VOICE RECOGNITION WITH ZEDBOARD

MOD002726

CONTENTS

1. INTRODUCTION

- ***Deliverable***: Demo of a household controlled by voice using the microphone-line of the Zedboard.

- A client of Alexa Voice Service (AVS) is available for the Raspberry Pi.
  - ***AlexaPi*** [2] may run on the Zedboard as well (Debian)
  - However, it simply sends and receives voice from the cloud.
  - Voice recognition is used only to trigger the interaction (keyword "alexa")

2. ALTERNATIVES

1. Cloud Computing: Voice recognition to trigger AlexaPi (desktop mode) and an Amazon ***skill*** that connects back to a local script (i.e, relay controller).

2. Local Processing: Alternative way to trigger AlexaPi (platform dependant) and "hack" the voice recognition to create a separated ***listener***.

Option 2: The listener application will be based on the module used in AlexaPi for voice recognition: CMUSphinx [3]

3. CMUSPHINX

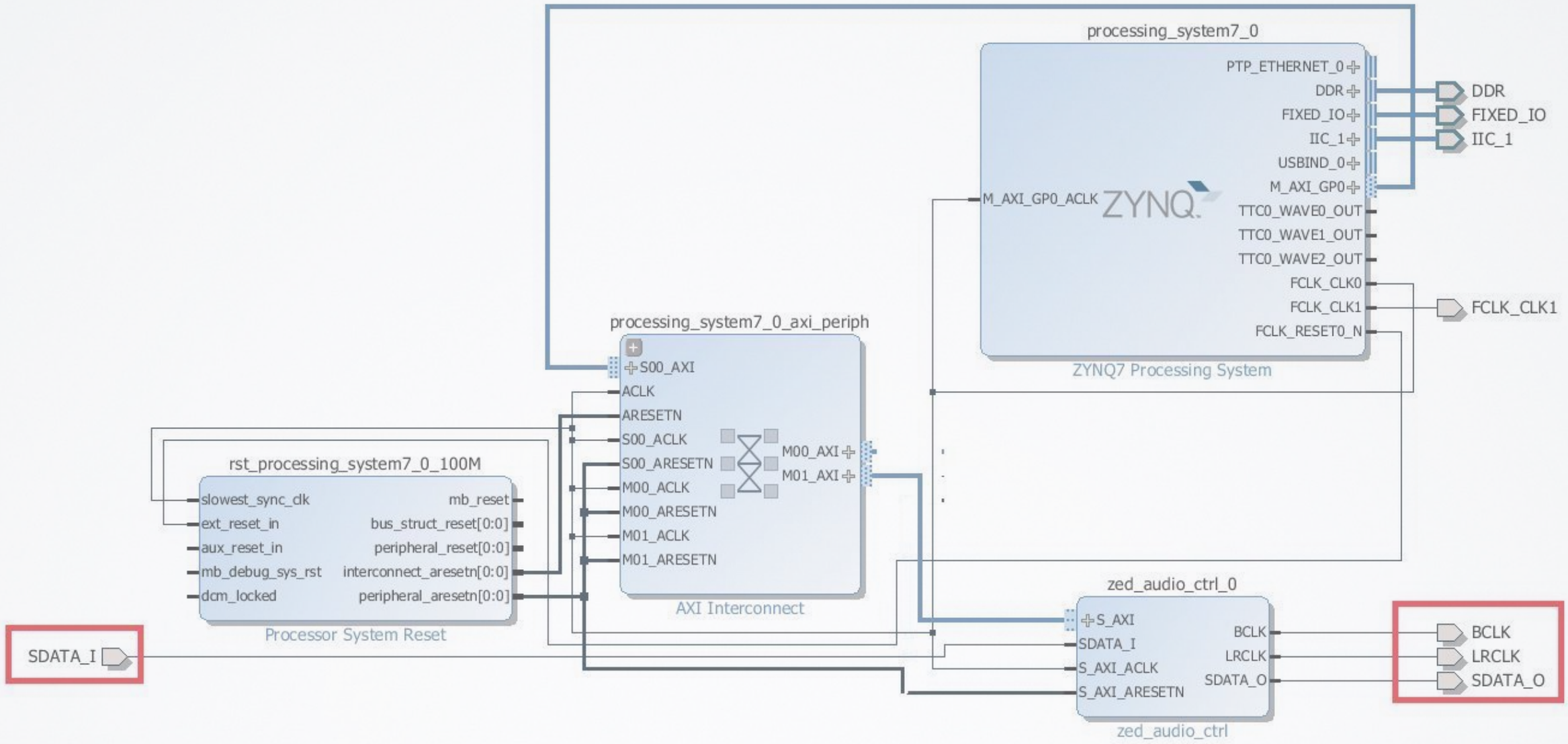
- Open-source ***Speech To Text*** library (STT) from Carnegie Mellon University (CMU) with an optimised version for ARM processors (PocketSphinx).

