FP7- Grant Agreement no. 283393 – *RadioNet3*

Project name: Advanced Radio Astronomy in Europe
Funding scheme: Combination of CP & CSA
Start date: 01 January 2012 Duration: 48 month

Deliverable 8.2

Hardware Design Document

Due date of deliverable: January 2014
Actual submission date: January 2014
Deliverable Leading Partner: ASTRON

ASTRON
Netherlands Institute for Radio Astronomy

An European project supported within the 7th framework programme

SEVENTH FRAMEWORK PROGRAMME

1 Document information

Document name:	UniBoard^2 Architecture
Type	Document
Revision	0.8
WP	8
Authors	Gijs Schoonderbeek
ASTRON Doc. Nr.	ASTON-TN-040 1.0

1.1 Dissemination Level

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

1.2 Document history

Revision	Date	Chapter / Page	Modification / Change
0.1	2013-06-03	Schoonderbeek	Creation
0.2	2013-09-09	Schoonderbeek	Update
0.3	2013-09-13	Schoonderbeek	Update after Review Arpad
0.4	2013-09-18	Schoonderbeek	Power^{MX}, Distribution list, References, breakout boards
0.5	2013-09-23	Schoonderbeek	Remark Alain (multipliers in Cor.)
0.6	2013-09-27	Schoonderbeek	4 vs 2 DDR modules per node HMC info added.
0.7	2013-10-07	Schoonderbeek	Update after FPGA decision.
0.8	2014-01-13	Szomoru	Minor text edits (mainly to first section), this version sent to RadioNet3 project management as deliverable D8.2

1.3 Distribution list

Group:	**Others:**
ASTRON: Andre Gunst Eric Kooistra Sjouke Zwier	ASTRON Albert Jan Boonstra SKA SA Francois Kapp

Daniël van der Schuur Harm Jan Pepping JIVE Arpad Szomoru Jonathan Hargreaves Salvatore Pirruccio Sergei Pogrebenko Paul Boven Harro Verkouter UMAN Aziz Ahmedsaid Ben Stappers INAF Gianni Comoretto BORD Benjamin Quertier Alain Baudry Stephane Gauffre UORL Cedric Dumez-Viou Rodolphe Weber Nicolas Grespier MPG Guenter Knittel Reinhard Keller	Jason Manley Henno Kriel CNRC-NR Brent Carlson CSIRO John Bunton Oxford Kristian Zarb Adami Oxford

1.4 Terminology

ADC	Analogue to Digital Converter
BF	BeamFormer
bps	Bits per second
BW	BandWidth
COTS	Commercial Of The Shelve
DDR	Double Data Rate
EMI	Electro-Magnetic Interference
Firmware	Embedded or real-time code that runs on a microprocessor (e.g. written in C)
FPGA	Field Programmable Gate Array
GMII	GbE media independent Interface 8 bits @125MHz
Hardware	Boards, subracks and COTS equipment
HDL	Hardware Description Language
IO	Input Output
IP	Intellectual Property
IPC	Association Connecting Electronics Industries (formally Institute for Interconnecting and Packaging Electronic Circuits)
LDO	Low DropOut regularo
LUT	Look Up Table
MAC	Multiply and Accumulate, Medium Access, Monitoring and Control, Media Excess Controller (layer 2 of OSI model)
PCB	Printed Circuit Board
POL	Point of Load

PHY	physical interface (layer 1 of OSI model)
RF	Radio Frequency
RSP	Remote Station Processing (in LOFAR)
SFP+	SFP for 10GbE
Subband	Frequency band, unit output of the filterbank
XAUI	10G attachment Unit Interface (4x 3.125Gbps) interface between MAC and PHY
XGMII	10G Media Independent Interface
XFP	10G Small form-factor Pluggable transceiver

1.5 References

[1] UniBoard2 Work Package description, Arpad Szomoru, 2011-10-28, RadioNet3 283393
[2] UniBoard Evaluation, Gijs Schoonderbeek, March 2012, ASTRON-RP-1316 1.0
[3] Arria 10 Device Overview, Altera 2013.09.04, http://www.altera.com/literature/hb/arria-10/arria_10_aib.pdf
[4] UltraScale Architecture and Product Overview, Xilinx, DS890
[5] QSFP+ Cage and PT Connector Assembly, TE Connectivity, 108-127010
[6] Product Specification 40BASE-SR4/10GBASE-SR 300m QSFP+ Gen2 Optical Transceiver Module FTL41QD2C, Finisar, 24-Jul-13 Rev B5
[7] CXP Connector and Housing Assembly, TE Connectivity, 114-13283, 20 AUG 12 RevD
[8] Impact™ Backplane Connector and Cable Assembly System, Molex, http://www.molex.com/molex/products/family?key=impact_backplane_connector_system&channel=PRODUCTS&chanName=family&pageTitle=Introduction#related
[9] ERmet ZDplus, ERNI Electronics GmbH
[10] MiniPOD™ Datasheet, Avago, AV02-3438EN, July 12, 2012
[11] Experimental evaluation of Avago MicorPOD 120Gbs module, Peter Maat, December 2012, ASTRON / DOME
[12] Cyber SKA website: http://www.cyberska.org/pg/groups/27370/powermx-platform-working-group/
[13] InfinX High Speed Mezzanine Connector, Amphenol TCS
[14] VSC7389 SparX-G16 Datasheet, Vitesse Semiconductor Corp, VMDS-10187 Revision 2.4 December 2006
[15] AVSP-1104 Datasheet, Avago Technologies, version 1.4.3, November 2012
[16] VSC3144-12 Product brief, Vitesse Semiconductor Corp. VPPD-02743 Revision 1.0 2011
[17] BCM56540 Product brief, Broadcom, 56540-PB01-R April 28,2011
[18] Micron HMC: http://www.micron.com/products/hybrid-memory-cube/short-reach-hmc
[19] Hybrid Memory Cube Specification 1.0, Hybrid Memory Cube consortium, 2013
[20] Hybrid Memory Cube consortium: http://www.hybridmemorycube.org/

1.6 Content

2 Introduction

2.1 Scope

UniBoard^2 is a Joint Research Activity (JRA) in the RadioNet3 project [1], funded by the EC through the FP7 programme, under grant agreement no. 283393. The partners in this JRA are the Universities of Bodeaux and Orleans, INAF, MPG Bonn, the University of Manchester, ASTRON and JIVE. This document started out as an overview of several options for the high-level hardware design of UniBoard^2 and has served as a discussion piece in the collaboration. It summarizes a number of board design options, incorporating the inputs and suggestions of several of the collaborators. This final version, which is submitted to the RadioNet3 project management as deliverable D8.2, reflects the consensus on the high-level design and functionality of the hardware platform among the UniBoard^2 partners. This being a high-level design document, throughout the actual hardware design phase several more in-depth technical documents will be produced and published on the RadioNet3 wiki.

2.2 System functionality

In Figure 1 the block diagram of a general beamforming or correlator architecture is shown. This functionality can be mapped on single UniBoards, or mapped on to multiple boards on a backplane.

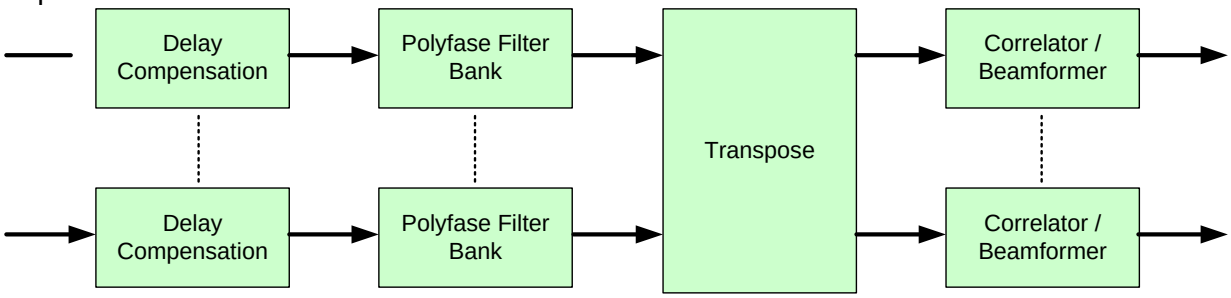


Figure 1 System block diagram

At the input side, data from multiple sources are aligned and fed into a filterbank. In the filterbank the frequency band is split into multiple channels. The channels are transposed from all frequency channels of one signal input to all signal inputs of one frequency channel. This transpose function can be partly implemented in firmware (on the FPGA) and in the routing between the FPGAs. In the last stage the data from the different inputs is processed (beams are formed or correlation products calculated).

2.3 General Design Standpoints

For the design of UniBoard^2, a processing system for antenna signals, we want to use the >50% principle. This means that:

- >50% of the power consumption goes to the processing of the antenna signals
- >50% of the cost goes to the goes to the processing of the antenna signals
- >50% of the design time goes to the development for the processing

2.4 UniBoard1

In Figure 2 the first concept is shown. This concept is based on the first UniBoard [2], which has the maximum possible number of FPGAs on a reasonably sized board with a semi-full mesh connecting the FPGAs. On one side off the board standard optical interconnections are placed, on the other backplane connectors. Through these connectors board to board connections can be established via a backplane.

