3D Photonic integration platform based on multilayer PolyBoard and TriPleX technology for optical switching and remote sensing and ranging applications

Grant Agreement no. 780502
Research and Innovation Action (RIA)
Information & Communication Technologies (ICT)

Deliverable D6.3

Development of SDO agent and evaluation of 4×4 optical switch (Module-1) using bench-top and system tests

Lead beneficiary for the deliverable:	ICCS/NTUA
Contact Person:	Lefteris Gounaridis
Phone:	+30 210 772 2057
e-mail:	lgou@mail.ntua.gr
Date due of deliverable:	31 July 2022
Actual submission date:	15 May 2022
Authors:	L. Gounaridis, A. Raptakis, C. Kouloumentas, P. Groumas, K. Tokas, P. Bakopoulos
Participants:	ICCS/NTUA, OPTAGON, MELLANOX
Work-package:	WP6
Dissemination level:	Public
Nature:	Demonstrator
Version:	1.0
Total number of pages:	27

Executive Summary

The document presents the characterization and the evaluation results of the Module-1. Starting with a brief description of the first active 3PEAT module, the assembly and the packaging approach that have been applied is presented and lastly the testing and the characterization results in lab and quasi-real settings that have been measured in ICCS and Mellanox premises. Finally, the development of the SDO agent is described.

Keywords: photonic integration, optical routing, photonic integrated circuit, SDO, active switch

List of Acronyms

AWG	Arrayed waveguide grating
CW	Continuous Wave
EDFA	Erbium-Doped Fiber Amplifier
ER	Extinction ratio
MMI	Multi-Mode Interferometer
MZI	Mach-Zehnder interferometers
PIC	Photonic Integrated Circuit
PZT	Piezoelectric lead zirconate titanate
SDO	Software Defined Optics
TEC	Temperature Controller

Table of Contents

1. Introduction

This deliverable presents the design and the fabrication of Module-1. Then the characterization and the evaluation results of Module-1 are presented. Two types of deposition of the PZT actuators were used (the top-top and the top-bottom approach), and 2 different versions of the PolyBoard with 2 and 4 waveguide layers.

2. Design of Module-1

Novel intra-data centre connectivity optical switches have been envisaged as part of 3PEAT, exploiting the advantages of the low-loss TriPleX platform on the one hand, and the multi-layer PolyBoard platform on the other. 3PEAT created the first active optical switch (Module-1) using an array of Mach-Zehnder interferometers with PZT actuators in both arms.

Module-1 is a 4x4 switching module with 4 input and 4 output fibers. Module-1 is also divided into two parts. The first part is built into TriPleX platform and consists of four fiber inputs, each of which may be switched to four distinct outputs through Mach-Zehnder interferometers and a PZT-based phase shifter. As a result, the first half consists of 4 sets of 3-MZIs in “tree topology”, resulting in 16 waveguide channels at the output. Using PolyBoard multilayer technology, the second part is completed in PolyBoard by integrating all 16 channels into 4 output fibres with 2 (in version 1) or without (in version 2) waveguide crossings. Figure 3 depicts the functional design proposed in the 3PEAT proposal.

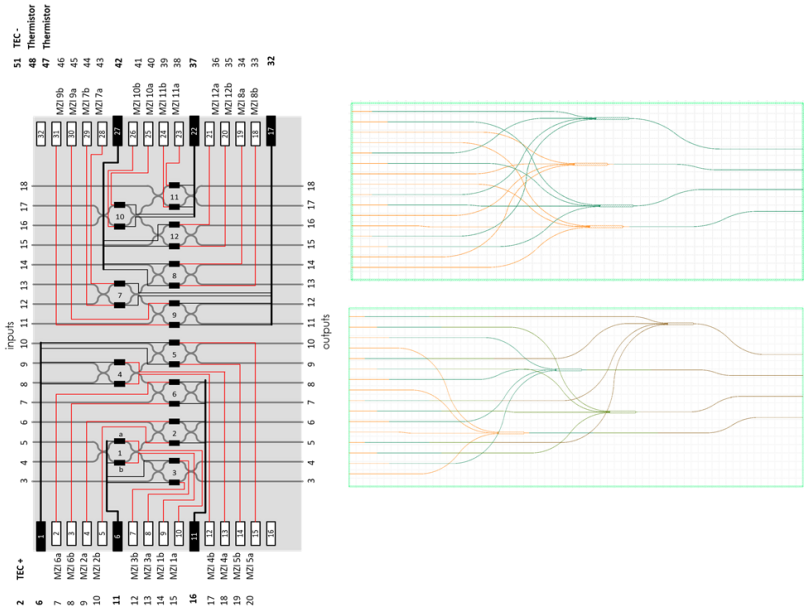

Figure 1. Representative sketch of 3PEAT Module-1.

TriPleX part

Because the efficiency of the PZT-based phase shifter in a top-top configuration was not as expected theoretically, it was decided to employ a 2 cm electrode length and a 5-um spacing between the electrodes for the PZT-based phase shifter. Because there was enough space on the mask, the MZIs were made in a push-pull arrangement, with an electrode in each of the MZI arms. Figure 2 shows a screenshot from the mask design of a MZI with two PZT-based phase shifters. Using two electrodes in MZIs has the advantage of allowing a second electrode to be employed in the event of low phase actuation or low yield.

Figure 2. Mach-Zehnder interferometer as used in Module-1. The waveguides are presented in red and the electrodes in yellow.

In the next figure, It can be seen the final design of the TriPleX part of Module-1, presenting the 12 MZI in a “tree topology”.

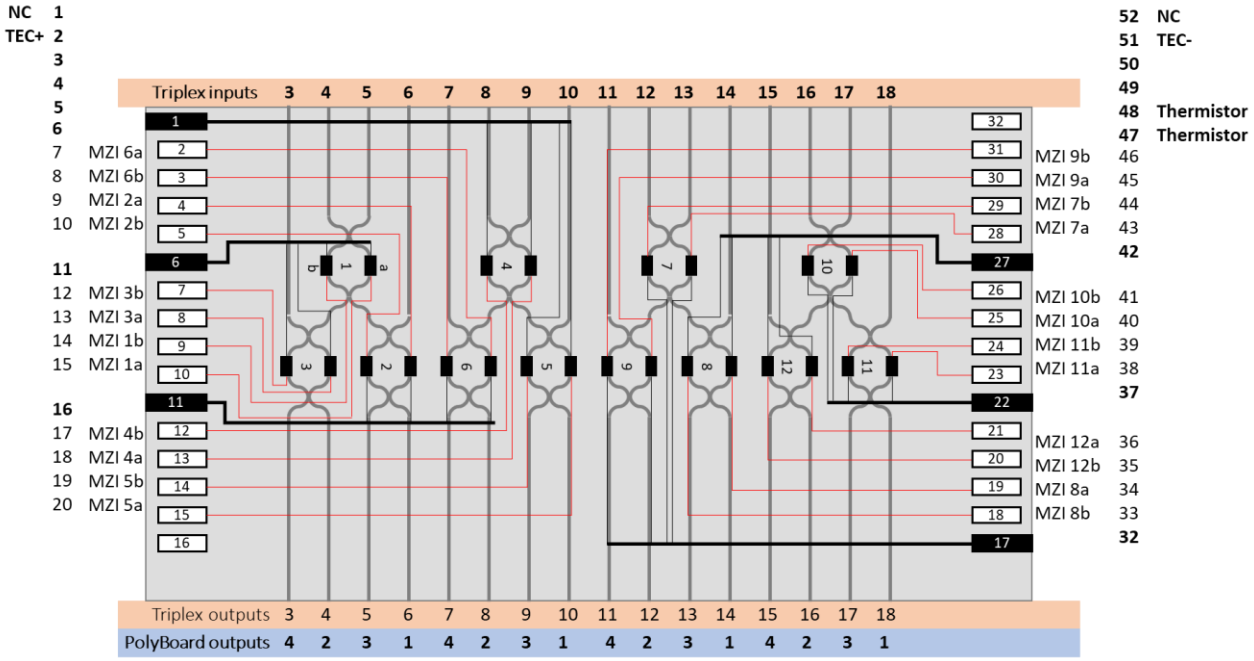

Figure 3. Final sketch of the TriPleX part of Module-1.

PolyBoard part

The envisioned concept of 3PEAT was to develop integrated optical switches without waveguide crossings. The optical crosstalk between the possible light paths is reduced when

waveguide crossings are not present. Unlike optical losses, which can be compensated for with in-line amplifiers, optical crosstalk is difficult to manage after it has occurred.

