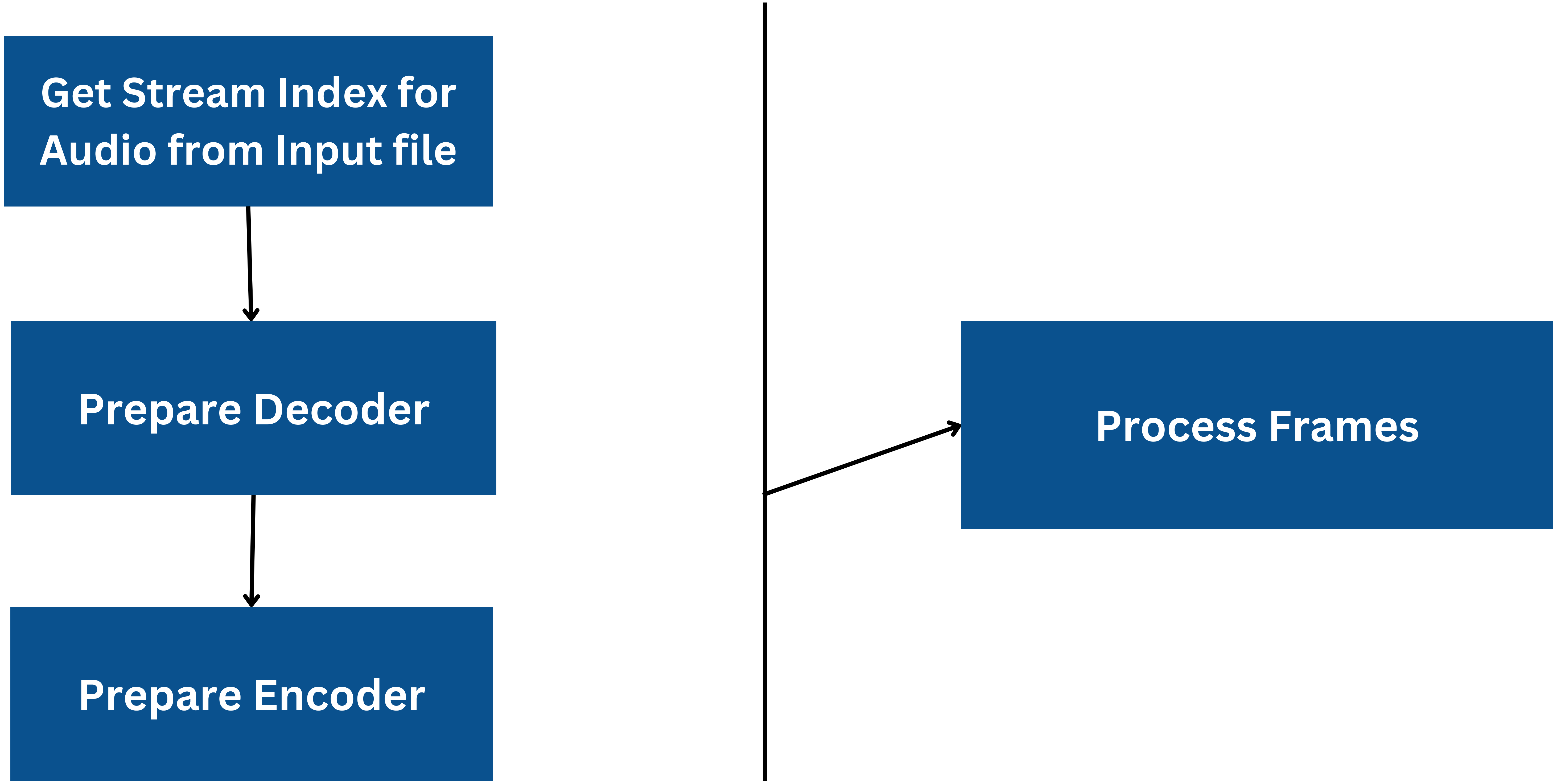
- AVFormatContext
- AVCodecContext
- AVCodec
- AVStream
- AVPacket
- AVFrame

raw data

Decode the audio Stream → **Encde the raw data or frame with different audio codec(aac)** → **write to file or send to server over network**

We need to populate the structures in such a way we can do decoding , encoding accurately. To achieve we need to follow a sequence adn then we can start processign the samples.

FFMPEG SEQUENCE



FFMPEG SEQUENCE

Get Stream Index for Audio from Input file

```
AVFormatContext* input_format_context = avformat_alloc_context();
```

```
int avformat_open_input(AVFormatContext **ps, const char *url,
                        const AVInputFormat *fmt, AVDictionary **options);
```

```
int avformat_find_stream_info(AVFormatContext *ic, AVDictionary **options);
```

```
or (int i = 0; i < input_format_context->nb_streams; i++)
   if (codec_type == AVMEDIA_TYPE_AUDIO) {
       audioStreamIndex = i;
   }
```

FFMPEG SEQUENCE

Prepare Decoder

```
AVCodec *avcodec_find_decoder(enum AVCodecID id);

AVCodecContext *avcodec_alloc_context3(const AVCodec  *codec);
```

```
int avcodec_parameters_to_context(AVCodecContext *codec, const AVCodecParameters *par);
```

```
int avcodec_open2(AVCodecContext *avctx, const AVCodec *codec, AVDictionary **options);
```

FFMPEG SEQUENCE

Prepare Encoder

```
int avformat_alloc_output_context2(AVFormatContext **ctx, const AVOutputFormat *oformat,
                                   const char *format_name, const char *filename);
```

```
AVStream *avformat_new_stream(AVFormatContext *s, const struct AVCodec *c);

AVCodec *avcodec_find_encoder(enum AVCodecID id);

AVCodecContext *avcodec_alloc_context3(const AVCodec *codec);

//Set the basic encoder parameters. (not API)
```

```
int avcodec_open2(AVCodecContext *avctx, const AVCodec *codec, AVDictionary **options);
```

```
int avio_open(AVIOContext **s, const char *url, int flags);
```

FFMPEG SEQUENCE

Process Frames

```
//Return the next frame of a stream
while (av_read_frame(input_format_context, input_packet) >= 0) {
  int response = 0;
  while (response >= 0) {
    //Return decoded output data from a decoder
    response = avcodec_receive_frame(input_codec_context_audio, iframe);
    //Supply a raw audio frame to the encoder
    response = avcodec_send_frame(output_codec_context_audio, iframe);
    while (response >= 0) {
      //Read encoded data from the encoder.
      response = avcodec_receive_packet(output_codec_context_audio, output_packet);
      //Convert valid timing fields (timestamps / durations) in a packet
      response = av_interleaved_write_frame(output_format_context, output_packet);
    }
  }
}
```

SERIES ON AUDIO

- Audio Raw Capture (Port Audio + PCM data)
- PCM Data to Encoded data using OpencoreAMR
- Audio Raw Capture + Encode (OpenAMR)
- Audio Raw Capture + Encode (AAC)
- Mix PCM's
- Filestream / upstream from the file

THANK YOU

Presented by Phani