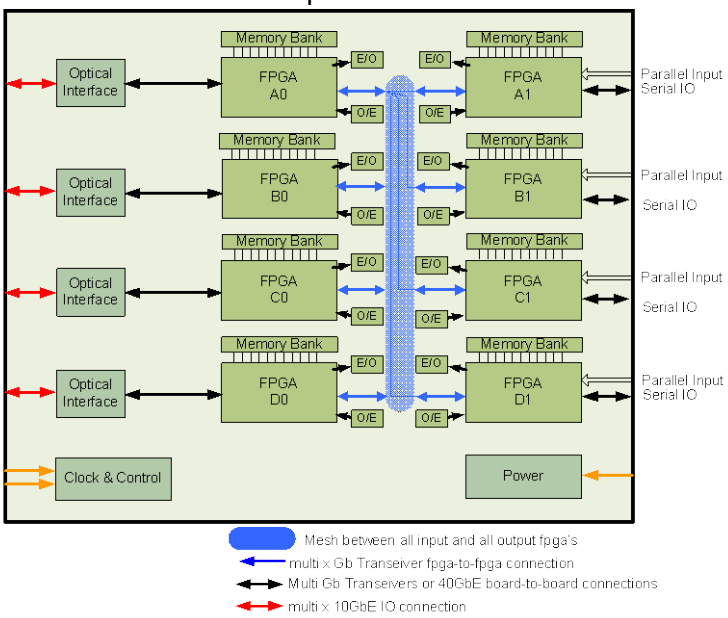


Figure 2 Concept 1 UniBoard like

Pro's

- Minimize external interconnections
- Reuse architecture from UniBoard
- Full speed FPGA-to-FPGA interconnections on the board.

Con's

- Scaling UniBoard^2's to large systems limited
- Fixed routing
- Fixed ratio between font and back nodes.

2.5 Mezzanine

The second concept is the mezzanine concept. In this concept the fixed mesh is replaced by a mezzanine board. Mezzanines can have different mesh structures. By replacing the mesh with active components like a switch, flexible routing or data transport can be offloaded from the main board. A parallel optics module placed on a mezzanine board would allow short range optical interconnections.

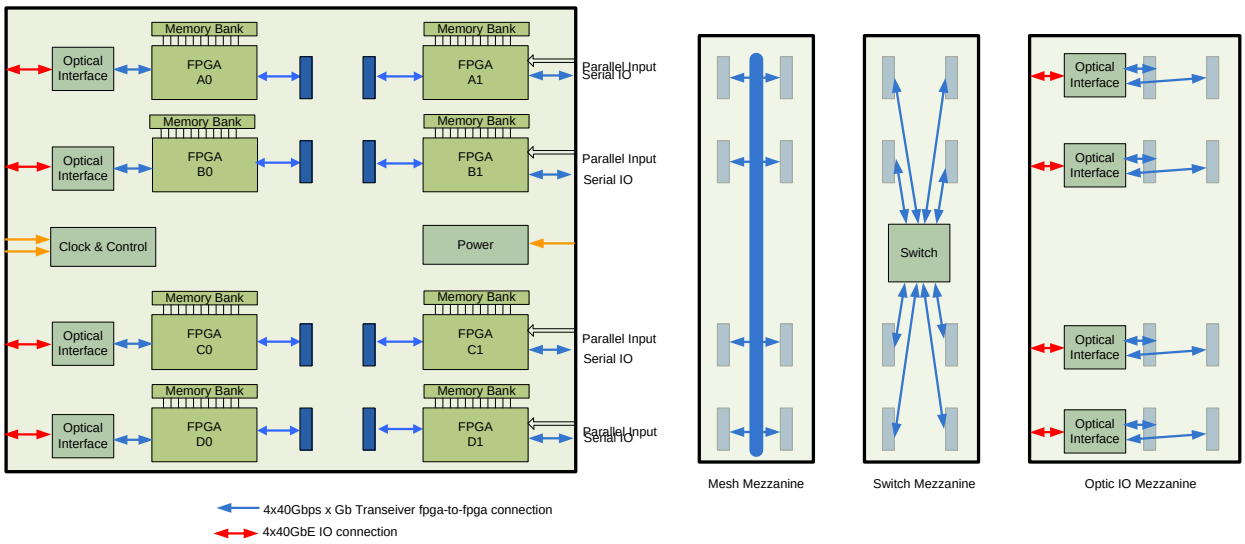


Figure 3 Concept 2 Mezzanine

Pro's

- Flexible

Con's

- Lot of connectors, and crossings introducing risk and SI issues
- Height of the combination limits rack density
- Cooling more complicated

2.6 Single Column

The third concept is the single column concept. With the mezzanine connectors replaced by backplane connectors, a single column of FPGAs remains. In this concept all FPGAs have the same interconnection scheme.

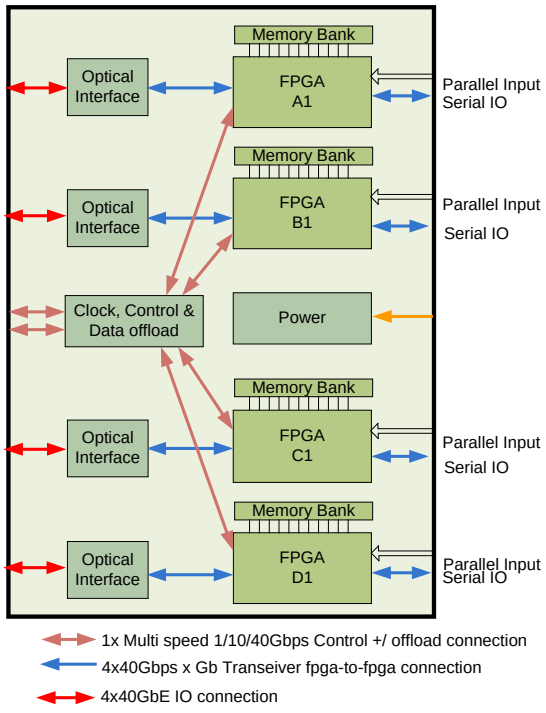


Figure 4 Concept 3 Single Column

Pro's

- Flexible

- Smaller boards are less expensive
- All FPGAs have the same IO interfaces.

Con's

- Connectors and crossings introducing risk and SI issues

2.7 Overview

On overview of the architectures are shown in Table 1.

Table 1 Overview architecture selection

	UniBoard like	Mezzanine	Single Column
Density	8 FPGAs	8 FPGAs	4-5 FPGAs
FPGA-FPGA speed	Optimal no connector crossings	Connector crossing limiting maximal throughput	Connector crossing limiting maximal throughput
Cost	Lowest cost per MAC	Extra connectors needed	For 8 FPGAs extra control logic and PCB's needed.
Flexibility	Ratio first stage / second stage processing fixed 1:1	Ratio first stage / second stage processing programmable with mezzanine.	Ratio in multiple of fours.
# Hops from one node to any other	2, within subrack	2, with mesh of 1 with parallel optics	1
Parallel ADC BW	4 FPGA (50% of the FPGAs)	4 FPGA (50% of the FPGAs)	4 FPGA (100% of the FPGAs)

3 Technologies

3.1 FPGA

3.1.1 Configuration

The requirements for the configuration are:

- Remote configuration (downloading new images in the flash) must be possible.
- Configuration within 5 sec

3.1.2 Altera Options

Some of the options are summarized in Table 2.

Table 2 Altera FPGAs

Device	Multipliers	KLE	Mbits	CPU	Max. Trans. Speed
GX660	3356	660	42	No	17.4 Gbps
GX900	3036	900	47	No	17.4 Gbps
GX1150	3036	1150	53	No	17.4 Gbps
GT900	3036	900	47	On	28 Gbps
GS660	3036	660	42	yes	17.4 Gbps

Table 3 IO per package

	F40 (1517 pins)		F45 (1932 pins)	
	Transceivers	IO	Transceivers	IO
GX900	48	624	96	480
GT900	48	624	96	480
GX1150	48	624	96	480
SX660	48	588	-	

More information can be found in [3].

3.1.3 FPGA pinning

To get some feeling of the pinning needed for this setup in the tables Table 4 and Table 5 are used to explore the pinning for the shown setup.

Table 4 Preliminary standard IO pinning requirements

Function	Pins per block	Blocks	Total pins
DDR4	140 (SE)	4	560
LVDS	9 (DIFF) 18 (SE)	4	72
Test IO	24	1	24
Monitoring & Control	4	1	4
Total			660

As can be seen in Table 3, the required pins are more than the available pin's. Some choices have to be made!

Table 5 Preliminary Serial IO pinning requirements

Function	Pairs per block	Blocks	Total pairs
QSFP+	4	4	16
Backplane	4	8	32
Power Mx	12	1	12
Total			60

3.2 In- and output interfaces

For the in- and output interface standard 40/100GbE interfaces will be used.

3.2.1 Front panel interface

On the front side an optical interface will be placed. In Figure 5 an overview of the interface is shown. For UniBoard^2 no media converter is needed. The 10/25Gbps interface can interface directly to the optical module.

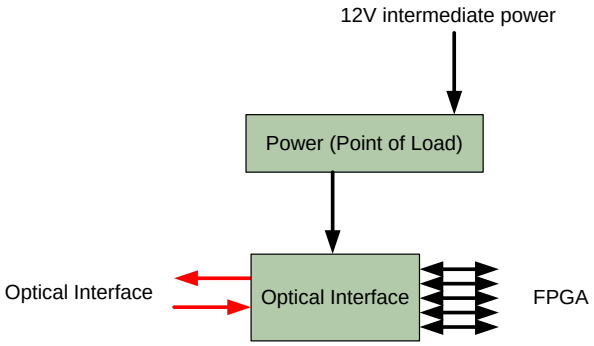


Figure 5 Block diagram optical interface

The length of the connection outside the board is not known on forehand, as this can be anywhere from <10m, in order to connect multiple cabinets, up to 10km, for connecting nearby telescopes. Different modules can be plugged into the cage on the board to accommodate both short and long range applications.

An overview of industry standard IO solutions is shown in Table 6.

Table 6 Standard IO solutions

Technology	Pro	Con	Lanes	Board area
QSFP+	- Pluggable module - Industry ready (Finisar / Avago module available for 40GBASE-SR4) - MSA (IEE802.3ba)		CAUI-4 (4x up-to 10.5Gb/s→ 40Gbps)	Stacked 25x73mm
CXP	- Pluggable module - Industry ready (Finisar module available for 100GBASE-SR10) - IEE802.ba	- Short range	CAUI (12x12.5Gb/s → 100Gbps)	stacked 27x69mm
CFP	- Pluggable module - Industry ready (Finisar)	Large footprint	In: CAUI Out:4x25	
SFP+	- Pluggable module	Single channel takes more board space	1x 25Gbps	

The Technologies are described in more detail in the following subsections.