Coupled out of the TriPleX part, the light is linked at 16 locations (4 from each input) into the PolyBoard part. Waveguides connected to the 4 inputs are combined and routed through four 4x1 MMIs. A 4x4 switch can be realized in this manner since each input can be connected to each output. 2 versions of multilayer PolyBoard parts have been designed and fabricated. Version-1 (Figure 4) has 4 waveguide layers without any crossing, and version-2 (Figure 5) has maximum 3 crossing per wavelength path.

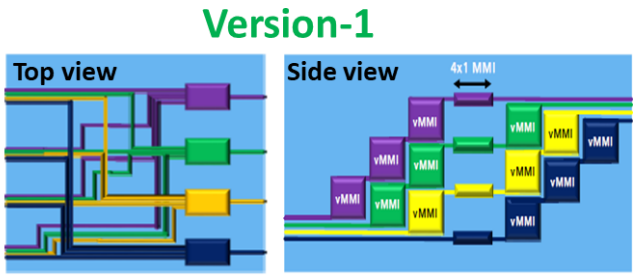

Figure 4. PolyBoard part of Module-1 with 4 waveguide layers.

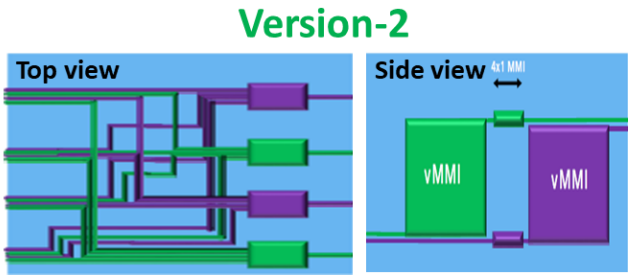

Figure 5. PolyBoard part of Module-1 with 2 waveguide layers.

For both implementations, the key photonic component was the vertical MMI, which allows the routing across the different waveguide layers. Details on the vMMIs can be found in deliverable D4.3.

On the right-hand side of the PolyBoard part, because it is wavelength independent, the 4x1 MMI was implemented as a combiner instead of the initially envisioned arrayed waveguide grating (AWG), a feature desired from the systems application perspective of the end-user Mellanox. The 4x1 MMI suffers an intrinsic loss of 6 dB as a result of its splitting ratio. In a realistic 4x4 AWG, however, losses of 4 dB are expected, reducing the loss penalty to 2 dB.

3. Fabrication

The waveguide geometry for the TriPleX part, was chosen to be the planarized ADS, as shown in Figure 6. Because it is preferable to fabricate PolyBoard on top of TriPleX, planarized ADS was chosen to achieve an efficient vertical coupling to PolyBoard.

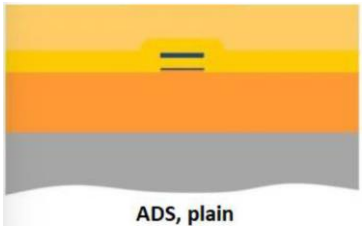

Figure 6. TriPleX waveguide geometry for Module-1.

The final mask design of the TriPleX part of Module-1 is presented in Figure 7. The TriPleX part has a total of 22 input waveguides, including 4 alignment loops, 2 has been used to determine the TriPleX to PolyBoard coupling losses, 4 as input ports, and 12 test ports for testing the performance of the PZT-based phase shifters in the MZIs.

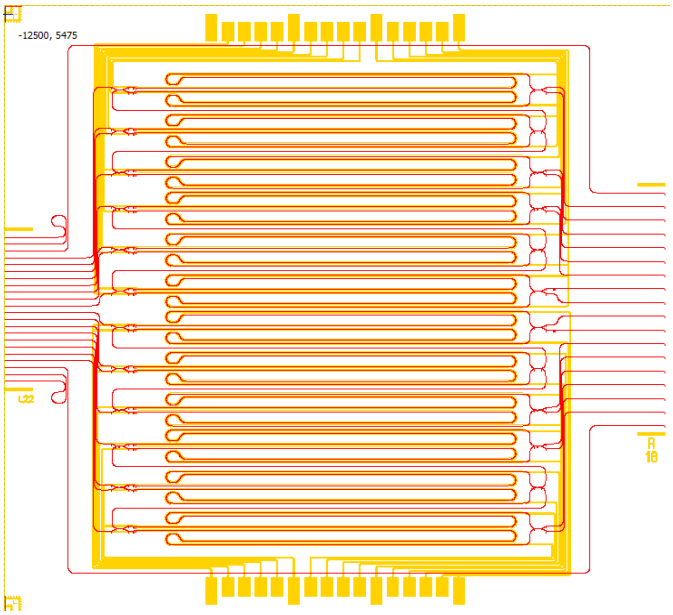

Figure 7. Mask design of the TriPleX part of Module-1.

Regarding the PolyBoard part, the final masks are presented in the next figure.

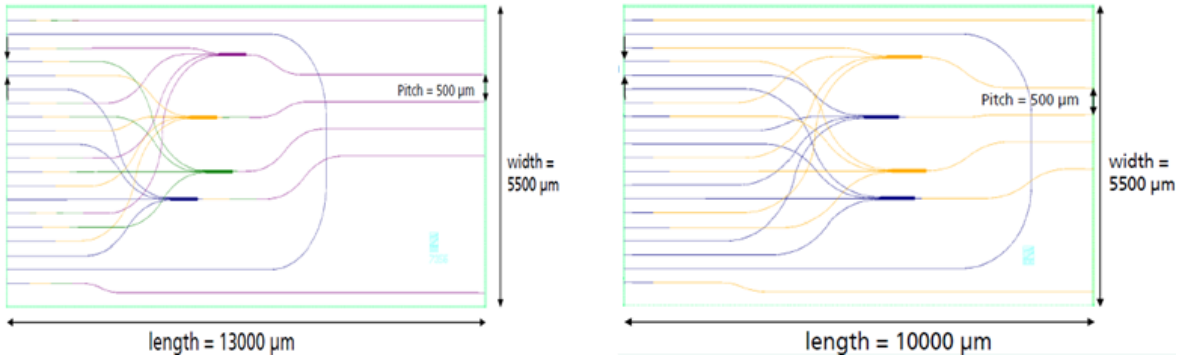

Figure 8. Mask layouts for the 4-waveguide layer (left) and 2-waveguide layer (right) -based PolyBoard part of Module-1.

Regarding version-1 for example, the first waveguide layer is represented by blue, the second by green, the third by orange, and the fourth by red. The bottom layer contains all 16 coupling points to the TriPleX part, while the top layer contains all 4 output ports. The chip is 13 mm in length and 5.5 mm in width. Two waveguides are located at the bottom and top of the screen for reference and alignment. The shifting colour along the light channel implies that these are also interconnected by vMMIs in different waveguide levels. A chip-to-chip connection was realized using an extra U-shaped waveguide in the lower layer.

4. Packaging

The "gold box" technique was employed for the switching Module-1. In this case, the PIC that needed to be accessed both electrically and optically are housed in a metallic casing with proper connectors for external instrumentation. The module provides great reliability and protection for the PICs while allowing for easy optical and electrical connections. Because of the limited space, this method is best suited for medium-complexity assemblies of a small size where the processing electronics do not need to be close to the chips. The assembly is thermally managed using a Peltier thermoelectric cooler (TEC) located inside the module. Because the power dissipation of the PIC is usually insignificant, a specialized heat sink is not required, and the heat is ejected from the system by the metal package itself. A thermistor is installed inside the "gold box" near the PIC to measure the temperature in real time. Figure 9 shows a 3D view of a "gold box" package.

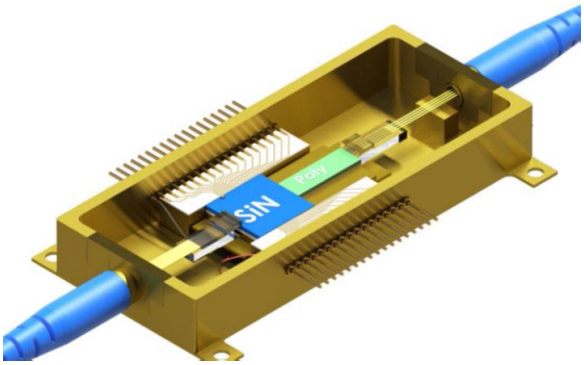

Figure 9. Schematic of the “gold-box” packaging approach.

