PLASMA: Pushing the Limits of Audio Spatialization with eMerging Architectures

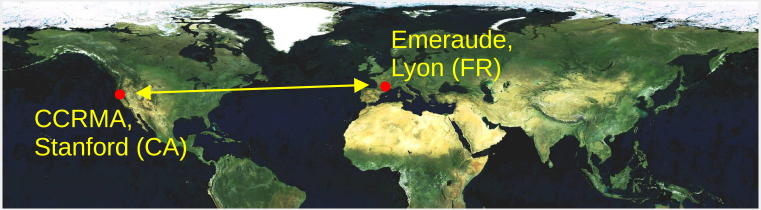
Associate Project-Team Between CCRMA and Emeraude (INRIA/INSA, France)

Romain Michon, Chris Chafe, Pierre Lecomte,
Fernando Lezcano-Lopez, Mike Mulshine, Tanguy Risset,
Dirk Roosenburg, Maxime Popoff, and Julius Smith

CCRMA Open House
October 21st, 2022

https://ccrma.stanford.edu/~rmichon/talks/ccrma-oh-22.pdf

PLASMA: a Collaboration Between CCRMA and INRIA/INSA



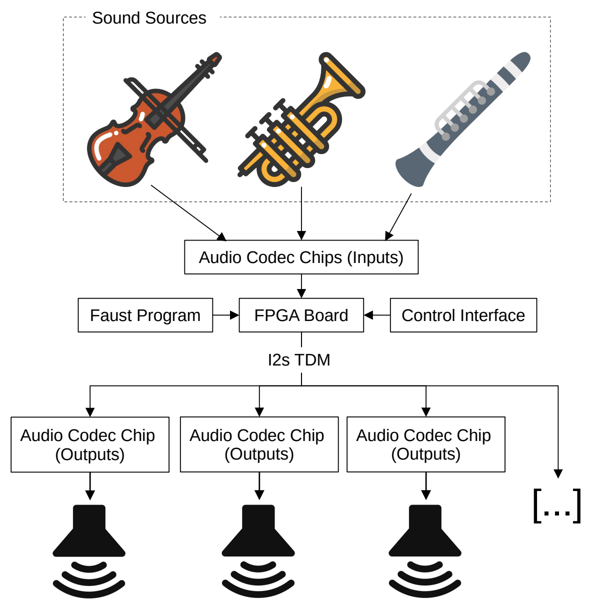
- INRIA is the French National Institute for Research in Digital Science and Technology. INSA is an engineering school.
- The Emeraude Team at INRIA/INSA is based in Lyon (France). It focuses on embedded audio systems and their programming.
- PLASMA is co-funded by INRIA/INSA (through its international associate team program), CCRMA, and the France-Stanford Center.
- PLASMA aims at Pushing the Limits of Audio Spatialization with eMerging Computer Architectures.
- In 2022, 4 trips between France and the US were funded. In 2023, Maxime Popoff and Romain Michon will visit CCRMA for one quarter and CCRMAlites will likely visit INRIA/INSA in Lyon.

Two Approaches, One Goal: Exploring the Use of Emerging Computer Architecture In the Context of Spatial Audio

- Traditional computer architectures used for real-time audio processing (i.e., PC + audio interface) are limited in terms of throughput and computational power: only that many audio channels through a USB cable and only that much CPU power.
- PLASMA aims at exploring the use of emerging embedded architectures to push the limits of spatial audio.
- For that, two approaches are being considered:
  - FPGA-Based **Centralized Approach**: Processing happens on an FPGA; A large number of audio codec chips are connected to the FPGA using low-level interfacing.
  - Microcontroller Network-Based **Distributed Approach**: Processing is distributed between small inexpensive microcontrollers; Audio streams are sent to each of them using TCP/UDP; Individual micontrollers are controlled using OSC.
- Open-source first!

Centralized FPGA-Based Approach

- A large number of audio codec chips are connected to the FPGA using i2s Time Division Multiplexing (TDM).
- The FPGA is fully programmable in Faust (i.e., Wave Field Synthesis, Ambisonics, etc.).
- Sound sources are provided to the system as analog audio inputs.
- The system is controlled using a laptop connected to the FPGA, OSC, etc.



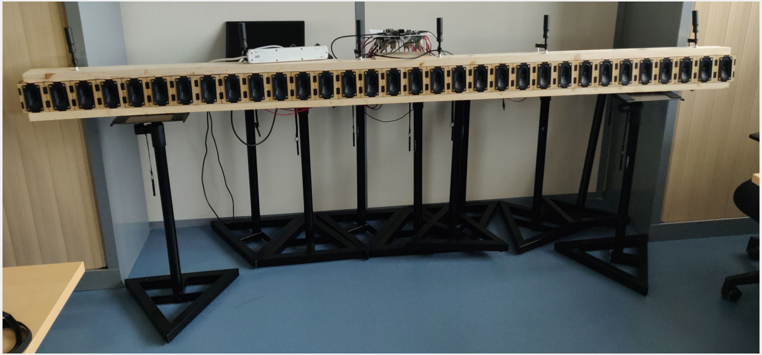
What's an FPGA?