3.2.1.1 QSFP+

UniBoard^2 will have a cage for QSFP+ pluggable modules. In Figure 6 an example of a TE cage is shown [5]. Multiple suppliers like Molex, TE, Samtec make these cages. Instead of QSFP+ a zQSFP+ cage could be placed, optimised for higher data rates like 100Gbps (4x28Gbps).

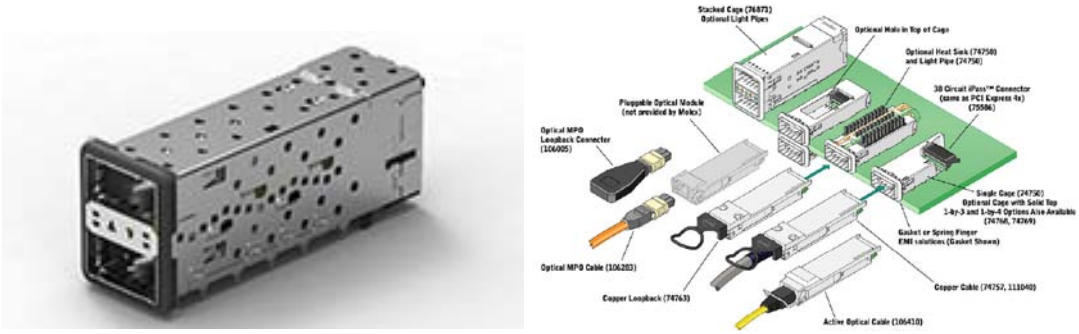


Figure 6 TE QSFP+ cage

An example of optical modules that could be used is the Finisar FTL410QD2C 40GBASE-SR4 [6]. Besides optical modules direct attach copper cables can be plugged in as well. An example is the Cisco QSFP-H40G-CU50CM, shown in Figure 7.

Figure 7 QSFP pluggable 10GBASE-SR4 module (left) and passive direct attach copper (right)

The power consumption of the modules is of the order of 1.5W. With eight modules on the board a total of 12W will be consumed

3.2.1.2 CXP

CXP is a 12 fibre standard in which each fibre can handle 10Gbps data rates, resulting in 120Gbps per interface. This standard is used for short distances. Cages are available through Molex and TE [7], module by Finisar.

Figure 8 CXP cage and optical connectors

3.2.2 Backplane interface

For the backplane side a copper interface will be made. Through this interface multiple UniBoards can be connected to a backplane. The backplane interface will be used for:

- ADC input
- Serial connectivity
- Power input
- Control signals

3.2.2.1 Serial connectivity

As described by Altera [3] speeds of up to 14Gbps can be achieved on a backplane. With 24 transceivers the maximal throughput per FPGA will be of the order of 300Gbps.
The backplane connector used for UniBoard1 is capable up to 12 Gbps. For UniBoard^2 another type of connectors might be more advantageous. In Figure 9 an example of an orthogonal connector from Molex [8] capable of data rates up to 25Gbps is shown. With liquid cooling boards do not need to be mounted vertically anymore, which makes different board arrangements possible.

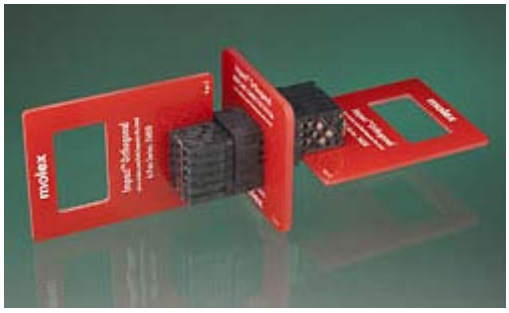
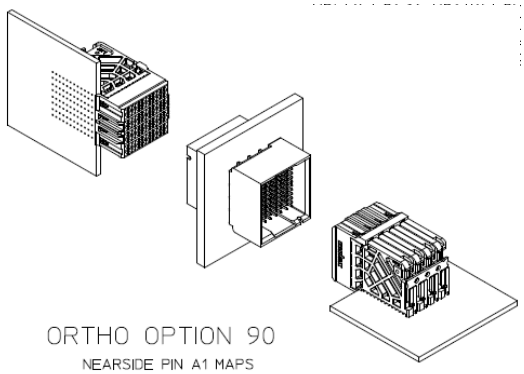


Figure 9 Molex Orthogonal connector

Orthogonal plugging
More details about the orthogonal pinning are shown in Figure 10. From this detail it can be seen that when orthogonal board are to be used a small backplane pass connector could be used. The pinning is point symmetric. This should constrain the pinning of the connector (B1-C1 for transmit T11-S11 for receive).



ORTHO PIN MAPPING

OPTION 90 (NEARSIDE – FARSIDE)

A1–T12	A2–S12	A3–Q12	A4–P12	A5–M12	A6–L12	A7–J12	A8–H12	A9–F12	A10–E12	A11–C12	A12–B12
B1–T11	B2–R12	B3–Q11	B4–N12	B5–M11	B6–K12	B7–J11	B8–G12	B9–F11	B10–D12	B11–C11	B12–A12
C1–S11	C2–R11	C3–P11	C4–N11	C5–L11	C6–K11	C7–H11	C8–G11	C9–E11	C10–D11	C11–B11	C12–A11
D1–T10	D2–S10	D3–Q10	D4–P10	D5–M10	D6–L10	D7–J10	D8–H10	D9–F10	D10–E10	D11–C10	D12–B10
E1–T9	E2–R10	E3–Q9	E4–N10	E5–M9	E6–K10	E7–J9	E8–G10	E9–F9	E10–D10	E11–C9	E12–A10
F1–S9	F2–R9	F3–P9	F4–N9	F5–L9	F6–K9	F7–H9	F8–G9	F9–E9	F10–D9	F11–B9	F12–A9
G1–T8	G2–S8	G3–Q8	G4–P8	G5–M8	G6–L8	G7–J8	G8–H8	G9–F8	G10–E8	G11–C8	G12–B8
H1–T7	H2–R8	H3–Q7	H4–N8	H5–M7	H6–K8	H7–J7	H8–G8	H9–F7	H10–D8	H11–C7	H12–A8
J1–S7	J2–R7	J3–P7	J4–N7	J5–L7	J6–K7	J7–H7	J8–G7	J9–E7	J10–D7	J11–B7	J12–A7
K1–T6	K2–S6	K3–Q6	K4–P6	K5–M6	K6–L6	K7–J6	K8–H6	K9–F6	K10–E6	K11–C6	K12–B6
L1–T5	L2–R6	L3–Q5	L4–N6	L5–M5	L6–K6	L7–J5	L8–G6	L9–F5	L10–D6	L11–C5	L12–A6
M1–S5	M2–R5	M3–P5	M4–N5	M5–L5	M6–K5	M7–H5	M8–G5	M9–E5	M10–D5	M11–B5	M12–A5
N1–T4	N2–S4	N3–Q4	N4–P4	N5–M4	N6–L4	N7–J4	N8–H4	N9–F4	N10–E4	N11–C4	N12–B4
P1–T3	P2–R4	P3–Q3	P4–N4	P5–M3	P6–K4	P7–J3	P8–G4	P9–F3	P10–D4	P11–C3	P12–A4
Q1–S3	Q2–R3	Q3–P3	Q4–N3	Q5–L3	Q6–K3	Q7–H3	Q8–G3	Q9–E3	Q10–D3	Q11–B3	Q12–A3
R1–T2	R2–S2	R3–Q2	R4–P2	R5–M2	R6–L2	R7–J2	R8–H2	R9–F2	R10–E2	R11–C2	R12–B2
S1–T1	S2–R2	S3–Q1	S4–N2	S5–M1	S6–K2	S7–J1	S8–G2	S9–F1	S10–D2	S11–C1	S12–A2
T1–S1	T2–R1	T3–P1	T4–N1	T5–L1	T6–K1	T7–H1	T8–G1	T9–E1	T10–D1	T11–B1	T12–A1

NOTE: PINOUTS SHOWN IN BALLOONS ARE GROUNDS.
PINOUTS SHOWN ARE FOR SQUARE (12 COLUMN) PATTERN ONLY.

Figure 10 orthogonal pinning

In section 5.1 an example of using orthogonal connectors is shown.

Coplanar plugging
Instead of orthogonal plugging back-to-back or coplanar plugging can be used as well, see Figure 11.

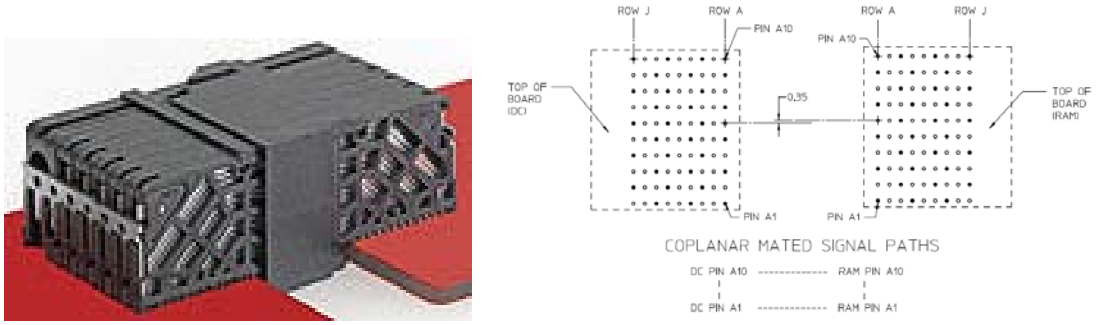


Figure 11 Coplanar plugging

Due to the ground pin's it is however not possible to use the same board for the normal daughter card side and the RAM assembly. This connection type can be used between UniBoard and breakout board (not between two UniBoards).