PZTs and thermal actuators incorporated on the Triplex chip are the components to be electrically connected in Module-1. The heaters are controlled by DC-like signals, while the PZTs actuators are updated on a millisecond period, with rise/fall (10-90%) time of around 300 nanoseconds. The usage of RF connectors and feedthroughs is not required due to the low frequency of these signals. External electronics signals are thus introduced inside the “gold-box” (Kovar) package through normal metal pins on the north and south sides. All connections were made with a manual ball-wedge machine, with the ball realized on the PIC side to reduce pressures on the pads, due to the size and number of chip pads. Despite the fact that the length of the connections is not crucial due to the sluggish speed of the signals to be delivered, short wire-bondings of no more than a few millimetres are preferable to increase the assembly's durability.

Electrical interposers constructed of aluminium nitride are used to match the pitch of the metal pins to that of the Triplex chip pads. The introduction of interposers also allows for a significant reduction in wire-bonding length. To link the actuators on the chip, a series of two wire-bondings is necessary. The package's metal pins are first attached to the interposer, which has tracked that fan out to match the pitch of the chip pads and metal pins. The signals from the interposer are subsequently delivered to the chip via a second wire-bonding.

A Peltier TEC is placed below the chips in the package, and the temperature is measured using a thermistor on top of the Triplex component. Between the chips and the TEC, a highly thermally conductive copper-tungsten submount enables for efficient heat transfer. The connections for driving the TEC and reading the thermistor are made to specialised metal pins on the package's north/south sides.

Because there were no sensitive gadgets in the assembly, there were no crosstalk difficulties. As a result, no special electro-magnetic shielding measures were required throughout the construction.

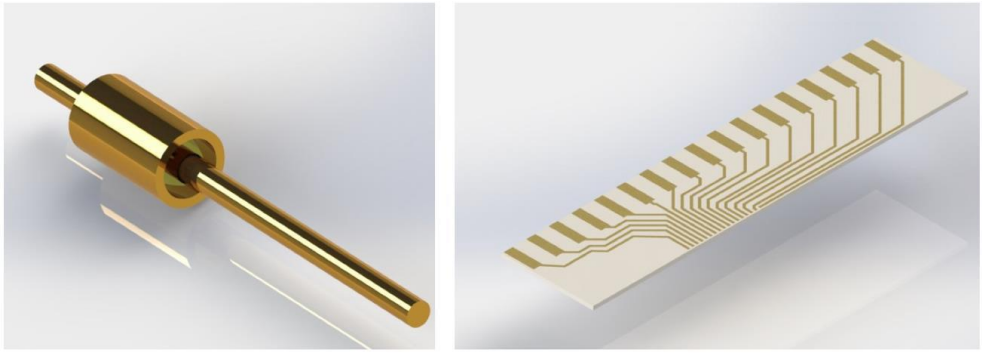

Figure 10. Metal pins of the Kovar package (left) and the aluminium nitride electrical interposer (right) in 3D (right).

5. Electronics

The first generation of active switching module driver boards was based on an array of high voltage driving amplifiers with 16 output channels that were incorporated on a breakoff PCB board. For waveform synthesis, this break out board was connected to a DAC evaluation board operated by an FPGA chip and an ARM coprocessor. The graphic below shows a block diagram of the driving electronics.

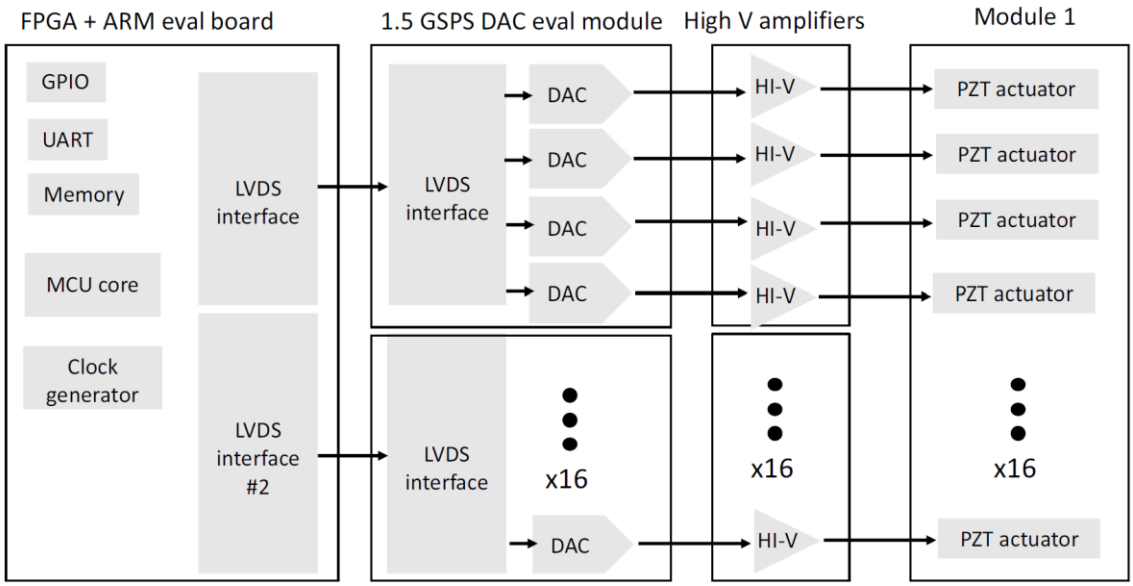

Figure 11. Block diagram of the first-generation driver board for active switching modules.

In an updated version, an architecture was created using the same FPGA evaluation board. All the channels that would be used in Module-1 could be accommodated. The digital element consists of two firmwares, a graphical user interface written in Python that runs on a desktop and controls the setting of each channel of the module to ON or OFF, which corresponds to a CROSS or BAR setting. The settings are converted to bytes and sent to the FPGA evaluation board via UART. The processor side (ARM) of the Zynq SoC is employed on the FPGA side. The bytes are received by this SoC UART, which is decoded to the precise setting for each switch, which is then translated to a LOW or HIGH state for the Zynq's GPIOs. The firmware for the Zynq ARM is written in C++ and uses Xilinx libraries.

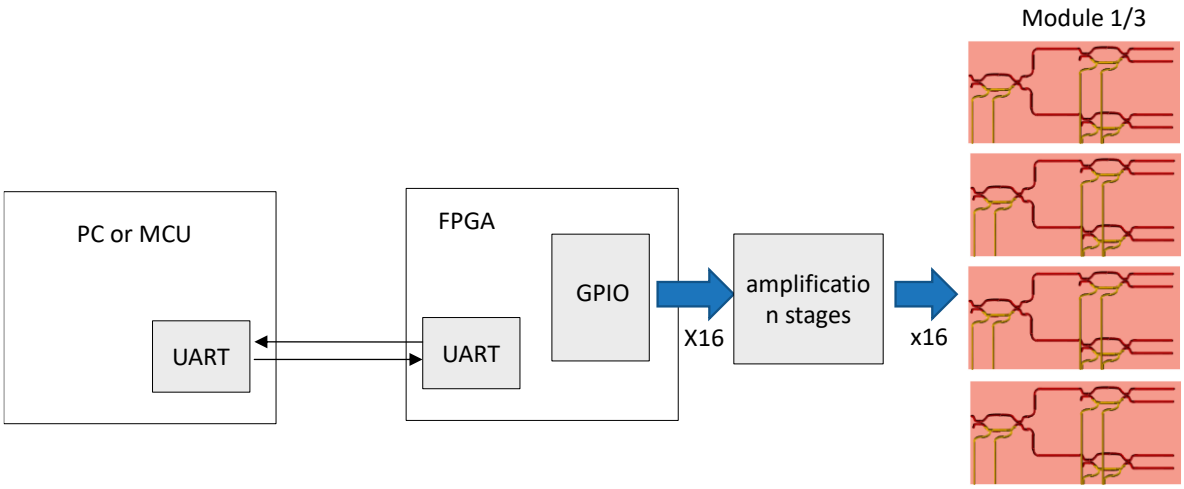

Figure 12. Block diagram of the implemented architecture for the control of the 3PEAT switching Module-1.

