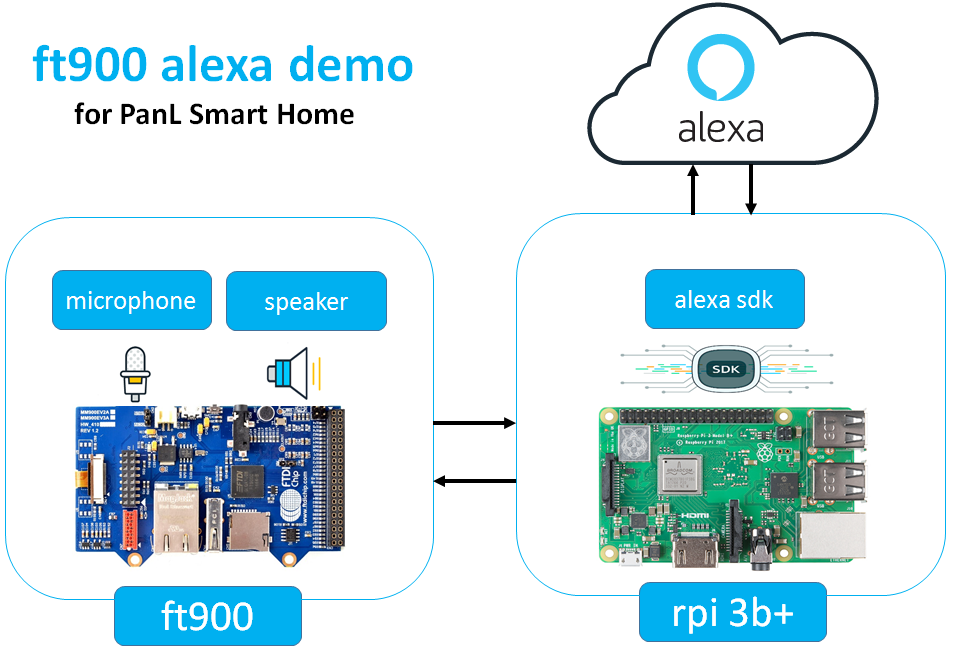
**FT900 Alexa Demo**

**FT900 Alexa Demo**

This PoC application demonstrates using FT900 microcontroller as an Amazon Echo Dot device, where users can issue voice commands to Alexa and hear Alexa’s voice responses generated by Alexa’s complex speech recognition and natural language processing in the cloud. To make this possible, FT900 (Alexa client) communicates with a Raspberry PI 3B+ (Alexa gateway), which relays voice requests and voice responses to and from the Alexa cloud. The RPI runs a customized version of Amazon’s official open-source Alexa Voice Service (AVS) SDK, written in C++.

****

**Github Code:** [**https://github.com/richmondu/FT900/tree/master/Alexa/Amazon%20Alexa%20Client**](https://github.com/richmondu/FT900/tree/master/Alexa/Amazon%20Alexa%20Client)