- How to use it in Zynq devices? (see chapter 5 - PS/PL tools):

  - *like a **bare-metal** application (Vivado / SDK)*
    An IP core implements a I2S link for sending samples to the audio codec. The samples are transferred between the PS part and the audio codec via the AXI bus. Then a SDK project should be created to process the samples by the processor [1].

  - *like a process inside of **Xillinux®***
    All the logic to interface the audio codec is integrated in a single IP core [5].
    The OS allows multiple applications at the same time, video interface, memory management, …

VIVADO/SDK



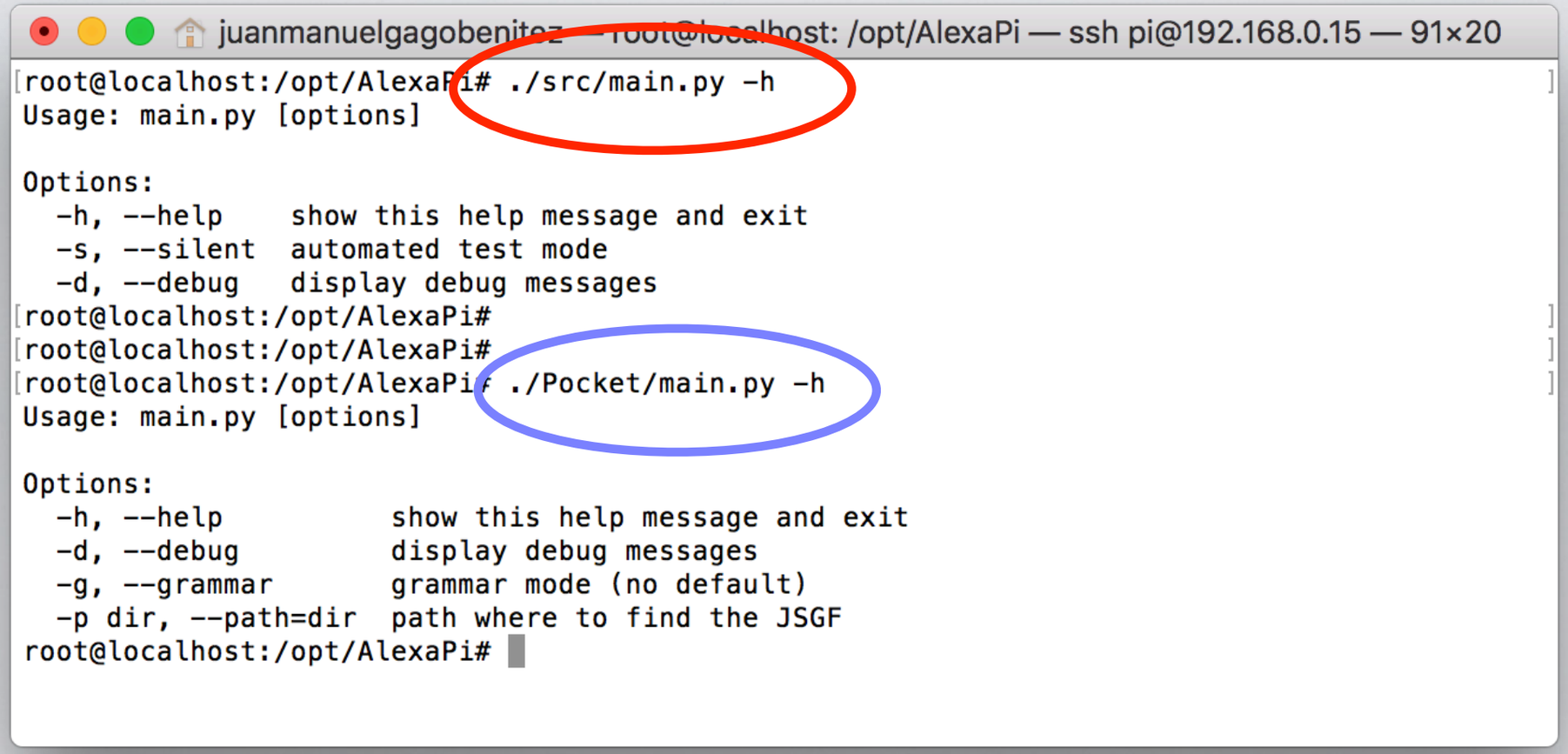
Source [1]