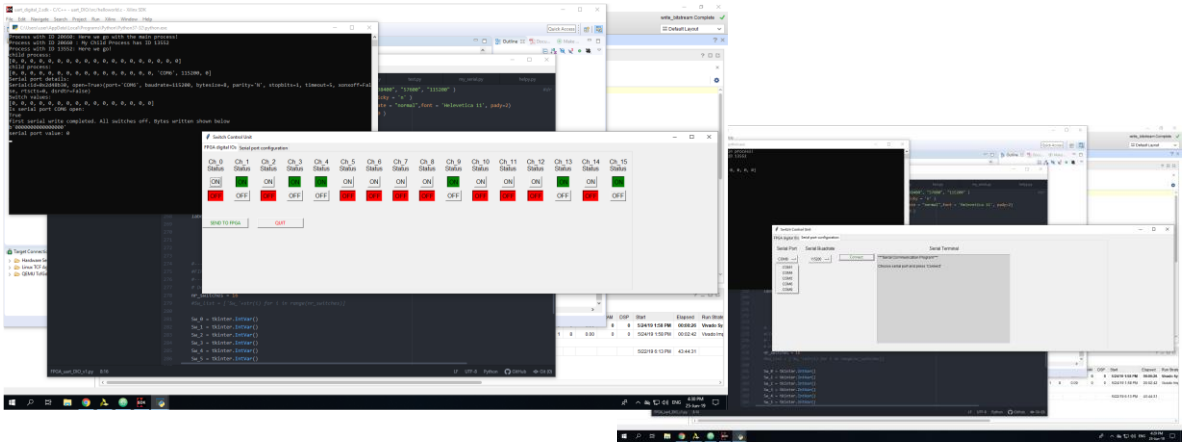

Figure 13. Example of the Graphical User Interfaces developed in Python.

The initial circuits investigated were low-speed (5 kHz) circuits with output voltages greater than 100V. These were built around a 32-channel driving chip that came in the form of five separate boards that could be stacked to drive up to 160 PZT components. A microprocessor and a DAC gave digital and analogue signals to control them. The figure below shows a snapshot of the developed driver boards.

Figure 14. 32-channel High-voltage, low-speed driving circuits

These were combined with the uController and the DAC on a single PCB board to create a standalone driving solution. The figure below depicts the combined version.

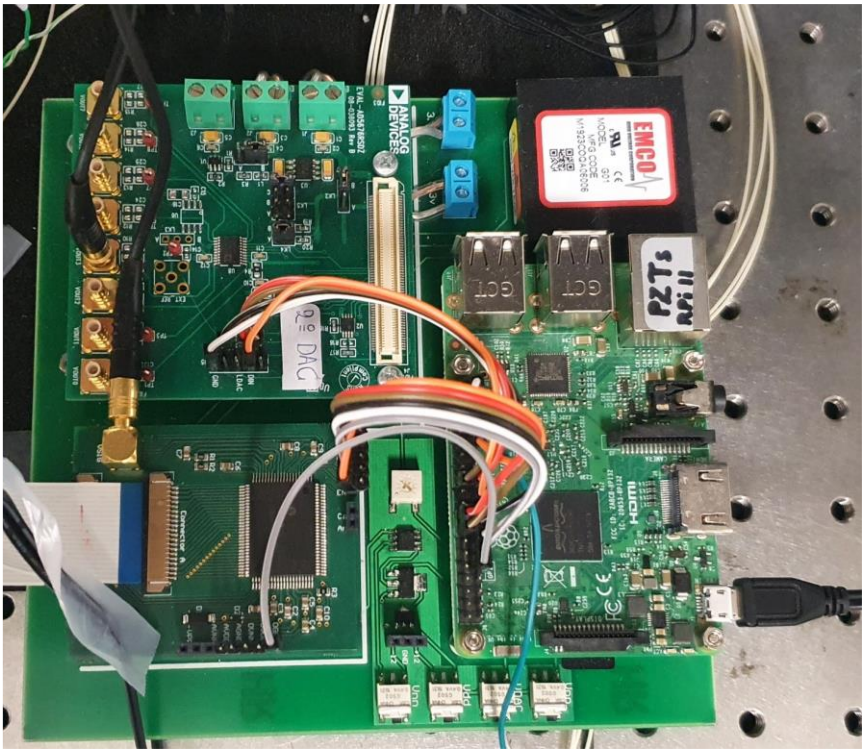

Figure 15. Integrated version of of high-voltage drivers for 160 channels.

Another driver for top-bottom PZT elements was created that could produce up to 60 V of output voltage and had a better bandwidth of around 100 kHz, allowing it to drive components with capacitance of roughly 5 nF. The rise and fall times were measured in tens

of nanoseconds. These were made in eight-channel arrays, with four such arrays totaling 32 channels.

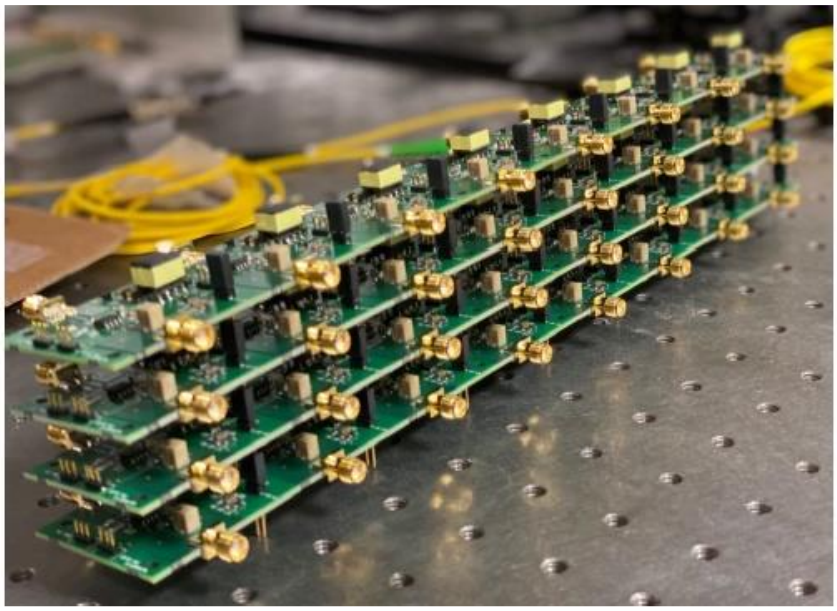

Figure 16. Intermediate solution for the PZT driver circuits (four arrays of 8 drivers each).

For the above drivers, a breakout board with inverting circuits at the inputs was constructed, allowing them to accept the LVCMOS signal (0-3.3V) from the digital outputs GPIOs of the FPGA as an input and invert it (-3.3 – 0V) before entering the driver-amplifiers. As a result, the output signal ranges from 0 to -60V. This solution was created in the event that driving the PZT elements with positive signals had an impact on their longevity, which was later discovered not to be the case. The breakout board that was created as well as the findings of an oscope may be seen in the diagram below.

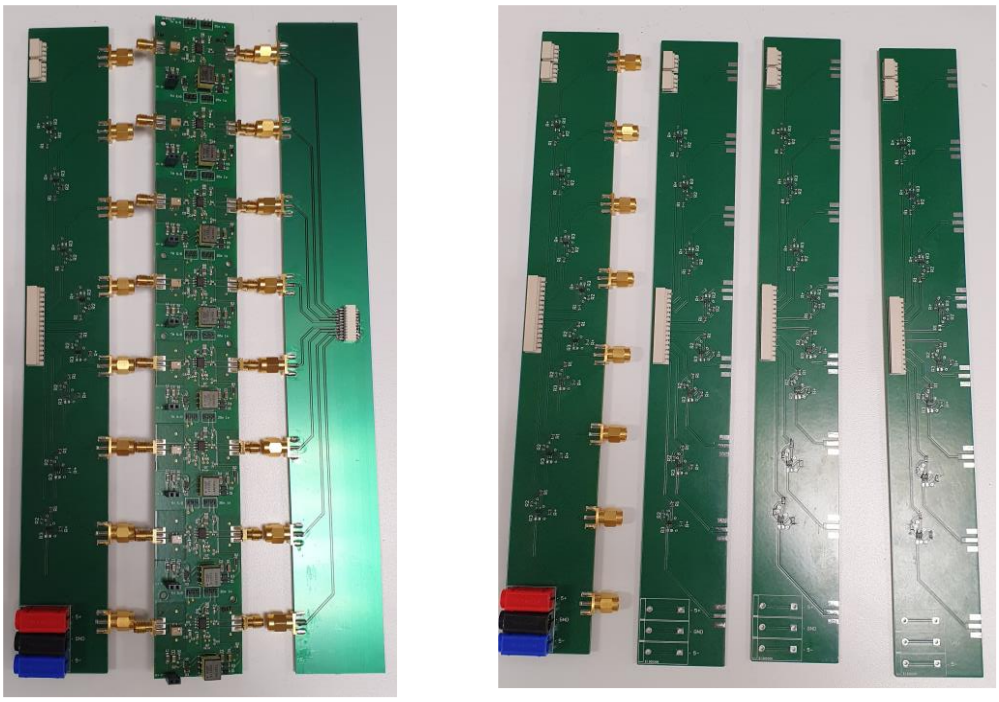

Figure 17. Fabricated breakout board with 4x8-fold inverting circuits (left) coupled to an 8-fold driver array (right) the many PCBs that were fabricated.