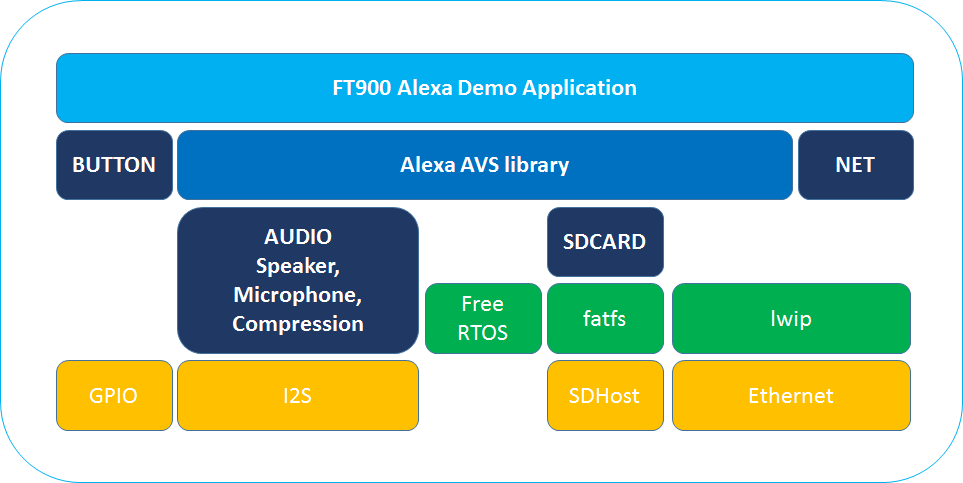
**PanL Smart Home with built-in Alexa**

This demo is targeted for FTDI/Bridgetek’s smart home devices, PanL Hub and PanL Display. Having Alexa built-in to PanL products allow customers to talk directly to Alexa via PanL without needing to buy Amazon Echo devices. PanL Hub, which runs on RPI, will act as the Alexa gateway while the PanL Display, which runs on FT900 microcontroller, will act as the Alexa client. Customers will be able to use both PanL Hub and PanL Display to issue voice commands and operations to Alexa.

****

**FT900 Alexa Client – Block Diagram**

Below is a block diagram of the Alexa Demo on the FT900 side.

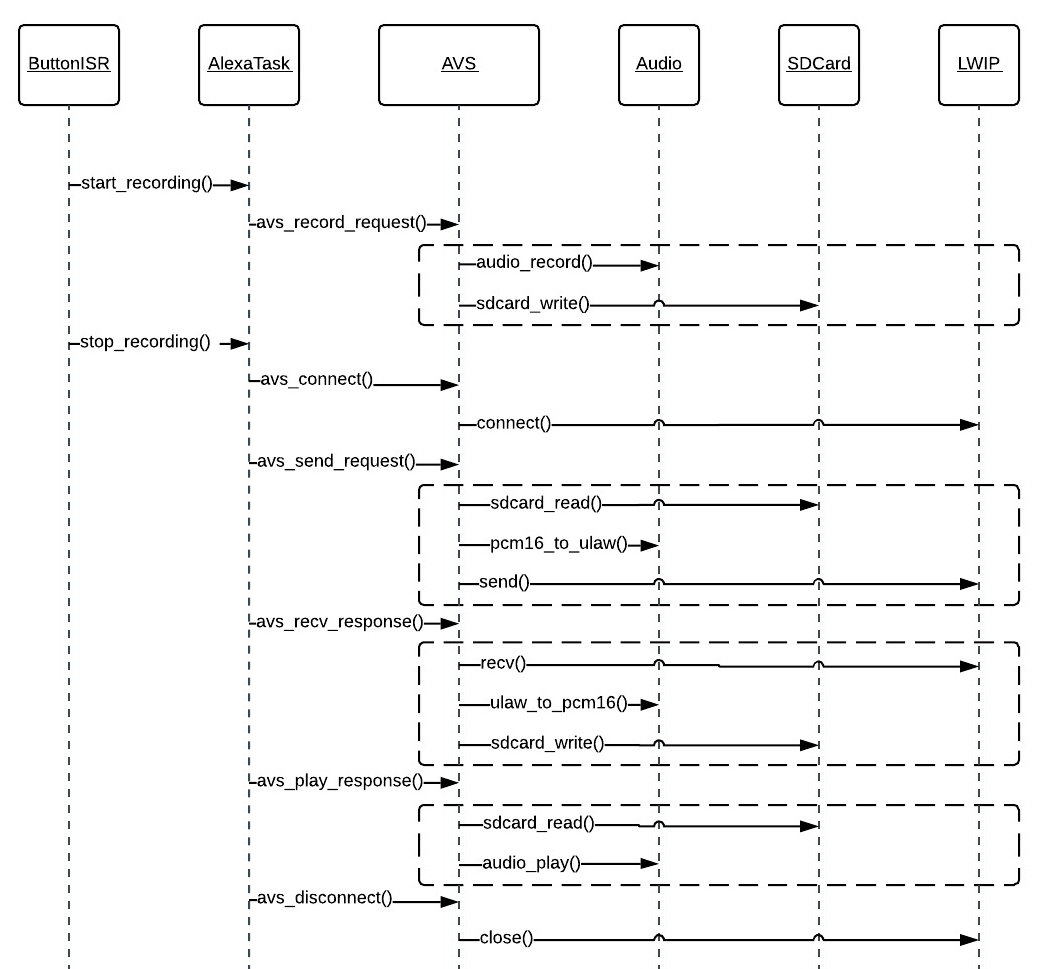


The main component of the Alexa Demo on the FT900 side is the **Alexa AVS library**. The library abstracts the audio, the SD card and the network communication by providing easy to use API interfaces. The Alexa AVS library contains the following functions:

|  |  |
| --- | --- |
| **Functions** | **Descriptions** |
| avs\_connect | Establishes connection to the RPI Alexa Gateway using configurations in avs\_config.h configuration file. |
| avs\_disconnect | Closes connection with RPI Alexa Gateway. |
| avs\_record\_request | Records voice request from microphone and saves it to SD card given the provided filename.  **Audio recorded**: 16-bit PCM, 16KHZ, stereo (2-channels)  **Audio saved**: 16-bit PCM, 16KHZ, mono (1-channel) |
| avs\_send\_request | Sends the voice request to the RPI Alexa Gateway provided the filename of the voice recording in the SD card.  **Audio read**: 16-bit PCM, 16KHZ, mono (1-channel)  **Audio sent**: 8-bit u-law, 16KHZ, mono (1-channel) |
| avs\_recv\_response | Receives the voice response from the RPI Alexa Gateway and saves to SD card given the provided filename.  **Audio received**: 8-bit u-law, 16KHZ, mono (1-channel)  **Audio saved**: 16-bit PCM, 16KHZ, mono (1-channel) |
| avs\_play\_response | Plays the voice response from the SD card to the speaker.  **Audio read**: 16-bit PCM, 16KHZ, mono (1-channel)  **Audio played**: 16-bit PCM, 16KHZ, stereo (2-channels) |

**FT900 Alexa Client – Sequence Diagram**

Below is a sequence diagram of the Alexa Demo on the FT900 side.

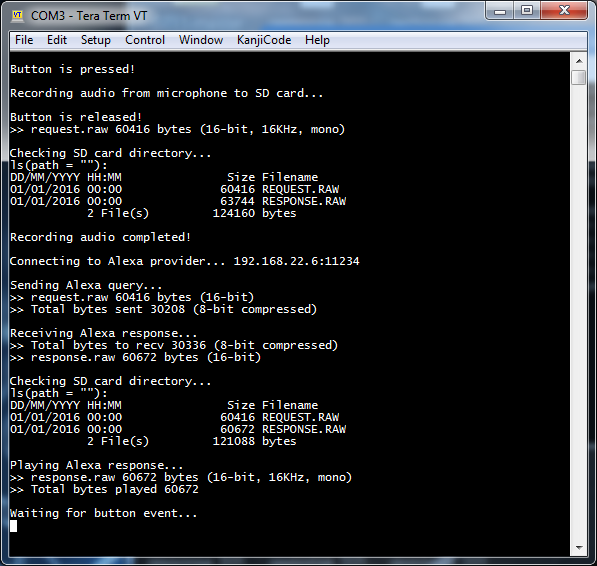


Notes:

1. **Wakeword detection** is not supported. To trigger voice recording, user has to press down a button and release it to stop recording, similar to remote control for Amazon Fire Stick TV.
2. Data sent is compressed from **16-bit to 8-bit** using **u-law G711 compression algorithm**. Data received is expanded from 8-bit to 16-bit using u-law G711 expanding algorithm. Compressing the data before transmission reduces the data bandwidth by half.
3. Alexa cloud expects **16-bit, 16KHZ, mono (1-channel)** **PCM** voice recording. For response, Alexa cloud sends out an **MP3** data stream. RPI converts the MP3 data stream to **PCM** before compressing to 8-bit and sending to FT900.
4. Converting stereo to mono data stream is a matter of removing alternating 16-bit WORD.