3.2.2.2 Standard backplane connectors

The standard backplane connectors as used by UniBoard are shown in Figure 12. With proper layout this connector is capable of 10Gbps. For a modified version (with the same pinning) the speed in increased to 20Gbps (including back drilling) [9].

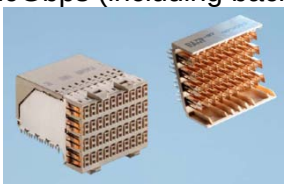


Figure 12 ERmet ZD backplane connector

3.2.2.3 ADC interface

Part of the backplane interface is an ADC interface. For UniBoard1 an LVDS type interface is implemented, with four times 8 bits per FPGA. More pins may be used for the ADC interfaces, depending on the connector pinning.
More and more ADCs nowadays are equipped with the JESD2004 serial interface. These types of ADCs can be connected to the UniBoard^2 through the serial interfaces.

3.2.3 Short range optical interface

If not all the transceivers are used for the default connections, a short range optical interface could be placed next to each FPGA. As an example the miniPOD [10] and MicroPOD are shown in Figure 13. On these connectors 12 receive or 12 transmit serial interfaces can be made at a maximum of 10Gbps each, resulting in a 120Gbps aggregate bandwidth. By placing a transmitter and receiver near each FPGA other interconnections to and from the FPGAs can be made, creating for example a ring architecture on the board, or giving the possibility to connect an optical backplane or extra optical IO. An evaluation of these modules is described in [11].

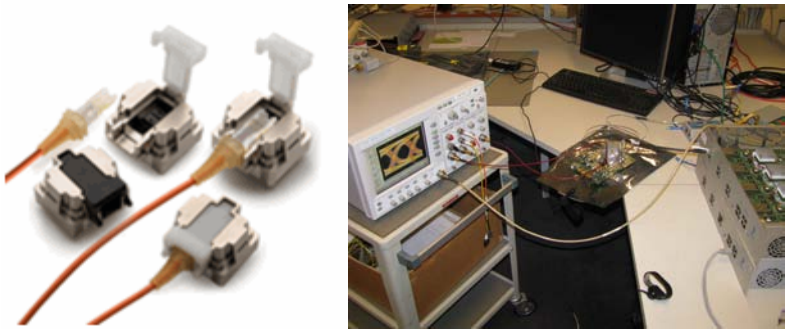


Figure 13 Avago Micro Pod (left) with UniBoard (right)

3.2.4 Power^{MX}

Instead of a dedicated IO connector a PowerMX IO Cluster can be placed. The smallest Atomic is a Qtr I/O Cluster which exists of two 100 pin's MegArray connectors, see Figure 14. On one of these connectors 12 serial links can be placed, on the other 4 serial links and control signals. More information about the PowerMX can be found on the CyberSKA website [12].

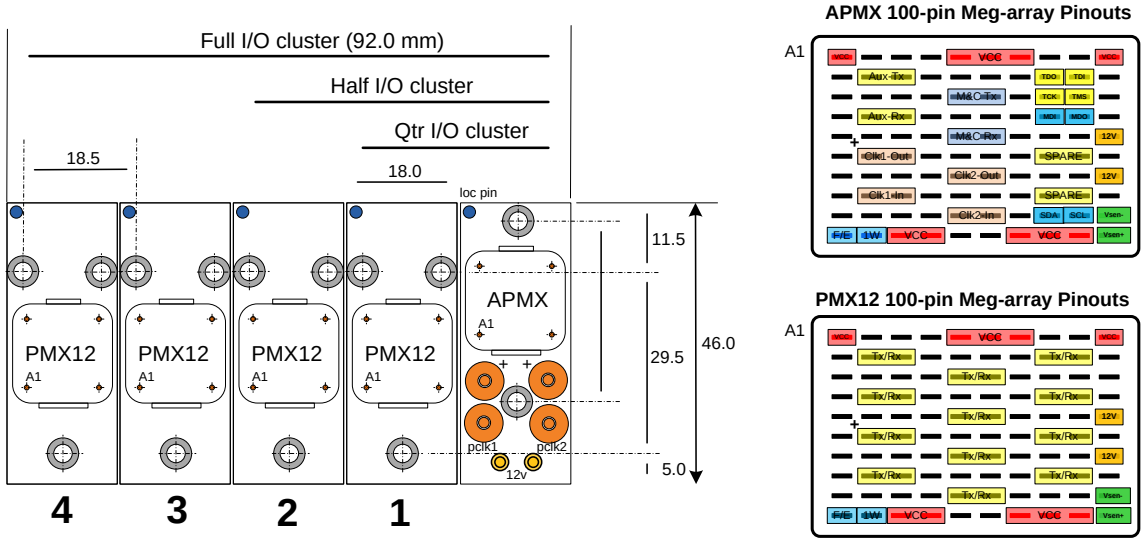


Figure 14 Power^{MX} IO cluster

3.3 Memory

3.3.1 Persistence Memory

Persistence memory can be used to store information like

- IP/MAC addresses
- Board parameters (serial number)

3.3.2 Image Memory

Serial Flash can be used for the FPGA image.

3.3.3 Application Memory

The memory requirements for UniBoard^2 are:

- 1 second of data storage
- FIFO (write and read within sample rate)

Solutions for the memory are:

- DDR3
- DDR4 (Standard DIMMs)
- GDDR5
- QDRII / QDRII+
- Hybrid Memory Cube

Some major technologies are described in the following subsections:

3.3.3.1 DDR3 (DIMM / SODIMM)

Double Data Rate memory is a SDRAM based memory. DDR3 chips can be placed on JDEC standard modules. The same 64 or 72 bits wide bus is used for writing and reading. The transfer rate of a
SODIMM memory module is up to 1866MT/s (Micron's maximal speed grade) which for a 64 bit wide module results in 14.9GByte/s (119Gbps). The power consumption of the module can be reduced by lowering the supply voltage to 1.35V. With a bus width of 72 bit and a maximal speed of 1866MT/s, DDR3 can reach a speed of 17GByte/s (without overhead). UniBoard^2 could hold at least two modules per FPGA, resulting in a maximum speed of 238.8Gbps.

3.3.3.2 DDR4 (DIMM / SODIMM)

DDR4 is the next DDR generation. Compared with DDR3 power consumption is reduced and speed increased. With a theoretical maximal transfer rate of 3200MT/s the speed of DDR4 is 33% higher than that of DDR3. Reducing the voltage from 1.5V to 1.2V further reduces power consumption. In the future it should be possible to drop the voltage to 1.05V. First engineering samples have been produced, however Micron does not have any SODIMM DDR4 modules available yet, although production should have started Q4 2012.... At an speed of 3200MT/s and a bus width of 64 bit, DDR4 can reach transfer speeds of up to 25GByte/s (without overhead).
Pinning of the DDR4 is shown in Table 7.

Table 7 Preliminary DDR4 SODIMM pinning

Function	Pins
Data	72
Address	3 (GA) + 3 (BA) + 18 (A)
Data Mask	9
DQS	9 (differential
Control	5 (RAS+CAS+WE+ODT+ACT)
CKE	4
Clock	4 (differential)
Total	136 IO lines needed

For UniBoard2 two or four DDR4 modules can be placed. When four modules are placed per FPGA a memory bandwidth of 688 Gbps Half duplex can be achieved. In this case a device with more IO is needed, compromising the number of transceivers to the front panel and the backplane. When two modules are placed per FPGA the memory bandwidth to the DDR is approximately 344Gbps (90% efficiency, 72 bits). This is not sufficient for the input bandwidth.

3.3.3.3 GDDR5

Graphics Double Data Rate memory is mainly used on GPU cards. This memory uses different clocks for commands and data. The input is split into two 32bit wide busses with each a separate clock for writing. The top speed is 7Gbps, for normal applications the speed is 3.2Gbps, resulting in 25.6GByte/s (without overhead). However GDDR5 is not supported by Altera or Xilinx.

3.3.3.4 QDRII+

Quad Data Rate SRAMs are dual port Double Data Rate SRAM memories. With a data bus width of 36 bit and a speed of 1100Mbps (550MHz) the double bus can reach data speeds of up to 10GByte/s (without overhead). The maximal size of a QDRII+ device is of the order of 18MByte

3.3.3.5 Hybrid Memory Cube

Hybrid Memory Cube is a new type of memory which became possible by the use of TSV (Through Sillicon Vias). In this technology memory chips are stacked onto a serial interface chip. The technology is under development, close to producing engineering samples. First pictures are available of two devices with four or two serial links each link build round 16 lanes reaching 160 / 120Gbps. With the use of serial interfacing to the memory the memories can be placed for near memory close to the FPGA or far memory on another module. With four lanes and a total data rate of 160Gbps HMC can reach speeds of up to 20 GByte/s (without overhead, counting send and receive and the same time).
To make deeper memories, memory bandwidth can be exchanged with size. By linking HMC devices as shown in Figure 15 for a 4-lane device 4 HMC's can be chained. With an eight lane device 8 HMC's can be chained.
More information about HMC can be found in [18] [20] and [20].