6. Static characterization

The static characterization was completed in ICCS premises. In Figure 20 the experimental setup is presented. The first step of the characterization was the voltage apply with up to 130 volts on the 2 arms of each MZI (where the PZT have been deposited), in order to measure and capture the transfer functions of the different photonic components. In Figure 18, the corresponding results are presented, referring to the version of Module-1 that contains PZTs with “top-top” deposition, and multilayer PolyBoard with 4 waveguide layers. The insertion losses that have been measured were 18.85 dB. As it can be observed in the figure, the maximum voltage that has been applied was 130 volts. Based on the results, the Vπ was estimated around 100 volts.

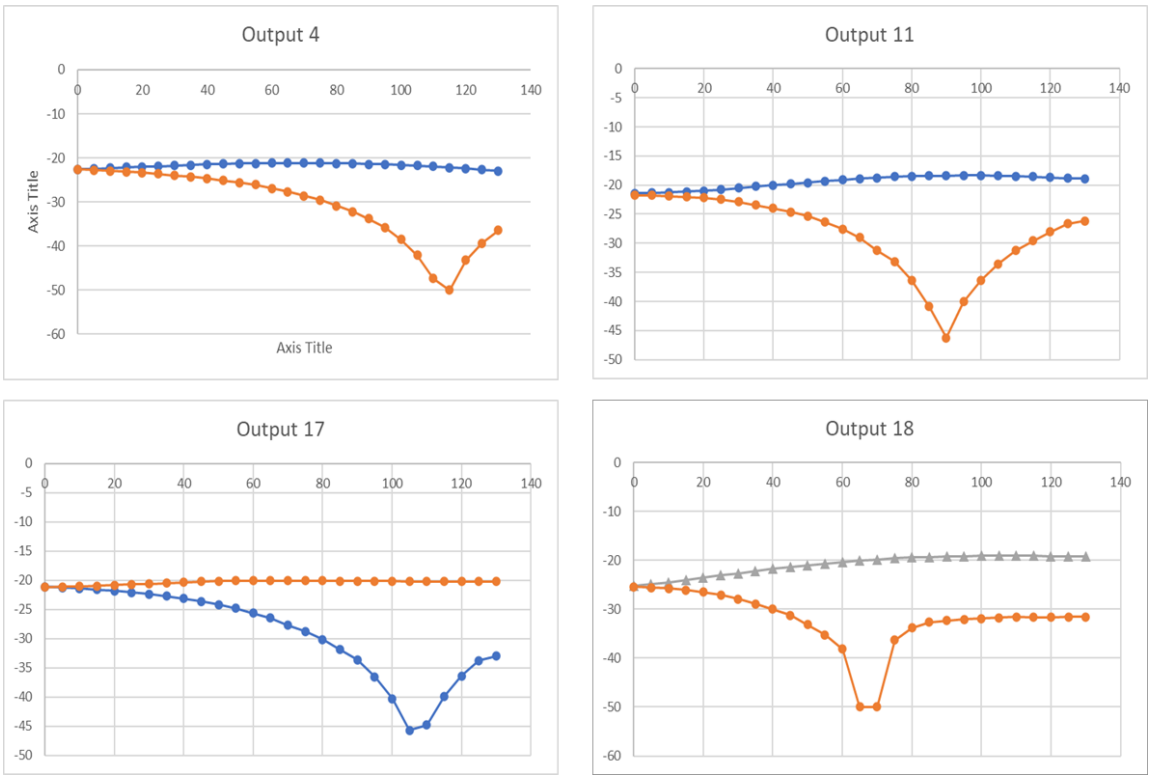

Figure 18. Representative transfer functions of different MZIs on Module-1, based on "top-top" PZTs.

Same as before, the next figure presents representative results of the MZIs' transfer functions, that are integrated in a second version of Module-1, which contained PZT with "top-bottom" configuration, but again, based on a 4-waveguide layer PolyBoard PIC.

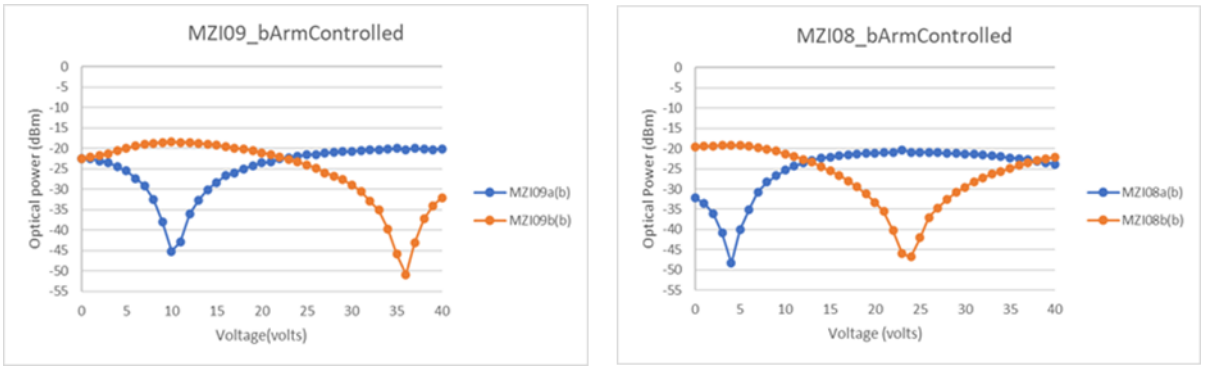

Figure 19. Representative transfer functions of different MZIs on Module-1 based on "top-bottom" PZTs.

In the case of "top-bottom" PZT based Module-1, similar insertion losses have been measured, in the order of 19.51 dB per selected path. Due to the different type of the PZTs, lower voltage was needed in order to achieve the switching state. In this case, a Vπ of almost 40 volts was needed.

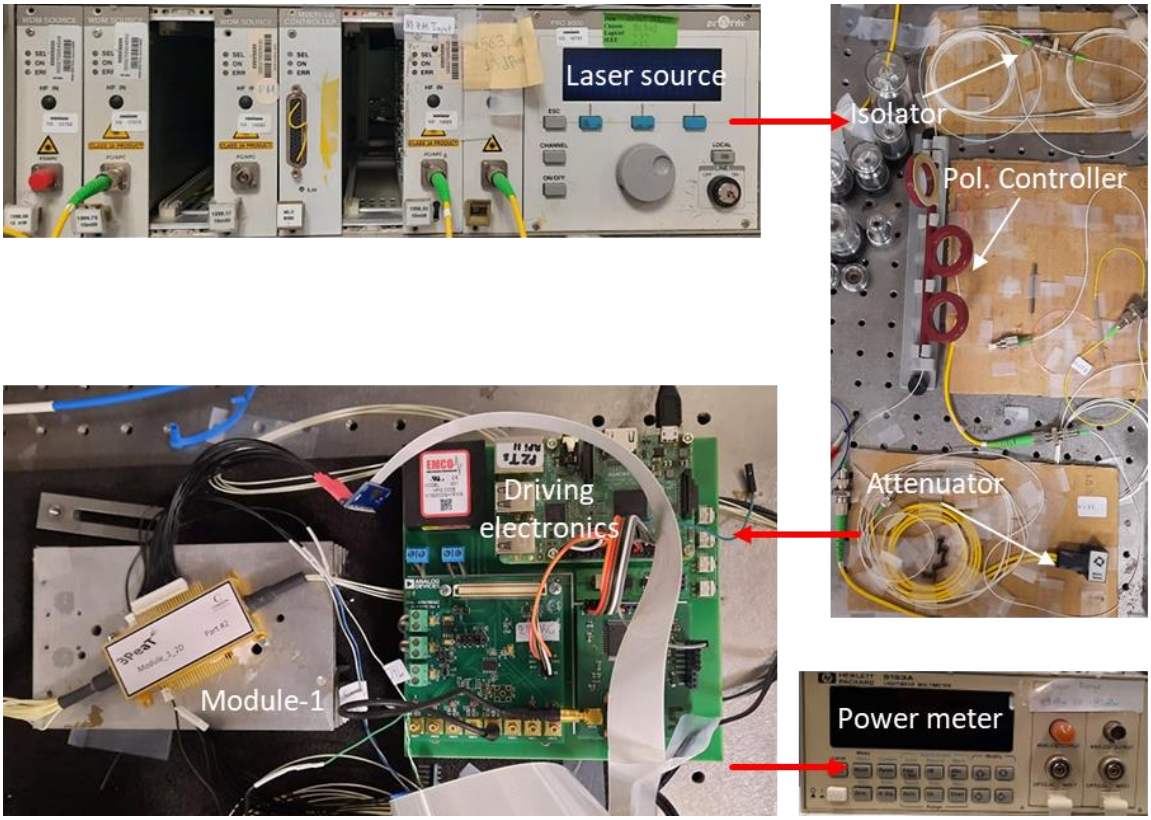

Figure 20. Experimental setup, developed in ICCS premises for the static characterization of Module-1.