XILLINUX®

- The Xillinux® bundle includes the Xillybus® core (PL part) and the Linux image with the drivers (PS part).
    - It is based on Ubuntu 16.04 so it's open-source and Python is already installed.
    - Supports sound by interfacing dedicated Xillybus® streams with the ***Pulseaudio*** framework [4]
- Btw, AlexaPi is written in Python and uses the Pulseaudio framework [2]
- This option doesn't require to have much experience with Vivado

CONTENTS

4. APPLICATION



Two Python scripts: **AlexaPi** (modifed) and **Pocket** (new design)

Modes: secret(trigger), grammar (turn on the lamp) or keyword (alexa)

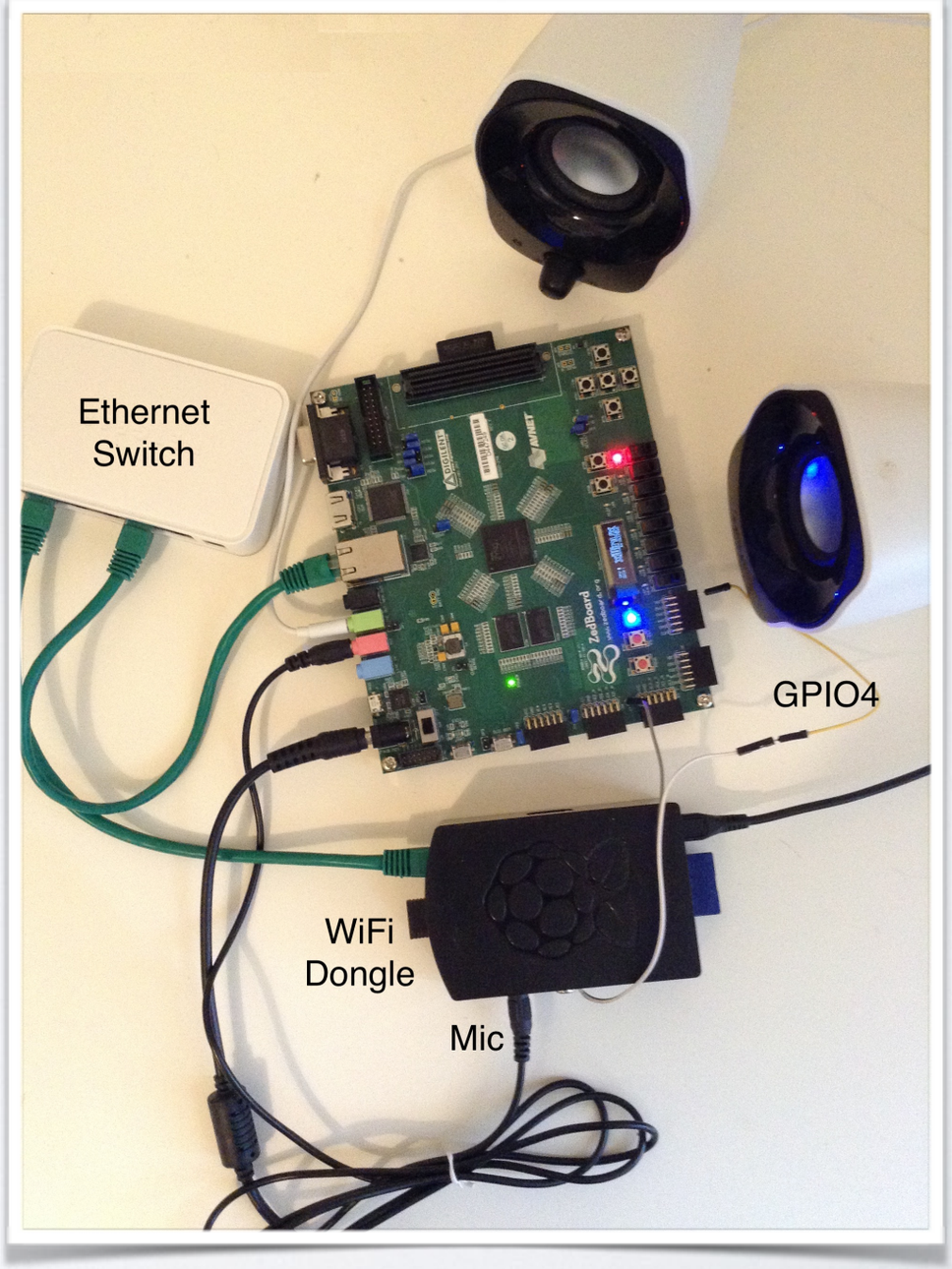
AUTOMATED TEST

Raspberry Pi configuration:

1. Keyword detection: audio generation with TTS module (***Text To Speech***)

2. Grammar recognition: monitoring led status (GPIO4 input)

3. AlexaPi on Zedboard: button simulation (GPIO4 output)



TEST RESULTS

	Test Case	Input	Expected	Output	Result
1	**Pocket** - keyword detection	10 random commands starting with “alexa”	10 detections	10 detections	Pass
2	**Pocket** - grammar recognition	20 turn on and off commands	20 led switching	20 led switching	Pass
3	**AlexaPi** - autotest mode	10 random commands starting with “alexa”	10 answers from AVS	8 answers, 2 unknown	Fail

5. CONCLUSIONS

- AlexaPi customised to use the leds and push-buttons on the Zedboard [2]
- Methodology with ope-source components:

PS	PocketSphinx [3] and Pulseaudio [4]
PL	Xillybus® integrated in the Xillinux® bundle [5]

5. CONCLUSIONS

- A listener application has been developed with PocketSphinx [3]
- Grammar mode to control a GPIO in the PL part: led or a relay to switch on a household (see deliverable)
- Voice recognition using cloud computing vs local processing

Cloud	Two failures in AlexaPi
Local	No failures with Pocket

6. REFERENCES

[1] Crocket, Louise et al. 2014. "The Zynq Book".

[2] AlexaPi, 2017. "Alexa client for all your devices". Available at https://github.com/alexapi/AlexaPi

[3] CMU, 2017. "Open source speech recognition toolkit". Available at https://cmusphinx.github.io

[4] Pulseaudio, 2017. "Welcome to PulseAudio!". Available at https://www.freedesktop.org/

[5] Xillybus, 2018. "Getting started with Xillinux for Zynq-7000"

QUESTIONS