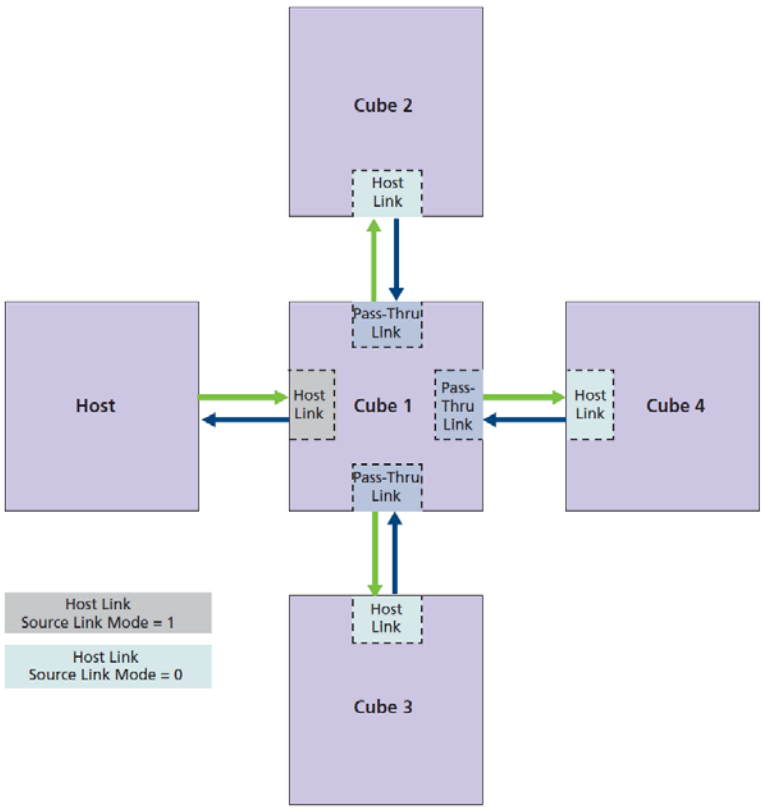


Figure 15 Chaining multiple MHC devices

3.3.3.6 Conclusion

An overview of the types of memory is given in Table 8.

Table 8 Overview of application memory solutions

Techn.	Pro	Con	Speed
DDR3	SO-DIMM Module (limited board space)	Latency	2133Mbps*8=17Gbyte/s/mod.
DDR4	DIMM module Low power Standard	Modules not in production yet	2666Mbps*8=21Gbyte/s/mod.
GDDR 5	Fast	BGA-chips Capacity up-to 256MByte IO not supported by Altera	
QDR II+		BGA-chips Price memory size	1400Mbps
QDRIV			2666Mbps
HMC	Fast	Engineering state Limited information	20Gbytes/s/chip

QDR is small in memory size and uses a lot of board space and IO. For DDR more overhead is needed compared to QDR but the total BW is the same.
GDDR is not supported by the FPGA vendors making memory interface firmware complex. Therefore on UniBoard^2 two DDR4 memories will be placed next to every FPGA. This will enable a small memory interface while maintaining maximal IO bandwidth with 96 transceivers which are needed to achieve the best processing to IO-bandwidth ratio. By making an HMC break-out board the memory bandwidth will be increased.

3.3.4 Background information:

http://www.altera.com.cn/literature/wp/wp_memoryselect.pdf
http://www.micron.com/products/dram/ddr3-to-ddr4
http://www.micron.com/about/blogs/2012/december/ddr4-gathers-momentum
http://www.cypress.com/?id=107&addcols=&parametric=html#parametric
http://www.micron.com/products/hybrid-memory-cube

3.4 PCB

With the high-speed traces (up to 25Gbps) the PCB is part of the design. Losses of serial links increase with increasing speeds as skin depth, the area of copper used by the high-speed signal, becomes smaller. This can be compensated by using wider traces, which implies that the distance between the layers has to be increased (fewer layers in a 2.5mm board), or by decreasing the dielectric constant of the PCB material. Another source of loss is dielectric loss. New materials focus on both parameters.

3.4.1 Material

One examples of PCB material is:

- Panasonic Megtron-6 Df 0.002 used for >25Gbps

Before the layout of UniBoard^2 will start, a discussion with the PCB supplier will take place. In this discussion the material selection and the technologies used (microvia's) will be covered.

3.4.2 Trace impedance

By using 85 Ω instead of 100 Ω the traces and the layers can be moved closer, allowing more layers and more traces.

3.5 Mezzanine connector

For concept 2 a mezzanine connector is needed which can handle the serial interconnections of the FPGA. This means that for a full duplex connection of four 40Gbps interfaces, each built around four 10Gbps interfaces, at least 32 pairs are needed. In Figure 16 the Amphenol InfinX^{TM} High Speed Mezzanine connector [13] is shown. This connector has pairs capable of 25 Gbps. The maximum number of pairs per connector is 108 (for the largest connector).

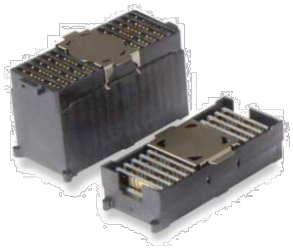


Figure 16 Mezzanine connector

3.6 Clock and control

3.6.1 Clock and PPS

Tuning of the clock and PPS traces will be used for all nodes with ADC interfaces. This should ease firmware design for synchronic ADC sampling.

3.6.2 Control

To reduce the number of interconnections to the board a 1 GbE switch will be placed on the board, which will distribute one 1GbE to all FPGAs. For the switch the Vitesse VSC7389 [14] will be used. This switch has 8 SGMII interfaces (each interface consist of two pairs, one for transmit and one for receive) for connecting the FPGAs, and 8 integrated tri-speed copper

transceivers of which four will be connected to the front side of the board. The VSC7389 has an EEPROM for storing the controlling software of the switch.
Through I2C interfaces the temperature of the FPGA can be monitored. The last FPGA on the board (lower backside FPGA) will use the I2C interface for monitoring not only the FPGA temperatures but also the board power levels. All front FPGAs have dedicated status information about the optical interface. This information includes whether fibres are connected and any errors during communication. In Figure 17 an overview of the control is given.

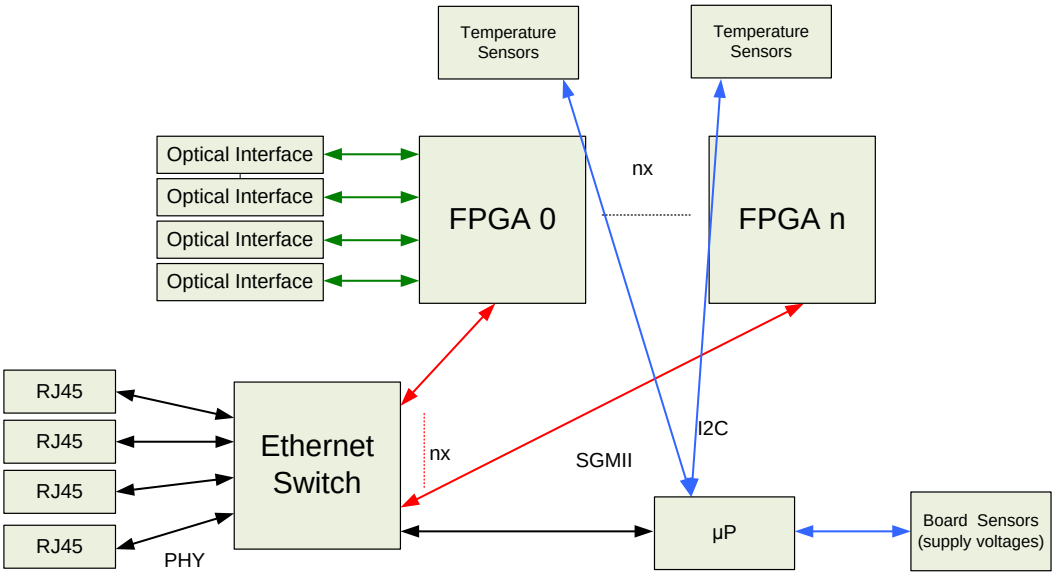


Figure 17 UniBoard control

3.7 CPU

UniBoard1 has a CPU embedded in the FPGA. On UniBoard^2 an external CPU could be placed for board control. The tasks of the CPU are:

- Remote configuration
- Reset
- Register access

Nice to have

- TCP/IP connection to UniBoard
- Post processing data
- High level control instead of register control
- Processing settings

Interfaces from the CPU to the FPGAs and outside

- Ethernet connection
- I2C for temperature monitoring
- JTAG programming FPGA's / Flash

3.7.1 Possible solutions

The CPU can be located on the UniBoard^2, on a module or in the form of an external PC (Local Control Unit, LCU). The options are discussed in the following subsections. A table summarizes the pros and cons.

3.7.1.1 FPGA solutions

The CPU can be located inside the FPGA. This could work together with an external LCU when needed.

Hardware implementation
For the basic functionality a processor might not be needed. Using dedicated hardware blocks and efficient fast access, the memory map/control bus and the configuration can be created. On the LCU all processing can be done for all nodes. When more processing would be needed the LCU could be upgraded.

Embedded Softcore Processor in FPGA
The UniBoard uses an embedded CPU (Nios). The logic needed for the UniBoard NIOS is of the order of 1-2% of the total.

Embedded ARM processor
The current FPGA families have parts with and without hardcore ARM processor. Within the Arria family the SoC devices are footprint compatible.

3.7.1.2 Processor on UniBoard

Small processor (uP)
On UniBoard^2 a small processor could be placed for basic control functionality. This type of processor has a 100MBit interface. Examples are a PIC processor, Arduino, Raspberry Pi. These processors have limited functionality but have a user-friendly interface. The processor can also be used for general board control like reset, setting and monitoring the power supplies. A PIC processor is used in the Apertif system to control a subrack with four UniBoards.

Full scale processor on UniBoard
When the functionality must be increased a bigger CPU could be placed on UniBoard^2 with a proper operating system. This CPU has dedicated memory. A possibility is the Freescale P4080 (used in the Dome project).
Pro's

- Don't need processor in the FPGA (a Nios-processor might still be needed to setup the transceivers, DDR and small control loops)
- TCP/IP connection to UniBoard instead of UDP (although UDP is now commonly used in the financial networking due to its limited overhead (reduction in transport time))
- Linux on UniBoard
- FPGA configuration

Con's

- Board space for CPU and DDR memory comparable with FPGA (in other words a CPU instead of an FPGA)
- Development time (as experienced in the DOME project, a second spin might be needed to get the CPU running), CPUs have a variety of busses and more clock sources compared to an FPGA. (e.g. CASPER)
- The life time of a CPU is shorter than that of an FPGA or even the board life time of the UniBoard (the time between the design and the last production batch of UniBoard^2 is expected to be > 5 year). This means that the CPU will be obsolete by the time the board is in production.
- Extra tool flow needed for Processor software (e.g. CASPER).
- Solder connections and components needed for the CPU introduce more risk in the production of the board.