**FT900 Alexa Client – Sample Execution Log**

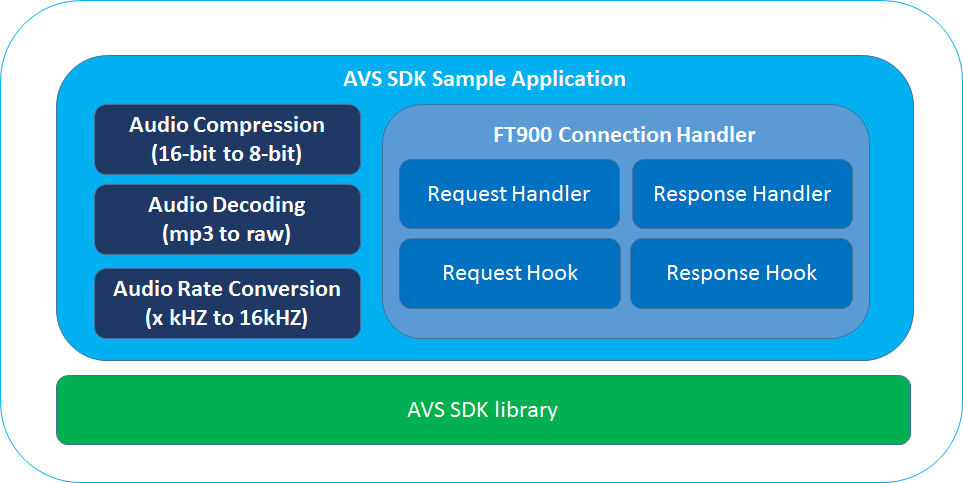
Below is a sample execution log of the Alexa Demo on the FT900 side.



**RPI Alexa Gateway – Block Diagram**

Amazon provides an official **Alexa Voice Service (AVS) SDK**, (written in C++) <https://github.com/alexa/avs-device-sdk> (Note that the Java, Python and NodeJS SDKs are already obsoleted and are no longer publicly available.) The version I am using is **AVS SDK is v1.11.0**, (**12-19-2018)**. Instructions to install on the AVS SDK on RPI can also be found on the github link above.

Below is a block diagram of the Alexa Demo on the RPI-side.

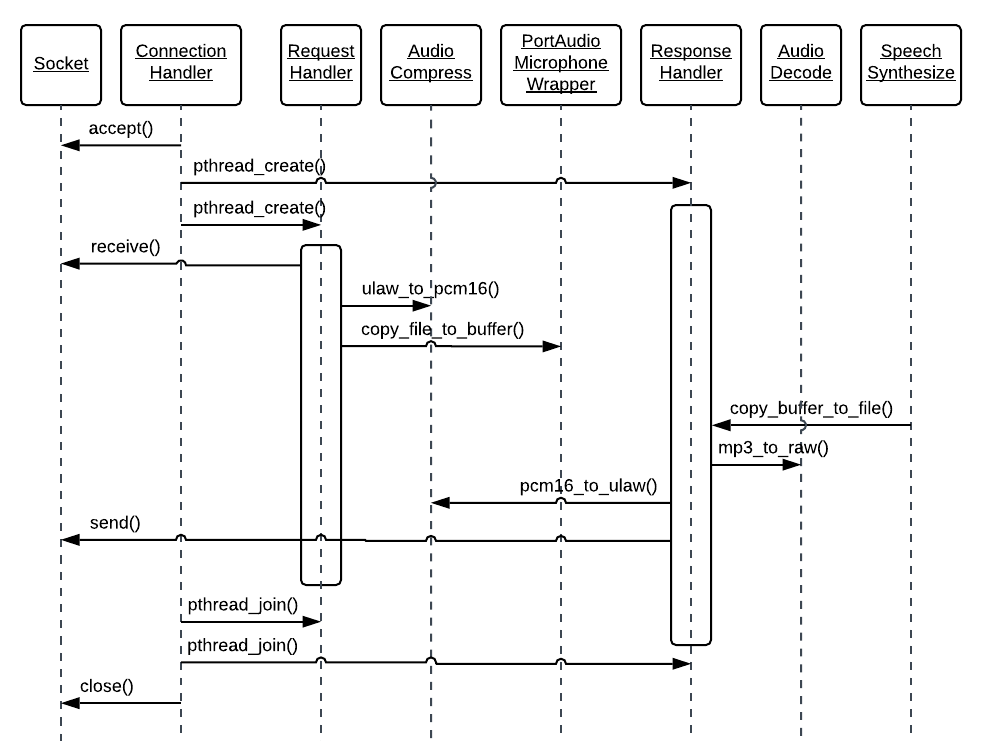


Below are the Alexa Demo modules that were implemented and modified on the RPI-side.