The best operating MZI from the version of Module-1 with “top-top” PZTs and multilayer PolyBoard with 4 waveguide layers, in terms of extinction ratio (ER), insertion losses and applying voltage that is needed for reaching the Vπ, were chosen in order be part of the system test that is described in the next paragraph.

As a next step of the testing and characterization procedure that took place in ICCS premises, was the apply of driving signals on the best performing MZI of the before mentioned Module-1. In the next figures, the corresponding results are presented.

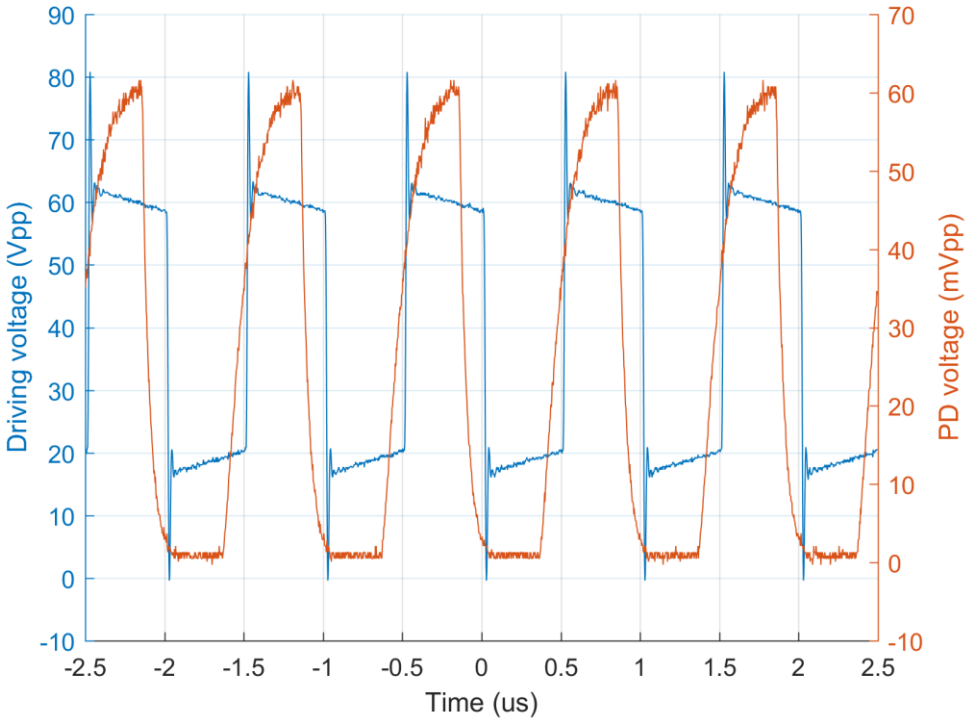

Figure 21. Driving square signal with frequency of 1 MHz (blue). In orange, the switching response of Module-1 on an external PD is presented, with ER of ~60 volts.

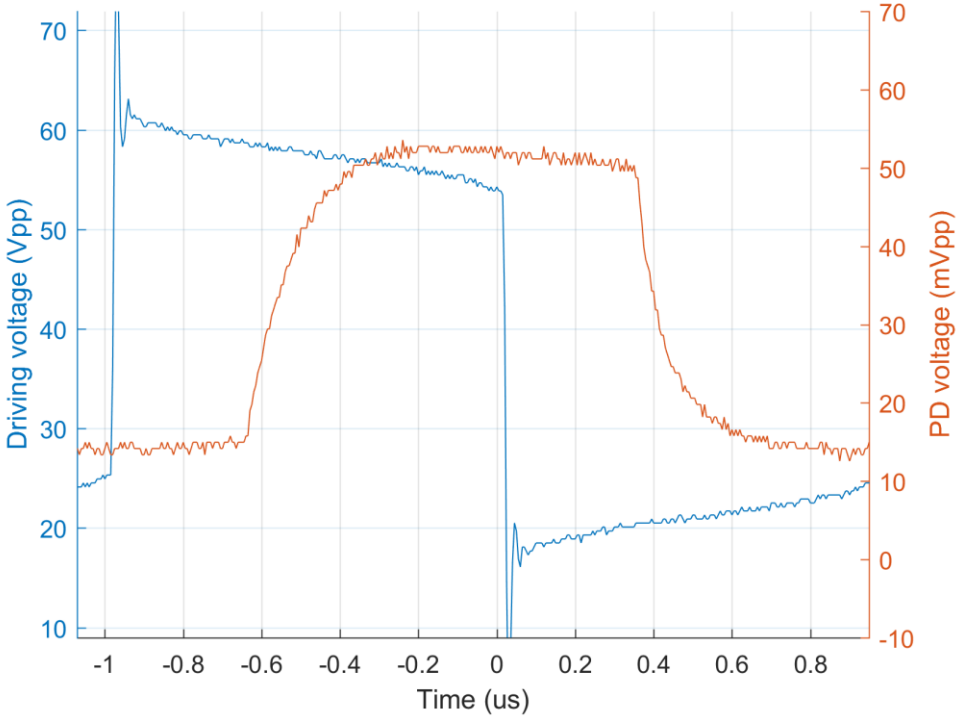

Figure 22. Zoomed-in section of the previous graph, that presents a 10-90% switching time of 300 usec.

As the last step of the characterization, was to drive 2 cascaded MZIs simultaneously, and observing the switching response of Module-1 again in an external PD. The next figure presents the corresponding results.

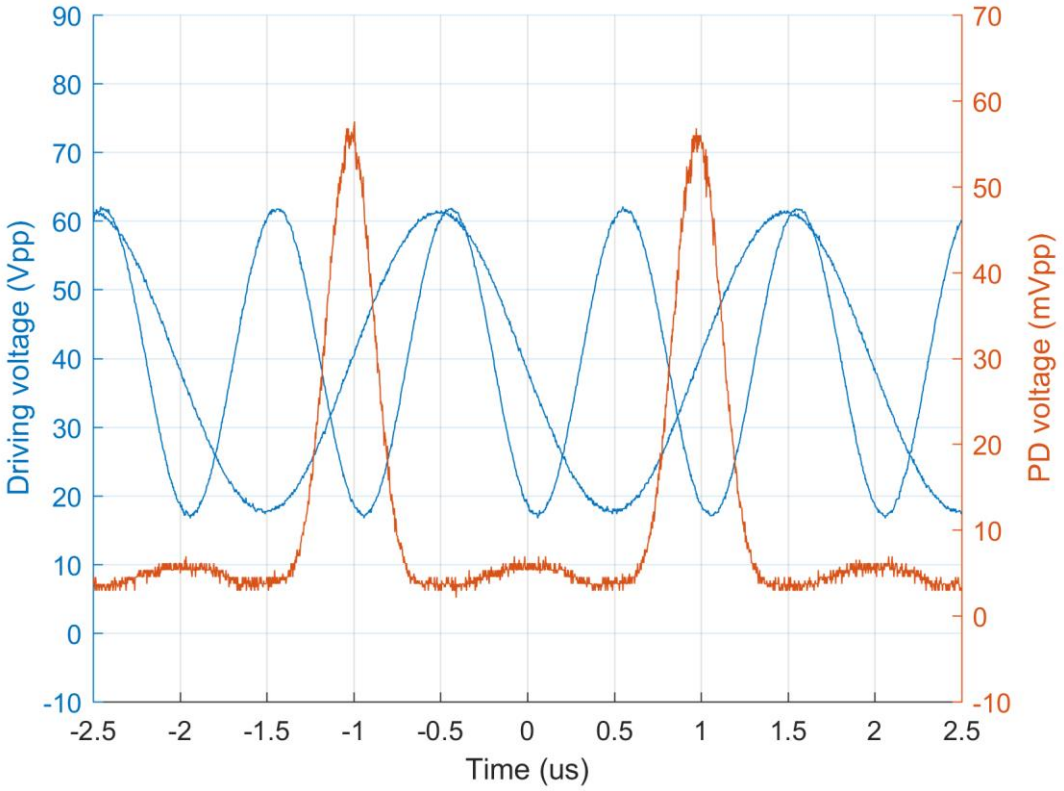

Figure 23. Switching response of Module-1 when 2 cascaded MZIs are driven with sinusoidal signals of different frequencies.