3.7.1.3 Processor on Module

Instead of placing a CPU on UniBoard^2 and increasing the complexity of routing and fabrication, a CPU can be placed on a module. In this ways the complexity of UniBoard can be limited while enabling the flexibility of the CPU for control.

COM (CPU on module)
A standard form factor can be used. Standard CPU modules are :

- COM Express
  Module size ranging from 84x55mm till 155x110mm (Basic module 95x125mm). Connector has 220 pins which has, USB ports, 6 PCIe lanes, 24 bit LVDS channels Gigabit Ethernet and 8 GPIO.
- ESMexpress
  Module size of 85x115m. Connectors with LDVS, SATA, USB, 3x GbE PCIe PEG
- Qseven
  Modules with a size of 70x70mm. Board edge connector → limited hight of components underneath module.

Background
http://emea.kontron.com/com-express-r-from-kontron/

CPU on XGB
Instead of placing the CPU on UniBoard^2 a Backplane board can be made holding the CPU.

3.7.1.4 Connections to the FPGA

The connection between the CPU and the FPGA can be through

- Ethernet:
  Ethernet is usually used for mid to long distances. This serial interface can handle speeds of up to 10Gbps. For control interfaces speeds of up to 1Gbps can be used. The FPGA will need logic to handle the traffic. In the newer FPGAs hard IP in combination with firmware can be used to implement the control interface.
- PCIe:
  PCI express is a low level (level 1 and 2 of the OSI model) communication protocol which is used for chip-to-chip communication in a memory mapped local bus structure. The bus exists of a master, usually the CPU, and a slave. Between a master and multiple slaves a bridge needs to be established. Some processors however (like the Freescale P5020) have multiple busses (masters). Hard IP and firmware code examples are available from Altera for the configuration of their FPGAs. PCIe is used with OpenCL for the communication between the CPU and the FPGA.

3.7.1.5 Conclusion CPU

A summary of the pro's and con's of the different CPU options are shown in Table 9.

Table 9 CPU pro's and con's

	Low level Firmware implementation	Soft core (Nios)	CPU	Emb. CPU (SoC)	CPU on module
Hardware design	Simple PCB design	Simple PCB Design	Extra time and space needed on the PCB	Simple PCB Design	Connector with routing needed
Hardware production risk			Increase in components and therefore higher risk.	Increase in components and therefore higher risk.	Limited increase in components.
SW/FW implemen-tation	More firmware development time needed.	Reuse of UniBoard's UNBOS	Firmware design shifted to software. Software development time needed to get CPU running	Firmware design shifted to software.	Firmware design shifted to software.
Boot time	Simple booting.	Simple booting	First CPU Linux boot, than FPGA. This will take time.	First CPU Linux boot, than FPGA. This will take time.	First CPU Linux boot, than FPGA. This will take time.
Interface speed	Implementation made for speed.	Soft core limits the control speed.	CPU handles TCP/IP. Low level control from CPU to FPGA.	CPU handles TCP/IP. Low level control from CPU to FPGA.	CPU handles TCP/IP. Low level control from CPU to FPGA.
TCP/IP	Development time needed to implement.	Development time needed to implement.	Supported by kernel.	Supported by kernel.	Supported by kernel.
Memory bus	Part of implementa-tion	Part of implementa-tion	Part of implementa-tion	Part of implementa-tion	Part of implementa-tion
Post processin g	Not possible	Not possible	On CPU.	On CPU.	On CPU.
Extra tool needed for develop-ment	Not needed	Not needed	Compiler needed. Extra tools might be needed to use all function of the controller.	FPGA tools needed for the CPU.	Compiler incl. Module tools needed.

For all situations a control interface has to be implemented on the FPGA. This interface must accept PCIe (for an external large CPU on UniBoard^2) or standard Ethernet. Only a small number of IO-lines can be made from the CPU to the FPGA (not a complete register map). To place a CPU on UniBoard^2 for only a 'hello world' might be a bit of overkill. Therefore a

firmware implementation will be taken as the starting point for UniBoard2. In this case a hardware implementation, soft core or an embedded CPU can be used. By using PCIe lanes to the break out board experiments with an external CPU can be done as well. A PIC processor with an Ethernet interface will be placed on UniBoard^2 for reset and power settings.

3.8 Ethernet Switch

To increase the flexibility of UniBoard^2 an onboard switch will be needed. In this section some examples are discussed in more detail.

3.8.1 10Gbps Gearbox

The Avago Technologies Vortex Gearbox^{TM} AVSP-1104 is a 10:4 gearbox [15]. This means that ten 10Gbps lanes can be multiplexed (full duplex) to four 25Gbps channels. The chip can also be used as an repeater to increase the maximum distance between two points.

3.8.2 10Gbps Switch

The Vitesse Semiconductor VSC3144-12 [16] is a 144x144 Cross point Switch with data rates of up-to 10.7Gbps. Smaller variants (16x16) crossbars are available as well. A proprietary 5 wire bus is used to set the switch.

3.8.3 40Gbps Switch

An example of a 40Gbps switch is the Broadcom BCM56540 [17], 16x10G (4x40G) + 48x3G. This switch need a external processor to set the switch.

3.8.4 Conclusion

Like the CPU the switch can be placed on an extension board or on a dedicated switching board connected to the backplane.

3.9 Test strategy

3.9.1 Boundary scan

The board will be equipped with a boundary scan interface. Via this interface access will be provided to all boundary-scanable devices. This will enable testing the board at low speeds

3.9.2 Monitoring

All FPGA temperatures will be monitored. Where possible the power of a supply will be monitored as well.

3.9.3 Operational status information

UniBoard^2 will be capable of monitoring:

- FPGA temperature
- Supply voltages
- Power supply status

4 System examples

4.1 Making a beamformer with UniBoard^2

How could a beamformer be made using UniBoard^2? In the figures Figure 18, Figure 19 and Figure 20 some examples are shown.

In the ring beamformer of Figure 18 the partial sums are made close to the input in each node. The output is transferred to the next FPGA on the board or the next UniBoard. With this structure nearly unlimited number of inputs can be processed. The bandwidth between nodes determines the output bandwidth.

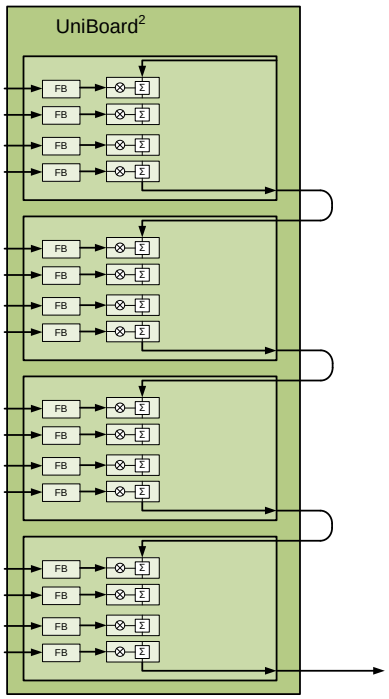


Figure 18 Ring beamformer

The FFT beamformer shown in Figure 19 can be used to make the same number of beams as there are inputs. Increasing the number of inputs does not only increase the number of input boards (vertically) but the processing boards (horizontally) as well.

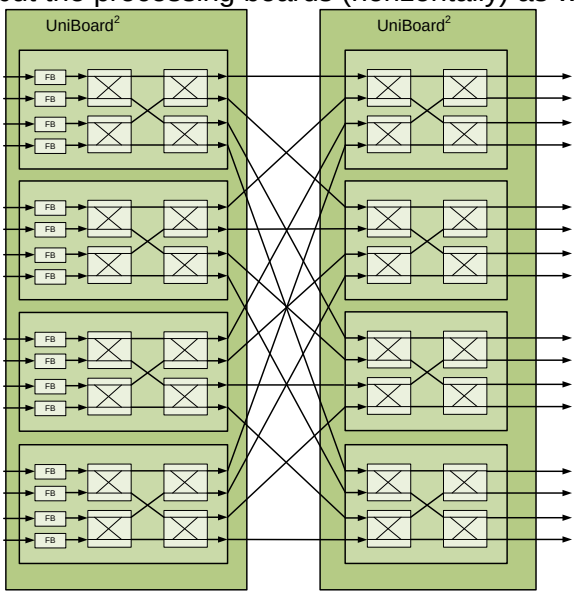


Figure 19 FFT beamformer

When not all beams need to be formed, but still more than can be formed in the ring structure, a subband beamformer can be used, see Figure 20. In this beamformer subbands from the input nodes are transferred to beamforming nodes. The number of input nodes does

not have to be equal to the number of beamformer nodes. If sufficient resources are available the beamformer can be implemented on the input node FPGA.

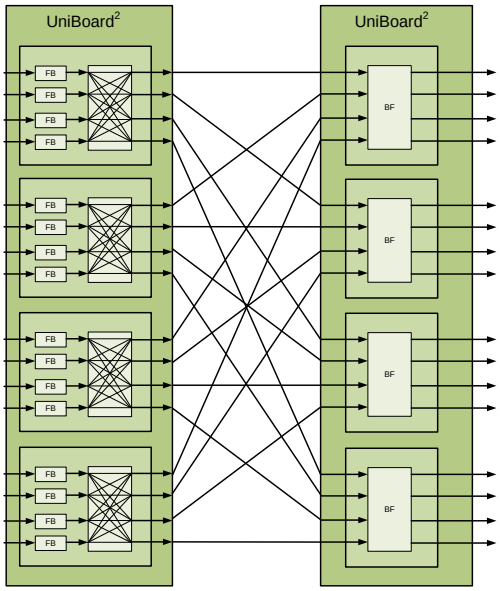


Figure 20 Subband beamformer

In practice the beamformer could look like shown in Figure 21, with on one side of the backplane the sensitive analog channels and on the other side the input and processing nodes.