|  |  |  |
| --- | --- | --- |
| **Files Implemented** | | **Descriptions** |
| FT900ConnectionHandler.cpp | | Handles connection with FT900 and creating threads for both FT900RequestHandler and FT900ResponseHandler. |
| FT900RequestHandler.cpp | | Handles processing of Alexa request from FT900. |
| FT900RequestHook.cpp | | Handles hook of request to PortAudioMicrophoneWrapper. |
| FT900ResponseHandler.cpp | | Handles processing of Alexa response to FT900. |
| FT900ResponseHook.cpp | | Handles hook of response from Speech Synthesizer of library. |
| FT900AudioCompression.cpp | | Handles ulaw 8-bit compression and expansion. |
| FT900AudioCompressionHelper.cpp | | Helper function for FT900AudioCompression.cpp. |
| FT900AudioDecoding.cpp | | Handles decoding of mp3 response to raw pcm. |
| FT900AudioRateConversion.cpp | | Handles conversion from x KHZ to 16 KHZ. |
| **Files Modified** | **Descriptions** | |
| SampleApplication.cpp | Initialize FT900ConnectionHandler in separate thread. | |
| PortAudioMicrophoneWrapper.cpp | Feed in request data to microphone datastream. | |
| SpeechSynthesizer.cpp | Call the callback function when response is received. | |
| UserInputManager.cpp | For testing of pre-recorded audio requests. | |
| MediaPlayer.cpp | Fix audiosink issue. | |
| DefaultClient.h | Provide access to Speech Synthesizer handle. | |

**RPI Alexa Gateway – Sequence Diagram**

Below is a sequence diagram of the Alexa Demo on the RPI-side.



Notes:

1. The primary modifications for the AVS SDK application are contained in the following classes:

**PortAudioMicrophoneWrapper**: PortAudioCallback() contains the data stream for Alexa request

**SpeechSynthesizer**: startPlaying() contains the data stream for Alexa response

1. Audio properties for Alexa Demo on the RPI-side:

**Audio received** (from FT900): 8-bit u-law, 16KHZ, mono (1-channel)

**Audio sent** (to Alexa cloud): 16-bit PCM, 16KHZ, mono (1-channel)

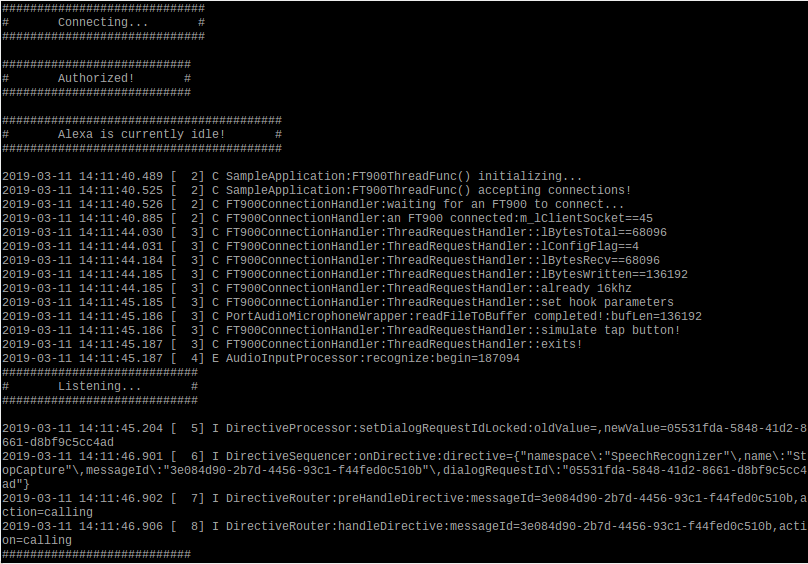
**Audio received** (from Alexa cloud): MP3