7. System characterization

The system characterization was realized in Mellanox premises. In Figure 24, the experimental setup is presented. ICCS, Optagon, and MLNX worked together to install the switch prototype (Module-1) in Mellanox's Athens headquarters. Module-1 was tested and shown with commercial networking equipment in this activity. The overview of the demo is shown in Figure 25. Optagon's controller FPGA board controls the optical switch and periodically allows alternate pathways. As a result, for the timeslot t, "server 1" is connected to "server 2," and for the timeslot t+1, "server 3." A file transfer application is running on the three servers depicted in Figure 25.

A MLNX ConnectX NIC and a C-band pluggable transceiver are attached to each server. The experiment started with demonstrating link bring up through the optical switch: commercial NICs and pluggable TRxs have certain link quality requirements that the optical

switch must meet. The link was established using the two pathways needed to connect "server 1" to "server 2" and "server 3," respectively.

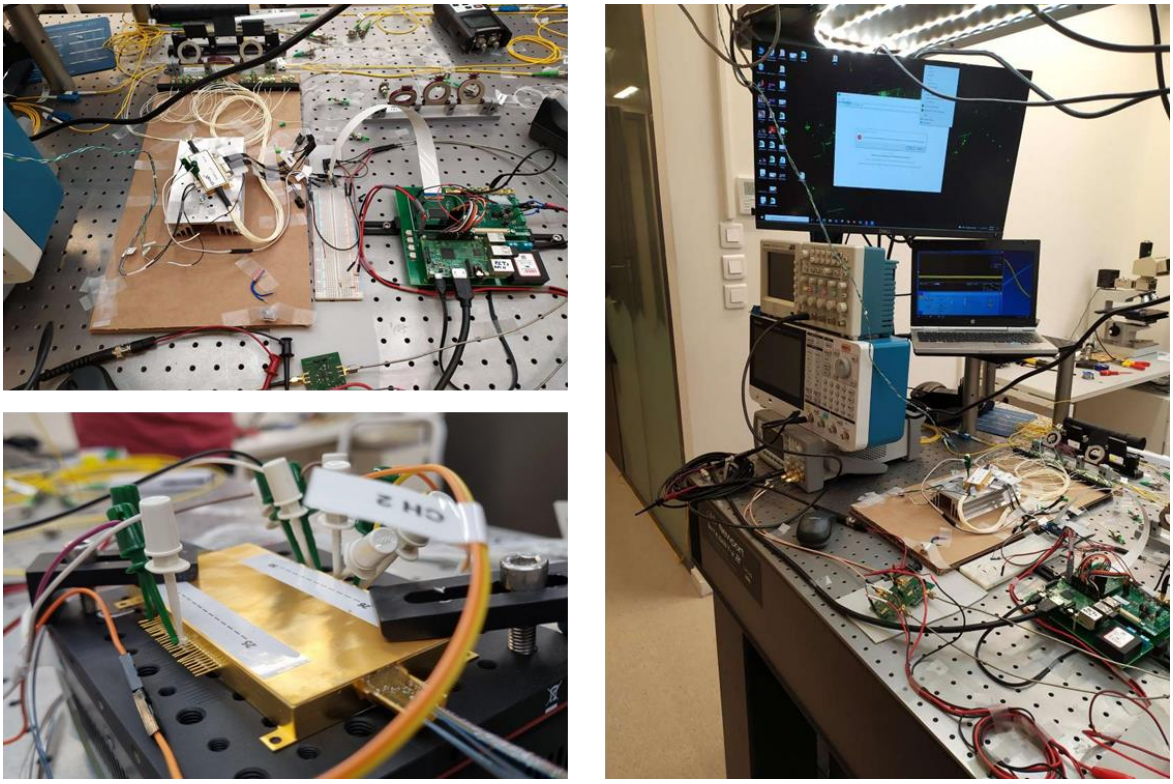

Figure 24. Experimental setup in Mellanox premises for the system test of Module-1.

Because the optical switch's insertion loss is greater than the TRxs' margin, EDFAs (Erbium-Doped Fiber Amplifiers) were added to the pathways between the servers to boost the optical signal. The experiment's second step required switching between servers.

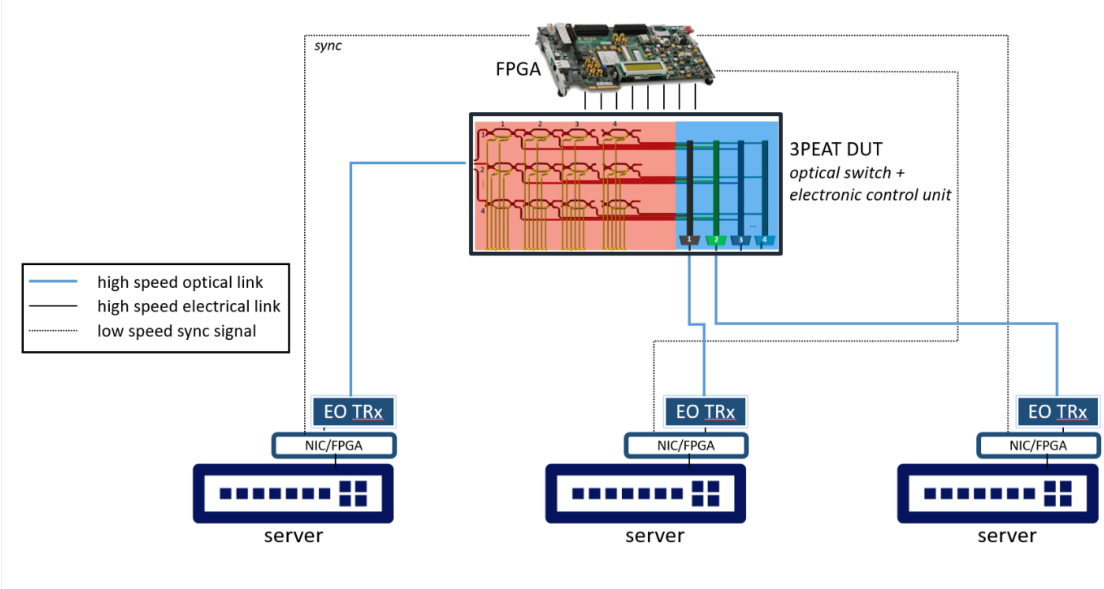

Figure 25. Scenario for the system evaluation of Module-1.

Figure 26 illustrates the incoming and outgoing traffic recorded on the servers' network interfaces. The incoming and outgoing traffic for "server 1" is shown in the top panel, with IP address "10.0.1.1." The data gathered on "server 2" interfaces is shown in the middle of the diagram, while "server 3" is shown at the bottom. As seen by the sharp symbols that represent the data in Figure 86, "server 3" initially transmits data that is received by "server 1." Following the outgoing and receiving data for the three servers can reveal the switching. Servers 2 and 3 send data one after the other, and Module-1 is set up correctly for the data to reach server 1.