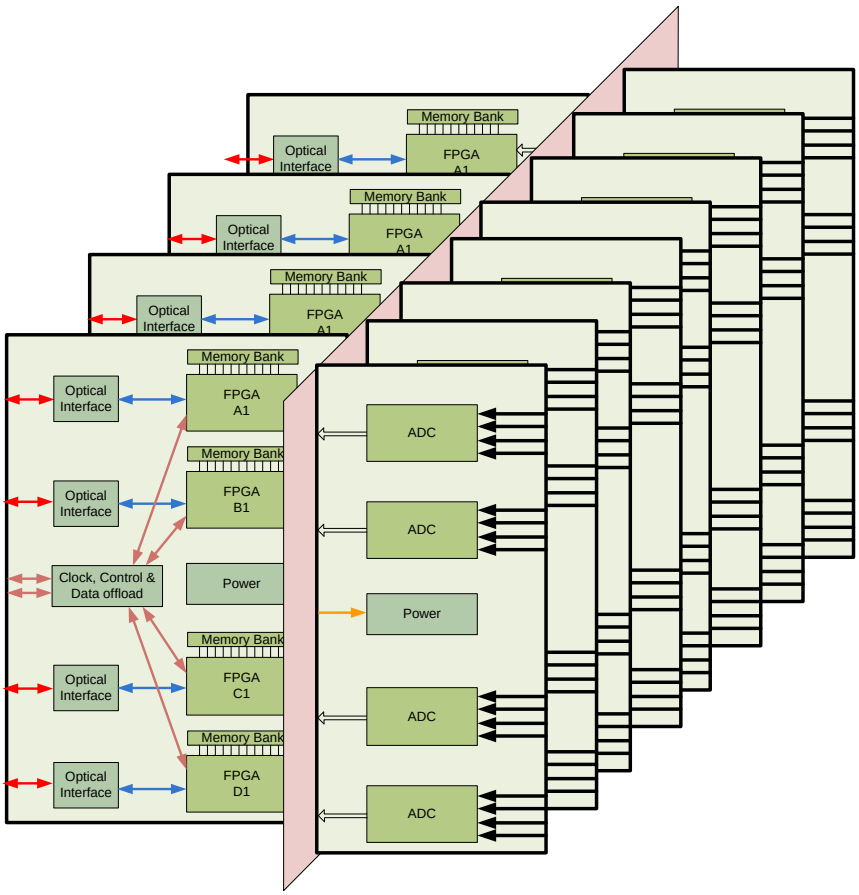


Figure 21 Beamformer system build up

4.1.1 Apertif beamformer

What could the Apertif Beamformer be constructed with UniBoard^2?
The analog bandwidth of the 64 inputs for the Apertif beamformer is 400MHz. From this input an average of 42 beams with a bandwidth of 300MHz (384 subbands) are made. This yields an input data rate of 410Gbips, after the filterbank 1152 Gbips on the board mesh, 1152Gbps on the backplane mesh. The output data rate is 160Gbps. The processing load is approximately 8.2TMAC.
If a beamformer were to be made using four UniBoard^{2}s each with 4 Arria10 FPGAs the processing capacity would be 20TMAC, more than double that which is needed. This would result in a 1GHz input bandwidth or 128 inputs at 500MHz bandwidth. This could be handled by the doubling of speed in the LVDS pairs and the more than doubled IO at the front node.

4.2 Making a Correlator with UniBoard^2

UniBoard^2 can be used to construct a correlator as well. In the figures Figure 22 and Figure 23 some examples are shown. For a ring correlator shown in Figure 22 correlation products are calculated near the signal input. All inputs are further transferred over the ring to the next node to calculate the next correlation product. This is comparable with the correlator structure as described in (CASPER, APERTIF AARTFAAC)

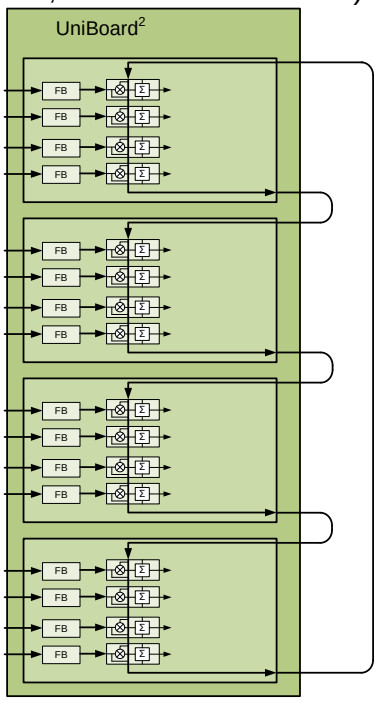


Figure 22 Ring correlator

In Figure 23 a subband correlator is shown. In the input nodes on one UniBoard the signals are aligned and passed through a filterbank. The subbands are distributed to correlator nodes where the same subbands from all signals are processed. The number of input nodes does not have to be the same as the number of correlation nodes. This is similar to the EVN UniBoard VLBI correlator.

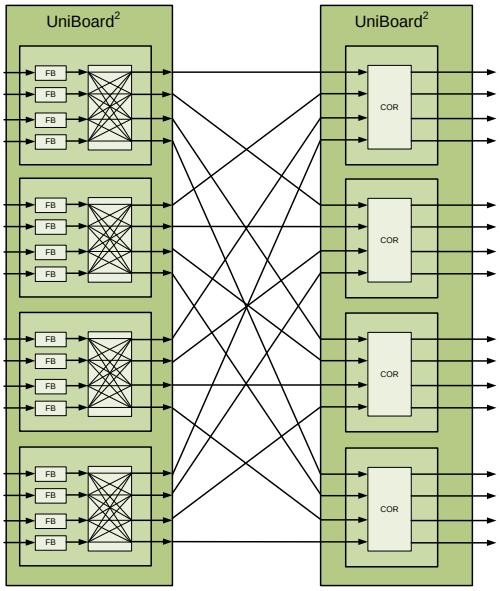


Figure 23 Subband Correlator

In practice a multi UniBoard correlator system could look like shown in Figure 24, with the input and filter banks on one side and the correlation nodes on the other side.

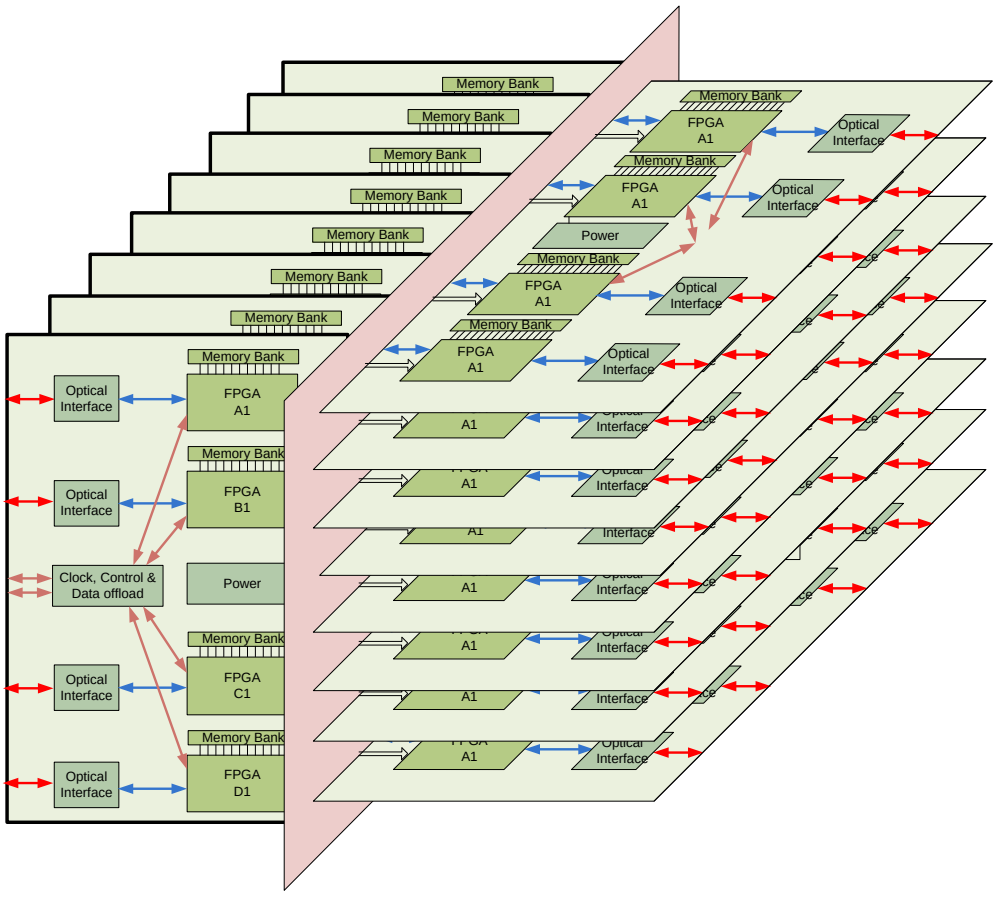


Figure 24 UniBoards in a correlator system

A system like this would have a capacity of 8x4 FPGAs for the filterbank, a maximum of 8x4 FPGAs for the correlator resulting in 29 TMAC/s for the filterbank and 29 TMAC/s for the correlator, assuming Arria10 FPGAs. To transfer data directly from a filterbank node to a correlator node 32 transceivers (8x4) would be needed.

4.2.1 Apertif Correlator with UniBoard^2

4.2.1.1 IO

The input of the correlator has 16*24=384 10Gbps links.
The input at the UniBoard^2 front side has 4 FPGA, each FPGA has 4 QSPFs which have 4 times 10Gbps lanes, for a total of 64. Using the front panel connections 6 UniBoard^{2}s are needed. The input at the UniBoard^2 backplane side has 4 FPGAs with 32 lanes at 10Gbps each. Using the backplane connections 3 UniBoard^{2}s will be needed. When both the QSFP on the front panel and the backplane interfaces are used only 2 boards are needed to make the correlator.