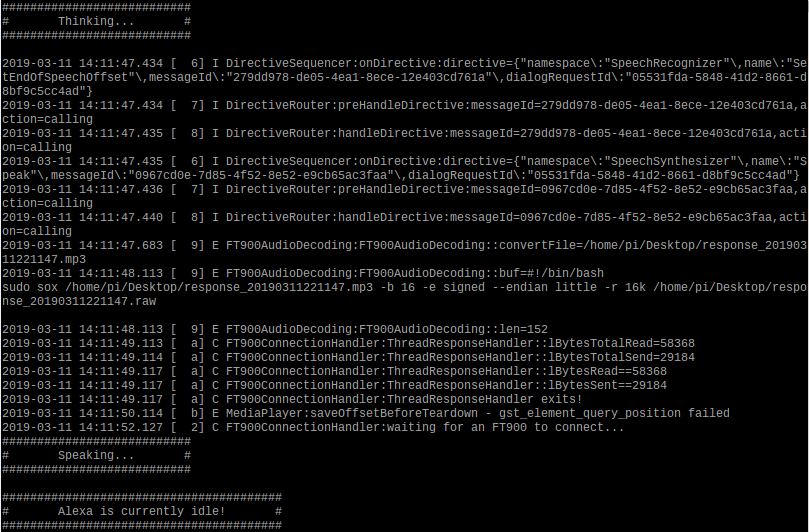
**Audio sent** (to FT900): 8-bit u-law, 16KHZ, mono (1-channel)

1. **SOX** utility is used to decode MP3 data stream to raw PCM16 data stream.
2. **G711 u-law** lossless companding (compression/expanding) algorithm is used to convert data stream from 16-bit to 8-bit and vice versa, without affecting audio quality. Compressing the data before transmission reduces the data bandwidth usage by half.

**RPI Alexa Gateway – Sample Execution Log**

Below is a sample execution log of the Alexa Demo on the RPI-side.





**Action Items**

Below are the action items for the Alexa Demo.

1. Support for **wake-word detection in FT900.** (Currently, user has to press down a button to start voice recording.)
2. **Use 8KHz instead of 16KHz** to reduce audio file size by half. (However, note that RPI will have to convert the 8KHz to 16KHz as Alexa cloud requires 16KHz).
3. **Performance/speed optimization**
4. Support for **alarms or notification-based messages**
   1. Currently, only responses triggered by requests are supported.
5. Support for **very long Alexa responses**.
   1. Need to test requests that have very long responses.
6. Support for **queuing Alexa requests from multiple FT900 clients**.
   1. Multiple FT900 can simultaneously send requests to RPI. RPI should queue the requests and only issue a request when a response for previous request is processed.
7. **RPI should not play response on its speaker** when the request is from FT900.
8. **Audio decoding implementation** currently uses bash scripts using **SOX** utility.
   1. Should be replaced with C/C++ code
9. **Upgrade to latest AVS SDK version.**
   1. Currently using **AVS SDK 1.11.0**, (12-19-2018). As of today, the latest version is **AVS SDK 1.12.0** (02-28-2019).

**References**

Below are links used for this FT900 Alexa Demo:

1. Alexa Voice Service (AVS)

<https://developer.amazon.com/alexa-voice-service>

1. Getting Started with the AVS Device SDK <https://www.youtube.com/watch?v=F5DixCPJYo8&feature=youtu.be>
2. Amazon’s Official AVS Device SDK (written in C++)

<https://github.com/alexa/avs-device-sdk>

1. AVS SDK Installation Guide on Raspberry PI (RPI)

<https://github.com/alexa/avs-device-sdk/wiki/Raspberry-Pi-Quick-Start-Guide-with-Script>

1. SOX Sound Exchange utility (used for MP3 audio decoding)

<http://sox.sourceforge.net/>

1. G711 Audio Companding algorithms (used for u-law compression/expanding)

<https://en.wikipedia.org/wiki/G.711>

The modified AVS SDK is located at

<https://github.com/richmondu/FT900/tree/master/Alexa/Amazon%20Alexa%20Gateway/avs_sdk>