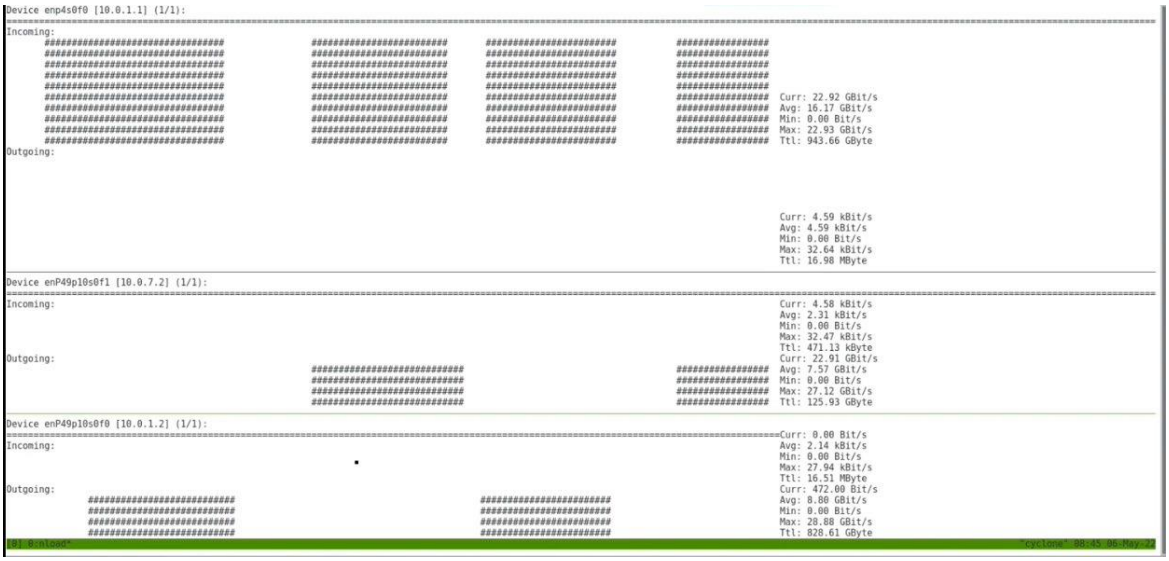

Figure 26. Data transmission among the servers, top: "server 1 - 10.0.1.1", middle "server 2 - 10.0.7.2" and bottom "server 3 - 10.0.1.2". The sharp symbols on the figure show the incoming and outgoing data captured by nload on the interfaces of the servers.

The acquired bandwidth is shown in the right-hand columns. We were able to transmit usable data at a rate of 23 Gbps, demonstrating that the experimental setup's link quality met the standards of commercial pluggable transceivers (25 Gbps pluggables were used for the experiment). It should be mentioned that in these studies, we used commercial networking equipment (NICs, pluggable transceivers) to demonstrate the 3PEAT switch prototype's compatibility and thus its practical applicability. Due to the specifications of these devices, which are designed to interact with commercial network protocols, using commercial equipment in the experiment imposes certain limits. The key impact is the sluggish link bring-up time imposed by both pluggable transceivers and NICs, which is caused by the combined effect of CDR lock, link training, PCS layer, and network protocol that are not built for slotted traffic. The network was run with extended timeslots (1-2s duration) to circumvent these limits, simulating the bandwidth steering use case described in the publication with DOI: [10.1145/3295500.3356145](https://doi.org/10.1145/3295500.3356145).

8. Development of the SDO agent

Modules 1 and 3 were conceived as fast (ns range) active optical circuit switches , based on the architecture presented above. Placing them in a real network environment would mean that the optical switch would be integrated with Software Defined Network architecture in the form of the following figure.

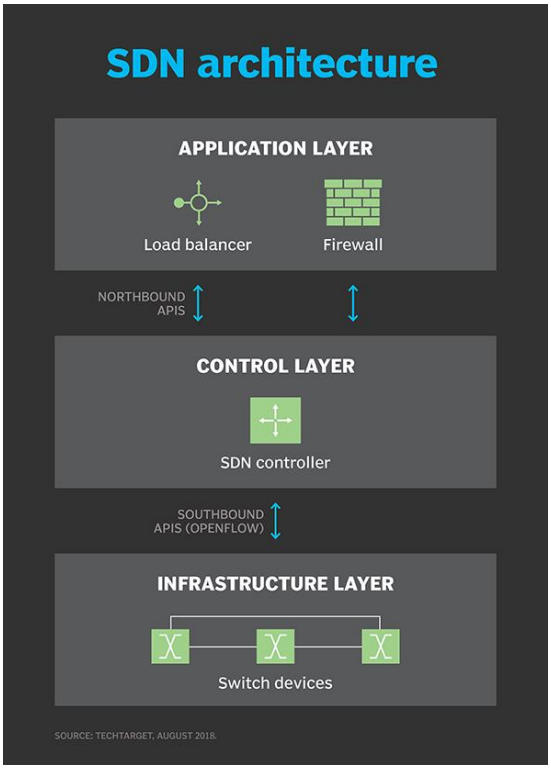

Figure 27. Typical SDN architecture. 3PEAT Module-1 would be placed in the Infrastructure layer.

Without going into detail, the above figure presents that Module-1 would find its place in the infrastructure layer (or data plane) and would be controlled by a SDN controller on the control layer (or control plane). Usually the devices in the infrastructure layer run SDN agents, whose purpose is to provide interface with the control layer (SDN controllers) and efficiently communicate with them. In the case of optical switches and devices that are closer to the physical layer the term Software Defined Optics (SDO) is used instead of SDN agent, and essentially performs the same task. To provide an interface to the control layer so that it can communicate and be controlled. The firmware that was developed for the control electronics, could be interfaced with slight modifications to interface with the SDN controllers (including OpenFlow protocols) as it will be explained below.

The PZT development underwent a lot of iterations, and two different electrode configurations were investigated in this project, namely “top-top” and “top-bottom” with each one having different driving requirements. The electronic drivers for the PZT electrodes that were presented in the above sections were based on different implementations depending on the driving voltage requirements and speed. Module-1 and switching speed was specified at 2.5 MHz due to the bandwidth targeted for the PZT electrodes. This would require an electronic control architecture based on FPGA, fast DACs and analog amplifiers/drivers that could amplify the signal of the DACs to a high amplitude (max 70 V for up to tens of MHz according to the analog drivers presented in Deliverable D6.2. Regardless the bandwidth of the PZTs and depending on the switching speed requirements of the optical switch, meaning how often does the optical switch need to change states and reconfigure its ports, this switching speed is also important in selecting the appropriate driving solution. In case MHz switching speed is required an FPGA solution should be used. The optical switch elements could be controlled by either the GPIO pins of the FPGA (after amplifying the LVCMOS signal accordingly) or the signals coming out of the DAC channels. In the first case the GPIO state HIGH or LOW would directly translate to a switch state CROSS or BAR. This approach is straightforward and takes advantage of the many GPIO pins present on the FPGA evaluation board. In the second case again the same applies, but the extra capability of defining the analog amplitude of the DAC provides more flexibility if needed. As it was already presented in the previous sections of this deliverable a firmware was developed in the FPGA for controlling the optical switch. It consisted of a VHDL firmware written that required an input in the form of bytes received from the UART port on the FPGA and according to the data bytes received, it would switch the GPIO pins accordingly to realise the routing state required. The firmware would run mainly on the processor of the FPGA and would communicate through UART, with a python GUI that run in a computer that was connected to the eval. Board. The GUI python controlled each switch crossbar switch and would then send the configuration through the serial protocol. The SDO agent in that case would substitute the GUI and would interface with the FPGA directly

through UART sending each time the necessary bytes that would configure the crossbar switches of the optical switch. Communication though UART is simple and efficient and very commonly used, as all platforms/microcontrollers provide at least one UART interface. In the case of using DACs for more flexibility, the FPGA firmware becomes more complex as a state machine is implemented that controls the output states of each DAC channel for each cross or bar configuration of the crossbar switches. The DAC sample information required to generate the necessary states must be fed to the FPGA and it is simpler to have this information hardcoded in the FPGA fabric so that the only information required to be given to the FPGA is the configuration state. This can be done through UART again, if the ARM processor of the Zynq SoC is included in the firmware, or it can be realized only in pure logic only in the FPGA fabric, by hardcoding also the optical switch routing states to the firmware. In the last case, a simple digital signal in the form of a trigger pulse is only required to switch between states.

In case other driving solutions need to be used, like the slower ones presented in the previous sections, these have firmware and a GUI that both were developed in python and run on lower spec open source microcontrollers like RPi, and essentially communicate with the analog driving parts (DACs or amplifier chips) with SPI protocol or digital signals. The firmware in that case would be slightly modified to receive the routing states and translate them to the signals that need to be applied to the crossbar switches, essentially obviating the need for a GUI.

Conclusion

The document presents an overview and the characterization and the evaluation results of Module-1. A description regarding the design, the fabrication steps and the packaging procedure have been provided. Moreover, the operation of the module has been successfully demonstrated in quasi-real settings in Mellanox premises, within an evaluation scenario as it was defined in the first period of the project. The steps that have been followed for the development of the SDO agent, are finally described.

List of figures