- Field-Programmable Gate Array.
- Integrated circuit designed to be configured "on the field" using a Hardware Description Language (HDL).
- FPGAs contain an array of programmable logic blocks, and a hierarchy of reconfigurable interconnects allowing blocks to be wired together.
- FPGA performances are limited by: (i) the amount of resources available on the chip, (ii) the maximum clock at which it can be ran.
- FPGAs provide a high level of parallelization.
- The two main manufacturers of FPGAs are Xilinx/AMD and Altera/Intel.

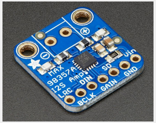
Why are FPGAs Good for Spatial Audio?

- Spatial audio algorithms (i.e., WFS, ambisonics, etc.) can usually be parallelized a lot.
- Some audio codec chips can be multiplexed using TDM through a very low number of GPIOs, i.e., 2 + 1 GPIOs for 16 audio channels on the most powerful codecs such as Analog Devices' ADAU 1787 as long as very fast clocks can be produced.
- FPGAs have lots of GPIOs (about 30 on basic FPGAs, more than 70 on high-end FPGAs), can produce very fast clocks, and can compute large numbers of digital audio streams in parallel.
- Hence, in theory, hundreds of audio channels can be processed in parallel (i.e., theoretically more than 1000x1000 on a Xilinx Ultrascale FPGA).
- Initial experiments show that the main bottleneck is the potential number of memory accesses done by the system in DDR rather than actual computational power. Hence, algorithm with small memory footprints can run without a problem.

Low-Cost WFS System Prototype

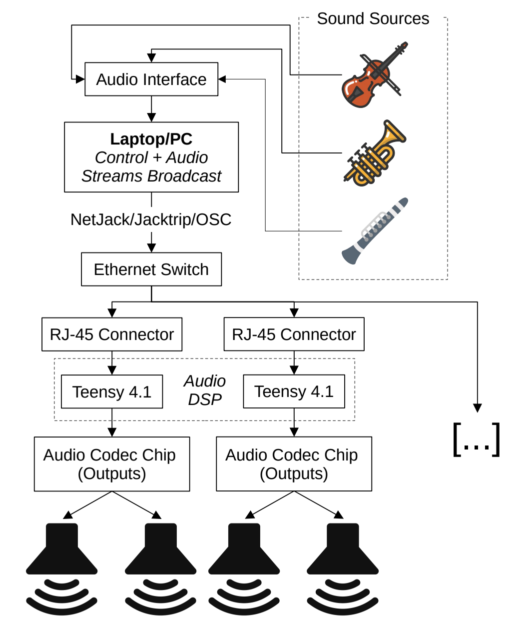


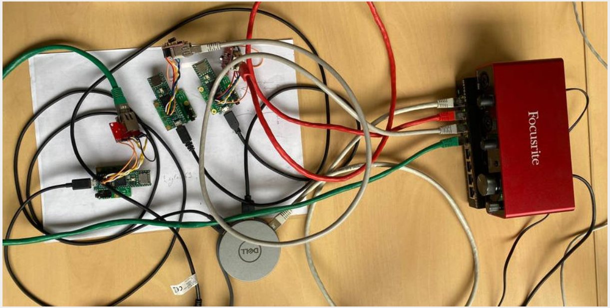
Prototype of an FPGA-based WFS system (programmed in Faust) using cheap ($6, well $3 a year ago lol) Adafruit i2s amplifiers (MAX98357A):



Distributed Network-Based Approach

- Audio processing is distributed between a series of Teensy 4.1 microcontrollers equipped with a stereo audio codec chip (each Teensy processes 2 audio channels).
- Sound sources are broadcast to all microcontrollers using NetJack or JackTrip.
- Audio DSP takes place on the microcontrollers themselves.
- Microntrollers are controlled using OSC (Open Sound Control).
- This type of approach can be infinitely scaled up.





- We currently have a grad-level intern working full time on this.
- JackTrip has been ported to the Teensy.
- We're in the process of comparing the performances of Jacktrip with that of NetJack in this context.
- A prototype is currently being assembled.
- Programming the Teensys and distributing audio DSP should be very straighforard thanks to Faust.

Thanks! :)

CCRMA Colloquium on Wednesday (Oct. 26th) on “High-Level Programming of FPGAs for Audio Real-Time Signal Processing Applications”

DSP seminar on Friday (Oct. 28th) on “Compiling Audio DSP for FPGAs Using the Faust Programming Language and High Level Synthesis”

Send your questions to: `romain.michon@inria.fr`

More @: `https://team.inria.fr/emeraude/plasma/`

Slides @: `https://ccrma.stanford.edu/~rmichon/talks/ccrma-oh-22.pdf`