4.2.1.2 Processing

The current Apertif Correlator design has 16 UniBoards. On UniBoard^2 the processing capacity is doubled which results in 8 boards for processing

4.2.1.3 System Design

Each optical link from the beamformer contain 1/16 of the frequency band (24 subbands). To combine all subbands from all telescopes for both polarisations would require 24 10G links for a single correlator cell. If the backplane interface is used all data can be received on a single node. The 96 SFP+ cages per UniBoard^2 (using single mode fibres) can be divided over multiple input boards. In this case 16 modes are needed or 4 boards.

4.2.1.4 Power

Using UniBoard^2 will reduce the power needed for the correlator. Using half the number of board the estimated power reduction is 50%. The FPGA son UniBoard^2 have 4 times more capacity at 2 times the power, but UniBoard^2 has only half the number of FPGAs. If we assume 300W per UniBoard and 1W/year at 2€/W, switching over to UniBoard2 could save 1k€ per year of operation

4.2.2 AARTFAAC Correlator

4.2.2.1 IO

The input of the AARTFAAC correlator comes from 3 subracks with each 4 RSP processing boards with an infiniband interface existing of 4 lanes, resulting in 48 lanes at 10Gbps. UniBoard^2 has 32 lanes per node which means that only 1/3 of the interfaces are used.

4.2.2.2 Processing

For processing each station uses two UniBoards, these could be replaced by a single UniBoard^2.

4.2.2.3 System design

The lanes from the subracks are subband oriented. This would mean that each node on UniBoard^2 could process the data from a single lane, easing data communication.

4.2.2.4 Power

Using half the number of board the power reduction will be about 50%.

5 What will it look like?

In Figure 25 a proposal for the setup of UniBoard2 is shown. In this setup there are two DDR4 modules and six QSFP+ cages per FPGA. Each FPGA has up to 96 transceivers, 24 transceivers are used for the front panel interface and 72 connected to the backplane connector. On the backplane serial ADC can be used for data input, or breakout boards for application specific needs. On a default breakout board as shown in Figure 25 Hybrid Memory Cube (HMC) devices will be placed.
With this setup a platform will be made for the next version of devices where interconnection with busses of serial links are expected. This is e.g. seen in the memory area with HMC and with the serial standard for ADC JESD204B.

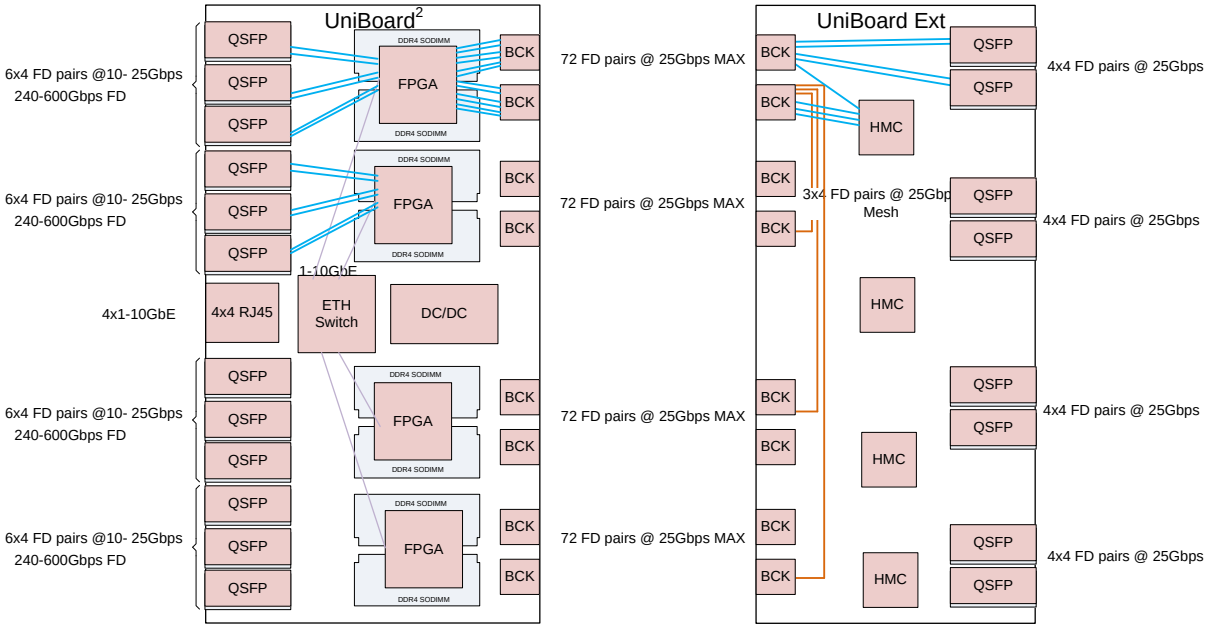


Figure 25 First setup UniBoard^2

5.1 Meshing on UniBoard

By placing a mesh from the FPGA to the connectors, a simple orthogonal backplane connector (section 3.2.2.1) can be used to make all connection from one FPGA on the vertical boards to all FPGA's on the horizontal boards, see Figure 26.

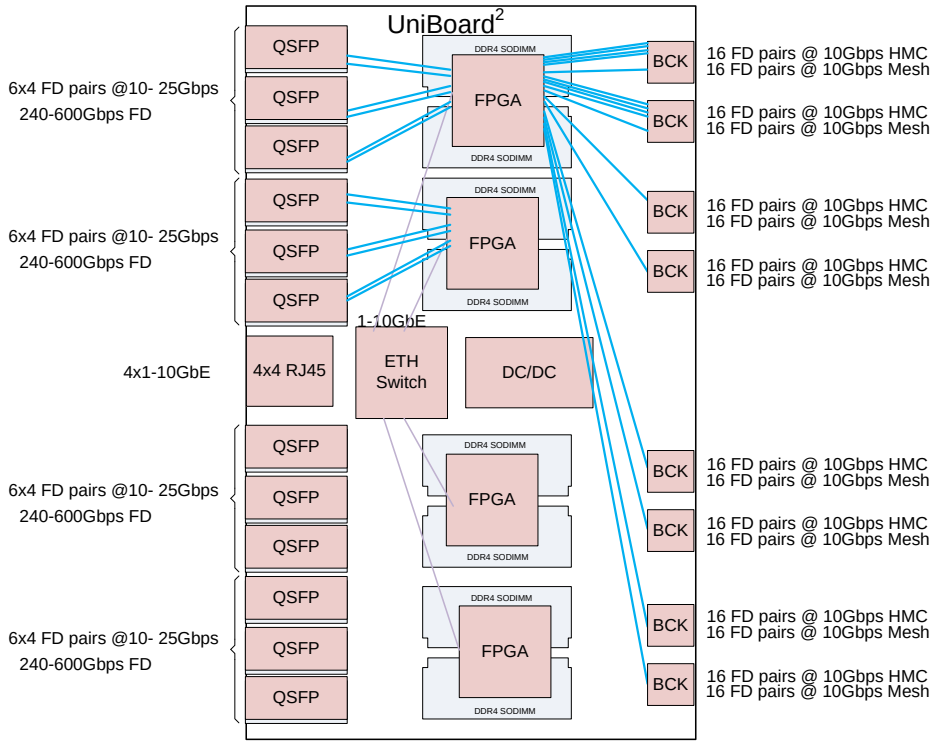


Figure 26 UniBoard with connector mesh

The pinning reserved for the mesh is shown in Table 10. In this table T is transmitter, F is FPGA, M is mesh and R is receiver. In this case four lanes per FPGA are used to a single backplane connector. By placing two backplane connectors a full 8x8 board structure can be made. The bandwidth from each FPGA to the other 32 FPGA's is expected to be 320Gbps in total.

Table 10 Pinning for the mesh

	1	2	3	4	5	6	7	8	9	10	11	12	
BC	TFOMO	TFOM1	TFOM2	TFOM3									AB
EF	TF1MO	TF1M1	TF1M2	TF1M3									DE
HJ	TF2MO	TF2M1	TF2M2	TF2M3									GH
LM	TF3MO	TF3M1	TF3M2	TF3M3									KL
PO					RF2M3	RF3M3	RF2M2	RF3M2	RF2M1	RF3M1	RF2M0	RF3MO	NP
ST					RF0M3	RF1M3	RF0M2	RF1M2	RF0M1	RF1M1	RFOM0	RF1M0	RS

A dedicated HMC-module can be made or systems where memory is needed or for single board systems. To enable these system the empty pairs can be used for the HMC interfaces. With two connectors per FPGA two HMC links can be made per FPGA enabling a storage bandwidth of <320Gbps (40GByte/s). The pinning for a single connector containing a single HMC link is shown in Table 11.

Table 11 Pinning for single HMC link

	1	2	3	4	5	6	7	8	9	10	11	12	
BC					TD0	RD0	TD4	RD4	TD8	RD8	TD12	RD12	AB
EF					TD1	RD1	TD5	RD5	TD9	RD9	TD13	RD13	DE
HJ					TD2	RD2	TD6	RD6	TD10	RD10	TD14	RD14	GH
LM					TD3	RD3	TD7	RD7	TD11	RD11	TD15	RD15	KL
PO													NP
ST													RS

The pairs used for the mesh can have other functions in the single board application e.g. QSFP+ IO. A full mesh can be made by connecting transmit to receive pins of a single connector on the extension board.

For Serial ADC applications the HMC receiver pins can be used for ADC input. The mesh pins can be used to make vertical connections form UniBoard-to-UniBoard. In this case two 320Gbps of ADC data can be fed into UniBoard.

The pins that have a function in neither Table 10 nor Table 11 can be used for general IO or other serial links.

REMARKS
With the long links for the mesh, the speed from FPGA-to-FPGA might be limited. The total expected trace length is in the order of 600mm.
In view of the added complexity of the board and longer links that such a mesh would imply, the decision was made to use the design as shown in figure 25.